

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



LSHTM Research Online

Jarvis, C; (2018) Spatial Analysis of Cluster Randomised Trials. PhD (research paper style) thesis, London School of Hygiene & Tropical Medicine. DOI: <https://doi.org/10.17037/PUBS.04648971>

Downloaded from: <https://researchonline.lshtm.ac.uk/id/eprint/4648971/>

DOI: <https://doi.org/10.17037/PUBS.04648971>

Usage Guidelines:

Please refer to usage guidelines at <https://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license. To note, 3rd party material is not necessarily covered under this license: <http://creativecommons.org/licenses/by-nc-nd/3.0/>

<https://researchonline.lshtm.ac.uk>

LONDON
SCHOOL *of*
HYGIENE
& TROPICAL
MEDICINE



Spatial Analysis of Cluster Randomised Trials

Christopher Jarvis

Thesis submitted in accordance with the requirements for the degree of
Doctor of Philosophy of the University of London

April 2018

Department of Infectious Disease Epidemiology
Faculty of Epidemiology and Population Health
London School of Hygiene and Tropical Medicine

Funded by the Medical Research Council London Hub for Trials Methodology Research,

Grant code MR/L004933/1- Q42

Declaration

Statement of Own Work

I, Christopher Jarvis, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, this has been indicated in the thesis. I have read and understood the School's definition of plagiarism and cheating given in the Research Degrees Handbook.

Christopher Jarvis

April 2018

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Abstract

Cluster randomised trials (CRTs) often use geographical areas as the unit of randomisation. Despite this, explicit consideration of the location and spatial distribution of observations is rare. In many trials, the location of participants will have little importance, however in some, especially against infectious diseases, spillover effects due to participants being located close together may affect trial results. This PhD takes a multidisciplinary approach to apply and evaluate spatial analysis methods in CRTs, furthering understanding of how spatial analysis can complement traditional evaluation of CRTs.

I began by conducting a systematic review of CRTs that used spatial analysis techniques. I found only 10 published papers, most of which being supplementary analyses of the main trial. I then conducted a spatial analysis of an Oral Polio Vaccine (OPV) transmission household CRT. This provided additional insights into the underlying mechanism of polio transmission that support the global cessation of OPV and emphasises the difficulties of the global eradication of polio. Following this, I performed a spatial reanalysis of an insecticide-treated bed net CRT, applying approaches from the systematic review and a new method I developed called *cluster reallocation* to assess the presence and impact of spatial spillover in the trial. This analysis confirmed the previous estimate of intervention effect while showing evidence of a spillover effect.

I carried out simulation studies to evaluate the impact of spillover and spatial effects on the standard CRT model and compared spatial regression to non-spatial models. These simulations focus on how to generate spatial spillover effects and the magnitude needed before spatial consideration becomes important to CRTs. I found that non-spatial CRT models are relatively robust to spatial effects and that the use of spatial models does not appear to improve upon the non-spatial model.

The collective findings of this thesis highlight that standard CRT approaches are typically robust to small scale spillover effects and consideration of the spatial distribution of observations appears to provide little utility in the main analysis of a trial. Despite this, spatial methods can provide additional insights into the mechanism of interventions and are well suited to secondary analyses of CRTs, especially with the increasing collection of GPS data in CRTs.

Acknowledgements

I would like to thank my supervisors John Edmunds, Neal Alexander, and Daniel Lewis, for their support, feedback and critique. I thank my original supervisor Gian Luca Di Tanna for his supervision during the first year of my PhD, and I am very grateful for John taking over supervision of my PhD and helping me complete within the original time frame. I have been fortunate to have had many beneficial conversations that have helped focus the direction of the thesis. I thank Karim Anaya-Izquierdo for hosting me in my second year and helping me get to grips with Gaussian Markov random fields. I thank Jo lines for his passionate conversations around spillover effects in mosquitoes trials and his insights into potential mechanisms. I very much enjoyed and am grateful to have collaborated with Yvonne Maldonado, Jonathan Altamirano, and Clea Sarnquist (Stanford), and Thomas Smith and Lea Multerer (Swiss TPH). I also thank Fred Binka, and for giving permission for data used in the PhD, and Emily Webb for providing data which unfortunately was unable to be used in this thesis.

The PhD experience has been greatly enriched due to many individuals at LSHTM and beyond. I owe a great deal to my dear friend Ellen White, who helped to proof read and critique this thesis, a service I will be repaying soon, as she will finish her PhD in nearly a year. To Christopher Rentsch for much needed and much unneeded gin. Chris Grundy for helping to develop my initial interest in spatial analyses. Tim Clayton, Stuart Bedston, and Neil Turner for their great guidance, professionalism, and their similar but very different straight talking attitudes that provided me with the skills to embark on the PhD.

I am very lucky to have had the support of my family. My parents have been wonderfully unique in never asking me to explain my thesis in detail, and providing unwavering support. My niece Mollie has been brilliant at closing my laptop and arresting me if I do not give her my full attention. This led to many hours in her *jail*, on multiple occasions. In particular, I am extremely indebted to my partner Bianca, who has been a great source of support, has been planning many trips away, has counted down the days until I hand this work in, and was not involved in typing this manuscript.

Acronyms

BYM	Besag, Yorke, and Mollie Model
CAR	Conditional autoregressive
CI	Confidence interval
CrI	Credible interval
CRT	Cluster randomised trial
cVDPV	Circulating vaccine-derived poliovirus
GIS	Geographical Information Systems
GLM	Generalised linear model
GLMM	Generalised linear mixed model
GMRF	Gaussian Markov random field
GP	Gaussian process
GPm	Gaussian process model
GPS	Global Positioning System
GWR	Geographically weighted regression
ICAR	Intrinsic conditional autoregressive
ICC	Intra-cluster correlation coefficient
IDW	Inverse distance weighting
IID	Multi-level model with independent random effect
INLA	Integrated nested Laplace approximation
IPV	Inactivated Polio Vaccine
ITN	Insecticide treated nets
LGM	Latent Gaussian model
MAUP	Modifiable Areal Unit Problem

MCMC	Markov chain Monte Carlo
MRF	Markov random field
MVN	Multivariate normal distribution
NIW	National Immunization Week
OPV	Oral polio vaccine
OR	Odds ratio
RR	Risk ratio
rt-QPCR	Quantitative reverse transcription polymerase chain reaction
SAR	Simultaneous autoregressive
SMR	Standardised mortality ratio
SP	Stochastic process
SPDE	Stochastic partial differential equation
SUTVA	Stable Unit Treatment Value Assumption
UTM	Universal Transverse Mercator
VAPP	Vaccine-associated paralytic polio
WHO	World Health Organisation
WT	Wild type

Notation

Statistics and parameters

s	A location in space-time
H	An attribute or property
$h(s)$	A property H for a given location s
d	Spatial variable usually involving distance or proximity weighted values
a, b	Coordinates (a, b)
W	Spatial weights matrix
c	Number of clusters
n	Total number of individuals
m	Number of individuals per cluster
π	Population proportion or pi when used in Gaussian formula
λ	Population rate
θ	Population parameter used when defining confidence intervals where $(\hat{\theta}_{low}, \hat{\theta}_{upp})$ are the lower and upper bounds of the interval and $\hat{\theta}$ is an estimate of the parameter
μ	Population mean
$\mu(\cdot)$	Mean function
σ	Population standard deviation
σ_b	Between-cluster standard deviation
σ_w	Within-cluster standard deviation
σ_v	Between-area standard deviation (when area is equal to cluster then $\sigma_v = \sigma_b$)
Σ	Variance-covariance matrix
$\Sigma(\cdot)$	Covariance function
ρ	Spatial correlation parameter
ω	Parameter used in distance weighting functions
ξ	Parameter used in distance weighting functions

Notation (continued)

Model terms

y	Outcome variable
z	Covariate
Z	Matrix of covariates with the first column constrained to be a vector of ones, also called a design matrix
x	Covariate reserved for intervention status
α	Intercept parameter
β	Regression parameter representing the intervention effect
γ	Regression parameters representing effect of a covariate
ψ	Regression parameter representing a spatial spillover effect
u	Random effect, typically $u \sim N(0, \sigma_u^2)$
v	Spatially correlated random effect
η	Linear predictor
ϵ	Error term typically $\epsilon \sim N(0, \sigma^2)$

Subscripts

i	Intervention arm ($i = 0$: control, $i = 1$: intervention)
j	Cluster ($j = 1, \dots, c$)
k	Participant ($k = 1, \dots, m$) so that y_{ijk} represents the outcome for the k^{th} participant in the j^{th} cluster in the i^{th} treatment arm
m	Observation where m^* is used to denote a set of neighbours
l	Covariate ($l = 1, \dots, L$)
s	Stratum or location ($s = 1, \dots, S$)
w	Within-cluster
b	Between-cluster
v	Geographical areas, the area used in the spatially correlated random effect v

Notation (continued)

Superscripts

t	Transpose of a matrix e.g. A^t
$*$	Denote neighbours where j
\dagger	conjugate transpose of a matrix e.g. A^\dagger
$G(\cdot)$	Link function. Where $y = G^{-1}(\eta)$

General notation

I	Identity matrix
$f(\cdot), g(\cdot)$	Functions
A, B	Matrices
L, U	Lower and upper triangular matrices. (typically used when performing matrix decomposition)
$\mathbb{E}(\cdot)$	Expectation of a random variable
$\mathbb{V}(\cdot)$	Variance of a random variable
Σ	Summation sign

Distributions

$N(\mu, \sigma)$	Gaussian (Normal) distribution
$Bin(\pi, n)$	Binomial distribution
$Pois(\lambda)$	Poisson distribution
$MVN(\mu, \Sigma)$	Multivariate normal distribution
$GP(\mu(\cdot), \Sigma(\cdot))$	Gaussian process with mean and covariance functions

Note: Extra notation for algorithms are defined in the inputs of the algorithms. The notation for Figure 3.2 differs slightly than defined here but a legend is provided detailing the notation for the figure.

Contents

Declaration	i
Abstract	ii
Acknowledgements	iii
Acronyms	iv
Notation	vi
List of Figures	xv
List of Tables	xvii
List of Algorithms	xviii
1. Introduction	1
1.1. Motivation for PhD	2
1.2. Aim	3
1.3. Objectives	3
1.4. Thesis structure	3
1.5. Bibliography	6
2. Background	7
2.1. Overview	7
2.2. Clinical trials	8

2.3. Cluster randomised trials	10
2.4. Spillover effects	16
2.5. Spatial data analysis	27
2.6. Spatial modelling	33
2.7. Linking spatial data types with CRTs	39
2.8. Summary	45
2.9. Bibliography	47
3. Paper A: Spatial analysis of CRTs: A systematic review	55
3.1. Overview	57
3.2. Role of candidate	57
3.3. Abstract	58
3.4. Background	60
3.5. Methods	62
3.6. Results	63
3.7. Discussion	72
3.8. Bibliography	76
4. Paper B: Spatial analyses of OPV transmission in an IPV-vaccinated community	80
4.1. Overview	83
4.2. Role of candidate	83
4.3. Abstract	84
4.4. Background	85
4.5. Methods	86
4.6. Results	92
4.7. Discussion	100
4.8. Bibliography	103

5. Paper C: Spatial effects of permethrin-impregnated bed nets	106
5.1. Overview	109
5.2. Role of candidate	109
5.3. Abstract	110
5.4. Background	112
5.5. Methods	113
5.6. Results	120
5.7. Discussion	125
5.8. Extended discussion of applied spatial methods	127
5.9. Bibliography	131
6. Simulating spatial effects in CRTs	137
6.1. Overview	137
6.2. Introduction	138
6.3. Creating spatial CRT data	139
6.4. Simulating a CRT	140
6.5. Simulating CRTs with spatial effects	143
6.6. Simulating a CRT with two clusters and spatial effects	153
6.7. Summary	162
6.8. Bibliography	164
7. Simulation study of the impact of spatial effects in CRTs	167
7.1. Overview	167
7.2. Introduction	168
7.3. Methods	169
7.4. Results	177
7.5. Discussion	185
7.6. Bibliography	193

8. Cluster reallocation: Exploring spatial spillover in CRTs	196
8.1. Overview	196
8.2. Introduction	197
8.3. Cluster reallocation	198
8.4. Applying the method to simulated data	204
8.5. Summary	215
8.6. Bibliography	217
9. Discussion	218
9.1. Overview	218
9.2. Summary of findings	219
9.3. Strengths	227
9.4. Limitations	232
9.5. Implications	235
9.6. Future work	237
9.7. Concluding remarks	242
9.8. Bibliography	243
A. LSHTM ethics approval	246
B. License agreement for paper A	249
C. Spatial modelling with random fields	251
C.1. Introduction	251
C.2. Origins of SAR and CAR	252
C.3. Random variables, processes and fields	252
C.4. Gaussian processes and Gaussian Markov fields	255
C.5. Non-Gaussian outcomes	255
C.6. Model fitting	256
C.7. Summary	260
C.8. Bibliography	261

List of Figures

1. Introduction	1
2. Background	7
2.1. Diagrams of different types of effects that can be estimated in cluster randomised trials	15
3. Paper A: Spatial analysis of CRTs: A systematic review	55
3.1. Flow chart of search results	64
3.2. Spatial analysis methods	70
4. Paper B: Spatial analyses of OPV transmission in an IPV-vaccinated community	80
4.1. Map of study area	89
4.2. Distance from nearest vaccinated shedding household for unvacci- nated individuals	96
4.3. Median distance from vaccinated shedding over time	97
4.4. Legend for Figure 4.5	97
4.5. Spatial mapping of OPV transmission shedding over time	98
5. Paper C: Spatial effects of permethrin-impregnated bed nets	106
5.1. Study area, spatial variation of SMR, and spatial heterogeneity of intervention effect from GWR model.	122

5.2. The change in effect estimate calculated by cluster reallocation of intervention participants to the control arm and vice versa.	123
6. Simulating spatial effects in CRTs	137
6.1. Study area showing spillover area	157
6.2. Overview of the spatial spillover simulations	159
6.3. Overview of the spatial correlation simulations	160
7. Simulation study of the impact of spatial effects in CRTs	167
7.1. Map of study area	177
7.2. The effect of spatial effects on the bias of standard cluster randomised trial analysis model (true effect is 2.0)	179
7.3. Zipplot of the effect of a spatial spillover effect on the coverage of cluster randomised trial model	180
7.4. Zipplot of the effect of a simultaneous autoregressive spatially correlated intervention effect on the coverage of a cluster randomised trial model	181
7.5. Zipplot of the effect of a simultaneous autoregressive spatially correlated error term on the coverage of a cluster randomised trial model	182
7.6. The effect of the spatial effects on the bias of the spatial models	184
7.7. Zipplot of the effect of a spatial spillover effect on the coverage of a BYM model	189
7.8. Zipplot of the effect of a spatial spillover effect on the coverage of a GPm model	190
7.9. Zipplot of the effect of a spatial spillover effect on the coverage of a BYM model	191
7.10. Zipplot of the effect of a spatially correlated intervention effect on the coverage of a GPm model	192

8. Cluster reallocation: Exploring spatial spillover in CRTs	196
8.1. Diagram of the cluster reallocation method	200
8.2. Cluster reallocation plot when no spillover is present	202
8.3. Cluster reallocation plot when positive spillover is present	203
8.4. Map of study area	207
8.5. Cluster reallocation plots for simulated scenarios with different strengths and range of spatial spillover effects	211
8.6. Cluster reallocation plot with bar chart for displaying imbalances in trial arms (Taken from Chapter 5)	214
9. Discussion	218
A. LSHTM ethics approval	246
B. License agreement for paper A	249
C. Spatial modelling with random fields	251
C.1. The connection between random variables and Gaussian Markov ran- dom fields	254

List of Tables

1. Introduction	1
2. Background	7
2.1. Alternative terms that are similar to spillover and/or contamination .	20
2.2. Overview of distance weighting functions	35
2.3. Eighteen possible types of spatial cluster randomised trials	40
2.4. Four common types of spatial cluster randomised trials with examples	41
3. Paper A: Spatial analysis of CRTs: A systematic review	55
3.1. Search terms	63
3.2. Characteristics of the included papers	66
4. Paper B: Spatial analyses of OPV transmission in an IPV-vaccinated community	80
4.1. Characteristics of unvaccinated individuals by coverage area	94
4.2. Spatial characteristics comparing unvaccinated shedding with unvac- cinated non-shedding individuals	95
4.3. Model considering effects of spatial variables on unvaccinated shedding	99
5. Paper C: Spatial effects of permethrin-impregnated bed nets	106
5.1. Summary of analyses of spatial patterns of dependence, heterogeneity, and spillover	121
5.2. Estimate of intervention effect by spatial model type	124

5.3. Spillover effect of distance to discordant pair presented by model type and interaction of distance with bed net intervention.	125
6. Simulating spatial effects in CRTs	137
7. Simulation study of the impact of spatial effects in CRTs	167
7.1. Study values that vary in the simulations	173
8. Cluster reallocation: Exploring spatial spillover in CRTs	196
9. Discussion	218
A. LSHTM ethics approval	246
B. License agreement for paper A	249
C. Spatial modelling with random fields	251

List of Algorithms

6.1. Simulating spatial data for a cluster randomised trial	140
6.2. Simulating cluster randomised trial data	142
6.3. Creating distance based spillover effects	147
6.4. Distance weighting functions	148
6.5. Simulating a simultaneous autoregressive spatially correlated value . .	150
8.1. Cluster reallocation	204
C.1. Metropolis-Hastings	257

1. Introduction

1.1. Motivation for PhD

Cluster randomised trials (CRTs) often use geographical areas as the unit of randomisation. Despite this, explicit consideration of locations and the spatial distribution of observations is rare [1]. When participants are close to one another there is the potential for effects to ‘spill over’ from one person to another. This is likely to affect CRTs when clusters are close to one another as there is the potential for these spillover effects to cross cluster boundaries. For example, an infected participant may travel across a cluster boundary and infect participants of that cluster. This movement, will violate the assumption of between-cluster independence, which is typically assumed in the analysis of CRTs.

Designing trials with well-separated clusters is a possible route to avoiding spillover effects. However, the range of a spillover effect is often unknown during the design phase of a trial, and disease autocorrelation can range over distances of several kilometres [2]. Furthermore, due to pragmatic reasons it may not be possible to design CRTs with clusters that are far apart. The existence of spillover also presents an opportunity to measure additional effects of an intervention, thus furthering understanding of the mechanism of the intervention. Therefore, there is a need to assess and develop analysis methods that can be used to adjust for and measure spillover effects.

Spatial statistics contains a large body of research which focuses on how to incorporate location into the analyses of data [3]. The use of spatial statistics in medical research is sporadic and in CRTs only a few examples exist at present [1, 4]. This PhD builds upon the work of co-supervisor Neal Alexander [5–7]. This thesis contributes to bridging trial methodology and spatial statistics, in an attempt to improve understanding of how to use the location of data for the benefit of CRTs.

1.2. Aim

The aim of this research is to explore the use of spatial methods within the analysis of CRTs. Specifically, to improve and develop knowledge of: methods that can be applied to CRTs, the impact of spatial effects on trial results, and the additional utility that can be gained from considering the spatial context of a CRT.

1.3. Objectives

The aim will be met by fulfilling the following objectives:

1. Describe and frame CRTs in relation to spatial data and summarise implications for spatial analyses.
2. Describe and identify spatial analysis methods that have been previously used in CRTs by conducting a systematic review.
3. Apply and assess a range of appropriate modern spatial methods to existing CRT data, in order to analyse the effect of spatial autocorrelation and spatial spillover effects on CRT results.
4. Evaluate the impact of spatial effects and the utility of spatial models in the analysis of a CRT by means of a simulation approach.
5. Develop methodology to assess the presence of spatial spillover in CRTs.

1.4. Thesis structure

This *research paper style thesis* consists of three parts, containing nine chapters in total. A research paper style thesis contains chapters that are written in the style of a journal article, and more traditional PhD chapters. The research paper chapters are prefixed with “Paper X.”, and the traditional PhD chapters have no prefix. Research papers chapters were chosen for applied work, and traditional PhD

chapters for the methodology and simulation studies. This removes the often short word count enforced by journals, allowing deeper exploration in the methodology and simulation study chapters.

The present chapter is a brief introduction, detailing the aims and original contributions of the PhD. Chapter 2 provides an overview of CRTs and spatial analysis, providing a framework from which to consider spatial data in the context of CRTs. Chapter 3 is a systematic review of spatial methods that have been used in CRTs and was published in September 2017 [1].

Chapter 4 is an applied analysis of an oral polio vaccine (OPV) trial conducted in Mexico. In this CRT, the investigators wanted to explore whether living near an individual who receives the OPV increases an individual's risk of poliovirus shedding. This analysis provides the first opportunity in the PhD to apply spatial methods to a CRT. Chapter 5 is an applied spatial reanalysis of the earliest CRT found in the systematic review (Chapter 3). Chapter 5 explores and applies a range of spatial methods to a CRT, including a novel method I developed, called *cluster reallocation*, which explores the presence of spillover.

Chapters 6 and 7 are simulation studies. Chapter 6 demonstrates the algorithms used to simulate CRT data with spatial effects. This provides an opportunity to evaluate and develop understanding of these algorithms, in a simplified setting with only two clusters. Chapter 7 is a larger simulation study, focusing on two aspects: determining the magnitude of how much spatial spillover affects the standard analysis of a CRT, and assessing whether spatial analysis methods can be used to overcome and adjust for spillover effects. Chapter 8 describes and tests the ability of the cluster reallocation method to help identify spillover under a range of simulated conditions. A detailed discussion of spatial method is presented in Appendix C, this section is not needed to understand the thesis but does provide background information that may help to provide a more holistic view and understanding of spatial methods. In chapter 9, I review and synthesise the key points of the PhD, considering the

impact on current practice, possible directions for future work, and limitations of the PhD.

The original contributions of this thesis are: the systematic development of knowledge about the use of spatial analysis methods in CRTs, the assessment of the type of spatial methods that are appropriate to use in CRTs, and the development of a new method called cluster reallocation, which explores the presence and magnitude of spatial spillover.

This thesis provides an in-depth consideration of spatial methods in CRTs, by combining two rarely overlapping disciplines, spatial statistics and CRT.

1.5. Bibliography

- [1] Jarvis, C, Di Tanna, GL, Lewis, D, Alexander, N, and Edmunds, WJ. Spatial analysis of cluster randomised trials: A systematic review of analysis methods. *Emerging Themes in Epidemiology* 2017 14. (12).
- [2] Howard, SC, Omumbo, J, Nevill, C, Some, ES, Donnelly, CA, and Snow, RW. Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2000 94. (4) 357–360.
- [3] Cressie, N. *Statistics for Spatial Data*. United States of America: John Wiley & Sons Ltd, 1991.
- [4] Auchincloss, AH, Gebreab, SY, Mair, C, and Diez Roux, AV. A review of spatial methods in epidemiology, 2000-2010. *Annual review of public health* 2012 33 107–22.
- [5] Alexander, N, Moyeed, R, and Stander, J. Spatial modelling of individual-level parasite counts using the negative binomial distribution. *Biostatistics* 2000 1. (4) 453–463.
- [6] Alexander, N, Moyeed, Ra, Hyun, PJ, Dimber, ZB, Bockarie, MJ, Stander, J, Grenfell, BT, Kazura, JW, and Alpers, MP. Spatial variation of Anopheles-transmitted *Wuchereria bancrofti* and *Plasmodium falciparum* infection densities in Papua New Guinea. *Filaria journal* 2003 2 14.
- [7] Lenhart, A, Orelus, N, Maskill, R, Alexander, N, Streit, T, and McCall, PJ. Insecticide-treated bednets to control dengue vectors: Preliminary evidence from a controlled trial in Haiti. *Tropical Medicine and International Health* 2008 13. (1) 56–67.

2. Background

2.1. Overview

Chapter 1 provided a brief summary of the motivation, aim, and objectives of the PhD. It outlined the structure and original contributions of the thesis. In this chapter, I provide a more detailed account of cluster randomised trials (CRTs) and spatial statistics. I also address objective 1 by proposing a framework for relating spatial data to CRTs, and consider implications for spatial analyses.

Objective

1. Describe and frame CRTs in relation to spatial data and summarise implications for spatial analyses.

2.2. Clinical trials

Randomised clinical trials often demand a substantial investment of money, time, and expertise, and can place a considerable burden on both trial personnel and participants [1]. This substantial investment is made because well-conducted trials are recognised as the most reliable way to evaluate the efficacy and safety of new treatments [2]. Increasingly, major public health decisions are formed based on a careful review of available evidence, wherein rigorously conducted trials are considered to be the gold standard [3, 4]. Therefore, in order for trials to produce the highest standards of evidence, it is crucial that they are designed and analysed to minimise the risk of bias, ensure the safety of participants, and adhere to ethical requirements for research.

Several guidelines relating to ethics, conduct, and reporting of trials have been developed [5–7]. However, trial conduct changes over time and assuring quality requires scrutiny and an understanding of established best practices. This has led to an increased focus on trial methodology with the goal of improving the way trials are conducted.

In perhaps the earliest example of a clinical trial, James Lind performed a comparative study of treatment for scurvy [8]. The trial was conducted on a sample of twelve patients with six different treatments. Although effort was made to make the participants ‘*as similar as I could have them*’, there was no randomisation or blinding used [8]. It would take nearly 200 years of development, when in 1948 the Medical Research Council conducted the first blinded randomised clinical trial with a control group [9]. In the subsequent 70 years, the rate of methodological development in trials has increased, moving greatly beyond the initial ‘drug trials’ to consider structural, organisational, and behavioural change interventions [10]. However, this expansion has implications on the complexity of design and analysis. Although much progress has been made, the main developments of comparison groups, randomisation, and blinding remain fundamental to the conduct of modern

day trials.

Comparison groups: To establish the effect of an intervention, participants receiving the intervention must be compared with participants who did not receive the intervention. The two comparison groups are usually known as the intervention and control arms. Without a comparison group, we cannot be sure that a suspected intervention effect has occurred. Even with a comparator group, bias may still arise if systematic differences exist between participants in control and intervention arms. For instance, if the control arm is much older and sicker than the intervention arm then they will likely have worse outcomes irrespective of treatment efficacy. A trial may have more than two arms and compare multiple interventions.

Randomisation: Randomisation involves the random allocation of participants. Random assignment helps to ensure that there are no systematic differences in known or unknown characteristics between the groups. This is a considerable strength of randomised controlled trials as investigators can even account for prognostic factors that they are unaware of [11]. It should not be possible to predict in advance which arm of a trial a future patient would be enrolled in.

Blinding: Where possible, individuals involved in a trial should have no knowledge as to who is allocated to which comparison group, and hence who receives which intervention. In a ‘double’ blind trial, both the investigator and participants are unaware of who receives treatment and who receives the control [2]. If individuals and investigators are aware of who is allocated treatment, then this can result in bias. For instance, consider a trial where lung cancer patients receive a non-standard drug treatment. If an investigator decides to pay extra attention to participants on the non-standard treatment, then they may have better outcomes, due to an increased level of care. Blinding may be straightforward in simple settings such as drug trials, but can be more complicated or impossible for complex interventions such as mass

education programmes.

Complex interventions include several interconnected components as an intervention [12]. For instance, rather than a single treatment, such as administering a drug, a complex intervention might consist of simultaneous components, including structural changes (e.g. improved green space), health promotion (e.g. positive messaging about physical activity), and behavioural change interventions (e.g. free gym membership). In this situation it can be difficult to disentangle the exact mechanisms by which the treatment succeeded or failed. They often include community and group based interventions where groups of connected individuals will receive the same intervention. For example, all patients attending the same clinic may receive a new type of surgery, while patients attending a separate clinic all receive the standard surgery. Trials involving complex interventions are key to establishing the efficacy of such interventions, however they tend to require greater thought and pose greater challenges to those who conduct them.

Utilising comparison groups, randomisation, and blinding can be difficult, if not impossible to achieve for complex interventions at the individual level. If we imagine a sexual health campaign where adverts are placed on buses; it would be difficult to establish who saw the adverts and who did not and thus defining controls and intervention participants would be hard. It is also implausible to randomly allocate people either to look at or ignore the adverts. Lastly, it seems unlikely that we can conceal from individuals whether they have seen the advert or not. Therefore, further design choices need to be made to investigate the effectiveness of complex interventions and a popular choice of design is the cluster randomised trial (CRT).

2.3. Cluster randomised trials

In CRTs, individuals are first grouped (or clustered) together, then each cluster is randomly assigned to intervention or control, allocating all the individuals within

the same cluster together to a specific trial arm [13]. A cluster can take many forms, which are often socially meaningful, for example, a household, workplace, or geographical area. A CRT consists of two levels: the cluster at randomisation level, and the participants, at data collection or observation level. CRTs are also known as group, community, and place-based trials.

An example of a geographical CRT is a primary care trial, where geographical areas (e.g. local authority districts, municipalities, regions etc.) are randomised to provide different primary care services. The clusters are the geographical areas responsible for providing primary care services and the participants are the people served by the primary care centres in each geographical area. A household CRT, involves randomising households to receive the intervention or to not receive the intervention. The clusters are the households and the participants are individuals living within the households.

Participants in individually randomised trials are typically assumed to be statistically independent, that is to say, the specific characteristics of any one observation has no bearing on any other. By comparison, in CRTs, individuals within the same cluster tend to be correlated, owing to the clustered nature of the trial design [4]. Within-cluster variability represents how different the people in the same cluster are to one another. Between-cluster variability represents how different the clusters are to one another. The efficiency of a CRT increases as the within-cluster variance reduces relative to the between-cluster variance [4]. The within-cluster variance will be zero when there is one participant per cluster (an individually randomised trial). Therefore, for a fixed sample size, an individually randomised trial is often more efficient than a CRT design [13]. Due to this comparative inefficiency, CRTs tend to be larger, costlier, and sometimes harder to manage compared to individually randomised trials.

CRTs are also at high risk of selection bias, especially if participants are aware of the intervention before trial enrolment. For instance, individuals may travel to a

different cluster if they know an otherwise unavailable treatment is available there [13].

Analysing CRTs tends to be more challenging than analysing individually randomised trials. This requires statistical methods which account for the fact that observations come from a number of predefined groups rather than each one being an independent observation. There are two approaches to analysing CRTs, the one-stage and the two-stage method:

The one-stage method accounts for correlation within clusters by using either a generalised linear mixed effects model with a random effect for clusters, or generalised estimating equations. This method provides an estimate of the size of the intervention effect whilst also accounting for dependency (the non-independence of observations) within clusters.

The two-stage method involves aggregating the observations at a cluster level to give one summary statistic per cluster, such as a risk measure or mean. The control and intervention clusters are then compared using a t-test or regression model.

The two-stage method provides results that are more reliable compared to the one stage method when there are few clusters, as using a mixed effects model can give unreliable results when there are fewer than 15-20 clusters per trial arm [4]. However, the one-stage method treats each cluster as providing a single observation, thus resulting in a loss of information due to ignoring the multiplicity of data points. Neither the one- nor the two-stage methods make specific adjustment for between-cluster dependence.

2.3.1. Reasons for choosing a CRT

The main reasons for choosing a CRT design are [4, 13, 14]:

- Type of intervention
- Logistics

- Ethics
- Avoiding contamination

Some interventions are better delivered to groups and are difficult or impossible to evaluate with an individually randomised trial [14]. For instance, mass education campaigns involving TV or radio will be seen by many individuals and it is very difficult to prevent control individuals from watching TV [13]. However, it is possible to conduct a CRT where different areas of a country receive different TV or radio shows.

When data collection requires visiting numerous locations, a CRT is often more logistically convenient compared to an individually randomised trial. For instance, in a trial evaluating bed nets, field workers may need to check whether the bed nets are being used. Having many participants close together will reduce the time, cost, and difficulty of collecting the data.

A further reason for choosing to adopt a CRT design relates to ethics and acceptance within a community. It may be unethical to only provide the intervention to some members of a community. In addition, it may be easier to convince a community to receive an intervention, if all members receive it, particularly when a study is unblinded.

One of the most common reasons for choosing a CRT over an individually randomised trial is to minimise the risk of contamination. Contamination is when individuals in one arm receive or are exposed to the intervention of another arm [4]. Contamination can be very difficult to prevent in individually randomised trials. For example, in a study of Pre-exposure Prophylaxis for HIV prevention, it was found that a small number of participants shared the intervention with their partners [15]. When contamination is present in a trial, the effect of the intervention will be underestimated as differences between the intervention and control arms are reduced, due to both being exposed to the intervention.

Allocation of interventions to groups rather than individuals can help reduce the risk

of contamination. However, if individuals are connected through some mechanism then it can still pose a significant problem within CRTs [4].

2.3.2. Direct, indirect, total, and overall effects

One of the main reasons for conducting an individually randomised trial, is to measure the direct effect of the intervention: how much an intervention affects an individual who receives it [16]. CRTs allow for estimation of additional types of effects. The four effects that can be calculated as part of analysing a CRT are: indirect, direct, total, and overall effect [17] (Figure 2.1). Of these four effects, only the overall effect is a randomised comparison [4].

In a conventional two arm CRT, the direct effect is measured by comparing those who receive the intervention and individuals who do not receive the intervention from the intervention clusters. This measure is the same as the intervention estimate that would be calculated if an individually randomised trial was conducted instead of a CRT [4].

The indirect effect as defined in Hayes & Moulton is measured by comparing the individuals not receiving the intervention in intervention clusters to the individuals in the control clusters [4]. An indirect effect refers to the effect of an intervention on individuals who do not receive the intervention. The indirect effect is of particular importance when investigating interventions that aim to reduce the risk of acquiring or transmitting infectious diseases [18]. Herd immunity, is an example of an indirect effect of vaccination, where individuals who are not vaccinated (i.e. have not received the intervention) have a lower risk of disease when living in a community with high vaccination coverage.

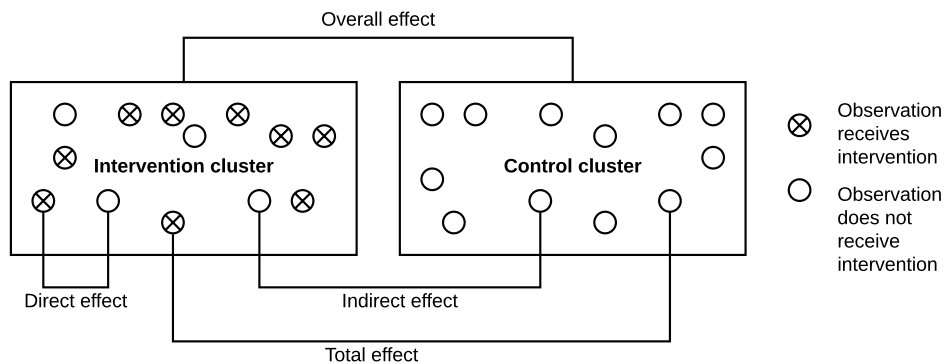
The total effect can be calculated by comparing individuals who receive the intervention in the intervention clusters and the entire control cluster [4]. This is the effect of the intervention compared to individuals who have not been exposed to the

intervention. Therefore, it represents the effect of the intervention, in an individually randomised trial when no contamination is present.

The overall effect is calculated by comparing participants in the intervention clusters with those in the controls clusters. The overall effect resembles the effect of the intervention in real life settings [17]. When an intervention is implemented in a community, it is likely that some individuals will receive the intervention, and some will not. Therefore, the overall effect provides an estimate of the change in outcome due to introducing the intervention throughout a community .

The four effects help to provide a more comprehensive view of how well an intervention may work when used in a real-world setting. In practice, it can be hard to measure these effects because we may not know who received the intervention and therefore, it is not possible to determine which observations to compare.

Figure 2.1. Diagrams of different types of effects that can be estimated in cluster randomised trials



Adapted from Halloran, Longini, and Struchiner 2009.

The definition of indirect effect presented here assumes that there are individuals in the intervention clusters who do not receive the intervention, and that the control clusters are unaffected by intervention clusters. These two assumptions may not hold. Individuals in the intervention cluster may all receive the intervention, or we may not be aware of such individuals, or have recorded their data. Additionally,

individuals in the control clusters may also be affected by the intervention clusters.

The more general definition of indirect effects is the effect of an intervention on individuals who do not receive the intervention [4]. This term appears to be used more frequently in CRT literature in both public health and in economics [19] and it is often used in medical CRTs involving infectious diseases [18]. Unfortunately, the term indirect effect is also used to refer to the effect of a mediator variable [20], but this is not what is meant by an indirect effect in this thesis.

2.4. Spillover effects

Spillover effects are the effect of an intervention on individuals who are in physical or social proximity to other intervention recipients [21, 22]. The distinguishing feature between contamination and spillover is that contamination involves individuals unintentionally receiving or being exposed to the intervention, whereas spillover affects individuals who do not receive the intervention that is the source of the spillover. Furthermore, spillover effects can occur without individuals in the control arm having received the intervention. Contamination, so defined, is not the main interest of this thesis, however many of the issues relating to spillover are also relevant for contamination.

Spillover as defined in this thesis, need not be restricted only to individuals within the control arm. Although a spillover effect implies a recipient did not receive the intervention, it could also refer to not receiving the intervention that is the source of the specific spillover effect on the recipient. For example, imagine a scenario where an individual receives an intervention that directly benefits them and results in spillover. If no other intervention participants are nearby then they will not be subject to spillover. If other intervention participants are nearby, then in addition to the direct benefit, they may be affected by spillover from the nearby intervention participants. In this case, although they received the intervention, the spillover

effects stem from interventions that were given to other participants and not from the intervention the individual specifically received.

This scenario can be adapted so that the individual does not receive the intervention. Then, if intervention participants are nearby, the individual will still be subject to the spillover effect. Therefore, the spillover effect is unrelated to their intervention status. Moreover, they would have been subject to the spillover regardless of whether they were given the intervention or not. This is because the spillover effect stems from their proximity to participants.

2.4.1. Positive and negative spillover

Spillover effects can be distinguished by whether they are positive or negative [23, 24]. In this thesis, a *positive spillover effect* benefits individuals who are affected by the spillover, and a *negative spillover effect* harms individuals who are affected by the spillover. It may be possible for positive or negative spillover to occur with beneficial or harmful interventions.

An example of a beneficial intervention could be insecticide-treated nets (ITNs) that reduce risk of malaria and provides a positive spillover effect due to reduced mosquito populations nearby. Alternatively, if a mosquito repellent is used it may be beneficial for the individual as mosquitoes are repelled away from an area, but result in negative spillover due to diverting mosquito populations to nearby locations.

An example of a negative intervention could be reducing the number of free years of education in an area. This intervention could result in a positive spillover for nearby areas due to less competition for university places or jobs. Alternatively, reducing access to education may result in people relocating to nearby control areas, which increases class sizes and may reduce the quality of teaching. In this case the negative intervention would result in a negative spillover effect. In practice, it is very unlikely that a trial would be conducted if the intervention is known to be harmful, however, this does not preclude the fact that an *a priori* assumed beneficial intervention

can turn out to be harmful after randomised evaluation [25, 26]. In addition, the consequences of an intervention such as a spillover effect can be unintended and difficult to foresee whether it would be positive or negative.

2.4.2. Mechanisms for spillover

In CRTs, there is frequently an assumption of an absence of between-cluster spillover; that movement of people and diseases occurs freely within a cluster but movement between clusters is negligible. Spillover between clusters can occur when clusters are connected through some mechanism. A possible mechanism for connectivity is location, where clusters that are located near to each other may affect each other. For example, in a trial of insecticide treated bed nets individuals in control clusters who live near intervention households may have a reduced risk of malaria due to a reduced population of mosquitoes in their surrounding area. Within-cluster spillover can also occur in a trial, however as long as clusters remain unconnected then this is likely less impactful on comparisons between the intervention and control arm.

Benjamin-Chung et al. identify and propose the following mechanisms for spillover effects in their review of health-related spillovers [22] :

1. **Distance-based:** spillover effects that occur within a specified distance of an intervention participant.
2. **Conditional on exposure to other participants' outcomes:** spillover due to the proportion of disease cases within a specified distance.
3. **Conditional on intervention density:** spillover related to the number of treated individuals within a specified distance.
4. **Treatment coverage mean/effect:** treatment coverage spillover refers to when a greater level of treatment coverage is associated with a reduced risk of outcome.

5. **Within-cluster:** spillover that occurs over small distances and only within clusters.
6. **Social network:** spillover between members that are connected socially.

Spillover effects due to location are the main interest of this thesis, specifically spatial spillover effects that cross cluster boundaries. Spillover effects that cross cluster boundaries are also known as between-cluster spillover,. Unless specified the term spillover will be used synonymously with between-cluster spillover, and within-cluster spillover will be reserved for when distinction of the type of spillover is required. Of the six mechanisms identified by Benjamin-Chung et al. (1) distance-based, (2) conditional on exposure to outcome, and (3) conditional on treatment density are due to physical proximity and are most relevant for this PhD [22] .

Spatial spillover as defined in this thesis, refers to when the intervention effect increases or decreases due to proximity of control participants to intervention participants .

2.4.3. Similar terms for spillover

There are many terms used to describe concepts that are similar to spillover. Confusion arises from there being multiple meanings for the same terms; such as the example as we have seen with indirect effects being both a mediator and an effect on people who do not receive an intervention. Further difficulties arise from the words with slightly different meanings being used interchangeably, for instance, sometimes contamination is used when spillover is meant and vice versa.

In this section I list and detail some of the complementary terms used for spillover (and/or contamination). Although there are slight differences in definition and meaning, they all refer to the idea of individuals being indirectly or unintentionally affected through some mechanism.

Table 2.1. Alternative terms that are similar to spillover and/or contamination

Term	Definition
Edge effects*	Edge effects refer to changes in populations or communities that occur near the boundaries of a habitat. They represent the impact that events outside an area have on that area. This term appears in geography and ecology and is revisited in section 2.5.3 [27, 28].
Externalities	An externality refers to the effect on an individual who did not choose receive that effect [29]. This term appears to be more common in the economics literature [22] and differs from spillover in that it refers to the choice of the individual affected.
Herd effects	Herd effects refer to the reduction in risk of infection among susceptible individuals due to the presence and proximity of immune individuals [30]. This term is mainly used in literature related to infectious diseases and could be considered as a subset of indirect effects.
Interference	Interference refers to when a treatment given to a participant affects not only that person, but also other participants [31]. This term appears to be more common in causal inference literature [31–34].
Peer Effects	Peer effects refer to the impact that peers have on an individual. This is when measures are correlated among peers. This term appears to be common in social science and economics literature [35].

*Edge effects are also used to refer to a spatial analysis boundary problem where information does not exist outside a study area and therefore the analysis is ignorant of how objects outside the study area may affect observations within the study area. [36]

2.4.4. Spillover in CRTs

Spillover can be both a design and an analysis consideration for CRTs. When control and intervention clusters are sufficiently far apart, the risk of spillover is negligible, and a conventional analysis will produce unbiased effect estimates. However, the reality of trial design is that clusters may not be well-separated in geographical space, social space. This is usually for eminently practical reasons, such as logistics, cost or resource and personnel constraints. Further, it can be difficult to assess *a priori* whether a particular design produces well-separated clusters in practice. For these reasons, it is prudent to give consideration to spillover and contamination throughout the life course of a CRT.

Valid inferences in randomised experiments rely on the Stable Unit Treatment Value Assumption (SUTVA). SUTVA requires that differences between the outcomes of control and intervention participants only depend on the participant's intervention status and not on what intervention others receive [37]. SUTVA can be violated when spillover is present [21]. The underlying reason for spillover is the proximity of individuals to one another. Violation of SUTVA is more likely if individuals are physically or socially close to one another as there is a greater chance that they may affect each other's outcomes.

Social proximity may refer to friendship or work groups and these groups provide networks through which information or diseases could spread. Knowledge-based social proximity may refer to connections through the internet and would be important when the intervention is information based. Social contacts can be important in infectious disease outbreaks where contact tracing is often used to infer a network through which infection might spread [38].

Physical proximity relates to the closeness in location between individuals or individuals to an exposure, and may refer to individuals who live near to each other, or whose daily activity spaces intersect. This type of proximity will be important when there is some impact on the surrounding environment. For instance, if a large number of people in a village use ITNs then this may reduce the population of mosquitoes in the village, providing a protective effect to individuals who live nearby but do not have a bed net.

Social and physical proximity are not mutually exclusive. Both may be present in a trial and affect results [39]. Furthermore, social and physical proximity may in some cases be equivalent, such as friends who live near one another. This PhD focuses on spillover stemming from physical rather than social proximity.

2.4.5. Consequences and opportunities

When spillover effects are present, individuals in the control clusters are indirectly affected by the intervention clusters. Estimates of the intervention effects may increase or decrease when affected by spillover. If the intervention is beneficial and positive spillover is present, then effect of the intervention will be diluted. Control participants receive a benefit from proximity to the intervention, and the difference between the control and intervention arms are reduced. This could result in effective interventions being missed due to underestimated effect estimates [14]. If the intervention is beneficial and negative spillover is present, the effect of the intervention will decrease. Control participants will receive a harmful effect from proximity the intervention, and differences between the control and intervention arms are enlarged. This will result in overestimating the intervention effect.

Spillover between clusters in a CRT may violate SUTVA, biasing the causal effects estimated from randomised comparisons. Spillover can be considered as a problem for trials, or could be viewed as an opportunity to measure the additional effect of the intervention for individuals who do not even receive the intervention. In some contexts, particularly in infectious diseases, measuring spillover or more generally indirect effects (such as herd effects) might be one of the main aims of the study [4, 18].

In vaccine trials, direct and indirect benefits of vaccines, are both of interest, to determine how much of the population need to be vaccinated to control the disease [30]. Furthermore, measuring herd effects is of great importance when trying to eradicate infectious diseases [30]. Therefore, designing a vaccine trial requires careful thought about how to ensure measurement of indirect effects without losing the ability to estimate unbiased direct causal effects. CRTs may well be the design of choice, for vaccine trials, for these reasons.

The causal mechanism for an intervention with spillover effects will be more complex. Exploring the type of spillover may help further unpack the mechanism of

an effect and help inform implementation of the intervention. Consider a study evaluating an intervention of insecticide spraying. If we are able to identify that a spillover effect is present, whereby individuals within 500 meters of the spraying receive a beneficial effect, then the mechanism may be distance-based. From this study, we could conclude that it would be of greater benefit to spray multiple locations with gaps of less than 500m compared to spraying one smaller contiguous area. In this situation, protection is provided to a larger region because of the additional knowledge of a spillover effect.

Alternatively, a spillover effect that is conditional on treatment density may require a certain level of coverage before spillover is present. An example is an insecticide spray whereby a level of 80% coverage is needed for spillover effects to be present. In this case, the implementors may decide to ignore the spillover effect, and focus on the direct benefit of the intervention, by spraying a larger area with lower coverage, instead of a smaller area with high coverage.

Spillover in CRTs creates an opportunity to further explore the underlying mechanism of the intervention, but they are not the only choice of trial design, that allow estimation of spillover effects [17, 40].

2.4.6. Design solutions

The chance of spillover between clusters will usually decrease when the geographical extent of a cluster increases, as travel across cluster borders becomes less likely and a lower proportion of individuals will live near a boundary [4]. Spillover due to spatial proximity can therefore be reduced by having large or well-separated clusters. The definition of a well-separated cluster will depend on the type of intervention and outcome, but in general refers to clusters that are a minimum distance apart and share no common boundaries.

One disadvantage of this design approach is that the study area for the trial can become very large. In addition, within-cluster variance may be increased, as the

chance systematic differences between participants of a cluster are more likely, if the cluster contains more people. Furthermore, for a given number of individuals, larger clusters leads to lower power compared to having more clusters with fewer individuals in each cluster [14]. Well-separated clusters may also limit the potential to measure spillover, as participants who might be subject to spatial spillover effects will not be near the boundaries of a cluster.

A similar approach which uses a buffer zone is called the *fried egg* design [4]. This design only compares participants from the centre of the clusters in the main trial analysis. This approach helps reduce the impact of edge effects and spillover, as individuals in the centre are less likely to move across cluster boundaries or be affected by those in neighbouring clusters. However, results will be biased if individuals living in the centre of a cluster are not representative of the rest of the cluster. Moreover, information is lost by only using the centre of the cluster.

Pseudo or double randomised trials are halfway between a CRT and an individually randomised trial. First, clusters are randomised to receive the intervention or control, and secondly, participants within each cluster are randomised to receive the intervention or control [41]. This allows the level of coverage for the intervention and control clusters to be varied and has been proposed as a method for measuring indirect effects and within-cluster spillover [33]. This provides considerable flexibility for estimating the effect of various levels of coverage. For instance, 80% of participants in intervention clusters may receive the intervention and 30% of participants in control clusters may receive the intervention.

In contrast to the previous designs which focused on reducing spillover, this approach can help to estimate it. Unfortunately, a double randomised approach negates the ethical and logistical convenience of a standard CRT and adds complexity to running the trial. Furthermore, they assume only small-scale spillover effects, and are not possible to conduct with interventions where you cannot precisely restrict who receives the intervention, such as with mass education programmes [33].

Stepped wedge CRTs [42] are an alternative form of CRT where all of the clusters eventually receive the intervention, and the randomisation focuses on the order in which cluster receive the intervention. This design has been proposed for measuring spillover in trials where the intervention disrupts the transmission of diseases [43]. This approach may allow for estimation of the spillover effects of an intervention. An advantage of this design is that it may gain greater acceptance within communities as no area is ultimately denied the intervention. One of the main challenges of this approach is the increased complexity of the analysis.

Torgerson found that in terms of total sample size, CRTs are more efficient than individually randomised trials when contamination is greater than 30%. This means more than 30% of the control participants receive or are exposed to the intervention [44]. In this context, Torgerson refers to individuals crossing over from the control arm to the intervention, which he states is similar to spillover [44]. Therefore, issues with spillover may be reduced by increasing sample size, without the need for greater complexity in design. Individually randomised trials tend to be easier to conduct, and typically have increased power compared to CRTs. However, this approach will not be possible when the intervention can only be given at a group level or when the level of spillover is unknown or suspected to be high. Furthermore, it is difficult to measure spillover effects in individually randomised trials [21].

Adaptations of the CRT can help avoid or measure spillover, but often add further complexity to running or analysing the trial. If a trial has already begun, or a reanalysis is intended, then it can be difficult or impossible to change the original trial design, for example to move clusters further apart. An alternative approach is to attempt to account for spillover in the analysis of a trial. The advantage of this approach is that it can be applied to existing trial data but it will add complexity to the analysis and may require additional data collection.

Accounting for spillovers that derive from mechanisms of physical proximity requires recording the location of individuals and the spatial definition of the clusters.

Global Positioning System (GPS) data are often collected as part of CRTs for trial management and the indexing of participants, in addition to providing an objective and absolute measure of location. A GPS coordinate provides a basis for measuring the physical proximity of individuals within a trial to a range of other spatially located features, including other individuals, cluster centres/edges, and environmental/structural resources, such as clinics, water sources, businesses etc. If collected, CRT participants are often related to a single coordinate, such as their household. This need not be the case and in future we may see the use of tracking devices to record richer information of the spatial location of individuals.

This PhD explores the utility of using spatial data and considers how spatial statistical analysis methods can use location information to measure and account for spatial spillovers in CRT.

2.5. Spatial data analysis

In this section, I review the key aspects of spatial data and spatial statistics, providing an overview of how these topics broadly fit in the context of CRTs established previously.

2.5.1. Spatial methods in epidemiology

The importance of location in epidemiology has a long history with probably the earliest example being John Snow's mapping of the 1854 Cholera outbreak in Soho [45]. Despite this early success, the use of spatial methods in epidemiology remains infrequent, even in infectious disease outbreaks [46]. Although it still represents a small amount, the use of spatial methods in the analyses of health data is increasing [47]. According to a recent review of spatial methods in epidemiology the most common methods were proximity calculations, spatial aggregation, cluster detection, spatial interpolation and smoothing, and spatial regression [47].

- **Proximity calculations:** measuring distances between a location and an exposure or resource. For example, the distance between a household and the nearest hospital.
- **Spatial aggregation:** aggregating features within a spatial area. For example, calculating the number of doctors per neighbourhood.
- **Cluster detection methods:** assessing non-random spatial patterns. These methods are often used to detect clusters of diseases and whether the patterns are non-random. These methods are split into global and local clustering. Global clustering looks for patterns but does not determine specific cluster location. Local clustering looks for hotspots of diseases. For example, assessing for areas of high prevalence of Zika virus.
- **Spatial interpolation and smoothing:** imputing or adjusting values in a location based on the values of nearby locations. These methods are mainly

used to fill in information where it is not observed or to improve the accuracy of disease rates. For example, smoothing disease incidence rates to avoid spike in prevalence due to lack of information.

- **Spatial regression:** incorporation of spatial structure into regression models. These methods extend standard statistical regression models and are used to account for spatial effects. For example, analysing the effect of the prevalence of Ebola in one region compared to the regions surrounding it.

2.5.2. Spatial data

Spatial data are observations to which labels have been added to show where the observations were collected [48]. Spatial data may contain information about attributes of interest and their locations. The location information can take different forms, for instance, an exact location such as the locations of types of shops, or a relative location such as the distance between London and Paris. Locations are commonly represented through coordinate systems, such as longitude and latitude [49] but could also be a named location such as Sweden or New York. The atomic form of geographic information, introduced by Goodchild et al. [50] posits that spatial data can be represented simply as a tuple:

$$\langle s, H, h(s) \rangle$$

In which, s is a location in space-time, H is an attribute or property, and $h(s)$ defines that given property for the given location. For example, the attribute may be temperature, and the location is Greenwich, London, then $h(s)$ represents the temperature in Greenwich. This ‘geo-atom’ can then be generalised to any of a number of conceptual models used to represent spatial data. The two most common models are known as the ‘*field*’ and ‘*entity*’ (or ‘*object*’) models [51].

In Geographical Information Systems (GIS), fields are more commonly known as

raster data and entities are referred to as vector data. Raster data are effectively lattices of geo-atoms, but are more commonly conceptualised as digital images (regular grids of cells) or as a collection of points arranged on a regular grid [49]. Raster tends to be used to represent continuously varying phenomena, such as elevation, types of land cover or rainfall. As raster data are effectively digital images, geographical scale is fixed to the resolution of the image. The finer the resolution, the smaller the area each pixel represents.

Vector data consists of the following spatial types:

- **Points:** a single point location such as a house. A single geo-atom.
- **Lines:** a collection of ordered points that are connected, such as a road or river. An ordered string of geo-atoms, in which properties or attributes are effectively constant for the aggregate object.
- **Polygons:** a series of connected points where the first and last point are in the same location, such as a lake or an administrative area. Effectively a closed line object. A polygon could contain holes.

Unlike the raster data model, vector data can define distinct or discontinuous objects in space with multiple attributes, for instance, the locations and characteristics of vaccinated adults in a city. Similarly, vectors do not have a fixed spatial scale, the lowest resolution is effectively set by the instrument of measurement. For instance, a GPS location may be accurate to 5-10 meters thus setting a lower spatial bound to inference but not an upper one. Observations in CRTs tend to relate to distinct locations in space rather than continuous surfaces and thus the vector data model will be used to describe the spatial representation of CRT data.

2.5.3. Challenges of Spatial data

2.5.3.1. Boundary problems or edge effects

Edge effects relate to how events outside of a boundary can impact the environment within the boundary. For example, the existence of a large polluting factory near to a town may affect the health of people who live there. Geographical clusters in CRTs are bounded and may be susceptible to edge effects, particularly if there are many small clusters [4]. The edge effects may be from outside the study area or if each cluster is considered as a separate boundary the edge effects could refer to spillover between clusters.

2.5.3.2. Modifiable areal unit problem

Areal data is spatial data that is collected or presented at an aggregated level, such as voting districts, states, or countries. The aggregation of information results in statistical bias stemming from how the aggregation is chosen. The modifiable areal unit problem (MAUP) arises from the arbitrary and (theoretically) modifiable choice of aggregation of the data [52]. For instance, when analysing data from an area such as London, information could be grouped at street level, by borough, or a by a newly defined group, such as by nearest park. The consequence of MAUP is that different choices of aggregation, can give different results. How clusters are defined in a CRT may impact the results of a trial. MAUP is more relevant when considering the design of a CRT, as during analysis it is unlikely the cluster definitions can be changed.

2.5.3.3. Spatial heterogeneity and dependence

Spatial heterogeneity (also called 1st order spatial effects) refers to the uneven distributions of values over space. This usually manifests as a patchy distribution of events or values over a broad area. It is usually symptomatic of an underlying spatial

process invalidating the assumption that data are independent. Spatial dependency (2nd order spatial effects) is a fundamental concept in spatial statistics [53] and stems from Tobler's 1st law of geography that '*everything is related to everything else, but near things are more related than distant things.*' [54]. Spatial dependency is frequently studied explicitly as a spatial autocorrelative process. There is a wide array of spatial correlation methodology [48, 53, 55] and a concise definition of this concept is provided by Hubert, Golledge, and Constanza [56]:

Given a set S containing n geographical units, spatial autocorrelation refers to the relationship between some variable observed in each of the n localities and a measure of geographical proximity defined for all $n(n-1)$ pairs chosen from n .

This definition describes the idea that spatial autocorrelation is a relationship based on proximity or some form of distance.

In agricultural field trials, it is long established that the location of the data can impact on trial results [57], and incorporation of spatial methodology is common [58]. In contrast, the impact of spatial effects in human CRTs have received little focus. The presence of spillover effects may manifest itself as spatially autocorrelated data, and therefore methods to deal with spatial correlation could be useful for accounting for spillover effects in CRTs. Spatial dependency may also be a cause of spillover effects in a trial and could bias results.

2.5.4. Spatial statistics

Spatial statistics is a field of research which takes into account the location of data, and it is based on the non-independence of observations; that is the assumption that nearby units are in some way associated [54, 59]. If the observations were spatially independent, then the location of the data would not matter and there would be no need to consider this in the analysis. Spatial statistics is traditionally categorised into three different fields, which are [53]:

- **Point Process** - analysis of data that consists of a set of point locations where events of interest have occurred, such as cases of cholera.
- **Geostatistical** - analysis of data that consists of point samples from a (conceptually) continuous distribution in space, such as temperature over an entire city.
- **Areal** - analysis of data that consists of values from set of areas which may be a regular lattice or an irregular set of areas, such as voting areas, districts, or states.

Trials of infectious diseases might be treated as point process data, but typically these types of spatial analyses aim to identify whether the pattern of disease is random or clustered [53]. The important variable to be analysed is the location of the event and typically in a CRT the locations where data will be recorded are fixed before the study begins. Therefore, the selection of sites may introduce biases which affect point process analyses. Furthermore, treating CRT data as a point process may also ignore the clustering which is introduced via the design of a CRT.

An alternative could be to treat the individual observations of a CRT as a set of points from a continuous distribution in space. However, geostatistical methods are predominantly focused on predicting the values in unmeasured locations [60]. This has obvious value when trying to predict where oil is, but is probably of less use for trials where we aim to estimate and establish the causal effectiveness of interventions..

CRTs appear to fit most naturally into an areal structure, especially for geographical trials where the observations are points, and the clusters are polygons. Aggregating the observations at a cluster level would generate areal data as it performed in the two-stage analysis approach. However, this would result in a loss of information and may hinder analysis of spillover. When the number of clusters is small, the two-stage approach is preferred, as multi-level models can provide unreliable results. Furthermore, this leads to fewer observations (one per cluster) and, therefore, it may be difficult to fit spatial models to the data, .

Based on these three categories it would appear the spatial structure of CRTs does

not fit neatly into any of the three major spatial statistics fields. The spatial analysis of CRTs will thus require applying spatial methods to data which differs from the original intended use of the methods. Areal analysis methods could be applied to individual point level data instead of aggregated areas. Geostatistical methods might be applied to a discrete spatial phenomenon, which is grouped into dependencies, as opposed to the typical continuous spatial phenomenon over an area.

2.6. Spatial modelling

When spatial correlation is present then nearby observations are related to one another, often manifesting as the presence of spatially correlated residuals in regression models. This violates the assumption of an independent error term in linear regression. The mechanism through which spatial correlation occurs is rarely observable, so spatial models often use latent variables and structures to incorporate spatial effects [61]. The standard approach is to use a mathematical representation of the spatial structure of the data, termed a spatial weights matrix.

2.6.1. Spatial weights matrix

To describe spatial weights matrices we will first consider their simplest form: binary weights. For n observations a spatial weights matrix W is an $n \times n$ matrix, where for every observation k with a set of neighbours m then $W_{mk} = \begin{cases} 1 & \text{if } m \in m^* \\ 0 & \text{if } m \notin m^* \end{cases}$ and $k \neq m$. In this form, two observations are related if and only if they are neighbours, and all neighbours have the same weight of relationship regardless of proximity. Generally, this W matrix is symmetric about the main diagonal such that $W_{mk} = W_{km}$, however, this will not be the case for all matrices. This type of matrix is also termed an adjacency matrix and is typically used when analysing areal data.

For point data, a spatial weights matrix can also be defined with a distance threshold that indicates which points are neighbours (effectively a fixed buffer), or a specified number of nearest neighbours could be used to represent who is related (effectively a type of variable buffer). Alternatively, the cells of the matrix can store values that weight connectivity on a continuous rather than binary basis. A distance weighting is typically used so that observations that are further apart are less connected as a function of distance. Distance is often transformed to be the inverse of a function such as a power law.

The choice of spatial weights matrix can have a large impact on the analysis and the choice of matrix form is an area of active research [62].

2.6.2. Distance weighting functions

Distance weighting functions (also known as kernels or covariance functions) are functions that map a positive distance value to an inverse distance weight. A number of functions exist to reflect possible relationships over distance, but no specification is *a priori* better than any other. Four distance weighting functions are detailed in Table 2.2.

Table 2.2. Overview of distance weighting functions

Name	Formula	Examples
Linear	$f(d) = \begin{cases} \frac{\omega-d}{\omega} & \text{if } d < \omega \\ 0 & \text{if } d \geq \omega \end{cases}$ <p>Where ω is that maximum distance that the spatial effect is present.</p>	
Inverse distance	$f(d) = \frac{\xi}{d^\omega}$ <p>Where ω and ξ are constants.</p>	<p>$\xi = 1$</p>
Exponential	$f(d) = \xi e^{-\omega d}$ <p>Where ω and ξ are constants.</p>	<p>$\xi = 1$</p>
Gaussian	$f(d) = \xi e^{-\frac{(d-\mu)^2}{\sigma^2}}$ <p>Where μ and σ^2 are the mean and variance of a normal distribution and ξ is a constant.</p>	<p>$\mu = 0$</p>

* d is distance or some measure of proximity.

2.6.3. Extending linear regression

A standard linear regression model can take the form

$$y = Z_l \gamma_l + \epsilon$$

where y is a vector representing the outcome variable, Z_l is a design matrix of l predictor variables, with the first column usually constrained to be a vector of ones, γ_l is a vector of coefficients for Z_l and ϵ is an error term which is typically assumed independent and identically distributed (i.d.d) as $\epsilon \sim N(0, \sigma^2)$ [63]. Generalised least squares (GLS) extends linear regression by relaxing the assumption of the independent error term. GLS is a technique that allows for correlation in the residuals of the model and this can be used to account for spatial correlation [64]. This is achieved by redefining the error term in the linear regression equation as $\epsilon \sim N(0, \sigma^2 \Sigma)$, where Σ is a variance-covariance matrix representing the dependency in the error term [65]. GLS is commonly used for times series analyses with an autoregressive lag structure relating the present value to values from previous time periods [66]. The dependency structure can be extended from one dimension (time) to two or more dimensions (space) by using a spatially weighted covariance structure. This approach is the basis for many spatial models, where correlations in the model are assumed to be a function of a spatial structure.

2.6.4. SAR and CAR models

Two common approaches for constructing the spatial covariance structure are the simultaneous autoregressive (SAR) and the conditional autoregressive (CAR) [61]. Taking the linear regression model above $y = Z\gamma + \epsilon$ with $\epsilon \sim MVN(0, \sigma^2 \Sigma)$. Defining $\Sigma = (I - \rho_1 W_1)^{-1} (I - \rho_1 W_1^t)^{-1}$ gives a SAR model and defining $\Sigma = (I - \rho_2 W_2)^{-1}$ gives a CAR model, where W is a spatial weight matrix, and ρ is a parameter representing the level of spatial correlation. The SAR and CAR models are equivalent

when $(I - \rho_1 W_1)^{-1}(I - \rho_1 W_1^t)^{-1} = (I - \rho_2 W_2)^{-1}$.

The SAR and CAR models are a family of spatial models that can be fitted using maximum likelihood. The SAR model was introduced by Whittle [67] and is commonly used in spatial econometrics [68]. CAR models were introduced by Besag [69], and provide the basis for a wide range of spatial models with diverse applications such as disease mapping [70], image restoration [71], and machine learning [72]. The SAR and CAR models are a subset of Markov random fields (MRF) [73, 74]. An in depth look at the connections between spatial models and random fields is presented in appendix C.

2.6.5. Spatial filtering

A spatial filtering model, attempts to translate the spatial structure of the data into explanatory variables. It involves eigenfunction decomposition of a spatial weights matrix and hence calculation of the eigenvectors. It is similar to the use of principal components in a regression model as a dimensionality reduction method. However, instead of the principal components, the eigenvectors themselves are used. The eigenvectors are chosen through stepwise model selection, so as to include those eigenvectors which most reduce spatial correlation in the error term [75]. Spatial filtering specifically aims to remove residual spatial autocorrelation by the addition of covariates [65].

Spatial filtering provide a mechanism for removal of spatial variation, and this model has also been applied in a multi-level setting [76]. Potentially it might provide a basis through which spatial autocorrelation can be removed during the analysis of a CRT whilst also accounting for the dependency due to the trial design. However, it does not allow for estimating of spatial effects or spillover effects, and therefore, spatial filtering methods will not be explored in this thesis.

2.6.6. Geographically weighted regression

Geographically weighted regression (GWR) is a spatial regression method that allows the coefficients of the model to vary spatially [77]. The model can be used to explore local spatial patterns in data. This is achieved by repeatedly applying a regression model to a locally weighted spatial subsets of the data. As per Brunson & Comber [78], the GWR can be represented as

$$y_s = \alpha(a_s, b_s) + Z_l \gamma_l(a_s, b_s) + \epsilon$$

where y_s is a measure attribute at location s , (a_s, b_s) are the coordinates of this location, Z_l is a matrix of l predictor variables associated with location s and α and γ_l are coefficients that are a function of the coordinates. This method involves applying a regression model to a spatial subset of the data (a neighbourhood). The coefficients of the model are then recorded for that subset. A different neighbourhood is then chosen, and the model reapplied. This process is repeated over the study area to estimate the coefficient for each neighbourhood. The neighbourhood is typically a radius around each point and can be fixed or vary for each iteration. The distribution of the coefficients can then be explored visually and through summary statistics to help to determine sources of heterogeneity in the data.

GWR could be used to explore heterogeneity in intervention estimates, it may help to assess locations where the intervention was less effective, or more effective. A map of the GWR estimates could be combined, with a map of the intervention allocation, and outcome, to explore how the different characteristics interact with one another. For example, the intervention may be more effective in areas of high prevalence of the outcome, suggesting that a low number of events would result in an underestimate of the true intervention effect. Alternatively, the intervention may be less effective when many intervention cluster are nearby, which could be suggestive of a spillover effect.

2.7. Linking spatial data types with CRTs

To help determine the types of spatial analyses that may be appropriate for CRTs, it is useful to represent CRTs within a typology of spatial data. In this section, I propose a conceptual model for linking spatial data to CRTs. I posit that through only classifying the spatial type of the cluster and observation levels, a wide range of CRTs can be formed. Furthermore, consideration of the spatial structure helps elicit which spatial measurements have meaning in different trial settings.

2.7.1. Types of spatial data in CRTs

The clusters and observations in a CRT can be conceptualised as point, line, or polygon features. The nature of the clusters and participants in a CRT may mean that their respective representation is either equivalent or different. In a school based CRT, the location of the clusters (school) and participants (residential location) can both be represented by a point. By comparison, a primary care trial may represent participants as individual residential locations, but define clusters as polygons, covering a predefined administrative zone.

In CRTs, clusters are often defined geographically, lending themselves to spatial representation. In contrast, the participants or observations may not have a spatial representation. For instance, there is no single way to spatially reference a child's exam score, in these situations the observations would most likely be linked to the location of the child's home but could equally be justified as linked to the child's school location. The distinction between these two locations is not obvious, and it is often unclear *a priori* which is more desirable from an analysis perspective.

The cluster and observation can be a point, line, or polygon, and need not be in the same location. The observation locations may be contained within the clusters or they may lie in practice outside of a cluster. There are 18 ways in which spatial data may describe CRTs, given in Table 2.3.

Table 2.3. Eighteen possible types of spatial cluster randomised trials

Observation Type	Cluster Type	Same or different location
Point	Point	Same
	Point	Different
	Line	Same
	Line	Different
	Polygon	Same
	Polygon	Different
Line	Point	Same
	Point	Different
	Line	Same
	Line	Different
	Polygon	Same
	Polygon	Different
Polygon	Point	Same
	Point	Different
	Line	Same
	Line	Different
	Polygon	Same
	Polygon	Different

Restricting observations (or participants) to points, and clusters to points or polygons allows for the spatial representation of most CRTs. Therefore, in practice, only the four combinations given in table 2.4 are required to represent almost all CRTs.

Typically, the analysis models for the different spatial combinations of CRTs are identical, the observations are labelled according to clusters and then a model is fitted that accounts for the dependency within clusters. Explicit consideration of the spatial types that make up a CRT helps to elicit the types of processes that may be occurring during a trial. Furthermore, it supports the evaluation of which spatial metrics may be meaningful for different trials. Although CRT share similarities in design, from a spatial perspective they can have very different structures.

Table 2.4. Four common types of spatial cluster randomised trials with examples

Spatial Type			Example		
Cluster	Observation	Location	Trial	Observation	Cluster
Point	Point	Same	Household	House	House
Point	Point	Different	School	House	School
Polygon	Point	Same	Geographical	House	Surrounding area
Polygon	Point	Different	Primary Care	House	Primary care area

2.7.1.1. Clusters as point features

Clusters that are of a similar scale to the observations themselves can be generalised and represented as points. Such clusters may include the location of households, schools, or workplaces where the cluster can be conceptualised as a destination for individuals. The clusters may contain multiple observations per location, such as several members of a household. When clusters and observations are both represented by a point, the trial is most likely a household, workplace or school based CRT.

A household CRT is an example of a trial where the clusters and observations are points that may share the same location. The clusters are the houses and the observations are recorded from the people who live in the houses. In a household CRT it may not be possible to differentiate between the location of the cluster and the observations. This may lead to difficulties when measuring spatial spillover effects because the measured distance within a cluster is zero. Thus, there are no participants within a cluster who are closer to the intervention compared to others. Spillovers in this context are potentially more associated with daily activities and spatial interactions based on individual mobility, which is challenging to capture and analyse.

Spatial spillover effects can be calculated at a cluster level in a household CRT. Intervention households could be treated as a point source and then distance to nearest intervention households could be analysed for control clusters only. The

estimate of spillover is clearly not a randomised comparison and could be confounded by other factors. Furthermore, this approach aggregates the spatial extents of all observations to the cluster level which may not represent the spatial effects at an individual level. We have no way to distinguish between a control household where all individuals are affected by spillover effects and control households where one person was initially affected by spillover and then affected other people in their household. Therefore, estimating spatial spillovers in household CRTs may conflate between-house spillover and within-house spillover effects.

A school based CRT is an example of a trial where the clusters and observations are points that can be in distinct locations. The clusters are the schools and the observations could be exam performance of the children within the schools. When only school locations are present, then this trial becomes identical to the household CRT in terms of spatial information. When spatial information on the observations such as the location of a child's home is collected, then further spatial analyses are possible.

This make it possible to distinguish between the locations of observations in a school trial. There will be individuals who live closer to schools and individuals who live closer to intervention participants. Therefore, spatial spillover effects can be measured at an individual level. Although it is likely that siblings live in the same house and attend the same school, conflation of within- and between-cluster spillover is unlikely unless there are multiple clusters within each school (i.e. where clusters are classrooms).

A school based trial offers different spatial information and it is possible to calculate the effect of distance to nearest school, distance to nearest intervention school, and distance to nearest intervention household. The measurement of interest will depend on the mechanism of the spillover. If the intervention in the school is expected to impact on the community surrounding the school then distance to nearest intervention school will be of interest. If it is suspected that the intervention pro-

vided at school may provide a protective effect to people who live with or near the participants, then distance to nearest intervention household would be meaningful.

2.7.1.2. Clusters as polygon features

Larger clusters may be conceptualised as polygons in cases such as administrative zones, primary care trust areas, or other bounded geographical areas.

Geographical CRTs are an example of trials where the clusters are polygons and the observations are points contained within the polygons. The clusters are the geographical areas and observations may relate to the participant's household located within the cluster. In this type of trial, the observations may have a single membership (only related to one cluster) or multi-membership (related to multiple clusters). In a geographical CRT, it is possible to distinguish between the location of the observations and therefore, cluster level and individual level spatial effects can be measured in these types of CRTs.

Cluster level spatial effects can be estimated by aggregating the data and fitting a spatial model. A spatial model could also be fitted on the individual level data using a random effect that connects the clusters based on spatial contiguity. Individual level spatial spillover could also be analysed ignoring the clusters. Cluster boundaries could also be used to estimate the effects of spillover. The cluster boundaries and distance to nearest intervention household create a way of separating individuals due to proximity of exposure.

Multi-membership can occur in Primary Care trials. When the clusters are polygons, and the participant's location is a point, it could fall within a cluster where they are receiving treatment/care, or could reside in the cluster they live in. These location need not be the same, with individuals receiving treatment in a different cluster from where they live. When trial participants reside out of the cluster the observations can be a member of more than one cluster referred to a multi-membership [79].

In this context, a control observation may be spatially close to intervention observations meaning they live in a control cluster. However they may be receiving treatment outside of the cluster they reside in and from a trial perspective they are part of another cluster. This would mean a standard CRT analysis would ignore that the observation will be correlated with individuals they live near but also correlated with individuals where they receive their primary care. The multi-membership may have implications for the analysis of intervention, and spillover effects.

In this setting, distance to nearest intervention observation or distance to nearest intervention area is possibly not a meaningful metric and a more appropriate analysis strategy might be to assess the multiple memberships of the participants.

Examining the spatial representation of CRTs exposes how trials which are similar in design, and often analysed in the same way, could be conceptualised as having distinct spatial structures. This is motivated by different underlying hypotheses as to the nature of the spillover mechanisms. Unless the design is adapted, then all estimates of spatial spillover are non-randomised comparisons. The spatial context of a CRT provides a way of approximating the proximity of individuals to different exposures and may provide a way of measuring spillover effects. A standard CRT analysis would treat a household CRT in the same way as a workplace CRT. Thus, estimating spatial effects requires careful thought about which metrics are meaningful.

2.7.2. Spatial effect modification of the intervention

When a spillover is present, this could manifest as the presence of interaction between proximity and the intervention. For example, consider a CRT of insecticide treated bed nets where a distance-based positive spillover effect is present over a distance of 200m. Then intervention effect will be reduced, when comparing the intervention arm to control participants within 200m of the intervention. The intervention effect will be larger when comparing the intervention arm to control participants further

than 200m from the intervention. In this scenario, distance modifies the effect of the intervention.

Alternatively the intervention status may modify the effect of proximity. Using the bed net CRT example, the spillover direction is from the intervention towards control clusters. Therefore proximity to the intervention will have an effect, but proximity to a control household should not have an effect.

Recording measurements of proximity such as GPS coordinates, allows the investigator to describe properties of the mechanism behind the intervention effect. Thus, helping to further understand how the intervention works.

2.8. Summary

This chapter provided an overview of the main concepts of CRTs and spatial analysis, setting forth a framework for spatial data in the context of CRTs.

In comparison to individually randomised trials, CRTs are a good choice of design to reduce contamination between control and intervention participants. Spillover effects may be present in CRTs, especially when clusters are near to each other. Although there are several design approaches which can be used to eliminate and account for spillover effects, these can result in increased complexity and may not allow spillover effect to be measured. A potential way of assessing spillover in the analysis of CRTs is to use spatial methods. This thesis focuses on exploring different approaches to incorporate spatial effects into the analyses of CRTs.

I have outlined a relationship between spatial data types and CRTs and determined that most CRTs can be represented by treating clusters as points or polygons and observations as points. CRTs do not fit neatly into the traditional fields of spatial statistics (1. Point process, 2. Geostatistical, 3. Areal). This poses further difficulties for spatial analysis of CRTs

There are several aspects of spatial data such as MAUP, and spatial correlation

that are of relevance for CRTs. MAUP is of greater importance in the design stage, and therefore, will not be addressed by this thesis. Spatial correlation is typically included through modelling approaches. The extension of linear regression to incorporate spatial correlation has been demonstrated by inclusion of a spatial weighted variance covariance matrix. Two of the main types of models (SAR and CAR) were described, and these and related spatial models will be revisited in subsequent chapters.

The discussion of spatial analysis in CRTs has so far been abstract, attempting to bring together different concepts from spatial statistics and trial methodology. In the next chapter (3) examples of spatial methods in CRTs are identified by conducting a systematic review of spatial analysis methods in CRTs.

2.9. Bibliography

- [1] Wang, D and Bakhai, A. Clinical Trials A Practical Guide to Design, Analysis, and Reporting. Remedica, 2006.
- [2] Pocock, SJ. Clinical Trials A Practical Approach. West Sussex: John Wiley & Sons Ltd, 1983.
- [3] Burns, PB, Rohrich, RJ, and Chung, KC. The Levels of Evidence and their role in Evidence-Based Medicine. *Plast Reconstr Surg* 2011 128. (1) 305–310.
- [4] Hayes, RJ and Moulton, LH. Cluster randomised trials. 2nd. Boca Raton: Chapman & Hall/ CRC, 2017.
- [5] Schulz, KF, Altman, DG, Moher, D, and Group, C. CONSORT 2010 Statement : Updated Guidelines for Reporting Parallel Group Randomized Trials OF TO. *Annals of Internal Medicine* 2010 152. (11) 1–8.
- [6] World Medical Association. World Medical Association Declaration of Helsinki. *Bulletin of the world health organization*. 2001 79. (4) 373–374.
- [7] Campbell, MK, Piaggio, G, Elbourne, DR, and Altman, DG. Consort 2010 statement: extension to cluster randomised trials. *Bmj* 2012 345 e5661–e5661.
- [8] Lind, J. A Treatise of the Scurvy. Edinburgh: Sands Murray & Cochran, 1753.
- [9] Medical Research Council. Streptomycin Treatment of Pulmonary Tuberculosis: A Medical Research Council Investigation. *BMJ* 1948 2 769–782.
- [10] Peace, KE and Chen, DG. Clinical Trial Methodology. Chapman & Hall/CRC Biostatistics Series. CRC Press, 2010.
- [11] Kirkwood, B and Sterne, J. Essential Medical Statistics. Essentials. Wiley, 2003.

- [12] Campbell, M, Fitzpatrick, R, Haines, A, Kinmonth, AL, Sandercock, P, Spiegelhalter, D, and Tyrer, P. Framework for design and evaluation of complex interventions to improve health Framework for trials of complex interventions. *British Medical Journal* 2000 321. (7262) 694–6.
- [13] Eldridge, S and Kerry, S. A Practical Guide to Cluster Randomised Trials in Health Research. Cornwall: John Wiley & Sons Ltd, 2012.
- [14] Campbell, MJ and Walters, SJ. How to design, analyse and report cluster randomised trials in medicine and health related research. Singapore: John Wiley & Sons Ltd, 2014.
- [15] Thomson, KA, Haberer, JE, Marzinke, MA, Mujugira, A, Hendrix, CW, Celum, C, Ndase, P, Ronald, A, Bangsberg, DR, and Baeten, JM. Brief Report: Medication Sharing is Rare among African HIV-1 Serodiscordant Couples Enrolled in an Efficacy Trial of Oral Pre-exposure Prophylaxis (PrEP) for HIV-1 Prevention. *Journal of Acquired Immune Deficiency Syndromes* 2017 75. (2) 184–189.
- [16] Wang, X. Spatial Gaussian Markov Random Fields: Modelling, Applications and Efficient Computations. *Journal of Biometrics & Biostatistics* 2013 05. (02) 2–5.
- [17] Halloran, ME, Longini, IM, and Struchiner, CJ. Design and Analysis of Vaccine Studies. Statistics for Biology and Health. Springer New York, 2009.
- [18] Hayes, RJ, Alexander, ND, Bennett, S, and Cousens, SN. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Statistical methods in medical research* 2000 9 95–116.
- [19] Benjamin-chung, J, Berger, D, Clark, A, Falcao, L, Tran, D, Hubbard, A, Miguel, E, and Colford, J. The identification and measurement of health-related spillovers in impact evaluations : a systematic review. Tech. rep. London, 2015.

- [20] Vanderweele, TJ, Hong, G, Jones, SM, and Brown, JL. Mediation and spillover effects in group-randomized trials: a case study of the 4Rs educational intervention. *Journal of the American Statistical Association* 2013 108. (502) 469–482. arXiv: NIHMS150003.
- [21] Benjamin-Chung, J, Arnold, BF, Berger, D, Luby, SP, Miguel, E, Colford Jr, JM, and Hubbard, AE. Spillover effects in epidemiology: parameters, study designs and methodological considerations. *International Journal of Epidemiology* 2017. (March) 332–347.
- [22] Benjamin-Chung, J, Abedin, J, Berger, D, Clark, A, Jimenez, V, Konagaya, E, Tran, D, Arnold, BF, Hubbard, AE, Luby, SP, Miguel, E, and Colford, JM. Spillover effects on health outcomes in low-and middle-income countries: A systematic review. *International Journal of Epidemiology* 2017 46. (4) 1251–1276.
- [23] Klöckner, CA, Nayum, A, and Mehmetoglu, M. Positive and negative spillover effects from electric car purchase to car use. *Transportation Research Part D: Transport and Environment* 2013 21 32–38. arXiv: 1405.6230.
- [24] Truelove, HB, Carrico, AR, Weber, EU, Raimi, KT, and Vandenberg, MP. Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental Change* 2014 29 127–138. URL: <http://dx.doi.org/10.1016/j.gloenvcha.2014.09.004>.
- [25] Blackwell, N, Oumarou, M, Dembelé, Y, Almou, I, and Illa, H. The FEAST trial of fluid bolus in African children with severe infection. *The Lancet* 2012 379. (9816) 614.
- [26] Todd, J, Heyderman, RS, Musoke, P, and Peto, T. When enough is enough: How the decision was made to stop the FEAST trial: Data and safety monitoring in an African trial of Fluid Expansion As Supportive Therapy (FEAST) for critically ill children. *Trials* 2013 14. (1) 1.

- [27] Levin, SA, Carpenter, SR, Godfray, HCJ, Kinzig, AP, Loreau, M, Losos, JB, Walker, B, and Wilcove, DS. *The Princeton Guide to Ecology*. Princeton University Press, 2009.
- [28] Rogerson, PA. *Statistical Methods for Geography: A Student's Guide*. SAGE Publications, 2006.
- [29] Buchanan, JM and Stubblebine, WC. Externality. *Economica* 1962 29. (116) 371–384.
- [30] Fine, P, Eames, K, and Heymann, DL. "Herd immunity": A rough guide. *Clinical Infectious Diseases* 2011 52. (7) 911–916.
- [31] Rosenbaum, PR. Interference Between Units in Randomized Experiments. *Journal of the American Statistical Association* 2007 102 191–200.
- [32] VanderWeele, TJ, Tchetgen, EJT, and Halloran, ME. Interference and Sensitivity Analysis. *Statistical Science* 2014 29. (4) 687–706.
- [33] Hudgens, MG and Halloran, ME. Toward Causal Inference With Interference. *American Statistician* 2008 103. (482) 832–842.
- [34] Tchetgen, EJ and Vanderweele, TJ. On causal inference in the presence of interference. *Statistical Methods in Medical Research* 2012 21. (1) 55–75. arXiv: NIHMS150003.
- [35] Angrist, JD. The perils of peer effects. *Labour Economics* 2014 30 98–108. arXiv: arXiv:1011.1669v3.
- [36] Griffith, D. Towards a theory of spatial statistics. *Geographical Analysis* 1980 12. (4) 325–329.
- [37] Rubin, DB. Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology* 1974 66. (5) 688–701.
- [38] Armbruster, B and Brandeau, ML. Contact tracing to control infectious disease: When enough is enough. *Health Care Management Science* 2007 10. (4) 341–355.

- [39] Chao, DL, Park, JK, Marks, F, Ochiai, RL, Longini, IM, and Halloran, ME. The contribution of neighbours to an individual's risk of typhoid outcome. *Epidemiology and infection* 2015 143. (16) 3520–7.
- [40] Vaucher, P. Designing phase III or IV trials for vaccines: Choosing between individual or cluster randomised trial designs. *Vaccine* 2009 27. (13) 1928–1931.
- [41] Borm, GF, Melis, RJ, Teerenstra, S, and Peer, PG. Pseudo cluster randomization: A treatment allocation method to minimize contamination and selection bias. *Statistics in Medicine* 2005 24. (23) 3535–3547.
- [42] Hemming, K, Haines, TP, Chilton, PJ, Girling, aJ, and Lilford, RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ* 2015 350 h391–h391.
- [43] Silkey, M, Homan, T, Maire, N, Hiscox, A, Mukabana, R, Takken, W, and Smith, TA. Design of trials for interrupting the transmission of endemic pathogens. *Trials* 2016 17 278.
- [44] Torgerson, DJ. Contamination in trials: is cluster randomisation the answer? *BMJ (Clinical research ed.)* 2001 322 355–357.
- [45] Snow, J. On the mode of transmission of cholera. 2nd ed. London: John Churchill, Princes Street, Soho, 1855:162.
- [46] Smith, CM, Le Comber, SC, Fry, H, Bull, M, Leach, S, and Hayward, AC. Spatial methods for infectious disease outbreak investigations: systematic literature review. *Eurosurveillance* 2015 20. (39).
- [47] Auchincloss, AH, Gebreab, SY, Mair, C, and Diez Roux, AV. A review of spatial methods in epidemiology, 2000–2010. *Annual review of public health* 2012 33 107–22.
- [48] Cressie, N and Wikle, CK. *Statistics for Spatio-Temporal Data*. Wiley, 2011.

- [49] Bivand, RS, Pebesma, E, and Gomez-Rubio, V. Applied Spatial Data Analysis with R. 2nd. New York: Springer, 2013.
- [50] Goodchild, MF, Yuan, M, and Cova, TJ. Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science* 2007 21. (3) 239–260.
- [51] Pfeiffer, D, Robinson, TP, Stevenson, M, Stevens, KB, Rogers, DJ, and Clements, AC. Spatial Analysis in Epidemiology. OUP Oxford, 2008.
- [52] Openshaw, S. The modifiable area unit problem. *Concepts and Techniques in Modern Geography* 1984 38 1–41.
- [53] Cressie, N. Statistics for Spatial Data. United States of America: John Wiley & Sons Ltd, 1991.
- [54] Tobler, AWR. A Computer Movie Simulation Urban Growth in Detroit Region. *Economic Geography* 1970 46. (332) 234–240.
- [55] Getis, A. A history of the concept of spatial autocorrelation: A geographer’s perspective. *Geographical Analysis* 2008 40. (3) 297–309.
- [56] Hubert, LJ, Golledge, RG, and Costanzo, CM. Generalized Procedures for Evaluating Spatial Autocorrelation. *Geographical Analysis* 1981 13. (3) 224–233.
- [57] Wiebe, GA. Variation and Correlation in Grain Yield among 1,500 Wheat Nursery Plots. *Journal of Agricultural Research* 1935 50. (4) 331–357.
- [58] Besag, J and Kempton, R. Statistical Analysis of Field Experiments Using Neighbouring Plots. *Biometrics* 1986 42. (2) 231–251.
- [59] Getis, A. Spatial statistics. *Geographical Information Systems Volume 1*. 2nd. John Wiley & Sons Ltd, 2005. Chap. 16239–252.
- [60] Diggle, PJ, Tawn, JA, and Moyeed, RA. Model-based geostatistics. *Journal of the Royal Statistical Society: Series C* 1998 47. (3) 299–350. arXiv: MR1626544.

- [61] Gelfand, A, Diggle, P, Fuentes, M, and Guttorp, P, eds. Handbook of Spatial Statistics. 2010.
- [62] Duncan, EW, White, NM, and Mengersen, K. Spatial smoothing in Bayesian models: a comparison of weights matrix specifications and their impact on inference. *International Journal of Health Geographics* 2017 16. (1) 47.
- [63] Bailey, TC and Gatrell, AC. Interactive Spatial Data Analysis. Longman, 1995.
- [64] Aitken, AC. IV. On Least Squares and Linear Combination of Observations. *Proceedings of the Royal Society of Edinburgh* 1936 55 42–48.
- [65] F. Dormann, C, M. McPherson, J, B. Araújo, M, Bivand, R, Bolliger, J, Carl, G, G. Davies, R, Hirzel, A, Jetz, W, Daniel Kissling, W, Kühn, I, Ohlemüller, R, R. Peres-Neto, P, Reineking, B, Schröder, B, M. Schurr, F, and Wilson, R. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography* 2007 30. (5) 609–628.
- [66] Chatfield, C. The Analysis of Time Series: An Introduction, Sixth Edition. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, 1996.
- [67] Whittle, P. On Stationary Processes in the Plane. *Biometrika* 1954 41. (3) 434–449.
- [68] Anselin, L. Spatial econometrics. *A companion to theoretical econometrics*. Ed. by Baltag, BH. Blackwell, 2001 310–330.
- [69] Besag, J. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B* 1974 36. (2) 192–236.
- [70] Lawson, A. Bayesian Disease Mapping. CRC Press, 2013.
- [71] Besag, J, York, J, and Mollié, A. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 1991 43. (1) 1–20.
- [72] Barber, D. Bayesian Reasoning and Machine Learning. 2011.

- [73] Rue, H and Held, L. Gaussian Markov Random Fields: Theory and Applications. Monographs on Statistics and Applied Probability. London: Chapman & Hall, 2005.
- [74] Arab, A, Hooten, MB, and Wikle, CK. Hierarchical Spatial Models. *Encyclopedia of GIS*. Ed. by Shekhar, S and Xion, H. Springer, 2008.
- [75] Griffith, D. A linear regression solution to the spatial autocorrelation problem. *Journal of Geographical Systems* 2000 2 141–156.
- [76] Park, YM and Kim, Y. A spatially filtered multilevel model to account for spatial dependency: application to self-rated health status in South Korea. *International journal of health geographics* 2014 13. (1) 6.
- [77] Brunson, C, Fotheringham, aS, and Charlton, ME. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis* 1996 28. (4) 281–298.
- [78] Brunson, C and Comber, L. An Introduction to R for Spatial Analysis and Mapping. SAGE Publications, 2015.
- [79] Silcocks, P and Kendrick, D. Spatial effects should be allowed for in primary care and other community-based cluster RCTS. *Trials* Jan. 2010 11. (1) 55.

3. Paper A: Spatial analysis of cluster randomised trials: A systematic review of analysis methods

Christopher Jarvis^{1,2}, Gian Luca Di Tanna³, Daniel Lewis¹, Neal Alexander¹,
W John Edmunds¹

1. London School of Hygiene and Tropical Medicine, London, UK
2. MRC London Hub for Trials Methodology Research, London, UK
3. Queen Mary University of London, London, UK

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Christopher Jarvis
Principal Supervisor	John Edmunds
Thesis Title	Spatial analysis of cluster randomised trials: A systematic review of analysis methods

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	Emerging Themes in Epidemiology		
When was the work published?	September 2017		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I defined the search strategy, performed the search and screening of articles, and drafted the article with feedback from co-authors.
--	---

Student Signature: _____

Date: 28/3/2018

Supervisor Signature: _____

Date: 28/3/2018

Improving health worldwide

www.lshtm.ac.uk

3.1. Overview

In Chapter 2, cluster randomised trials (CRTs) and spatial statistics were reviewed, and a framework for considering spatial data in CRTs proposed. In this chapter, I conduct a systematic review to determine to what extent the methods of spatial statistics have been incorporated into the analyses of CRTs. This paper was published in September 2017 in *Emerging Themes in Epidemiology*.

Objective

2. Describe and identify spatial analysis methods that have been previously used in CRTs by conducting a systematic review.

3.2. Role of candidate

I defined the search strategy with feedback from Gian Luca Di Tanna (GLDT) and Daniel Lewis (DL). I performed the search and screening of articles with GLDT as second reviewer. I drafted the initial article and revisions were made with feedback, input, and guidance from GLDT, DL, Neal Alexander, and W John Edmunds.

3.3. Abstract

Background

Cluster randomised trials (CRTs) often use geographical areas as the unit of randomisation, however explicit consideration of the location and spatial distribution of observations is rare. In many trials, the location of participants will have little importance, however in some, especially against infectious diseases, spillover effects due to participants being located close together may affect trial results. This review aims to identify spatial analysis methods used in CRTs and improve understanding of the impact of spatial effects on trial results.

Methods

A systematic review of CRTs containing spatial methods, defined as a method that accounts for the structure, location, or relative distances between observations. We searched three sources: Ovid/Medline, Pubmed, and Web of Science databases. Spatial methods were categorised and details of the impact of spatial effects on trial results recorded.

Results

We identified ten papers which met the inclusion criteria, comprising thirteen trials. We found that existing approaches fell into two categories; spatial variables and spatial modelling. The spatial variable approach was most common and involved standard statistical analysis of distance measurements. Spatial modelling is a more sophisticated approach which incorporates the spatial structure of the data within a random effects model. Studies tended to demonstrate the importance of accounting for location and distribution of observations in estimating unbiased effects.

Conclusions

There have been a few attempts to control and estimate spatial effects within the context of human CRTs, but our overall understanding is limited. Although spatial effects may bias trial results, their consideration was usually a supplementary, rather than primary analysis. Further work is required to evaluate and develop the spatial methodologies relevant to a range of CRTs.

3.4. Background

Randomised controlled trials assess the efficacy and safety of interventions [1, 2]. When it is difficult to allocate interventions at the individual level, for example due to logistical or financial restrictions, randomisation and allocation of interventions at a group level may be preferred, this is a cluster randomised trial (CRT) [3]. CRTs also allow for estimation of spillover and herd effects; the apparent treatment effect on individuals who do not receive the intervention [4]. Failure to account for spillover effects can result in biases that reduce the quality of trials and mean that absence of bias is no longer guaranteed by the randomisation, especially when the relationship between intervention and outcome is complex [5, 6].

Spatial effects are effects stemming from locational variation in the distribution of phenomena of interest or in the intensity of interaction between phenomena of interest. Such effects manifest as local variation in the estimated treatment effect over a study area. The existence of spatial effects suggests that global effects estimates are uncertain, and may under- or over- estimate the true effect dependent on location. When values of a single variable are related to nearby values of the same variable this is called spatial dependence, the existence of which is usually captured using spatial autocorrelation measures [7]. Spatial dependence is a fundamental concept in spatial statistics [8] and stems from Tobler’s [9] 1st law of geography that “everything is related to everything else, but near things are more related than distant things.” In agricultural field trials it is long established that the location of the data can impact on trial results [10]. Incorporation of spatial methodology in agricultural trials is common, [11] but the impact of spatial effects in human CRTs have not been researched extensively.

Clusters in CRTs are often defined geographically and valid inference relies on the assumption that the clusters are independent irrespective of their nearness to one another [3]. There is frequently an assumption of an absence of spillover; that movement of people and diseases occurs freely within a cluster but movement be-

tween clusters is negligible, non-existent, or not relevant [12]. This assumption can be violated when there is movement of people or diseases across borders, such as mosquitoes flying between control and intervention households. If an intervention such as insecticide-treated bed nets provide a protective effect to nearby control households then ignoring mosquito mobility will result in underestimating the intervention effect of the trial. This could result in trials discarding effective interventions because the control and intervention are both receiving the benefit of the treatment. Spillover may also be due to connections in social networks, [13] also violating the assumption of independence. In this paper, we consider spillover that can be estimated using GPS data which is often collected as part of trials and therefore do not consider social networks. Spillovers are more likely in trials that have spatially close clusters, and the effect is especially important when an individual's outcome is affected by their proximity to other individuals with different exposure statuses. One way to minimise the potential for spillover is to design a trial with well-separated clusters. In practice this may not be logistically or financially feasible as spatial effects can be present over distances of several kilometres [14]. Furthermore, greater distances between clusters removes our ability to measure spillover spatial effects. To be able to measure spatial effects we need to have data on people nearby to one another and by separating the clusters we may no longer have such information. If proximity to an intervention affects non-treated individuals this is of usually scientific interest and something we should measure. There is a need to control for and estimate spatial effects in the analysis of a trial without adding extra complexity to the design.

We therefore focus only on spatial analysis methods used in CRTs in this review and do not consider alternative trial designs. A further reason for focusing on analysis methods is that this may enable analysis of existing and previous CRTs where redesign is not possible. This review is a diagnostic review of spatial analysis methods that have been used in CRTs. As such, it does not attempt to pose statistical

solutions or determine the best way to account for spatial effects within CRTs. We will describe the state of the literature and aim to improve understanding and help inform further research into spatial effects within CRTs by (i) Identifying spatial analysis methods used in CRTs. (ii) Summarising and grouping spatial methods (iii) Assessing the impact of spatial effects.

3.5. Methods

3.5.1. Search terms and review process

The PRISMA guidelines [15] for systematic reviews and meta-analyses were followed for this review. It was conducted between January 2016 and September 2016. Ovid/Medline, Pubmed, and Web of Science databases were electronically searched and Mendeley was used to store articles. Search terms for CRTs and spatial effects are detailed in Table 3.1. Studies up to end of 2015 were included and only English language articles and search terms were considered.

Papers from each database were combined into a single spreadsheet containing the title, authors, journal, and year. Duplicates were removed automatically within the software and then manually during the title screen. The titles were screened to remove irrelevant papers such as individually randomised trials. Following this, abstracts were screened and the full texts of potentially relevant papers were independently reviewed by two reviewers and disagreements resolved. After selecting relevant papers the references of the articles were screened.

3.5.2. Inclusion and exclusion criteria

The inclusion criteria for studies were: (i) The study is a CRT. (ii) Spatial methods are used in the analysis of the study. We categorised a spatial method as one which accounts for the structure, location, or relative distances of the data. This includes

direct estimation of an effect such as, the change in risk for those within 100 meters of an intervention household or the use of spatial models which account for spatial structure.

The exclusion criteria were (i) Non-randomised studies (ii) Individually randomised trials and hybrid CRTs such as Double or pseudo-randomised studies as they were considered not to be cluster randomised trials (iii) Grey literature (iv) Studies where spillover effects are measured in a non-spatial way, for example comparison of vaccinated and non-vaccinated individuals within an intervention cluster. (v) Studies that account for spatial effects at the design stage only; for instance, using buffer zones or well-separated clusters (vi) Articles which were study protocols and therefore had not applied their methods yet.

3.5.3. Data extraction

The following variables were collected on each paper: title, year, journal, author, intervention, outcome, whether a map was presented, and spatial analysis method.

Table 3.1. Search terms

Databases	CRT terms	Spatial terms	Search string
Pubmed Medline/Ovid Group WebOfScience	randomi*ed trial group community cluster place	spatial* indirect effect* spillover* contamination* externalit*	(randomi*ed trial) AND (group OR community OR cluster OR place) AND (spatial* OR indirect effect* OR spillover* OR contamination* OR externalit*)

A star (*) represents a wildcard character

3.6. Results

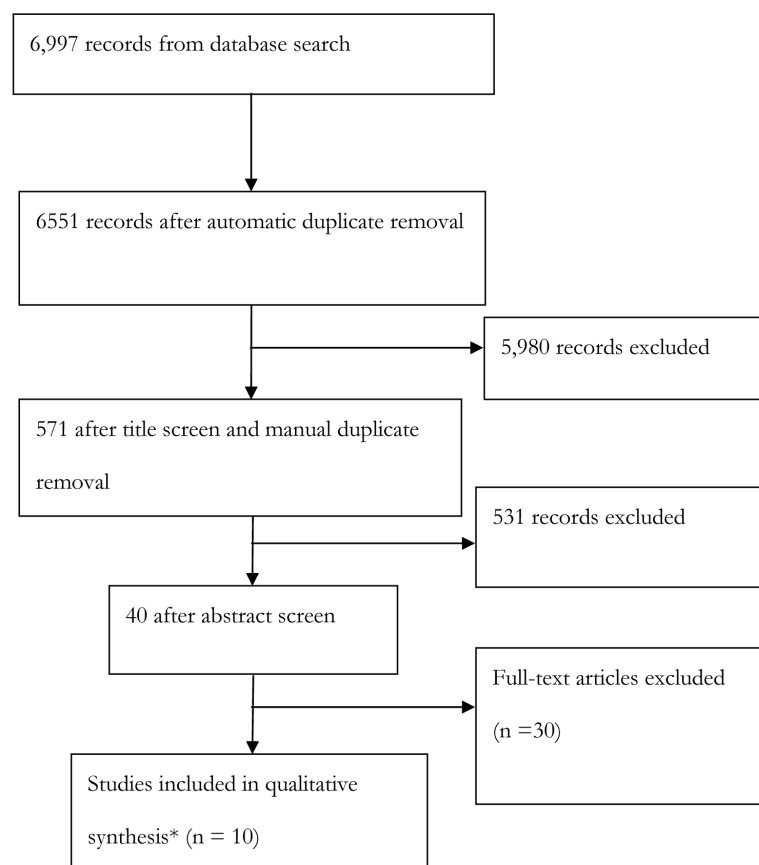
3.6.1. Search results

A flow chart of the search process can be seen in Figure 3.1 and the search terms in Table 3.1. The search terms returned 6,997 records, reducing to 571 records after the

title screen and duplicate removal. Of the 571 records, 40 abstracts were considered relevant for full text review by the reviewers. One study [16] was a replication analysis and it was decided to include the original study in the review instead of the replication. There are ten papers and thirteen trials in this review as some papers include multiple trials.

Whilst this review was being conducted a systematic review on health-related spillover in impact evaluations was released [13]. It has the more general aim of attempting to summarise methods to estimate health related spillover in low and middle income countries. Our review is different as it only includes spatial methods used within CRTs and does not restrict by type of country. The results from both reviews were compared and did not lead to additional records being included.

Figure 3.1. Flow chart of search results



3.6.2. General characteristics

This review contains ten papers published between 1998 and 2015, they relate to thirteen trials as some papers contained more than one trial. The trials took place around the globe with three taking place in Kenya, three in the United Kingdom, and the rest in Mexico, Venezuela, Ghana, Papua New Guinea, Vietnam, Haiti, and India. There is one stepped wedge trial [17] and all others are parallel cluster randomised trials. Six of the papers were a spatial reanalysis of a previously reported trial.

Seven trials focused on pathogens carried by mosquitoes. The intervention for six of these was insecticide-treated bed nets or curtains and the other intervention was a drug. Two of the seven mosquito trials considered all cause child mortality as an endpoint, the other five looked at entomological endpoints.

There were two vaccine trials, the first looked at vaccine uptake in response to a mass campaign and the second evaluated vaccine effectiveness for a typhoid vaccine. One paper considered primary care and community based trials, within this they applied spatial methods to three different trials, one simulated and two real trials. The final trial looked at the impact of deworming on education and health within schools. Further details of the trials in the review are in Table 3.2.

Table 3.2. Characteristics of the included papers

Author & Year	Location	Intervention	Outcome	Map	Spatial Method Type	Spatial Method
Binka 1998 [18]	Ghana	Permethrin-impregnated bed net	All-cause child mortality	Yes	Straight line distance	Distance to discordant observation and points of interest. Standardised mortality rates calculated at several distances.
Alexander 2003 [27]	Papua New Guinea	Diethylcarbamazine (DEC) plus ivermectin versus DEC alone	Spatial distribution of <i>Wuchereria bancrofti</i> and microfilariae	Yes	Spatially structured random effect	Negative binomial model with a distance parameter in the covariance structure of a random effect. Measures half distance of spatial correlation.
Gimnig 2003 [20]	Kenya	Permethrin-treated bed net	Spatial distribution of mosquitoes	Yes	Straight line distance	Distance to discordant observation and to points of interest. Poisson regression model with a random effect for cluster.
Hawley 2003 [19]	Kenya	Permethrin-treated bed net	All-cause child mortality, anaemia, and density of mosquitoes	Yes	Straight line distance	Distance to discordant observation and to points of interest. Cox regression model with a random effect for adjusted for cluster.
Miguel & Kremer 2004 [23]	Kenya	Deworming	Helminth infection	No	Density	Total number and proportion of treated students within 6 km of a school. Included in primary analysis random effects model adjusted for cluster (School).
Kroeger 2006 [21]	(1) Mexico (2) Venezuela	Insecticide-treated curtains and water container covers	Reduction in entomological indices for Dengue	No	Straight line distance	Distance to nearest participant with outcome at the beginning of the study. Odds ratio of outcome for nearby houses compared to houses further away.
Ali 2007 [22]	Vietnam	Vaccine campaign	Vaccine uptake	Yes	Straight line distance and density	Distance to points of interest and density. Random effects model including typhoid prevalence and private practitioner density
Lenhart 2008 [24]	Haiti	Insecticide-treated bed nets	Reduction in entomological indices for Dengue	Yes	Density	Total number of bed net households within 100 meters. Spearman's correlation of number of bed net houses within 100m compared with change in entomological measures.
Silcocks 2010 [29]	UK (3 trials)	(1) Sun exposure (2) home safety intervention (3) intervention to reduce baby walker use	(1) Lip Cancer (2) number of injuries per individual (3) ownership of baby walker	No	Spatially structured random effect	Spatial weights matrix in covariance of random effect. A Multiple membership spatial random effects model with fixed North/South or East/West gradient covariate effect for gradient
Chao 2015 [25]	India	Typhoid vaccine	Vaccine effectiveness	Yes	Density	The sum of the risk of those within 100 meters of a participant called the potential exposure. Included in a model with a random effect for cluster.

3.6.3. Spatial methods

The studies took two approaches for analysing spatial effects, referred to in this paper as spatial variables and spatial models, expanded upon in Figure 3.2. Nine of the trials analysed the effect of a spatial variable which is a measurement that relates to where the observations are located. The two types of spatial variables found in this review were straight line distance such as distance between participants, and density, for instance the number of treated participants within a 100 meter radius. Four trials used spatial models by including the spatial structure of the participants using random effects statistical models. The spatial models are specified by measuring how participants are connected to one another, for example recording participants who are neighbours, described further in Figure 3.2. Incorporating spatial structure into a model this way treats a spatial effect as an underlying unobserved process which may not be directly measurable. The two approaches make different assumptions about the type of spatial relationship and a variety of methods were used for each approach.

3.6.3.1. Spatial variables

Straight line distance

Five trials [18–22] estimate spatial effects by measuring the distance between participants and a location of interest. In these studies, the location is either another participant or a feature which may affect the outcome, such as a health facility. Several studies analysed the effect of distance to more than one type of location. The distance between each control participant to their nearest intervention participant was analysed in three trials [18–20], termed distance to nearest discordant observation. They also analysed proximity to nearest reservoir or health facility. Distances were categorised and the effect measured for each category. For example, Binka et al. [18] calculated a standardised mortality rate at five separate distance categories.

Kroeger et al. [21] considered whether distance to a participant with the outcome at the beginning of the study affects the odds of having the outcome at the end of the study. They tested at four separate distances and corrected for multiple testing. Ali et al. [22] included distance to school and nearest hospital in a model which assessed the intervention effect and accounted for cluster effects. This was the only trial that included straight-line distance in the primary analysis of their trial.

Density

Four trials [22–25] analysed the effect of density in the area surrounding the participants. They analysed the density of factors that may affect the outcome, for example the number of people vaccinated within 100 meters. The methods differ by whether they used a count or a proportion and whether they focused on the treatment density or the risk of infection from surrounding individuals. Including density as a spatial variable assumes that number of objects within a certain distance is important as well as the distance to the nearest object.

Lenhart et al. [24] measured intervention density as the number of households with bed nets within 100 meters of an observation. The study assessed spatial spillover through the correlation of change in baseline of outcome measure with number of bed net households within 100m. In contrast, Miguel & Kremer [23] measured density as the proportion of children treated within 6km of a school, as well as the total number and accounted for this in their primary analysis. Ali et al. [22] included a proportion and count density in their primary analysis model by accounting for typhoid prevalence and number of private practitioners for the neighbourhoods of each participant.

Chao et al. [25] differs from the previous applied papers because they develop and test a new method to deal with spatial effects in CRTs. They define a variable called '*potential exposure*' which is 'the sum of the relative risks of all who live within 100 meters of each person' [25]. The potential exposure controls for the spatial variation

in risk surrounding an observation. They demonstrate that this variable can be used to account for spatially heterogeneous risk factors in the primary analysis of a trial.

Figure 3.2. Spatial analysis methods

Spatial variables

Nine of the trials measured spatial effects by including a spatial variable in their analysis, denoted as D :

$$Y = \beta D + \epsilon$$

β = the effect of a change in variable D

D was defined in the following two ways:

Straight Line Distance

D = straight line distance. For instance, distance to nearest discordant observation or distance to nearest health facility. This variable may be treated as continuous or categorical.

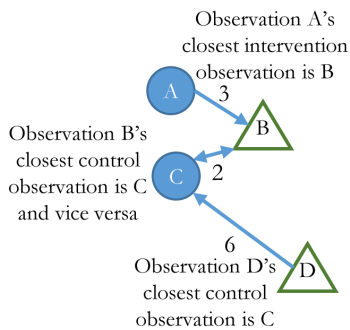
Studies: [18,19,20,21,22]

Density

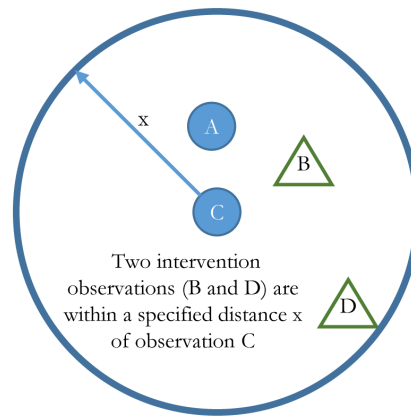
D = the density of observations of interest within a specified distance. For instance, the number of intervention observations or the total risk of observations within the specified distance.

Studies: [22,23,24,25]

Nearest discordant observation



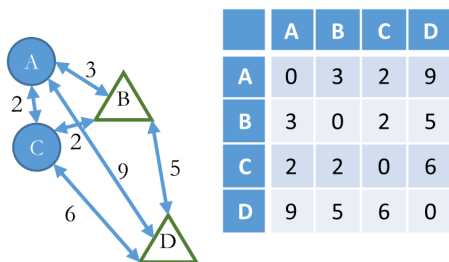
Number of intervention observations



Spatial Models

Four of the trials model the spatial relationship of the data using a random effect, denoted as U below. The spatial structure is represented as a matrix that defines how the observations are connected to one another.

Spatial Structure Spatial Weights Matrix



$$Y = X\beta + U + e$$

Where $U \sim MVN(0, \sigma_u^2 \Sigma)$

Σ was defined as:

$\Sigma_{ij} = e^{-d_{ij}/\phi}$ where d_{ij} represent the distance between observation i and j and ϕ is a scale parameter

Study: [27]

$\Sigma = (I - \rho W)^{-1}$ where W is a spatial weights matrix I is the identity and ρ represent the degree of spatial correlation

Study: [29]

The spatial weights matrix can be formed in many ways. Here we have used distance. A common approach is to define observations within a specified distance as neighbours and create a spatial adjacency matrix.

Control Intervention

Distance of x units

Key

Y = a vector representing the outcome β = a vector of coefficients

X = design matrix U = a random effect

ϵ = an error term D = a spatial variable

where $\epsilon \sim N(0, \sigma^2)$

3.6.4. Spatial models

The methods presented in the previous section assume that the underlying spatial process can be measured and spatial effects can be estimated from this measure. The remaining two papers include four trials that model spatial effects using a spatially structured random effect. This approach makes fewer assumptions about the mechanism of spatial process and allows for a range of local and global dependency structures [26].

Alexander et al. [27] investigated the spatial pattern of mosquito borne vectors. Adapted from a previous paper they incorporated a distance parameter in the covariance matrix of a random effect within a negative binomial model [28]. This distance parameter allows participants who are closer together to be more similar than participants that are further apart as shown in Figure 3.2. They also estimate a distance decaying parameter in the covariance structure of the random effect, which they use to estimate the ‘*half distance*’, which is the distance at which spatial correlation halves.

Silcocks & Kendrick [29] applied several types of spatial models to two primary care trials and one simulated trial. The model was a variation of a Besag, York, and Mollie model [30] which contains a spatially structured random effect and a random effect for cluster. The spatial structure of the participants was represented using a spatial weights matrix and is included in the covariance of the random effect, further described in Figure 3.2. In primary care or community based trials people may reside in one area and receive treatment in another [29]. Having a spatial and non-spatial random effect allowed participants to have membership to multiple clusters which they called a ‘*multiple membership model*’. They also consider a fixed north/south and east/west gradient covariate in their model evaluations.

3.6.5. Impact of spatial effects

All thirteen trials found evidence of a spatial effect within their studies. Seven trials report a protective spillover effect for participants who live close to an intervention. There is evidence that adjusting for spatial effects affects the precision and value of the estimated intervention effect. Chao et al. [25] saw that adjustment for spatial effects lowered the effect estimate of the intervention. The precision and intervention estimate changed in the three trials analysed by Silcocks & Kendrick [29]. The study demonstrated that spatial models fitted better than a standard CRT random effect model by comparing the Akaike information criteria of the models. They are explicit that this is just illustrative but both studies conclude that spatial effects may need to be adjusted for in CRTs and that further research into methods is required. Despite this only two of the nine applied trials adjusted for spatial effects in the primary analysis of their trial.

3.7. Discussion

This review has found multiple approaches to incorporating or measuring spatial effects in the context of CRTs, however these stem from only a few examples in the literature. Further, no conventional or standard approach was found. The approaches differ by whether they directly analyse a spatial variable or model the spatial structure. Spatial variables were either straight-line distance from a participant to a place of interest, or a measure of the density surrounding a participant. For instance, distance between a control participant and an intervention participant. Spatial models included spatial structure in the covariance of the random effects model using a distance parameter or spatial weights matrix. Accounting for spatial structure affects both the precision and point estimates of treatment effects and failure to do so could give inaccurate results [25, 29].

The papers in this review are only a small proportion of the total number of CRTs

that have been published. That only ten records were found suggests that spatial effects are not often considered in this area. Furthermore, despite evidence of spatial effects, they were rarely adjusted for in the primary analysis of the trial. It appears that the impact of location on analytical approach is at best an afterthought and in most cases ignored. The trials come from a variety of domains and although predominantly focused in infectious diseases, there may be implications for a broader range of trials particularly in trials of health services organisation [29] which are becoming more common [31]. Therefore, a wide range of trials may not be accounting for spatial effects that bias results, however it is presently unclear as to what extent this may be an issue.

Further research is required to determine how much spatial effects impact trial results. Simulation studies may allow exploration of how the magnitude and extent of spatial autocorrelation may bias trial results. There have been some attempts to quantify how important spatial effects may be in trials more generally [32, 33]. Methods that allow for estimation of treatment effects whilst accounting for spatial effects could be investigated and tested under simulation. However, there are several challenges to overcome such as whether we can estimate the true randomised intervention effect using a spatial model, and how such an effect estimate should be interpreted.

An alternative use of spatial data is to conduct additional analyses to complement analysis of the main trial. These analyses explore spatial components of the trial, and could improve understanding of the mechanism of the intervention effect. Spatial methods could be applied to previous trial data and a toolbox of standard approaches defined allowing future trials to predefine spatial analyses. This would allow for quantification of the distance over which spatial effects are present for different disease areas. A further area of research is in alternative CRT designs, although they are not the focus of this review this might include double randomisation or pseudo randomised trials where clusters are first randomised and subsequently individuals

within clusters are randomised to allow for measurement of spillover effects [34, 35]. Multiple terms in the literature refer to spillover effects and many of them could refer to spatial effects with differing terminology between fields and researchers in the same fields. To add to confusion, these terms can have dual meanings for instance, an indirect effect can be the effect on an individual who does not receive the intervention or the effect of an intervention through a mediating variable. The search strategy attempted to include a broad range of terms for CRTs and spatial effects but at present, there is no established standard for citing the use of spatial data in the analysis of CRTs. Consequently, it is possible that trials have been missed. Although this is a weakness, comparison with a larger more general systematic review on health-related spillover in impact evaluations [13] did not result in the addition of any further trials. They had searched 19 databases and screened more than 34,000 records. Due to this the authors conclude the omission of further studies is likely to be minimal.

This review has focused only on the analysis stage of a CRT and it could be argued that adjusting for spatial effects is not necessary in a well-designed trial as clusters should be well-separated to minimise spillover effects [4]. Trial designs such as the fried egg design [3] which incorporates a buffer around clusters could be used to attempt to eliminate or measure spatial spillover. However, in cases where spatial correlation is present over large distances [14] this may not be possible and could lead to the inability to detect the difference between no effect and everyone having an effect [36]. Additionally, as spatial effects are rarely considered in trials, it may not be until after the design stage that the problem becomes apparent, if at all.

On the other hand, having clusters relatively close together could have advantages, because the measurement of spatial spillover effects is of scientific interest. Knowledge that an intervention provides indirect benefit based on proximity is useful to differentiate between interventions and to plan how to benefit the largest number of people. Furthermore, subjects who live further apart are more likely to be het-

erogeneous than those living close together due to cultural, geographical, and social differences which could make treatment differences harder to distinguish due to imbalance between clusters. In conclusion, although there have been a few attempts to control and estimate spatial effects within the context of human CRTs our understanding is limited. Although there are commonalities between approaches there is no consensus on how to account for spatial effects within CRTs and more work needs to be done to evaluate and develop spatial methodology within the context of a range of CRTs.

3.8. Bibliography

- [1] Pocock, SJ. *Clinical Trials A Practical Approach*. West Sussex: John Wiley & Sons Ltd, 1983.
- [2] Smith, PG, Morrow, RH, and Ross, DA, eds. *Field Trials of Health Interventions: A Toolbox*. 3rd. Oxford: Oxford University Press, 2015.
- [3] Hayes, RJ and Moulton, LH. *Cluster Randomised Trials*. Boca Raton: Chapman & Hall/ CRC, 2009.
- [4] Hayes, RJ, Alexander, ND, Bennett, S, and Cousens, SN. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Statistical methods in medical research* 2000 9 95–116.
- [5] Hahn, S, Puffer, S, Torgerson, DJ, and Watson, J. Methodological bias in cluster randomised trials. *BMC medical research methodology* 2005 5 10.
- [6] English, M, Schellenberg, J, and Todd, J. Assessing health system interventions: Key points when considering the value of randomization. *Bulletin of the World Health Organization* 2011 89. (12) 907–912.
- [7] Getis, A. A history of the concept of spatial autocorrelation: A geographer’s perspective. *Geographical Analysis* 2008 40. (3) 297–309.
- [8] Cressie, N. *Statistics for Spatial Data*. United States of America: John Wiley & Sons Ltd, 1991.
- [9] Tobler, AWR. A Computer Movie Simulation Urban Growth in Detroit Region. *Economic Geography* 1970 46. (332) 234–240.
- [10] Wiebe, GA. Variation and Correlation in Grain Yield among 1,500 Wheat Nursery Plots. *Journal of Agricultural Research* 1935 50. (4) 331–357.
- [11] Besag, J and Kempton, R. Statistical Analysis of Field Experiments Using Neighbouring Plots. *Biometrics* 1986 42. (2) 231–251.

- [12] Lines, J and Kleinschmidt, I. Combining malaria vector control interventions: some trial design issues. *Pathogens and global health* 2013 107. (1) 1–4.
- [13] Benjamin-chung, J, Berger, D, Clark, A, Falcao, L, Tran, D, Hubbard, A, Miguel, E, and Colford, J. The identification and measurement of health-related spillovers in impact evaluations : a systematic review. Tech. rep. London, 2015.
- [14] Howard, SC, Omumbo, J, Nevill, C, Some, ES, Donnelly, CA, and Snow, RW. Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2000 94. (4) 357–360.
- [15] Moher, D, Liberati, A, Tetzlaff, J, Altman, DG, and Grp, P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement (Reprinted from Annals of Internal Medicine). *Physical Therapy* 2009 89. (9) 873–880.
- [16] Aiken, AM, Davey, C, Hargreaves, JR, and Hayes, RJ. Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a pure replication. *International Journal of Epidemiology* 2015 44. (5) 1572–1580.
- [17] Hemming, K, Haines, TP, Chilton, PJ, Girling, aJ, and Lilford, RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ* 2015 350 h391–h391.
- [18] Binka, FN, Indome, F, and Smith, T. Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. *The American Journal of Tropical Medicine and Hygiene* 1998 59. (1) 80–85.
- [19] Hawley, WA, Phillips-Howard, PA, Ter Kuile, FO, Terlouw, DJ, Vulule, JM, Ombok, M, Nahlen, BL, Gimnig, JE, Kariuki, SK, Kolczak, MS, and High-tower, AW. Community-wide effects of permethrin-treated bed nets on child

- mortality and malaria morbidity in western Kenya. *American Journal of Tropical Medicine and Hygiene* 2003 68. (4.) 121–127.
- [20] Gimnig, JE, Kolczak, MS, Hightower, AW, Vulule, JM, Schoute, E, Kamau, L, Phillips-Howard, PA, Ter Kuile, FO, Nahlen, BL, and Hawley, WA. Effect of permethrin-treated bed nets on the spatial distribution of malaria vectors in western Kenya. *American Journal of Tropical Medicine and Hygiene* 2003 68. (4) 115–120.
- [21] Kroeger, A, Lenhart, A, Ochoa, M, Villegas, E, Levy, M, Alexander, N, and McCall, P.J. Effective control of dengue vectors with curtains and water container covers treated with insecticide in Mexico and Venezuela: cluster randomised trials. *BMJ (Clinical research ed.)* May 2006 332. (7552) 1247–52.
- [22] Ali, M, Thiem, VD, Park, JK, Ochiai, RL, Canh, DG, Danovaro-Holliday, MC, Kaljee, LM, Clemens, JD, and Acosta, CJ. Geographic analysis of vaccine uptake in a cluster-randomized controlled trial in Hue, Vietnam. *Health and Place* 2007 13. (3) 577–587.
- [23] Miguel, E and Kremer, M. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica* 2004 72. (1) 159–217.
- [24] Lenhart, A, Orelus, N, Maskill, R, Alexander, N, Streit, T, and McCall, P.J. Insecticide-treated bednets to control dengue vectors: Preliminary evidence from a controlled trial in Haiti. *Tropical Medicine and International Health* 2008 13. (1) 56–67.
- [25] Chao, DL, Park, JK, Marks, F, Ochiai, RL, Longini, IM, and Halloran, ME. The contribution of neighbours to an individual’s risk of typhoid outcome. *Epidemiology and Infection* 2015 143. (16) 3520–7.
- [26] Banerjee, S, Carlin, BP, and Gelfrand, AE. Hierarchical Modeling and Analysis for Spatial Data. Boca Raton: Chapman & Hall/ CRC, 2015.

- [27] Alexander, N, Moyeed, Ra, Hyun, PJ, Dimber, ZB, Bockarie, MJ, Stander, J, Grenfell, BT, Kazura, JW, and Alpers, MP. Spatial variation of Anopheles-transmitted *Wuchereria bancrofti* and *Plasmodium falciparum* infection densities in Papua New Guinea. *Filaria journal* 2003 2 14.
- [28] Alexander, N, Moyeed, R, and Stander, J. Spatial modelling of individual-level parasite counts using the negative binomial distribution. *Biostatistics* 2000 1. (4) 453–463.
- [29] Silcocks, P and Kendrick, D. Spatial effects should be allowed for in primary care and other community-based cluster RCTS. *Trials* Jan. 2010 11. (1) 55.
- [30] Besag, J, York, J, and Mollié, A. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 1991 43. (1) 1–20.
- [31] Eldridge, SM, Ashby, D, Feder, GS, Rudnicka, AR, and Ukoumunne, OC. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. *Clinical Trials* 2004 1. (1) 80–90.
- [32] Diamond, R, Imbens, GW, and Diamond, R. Clustering, Spatial correlations and randomization inference. 2010.
- [33] Baylis, K and Ham, A. How important is spatial correlation in randomized controlled trials? 2015.
- [34] Sinclair, B, McConnell, M, Green, DP, Sinclair, B, and Green, DP. Detecting Spillover Effects : Design Multilevel Experiments and Analysis of Chicago. *American Journal of Political Science*, 2014 56. (4) 1055–1069.
- [35] Baird, S, Bohren, JA, McIntosh, C, and Ozler, B. Designing Experiments to Measure Spillover Effects. *Policy Research Working Paper No. 6824* 2014. (March). URL: <http://www.ssrn.com/abstract=2402749>.
- [36] Rosenbaum, PR. Interference Between Units in Randomized Experiments. *Journal of the American Statistical Association* 2007 102 191–200.

4. Paper B: Spatial analyses of oral polio vaccine transmission in an inactivated polio vaccinated community

Christopher Jarvis^{1,2}, Jonathan Altamirano³, Clea Sarnquist³, W John Edmunds¹, Yvonne Maldonado³

1. London School of Hygiene and Tropical Medicine, London, UK
2. MRC London Hub for Trials Methodology Research, London, UK
3. Stanford University School of Medicine, Stanford, California, USA

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk



Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Christopher Jarvis
Principal Supervisor	John Edmunds
Thesis Title	Spatial analyses of OPV transmission in an IPV-vaccinated community

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Clinical Infectious Diseases
Please list the paper's authors in the intended authorship order:	Christopher Jarvis, Jonathan Altamirano, Clea Sarnquist, W John Edmunds, Yvonne Maldonado
Stage of publication	Submitted

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I planned and conducted the analysis, and drafted the article with feedback from co-authors.
--	--

Student Signature: _____

Date: 28/3/2018

Supervisor Signature:



Date: 28/3/2018

4.1. Overview

In chapter 3, a systematic review was conducted to determine to what extent the methods of spatial statistics have been incorporated into the analyses of cluster randomised trials (CRTs). In this chapter, I apply spatial methods to a CRT to consider the spatial aspects of poliovirus transmission. The CRT is a household CRT, which has implications for which spatial measurements are meaningful, as the clusters and the observations are represented by the same point locations. Specifically this paper is considers and attempts to measure the spillover effect of using OPV onto participants who do not receive the vaccine. This paper has been submitted to Clinical Infectious Diseases in combination with several other papers relating to this CRT, and is currently under review.

Objective

3. Apply and assess a range of appropriate modern spatial methods to existing CRT data, in order to analyse the effect of spatial autocorrelation and spatial spillover effects on CRT results.

4.2. Role of candidate

I created the statistical analysis plan which was agreed by the co-authors. I conducted the statistical analysis, and wrote the first draft of the paper with input from Jonathan Altamirano (JA) on the study design and knowledge relating to polio. Revisions were made with feedback, input, and guidance from JA, Clea Sarnquist, W John Edmunds, and Yvonne Maldonado.

4.3. Abstract

Background

Understanding spatial dynamics of oral polio vaccine (OPV) transmission will improve resource targeting. Mexico provides a natural laboratory as it uses both inactivated polio vaccine (IPV) routinely and OPV bi-annually. We performed a spatial analysis to consider whether living near an individual who receives the OPV increases an individual's risk of poliovirus shedding.

Methods

Children in three villages near Orizaba, Mexico were randomized to three levels (10%, 30%, 70%) to receive OPV. We measured distance to nearest OPV shedding, and the amount of shedding close to unvaccinated individuals. We used maps to show the proximity and amount of shedding. Distance and density of shedding was analyzed separately using mixed effects logistic regression with random effects for household and time, adjusted for age, gender, area, and running water.

Results

The median distance to nearest OPV shedding households was 85 meters (IQR 46, 145) and median number of vaccinees shedding OPV within 200m was 3 (2, 6). Transmission to unvaccinated household occurred by day one (Figure 4.5) and persisted in some cases up to 71 days. There was no association (Odds Ratio [OR] 1.04 95% Credible Interval [CrI] 0.92, 1.16) between distance from OPV shedding and odds of transmission. There was some suggestions of an association between the number of OPV vaccinees shedding within 200m with unvaccinated transmission (OR 0.93 95% CrI 0.84, 1.01) but not at 100 or 500m. Results were consistent across the three villages.

Conclusions

Household structure appears to have limited value in predicting transmission of poliovirus shedding. Use of OPV results in rapid but low levels of persistent transmission throughout the community and this would usually go undetected. The only way to avoid this is to not use OPV or to have strong controls such as quarantine, or strict hygiene protocols. After withdrawal of OPV worldwide the decision to reintroduce due to an outbreak should not be taken lightly as it appears a small amount of OPV is needed to result in transmission.

4.4. Background

The goal of global polio eradication may soon be reached. In 2016, there were 37 cases of wild type (WT) polio, a decrease from the 87 cases in 2015 [1, 2]. Currently, only Pakistan and Afghanistan continue with endemic transmission of WT serotype 1 polio, and Nigeria reported four cases in September 2016 [3]. The success of the polio eradication effort can be largely attributed to the use of oral polio vaccine (OPV) in developing countries, due to its low cost, easy administration, and ability to confer passive immunization to contacts, presumably via fecal-oral infection. However, the risks from OPV use are now complicating eradication efforts. Vaccine-associated paralytic polio (VAPP) is a rare, adverse reaction to OPV administration, occurring in every 900,000 doses [4]. Of more concern, prolonged circulation of OPV, primarily in communities with suboptimal sanitation, can result in OPV mutation and neuroreversion, leading to circulating vaccine-derived poliovirus (cVDPV). In addition, cVDPVs studied in Nigeria were found to be as virulent as WT polio [4]. To date, the spatial characteristics of OPV transmission are not well characterised. Understanding the pattern and extent of geographic variation in OPV circulation could help to predict and prepare for risk of OPV reintroduction, especially in undervaccinated communities. Areas of increased risk could be detected, and resources

accordingly deployed to reduce or prevent prolonged OPV circulation. In the current global setting of polio transmission, this is particularly important for serotype 2, which was declared eradicated as of September 2015 by the WHO [5].

In this study, we had the opportunity to identify household and community transmission of OPV because it was conducted in Mexico, where inactivated polio vaccine (IPV) is provided for routine immunization at 2, 4, 6 and 18 months of age, and OPV is only administered in two National Immunization Weeks (NIW) to children 5 years of age and under. We are able to investigate spatial transmission at a household level within three villages that received different levels of OPV vaccination during the February 2015 NIW. Furthermore, we record transmission based on stool samples allowing detection of OPV transmission from vaccinated children and their household and community contacts. Spatial analyses of polio transmission have been performed as far back as 1967 [6], however most involved aggregated case data over large areas of several kilometres [7–9]. This study allows investigation at a local level of what happens in a community when the OPV vaccine is introduced.

In this paper, we explore two aspects of between household transmission of polio. First, we consider whether living near someone who is shedding poliovirus affects an individual’s chance of shedding poliovirus. Second, we determine how the number of people shedding near an unvaccinated contact impacts that individual’s chance of acquiring and shedding poliovirus. We refer to the distance and density of shedding collectively as proximity to shedding.

4.5. Methods

4.5.1. Study design

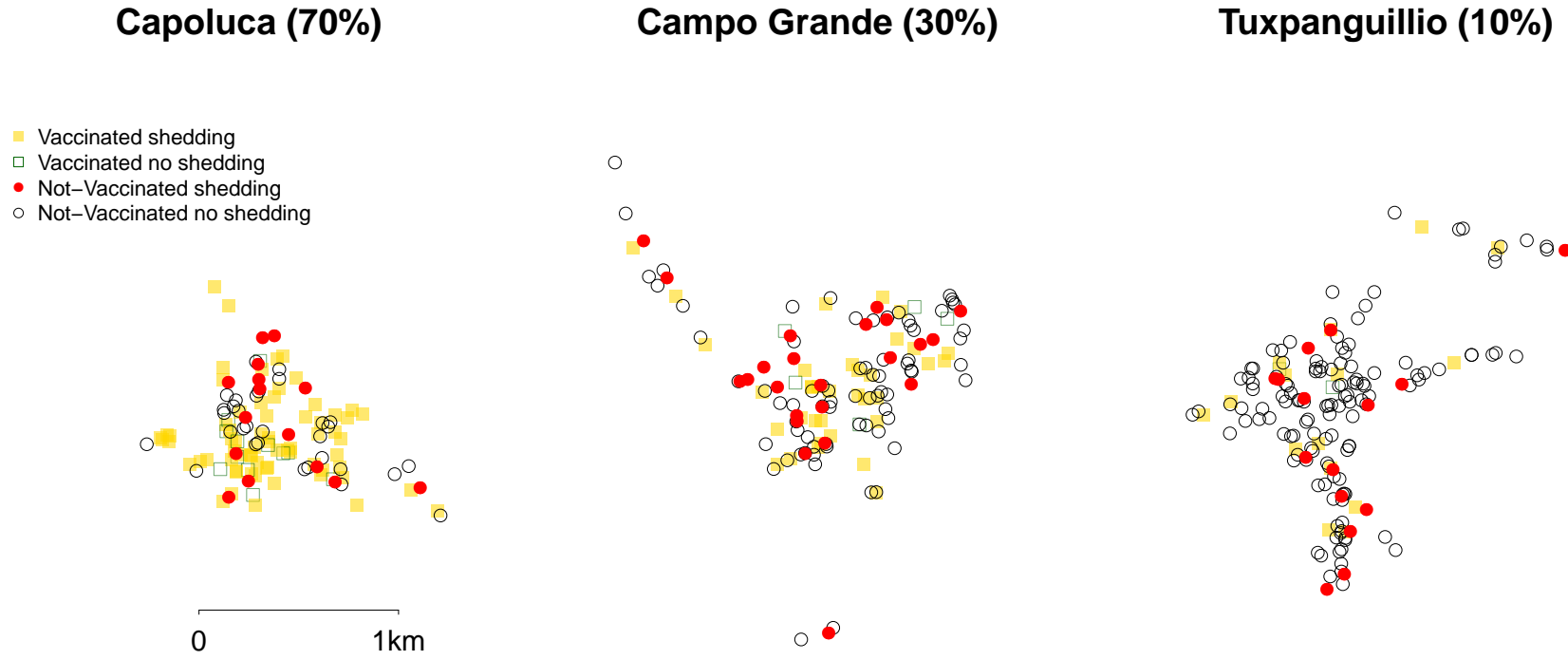
The study has already been described in detail elsewhere [10] but in brief this was a prospective household cluster randomised trial (CRT) in 3 indigenous localities in Orizaba, Veracruz, Mexico (Capoluca, Campo Grande, and Tuxpanguillo). Within

each community, approximately 150 households were enrolled in this study, and each community received a different amount of OPV coverage as part of the study; 70% of enrolled households in Capoluca, 30% in Campo Grande, and 10% in Tuxpanguillo. When enrolment began in February 2015, 155 households were randomised to receive OPV out of 466 households included across the three localities. No other households in any of the three communities received OPV until the May 2015 NIW. Only one child from each of the 155 households received OPV. Inclusion criteria for household enrolment was the presence of a child <5 years with an up-to-date IPV vaccination record that was eligible to receive OPV. All adult participants consented to participation, and guardians of minors consented for minors to participate. Exclusion criteria for children <5 years included presentation with illness (febrile, diarrhea, or respiratory), immunodeficiency caused by AIDS, disease, or medication, recent blood transfusions, and prior adverse reactions to OPV. Exclusion criteria for all other participants was refusal to participate or change in residence during the study period. Within our study population, GPS coordinates were collected from 423 households, 137 of the vaccinated households and 286 unvaccinated households. As a result, only the shedding and transmission results from these participants were considered in this analysis (Figure 4.1).

After enrolment, 10 stool samples were scheduled for collection from each member of all enrolled households, one baseline sample collected before vaccination, and then collected 1, 4, 7, 10, 14, 21, 28, 51, and 71 days after vaccination. During each visit, health information, travel and visit details, and records for any vaccines received during the study period for children <5, were collected via follow-up surveys. Exclusion criteria for follow up were individuals that refused to participate, change in residence during the study period, or absence during follow-up visits. Viral RNA was extracted from frozen stool samples utilizing the MagNA Lyser (Roche) and KingFisher Duo Prime (Fisher Scientific), using the bacteriophage MS2 as internal control for extraction efficiency.

Viral RNA then underwent quantitative reverse transcription polymerase chain reaction (rt-QPCR) in order to detect and quantify any Sabin OPV present in the samples. The probes and primers were adopted and adapted from Kilpatrick et al. [11] and the CDC protocol for Polio QPCR. Samples were run in triplicate and a sample was considered positive if two thirds of reactions had a Ct <37. Positive samples were re-run, to minimize false positives, and if positive again the RNA was Sanger sequenced for confirmation.

Figure 4.1. Map of study area



4.5.2. Sample

We use two types of participants from the study, vaccinated individuals, and individuals who live in unvaccinated households, referred to here as unvaccinated individuals. We are interested in the outcomes of unvaccinated individuals and their proximity to vaccinated individuals. Therefore, vaccinated participants are not analysed directly, instead we create spatial variables for unvaccinated participants by comparing the locations of vaccinated and unvaccinated households. For instance, we measure the distance from an unvaccinated person to their nearest vaccinated household. We do not include unvaccinated people who live in vaccinated households as the focus of our analyses is between-household transmission. Within-household transmission is analysed separately [12]. Vaccinated individuals are treated as sources because we can be more certain their shedding is due to the OPV and not transmission from other participants.

4.5.3. Descriptive analyses

Maps were used to visualise the spatial distribution of vaccination and shedding. We represented the density and location of vaccinated shedding over time using contour plots and overlaid the position of unvaccinated participants.

The distributions of key variables were assessed graphically and through calculation of summary statistics. Spatial variables were summarised using medians and interquartile ranges (IQR) as they were skewed. Coverage of vaccine was considered important to adjust for, as higher coverage led to both a smaller distance to nearest shedding household and a higher proportion of shedding.

4.5.4. Statistical analyses

4.5.4.1. Outcomes

Stool samples were collected from participants at ten time points. We only analysed data from the first 28 days of the study as almost all the shedding occurred before this end point. We first looked at the presence of poliovirus in stool samples and proximity to vaccinated participant shedding at any time point in the first 28 days. Second, we considered the presence of poliovirus and proximity at each given time point.

4.5.4.2. Spatial variables

For the first analyses the outcome was a binary variable for shedding at any time point throughout the study period. We used a spatial variable approach as identified in a recent systematic review [13], this approach is similar to previous spatial analyses of mosquitoes bed net trials [14, 15]. We measured proximity to any vaccinated participant shedding aggregated over the study period. Proximity was measured in two ways; distance to nearest vaccinated shedding household and the number of vaccinated individuals shedding within 100, 200, 500, and 1000 meters of an unvaccinated household. Sensitivity analyses were performed restricting the data to the first 14 days to reduce the risk of shedding being recorded due to secondary transmission.

For the second analyses, we calculated the same spatial variables but within each time point. Therefore, distance to nearest vaccinated shedding could change depending on the study day. We also calculated the spatial variables using calendar days after initiation of the NIW OPV administration and found no differences on the conclusions.

4.5.4.3. Modelling

The study is a CRT with a hierarchical structure giving multiple observations per individual per household. Mixed effects logistic regression with a random effect for household was used to assess the association of distance and density of vaccinated shedding with unvaccinated shedding over the 28-day period. Mixed effects logistic regression with a random effect for household, and an autoregressive lag one random effect for study day was used to look at the impact of distance and density of OPV shedding. The autoregressive random effect allows the previous study period to provide information to the next period. We adjusted for study area, age, gender, and whether the house had running water.

The spatial variables were included as continuous variables. Non-linearity was explored using quadratic terms and treating the variables as categorical, there was no suggestion of non-linear effects. The models were fitted using integrated nested Laplace approximation [16] and all analyses were performed using R 3.4.2 [17].

4.6. Results

4.6.1. Descriptive

There are 1,145 unvaccinated individuals included in the analyses, and they come from 286 households. The age distribution was positively skewed with a median age of 17 years ranging from 1 month to 95 years and 58.1% were female. Further details stratified by community can be seen in Table 4.1. Unvaccinated individuals provided 10,059 stool samples in total, 978 (85.4%) people contributed eight or more out of a possible ten samples and 57 (5.0%) individuals contributed only one sample. There was no missing data for the variables age, running water in household, and gender, the only missing data was the omission of stool samples. If we assume each participant could have provided 10 stool samples then we observed 87.8% of the

11,450 potential samples for unvaccinated individuals.

In unvaccinated individuals, there were 89 (0.9%) positive samples which came from 80 (7.0%) individuals, only six people contributed more than one positive sample. The number of positive samples varied over time, with 24 positive samples observed at day seven, 2 at day ten and 8 at day fourteen.

In the original study 155 children were vaccinated, of these 137 (88.4%) had GPS data which were used to calculate the spatial variables, the remaining 18 did not have GPS data and were not used in the analyses as we do not know their locations (Figure 4.1). The median distance to vaccinated households for unvaccinated individuals was 77.6 meters (IQR 42.0 to 126.1m) . All unvaccinated households were within 826.4m at some point in the study and no household was closer to shedding in another village than in their own (Figure 4.2).

The vaccinated children provided 1,237 samples, of which 342 (27.6%) were positive. There were 108 vaccinated children who shed OPV at any point in the study.

Shedding began quickly and between-household transmission occurred on the first study day in some cases. In addition, the vaccinated children tended to shed more consistently throughout the study, with 36 (26.3%) individuals providing four or more positive samples and two individuals providing eight positive stool samples. This can be seen in Figure 4.5 where the contours are present early on and are consistent throughout the study. Unvaccinated shedding occurs early in the study but the location of shedding is more variable over time.

Table 4.1. Characteristics of unvaccinated individuals by coverage area

	Coverage area			Total
	70% Vaccinated	30% Vaccinated	10% Vaccinated	
Households with GPS	126	136	161	423
Vaccinated households	80	40	17	137
Unvaccinated households, n (%)	46 (36.5%)	96 (70.6%)	144 (89.4%)	286 (67.6%)
Unvaccinated participants	161	396	588	1,145
shedding, n (%)**	24 (14.9%)	32 (8.1%)	24 (4.1%)	80 (7.0%)
Age (years), median (IQR)	12.2 (4.0 to 27.0)	17.0 (4.0 to 30.2)	18.0 (4.2 to 33.0)	17.0 (4.1 to 31.0)
Female, n (%)**	98 (60.9%)	223 (56.3%)	344 (58.5%)	665 (58.1%)
Running water	142 (88.2%)	343 (86.6%)	521 (88.6%)	1,006 (87.9%)
Samples provided n	1,294	3,202	5,563	10,059
Positive samples, n (%)	30 (2.3%)	33 (1.0%)	26 (0.5%)	89 (0.9%)
Spatial characteristics				
Distance to nearest vaccinated:				
household, median (IQR)	41.8m (19.7 to 63.3)	71.3m (38.1 to 99.8)	99.7m (55.3 to 151.8)	77.6m (42.0 to 126.1)
shedding household	45.0m (28.5 to 69.6)	82.4m (51.7 to 130.0)	112.1m (56.1 to 160.6)	85.0m (46.0 to 145.0)
Number of vaccinated shedding participants:				
within 100 meters, median (IQR)	3 (2 to 5)	1 (1 to 2)	1 (0 to 1)	1 (0 to 2)
200m	10 (8 to 17)	5 (2 to 8)	2 (1 to 3)	3 (2 to 6)
500m	58 (43 to 62)	22 (14 to 26)	8 (4 to 10)	10 (6 to 25)
1000m	78 (77 to 78)	38 (36 to 39)	15 (14 to 16)	16 (15 to 38)

*m = Meters ** IQR = Interquartile range

Table 4.2. Spatial characteristics comparing unvaccinated shedding with unvaccinated non-shedding individuals

Unvaccinated	N	Distance to nearest vaccinated (meters)		Number of vaccinated shedding households within:			
		Household	Shedding Household	100m	200m	500m	1000m
70% Coverage		Median (IQR)					
Positive	24	42.5 (19.7, 69.6)	44.9 (28.7, 69.6)	2.5 (2, 5)	10 (7, 19)	54 (43, 61)	78 (77, 79)
No Shedding	137	41.8 (18.5, 57.7)	45 (27.5, 71)	3 (2, 5)	10 (8, 16)	58 (43, 62)	78 (77, 78)
30% Coverage							
Positive	32	78.8 (53.6, 99.5)	83.5 (63.7, 124.1)	1 (1, 2)	6 (1.8, 8)	22.5 (16, 25.2)	38 (37, 39)
No Shedding	364	71.3 (37.9, 99.8)	82.4 (51.7, 130)	1 (1, 2)	5 (2, 8)	22 (13.8, 26)	38 (36, 39)
10% Coverage							
Positive	24	115.7 (56.1, 209.4)	115.7 (57, 211.8)	0.5 (0, 1)	2 (0, 3)	6 (3.8, 10)	15 (12.2, 15)
No Shedding	564	99.7 (55.3, 150.2)	112.1 (56.1, 160.1)	1 (0, 1)	2 (1, 3)	8 (4, 10)	15 (14, 16)

*m = Meters, IQR = Interquartile range

4.6.2. Spatial variables

The median distance to vaccinated shedding during the first 28 days was 85.0m (46.0 to 145.0m). The median distance at a given time point varied over the study period+ ranging from 126.1m on study day 4 to 1,626.3m on study day 71 when there were very few cases of vaccinated shedding (Figure 4.3). The median number of shedding individuals within 200 meters was 3 (2 to 6), further distances are shown in Table 4.1. There were no discernible patterns for proximity after stratifying by outcome and locality, these summary measures are presented in Table 4.2.

Figure 4.2. Distance from nearest vaccinated shedding household for unvaccinated individuals

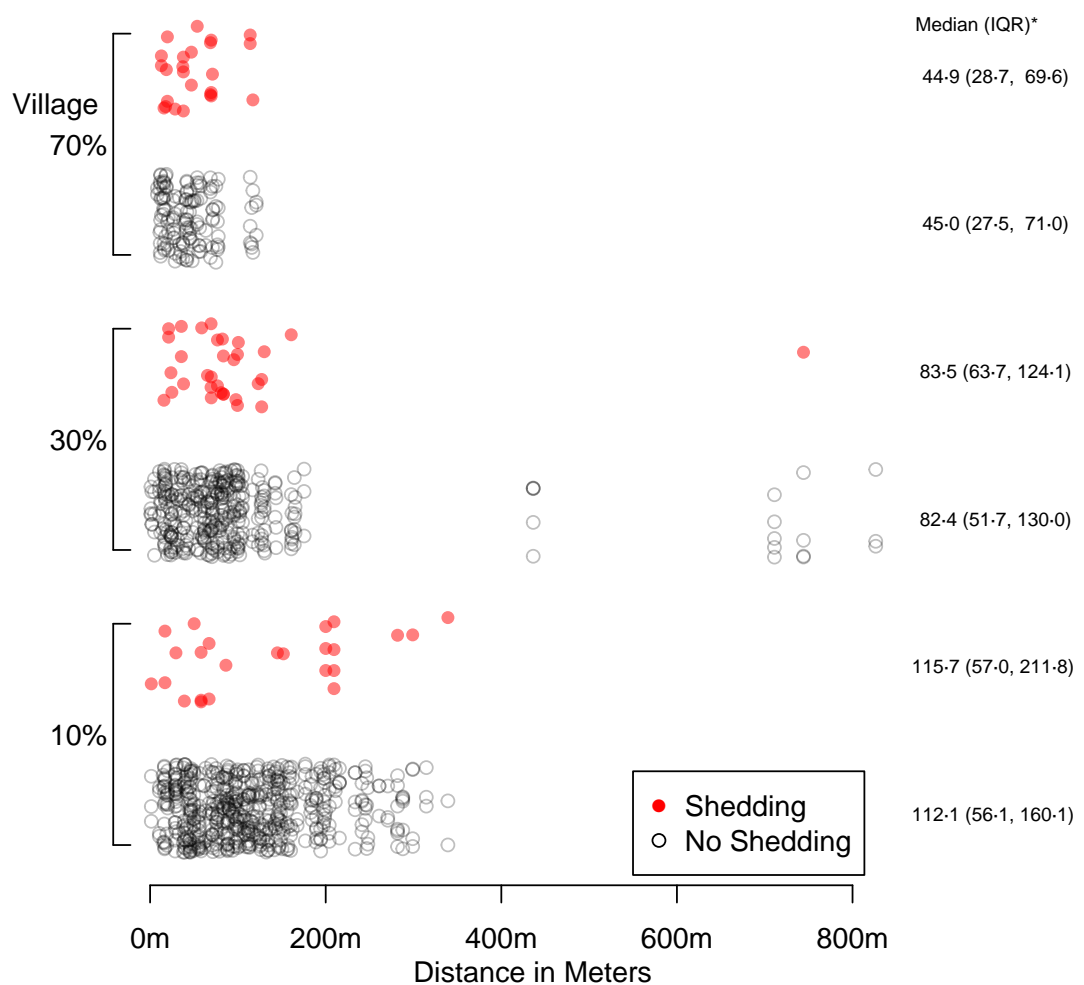


Figure 4.3. Median distance from vaccinated shedding over time

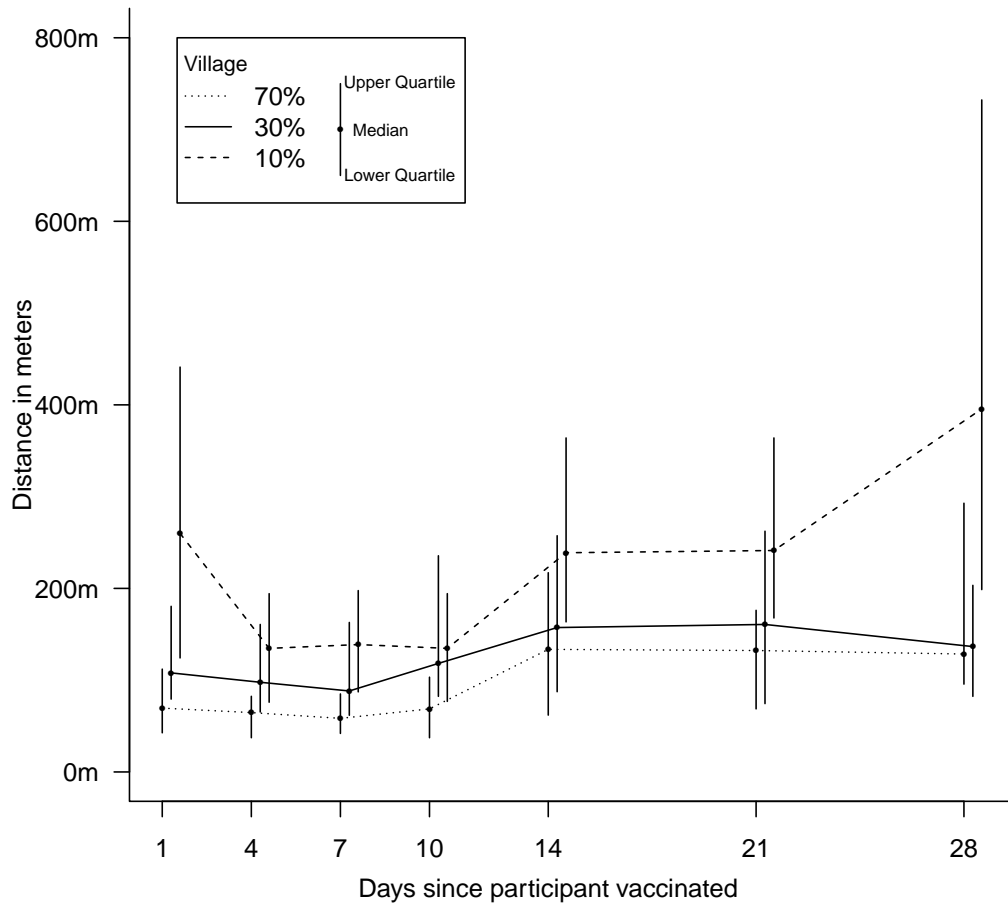


Figure 4.4. Legend for Figure 4.5

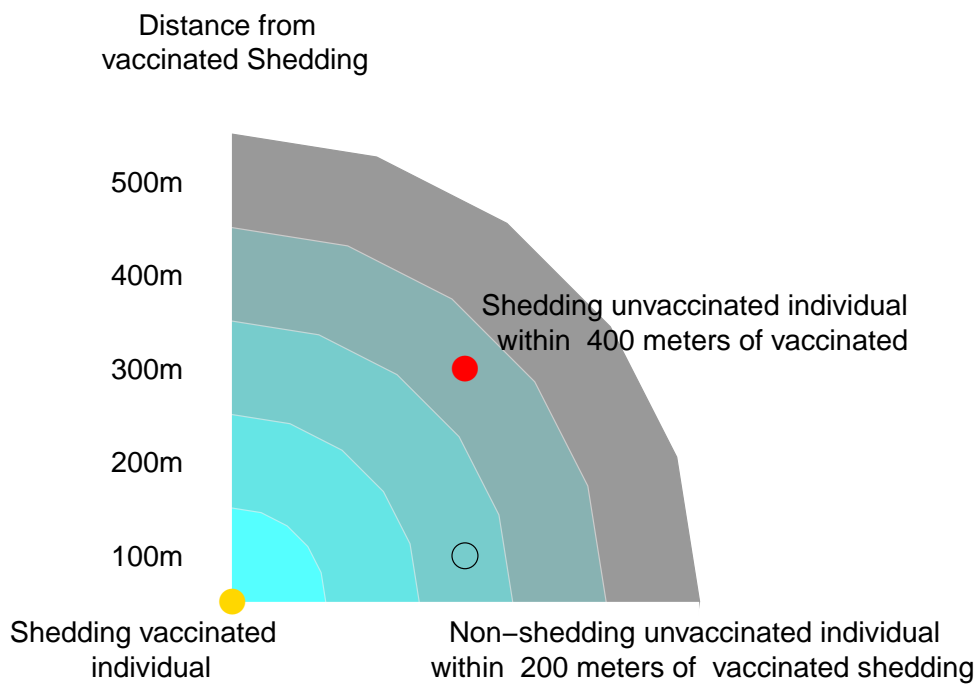
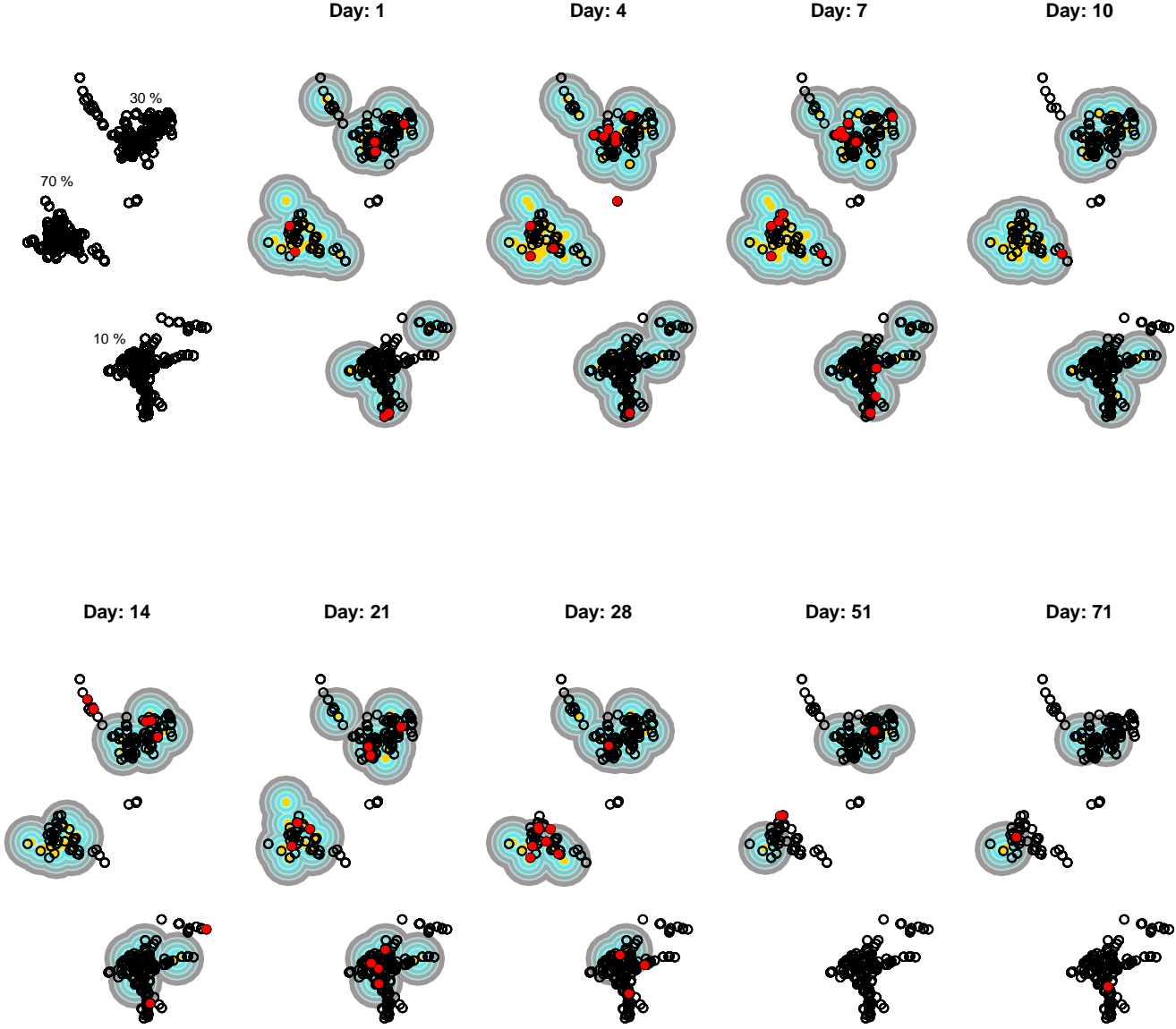


Figure 4.5. Spatial mapping of OPV transmission shedding over time



4.6.3. Models

After adjusting for age, gender, and running water there was very little suggestion (Odds Ratio [OR] 1.15 (95% Credible Interval [CrI] 0.86 to 1.46)) of an association between distance (per 100m) from vaccinated shedding and odds of shedding for unvaccinated individuals. This was consistent when considering non-linear effects of distance. Incorporating time into the analysis with an autoregressive lag of one resulted in comparable results with an OR 1.04 (CrI 0.92 to 1.16). There was also no indication that density of vaccinated shedding within 200m of an individual affects their odds of shedding with OR 0.99 (CrI 0.95 to 1.04). When including time, there was some suggestion that the number of shedding vaccinated individuals within 200m may have some effect on unvaccinated shedding with OR 0.93 (CrI 0.84 to 1.01). Results for other distances are displayed in Table 4.3. Restricting the analysis time to 14 days instead of 28 gave consistent results.

Table 4.3. Model considering effects of spatial variables on unvaccinated shedding

	Aggregated over time	Including Time
Distance to nearest vaccinated:	OR (95% CrI)*	
Household, per 100 meters	1.15 (0.86,1.45)	
shedding household, per 100m	1.15 (0.86,1.46)	1.04 (0.92,1.16)
Number of vaccinated shedding:		
100m	1.00 (0.87,1.13)	1.02 (0.82,1.21)
200m	0.99 (0.95,1.04)	0.93 (0.84,1.01)
500m	0.99 (0.98,1.01)	0.98 (0.96,1.01)
1000m	1.0 (0.97,1.03)	1.0 (0.98,1.02)

*Odds Ratio, 95% Credible Interval

4.7. Discussion

Through visualisation of transmission onto maps we were able to determine the dynamics of geospatial OPV transmission in a community with primary IPV-induced immunity. We found that shedding due to the introduction of OPV occurred rapidly, and was associated with between-household transmission on the first day of OPV vaccination. We found little evidence to suggest that living nearer to a household with a person who is shedding OPV affected the likelihood of shedding OPV up to the village dimensions of 850m. Indeed, there were no statistical differences in OPV acquisition among unvaccinated individuals based on distance from vaccinated individuals. In addition, the threshold for OPV dispersion appeared to be low; between household transmission in the 10% and 70% vaccination coverage communities were similar. Therefore, only a small amount OPV appeared to be needed for community transmission of OPV. This raises important implications about the impact of using OPV in future outbreaks and vaccination campaigns, especially as the transmission of OPV would usually be undetected, at least in highly vaccinated communities.

There are several strengths of this study. First, the study includes a large amount of individual level data with multiple observations per person. To the authors' knowledge previous spatial analyses have only been conducted at an aggregated level on WT polio, where the detection of cases was not through stool samples. As no other children were vaccinated with OPV in these communities until May 2015, we know the precise sources of OPV in these communities during the study period, giving us insight into what happens when OPV is administered at one time. Second, we were also able to consider variation of coverage as 10%, 30%, and 70% vaccinated cover of children was performed in three separate villages. Third, as one of the requirements of the study was that all vaccinated children had up-to-date IPV vaccinations, this mimics the transmission environment in future settings, as the Polio Endgame requires at least 1 dose of IPV in routine immunization schedules globally [2].

Using household location to represent a person’s location is at best an imprecise average of their movement throughout the study. Furthermore, it necessitates grouping people who live together to the same location. Information from contact tracing may have been useful to measure proximity of individuals, but might not be available in practice. It seems clear that transmission of polio is not purely spatial and when only household location is available it appears to have limited ability to predict transmission.

We tried to minimise misclassification of within-household transmission as between-household transmission by only using vaccinated individuals as point sources and unvaccinated households in the analyses. We excluded individuals who were unvaccinated and living with vaccinated individuals. Misclassification cannot be removed entirely due to the propagation of between- and within- household transmission. We also attempted to reduce the number of false positives for stool samples by using a two-step laboratory process to identify OPV by RT-QPCR described in more detail in the methods section, however, it is likely that there were small numbers of false negative and false positive samples. However, given the large number of individuals and stool samples collected, it is unlikely that small number of incorrectly identified samples would have affected the conclusions of this paper.

Our results show vaccinated children shedding as early as day 1 post-vaccination. This result is supported by prior OPV trials, where most vaccinated children shed within one week of vaccination [18–20]. That OPV can be transmitted to the contacts of vaccinated children has also been well-documented in these trials. Low levels of transmission also occurs as quickly as one week after vaccination as shown by transmissibility trials from the 1960s, results which are corroborated by more recent work in Zimbabwe looking at HIV-infected mothers with OPV-vaccinated children [18–21]. However, these studies collected samples on a weekly basis. Our samples were collected with more granularity and show that inter-household transmission occurred within one day of vaccination, even in the community with 10%

OPV coverage.

Household locations and spatial distribution appear to have limited use in predicting the transmission of poliovirus shedding. Therefore, in order to understand how OPV shedding occurs within a community alternative information such as contact patterns should be analyzed. The mechanism for predicting shedding is not well understood, including the role of number of prior IPV and OPV doses after vaccination.

The use of live poliovirus vaccine results in rapid dispersion and persistent transmission of polio virus throughout a community up to at least 71 days. The only way to avoid this is to not use OPV or to have strong controls such as quarantine, or strict hygiene protocols around the vaccinee. At present, what we observe in this study would not be detected through any clinical screening since all transmission was asymptomatic and detected by analysis of prospectively collected stool samples. Therefore, better methods such as collection and analysis of sewage samples are critically needed to ensure shedding has stopped within a community. Without this any conclusions about the eradication of circulating polioviruses are at best overconfident. These results further support The Polio Eradication and Endgame Strategic Plan 2013-2018 to withdraw all OPV vaccines by 2020 [2]. After withdrawal of OPV worldwide reintroduction of OPV due to an outbreak should be carefully considered as it appears a small amount of the vaccine may result in community transmission.

4.8. Bibliography

- [1] Global Polio Eradication Initiative and WHO. POLIO NOW. 2017. (Visited on 01/01/2017).
- [2] Morales, M and Rudolf H. Wassilak, SGT. Polio Eradication & Endgame Strategic Plan 2013-2018. *Global Polio Eradication Initiative* 2013 134.
- [3] Morales, M, Tangermann, RH, and Wassilak, SG. Progress Toward Polio Eradication - Worldwide, 2015-2016. *MMWR. Morbidity and Mortality Weekly Report* May 2016 65. (18) 470–473.
- [4] Burns, CC, Diop, OM, Sutter, RW, and Kew, OM. Vaccine-derived polioviruses. *Journal of Infectious Diseases* 2014 210. (Suppl 1) S283–S293.
- [5] Hampton, LM, Farrell, M, Ramirez-Gonzalez, A, Menning, L, Shendale, S, Lewis, I, Rubin, J, Garon, J, Harris, J, Hyde, T, Wassilak, S, Patel, M, Nandy, R, and Chang-Blanc, D. Cessation of Trivalent Oral Poliovirus Vaccine and Introduction of Inactivated Poliovirus Vaccine - Worldwide, 2016. *MMWR. Morbidity and Mortality Weekly Report* Sept. 2016 65. (35) 934–938.
- [6] Mustachhi, P, David, F, and Fix, E. Three Tests for Space-Time Interaction: A Comparative Evaluation. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Biology and Problems of Health*. 4. California: University of California Press, Berkeley, 1967:229–235.
- [7] Trevelyan, B, Smallman-Raynor, M, and Cliff, AD. The spatial structure of epidemic emergence: Geographical aspects of poliomyelitis in north-eastern USA, July-October 1916. *Journal of the Royal Statistical Society. Series A: Statistics in Society* 2005 168. (4) 701–722.
- [8] Trevelyan, B, Smallman-Raynor, M, and Cliff, AD. The spatial dynamics of poliomyelitis in the United States: From epidemic emergence to vaccine-induced retreat, 1910-1971. 2005.

- [9] Uptill-Brown, AM, Lyons, HM, Pate, MA, Shuaib, F, Baig, S, Hu, H, Eckhoff, PA, and Chabot-Couture, G. Predictive spatial risk model of poliovirus to aid prioritization and hasten eradication in Nigeria. *BMC medicine* 2014 12. (1) 92.
- [10] Sarnquist, C, Altamirano, J, and Maldonado, YA. Poliovirus vaccine transmissibility in communities after cessation of routine oral poliovirus vaccine immunization. *Clinical Infectious Diseases* 2018 Submitted.
- [11] Kilpatrick, DR, Yang, CF, Ching, K, Vincent, A, Iber, J, Campagnoli, R, Mandelbaum, M, De, L, Yang, SJ, Nix, A, and Kew, OM. Rapid group-, serotype-, and vaccine strain-specific identification of poliovirus isolates by real-time reverse transcription-PCR using degenerate primers and probes containing deoxyinosine residues. *Journal of Clinical Microbiology* 2009 47. (6) 1939–1941.
- [12] Altamirano, J, Purington, N, and Maldonado, YA. Characterization of Household Shedding and Transmission of Oral Poliovirus Vaccine in Mexican Communities with Varying Vaccination Coverage. *Clinical Infectious Diseases* 2018 Submitted.
- [13] Jarvis, C, Di Tanna, GL, Lewis, D, Alexander, N, and Edmunds, WJ. Spatial analysis of cluster randomised trials: A systematic review of analysis methods. *Emerging Themes in Epidemiology* 2017 14. (12).
- [14] Binka, FN, Indome, F, and Smith, T. Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. *The American Journal of Tropical Medicine and Hygiene* 1998 59. (1) 80–85.
- [15] Lenhart, A, Orelus, N, Maskill, R, Alexander, N, Streit, T, and McCall, PJ. Insecticide-treated bednets to control dengue vectors: Preliminary evidence from a controlled trial in Haiti. *Tropical Medicine and International Health* 2008 13. (1) 56–67.

- [16] Rue, H, Martino, S, and Chopin, N. Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical Society Series B* 2009 71. (2) 319–392.
- [17] R Core Team. R: A language and environment for statistical computing. Vienna, Austria, 2017. URL: <https://www.r-project.org/>.
- [18] Benyesh-Melnick, M, Melnick, JL, Rawls, WE, Wimberly, I, Oro, JB, Ben-Porath, E, and Rennick, V. Studies of the immunogenicity, communicability and genetic stability of oral poliovaccine administered during the winter. *American journal of epidemiology* 1967 86. (1) 112–36.
- [19] Paul, JR, Horstmann, DM, Riordan, JT, Opton, EM, Niederman, JC, Isacson, EP, and Green, RH. An oral poliovirus vaccine trial in Costa Rica. *Bulletin of the World Health Organization* 1962 26. (3) 311–29.
- [20] Fine, P and Carneiro, I. Transmissibility and Persistence of Oral Polio Vaccine Viruses: Implications for the Global Poliomyelitis Eradication Initiative. *American Journal of Epidemiology* 1999 150 1099–1102.
- [21] Holubar, M, Troy, SB, Nathoo, K, Stranix-Chibanda, L, Musingwini, G, Srinivas, N, Huang, C, Junn, A, Halpern, MS, and Maldonado, YA. Shedding of Oral Poliovirus Vaccine (OPV) by HIV-Infected and -Uninfected Mothers of OPV-Vaccinated Zimbabwean Infants. *Journal of the Pediatric Infectious Diseases Society* 2016 6. (1) 105–108.

5. Paper C: Spatial effects of permethrin-impregnated bed nets on child mortality: 25 years on, a spatial reanalysis of a cluster randomised trial

Christopher Jarvis^{1,2}, Lea Multerer^{3,4}, Daniel Lewis¹, Fred Binka⁵, W John Edmunds¹, Neal Alexander¹, Thomas A Smith^{3,4}

1. London School of Hygiene and Tropical Medicine, London, UK
2. MRC London Hub for Trials Methodology Research, London, UK
3. Swiss Tropical and Public Health Institute, Socinstrasse 57, CH-4002 Basel, Switzerland
4. University of Basel, Petersplatz 1, 4003 Basel, Switzerland
5. School of Public Health, University of Health and Allied Sciences, Ho, Volta region, Ghana

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk



Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Christopher Jarvis
Principal Supervisor	John Edmunds
Thesis Title	Spatial effects of permethrin-impregnated bed nets on child mortality: 25 years on, a spatial reanalysis of a cluster randomised trial

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	TBC
Please list the paper's authors in the intended authorship order:	Christopher Jarvis, Lea Multerer, Daniel Lewis, Fred Binka, W John Edmunds, Neal Alexander, Thomas A Smith
Stage of publication	Not yet submitted

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I planned and conducted the analysis, and drafted the article with feedback from co-authors.
--	--

Student Signature: _____

Date: 28/3/2018

Supervisor Signature:



Date: 28/3/2018

5.1. Overview

The previous chapter involved an applied spatial analysis of a household cluster randomised trial (CRT). In this chapter, I apply spatial methods to a geographical CRT where the clusters are polygons, and the observations are points. The data is from the earliest example found in the systematic review from chapter 3. I conduct a spatial reanalysis of the effect of insecticide-treated bed nets, applying a range of spatial methods. This chapter also includes the first application of the *cluster reallocation* method which is further detailed in chapter 8. This paper is yet to be submitted to a journal.

Objective

3. Apply and assess a range of appropriate modern spatial methods to existing CRT data, in order to analyse the effect of spatial autocorrelation and spatial spillover effects on CRT results.

5.2. Role of candidate

I conceived of and conducted the statistical analysis with the exception of the bootstrapping which was conducted by Lea Multerer (LM). Fred Binka and Thomas A Smith (TS) provided permission to use and access to the data. I drafted the initial paper with feedback, input, and critique provided from LM, TS, Daniel Lewis, W John Edmunds, and Neal Alexander. I developed the cluster reallocation method that is applied in this chapter.

5.3. Abstract

Background

Insecticide-treated nets (ITN) have been proven to be an effective intervention, reducing the risk of mortality and the burden of malaria. In addition to the direct effect, there is evidence for positive spillovers, or spatial indirect effects. Spatial analyses in cluster randomised trials (CRTs) are rare, but a large scale CRT from 1993 was one of the first to conduct a spatial analysis of ITNs in CRTs. We revisit this data to apply a broader range of spatial methods to further explore spatial spillover, and demonstrate the extra utility that spatial methods can provide for CRTs.

Methods

We conducted three broad sets of analyses: (1) Exploratory spatial analysis, considering spatial dependence, heterogeneity, and spillover in the data; (2) Spatial modelling of the intervention effect, to estimate the true intervention effect in light of any anticipated spatial effects; (3) Analysis of distance based spillover and interaction with the intervention, to characterise the functional distance over which the spillover effect is present.

Results

There was a consistent suggestion of spatial patterns from the exploratory analysis. Bed nets were associated with a 17% reduction in all-cause mortality for children aged 6-59 months, and the intervention estimate remained robust when allowing for the spatial structure of the data (standardised mortality ratio [SMR] 0.83 95% confidence interval 0.71 to 0.98). There was strong evidence of a spatial spillover effect, with every additional 100m a control household was from an intervention

household (and vice versa), the SMR increased by 1.7% (SMR 1.017 95% credible interval 1.006 to 1.026).

Conclusions

Despite evidence of a spatial spillover effect, the conclusions of the trial remain unaffected by spatial model specifications. Use of ITNs is clearly beneficial for individuals and there is compelling evidence that they provide an indirect benefit to individuals living nearby. This paper demonstrates the extra utility that spatial methods can provide when analysing a CRT.

5.4. Background

A series of cluster randomised trials (CRTs) carried out two decades ago in endemic areas in Africa demonstrated strong evidence that insecticide treated bed nets (ITNs) can reduce child mortality [1–5]. For instance, a large-scale CRT of ITNs in the Kassena-Nankana (Navrongo) district of northern Ghana found a 17% reduction in all-cause child mortality in children aged 6 months to 4 years (standardised mortality ratio [SMR] 0.83 95% confidence interval [CI] 0.69 to 1.00) [3]. This study began in July 1993 and provided 31,000 ITNs to intervention participants in 48 geographically defined polygon clusters. A meta-analysis of all the trials estimated an average reduction in all-cause mortality of 18% (risk ratio [RR] 0.82 95% CI 0.76 to 0.89) [6].

In addition to the direct effect, there is evidence for positive spillovers, or spatial indirect effects [6–8]. Using the data of the Navrongo CRT, a subsequent study by Binka et al. found decreases in mortality among individuals without ITNs who happened to live close to ones who did [9]. This study was the earliest example found in a recent systematic review of spatial analyses in CRTs [10]. The ‘positive spillover’ is evidence of a spatial indirect effect of the bed net intervention, in addition to the direct effect, that ITNs reduce mortality for those using them. Binka et al. used the locations of households to gain further value from the trial data allowing exploration of spatial indirect effects. Subsequent analyses have also demonstrated positive spillover with ITNs [5, 7, 8, 11, 12] although at least one study failed to find evidence of spatial indirect effects [13].

The approach to estimating the spatial indirect effect used by Binka et al. [9] was novel at the time, but in subsequent years, the emergence of a subdisciplinary focus on spatial epidemiology has led to a broader range of applicable methods [14, 15]. The specification of new spatial models, particularly in light of the continued growth of computational capacity, and refinement of optimisation methods has made advanced spatial regression approaches tractable. In this paper, we revisit the

Kassena-Nankana CRT using contemporary spatial methods to: explore the existence of positive spillovers; estimate the spatial indirect effects; and, consider their impact, if any, on the overall trial conclusions.

To the authors' knowledge, this paper provides the first time that many of these methods have been applied to a CRT. In addition, we propose a new method called *cluster reallocation* which allows trialists to consider if spatial spillover is present in a CRT. As well as demonstrating the application of contemporary spatial methods, this enhanced and extended reanalysis demonstrates the additional utility of collecting GPS coordinates during trials. We argue that beyond being a useful resource for trial management, or for mapping trial context, an explicit analysis of location can yield important additional information about the functioning of particular interventions in CRTs.

5.5. Methods

Study

The trial was conducted between July 1993 and June 1995 in the Kassena-Nankana district in the Upper East Region of Ghana. The study design has been described previously [3]. In short, a parallel CRT with 96 geographically contiguous clusters (Figure 5.1) with an average of roughly 1400 persons per cluster, and an average of 124 households per cluster.. The intervention of permethrin-impregnated bed nets was allocated to 48 clusters and the outcome was all-cause mortality in children aged 6 months to 4 years.

Data

The data contains one record per household, with variables for the location coordinates (Westings and Northings projected in WGS 84 / UTM zone 30N), and the

observed and expected number of deaths per household. As per Binka et al. [9] “The expected number of deaths for each cluster was calculated by applying age-specific death rates derived from preintervention population to the postintervention time at risk.” The distance from intervention households to the nearest control and vice versa, referred to as the distance to discordant pair, had been calculated previously [9]. These distances were verified by calculating Euclidean distance based on the UTM coordinates in QGIS 2.18.14 [16] and R 3.4.2 [17]. The outcome is a SMR and was calculated by dividing the observed deaths in each household by the expected number of deaths.

Statistical Analyses

We conducted three broad sets of analyses: (1) Exploratory spatial analysis, considering spatial dependence, heterogeneity, and spillover in the data; (2) Spatial modelling of the intervention effect, to estimate the true intervention effect in light of any anticipated spatial effects; (3) Analysis of distance based spillover and interaction with the intervention, to characterise the functional distance over which the spillover effect is present. R 3.4.2 [17] was used for all statistical analyses with the following packages [18–31].

Exploratory spatial analysis

We explore spatial patterns in the intervention assignment through the use of a join count statistic [32], and spatial correlation of the outcome is assessed using *Moran’s I* [33]. Spatial heterogeneity of the effect of bed nets over the study region was considered through the use of geographically weighted regression (GWR) [34]. Evidence of a spillover effect across cluster boundaries is assessed using a novel method we developed called cluster reallocation. The spatial patterns are also assessed visually through the use of maps. To the extent of our knowledge, with the exception of Moran’s I, this is the first time these methods have been applied to a CRT.

The join count statistic was calculated to assess whether there was a spatial pattern in treatment assignment. The join count method assesses spatial correlation for binary variables and involves counting all pairs of neighbouring (Queen's case) clusters in the trial by type of adjacency: intervention-intervention, control-control, or intervention-control [32]. Neighbours were defined using Queen's case where clusters that share boundary or vertex are considered neighbours. Using a hypothesis testing framework, we then assess whether the observed counts of these three possible adjacencies in the trial deviates from the expected counts based on a random pattern [33]. Although the allocation of intervention to clusters is based on a random process it could still result in a non-random spatial pattern. For example, randomisation could result in all control clusters being in one area and all intervention clusters being in another area of the study. In this case, the study area could be split into two sections, an area with only intervention cluster present and an area with only control cluster present, this may present issue when trying to measure spatial effects as few intervention clusters may border control clusters.

Moran's I statistic was used to assess the presence of global spatial autocorrelation in the SMRs at cluster level. Moran's I is an extension of Pearson's product-moment correlation into two dimensions, it considers the strength of association and the spatial lag over which it is present. The SMR was calculated for each cluster, with a binary spatial weights matrix (Queen's case) used to represent the connectivity between clusters. The spatial weights matrix takes a value of one if the clusters share a boundary or vertex, and zero otherwise. Moran's I was calculated for the whole study area, and for the control and intervention clusters separately. Moran's I was also calculated on the Pearson's residuals from a multi-level model with a random effect for cluster and a fixed effect for intervention. The residuals were aggregated to cluster level, this calculates Moran's I adjusted for the intervention effect and the clustering.

P-values for Moran's I were calculated using Monte Carlo simulation. This was

achieved through a permutation test where the values for each cluster are shuffled to different locations and the statistics recalculated. This process is repeated many times and the observed value is compared to the sampling distribution of the simulated values to test for evidence of deviation from a random spatial pattern.

Spatial heterogeneity of the intervention estimates were explored using GWR [34]. This method involves applying a regression model to a spatial subset of the data (a neighbourhood) and then recording the coefficients of that model, a different neighbourhood is then chosen and the model reapplied. This process is repeated over the entire study area to give one estimate of the coefficient for each neighbourhood. The neighbourhood is typically a radius around each point. The distribution of the coefficients can then be explored visually on a map, helping to determine sources of heterogeneity in the data. In situations where data are spatially heterogeneous, GWR produces coefficient estimates that vary over space, indicating local areas of departure from a global process. A Poisson regression model without a random effect for cluster was used for the GWR.

To assess for the presence of spatial spillover we developed a new method called cluster reallocation which considers how changes in the definition of cluster boundaries affects the intervention estimates of the trial (see chapter 8 for more details). Cluster reallocation is a computationally intensive method that involves reallocating individuals to either the intervention or control arm based on their proximity to cluster boundaries. At each step, a model is applied to estimate the intervention effect. In this analysis, clusters were dilated (buffered) incrementally between 0m (original case: no change to cluster) and 1000m in steps of 100m. The process was carried out independently for intervention and control clusters.

At each 100m increment, households were reassigned to intervention or control clusters, and the main trial model was refitted. In the absence of spillover, we hypothesised that as the size of either control or intervention clusters grew, estimates of the intervention effect should attenuate to the null as differences between the inter-

vention and control arms are diluted. However, if alternatively, a spatial spillover is present, then we would expect the magnitude of the intervention effect estimate to increase over the functional distance of the spillover, as the intervention cluster is dilated.

Modelling spatial dependence

This main result of the original trial paper [3] was replicated using a multi-level model with a random effect (IID) for cluster and a fixed effect for the intervention. Three approaches were considered for adjusting for spatial structure. A conditional autoregressive model (CAR) [35–37], a Besag, Yorke and Mollie model (BYM) [38], and a Gaussian Process model (GPm) [39, 40].

The three spatial models were chosen for the different ways they incorporate spatial dependency. The approaches differ by whether they assume the underlying spatial process is discrete, called a Gaussian Markov random field (GMRF), or continuous, called a GP. A GMRF is a collection of spatially indexed random variables with a Markov property, where all possible combinations of the random variables are multivariate normally distributed (MVN) [41]. GMRFs are commonly used when data is recorded for distinct areas covering an entire region, such as clusters in a CRT. A GP assumes that the spatial process is MVN typically with a mean of zero and a covariance function that incorporates distance, therefore, changes in outcome due to the spatial process are a function of distance. GPs are commonly used when data is recorded for some points in an area and information is missing in other locations, such as households in a CRT [42]. In classical spatial statistics, GMRFs refers to areal data and GPs to geostatistical models [15, 36].

The CAR and BYM models incorporate spatial structure at cluster level. Fitting a CAR model to the data requires the aggregation of households to clusters as the model is restricted to only one observations per spatial area. In contrast, the BYM model can be fitted to the individual (household) level data. The CAR model

includes a spatially structured random effect for the clusters. The BYM model extends the CAR model to include an additional independent random effect and relaxes the link between one observations per cluster. The spatially structured random effect relaxes the assumption of independence between clusters, and allows clusters that are adjacent to share information. A binary spatial weights matrix (Queens’s case) was used where clusters that share a common boundary or vertex are considered to be adjacent. The BYM model effectively reproduces the standard approach to analysing a CRT, but with an additional spatial effect at the cluster level.

The GPM incorporates spatial structure at an individual (here: household) level. GPMs tend to have dense spatial matrices which makes computation difficult. Fortunately a link between the continuously indexed Gaussian process (GP) and the discrete indexed GMRFs has been proposed which uses stochastic partial differential equations (SPDE) [43]. In short, Lindgren et al. demonstrate that a Matérn covariance model (A GP with a constant mean, and Matérn covariance) is a solution to an SPDE [43]. They show that approximating an area with a finite number of triangles or a ‘mesh’, allows the solution of the SPDE to be represented as the weighted sum of the vertices of the ‘*mesh*’. Then assuming a Markov property on the mesh, it can be modelled as a GMRF. The GP is a solution of an SPDE, and the SPDE can be approximately solved by using a GMRF. Thus, the GP can be modelled using GMRF methods through the use of the mesh. The choice of mesh is a trade-off between how accurately the area can be represented and computational costs. Further adjustments can be made so that the mesh is finer in locations with data and less fine where there is less data (or information). Further details of the SPDE approach are described by Simpson et al. and Blangiardo et al. [44, 45].

The four models in this section were fitted using integrated nested Laplace approximation (INLA), using noninformative priors [21]. INLA is a deterministic algorithm which has proven to be capable of providing fast and reliable results for a wide range

of models [45, 46]. INLA was well suited to this analysis due to the complexity of the model types and the size of data when accounting for spatial structure (12,000 observations giving a spatial weights matrix with 144 million elements). Once fitted the posterior distributions of all models were sampled and the mean and 95% credible intervals (CrI) calculated. Further details of the models and spatial modeling is provided in Appendix C.

Spillover effect and interaction

A frequentist approach was used for exploring spillover and interaction as this is more commonly used in analysis of CRTs [47]. Evidence for positive spillover is assessed directly, by including distance to discordant pair as a variable in the IID model. The form of the distance variable was explored further to account for non-linearity by using quadratic terms. In geographic literature, this form of variable is often referred to as ‘distance decay’ [48].

Effect modification of the intervention effect by distance to discordant pair was assessed by including distance as a continuous variable using the IID model with an interaction term. For ease of interpretation an interaction term with a binary distance variable (threshold 400m). The distance of 400m was chosen as the spillover effect appears to attenuate at this distance based on the exploratory spatial analyses.

Robustness to distributional assumptions for the spillover effect was assessed using bootstrapping [49, 50]. The opinion in the literature about how to bootstrap from hierarchical data (clusters and households) diverges [51]. The simplest approach is to ignore the hierarchy and just select bootstrap samples from the households. This might lead to an unbalanced design, but in this data set, the number of households per cluster is varying anyway. We checked whether results differ when we just resample on the cluster level (each cluster then keeps the same amount of households, but there is less variation). The results were the marginally different and we therefore chose the simpler approach. We calculated 1000 bootstrap samples, and running the

simulations for 100 samples gave consistent results, suggesting that 1000 iterations is adequate. These were performed at sciCORE scientific computing core facility, University of Basel (<http://scicore.unibas.ch/>).

Distance to discordant household was also fitted for the BYM and GP models (using INLA and therefore Bayesian inference) to see if accounting for spatial structure affected the estimate of the spillover effect of the intervention.

5.6. Results

General characteristics

Over the course of the trial there were 861 deaths, resulting in an SMR of 1.24. The SMR in the control arm was 1.37 and was 0.24 units higher than in the intervention arm (SMR 1.12). The median distance to nearest discordant household was 805m (interquartile range [IQR] 311 to 1,106m). The median distance to nearest bed net for control households was 788m (IQR 305m to 1010m). Just over 50 percent of control households were within 600m of their nearest intervention household. The main trial results and a spatial analysis have been presented previously [3, 9].

Exploratory spatial analysis

There was no suggestion of a spatial pattern for the allocation of treatment for either the intervention or the control, as can be seen in Table 5.1 (join count statistic). There was strong evidence ($p < 0.001$) to suggest a spatial pattern for the SMR aggregated at cluster level. This pattern was consistent when only considering the control clusters, and when tested on residuals after adjusting for intervention, but was no longer apparent when only the intervention clusters were tested.

The IQR for the GWR intervention estimate ranged from 0.77 to 1.09 suggesting some spatial heterogeneity. and that there are areas where the intervention estimate

is greater than one. Figure 5.1, presents maps of the intervention assignment, spatial distribution of SMR at cluster level, and the intervention estimates from GWR. There is no obvious pattern, with the effect of bed nets appearing unrelated to whether surrounding areas have high and low mortality ratios, or the density of intervention clusters nearby.

Table 5.1. Summary of analyses of spatial patterns of dependence, heterogeneity, and spillover

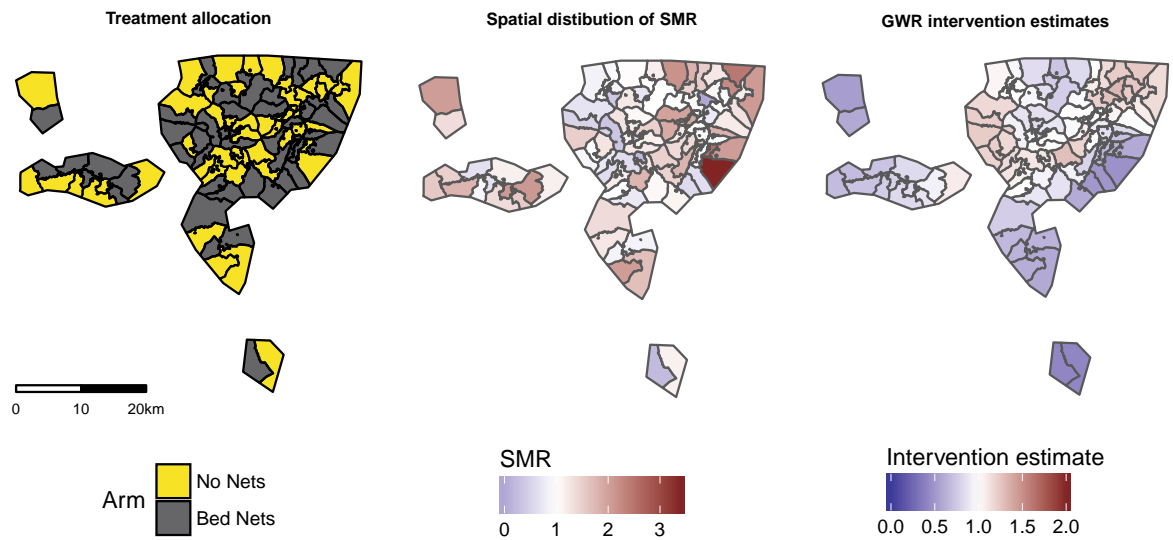
Method	Test Statistic	P-value
Spatial correlation of intervention allocation		
Join Count – Control	10.105	0.979
Join Count – Intervention	11.282	0.744
Spatial correlation of SMR		
Moran’s I – entire study area	0.237	<0.001
Moran’s I – control	0.396	<0.001
Moran’s I – intervention	0.095	0.176
Moran’s I – residuals	0.134	0.0198
Spatial heterogeneity of intervention		SMR
GWR – Median (IQR)		0.94 (0.77 to 1.09)
Impact of spillover on intervention effect*		SMR
Cluster reallocation method		
Original cluster definition		0.827
Controls cluster larger – Mean (Min, Max)		0.885 (0.855 to 0.916)
Intervention clusters larger – Mean (Min, Max)		0.781 (0.742 to 0.827)

GWR = Geographically weighted regression

SMR = Standardised mortality ratio

*This summary represent the mean of the intervention estimates that derive from increasing either control or intervention cluster boundaries based on the cluster reallocation method

Figure 5.1. Study area, spatial variation of SMR, and spatial heterogeneity of intervention effect from GWR model.

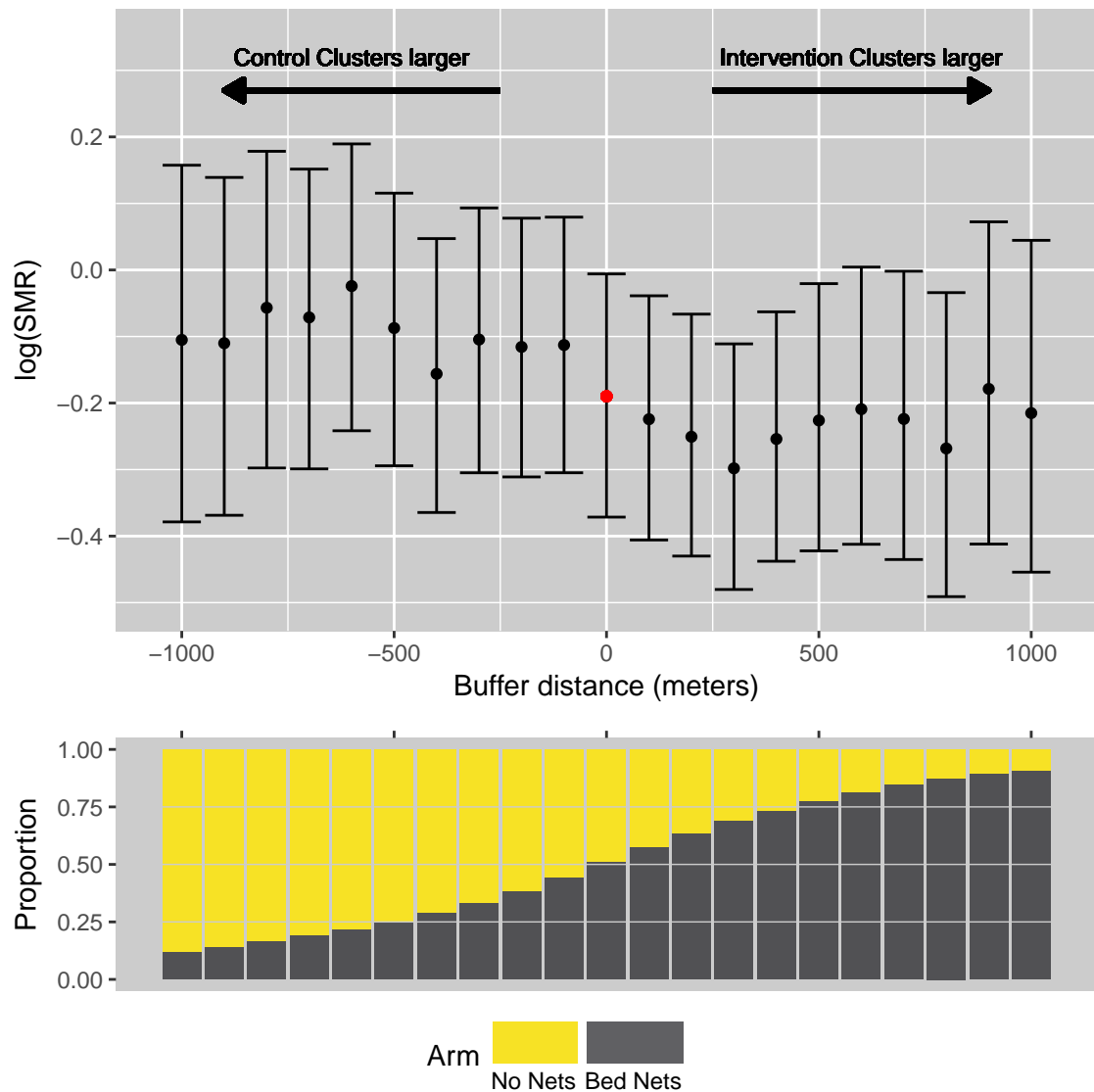


The cluster reallocation method provides a strong suggestion of a spillover effect from intervention clusters towards control clusters. Expanding the intervention cluster boundary resulted in stronger (bed nets more effective) intervention estimates compared to the original cluster definitions up to around 400m (Figure 5.2). This is consistent with individuals within 400m of the intervention receiving an indirect benefit due to proximity. Expanding the control cluster attenuated the effect estimate to the null. This is consistent with an absence of spillover from control to intervention as with each increase in buffer the newly defined intervention and control arms contain more similar participants. Past 500m the number of participants in each arm becomes very unbalanced and the point estimates from these models should be treated with caution (Figure 5.2 lower graph). However, even when the buffering was at 1000m, the smaller arm still had greater than 1,000 observations, which is reflected in the consistent size of the confidence intervals.

Analysis of the spatial autocorrelation of cluster level SMRs, the heterogeneity in the intervention estimate over space, and the behaviour of the intervention estimate under cluster dilution are all suggestive of a spatial spillover effect from the intervention to the control clusters. Further, this effect is unlikely to be an artefact of

the initial intervention allocation to the clusters, as there is no evidence of a spatial pattern in the cluster allocation.

Figure 5.2. The change in effect estimate calculated by cluster reallocation of intervention participants to the control arm and vice versa.



Spatial models

Bed nets were associated with a 17% reduction in all-cause mortality for children aged 6-59 months. For the CAR and BYM models, adjusting for cluster-level spatial effects made negligible difference to estimated effects, and no difference to the conclusions of the trial. The GPm model also resulted in near identical estimates

to that of the non-spatial model (Table 5.2).

Adjusting for distance to discordant pair did not greatly influence the estimate of the intervention effect and this was consistent even when taking account of spatial structure of the trial using the BYM and GP approaches.

Table 5.2. Estimate of intervention effect by spatial model type

Model	SMR (95% CrI) of Intervention*	Description
IID	0.83 (0.71 to 0.98)	Multi-level model with independent random effect for cluster
CAR	0.84 (0.72 to 0.98)	Conditional Autoregressive model
BYM	0.84 (0.72 to 0.98)	Besag ,Yorke and Mollie model
GPm	0.83 (0.70 to 0.97)	Gaussian process model fitted using Stochastic Partial differential equation approach
IID + Discordant distance	0.82 (0.70 to 0.95)	Distance to discordant pair included in the model
BYM + Discordant distance	0.84 (0.72 to 0.98)	
GPm + Discordant distance	0.82 (0.70 to 0.96)	

*CrI = Credible interval

Spillover effect and interaction

Distance to discordant pair was strongly associated with child mortality ($p = 0.005$), for every additional 100m a control household was from an intervention household (and vice versa), the mortality increased by 1.7% (SMR 1.017 95% CI 1.006 to 1.026). The result was consistent when adjusting for spatial structure, and was robust to distributional assumptions with a bootstrapped estimate of 1.014 (95 % CI 1.007 to 1.027) [Table 5.3].

There was no statistical evidence ($p = 0.214$) for effect modification between the distance to discordant pair and use of bed nets. However, the study was not powered to test for interaction, and thus likely has low power [52]. Contrastingly, to the statistical evidence, the stratum specific SMRs were suggestive of interaction.

Distance from bed net households was associated with reduced mortality, but distance from control household was not. Treating distance to discordant pair as binary (threshold of 400m), suggests that use of bed nets does not reduce mortality for individuals living within 400m of bed net households but does reduce mortality for

those living further than 400m away (Table 5.3). These results are consistent to what was found in the original and spatial analysis of the Binka et al. trial [3, 9].

Table 5.3. Spillover effect of distance to discordant pair presented by model type and interaction of distance with bed net intervention.

Variable	Model	SMR (95% CrI)
Distance to discordant pair (per 100 meters)	IID	1.017 (1.006 to 1.026)
	BYM	1.012 (1.004 to 1.020)
	GP	1.018 (1.005 to 1.029)
	Bootstrapped Model (95% CI)	1.014 (1.007 to 1.027)
Stratum specific SMRs	(Global test for interaction $p = 0.214$)	SMR (95% CI)
Distance to discordant	Intervention	
400m or nearer	No bed nets	1.00
	Bed nets	0.95 (0.71 to 1.25)
Further than 400m	No bed nets	1.00
	Bed nets	0.78 (0.64 to 0.95)
Intervention	Distance to discordant pair	
Bed nets	400m or nearer	1.00
	Further than 400m	1.11 (0.88 to 1.38)
	400m or nearer	1.00
No bed nets	400m or nearer	1.00
	Further than 400m	1.34 (1.08 to 1.67)

CI = Confidence Interval CrI = Credible Interval

5.7. Discussion

We examined the existence of spatial spillovers and the impact of spatial effects on the overall trial conclusions in a CRT of ITNs. Multiple approaches strongly suggest evidence of a positive spatial spillover effect due to being near households who use bed nets. Allowing for detailed spatial correlations and spillover effects did not change the primary conclusions of the trial [3].

Our analysis suggests, for every additional 100m a control household was from an intervention household (and vice versa), the standardised child mortality ratio increased by 1.7% (SMR 1.017 95% CrI 1.006 to 1.026). Bed nets were associated with a 17% reduction in all-cause mortality for children aged 6-59 months (SMR 0.83 95% CrI 0.71 to 0.98). This effect estimate remained robust to models with different spatial specifications and raises the question of whether standard CRT analyses are always robust to spatial spillover effects. This trial was over a large study area, with large clusters, meaning that spillover effects may need to be substantial in order

to impact the study results. An alternative explanation is that the spatial models were also subject to the same biases as the standard CRT model. Spatial effects of different strengths and distances could be tested through simulation studies to test the impact on CRT results.

Although there was no statistical evidence of interaction ($p=0.214$), the stratum specific effects suggest that distance from bed net households affects mortality, but distance from control households does not. Furthermore, it suggests that the use of bed nets was more effective when comparing individuals more than 400m apart. This interaction effect is plausible and assessment of interaction between distance from bed nets and use of bed nets should be assessed in other ITN CRTs.

We were surprised at the strength of the spillover effect, and used bootstrapping to test robustness to distributional assumptions of the model obtaining consistent results (SMR 1.014 95% CI 1.007 to 1.027). The spillover effect could be explained due to reduced populations of mosquitoes in areas near to ITNs and is consistent with ‘mass killing’ effects, found previously [7, 12]. Adjusting for distance to discordant household had negligible impact on the main trial result.

There are several limitations of this analysis, the data was aggregated at household level, which may have resulted in potential loss of information. However the results were consistent with an individual level analysis and any spatial analyses would require aggregating the spatial information at a household level. There were also gaps between some of the clusters where no spatial data was collected, which may have affected results.

A further weakness is the use of household coordinates to represent the spatial structure or mechanism of the spillover. Household location is at best one of the many locations that an individual visits during the study period. This results in omission of many of the spatial locations related to the participants. This is probably of minimal impact in the context of mosquitoes and ITNs where transmission is likely to happen at night when at home. Despite this, the value of household location in

other contexts should be considered. In future, the possibility of tracking movement of people or mosquitoes may provide improved insights into the mechanism behind spatial spillover effects.

These analyses demonstrate that collection of GPS data allows exploration of intervention mechanisms beyond the creation of maps. Our analyses only required GPS coordinates and did not require any new data collection. This approach could be used to reanalyse previous geographical CRTs that collected coordinates, thus gaining extra utility from previously collected data. We explored a range of spatial methods allowing for comparison of different spatial methods and the conclusions of our analysis remained consistent with the original analyses conducted in the 1990s. These analyses add to the ever expanding literature on the spatial indirect effects of ITNs on mosquitoes and spatial analyses of CRTs [6, 10].

In summary, despite evidence of a spatial spillover effect, the conclusions of the trial remain unaffected to spatial model specifications. Use of ITNs is clearly beneficial for individuals and there is compelling evidence that they provide an indirect benefit to individuals living nearby. Although this paper demonstrates robustness of CRT analyses to spatial effects, this is just one scenario, where the clusters may be large compared to the spillover. Simulation studies could be used to evaluate the robustness of intervention estimates in CRTs for differing distances and strengths of spatial effects.

5.8. Extended discussion of applied spatial methods

This section is an extended discussion of the spatial methods used in the paper and their utility in spatial analysis. This section will be omitted from the submitted paper, as it is of greater methodological focus, but is of relevance to the thesis.

This chapter has applied several spatial methods to examine the existence of spatial spillovers, and the impact of spatial effects, on the conclusions in a CRT of insecticide

treated nets (ITNs). The analyses were split into 3 parts: (1) Exploratory spatial analysis, considering spatial dependence, heterogeneity, and spillover in the data; (2) spatial modelling of the intervention effect, to estimate the true intervention effect in light of any anticipated spatial effects; (3) Analysis of distance based spillover and interaction with the intervention, to characterise the functional distance over which the spillover effect is present.

5.8.1. Exploratory spatial analysis

The join-count statistic was used to test for a spatial pattern in the randomised allocation of the clusters. Although the process of assignment is random, it could still result in spatial patterns. This method does not appear to have great utility in the analysis of CRTs. It could perhaps be used for restricted randomisation, to prevent the trial from ending up with effectively one large control, and one large intervention cluster.

Moran's I was used to assess for presence of spatial correlation in the SMR at a cluster level. It resulted in strong evidence of spatial correlation for the overall study area, the control clusters, and the residuals of a model adjusting for intervention. When only the intervention clusters were considered there was no evidence of spatial correlation. This method appears useful, especially when considering an intervention that interrupts transmission. In this case, the absence of spatial correlation, may be reflective of an intervention working.

GWR was used to explore spatial heterogeneity. This method resulted in a strong suggestion of a spatial pattern, with some areas having a harmful intervention estimate, and others areas a beneficial effect estimate. GWR could be used in future trials to explore for patterns, and may help to identify spatial interaction of the intervention. GWR may help identify situations where areas with low prevalence are less affected by the intervention. The results from GWR appear most useful when presented with a fuller picture of spatial features of a CRT. In this chapter

the intervention status, and the SMR in each cluster were presented alongside the GWR intervention estimates. There was no clear relation between the three plots, which suggests that other characteristics may be driving the spatial patterns.

Cluster reallocation was also applied in this paper, details of the method will be explained later in chapter 8. The method was suggestive of a spillover effect from intervention to control cluster. The estimate of uncertainty changed only slightly depending on the buffer size, and this is most likely attributed to the sample size of the study. The utility of this method beyond this application is hard to assess from this chapter alone, and will be left for chapter 8.

5.8.2. Modelling spatial dependence

Although the CAR, BYM, and GPm models were chosen because they differ in the level at which they model the spatial effects there was very little difference to the intervention estimate of all three models. Furthermore, the intervention estimates were near identical to the non-spatial model. It is therefore difficult to assess how appropriate these methods are for analysis of CRTs, and instead raises questions on the robustness of the IID model in the presence of spatial effects. The BYM and the GPm models will be tested under simulated conditions in chapter 7.

Spillover effect and interaction

In this study distance to discordant pair was fitted into a model to test for spillover. Although the variable was strongly suggestive of a spillover effect, distance to discordant pair represents the distance from intervention to control, and from the control to intervention. Therefore, this variable assumes spillover from both sides of cluster boundaries. Based on the analyses in this chapter, and the plausible mechanism of spillover, it is likely a positive spillover effect. This suggests that the association of distance to discordant pair with mortality is driven by the control observations dis-

tance from intervention. Exploring interaction was consistent with this as distance from control household was not associated with mortality. The spatial variable approach was useful in confirming the presence of spatial spillover, and using an interaction helped to further explore the mechanism of the intervention.

In conclusion, the spatial methods have provided extra utility beyond a typical CRT analysis. The range of methods applied here could be used in other contexts, and might help to better understand the different processes that occur during a CRT.

5.9. Bibliography

- [1] Habluetzel, A, Diallo, DA, Esposito, F, Lamizana, L, Pagnoni, F, Lengeler, C, Traore, C, and Cousens, SN. Do insecticide-treated curtains reduce all-cause child mortality in Burkina Faso? *Tropical Medicine and International Health* 1997 2. (9) 855–862.
- [2] D’Alessandro, U, Olaleye, B, Langerock, P, Aikins, MK, Thomson, MC, Cham, MK, Greenwood, BM, McGuire, W, Bennett, S, Cham, MK, and Cham, BA. Mortality and morbidity from malaria in Gambian children after introduction of an impregnated bednet programme. *The Lancet* 1995 345. (8948) 479–483.
- [3] Binka, FN, Kubaje, A, Adjuik, M, Williams, LA, Lengeler, C, Maude, GH, Armah, GE, Kajihara, B, Adiamah, JH, and Smith, PG. Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Trop Med Int Health* 1996 1. (2) 147–154.
- [4] Nevill, CG, Some, ES, Mung’ala, VO, Mutemi, W, New, L, Marsh, K, Lengeler, C, and Snow, RW. Insecticide-treated bednets reduce mortality and severe morbidity from malaria among children on the Kenyan coast. *Tropical medicine & international health* 1996 1. (2) 139–146.
- [5] Hawley, WA, Phillips-Howard, PA, Ter Kuile, FO, Terlouw, DJ, Vulule, JM, Ombok, M, Nahlen, BL, Gimnig, JE, Kariuki, SK, Kolczak, MS, and High-tower, AW. Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. *American Journal of Tropical Medicine and Hygiene* 2003 68. (4.) 121–127.
- [6] Lengeler, C. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database of Systematic Reviews* 2004. (2). arXiv: 14651858.CD000363.pub2. [Art. No.: CD000363. DOI: 10.1002].

- [7] Gosoni, L, Vounatsou, P, Tami, A, Nathan, R, Grundmann, H, and Lengeler, C. Spatial effects of mosquito bednets on child mortality. *BMC Public Health* 2008 8 356.
- [8] Hightower, AW, Ombok, M, Otieno, R, Odhiambo, R, Oloo, AJ, Lal, AA, Nahlen, BL, and Hawley, WA. A geographic information system applied to a malaria field study in western Kenya. *American Journal of Tropical Medicine and Hygiene* 1998 58. (3) 266–272.
- [9] Binka, FN, Indome, F, and Smith, T. Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. *The American Journal of Tropical Medicine and Hygiene* 1998 59. (1) 80–85.
- [10] Jarvis, C, Di Tanna, GL, Lewis, D, Alexander, N, and Edmunds, WJ. Spatial analysis of cluster randomised trials: A systematic review of analysis methods. *Emerging Themes in Epidemiology* 2017 14. (12).
- [11] Gimnig, JE, Kolczak, MS, Hightower, AW, Vulule, JM, Schoute, E, Kamau, L, Phillips-Howard, PA, Ter Kuile, FO, Nahlen, BL, and Hawley, WA. Effect of permethrin-treated bed nets on the spatial distribution of malaria vectors in western Kenya. *American Journal of Tropical Medicine and Hygiene* 2003 68. (4) 115–120.
- [12] Howard, SC, Omumbo, J, Nevill, C, Some, ES, Donnelly, CA, and Snow, RW. Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2000 94. (4) 357–360.
- [13] Quinones, ML, Lines, J, Thomson, MC, Jawara, M, and Greenwood, BM. Permethrin-treated bed nets do not have a ‘mass-killing effect’ on village populations of *Anopheles gambiae s.l.* in The Gambia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1998 92. (4) 373–378.

- [14] Auchincloss, AH, Gebreab, SY, Mair, C, and Diez Roux, AV. A review of spatial methods in epidemiology, 2000-2010. *Annual review of public health* 2012 33 107–22.
- [15] Gelfand, A, Diggle, P, Fuentes, M, and Guttorp, P, eds. Handbook of Spatial Statistics. 2010.
- [16] QGIS Development Team. QGIS Geographic Information System. 2017. URL: <http://qgis.osgeo.org>.
- [17] R Core Team. R: A language and environment for statistical computing. Vienna, Austria, 2017. URL: <https://www.r-project.org/>.
- [18] R Core Team. foreign: Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ... 2017. URL: <https://cran.r-project.org/package=foreign>.
- [19] Wickham, H, Francois, R, Henry, L, and Müller, K. dplyr: A Grammar of Data Manipulation. 2017. URL: <https://cran.r-project.org/package=dplyr>.
- [20] Bivand, R, Keitt, T, and Rowlingson, B. rgdal: Bindings for the 'Geospatial' Data Abstraction Library. 2017. URL: <https://cran.r-project.org/package=rgdal>.
- [21] Lindgren, F. Bayesian Spatial Modelling with R - INLA. 2015 63. (19).
- [22] Bates, D, Mächler, M, Bolker, B, and Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 2015 67. (1) 1–48.
- [23] McElreath, R. rethinking: Statistical Rethinking book package. 2016.
- [24] Gollini, I, Lu, B, Charlton, M, Brunsdon, C, and Harris, P. GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *Journal of Statistical Software* 2015 63. (17) 1–50.
- [25] Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016. URL: <http://ggplot2.org>.

- [26] Garnier, S. viridis: Default Color Maps from 'matplotlib'. 2017. URL: <https://cran.r-project.org/package=viridis>.
- [27] Pebesma, E. sf: Simple Features for R. 2017. URL: <https://cran.r-project.org/package=sf>.
- [28] Auguie, B. gridExtra: Miscellaneous Functions for "Grid" Graphics. 2017. URL: <https://cran.r-project.org/package=gridExtra>.
- [29] Bivand, RS, Pebesma, E, and Gomez-Rubio, V. Applied Spatial Data Analysis with R. 2nd. New York: Springer, 2013.
- [30] Canty, A and Ripley, BD. boot: Bootstrap R (S-Plus) Functions. 2017.
- [31] Dowle, M and Srinivasan, A. data.table: Extension of 'data.frame'. 2017. URL: <https://cran.r-project.org/package=data.table>.
- [32] Cliff, AD and Ord, JK. Spatial processes: models & applications. London: Pion, 1981.
- [33] Moran, PAP. Notes on Continuous Stochastic Phenomena. *Biometrika* 1950 37. (1/2) 17.
- [34] Brunsdon, C, Fotheringham, aS, and Charlton, ME. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis* 1996 28. (4) 281–298.
- [35] Besag, J. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B* 1974 36. (2) 192–236.
- [36] Cressie, N. Statistics for Spatial Data. United States of America: John Wiley & Sons Ltd, 1991.
- [37] Besag, J and Kooperberg, C. On conditional and intrinsic autoregression. *Biometrika* 1995 82 733–746.
- [38] Besag, J, York, J, and Mollié, A. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 1991 43. (1) 1–20.

- [39] Rasmussen, CE and Williams, CKI. Gaussian processes for machine learning. MIT Press, 2006. arXiv: 026218253X.
- [40] Gelfand, AE and Schliep, EM. Spatial statistics and Gaussian processes: A beautiful marriage. *Spatial Statistics* 2016 18 86–104.
- [41] Rue, H and Held, L. Gaussian Markov Random Fields: Theory and Applications. Monographs on Statistics and Applied Probability. London: Chapman & Hall, 2005.
- [42] Haran, M. Gaussian random field models for spatial data. *Handbook of Markov Chain Monte Carlo*. 2011. Chap. 18449–478. arXiv: arXiv:1206.1901v1.
- [43] Lindgren, F, Rue, H, and Lindström, J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 2011 73 423–498.
- [44] Simpson, D, Lindgren, F, and Rue, H. Think continuous: Markovian Gaussian models in spatial statistics. *Spatial Statistics* 2012 1. (1) 16–29. arXiv: 1110.6796.
- [45] Blangiardo, M and Cameletti, M. Spatial and Spatio-temporal Bayesian Models with R-INLA. Wiley, 2015.
- [46] Rue, H, Martino, S, and Chopin, N. Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical Society Series B* 2009 71. (2) 319–392.
- [47] Hayes, RJ and Moulton, LH. Cluster randomised trials. 2nd. Boca Raton: Chapman & Hall/ CRC, 2017.
- [48] Smith, MJ de, Goodchild, MF, and Longley, P. Geospatial analysis: A comprehensive guide to principles, techniques and software tools. 5th. London, UK: Winchelsea Press, 2015.

- [49] Efron, B. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* 1979 7. (1) 1–26. arXiv: arXiv:1306.3979v1.
- [50] Davison, AC and Hinkley, DV. *Bootstrap Methods and Their Applications*. Cambridge: Cambridge University Press, 1997.
- [51] Field, CA and Welsh, AH. Bootstrapping clustered data. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 2007 69. (3) 369–390.
- [52] Greenland, S. Tests for interaction in epidemiologic studies: a review and a study of power. *Stat Med* 1983 2. (2) 243–251.

6. Simulating spatial effects in cluster randomised trials

6.1. Overview

The previous two chapters involved applied spatial analyses of cluster randomised trials (CRTs), and explored different types of spatial methods that could be used in a CRT context. Despite a strong suggestion of spatial spillover, the intervention estimate in chapter 5 remained robust to spatial model specifications. In this chapter, I provide the foundations for testing the robustness of intervention estimates to spatial effects. Algorithms are proposed for simulating spatial effects in CRTs, and they are examined in a simplified CRT with only two clusters. This chapter describes in detail, the algorithms used to simulate data in subsequent chapters.

Relevant objective

5. Evaluate the impact of spatial effects and the utility of spatial models in the analysis of a CRT by means of a simulation approach.

6.2. Introduction

To evaluate the impact that spatial effects can have in cluster randomised trials (CRTs), a mechanism for simulating these effects must be chosen. Spatial effects, are effects stemming from locational variation in the distribution of phenomena of interest, or in the intensity of interaction between phenomena of interest. There are various approaches to simulating spatial effects and it appears that no standard way exists. Although there is guidance for simulation studies in general [1, 2], in my experience, often the steps of the process are not transparent, and more focus is given to the modelling process under evaluation as opposed to the data generating mechanism. Therefore, it was considered important to provide detailed steps of the algorithms used for the simulating spatial effects.

This chapter explores several ways of simulating spatial effects in CRTs, and includes special considerations needed in this context. I describe an approach to create spatially indexed CRT data, alongside a general algorithm that can be used for simulating a non-spatial (or spatial) outcome for a CRT. Algorithms are provided to create spatial spillover and spatially correlated effects.

The algorithms proposed are evaluated in a simplified simulated scenario with only two clusters. This helps to assess whether the spatial spillover effects results in spatial correlation, and whether the spatial correlation methods result in spillover effects across boundaries. Using a simplified setting also helps to explore differences between the spatial effects without the added complexity of having many clusters and observations.

The aim of this chapter is to provide step by step details and give a clear explanation of the algorithms used for simulations in subsequent chapters. Furthermore, this chapter helps to gain insight and understanding that may help explain the results of more complex simulation studies.

6.3. Creating spatial CRT data

Simulating spatial effects requires data with location information. One approach to create such data, involves simulating random numbers that correspond to coordinates; this method is easy to use for simulating point data but is more difficult for creating polygons. An alternative approach, is to assign values to the locations of existing spatial data. This process is require less work to create polygon data, as it uses an already existing structure. Additionally using an admin area has the advantage that it could actually be used for a real CRT. As CRT can have two spatial levels, a combination of randomly generated or existing locations could be used.

In a CRT, if real point or randomly generated point data is created first, then a further process is needed to form the clusters. Silkey et al. simulate random point locations inside a 9x9km grid, and then split the grid into 81 equal-area clusters [3]. This method creates a CRT with square clusters, with every cluster having the same shape. Using a regular shaped grid is useful for simulations but possibly does not represent a realistic trial structure. An alternative approach used by Baylis et al. where clusters are built by using a k-nearest neighbours approach [4]. For example, select several of the simulated locations, then calculate their nearest neighbours and assign them to the same cluster. This process can be repeated until everyone is assigned to a cluster. However, this approach could result in small, or perhaps non-contiguous clusters.

Simulations based on real spatial data have been demonstrated outside of a CRT context. Dormann et al. used a spatial dataset on a volcano in New Zealand, and then attached simulated values to the already existing locations [5]. Using real spatial data can advantageous because it reflects an existing non-random structure, for example actual household locations are most likely not randomly located. However, using real household locations does restrict the upper limit of the number of observations in the simulated CRT, and the structure of the locations may lack generalisability. Using random locations may loose some realism as to the structure of

the locations, but allows for greater control on the number of observations in the CRT. A possible way to rectify both disadvantages could be to simulate locations based on a road network on top of the study area, but this approach will not be considered further, as it assume either an existing road network exists, or requires creating one, which suffers from the same issues.

In the following simulation studies a hybrid of the existing and randomly generated spatial will be used to create the locations for CRT data. The clusters are derived from existing polygon data, and observations created by simulating points over the study area assuming spatial randomness. The steps taken are outlined in Algorithm 6.1.

Algorithm 6.1 Simulating spatial data for a cluster randomised trial

INPUTS

poly existing polygon shape file

c total number of clusters in the simulated CRT

n total number of observations

- 1: Perform a spatial clip by selecting c polygons (clusters) from *poly* and removing polygons that were not selected.
 - 2: Simulate n point locations (observations) under the assumption of spatial randomness.
 - 3: Perform a spatial merge to link each point to the cluster it resides in.
 - 4: Randomise the clusters to intervention or control.
 - 5: Assign each observation to the intervention status of their cluster.
-

6.4. Simulating a CRT

Simulating CRT data requires a number of study characteristics to be predefined.

These are:

- Total number of participants n
- Total number of clusters c
- Number of participants per cluster m

Let subscripts $i =$ treatment arm ($i = 0$: control, $i = 1$: intervention), $j =$ cluster ($j = 1, \dots, c$), and $k =$ participant ($k = 1, \dots, n$) so that y_{ijk} represents the outcome for the k^{th} participant in the j^{th} cluster in the i^{th} treatment arm. With equal numbers of participants in each cluster it follows, that m (the number of participants in a given cluster) is constant. With unequal numbers of participants per cluster the value of m will vary based on the value of j giving m_j .

The simulation approach will depend on the type of trial. Leyrat et al. provide details of an approach to simulate CRT data for a continuous outcome. In their paper, they consider how to analyse CRTs with a small number of clusters and therefore choose to vary the number of clusters [6]. Baio provides examples of simulating stepped-wedge CRTs in the documentation for the *SWSamp* R package [7]. Of relevance but not directly related to CRTs, Wang and Sabo consider how to simulate clustered and dependent binary data, and Crowther and Lambert present methods for simulating biologically plausible survival data [8, 9].

Further to the study characteristics, values describing the participants and the study intervention need to be decided. These are

- Between-cluster variance σ_b^2
- Within-cluster variance σ_w^2
- The size of the intervention effect β
- An intercept value α

The between- and within-cluster variance will affect the intra-cluster correlation of the CRT. Leyrat et al. fix $\sigma_b^2 + \sigma_w^2 = 1$. The size of the intervention effect and intercept is arbitrary but clearly a large intercept or intervention effect may overwhelm simulated CRT structure. Algorithm 6.2 describes the process used for simulating CRT in this thesis, it can be used to simulate CRT data for equal and unequal numbers of participants per cluster. Although extensions for binary and count outcomes are provided, they are not utilised in this PhD and are provided for generality.

Algorithm 6.2 Simulating cluster randomised trial data

INPUTS

c	total number of clusters in the simulated CRT
n	total number of observations
σ_b^2	between-cluster variance
σ_w^2	within-cluster variance
β	intervention effect
α	intercept

- 1: Create a cluster effect u_j by simulating c random samples from a distribution with variance σ_b^2 .
- 2: Create an individual level effect ϵ_k by simulating n random samples from a distribution with variance σ_w^2 . (optional if no individual level variation required).¹
- 3: Randomise the clusters to intervention or control.
- 4: Create a dummy variable x_i to represent whether an individual is in an intervention or control cluster. Where $x_i = \begin{cases} 1 & \text{intervention} \\ 0 & \text{control} \end{cases}$.
- 5: Calculate the linear predictor $\eta_{ijk} = \alpha + \beta x_i + u_j + \epsilon_k$.

Continuous outcome

- 6: Calculate the outcome $y_{ijk} = \eta_{ijk}$

Binary outcome

- 6: Calculate $\pi_{ijk} = \frac{e^{\eta_{ijk}}}{1+e^{\eta_{ijk}}}$. This transforms the linear predictor η_{ijk} from the domain $(-\infty, \infty)$ to the domain $[0, 1]$ giving a probability.
- 7: Simulate $y_{ijk} \sim Bin(1, \pi_{ijk})$ to give binary outcome values from a binomial distribution.

Count outcome

- 6: Calculate $\lambda_{ijk} = e^{\eta_{ijk}}$. This transforms the linear predictor η_{ijk} from the domain $(-\infty, \infty)$ to the domain $[0, \infty]$ giving a count.
- 7: Simulate $y_{ijk} \sim Pois(\lambda_{ijk})$ to give a positive integer outcome values from a Poisson distribution.

1. Individual level variation is not needed for the binomial and count outcomes as variation stems from the final steps of drawing from a distribution. Alternatively the continuous outcome can be drawn from $N(\eta_{ijk}, \sigma_w^2)$, and then generating ϵ_k is not needed.

Following this process gives an outcome value for each participant, which is the sum of an intercept, an intervention effect, a cluster level effect, and includes individual level variation. This process can be used to simulate CRT data for continuous, binary, and count outcomes.

Each simulated outcome relates to a single observation that has information about

cluster membership, intervention status, and location. To create a CRT with no spatial effects the process above is sufficient (Algorithm 6.2). To consider how the spatial structure or spillover effects may affect trial results requires adjusting the linear predictor in step 5. In the next section, several ways of creating spatial and spillover effects are proposed.

6.5. Simulating CRTs with spatial effects

There are many ways of creating spatial effects or spatial relationships [10]. Perhaps the simplest is to assign an effect to the coordinates of the data, for instance, where every unit increase north gives an increase in the outcome y . Silcocks et al. use this approach to create a fixed north/south and east/west gradient covariate in their simulation studies [11]. Alternatively, spatial effects can be created by assigning effects to relative distance between participants or using a spatial network for connectivity. [3] use an SIR (Susceptible, Infected, and Recovered) model based on spatial structure to create a spatially structured effect. Staples et al. [12] use a similar approach to create a effects based on the contact network structures in a CRT.

A different approach is to define a spatial weights matrix and then multiple the Cholesky decomposition of the weights matrix with a vector to create a spatially correlated vector. This approach can be used to create a simultaneous autoregressive (SAR) spatial effect [13]. Bivand et al. how to achieve this, an a variation of their approach is used to simulate SAR correlated values in this PhD. A similar but slightly different approach construct random fields by inducing spatial autocorrelation by enforcing a specific covariance structure to the data [14–17]. Schlather et al. provide a further details of the random fields approach, with examples in their paper describing their *RandomFields* package in R [18]. The approaches differ in ease of computation, interpretability, and flexibility of the underlying mechanism they

assume. Often the underlying mechanism is not observable, and therefore more flexible complex methods may be preferred. However, simple approaches can provide useful insights, and may be easier to interpret and understand.

The systematic review from chapter 3 found two main types of spatial analyses, a spatial variable and a spatial modelling approach [19]. In keeping with the review the simulated spatial effects considered in this section are categorised as spatial spillover effects (or spatial variables), and spatially correlated effects (or spatial modelling). Spatial spillover effects involve adding variables into the model that represent spatial proximity. Spatially correlated effects force a covariance structure on the data based on spatial proximity, resulting in correlation between observations that are spatially connected.

6.5.1. Spatial spillover effects

A spatial spillover effect can be created by including an extra variable in the CRT model as follows:

$$\eta_{ijk} = \alpha + \beta x_i + u_j + \epsilon_k + \psi d_k$$

where ψ represents the population spillover effect. d_k represents proximity to the intervention. d_k could be defined as the distance to the nearest intervention household, the number of intervention households within a certain distance, or another form representing the mechanism of the spillover effect.

6.5.1.1. Proportion of spillover effects and distance weights

One way to think of the variable d_k is as the proportion of the spillover effect for individual k . Restricting $\psi = \beta \times \pi_k$ where $\pi_k \in [0, 1]$, defines the spillover effect as a proportion of the intervention effect β , based on proximity. Defining d_k as a proportion, allows ψ to represent the maximum amount of spillover effect. The

spillover effect is achieved by adding ψd_k to the model. This requires a function that takes in a measure of proximity, and outputs values between zero and one. This is possible with slight adaptations of the linear, exponential, inverse distance, and Gaussian distance weighting functions presented previously (chapter 2). These functions are sometimes also referred to as kernels, or covariance functions [20].

The variable d_k can take three forms, binary, categorical, and continuous.

Binary and categorical: When d_k is binary, it represents the observations that are impacted by the spillover. The spillover effect ψ , is either present or absent, and everyone affected, is affected by the same amount of spillover. For example, if everyone within 200m of an intervention receives 50% of the interventions benefit, and anyone past 200m does not. A binary spatial effect is simple to simulate but is probably an unrealistic underlying mechanism. Despite this, it does provide a good starting point for the simulations. However, binary spatial variables have been used in the analysis of spatial data previously, where the effect of individuals within 100 meters may be compared to those further than 100 meters away [21].

When d_k is categorical, it extends the binary variable to represent several levels of homogeneous spillover effects. This approach allows different magnitudes of spillover effects to be applied to groups of individuals, giving greater flexibility compared to the binary spatial variable. Individuals within each category receive the same effect, and convention would be to create an effect that decreases as the distance to exposure increases. Categorical spatial variables have also been used in analyses previously [22].

A categorical spillover effect is probably unrealistic. It is difficult to imagine why a spillover effect would be constant over a certain space and then change size and be constant over another space. When only two categories exist d_k is equivalent to the binary spillover effect. Assuming that the spillover effect has no discontinuities then as the number of categories increase, the categorical spillover effect will tend

to a continuous spillover effect.

Continuous and distance weights: When d_k is continuous it represents the magnitude of the spillover effect ψ for each observation. As d_k increases, the magnitude of the spillover effect either increases or decreases. This allows the spillover effect to be distance decaying, where individuals close to the source of spillover are affected more than individuals that are far away. A continuous variable could be distance to nearest intervention observation or the number of intervention observations within a specified distance. Moving forward, d_k will be represented by distance to nearest intervention observation. A continuous spillover effect is probably more realistic compared to binary or categorical spillover, and a cut-off point could be used so that any observations past a certain distance are unaffected by the spillover effect. Linear continuous spatial variables have been used in analyses previously [23]. Algorithm 6.3 describes the process used to create distance based spillover effect in this thesis. The values created in the final steps of the algorithms can then be added to the linear predictor in step 5 of algorithm 6.2, this will create a CRT with a spatial spillover effect.

Algorithm 6.3 Creating distance based spillover effects

INPUTS

- $dist_k$ the distance to nearest intervention observation for each control observation
- $distmax$ the maximum distance that spillover is present
- π the maximum proportion of the intervention effect that the spillover effect can take
- β intervention effect
- s the number of stratum (categories) ranging from $(1, \dots, S)$
- $distmax_s$ the maximum distance that spillover exist for category s
- π_s the maximum proportion of the intervention effect that the spillover effect can take for for category s where $\boldsymbol{\pi} = \{\pi_1, \pi_2, \dots, \pi_s\}$.

Binary effect

- 1: Calculate a binary variable $d_k = \begin{cases} 1 & \text{if } dist_k < distmax \\ 0 & \text{if } dist_k \geq distmax \end{cases}$.
- 2: Calculate $\psi = \beta \times \pi$.
- 3: Calculate $\psi \times d_k$.

Categorical effect

- 1: Calculate a categorical variable where

$$d_k = \begin{cases} 0 & \text{if } dist_k < distmax_1 \\ 1 & \text{if } distmax_1 \leq dist_k < distmax_2 \\ \dots & \\ S & \text{if } distmax_{S-1} \leq dist_k < distmax_S \end{cases}.$$

- 2: Calculate $\psi_s = \beta \times \pi_s$.
- 3: Calculate $\psi_s \times d_k$.

Continuous effect

- 1: Calculate $d_k = f(dist_k)$ where $f(0) = 1$, where $f(\cdot)$ is a distance weighting function with domain $[0, \infty)$ and range $[0, 1]$.
 - 2: Calculate $\psi = \beta \times \pi$.
 - 3: Calculate $\psi \times d_k$
-

To use distance as the basis for a continuous spillover effect requires encoding distance so that observations that are closer together have larger values, and observations that are further apart have smaller values. For the purpose of these simulations it is also important for the functions to give outputs in the range $[0, 1]$ as spillover is being represented as a proportion of the intervention effect. Four distance weighting functions are given in Algorithm 6.4. These algorithms show some of the potential forms that the variable d_k can take for a continuous spillover effect.

Algorithm 6.4 Distance weighting functions

INPUTS

- $dist_k$ the distance to nearest intervention observation for each control observation
- $distmax$ the maximum distance that spillover is present
- ξ parameter in the inverse distance and exponential weighted function
- ω parameter in the inverse distance and exponential weighted function
- μ the mean for the Gaussian distance weighted function
- σ^2 the variance for the Gaussian distance weighting function

Linear weighted effect

- 1: Calculate $d_{lin} = \begin{cases} distmax - dist_k & \text{if } dist_k < distmax \\ 0 & \text{if } dist_k \geq distmax \end{cases}$.
- 2: Calculate $d_k = \frac{d_{lin}}{distmax}$.

Inverse distance weighting

- 1: Calculate a $d_k = \begin{cases} \frac{\xi}{dist_k^\omega} & \text{if } dist_k < distmax \\ 0 & \text{if } dist_k \geq distmax \end{cases}$.

Exponential weighted effect

- 1: Calculate a $d_k = \begin{cases} \xi e^{-\omega d} & \text{if } dist_k < distmax \\ 0 & \text{if } dist_k \geq distmax \end{cases}$.

Gaussian weighted effect

- 1: Calculate $d_k = \begin{cases} \xi e^{-\frac{(d-\mu)^2}{\sigma^2}} & \text{if } dist_k < distmax \\ 0 & \text{if } dist_k \geq distmax \end{cases}$.
-

6.5.2. Spatially correlated effects

The SAR approach, is a Markov random field, and creates a spatial effect by manipulating the covariance structure of a vector of values [13, 24, 25]. The SAR approach can be used to create a spatially correlated intervention effect or a spatially correlated error term. Simulating a SAR spatial effect assumes that the spatial effect can be defined within the covariance structure of the data. It uses a weight matrix, and a parameter ρ which refers to the level of spatial autocorrelation.

6.5.2.1. Spatially correlated intervention

A spatially correlated intervention effect can be simulated as follows

$$\eta_{ijk} = \alpha + u_j + \epsilon_k + v_k$$

where $v_k \sim MVN(\beta x_i, \sigma^2 \Sigma)$. For SAR models $\Sigma = (I - \rho W)^{-1}(I - \rho W^t)^{-1}$. W is a spatial weights matrix based on the distance between the observations, I is an identity matrix, and ρ represents the degree of spatial autocorrelation. In this model, the matrix that is used to create a spatially correlated value is multiplied by the intervention effect. This creates a weighted intervention effect with no upper bound on the size of the intervention. The process for creating an SAR correlated value is described in algorithm 6.5.

Algorithm 6.5 Simulating a simultaneous autoregressive spatially correlated value

INPUTS

ρ the level of spatial correlation

$maxdist$ the maximum distance below which observations are neighbours

\mathbf{a} a vector of values

- 1: Calculate $dist$ the distance between observations that are neighbours.
 - 2: Apply a distance weighted function to $dist$ so that nearby observations have a greater weight and further away observation has less weight on each other.
 - 3: Create a spatial weights matrix W based on the weighted distances of the neighbours. Where a value of zero is used for any observations that are not neighbours, and each observation cannot be its own neighbour.
 - 4: Calculate $A = I - \rho W$.
 - 5: Calculate $B = (A^t A)$ to give a symmetric matrix.
 - 6: Factorise $B = LL^\dagger$ using Cholesky decomposition where L is a lower triangular matrix and L^\dagger is the conjugate transpose of L .
 - 7: Calculate $\mathbf{a}L^\dagger$ to create a spatially correlated vector of values.
-

The SAR simulation can be used to create a spatially correlated error term, this would represent the presence of a spatial effect, potentially unrelated to the intervention in a CRT.

Spatially correlated error The following model represents a spatially correlated error term

$$\eta_{ijk} = \alpha + \beta x_i + u_j + \epsilon_k + v_k$$

where $v_k \sim MVN(0, \sigma^2 \Sigma)$. With Σ , defined as previously. The error consists of a cluster effect u_j , a spatially correlated value v_k , and residual error ϵ_k ($u_j + \epsilon_k + v_k$).

6.5.2.2. Cluster level spatial correlation

The spatial effects presented so far in the chapter have all been at the observation level. Changing the indexes used in the model formula allows the spatial covariance

error approach to be simulated based on the spatial structure of the clusters using the formula

$$\eta_{ijk} = \alpha + \beta x_i + v_j + \epsilon_k$$

where $v_j \sim MVN(0, \sigma_v^2 \Sigma)$. With Σ defined as previously. The error consists of a spatially correlated cluster effect v_k , and residual individual level error ϵ_k ($v_j + \epsilon_k$).

6.5.3. Considerations for simulating spatial effects in CRTs

Here I outline and propose several aspects of simulating spatial effects that need to be considered before conducting a simulation study. Some are perhaps unique to the CRT setting, and others are related to more general problems with simulating spatial effects.

In a CRT where individuals are points and clusters are polygons, if the intervention works, this induces a spatial pattern in the data at a cluster level. For example, if each intervention cluster has a lower risk due to the intervention, then intervention clusters are systematically different from their neighbouring control clusters. Therefore, when applying spatial methods to assess spatial correlation, it is important to adjust for intervention status.

For a continuous outcome, a variogram approach could be used, and adjustment for intervention is straight forward [26, 27]. When using *Moran's I*, adjusting for the intervention status can be achieved by first regressing the intervention on the outcome, and assessing the spatial correlation of the residuals [28, 29]. Alternatively, Moran's I can be calculated for the entire study area, and then the control and intervention study areas separately.

Another aspect to consider with spatial effects within CRTs, is the level at which the spatial effects are present. In areal data the spatial structure is assumed at cluster level, often because finer level data is not available. For point data the spatial effects

are assumed at the individual level. When data is present at two different levels this is called spatial misalignment [30].

In CRTs, spatial effects can be considered at the cluster and/or individual level. If the effect is at a cluster level then all individuals within a cluster are affected in the same way. If it is at the observation level, then the range of the effect could be smaller or larger than the cluster size. For example, you may have connectivity within clusters, and thus a small number participants are connected between clusters. Alternatively, if the range is large, then there is potential for all observations within the trial to be connected. If all observations in trial are connected, then this would range across the entire study area, and thus also across cluster boundaries.

Caution is needed not to overwhelm the values of the CRT simulation. For instance, if the magnitude of the simulated spatial effect is much larger than the simulated outcome values without the spatial effect. In this scenario, the simulation is no longer considering spatial effects within the context of a CRT, but rather considering how a multi-level model works on a spatially correlated variable. When the simulated spatial effect is very small compared to the intervention effect, the process only adds noise to the data generating model. Therefore, a balance between the size of the intervention effect and the magnitude of the spatial effect needs to be achieved. This avoids inferring results that are not actually related to a CRT.

Consideration of what type of connectivity is meaningful is also needed. When connecting points, a distance or a nearest neighbours based approach could be used. The most appropriate method will be based on the assumed form of the spatial effect. Further methods for modelling relations between points are also found in probabilistic graphical modelling [20].

The neighbours based approach may be useful for socially meaningful connections, whereas distance based might be more useful when trying to model the movement of a mosquito. For analysis of real data, households are generally clustered into neighbourhoods or groups of settlements. Therefore, heterogeneity of the types of

settlements in a CRT, will change the impact of using neighbours versus distances. For instance, in a densely populated area, a distance based approach will give a larger number of connected observations compared to a neighbour based approach. Whereas in a sparsely populated area, the distance based approach may result in no neighbours. The neighbour based approach will guarantee neighbours but may result in connectivity between observations that are (physically) far apart.

The simulations in this thesis will use a distance based approach, as the focus of the thesis is spatial effects due to physical proximity. Furthermore, in these simulations the difference between a neighbour approach and a distance based approach is likely minimal, as the locations were generated under a random process.

6.6. Simulating a CRT with two clusters and spatial effects

6.6.1. Introduction

The methods for simulating spatial effects are explored by applying them to a simulated ‘trial’ with two clusters. A simplified setting was used to compare the different simulation methods, and to improve understanding of the results achieved in subsequent simulations. A simplified setting also helps to clarify which methods to use in the larger simulation study, reducing unnecessary computing, and simulation time.

The cluster level models are excluded as they are not appropriate for data with only two clusters. The SAR error model is not simulated as the mechanism is similar to the SAR intervention model. Furthermore, it was decided to remove random error from the data generating mechanism so that noise would not preclude the impact of the spatial effects. The SAR error model will be assessed in Chapter 7.

6.6.2. Methods

6.6.2.1. Data generation

Two thousand points were simulated within a 1000 by 1000 unit square, assuming spatial randomness. The middle third of the area was then assigned to an intervention, creating a CRT with one intervention cluster surrounded by a control cluster (Figure 6.1). For each simulation, a continuous outcome was generated with an intervention effect β , fixed at one and an intercept α of 0.2. The model used was $y = \alpha + \beta x_i$ where y is a continuous outcome and $x_i = \begin{cases} 1 & \text{if } Intervention \\ 0 & \text{if } Control \end{cases}$. As there was no error term in the model, any changes in the outcome values are only due to the intervention and spatial effects.

The outcome is continuous and therefore the intervention effect estimates are an absolute change in value. For instance, an effect of 2 reflects that on average, intervention participants are expected to have an outcome of 2 units higher compared to control participants. This is highlighted because odds ratios and risk ratios are commonly used in medical research. A continuous outcome was chosen as it allows greater control on simulating the spatial effects, with binary and count outcomes left for work beyond the thesis.

The spillover and spatial effects were restricted to roughly a 200 unit radius. Therefore, there is a buffer zone of 200 units around the intervention cluster where we would expect the control observations to be affected (Figure 6.1), beyond that, the simulation should create minimal, if any, differences on the values of the control observations.

6.6.2.2. Spatial spillover

Adding spatial variables to the linear predictor can be used to create a spatial effect. The previous model is extended to generate an outcome with a spatial spillover effect

as follows:

$$y = \alpha + \beta x_i + \psi d_k$$

Where ψ is a spillover effect, and d_k is a spatial variable. In this simulation the maximum spillover effect was defined as 0.8, which is 80% of the intervention effect. Binary and categorical variables were created with a maximum distance of 200 units. Four categories of equal distances were used for the categorical variable. For the continuous spatial variables, linear, Gaussian, exponential, and inverse distance weighting (IDW) methods were used. To ensure that the different weights are comparable, they were simulated to give a value of 0.01 at 200 units. This means any observation at a distance of 200 units will receive 1% of the spillover effect. This is achieved by using non-standard values for the parameters of the weighted functions. The maximum distance of the linear effect was 202.034 units, the power for the IDW was 0.868, sigma equals 65.91 for the Gaussian weights, and the exponential models have a power of 0.023. A visual comparison between the different weights is given in the bottom left panel of Figure 6.2.

6.6.2.3. Spatial correlation

A SAR approach was used to generate a spatially correlated intervention effect. Neighbours were defined as observations within 200 units of one another and the spatial weights were inverse distance. This allows observations in different cluster to be neighbours. For the SAR simulation the values used for spatial correlation were $\rho = 0.1, 0.2, 0.6, 0.7, 0.9,$ and 0.99 which gives a range of weak to very strong spatial correlation. The model

$$y = \alpha + v_k$$

was used to create a spatially correlated effect where $v_k \sim MVN(\beta x_i, \sigma^2 \Sigma)$, and $\Sigma = (I - \rho W)^{-1}(I - \rho W^t)^{-1}$, W is a spatial weight matrix based on the distance between the observations, I is an identity matrix, and ρ represents the degree of spatial autocorrelation.

6.6.2.4. Assessing spillover and spatial correlation

Presence and impact of spillover was assessed visually by mapping the study area with the size of the points weighted by the outcome values. When spillover is absent, the point in the intervention cluster will be larger compared to the points in the control cluster, and the size will be homogeneous within each cluster. When spillover is present then control points should be larger, the closer they are to the intervention.

Two approaches were used to consider spatial correlation. The first, plots a variogram adjusting for the intervention effect. This approach allows estimation of the magnitude and range of spatial correlation. It is important to adjust for the intervention effect, otherwise the spatial correlation detected could be due to a spatially assigned intervention.

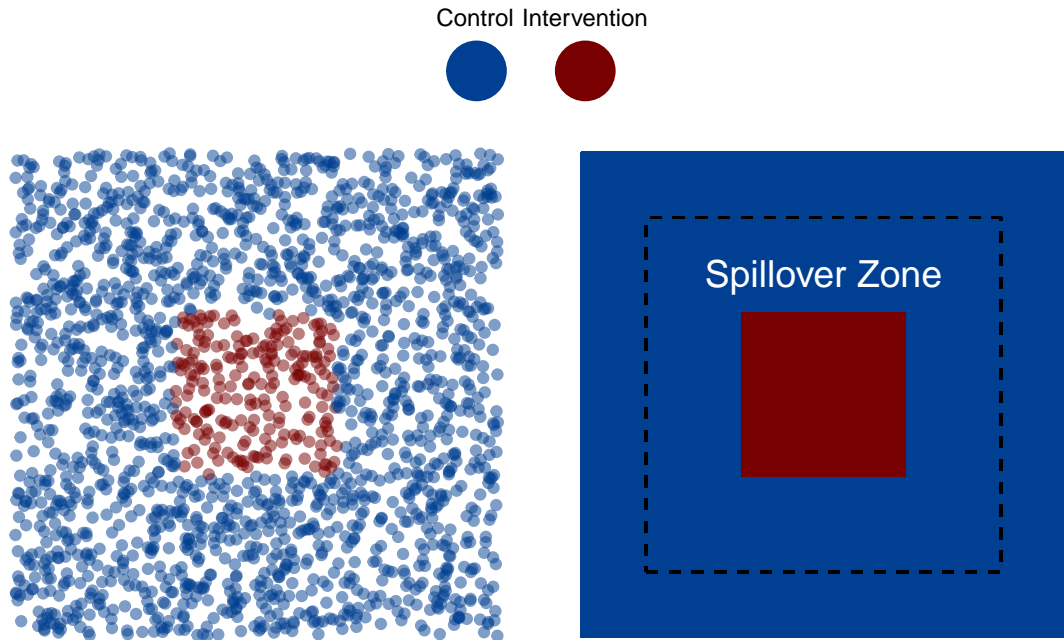
The second approach, formally tests for spatial correlation using Moran's I. Rather than test on the simulated data, the test was performed on the residuals of a regression model with intervention as a covariate. This method adjusts for the intervention effect. Moran's I was calculated using Monte Carlo simulation with 999 simulations used.

6.6.3. Results

The study area contained 2,000 points in a 1,000 by 1,000 unit square. There were 219 observations in the intervention cluster, and 1,781 points in the surrounding control cluster. Of the control observations, 724 (40.6%) were within 200 units

of the intervention cluster, and are therefore in the ‘spillover zone.’ The size and density of the study area can be observed in Figure 6.1.

Figure 6.1. Study area showing spillover area



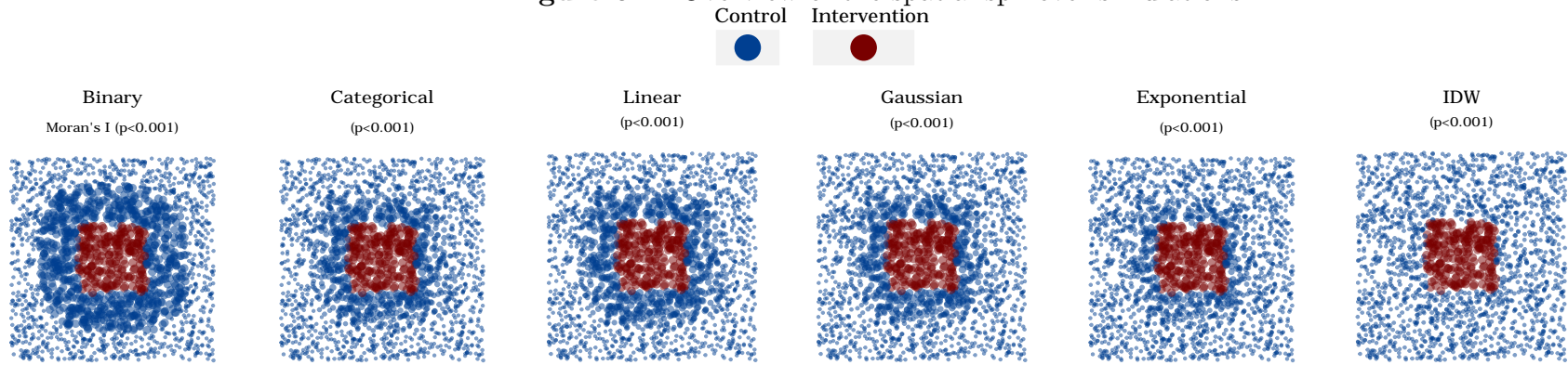
6.6.3.1. Spatial spillover

All of the spatial variable simulations resulted in strong evidence of spatial correlation, with $p < 0.001$ for Moran’s I (Figure 6.2). Unsurprisingly, the binary effect resulted in the largest amount of spillover. The variograms suggest that spatial correlation is present over a range of 300 units, and that the semivariance was largest for the binary spillover variable. The magnitude of spillover was similar for the linear, categorical, and Gaussian variables, with comparable variograms. The exponential and IDW variables had considerably smaller semivariance, and the spillover was much harder to ascertain from the plots of the study areas. The maximum value of the intervention observations was 1.2 (for all scenarios). The spatial binary approach results in the greatest magnitude of spillover and the maximum value in the control observations was one for this scenario.

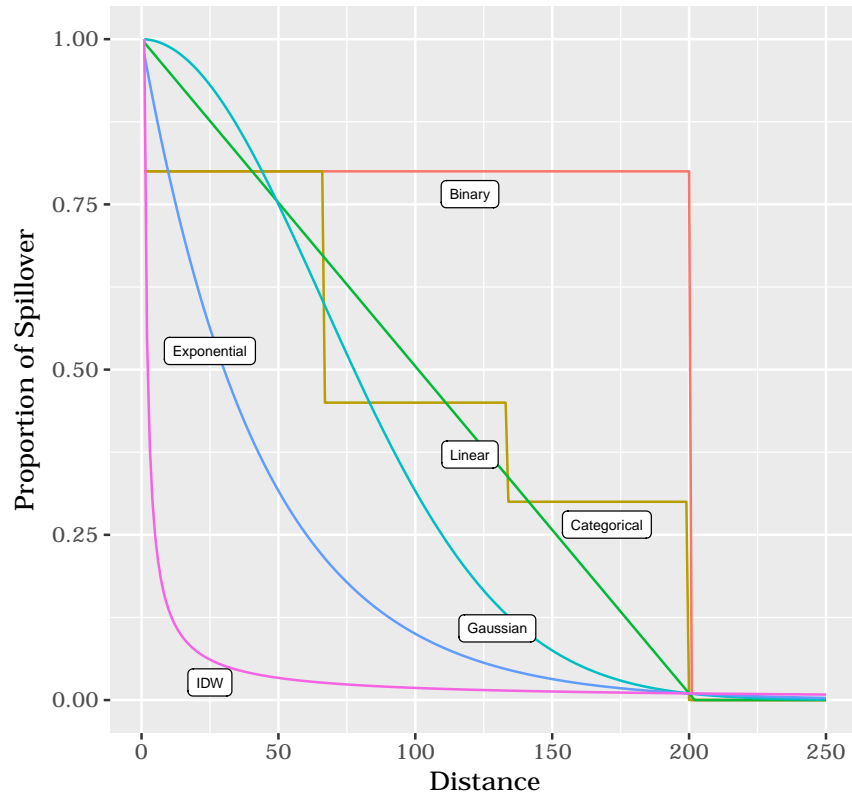
6.6.3.2. Spatial correlation

There was strong evidence of spatial correlation, with $p < 0.001$ for the SAR simulations. The amount of spillover increased as the magnitude of spatial correlation increased. The variograms show a suggestion of spatial correlation over a range of between 300 and 400 units, however in comparison to the spatial variable approach, the magnitude of the semivariance was small, being similar to the exponential and IDW variables. The variogram of the Gaussian spatial variable is included in the Figure 6.3 for reference. As can be seen in the lower left panel of Figure 6.3, the maximum value in the control and intervention arms was positively associated with the value of ρ . The maximum value of all the observations ranged from 1.29 for $\rho = 0.1$ to 2.87 for $\rho = 0.99$.

Figure 6.2. Overview of the spatial spillover simulations



Size of spillover for different distance weights



Variograms of spatial variables

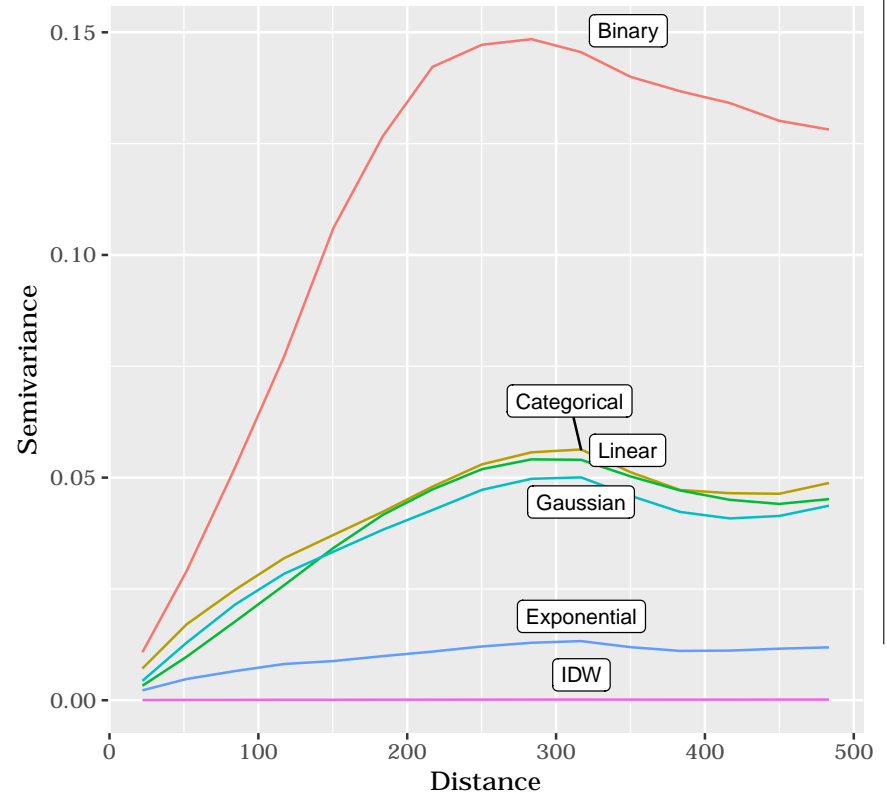
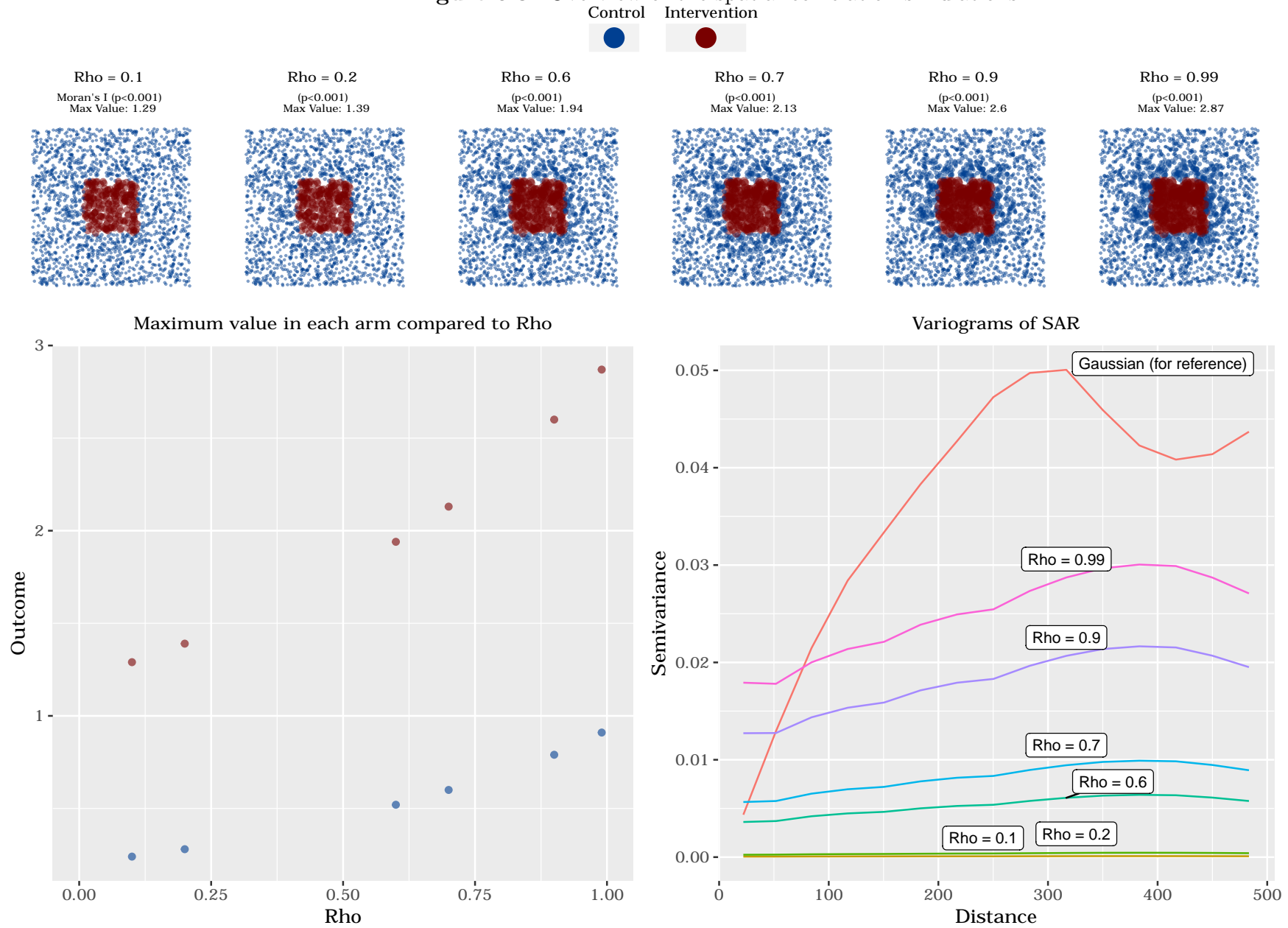


Figure 6.3. Overview of the spatial correlation simulations



6.6.4. Conclusions

The spatial spillover and spatial correlation approaches resulted in spatially correlated spillover effects. There was strong evidence of spatial correlation when tested using Moran's I, adjusted for intervention status. Variograms of the outcome, adjusted for intervention status, also suggested varying levels of spatial correlation present for a range of 300 to 400 units. Spillover effects on the control observations were present in all scenarios, with control observations near to the intervention, having higher values than those further away. However, the magnitude of spillover varied, and was in some cases (IDW spatial variable) negligible. These results are reassuring and indicate that the simulated approaches have work as desired and create spatial correlated effects.

For the spatial spillover effect, the choice of distance weighting function can have a large impact on the magnitude and range of the spillover effect, but does not appear to impact the ability for Moran's I to detect spatial correlation. Despite fixing a roughly equal level of spillover effect at 200 units, the Gaussian, linear, binary, and categorical functions resulted in a larger amount of spillover compared to the IDW and exponential weights. This makes sense when considering the difference in the curves of the different distance weighting functions.

In the spatial spillover variables, the maximum value of the intervention observations are by design unaffected by the strength of the spillover effect. Therefore, such simulations cannot result in a spillover effect that is greater in magnitude than the intervention. In contrast, the maximum value of the intervention is affected by the magnitude of spatial correlation in the SAR simulation. The values in the control and intervention arms were positively correlated with the magnitude of spatial correlation. This make sense as the spatially correlated effect represent a cumulative effect based on proximity to nearby intervention participants. In contrast to the spillover effects, proximity to multiple individuals will result in greater spillover compared to proximity to a single individual, and the total amount of spillover could

be greater than the size of the direct intervention effect. This may make it difficult to control the size of the intervention effect, and could result in the simulated values becoming very large. Caution is therefore needed to avoid creating very large simulated values, as the evaluation is no longer of a CRT with spatial correlation, but rather just a spatially correlated value.

6.7. Summary

This chapter has described an approach to simulate spatially structured CRT data. A prominent focus on the computational side of simulations was given, presenting algorithms for creating a spatial spillover variable based on distance to nearest intervention observations, and spatial correlation using an SAR approach.

Several considerations when simulating spatial effects in CRT are given, such as: spatially assigned interventions naturally inducing spatial correlation, the level that spatial correlation is present (observation or cluster), the magnitude of the spatial effect compared to the intervention effect, and what measure of connectivity is useful.

A simulation study with only two clusters was used to explore the types of results these methods create. The spatial variable approach allows for great control of the magnitude of effect, but forces a known mechanism for the spatial effect. The various distance weights impact to the magnitude of the spillover effect, but they all create spillover. The choice of values for the parameters for a distance weight function is more important than the distance weight itself.

The Gaussian weighted function has a potentially more realistic representation of distance decay compared to linear, binary, or categorical distance weights, and has an easier interpretation than the exponential or IDW. The range of the Gaussian distance weight can be defined in terms of standard deviations, a distance of three standard deviations results in a distance weight of one percent. Due to this, only the Gaussian weighted function will be used for spatial spillover in further simulations.

The covariance based SAR effect results in spillover and spatial correlation, but controlling the values that the outcome will take is difficult. Increasing the level of spatial correlation, increases the size of the outcome values. Therefore, if the spatial correlation is large, the outcome values of the trial may be overwhelmed by the correlation, and the simulation will no longer represent a CRT setting.

The Gaussian weighted spatial variable and the SAR intervention effect will be used in the next chapter.

6.8. Bibliography

- [1] Burton, A, Altman, D, Royston, P, and Holder, R. The design of simulation studies in medical statistics. *Statistics in medicine* 2006 25. (24) 4279–4292.
- [2] Morris, TP, White, IR, and Crowther, MJ. Using simulation studies to evaluate statistical methods. *ArXiv e-prints* 2017. arXiv: 1712.03198.
- [3] Silkey, M, Homan, T, Maire, N, Hiscox, A, Mukabana, R, Takken, W, and Smith, TA. Design of trials for interrupting the transmission of endemic pathogens. *Trials* 2016 17 278.
- [4] Baylis, K and Ham, A. How important is spatial correlation in randomized controlled trials? 2015.
- [5] F. Dormann, C, M. McPherson, J, B. Araújo, M, Bivand, R, Bolliger, J, Carl, G, G. Davies, R, Hirzel, A, Jetz, W, Daniel Kissling, W, Kühn, I, Ohlemüller, R, R. Peres-Neto, P, Reineking, B, Schröder, B, M. Schurr, F, and Wilson, R. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography* 2007 30. (5) 609–628.
- [6] Leyrat, C, Morgan, KE, Leurent, B, and Kahan, BC. Cluster randomized trials with a small number of clusters: which analyses should be used? *International Journal of Epidemiology* 2017. (August 2017) 321–331.
- [7] Baio, G. SWSamp : Simulation-based sample size calculations for a Stepped Wedge Trial (and more). *SWSamp (R package)* 2016.
- [8] Wang, A and Sabo, RT. Simulating Clustered and Dependent Binary Variables. *Austin Biom and Biostat* 2015 2. (2) 1–5.
- [9] Crowther, MJ and Lambert, PC. Simulating biologically plausible complex survival data. *Statistics in Medicine* 2013 32. (23) 4118–4134.
- [10] Lichstein, JW, Simons, TR, Shriener, Sa, and Franzreb, KE. Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs* 2002 72. (3) 445–463.

- [11] Silcocks, P and Kendrick, D. Spatial effects should be allowed for in primary care and other community-based cluster RCTS. *Trials* Jan. 2010 11. (1) 55.
- [12] Staples, PC, Ogburn, EL, and Onnela, JP. Incorporating Contact Network Structure in Cluster Randomized Trials. *Scientific Reports* 2015 5. (17581) 1–12. arXiv: 1505.0044.
- [13] Bivand, RS, Pebesma, E, and Gomez-Rubio, V. Applied Spatial Data Analysis with R. 2nd. New York: Springer, 2013.
- [14] Adler, RJ and Taylor, JE. Random Fields and Geometry. Springer Monographs in Mathematics. Springer New York, 2009.
- [15] Brouste, A, Istas, J, and Lambert-lacroix, S. On Gaussian random fields simulations. *Journal of Statistical Software* VV. (Ii) 26.
- [16] Chan, BG. An Algorithm for Simulating Stationary Gaussian Random Fields. *Statistical Algorithms* 1996. (1983) 171–181.
- [17] Chan, G. An effective method for simulating Gaussian random fields. *Proceedings of the statistical Computing section* 1999 1–6.
- [18] Schlater, M, Malinowski, A, and Menck, PJ. Analysis, Simulation and Prediction of Multivariate Random Fields with Package RandomFields. *Journal of Statistical Software* 2015 63. (8) 1–25. arXiv: arXiv:1501.0228.
- [19] Jarvis, C, Di Tanna, GL, Lewis, D, Alexander, N, and Edmunds, WJ. Spatial analysis of cluster randomised trials: A systematic review of analysis methods. *Emerging Themes in Epidemiology* 2017 14. (12).
- [20] Bishop, CM. Pattern Recognition and Machine Learning. Information Science and Statistics. Springer, 2006.
- [21] Kroeger, A, Lenhart, A, Ochoa, M, Villegas, E, Levy, M, Alexander, N, and McCall, PJ. Effective control of dengue vectors with curtains and water container covers treated with insecticide in Mexico and Venezuela: cluster randomised trials. *BMJ (Clinical research ed.)* May 2006 332. (7552) 1247–52.

- [22] Binka, FN, Indome, F, and Smith, T. Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. *The American Journal of Tropical Medicine and Hygiene* 1998 59. (1) 80–85.
- [23] Ali, M, Thiem, VD, Park, JK, Ochiai, RL, Canh, DG, Danovaro-Holliday, MC, Kaljee, LM, Clemens, JD, and Acosta, CJ. Geographic analysis of vaccine uptake in a cluster-randomized controlled trial in Hue, Vietnam. *Health and Place* 2007 13. (3) 577–587.
- [24] Whittle, P. On Stationary Processes in the Plane. *Biometrika* 1954 41. (3) 434–449.
- [25] Haran, M. Gaussian random field models for spatial data. *Handbook of Markov Chain Monte Carlo*. 2011. Chap. 18449–478. arXiv: [arXiv:1206.1901v1](https://arxiv.org/abs/1206.1901v1).
- [26] Cressie, N. *Statistics for Spatial Data*. United States of America: John Wiley & Sons Ltd, 1991.
- [27] Diggle, PJ, Tawn, JA, and Moyeed, RA. Model-based geostatistics. *Journal of the Royal Statistical Society: Series C* 1998 47. (3) 299–350. arXiv: [MR1626544](https://arxiv.org/abs/MR1626544).
- [28] Moran, PAP. Notes on Continuous Stochastic Phenomena. *Biometrika* 1950 37. (1/2) 17.
- [29] Rogerson, PA. *Statistical Methods for Geography: A Student's Guide*. SAGE Publications, 2006.
- [30] Banerjee, S, Carlin, BP, and Gelfrand, AE. *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton: Chapman & Hall/ CRC, 2015.

7. Simulation study of the impact of spatial effects in cluster randomised trials

7.1. Overview

In chapter 6, algorithms were proposed and tested for simulating spatial effects in cluster randomised trials (CRTs). In chapter 5, we saw that intervention estimates remained robust to spatial model specifications, despite a strong suggestion of spatial spillover. This chapter utilises the algorithms from the previous chapter to test this robustness observed in chapter 5. A simulation study is conducted to test the robustness of the standard CRT model to spatial effects, and consider whether spatial models can provide improvement when spatial effects are present.

Objective

5. Evaluate the impact of spatial effects and the utility of spatial models in the analysis of a CRT by means of a simulation approach.

7.2. Introduction

Spillover effects are the effect of an intervention on individuals who are in physical or social proximity to intervention recipients, but who do not receive the intervention themselves [1, 2]. When present, they may violate the Stable Unit Treatment Value Assumption (SUTVA) which is required for valid inferences in randomised experiments. When the borders of a geographical CRT are close together, there is a risk of between-cluster spatial spillover [3, 4]. Furthermore, when CRTs have geographical clusters they have an inherent spatial structure with some clusters being near to intervention clusters and others nearer to control clusters. The spatial structure is rarely accounted for in the main analysis of a CRT and there is little knowledge about the impact on trial results [5].

Chapter 5 presented the spatial reanalysis of an insecticide treated bed net CRT [3, 6]. The intervention estimate of the trial remained robust to spatial model specification, despite compelling evidence of a positive spatial spillover effect (from intervention clusters towards control clusters). This led to speculation as to whether spatial effects impact the results of CRTs, prompting consideration of how the magnitude of spatial effects impact intervention estimates and motivates this simulation study

This simulation study explores how different types of simulated spatial effects impact a standard one-stage CRT analysis, and whether spatial regression methods can be used to account for such effects. The study considers spatial spillover, a spatially correlated intervention, and spatially correlated errors on the bias and coverage of intervention estimates. The standard one-stage (non-spatial) CRT model is compared to a cluster and an individual level spatial model. This approach will help identify when the typical analysis of a CRT may fail, and consider if spatial modelling approaches offer improvement in such cases.

7.3. Methods

This simulation study involves three scenarios for different types of spatial effects, and compares three models. The three different scenarios of spatial effects simulated were: (1) spatial spillover effect; (2) spatially correlated intervention effect; (3) spatially correlated error term. The three models applied are: (A) standard one-stage CRT model; (B) Besag, Yorke, and Mollie (BYM) cluster level spatial model; (C) Gaussian process (GPm) individual level model. The standard CRT model is applied to all three scenarios, and the impact of the spatial effects on the bias and coverage is considered. Following this, the spatial models are applied to scenarios where the standard CRT model demonstrated bias or poor coverage.

7.3.1. Data generating mechanism

The data generating mechanism describes the process used to create simulated datasets using random number generators.

Creating spatial CRT data

A polygon shapefile of Nepal was used as the basis for the study area (Admin level 5). Spatial data for each country can be split into different administrative levels (e.g National, Regions, County), and in Nepal administrative level 5 is equivalent to wards. This file was spatially clipped to create a CRT with 30 contiguous polygons, meaning 30 polygons were selected and the remaining were deleted. Assuming spatial randomness, 5,000 points were simulated within the study area. The points were joined to the polygon areas using a spatial merge, so that each point was assigned to the cluster they resided within. The polygons were randomised at a ratio of 1:1 to intervention or control, and the points were assigned the same intervention status as their cluster. This gives the study area shown in Figure 8.4.

The number of observations and clusters were chosen for three reasons: (I) at least

15 cluster per arm are needed for the one-stage CRT analysis method to be valid [4]; (II) to have a dense study area that could be conducive to spillover effects; (III) it was computationally feasible.

Simulating a CRT

Let subscripts $i =$ treatment arm ($i = 0$: control, $i = 1$: intervention), $j =$ cluster ($j = 1, \dots, c$), and $k =$ participant ($k = 1, \dots, n$) so that y_{ijk} represents the outcome for the k^{th} participant, in the j^{th} cluster, in the i^{th} treatment arm. Define $\alpha, \beta \in \mathbb{R}$ as the intercept and the intervention effect, respectively. Then a CRT, can be generated from the model

$$y_{ijk} = \alpha + \beta x_i + u_j + \epsilon_k$$

where $x_i = \begin{cases} 1 & \text{if } i = \textit{intervention} \\ 0 & \text{if } i = \textit{control} \end{cases}$, u_j represents a cluster effect, and $\epsilon_k \sim N(0, \sigma^2)$. y_{ijk} is a continuous outcome for the k^{th} participant, in the j^{th} cluster, in the i^{th} arm, u_j is constant within each cluster, and ϵ_k provides variation within clusters. This is the base model used to simulate CRT data with spatial effects. Further details on the algorithm used for simulating a CRT are given in algorithm 6.2. Two approaches were taken to create spatial effects, a spatial spillover variable, and a spatial correlation effect.

Spatial spillover

Spatial spillover was simulated using the model

$$y_{ijk} = \alpha + \beta x_i + u_j + \epsilon_k + \psi d_k$$

Where ψ represents the spillover effect, and d_k is a spatial variable representing proximity to the intervention for the k^{th} participant. Defining $\psi = \pi \times \psi$ where π is a proportion (lying between zero and one), restricts the spillover effect ψ to be a

proportion of the intervention effect. In addition, restricting d_k to be a proportion based on spatial proximity defines the spatial variable d_k , to represent the amount of the spillover effect that individual k receives. Moreover, this means the spillover effect is defined as a proportion of the intervention effect, and individuals receive a proportion of the spillover effect, depending on their proximity to the intervention. d_k is defined as distance to nearest intervention observation, and therefore only defined for control observations. The distance was weighted using a Gaussian distance weighting function

$$f(d_k) = e^{-\frac{(d_k)^2}{\sigma^2}}$$

where σ is the standard deviation of the weighting function and relates to the effective range over which the distance effect is present.

The range was defined in terms of the standard deviation σ and was assigned values of 33.3 meters, 66.6m, and 166.6m as these correspond to the proportion of the spillover effect being roughly 1% of the intervention effect at a range of 100m, 200m, and 500m respectively. Further details can be seen in Algorithms 6.3 and 6.4.

Spatial correlation

A spatially correlated error term was simulated using the model

$$y_{ijk} = \alpha + \beta x_i + u_j + v_k + \epsilon_k$$

In this model, the error term is composed of $u_j + v_k + \epsilon_k$. The error consists of a cluster effect u_j , a spatially correlated effect v_k , and an error term ϵ_k . v_k is a spatial structured random effect where $v_k \sim MVN(0, \sigma_v^2 \Sigma)$ and σ_v^2 is the between-area variance. A simultaneous autoregressive (SAR) spatial effect was generated where $\Sigma = (I - \rho W)^{-1}(I - \rho W^t)^{-1}$, W is a spatial weights matrix based on the distance between the observations, I is an identity matrix, and ρ represents the degree of spatial autocorrelation.

This model definition generates data with an SAR error effect. A SAR intervention effect was also simulated with a slightly modified model defined as

$$y_{ijk} = \alpha + u_j + v_k + \epsilon_k$$

Where $v_k \sim MVN(\beta x_i, \sigma_v^2 \Sigma)$, with Σ defined as before. Further details of the algorithm used for simulating SAR correlated values are given in algorithm 6.5.

7.3.2. Value held constant during simulations

The following study values remained unchanged throughout the iterations of the simulations:

- Intervention effect ($\beta = 2.0$)
- Intra-cluster correlation coefficient (ICC) ($\sigma_b = 0.05$, $\sigma_w = 0.95$, giving and ICC = 0.05)
- Number of clusters ($c = 30$)
- Fixed sample size ($n = 5,000$)
- Location of observations (simulated under spatial randomness)

An intervention effect of 2.0 was chosen as based on chapter 6 and other preliminary work, it was large enough to not be overwhelmed by a spatial error term simulated from multivariate normal distribution, but small enough so as to be affected by spatial effect. The ICC of 0.05 was used with $\sigma_b = 0.05$ and $\sigma_w = 0.95$. This value was chosen based on a study of ICCs, which estimated that 90% of ICC may be less than 0.55, although this research was based on primary care trials [7]. Furthermore, an ICC of 0.05 corresponds to the median ICC value found in a study by Campbell et al. [8]. The locations were fixed to avoid conflating the impact of spatial effects and the spatial structure of a trial.

7.3.3. Values that are changed during simulations

The values that are varied in this study are; the strength of the spatial effect and the distance over which it is present. For spatial spillover, the strength of the spatial effect, was the proportion of the intervention effect that control participants receive, and the distance was the standard deviation of the Gaussian weighting function. For the spatial correlation effect, the strength of the spatial effect was the correlation parameter, and the distance represents the maximum distance for observations to be classified as neighbours. The full range of values varied in the simulation study for each type of spatial effect, are given in Table 7.1. The value for the spatially correlated intervention and spatially correlated error term are the same and hence one set of values is given/

Table 7.1. Study values that vary in the simulations

Spatial effect	Variable	Values
Gaussian weight spillover effect	Range* (meters)	100, 200, and 500m
	Proportion of intervention effect	0, 0.2, 0.4, 0.6, 0.8, 0.99
SAR spatial correlation	Distance of neighbours (meters)	100, 200, 500
	ρ (strength of the correlation)	0, 0.2, 0.4, 0.6, 0.8, 0.99

* $\sigma = 33.3, 66.6, \text{ and } 166.67\text{m}$, when the range is 100, 200, and 500m respectively.
Range is where the distance weight is 0.01

7.3.4. Methods

Three methods are applied, a standard CRT model and two spatial models.

Standard CRT model

A linear mixed effects model with a random effect for cluster will be applied. This model will be fitted using the lme4 package in R version 3.4.3 [9, 10] and is defined as

$$y_{ijk} = \alpha + \beta x_i + u_j$$

Where u_j is a random effect for cluster, α is an intercept, β is the effect of the intervention, y_{ijk} is a continuous outcome, and x_i is a binary variable representing the intervention status for the observations.

Spatial CRT models

Two types of spatial models were applied; a cluster level spatial model originally presented by Besag, York, and Mollie [11] and referred to as a BYM model, and a individual level spatial model, called a Gaussian process model (GPM). Further details of the spatial models and their relations to random fields are presented in Appendix C. The BYM model is defined as

$$y_{ijk} = \alpha + \beta x_i + u_j + v_j$$

Where the v_j is a spatially structured cluster level random effect, and u_j is a random effect for cluster, and $\alpha, \beta, x_i, y_{ijk}$ defined as before. This model has been widely used in disease mapping [12, 13].

The GPM model is defined as

$$y_{ijk} = \alpha + \beta x_i + u_j + v_k$$

$$v_k \sim N(f(z), \sigma_v^2)$$

$$f \sim GP(\mu(\cdot), \Sigma(d))$$

where u_j is a random effect for cluster, v_k is a spatially structured individual level random effect, $\mu(\cdot)$ is a mean function restricted to be zero, $\Sigma(\cdot)$ is a variance-covariance function incorporating distance measure d and $\alpha, \beta, x_i, y_{ijk}$ defined as

before. As the GP has a mean of zero, and covariance defined by the distance between the observations, the values of v_k are a function of the distance calculations. GPm models are typically computationally intensive to fit, due to dense covariance matrices. integrated nested Laplace approximation (INLA) is used to fit the spatial models [14]. For the GPm, the INLA stochastic partial differential equation approach (SPDE) is used [15–17]. Uninformative priors are used for the spatial models.

7.3.5. Number of simulations

For the standard CRT model each scenario was run for 2000 simulations. For the spatial models the number of runs was a trade off between computation length and demonstrating bias. When running the standard CRT model for 2000 simulations bias was clearly demonstrated when the number of simulations was less than 200. Therefore, 200 simulations were run for each scenario for the spatial models.

7.3.6. Estimand

For each simulation, the estimate of the intervention effect and the standard error (standard model) or standard deviation (spatial models) of the estimate will be recorded. A normal approximation 95% confidence/credible interval will be calculated using these estimands. For the standard model a 95% confidence interval (CI) will be calculated. The spatial models require a Bayesian approach, and therefore 95% credible intervals (CrI) will be calculated.

7.3.7. Performance measures

Bias

Bias will be measured by taking the difference between the average estimate of the treatment effect and the true treatment effect. Bias is defined as $\delta = \mathbb{E}[\hat{\theta}] - \theta$ [18].

Where θ is the true treatment effect and $\hat{\theta}$ is estimate of the treatment effect. The impact of bias will be presented visually by plotting all points from the simulations on the same graph against the true effect, allowing for quick comparison of the different methods.

Coverage

If we define a CI $(\hat{\theta}_{low}, \hat{\theta}_{upp})$ as the $P(\hat{\theta}_{low} \leq \theta \leq \hat{\theta}_{upp}) = \pi$ where $\pi \in [0, 1]$ then a 95% CI is when $P(\hat{\theta}_{low} \leq \theta \leq \hat{\theta}_{upp}) = 0.95$. It follows that coverage is the $P(\hat{\theta}_{low} \leq \theta \leq \hat{\theta}_{upp})$ [18, 19]. In practice, we can count the number of simulations where the true simulated value is contained in the 95% CI and divided by the total number of simulations. The coverage will be assessed visually through the use of Zip plots [19]. Zip plots are a new visualisation created by Morris et al. [19], which helps to assess coverage of a method by viewing the CIs directly. For each scenario and method, the CIs are centile-ranked according to their significance against the true intervention effect. The vertical axis is represents the ranking and is plotted against the intervals. Intervals that contain the true intervention effect are coloured blue, and intervals excluding the true effect coloured orange. Finally the Monte Carlo 95% CI for percentage coverage is represented by a red dashed line. Zip plots allow efficient comparison of the coverage between a range a methods and display the Monte Carlo error [20] of the coverage.

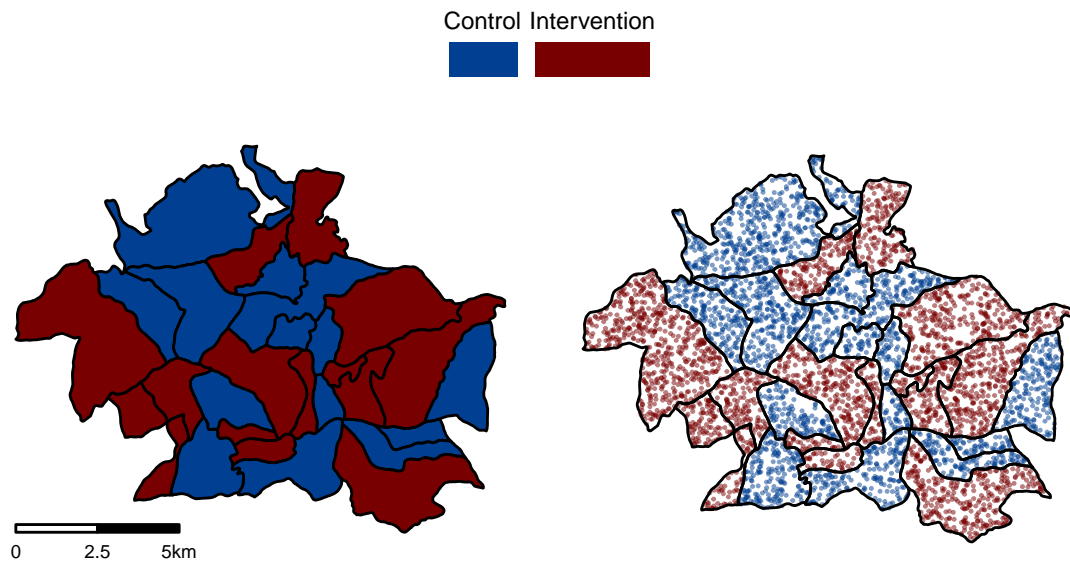
The simulations for the spatial models will calculate a 95% CrI, although the interpretation of the interval changes, the use of coverage remains the same.

7.4. Results

7.4.1. General characteristics

There were 2,359 (47.8%) observations in the control arm and 2,641 (52.2%) in the intervention arm. The number of observations per cluster was wide ranging, with a median of 128 (min = 26, max = 537). The median distance from control to intervention observations was 529.5m (min = 20.96, max = 2,661.5m). Sixty seven (2.8%) control observations were within 100m of an intervention observation, 315 (13.3%) were within 200m, and 1,118 (47.4%) were within 500m. The study area was densely populated, and the cluster shapes were irregular polygons as can be seen in Figure (7.1).

Figure 7.1. Map of study area



7.4.2. Standard CRT analysis

7.4.2.1. Spatial spillover

Spatial spillover has the capability to bias estimates from a standard one stage CRT model, Figure 7.2. In this simulation the spatial spillover effect biased the estimate towards the null. There was minimal bias for the spatial spillover effect with a range

of 100 meters. There was obvious bias when the range of spillover was increased to 200m, and the strength was above 40% of the intervention effect. When the range of spillover was 500m, the intervention estimates were strongly biased away from the true effect of 2, even when the maximum spillover effect was 0.2. The effect on the coverage of the model showed similar patterns to that of bias as can be seen in Figure 7.3. Consistent with bias, coverage is mostly unaffected when the range is 100m, with a very pronounced effect for ranges of 200 and 500m. The scenario of 500m and a spillover effect of 0.99 excluded the true intervention effect in almost all simulations.

7.4.2.2. Spatially correlated intervention effect

It was also clear from Figure 7.2 that a spatially correlated intervention effect can bias intervention estimates. In contrast to the spatial spillover, it biased results away from the null giving stronger effect estimates. Bias was present for all ranges of connectivity between neighbours, and with greater bias resulted from both a larger range, and stronger spatial correlation as would be expected. Coverage was also reduced when the intervention effect was spatially correlated (Figure 7.4) .

7.4.2.3. Spatially correlated errors

The presence of a spatially correlated error term had no observable effect on the bias of the intervention estimates (Figure 7.2). This was also reflected in the coverage of the models which was unaffected by the presence of spatial correlation in the error term (Figure 7.5). This contrasts with the spatial spillover and spatially correlated intervention simulations.

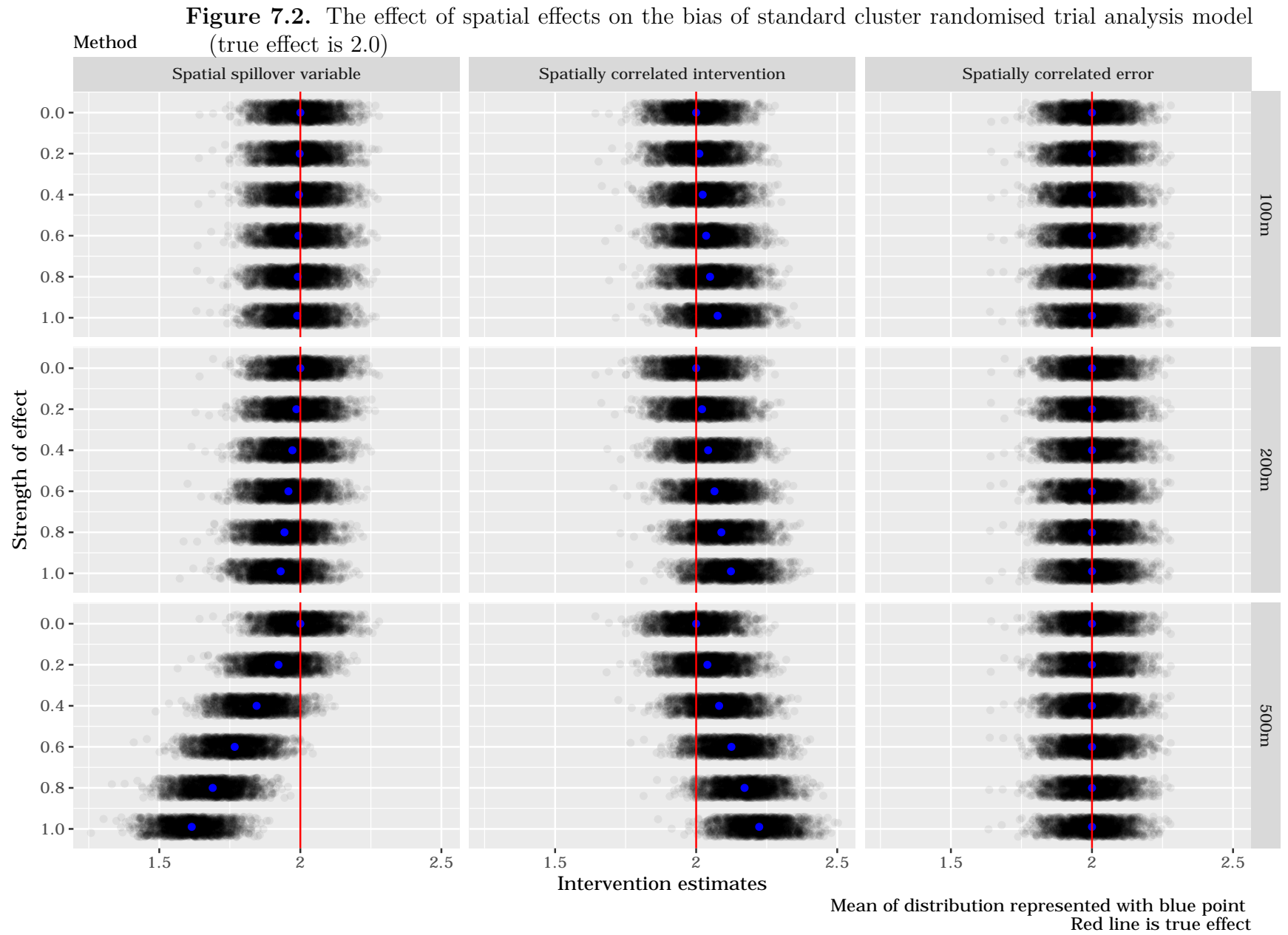


Figure 7.3. Zipplot of the effect of a spatial spillover effect on the coverage of cluster randomised trial model

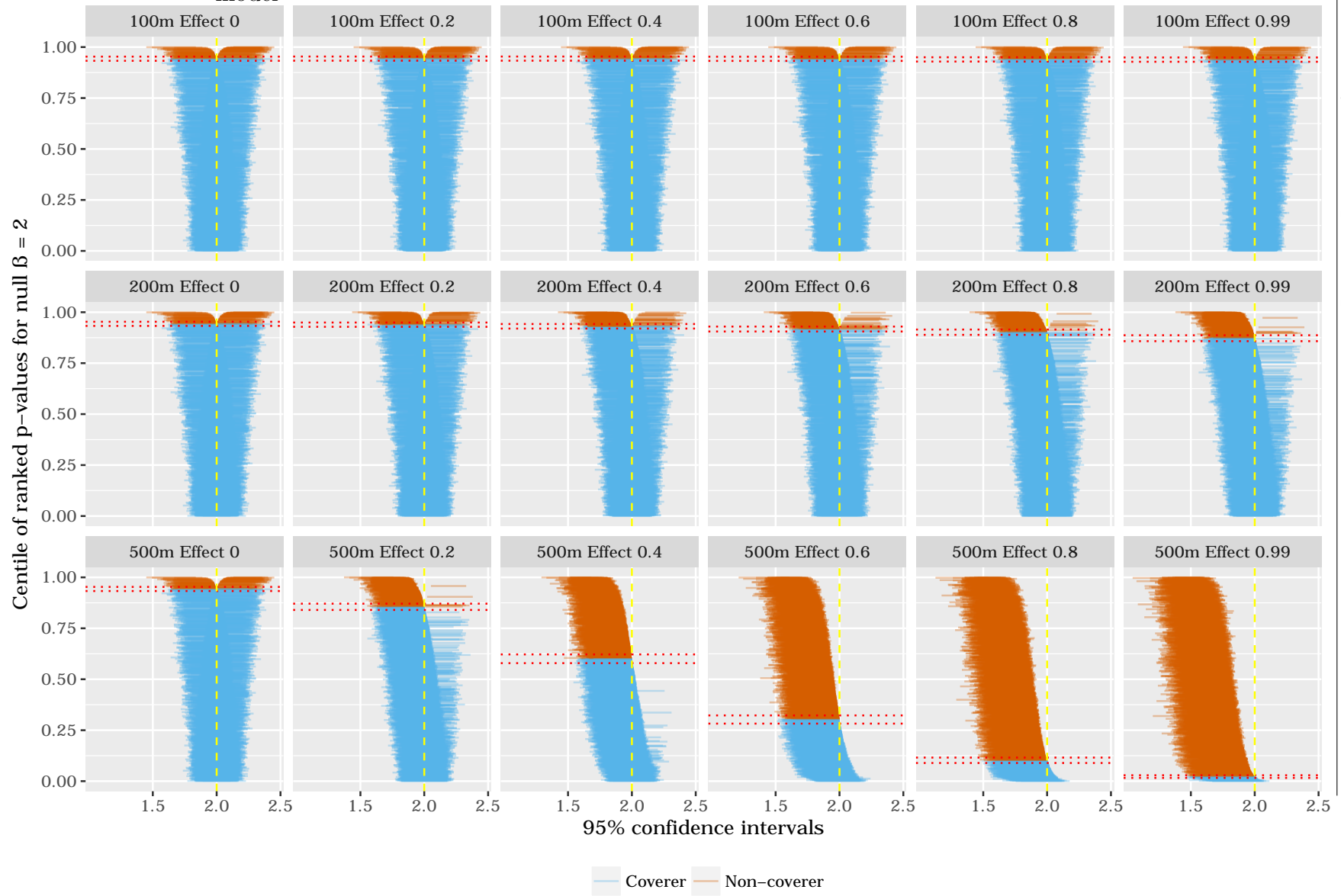


Figure 7.4. Zipplot of the effect of a simultaneous autoregressive spatially correlated intervention effect on the coverage of a cluster randomised trial model

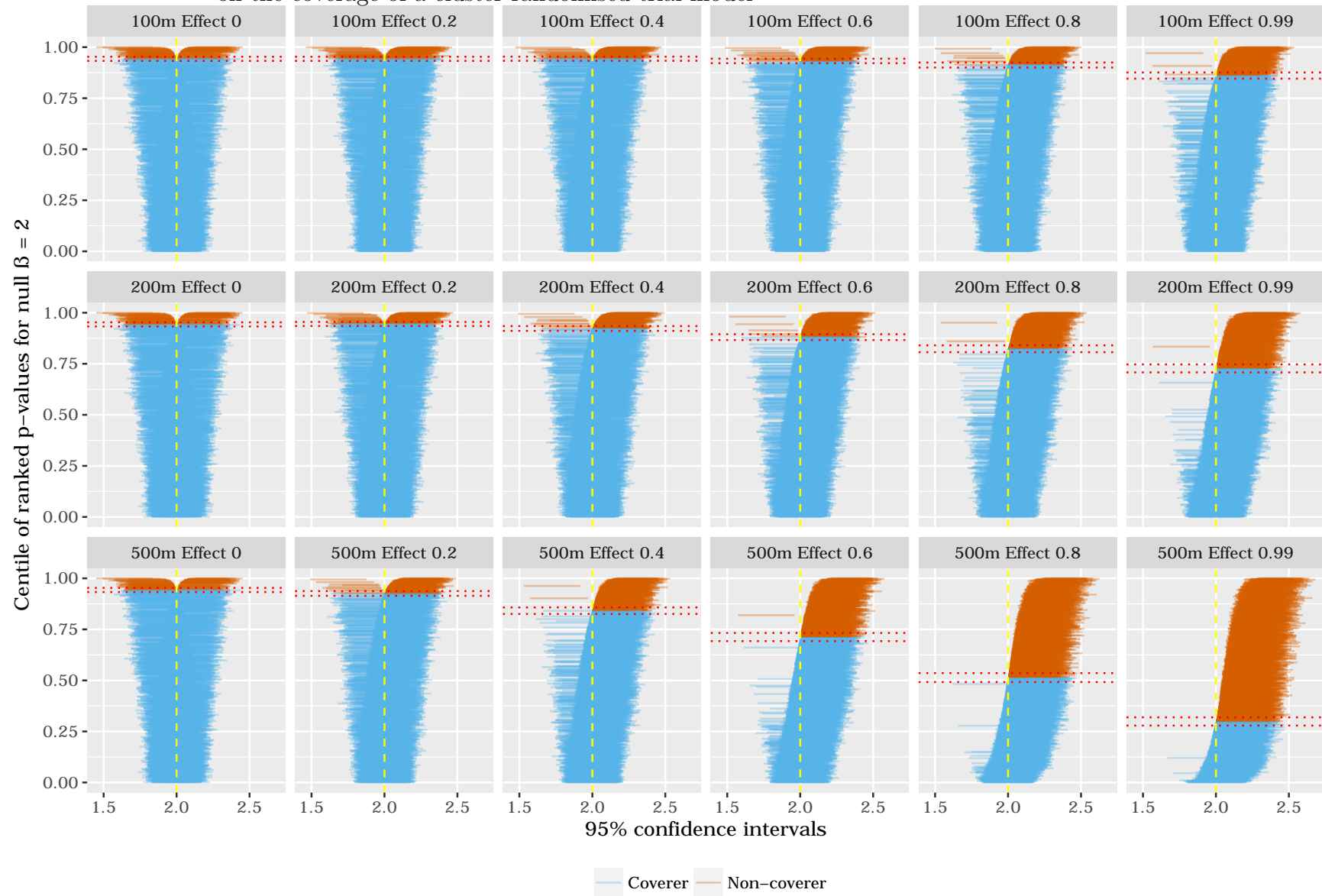
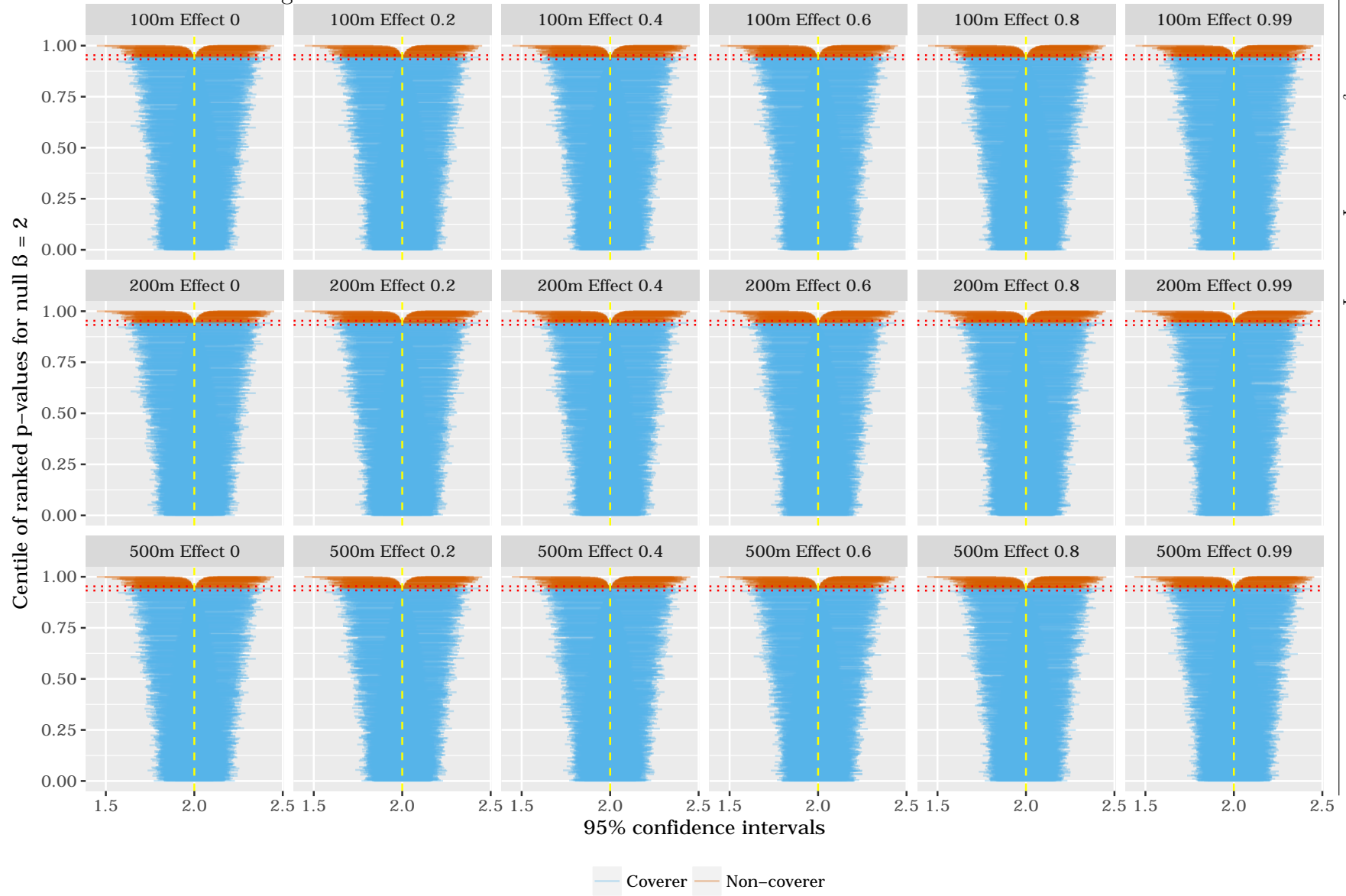


Figure 7.5. Zipplot of the effect of a simultaneous autoregressive spatially correlated error term on the coverage of a cluster randomised trial model



Simulation study of the impact of spatial effects in CRTs

7.4.3. Spatial analysis of CRTs

7.4.3.1. Bias

The BYM and GPm models also gave biased estimates of the intervention effect in the presence of a spatial spillover effect (Figure 7.6). Bias was less pronounced for a spillover effect of range 200m but very pronounced for an effect over a 500m range. The spatial spillover effect biased the intervention estimates towards the null. The estimates from the BYM and GPm models were also strongly biased when the intervention effect was spatially correlated. The estimates were both biased away from the null, resulting in larger estimates of the intervention effect. The results from the cluster and individual level spatial models were very similar despite accounting for spatial connectivity at different scales.

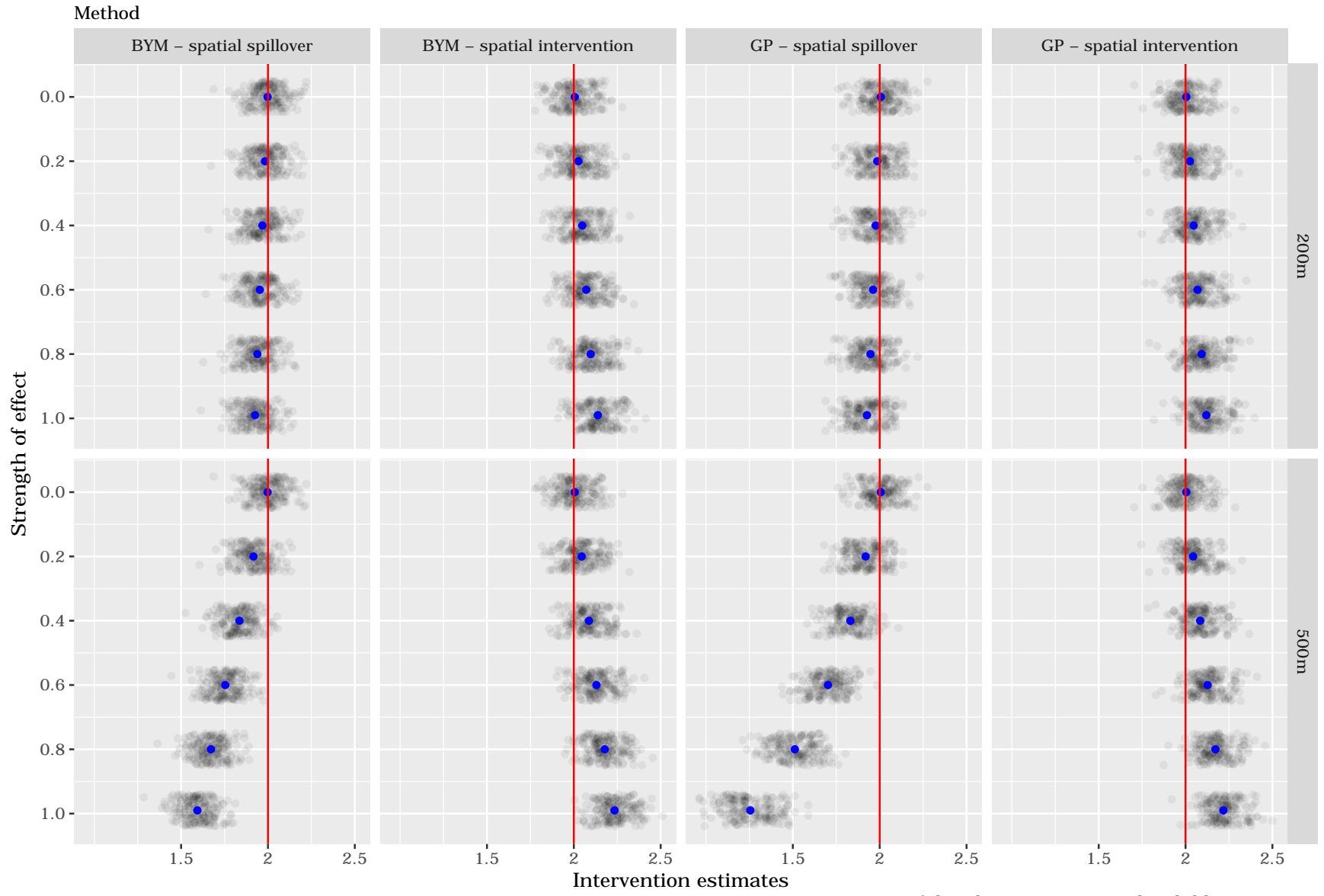
7.4.3.2. Coverage

The coverage of the BYM and GPm models were affected by spatial spillover and spatially correlated intervention effects. There was very low coverage for the spatial effects with a range of 500m. The Zipplots were consistent to the Zipplots of the non-spatial one-stage CRT model and are presented at the end of the chapter in Figures 7.7, 7.8, 7.9, and 7.10.

7.4.3.3. Comparison to non-spatial model

The results from applying spatial models to the simulated data were near identical to the results from the standard CRT model. This suggests that allowing for a spatially correlated covariance matrix, did not account well for the simulated spatial mechanisms.

Figure 7.6. The effect of the spatial effects on the bias of the spatial models



Mean of distribution represented with blue point
Red line is true effect

7.5. Discussion

The simulation results indicate that spatial effects can affect the bias and coverage of intervention estimates from CRTs. Specifically these simulations resulted in spatial spillover effects biasing intervention estimates towards the null, whereas spatially correlated intervention effects increased estimates of the intervention effect biasing away from the null. Contrastingly, spatially correlated errors had little impact on the results of the CRT.

The direction of bias for the spatial spillover effect is plausible. When positive spillover effects are present, individuals in the control cluster who are in proximity to the intervention receive a beneficial effect. Control participants affected by spillover, will be more similar to intervention participants compared to control participants who are not subject to spillover effects. Therefore, the control and intervention arm will be more similar, reducing the size of the intervention estimate. Conversely, if there is a negative spillover effect, then individuals near intervention boundaries will be negatively impacted and intervention estimates may be biased away from the null.

Simulation of a spatially correlated intervention effect biased results away from the null. In contrast to the spatial spillover, intervention observations were also affected by the spatial correlated intervention effect. This would be plausible when the spatial effect is additive. Intervention participants receive not only the direct benefit of the intervention, but also a cumulative effect from nearby intervention participants. In this case as the density of intervention participants is greater in the intervention clusters than in control clusters, the estimated intervention effect may increase. This is a result of the intervention participants being subject to a greater number of spatial effects than control participants. This suggests that within-cluster effects stemming from a spatially correlated intervention can outweigh the bias to the null stemming from between-cluster spatial spillover spatial effects. This would lead to an overestimation of the direct effect of the intervention, as the models struggle to

distinguish between the direct and the spatially correlated effects of the intervention.

The addition of a spatially correlated error term did not affect the intervention estimate regardless of range, or strength of spatial correlation. This may be due to randomisation adjusting for the underlying spatial correlation, and suggest that trialists need only be concerned with spatial effects, which stem from the introduction of their intervention or that interact with the intervention. This could be further tested by comparing varying the random allocation within the simulation study.

Unsurprisingly, increasing the magnitude and distance of the spillover or spatially correlated intervention effect resulted in greater bias. Spatial spillover had little impact at a range of 100m, but past 100m it was more marked. A spillover effect at 100m affects roughly 2.8% of the control participants in this setting. Bias was present for even when the intervention effect had a small level of spatial correlation.

These results, need to be considered in the context of analyses of real trial data to see if the level of spatial effects are plausible. Clearly an effect over a range of 500m, which provides 99% of the intervention effect to non-intervention participants is implausibly high. However, spillover effects in practice will be represented somewhere in the range of simulated scenarios (no spillover through to very strong spillover effects). Therefore, when designing a trial, consideration of how much spillover is expected and the proximity of individuals to one another is of importance. Such information, if known, could be incorporated into sample size calculations and therefore increase the power of trials.

The spatial models used, did not reduce the bias, or improve coverage, suggesting that they did not capture the spatial mechanism well. Perhaps a more fruitful direction of research, is to attempt to model the spillover effect directly through the use of covariates in a model. Work on different ways of specifying the spillover variables could be conducted and general forms developed. For example, Chao et al. propose a method that incorporates the surrounding risk of individuals in a study [21].

There are several limitations to the simulations presented in this chapter. Only two types of spatial methods were used, and they potentially do not capture the spatial mechanisms well. A further weakness stems from the use of locations simulated under spatial randomness. Households are typically grouped together in neighbourhoods, and this structure may impact the generalisability of the conclusions. In addition, the locations used for the simulations were fixed and the spatial distribution of the observations may impact the spatial effect in a given context. Further work, considering a broader range of spatial methods and simulating clustered locations, or using real household locations could be conducted to assess the impact of these weaknesses on the conclusions.

The ICC of the simulated trial was fixed at 0.05 and in the interest of not expanding number of scenarios, the impact of ICC was not considered. I hypothesise that decreasing the ICC, will result in less impact of spatial spillover as the power of the trial will increase, and increasing the ICC will increase the impact of spatial effects as power is reduced. In addition to this, only one size of intervention effect was considered. For a fixed sample size, if the intervention effect is small, the presence of a spatial effect will have a greater impact on the trial's ability to detect an effect, compared to a very large intervention effect. The impact of ICC and intervention effect size on spatial effects in CRT is another area to be considered in future simulations. Furthermore, the context of what ICC means in a CRT with between-cluster dependence, is worth consideration.

Although these limitations hamper the generalisability of the study, they provide plenty of areas for further research into how spatial effects impact CRTs. I consider this study as possibly the first step of many, that allow exploration of various aspects of how spatial effects may impact CRTs and consider solutions for estimating such effects.

In conclusion, this study supports that spatial effects can have a large impact on the results of CRTs, and the use of spatial models appears to do little to alleviate the

effects of spatial spillover or spatially correlated effects. However, the presence of underlying spatial correlation does not appear to impact trial results, and therefore greater focus should be placed on spatial effects that interact with or stem from the intervention.

Figure 7.7. Zipplot of the effect of a spatial spillover effect on the coverage of a BYM model

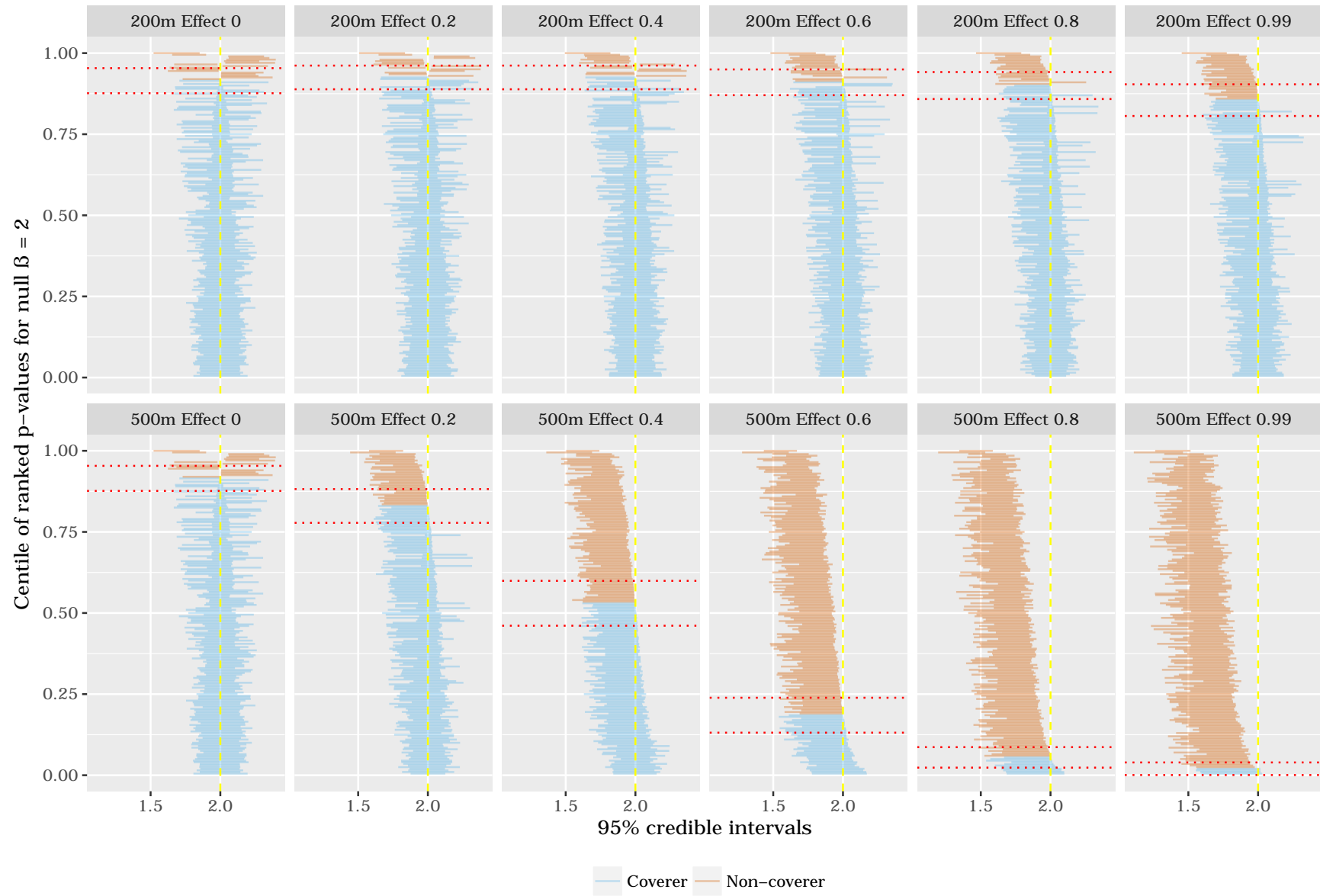


Figure 7.8. Zipplot of the effect of a spatial spillover effect on the coverage of a GPm model

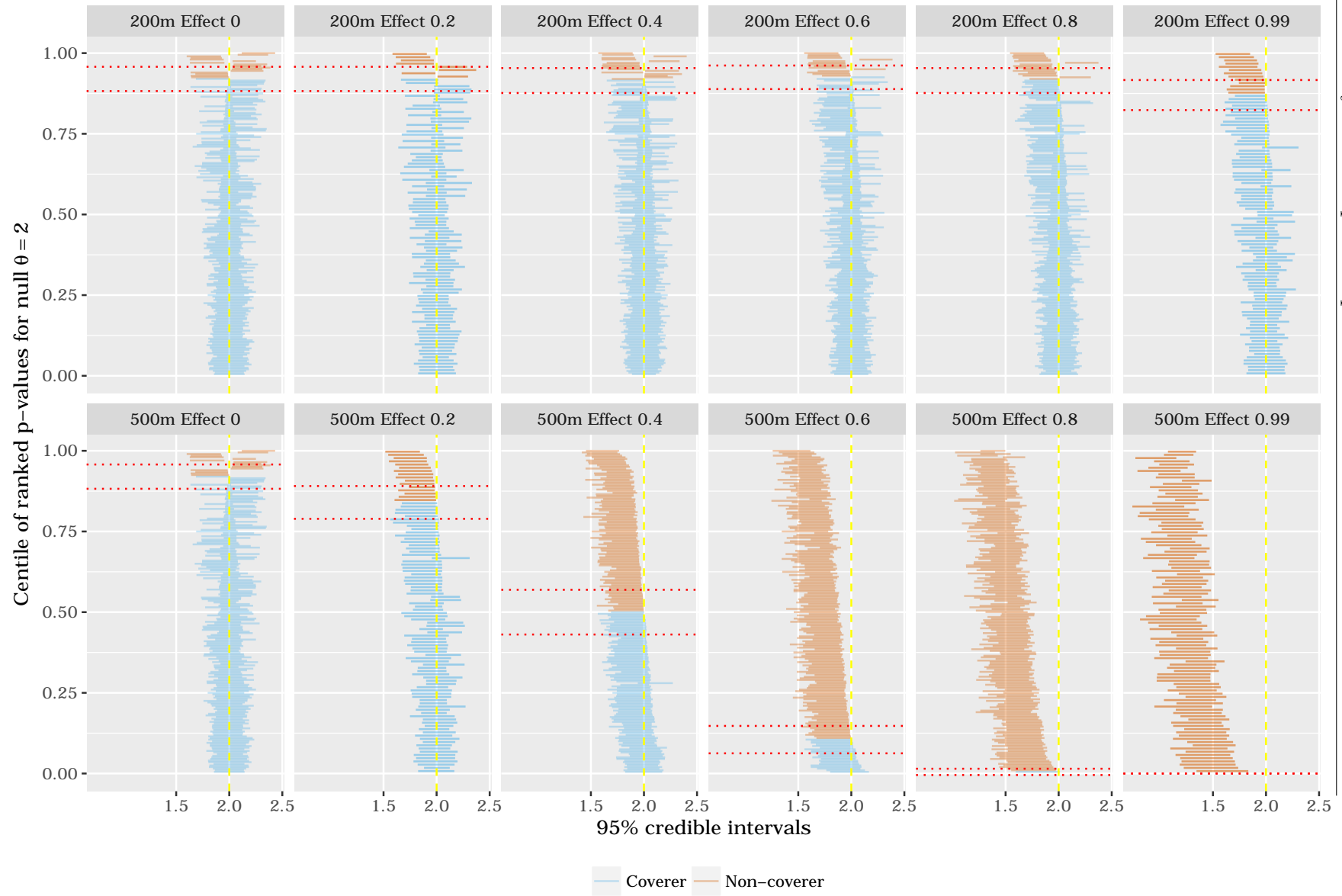


Figure 7.9. Zipplot of the effect of a spatial spillover effect on the coverage of a BYM model

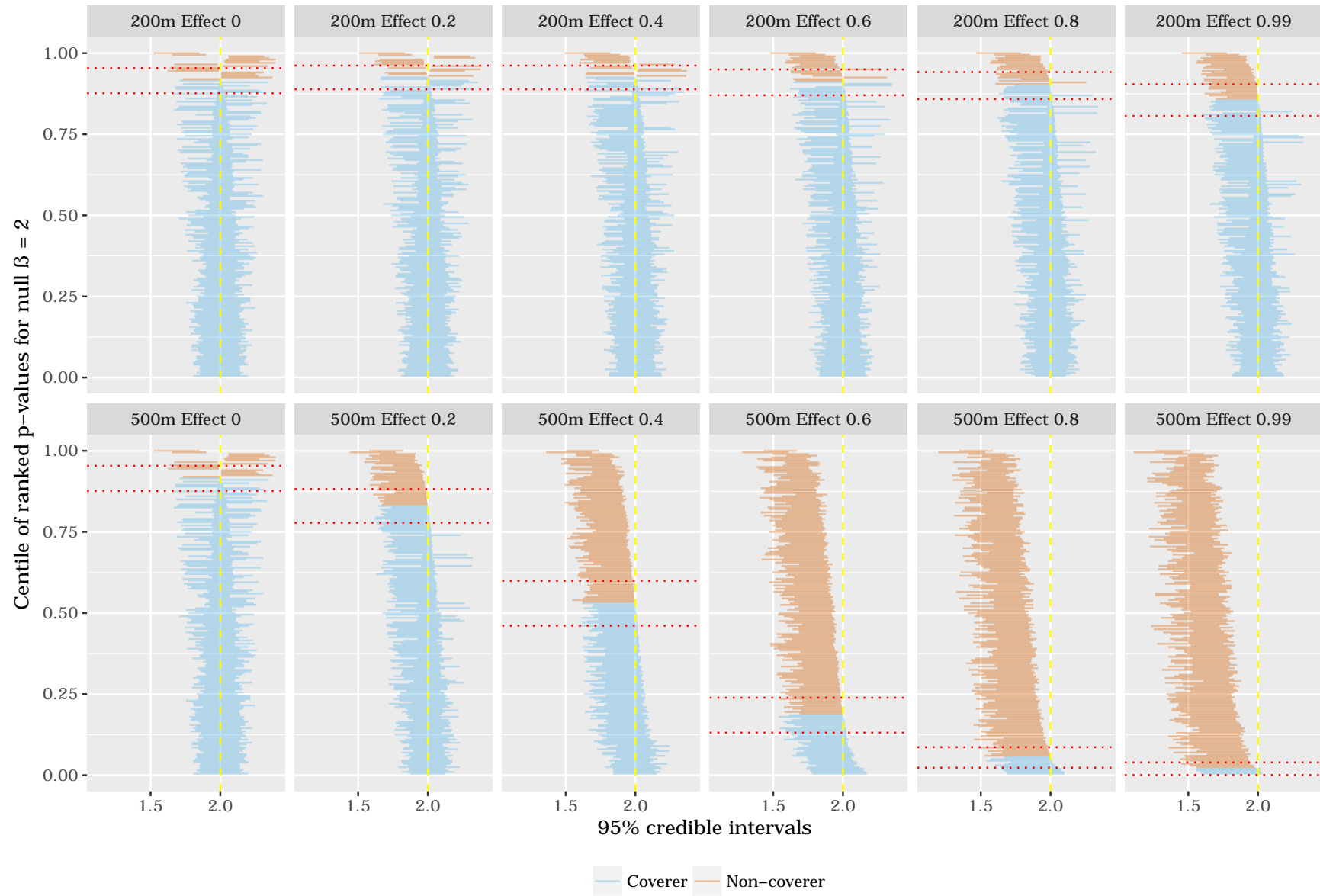
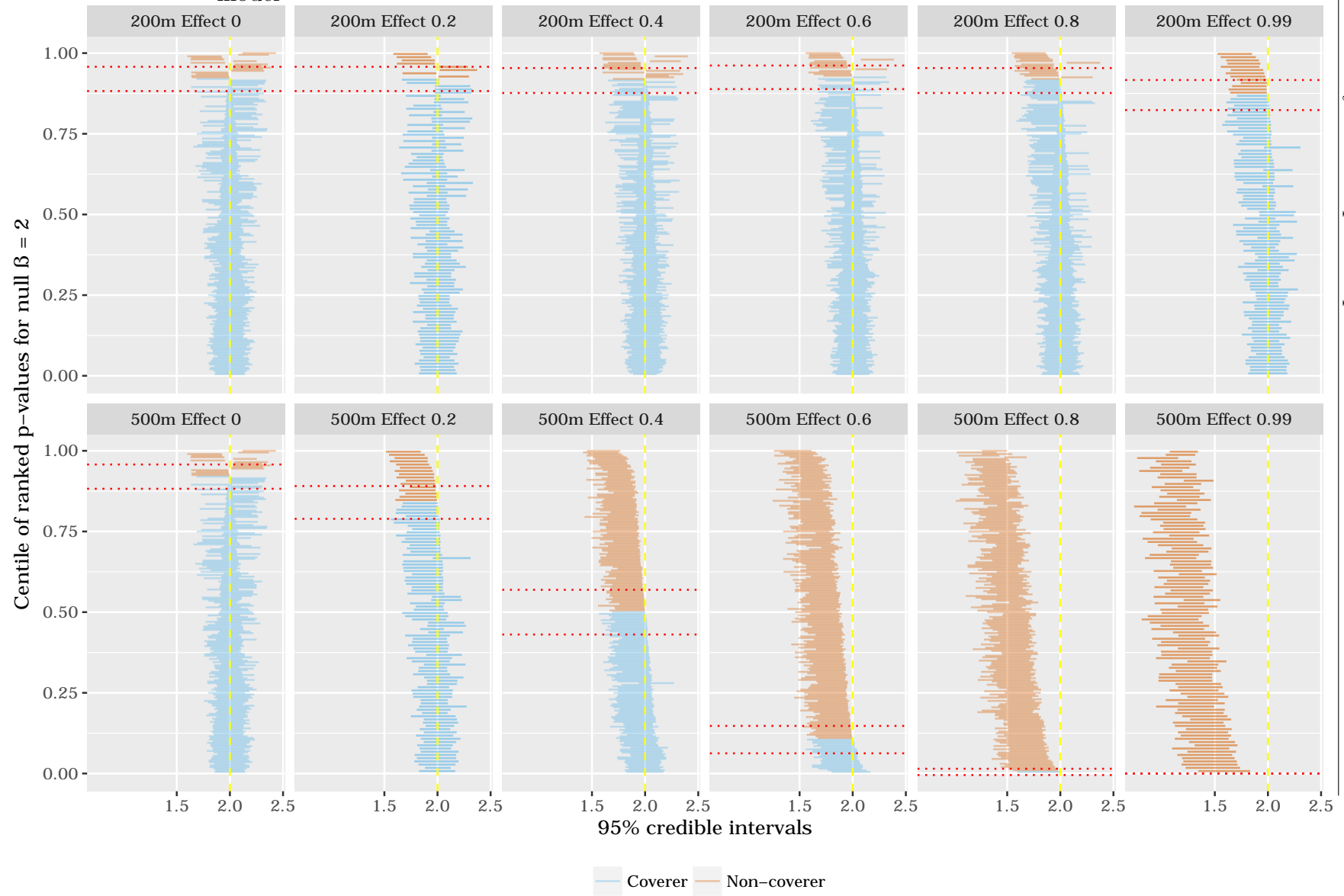


Figure 7.10. Zipplot of the effect of a spatially correlated intervention effect on the coverage of a GPm model



7.6. Bibliography

- [1] Benjamin-Chung, J, Arnold, BF, Berger, D, Luby, SP, Miguel, E, Colford Jr, JM, and Hubbard, AE. Spillover effects in epidemiology: parameters, study designs and methodological considerations. *International Journal of Epidemiology* 2017. (March) 332–347.
- [2] Benjamin-Chung, J, Abedin, J, Berger, D, Clark, A, Jimenez, V, Konagaya, E, Tran, D, Arnold, BF, Hubbard, AE, Luby, SP, Miguel, E, and Colford, JM. Spillover effects on health outcomes in low-and middle-income countries: A systematic review. *International Journal of Epidemiology* 2017 46. (4) 1251–1276.
- [3] Binka, FN, Indome, F, and Smith, T. Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. *The American Journal of Tropical Medicine and Hygiene* 1998 59. (1) 80–85.
- [4] Hayes, RJ and Moulton, LH. Cluster randomised trials. 2nd. Boca Raton: Chapman & Hall/ CRC, 2017.
- [5] Jarvis, C, Di Tanna, GL, Lewis, D, Alexander, N, and Edmunds, WJ. Spatial analysis of cluster randomised trials: A systematic review of analysis methods. *Emerging Themes in Epidemiology* 2017 14. (12).
- [6] Binka, FN, Kubaje, A, Adjuik, M, Williams, LA, Lengeler, C, Maude, GH, Armah, GE, Kajihara, B, Adiamah, JH, and Smith, PG. Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Trop Med Int Health* 1996 1. (2) 147–154.
- [7] Adams, G, Gulliford, MC, Ukoumunne, OC, Eldridge, S, Chinn, S, and Campbell, MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *Journal of Clinical Epidemiology* 2004 57. (8) 785–794.

- [8] Campbell, MK, Fayers, PM, and Grimshaw, JM. Determinants of the intra-cluster correlation coefficient in cluster randomized trials: The case of implementation research. *Clinical Trials* 2005 2. (2) 99–107.
- [9] Bates, D, Mächler, M, Bolker, B, and Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 2015 67. (1) 1–48.
- [10] R Core Team. R: A language and environment for statistical computing. Vienna, Austria, 2017. URL: <https://www.r-project.org/>.
- [11] Besag, J, York, J, and Mollié, A. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 1991 43. (1) 1–20.
- [12] Richardson, S and Guhenneuc-Jouyaux, C. Impact of Cliff and Ord (1969, 1981) on spatial epidemiology. *Geographical Analysis* 2009 41. (4) 444–451.
- [13] Lawson, A. Bayesian Disease Mapping. CRC Press, 2013.
- [14] Lindgren, F. Bayesian Spatial Modelling with R - INLA. 2015 63. (19).
- [15] Rue, H, Martino, S, and Chopin, N. Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical Society Series B* 2009 71. (2) 319–392.
- [16] Lindgren, F, Rue, H, and Lindström, J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 2011 73 423–498.
- [17] Simpson, D, Lindgren, F, and Rue, H. Think continuous: Markovian Gaussian models in spatial statistics. *Spatial Statistics* 2012 1. (1) 16–29. arXiv: 1110.6796.
- [18] Burton, A, Altman, D, Royston, P, and Holder, R. The design of simulation studies in medical statistics. *Statistics in medicine* 2006 25. (24) 4279–4292.

- [19] Morris, TP, White, IR, and Crowther, MJ. Using simulation studies to evaluate statistical methods. *ArXiv e-prints* 2017. arXiv: 1712.03198.
- [20] Koehler, E, Brown, E, and Haneuse, SJPA. On the Assessment of Monte Carlo Error in Simulation-Based Statistical Analyses. *The American statistician* 2009 63. (2) 155–162.
- [21] Chao, DL, Park, JK, Marks, F, Ochiai, RL, Longini, IM, and Halloran, ME. The contribution of neighbours to an individual’s risk of typhoid outcome. *Epidemiology and infection* 2015 143. (16) 3520–7.

8. Cluster reallocation: A method for exploring between-cluster spatial spillover in cluster randomised trials

8.1. Overview

Chapter 6 provided details for simulating spatial effects in cluster randomised trials (CRTs). Chapter 7 built upon this work utilising the methods to test the impact of spatial effects on CRTs through a simulation study. This chapter describes a novel method I have developed called *cluster reallocation*. The cluster reallocation method is tested on a range of simulated conditions using the approaches outlined in the previous two chapters.

Objective

6. Develop methodology to assess the presence of spatial spillover in CRTs.

8.2. Introduction

In the preceding chapter we have seen that spatial spillover effects can impact the results of cluster randomised trials (CRTs). When between-cluster spillover is present, it violates the independence assumption made in a CRT analysis. The impact of spillover may be reduced by having large or well-separated clusters, as then movement across cluster borders is less likely. However, spatial effects may be present over large distances, and designing a trial with well-separated clusters may not be feasible [1]. Furthermore, defining a ‘large’ cluster depends on the type of intervention and outcome, and requires knowledge about the spillover effect. Unfortunately, such knowledge is rarely available at the design stage of a trial when decisions about cluster size can be made. Therefore, detecting and measuring the magnitude of spatial spillover is important within current CRTs and is also important to inform the design of future trials.

There is no standard way of exploring whether spatial spillover is present, although several approaches have been used in the analysis of CRTs. Distance from nearest intervention household, and the number of surrounding intervention households have been used in insecticide treated net (ITN) trials [2, 3]. The risk of the individuals surrounding a participant has been used in a typhoid trial [4]. These three approaches all assume that the mechanism of spillover is based on distance but they also make different assumptions.

Distance to nearest intervention assumes that only distance to a single nearest intervention household is important. Therefore, this approach does not distinguish between proximity to one intervention household and proximity to several intervention households. In contrast, the number of intervention households surrounding an observation assumes that how many intervention participants are nearby is important. This method assumes that spillover is based on the density of the intervention observations nearby, but does not take into account the characteristics of the participants. Using the surrounding risk does includes attribute information about the

individuals surrounding an individual. This method assumes that the characteristics of individuals as well as the density of nearby observations is important. Therefore, the surrounding risk approach would distinguish between a high density of low risk individuals and a high density of high risk individuals, whereas the number of intervention households would not. However, it may be difficult to test whether these assumptions hold, and they will depend on the context of the study.

In this chapter, I propose a method called *cluster reallocation*, which helps to explore the presence of spatial spillover in CRTs. In contrast to other approaches, it assumes that spillover is based on proximity to cluster boundaries, but makes no further assumptions about the mechanism of the spillover. Cluster reallocation draws inspiration from local spatial regression models, such as Geographically weighted regression (GWR), and resampling methods such bootstrapping, the jackknife, and permutation tests [5–7]. In this chapter, I describe how the method works and test its ability to display evidence of spatial spillover in CRTs for a range of simulated spillover effects.

8.3. Cluster reallocation

Cluster reallocation is a relatively simple iterative method that explores the presence of spillover in a CRT. It does this by hypothetically reassigning participants to the intervention or control arm of the trial based on their proximity to cluster boundaries. It could be described as a geographically weighted resampling technique where the resampling is based on physical proximity to the intervention or cluster boundaries. The process for cluster reallocation is displayed visually for a CRT with and without spillover in Figure 8.1, and more formally in Algorithm 8.1.

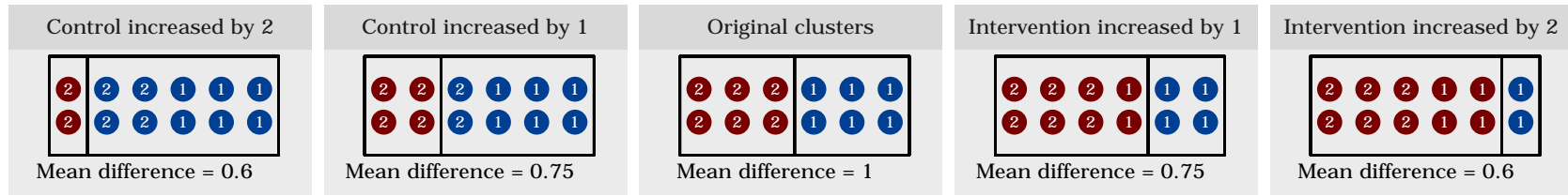
The method involves buffering the intervention cluster boundary, then reallocating control observations within the buffered boundary to the intervention arm. The intervention effect is then calculated using the newly defined trial arms. The process

is repeated for larger and larger buffers, and then repeated for the boundaries of the control clusters. This provides estimates of the intervention effect for hypothetical spatial definitions of the intervention and control arms, referred to as ‘buffered estimates’.

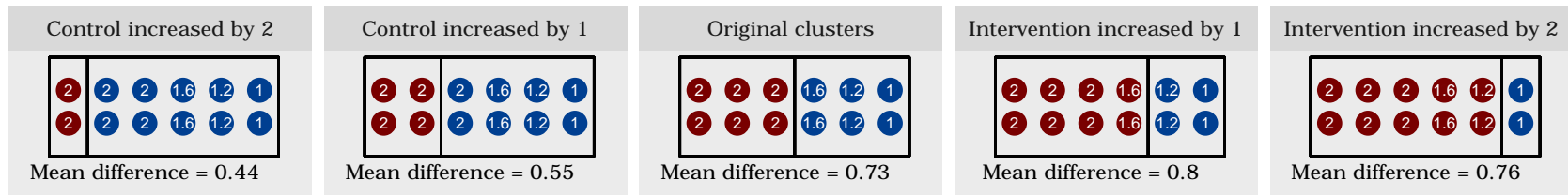
In Figure 8.1, the top part of diagram is a trial where no spatial spillover is present. The left side shows the intervention effect (mean difference) decreases when the control cluster boundaries are increased. The right side of the figure shows that the intervention effect also decreases when the intervention cluster boundary is increased. When spillover is absent, increases in the intervention or cluster boundaries result in weaker intervention effects (biased towards the null). I hypothesise, that in the absence of spillover and when the intervention works, the observed intervention estimate will be either the maximum or minimum estimate compared to the buffered estimates. This is because reallocation of the observations to different trials arms, will increase similarities between the arms, thus diluting the intervention effect.

In the bottom half of Figure 8.1, a trial with positive spatial spillover is presented. When the control clusters are expanded, the intervention estimates decrease. When the intervention cluster is increased, the estimates are larger compared to the original trial estimate. When spillover is present, the intervention estimates from the original cluster definition may not be a maximum or minimum. For example, imagine the intervention has a positive spillover effect on the control participants, meaning that control individuals affected by the spillover have a lower risk of the outcome. The control participants near an intervention boundary may be more similar to the intervention participants than to other control participants, who are further away from an intervention boundary. Reallocating participants near the intervention boundaries, from control to intervention, will increase the differences between the newly defined intervention and control arms, resulting in a stronger effect estimates (further from the null). I propose that this will continue based on the functional distance of the spillover.

Figure 8.1. Diagram of the cluster reallocation method
Study without spillover



Study with spillover



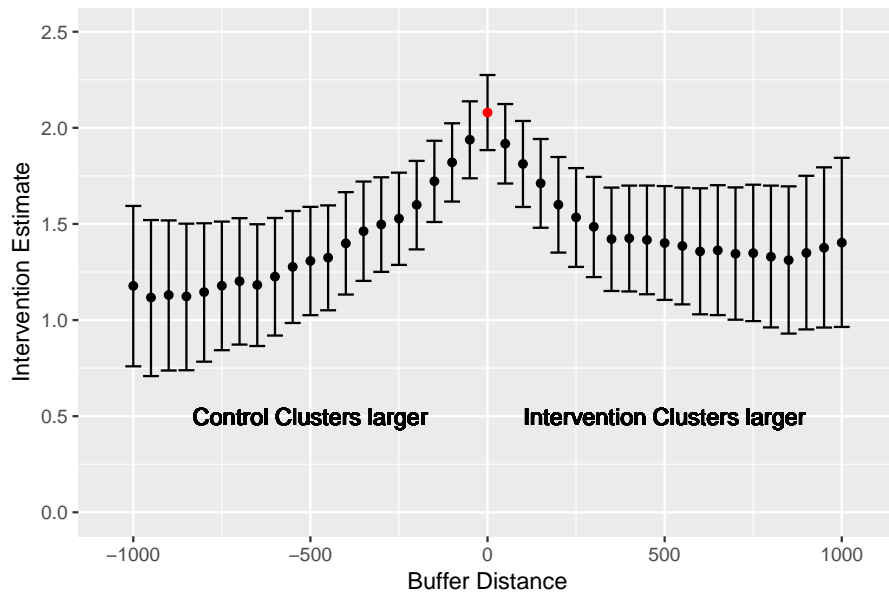
Control (Outcome value = 1) ● Intervention (Outcome value = 2) ●

8.3.1. Cluster reallocation plot

Cluster reallocation provides estimates of the intervention effect for hypothetical spatial definitions of the intervention and control arms. These can be explored graphically through a cluster reallocation plot. A cluster reallocation plot displays the intervention estimate against different buffers and the relation between the calculations involved in the method and the cluster reallocation plot was presented in Figure 8.1. In this section, I will provide examples of the cluster reallocation plot when spillover is absent, and when spillover is present.

8.3.1.1. Without spillover

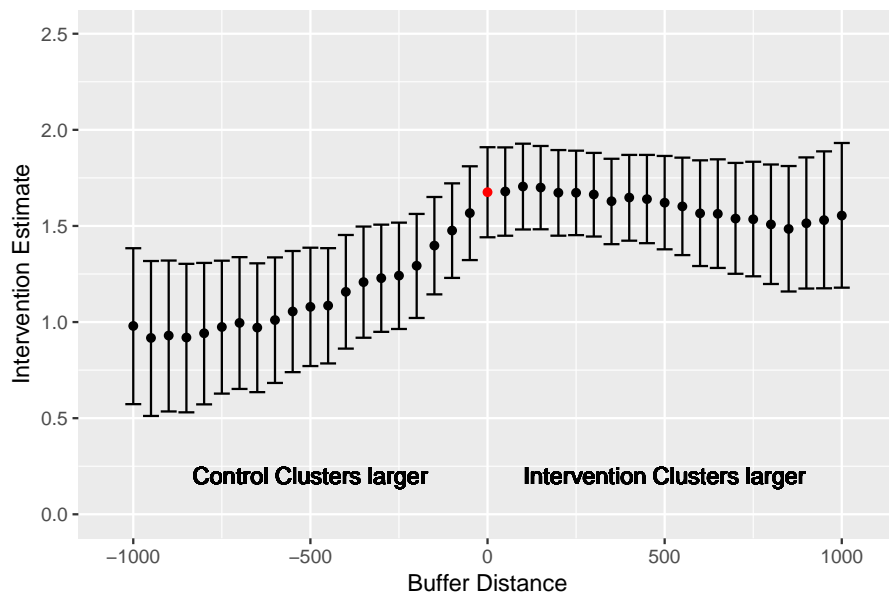
An example plot for a scenario without spillover is given in Figure 8.2. Here, the true intervention estimate is 2.0 and the cluster boundaries were increased to a distance of 1,000m in steps of 50m. In this figure, the red point symbolises the original trial estimate, and the confidence intervals (CIs) around the estimates are provided. As the intervention cluster boundary is increased (moving right, along the x axis) the intervention estimate decreases. As the control cluster boundary is increased (moving left, along the x axis) the intervention estimate also decreases. Therefore, any buffering results in a weaker intervention estimate. In this instance, the pattern is roughly symmetrical, which further suggests an absence of spillover. The CIs become larger as the buffers are increased, reflecting the imbalance between the control and intervention arms. In addition the intervention estimates for a buffer of 1,000m are higher than the 950m buffer, reflecting the impact of random error on the estimates.

Figure 8.2. Cluster reallocation plot when no spillover is present

8.3.1.2. With spillover

In contrast to the previous figure, Figure 8.3 is not symmetric around the original trial intervention estimate. Here, the true intervention estimate is 2.0, and in the presence of spillover the original trial estimate is biased towards the null. In addition, the estimates from buffering the intervention boundaries are larger compared to the intervention estimates from buffering the control cluster boundaries. The intervention estimates are at least as large as the original intervention estimate up to a buffer of 300m when increasing the intervention cluster boundaries. In contrast, increasing the control cluster boundary reduces the strength of the intervention estimates, and the left half of the plot looks consistent with the Figure 8.2 where no spillover was present. Again as the distance of the buffer is increased the confidence intervals become larger, reflecting uncertainty due to the imbalance in the trial arms.

Figure 8.3. Cluster reallocation plot when positive spillover is present



8.3.2. Algorithm

Algorithm 8.1 provides a description of the steps involved in the cluster reallocation method.

Algorithm 8.1 Cluster reallocation

INPUTS

poly the cluster boundaries.

points the locations of the intervention observations

maxdist the maximum distance to buffer the cluster boundaries by.

d_{step} the increase in cluster boundaries for each step.

- 1: Assign *Control* = 0
 - 2: **FOR** *dist* in **SEQUENCE**(BEGIN = 0, MAX = *maxdist*, STEPS = *d_{step}*) {
 - 3: **IF** *Control* = 1 **THEN**
 - 4: Create *poly_{control}* = control cluster boundaries
 - 5: Create *poly_{buff}* = **BUFFER**(*poly_{control}*, **BY** = *dist*)
 - 6: **ELSE IF** *Control* = 0 **THEN**
 - 7: Create *poly_{intervention}* = intervention cluster boundaries
 - 8: Create *poly_{buff}* = **BUFFER**(*poly_{intervention}*, **BY** = *dist*)
 - 9: **END IF**
 - 10: Check which *points* reside in which cluster in *poly_{buff}*
 - 11: Create *Intervention_{dist}* by assigning the *points* to the intervention status of their *poly_{buff}* cluster
 - 12: Fit model using *Intervention_{dist}* as the trial main effect
 - 13: Record intervention effect estimate β_{buff}
 - 14: Change *Control*=1 and repeat from steps 2 to 13
-

8.4. Applying the method to simulated data

In this section, the cluster reallocation method is tested against a range of simulated conditions. Five scenarios with spatial spillover over different distance were simulated, and a range of strengths of spillover were used within each of simulated scenarios. The results of application to simulated data, then lead to considerations for how to use cluster reallocation in practice.

CRT spatial data was simulated using the same process as described in chapter 7. In brief, the model $y_{ijk} = \alpha + \beta x_i + u_j + \epsilon_k + \psi d_k$ was used to generate a CRT with a spatial spillover effect. The spatial spillover effects were simulated using a Gaussian distance weighting function, $f(d_k) = e^{-\frac{d_k^2}{\sigma^2}}$.

For each simulated dataset, the cluster reallocation method was applied using a

maximum buffer distance of 1,000m and steps of 100m. The two parameters that control the magnitude of the spatial effect are σ , the standard deviation of the spatial effect, and ψ , the proportion of the intervention effect that the spillover relates to. The distance values used were $\sigma = 0, 33.3, 100, 200,$ and 400m . When $\sigma = 0$, there is no spillover effect. When $\sigma = 33.3, 100, 200,$ and 400m , the distance weighting functions give a value of 1.0% at 100, 300, 600, and 1,200m respectively.

The distances were chosen as they reflect different scenarios, that may be present when using the method. The scenarios are: (1) no spillover (2) spillover at less than the buffer steps (3) spillover at less than half of the total buffer distance (4) spillover at more than half of the total buffer distance (5) spillover at more than the total buffer distance.

In scenario 1, when $\sigma = 0\text{m}$, the spillover effect will be absent regardless of its strength ($\psi = 0$ was not used to avoid replication). It is expected that this will result in a symmetric plot, where the intervention estimate for original cluster definitions is a maximum.

In scenario 2, when $\sigma = 33.3\text{m}$, the range of spillover is equal to the first step used for cluster reallocation. Therefore, although spillover is present, it should not be reflected in the cluster reallocation plot.

In scenario 3, when $\sigma = 100\text{m}$, the spillover effect is present over a range of 300m. It is expected that this will be reflected in the figure when the intervention clusters are buffered up to 300m.

In scenario 4, when $\sigma = 200\text{m}$, the range of the spillover effect is 600m, which is just over half the distance of the total buffering. It is expected that this will be reflected in the cluster reallocation plot.

In scenario 5, when $\sigma = 400\text{m}$, the range of spillover is 1,200m, which and is greater than the maximum buffer size chosen. It is expected that this will be strongly reflected in the plot and that it may suggest that spillover is present past the maximum distance.

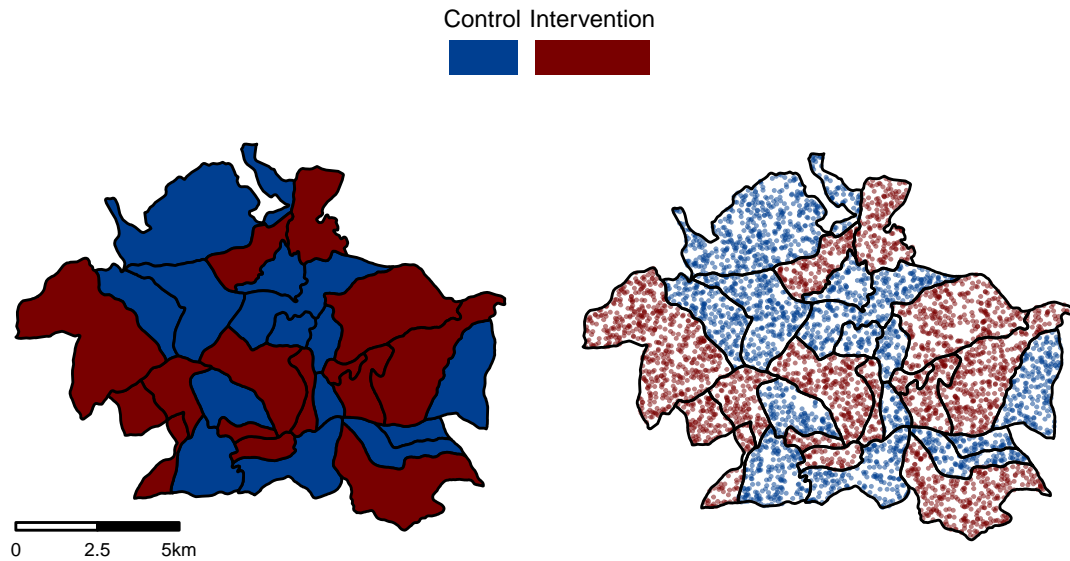
For each scenario, the strength of spillover was varied. The following values were used for $\psi = 0.2, 0.4, 0.6,$ and 0.8 . These values give a range of small to strong spillover effects. One simulated dataset was calculated per combination. Multiple simulations for each combination were not required since the effect of spillover on bias was very marked in the previous simulation study (chapter 7). Further details of the algorithms used can be seen in Algorithms 6.2, 6.3 and 6.4. The following values were kept constant for the simulated data: the intercept $\alpha = 0.2$, the intervention effect $\beta = 2$, the intra-cluster correlation coefficient (ICC) was 0.05 , the number of clusters was 30 , and the sample size was $5,000$ total observations. The locations of the observations were also fixed for all the simulations.

8.4.1. General characteristics of study area

There were $2,359$ (47.8%) observations in the control arm and $2,641$ in the intervention arm. The number of observations per cluster was wide ranging with a median of 128 ($\text{min} = 26, \text{max} = 537$). The median distance from control to intervention observations was 529.5m ($\text{min} = 20.96, \text{max} = 2,661.5\text{m}$). Sixty seven (2.8%) control observations were within 100m of an intervention observation, 315 (13.3%) were within 200m , and $1,118$ (47.4%) were within 500m . The study area was densely populated and the cluster shapes were irregular polygons as can be seen in Figure 8.4.

8.4.2. Cluster reallocation

Cluster reallocation was applied to the five simulated scenarios, as describe in the previous section. Figure 8.5 display the results from the different simulated spillover effects.

Figure 8.4. Map of study area

8.4.2.1. Scenario 1: No spillover

In the top row of Figure 8.5, the range of the spillover effect is zero. Due to this, all the cluster reallocation plots in the top row are identical. The cluster reallocation method behaves as expected, with the original trial estimate being the maximum intervention estimate, and the intervention effect decreasing as the buffers increase. The cluster reallocation plots are symmetric around the main intervention estimate which is marked in red. Though, the individual plots are not exactly symmetric due to random error. As the size of the buffers increase, the confidence intervals become larger, reflecting greater uncertainty in the estimates. This is because buffering results in imbalances between the hypothetical trial arms. Furthermore, in this scenario, as the buffering move past 500m the change in intervention estimates flattens out.

8.4.2.2. Scenario 2: Spillover at less than the buffer steps

In the second row of Figure 8.5, the range of the spillover effect is 100m. The range is equivalent to the first step used for buffering. Although the strength of the spillover effect is varied, there is little, if any difference between the plots within this row.

Furthermore, the cluster reallocation plots are consistent with scenario 1, where spillover is not present. This suggests that the method behaves as anticipated and raises implications for the choice of step size for the cluster reallocation method. In this scenario, a smaller step size may help display the spillover effect.

8.4.2.3. Scenario 3: Spillover at less than half of the total buffer distance

In the third row of Figure 8.5, the range of the spillover effect is 300m. Despite the presence of spillover, the original cluster intervention estimate is the maximum value for all values of ψ . There is very little suggestion of spillover when $\psi = 0.2$, where the cluster reallocation plot is roughly symmetric and consistent with the plots in scenario 1.

There is a slight suggestion of spillover when $\psi = 0.4, 0.6$, or 0.8 . The intervention estimates are slightly higher when increasing the intervention boundary compared to increasing the control boundary. However, the asymmetry is marginal and likely only noticeable as comparison can be made with a plots without spillover.

A further indication of spillover is seen when comparing the original intervention estimate to those in scenario 1. When $\psi = 0.6$ or 0.8 , then the original intervention estimate is lower compared to scenario 1, which reflects bias from spillover. In practice this comparison could not be made, as only one cluster reallocation plot would be created.

8.4.2.4. Scenario 4: Spillover at more than half of the total buffer distance

In the fourth row of Figure 8.5 the range of the spillover effect is 600m. The range of the spillover is roughly half of the overall buffering distance. Again when the spillover proportion is 0.2 , there is little suggestion of spatial spillover in the cluster reallocation plot. There is some suggestion of asymmetry when $\psi = 0.4$, though it is difficult to assess the presence of spillover in this case.

There is a strong suggestion of spatial spillover when ψ is 0.6 or greater. The intervention estimates are much higher when the intervention boundary is increase compared to increasing the control cluster boundaries. Furthermore, when the spillover proportion is 0.8, the intervention estimates for the original cluster definitions is no longer a maximum. The cluster reallocation method has behaved as expected in the presence of spillover. However it appears that a strong spillover effect over a large range is required for the impact of spillover to be noticeable in the cluster reallocation plot.

8.4.2.5. Scenario 5: Spillover at more than the total buffer distance.

In the bottom row of the figure, the range of the spillover effect is 1,200m. The range of spillover is larger than the total buffering distance used. When the proportion of spillover is 0.2 this is still little suggestion of spillover in the cluster reallocation plot. When the proportion of spillover is 0.4 or greater, there is a very strong suggestion of spatial spillover. In particular when the $\psi = 0.8$, the plot demonstrates a near linear relationship between buffering and the intervention estimates. The presence of spillover is clear, and using a maximum buffer distance smaller than the total range of spillover, does not impact conclusions in this scenario.

8.4.2.6. Summary of spillover scenarios

The cluster reallocation method has worked as expected under the simulated scenarios. Despite this, the simulations have raised several implications for interpreting the cluster reallocation plots, particularly when small spillover effects are present.

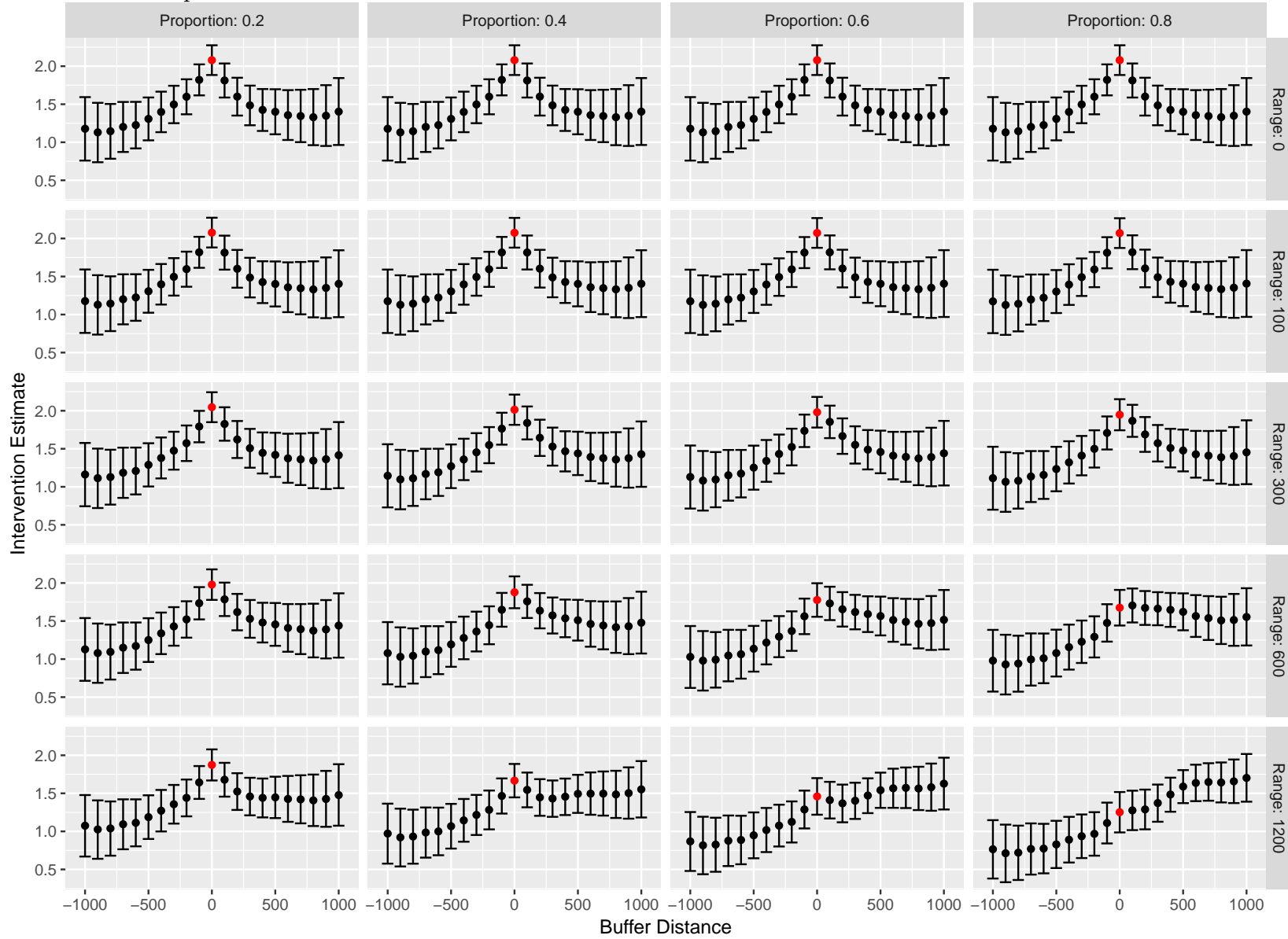
When no spillover is present the cluster reallocation plots are roughly symmetric and the original intervention estimate is a maximum or minimum. However, the original estimate can still be a maximum or minimum when spillover is present, as seen in scenario 3. Therefore, it is not sufficient to conclude spillover is absent based on the original trial estimate being a maximum or minimum.

The cluster reallocation plots in scenario 1 and 3 were very similar. Asymmetry due to spillover may be very slight and difficult to distinguish from asymmetry due to random variation. Therefore, caution is needed when making conclusions based on a lack of symmetry in a cluster reallocation plot.

The simulation study in in chapter 7 demonstrated that spatial spillover effects can bias intervention estimate in CRTs. The cluster reallocation method is consistent with this results and the bias is reflected in the . The intervention estimate of the original trial reduces in the presence of spatial spillover, and the cluster reallocation method can demonstrates spillover more markedly in the scenario 4 and 5 (Figure 8.5).

It appears that the cluster reallocation method is not very sensitive to small scale or strength spillover effects. These simulations suggest that the cluster reallocation method will not reflect spillover unless the range and strength of spillover is large. Combined with the evidence of bias from spatial spillover, this potentially suggests the presence of spillover in a cluster reallocation plot reflects bias in the means that the original intervention estimate.

Figure 8.5. Cluster reallocation plots for simulated scenarios with different strengths and range of spatial spillover effects



8.4.3. Cluster reallocation in practice

The results from applying the method to simulated data raise several implications for the use of the cluster reallocation method. Here I will discuss these points in turn and provide advice on how to apply the method in practice.

8.4.3.1. Determining the step size and maximum buffer distance

The increase in buffer distance for each step will need to be considered in relation to the size of the study area, and the potential spillover effect.

For the intervention estimate to change, individuals in one trial arm need to be reallocated to the other due to buffering. If the step size is too small, then this may result in no observations being reassigned to a different trial arm. In this case, each increase in buffer may not provide additional information. Thus, the step size would be ineffective and a larger step size needed. Furthermore, for a fixed maximum distance, a smaller step size will result in a greater number of steps being taken, increasing the computational burden.

On the other hand, choosing too large a step size may make it hard to determine whether a spillover effect is present. In the simulated scenarios, when a spillover effect was only present up to 100m, then using steps of 100m meant the cluster reallocation plot did not reflect any spillover. Clearly it is inappropriate to use a step size that is larger than the range of the spillover. Therefore, choosing too large a step size may be insensitive to the scale of the spillover effect and result in missing important changes in the intervention.

A further aspect that needs to be determined is the maximum buffer distance. In scenario 5, stopping the buffering before the functional distance lead to an underestimate of the range of the spillover. However, this implication did not affect the ability to detecting the presence of spillover. The maximum buffer distance needs to be larger than the scale of spillover, and needs to allow individuals to be reallocated

between trial arms. It appears that the choice of step size is of greater importance than the maximum buffer distance chosen.

In order to determine the step size for buffering, distance to nearest intervention participants can be calculated. From this, the minimum meaningful distance for cluster reallocation within that data can be calculated. For example, if no individuals is within 50m of the intervention, then the step size will need to be at least 50m. Considering the maximum distance will inform the maximum buffer distance. There is clearly no point in buffering past the maximum distance, as there will be no observations to compare. Categorising the distance distribution will help to determine how many people would be reallocated for each step size. This helps to avoid having steps where no one is reallocated.

The cluster reallocation method can then be used with a step size greater than the minimum distance, and total buffer distance less than the maximum distance. As the method is exploratory, multiple step sizes, and differing maximum distances could be attempted. If computational time is an issue, a larger maximum distance with large steps can be used to initially explore for spillover. Following this, smaller maximum distances with smaller step sizes could be used. If spillover is present when using larger steps, then consideration of smaller steps is likely not needed.

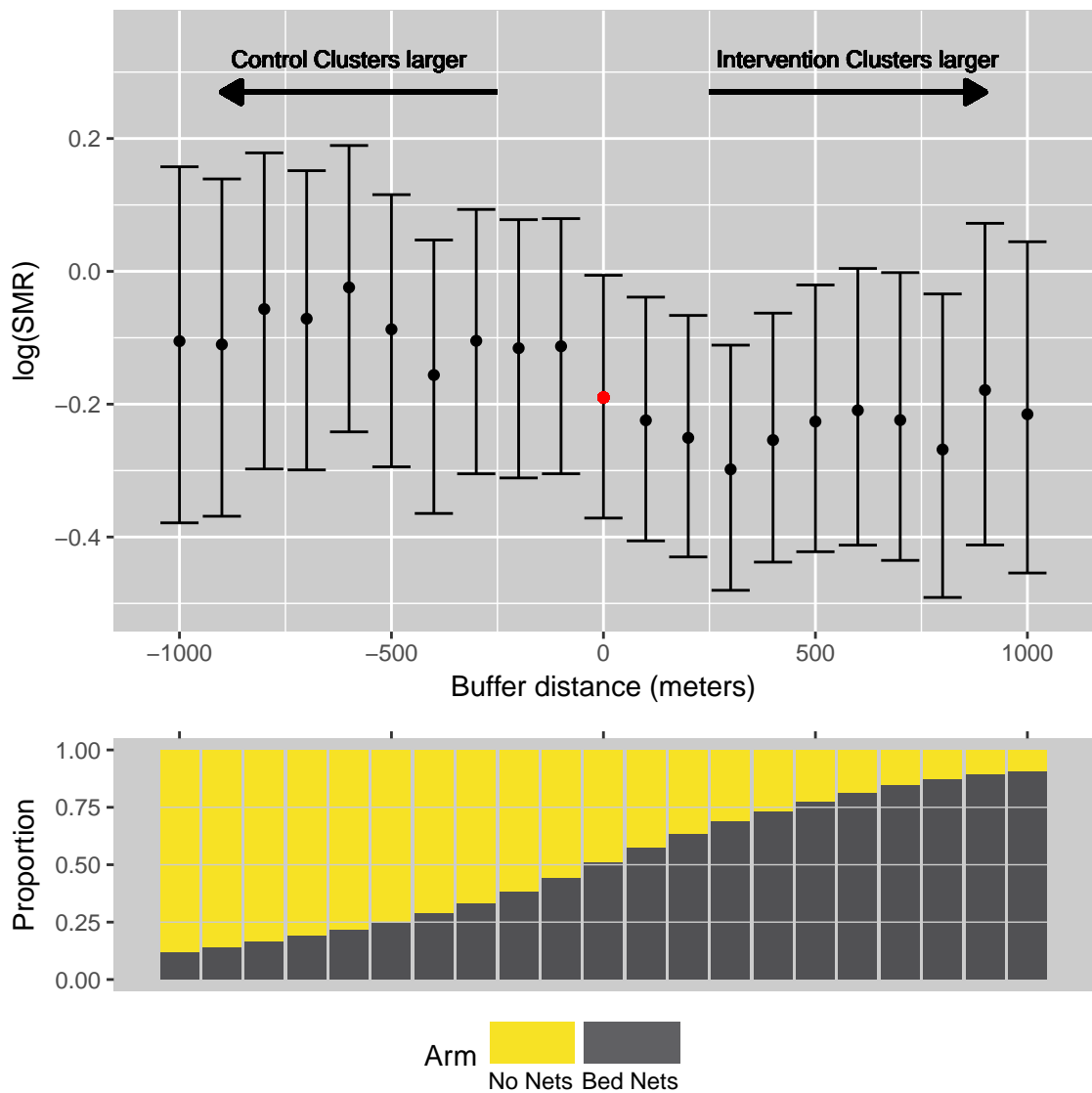
8.4.3.2. Imbalance in trial arms

A further factor that will also affect the maximum buffer distance is the imbalance in the hypothetical trial arms. Participants are either continually reallocated to the control or intervention arm, therefore, as the buffering increases, the imbalance between the trial arms increase. Eventually all participants will be in one arm of the trial, and a comparison can no longer be made.

It is important to consider this imbalance in the interpretation of the results. Changes in intervention estimates where the buffers are large, and imbalances severe, should be treated with caution. A possible adaptation for the cluster reallocation

is presented in Figure 8.6, which was taken from chapter 5. Here, a bar chart is presented below the cluster reallocation plot, each bar represent the proportion of the participant in the intervention and control arm. As can be seen from the lower part of Figure 8.6, the arms are very unbalanced when the buffer is at 1000m. In contrast to the other plots in this chapter, a lower intervention effect is a stronger effect as the trial estimated a standardised mortality ratio (SMR).

Figure 8.6. Cluster reallocation plot with bar chart for displaying imbalances in trial arms (Taken from Chapter 5)



8.5. Summary

This chapter has proposed a method called cluster reallocation for assessing spillover in geographical CRTs. It involves reassigning observations to the control or intervention arm based on their proximity to cluster boundaries, and repeatedly fitting the main trial model to the new trial arm definitions. The approach is agnostic towards the analysis method used in the CRT. Thus, the method can be applied regardless of the type of outcome. This also allows a one- or two-stage method to be used for the comparison. Therefore, it could be applied when there are a small number of clusters, although it may be difficult to measure spillover in such a setting. At present, it is only applicable for CRTs where the clusters are polygons and the observations are o points inside the polygons.

The method draws inspiration from computationally intensive methods such as bootstrapping, permutation tests, and GWR. The approach requires repeatedly applying models to data, and therefore is potentially not feasible for complex models that take a long time to run. Though, typically CRT models consist of a generalised linear mixed effects model with a single random effect for cluster or a generalised estimating equation which have a lower computational burden compared to permutation tests [8].

The main output of cluster reallocation is a graphical display showing the change in intervention estimates for different cluster boundary definitions. Several examples of cluster reallocation plots were given and described in this chapter. The plots provide an indication of the presence and range of spatial spillover effects.

A range of simulated conditions were used to test the method. Cluster reallocation performed well, demonstrating that it can be used to detect spatial spillover effects. However, the simulation identified that results are reliant on the buffer size, and step sizes chosen. When the step size is smaller than the range of spillover, the method will not reflect spillover. In addition, it is important to note that larger distances of buffering will result in imbalances between the control and intervention arms. This

is reflected in the size of the confidence intervals of the estimates with large trial arm imbalances. The step size and maximum buffer distance can be chosen by assessing the spatial distribution of the observations. Further to this, and an adaptation of including a bar chart in the cluster reallocation plot, helps to indicate when trial arms are imbalanced. Further tests will be needed to assess how well it reacts to other types scenarios such as negative spillover effects.

In conclusion, the cluster reallocation method provide a relatively simple approach for assessing for the presence of spatial spillover in CRTs. The method performed well in simulated conditions and further applications to real data and other simulated scenarios is desirable. Work is planned to create an R package so that the method can be used by a wider audience.

8.6. Bibliography

- [1] Howard, SC, Omumbo, J, Nevill, C, Some, ES, Donnelly, CA, and Snow, RW. Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2000 94. (4) 357–360.
- [2] Binka, FN, Indome, F, and Smith, T. Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. *The American Journal of Tropical Medicine and Hygiene* 1998 59. (1) 80–85.
- [3] Lenhart, A, Orelus, N, Maskill, R, Alexander, N, Streit, T, and McCall, PJ. Insecticide-treated bednets to control dengue vectors: Preliminary evidence from a controlled trial in Haiti. *Tropical Medicine and International Health* 2008 13. (1) 56–67.
- [4] Chao, DL, Park, JK, Marks, F, Ochiai, RL, Longini, IM, and Halloran, ME. The contribution of neighbours to an individual’s risk of typhoid outcome. *Epidemiology and infection* 2015 143. (16) 3520–7.
- [5] Brunsdon, C, Fotheringham, aS, and Charlton, ME. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis* 1996 28. (4) 281–298.
- [6] Davison, AC and Hinkley, DV. *Bootstrap Methods and Their Applications*. Cambridge: Cambridge University Press, 1997.
- [7] Efron, B. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* 1979 7. (1) 1–26. arXiv: [arXiv:1306.3979v1](https://arxiv.org/abs/1306.3979v1).
- [8] Eldridge, S and Kerry, S. *A Practical Guide to Cluster Randomised Trials in Health Research*. Cornwall: John Wiley & Sons Ltd, 2012.

9. Discussion

9.1. Overview

In the previous eight chapters, I have provided a detailed consideration of spatial effects and methods in relation to cluster randomised trials (CRTs). In this section, I review and synthesise the key points of the PhD, considering the strengths and limitations of the work. Following this, I will discuss implications for current practice, and possible directions for future work beyond the PhD.

9.2. Summary of findings

The main aim of this thesis was to explore the use of spatial analysis methods within cluster randomised trials (CRTs). Specifically, to improve and develop knowledge of: methods that can be applied to CRTs, the impact of spatial effects on trial results, and the additional utility that can be gained by considering the spatial context of a CRT.

This aim led to the following objectives:

1. Describe and frame CRTs in relation to spatial data and summarise the implications for spatial analysis (chapter 2).
2. Describe and identify spatial analysis methods that have been previously used in CRTs by conducting a systematic review (chapter 3).
3. Apply and assess a range of appropriate modern spatial methods to existing CRT data, in order to analyse the effect of spatial autocorrelation and spatial spillover effects on CRTs results (chapters 4 and 5).
4. Evaluate the impact of spatial effects and the utility of spatial models in the analysis of a CRT by means of a simulation approach (chapter 7 with chapter 6 providing background).
5. Develop methodology to assess the presence of spatial spillover in CRTs (chapter 8 and also applied in chapter 5).

In this section, I will discuss the main findings of the PhD in relation to these objectives. Throughout the thesis, and particularly in the applied papers, greater weight has been given to epidemiological findings due to journal conventions and the intended audiences of the publications. However, in this discussion I will give equal weight to the methodological findings, as methodology is a core consideration of the thesis objectives.

9.2.1. Objective 1: Describe and frame CRTs in relation to spatial data and summarise the implications for spatial analysis.

CRTs are well used in epidemiological studies and many are implicitly spatial. Trial design and management often relies on knowing where observations and clusters are located to administer an intervention or survey a participant. Although many published CRTs may include maps of the trial context, most do not conduct relevant analyses that are spatially explicit.

In chapter 2, a framework was proposed that related spatial data types to CRTs. Using a vector based spatial model, a CRT can be represented in terms of points, lines, and polygons. CRTs can take various forms dependent upon the units of observation and how they are clustered. A typical CRT that observes individuals within geographical areas could use points and polygons respectively, however in a school-based intervention a cluster might be best represented as a point. Examining the spatial representation of CRTs exposed how trials that are comparable in design, and often analysed in the same way, could be conceptualised as having distinct spatial representations. Similarly, a standard CRT analysis would analyse a household CRT and a workplace CRT in the same way, however when estimating spatial effects, the relevant spatial metric for these trials can differ.

Chapter 2 also described how CRT data differs from those traditionally analysed in the three fields of spatial statistics (1. Point process, 2. Geostatistical, 3. Areal). CRT data does not fit neatly within a single area of spatial statistics. Geostatistical, and areal methods were considered most appropriate for CRTs, as the locations of the data are typically fixed. This led to the application of cluster level (areal) and observation level (geostatistical) methods being applied in the PhD. Specifically, the Besag, York, and Mollie model (BYM) and Gaussian Process models (GPM) were tested further through simulation studies (see appendix C for a detailed overview of

these models).

Consideration of the types of spatial representation, common CRT data, and spatial modelling approaches helped to determine which methods were most appropriate for the PhD. A particular focus on the spatial representation of observations and clusters highlighted that tractable spatial CRT analysis models exist that have seen little or no use in the literature to date. For instance, a CRT can be conceived with line-type observations and clusters that are in points, perhaps representing the movement trajectories of individuals tied to particular households. Objective 1 also demonstrated that CRTs presented an uncommon analysis challenge, seemingly straddling several distinct areas of spatial statistical research. Understanding how this challenge had been tackled to date was the remit of objective 2.

9.2.2. Objective 2: Describe and identify spatial analysis methods that have been previously used in CRTs by conducting a systematic review.

The systematic review in chapter 3 demonstrates the relative paucity of published research that uses spatial analysis methods in CRTs. The ten papers found in the review represent a small proportion of CRTs that have been published. This suggests that spatial effects are not often considered in this area. Furthermore, despite evidence of spatial effects in the papers reviewed, they were rarely adjusted for in the primary analysis of the trial. Thus, there were few examples to draw upon when deciding how to conduct a spatial analysis of a CRT.

The review identified two approaches to analysing spatial data in CRTs; spatial variables and spatial modelling. The spatial variable approach includes a proximity based measurement as a covariate in a model. The spatial modelling approach incorporates spatially structured random effects in a model. On reflection, the spatial variable approach relates to measuring spatial spillover, and the spatial modelling

approach relates to accounting for spatial correlation.

All thirteen trials in the review found evidence of a spatial effect within their studies. There was evidence to suggest that accounting for spatial structure affects both the precision and point estimates of intervention effects, and failure to do so could give inaccurate results [1, 2]. Silcocks & Kendrick demonstrated that spatial models fitted better than a standard CRT random effects model, by comparing the Akaike information criteria of the models [1]. However, the study of Silcocks & Kendrick was mainly focused on multi-membership, where observations in CRTs are members to more than one cluster, which differs slightly to the focus of spatial effects in this thesis [1].

The literature review highlighted that the specific research question under study could influence the appropriateness of a given analysis method. The use of relevant spatial models can produce a better fitting model that reduces the biasing effect of spatial correlation on model estimates. However, when the research question asks specifically about the effect of spillover, the spatial variable approach allows for a separate estimate of the spillover effect to be made. The analysis methods found in the literature review were subsequently updated and applied to objectives 3 and 4.

9.2.3. Objective 3: Apply and assess a range of appropriate modern spatial methods to existing CRT data, in order to analyse the effect of spatial autocorrelation and spatial spillover effects on CRT results.

Chapters 4 and 5 applied spatial analysis methods to the analysis and reanalysis of real CRTs. The research in chapter 4 was part of a contemporary CRT of the oral polio vaccine (OPV), and demonstrated an applied spatial analysis of poliovirus shedding transmission from the OPV. The chapter involved mapping the spatial data of the trial over time, and the use of spatial variables in regression models.

In comparison, chapter 5 was a reanalysis of an existing CRT of insecticide treated bed nets, which had previously demonstrated evidence of spatial spillovers, using a range of modern spatial analysis methods.

In chapter 4, the spatial variables provided little extra insight, and it appears that household locations and spatial distribution have limited utility in predicting the transmission of poliovirus shedding. The spatial variables used were ‘distance to nearest shedding’ and ‘the number of individuals shedding within a certain distance’. It is possible that other measurements might better represent spatial proximity, particularly concerning individual mobility, however, with only household location available, this was as granular a spatial analysis as could be conducted on these data.

Despite the limited benefit of spatial variables, the value of spatial visualisation was very apparent. In this case, mapping the data over time showed that transmission occurred rapidly over the study area, and this may explain why the spatial models were less relevant. Since transmission in this context is faecal-oral, it implies that either more detailed spatio-temporal data are required, or that transmission is perhaps a spatial and a social process.

From an epidemiological perspective, the chapter supported the global cessation of OPV [3, 4]. It highlighted that poliovirus transmission shedding is likely more present than previously thought. This may explain why the year for eradication of polio has been delayed multiple times [5].

In chapter 5, a range of spatial methods were used to examine the existence of spatial spillovers. It also considered the impact of spatial effects on the conclusions of a CRT of insecticide treated nets (ITNs). The chapter demonstrated the benefit of collecting GPS data beyond the creation of maps. The analyses only required GPS coordinates, did not necessitate any new data collection, and could be applied to previous geographical CRTs.

The analyses of chapter 5, were split into 3 parts: (1) exploratory spatial data

analysis; characterisation of spatial dependence, heterogeneity, and spillover; (2) spatial modelling of the intervention effect to estimate the true intervention effect in light of any anticipated spatial effects; (3) analysis of distance based spillover and interaction with the intervention to characterise the magnitude of spillover and the functional distance over which the spillover effect is present.

Exploratory spatial data analysis demonstrated the utility of Moran's I and geographically weighted regression (GWR) to explore spatial patterns in the data. These approaches were strongly suggestive of a spatial pattern. The utility of the join count statistic was less clear.

Multiple approaches strongly suggested evidence of a positive spatial spillover effect due to being near households that use bed nets. Allowing for the detailed spatial correlations and spillover effects did not change the primary conclusions of the trial. It was unclear whether the spatial models were beneficial compared to the standard one stage CRT model. Consistent conclusions for spatial and non-spatial models may suggest that the intervention estimate was robust to the spatial effects. Alternatively, it may be that the spatial models were just as susceptible to bias as the non-spatial models. Chapter 7 suggests that if the magnitude of the spatial effect is large then the results of the CRT will have been biased. Considering that positive spillover was present, this would suggest that ITNs are actually more effective than previously thought. However, if the magnitude of the spillover was small, which it appears to be, then it likely had little impact on the intervention estimates.

The novel 'cluster reallocation' method was used for the first time and proved well suited to the bed net trial. Cluster reallocation was also suggestive of a spatial effect and the method will be discussed further in section 9.2.5.

These analyses add to the growing literature on both the effects of ITNs on mosquitoes and the spatial analysis of CRTs [6, 7].

Working with real CRT data suggested that the analysis of CRTs can benefit from a spatial approach, however benefits derived as much from exploratory spatial analysis

and visualisation as they did from spatial statistical modelling approaches. In fact, the practical application of spatial models revealed uncertainty as to the additional benefit of a spatial model over a standard approach to analysing CRTs. A greater understanding of the situations in which spatial spillover might bias results obtained by conventional approaches to CRTs was needed and became the focus of objective 4.

9.2.4. Objective 4: Evaluate the impact of spatial effects and the utility of spatial models in the analysis of a CRT by means of a simulation approach.

A simulation approach was used to establish whether spatial effects are an issue for CRTs and whether spatial models can help resolve bias brought about by unobserved spatial processes. As the spatial simulation of CRTs is novel, as far as I am aware, and approaches to simulating spatial correlation in general are not well known, this objective is split between two chapters. Chapter 6, introduces the methods for simulating a spatial CRT and provides a simple exemplar, while chapter 7 simulates a realistically complex CRT and focusses on evaluating the implications of differing spillovers.

In chapter 7, several algorithms were described to simulate spatial effects. Following the outcomes of the literature review (objective 2), approaches focused on both spatial variable and spatial modelling methods. A range of possible spatial variable metrics were considered, however when creating spatial spillover effects, it was evident that the parameterisation of a distance weighting function was more important than the choice of function itself. The Gaussian weighting function has a straightforward interpretation and therefore was chosen for future simulations. When considering the simultaneous autoregressive (SAR) spatial modelling algorithm, increasing the level of spatial correlation resulted in larger simulated values.

This meant that for large values of spatial correlation the simulated values in the control and intervention arms increased, as opposed to the simulated spillover effect which only increased the values of the control arm. Despite this, it was considered important to include a spatial correlated scenario in the simulations. Furthermore, although a spillover effect is generally conceptualised as the proportion of an intervention effect, the presence of spillover from multiple participants could result in the spillover being larger than the intervention effect itself. Therefore, the SAR model may be a more realistic way of representing the spatial effects.

The realistically complex simulation study in chapter 7 showed that CRT results can be affected by the presence of spatial spillover or spatially correlated intervention effects. The simulated spatial effects resulted in biased estimates and reduced the coverage of the models. The spatial spillover biased the estimates towards the null, and the spatially correlated intervention effect biased the estimates away from the null. In contrast, spatially correlated errors did not impact on the bias or coverage of the standard one stage CRT model. This implies that when a spatial process is unrelated to the intervention it will not affect the results of a CRT, which is an encouraging finding for the validity of CRTs in general.

Contrary to expectations, the use of spatial models did not help alleviate issues of bias and coverage in the presence of spatial effects. The models were consistent with the non-spatial model, and did not appear to provide improvements. This has implications for the conclusion of the ITN trial, suggesting that ITNs are potentially more effective than previously thought, as discussed in the previous objective. It also suggests that extending the spatial variable approach may be a more worthwhile focus for including spatial effect in analyses of CRTs in comparison to spatial modelling. For this reason, a simple diagnostic analysis tool was developed in response to objective 5 that would be sensitive to spatial spillovers.

9.2.5. Objective 5: Develop methodology to assess the presence of spatial spillover in CRTs.

Cluster reallocation was proposed as a method to explore spatial spillover effects in CRTs. Spatial models can be complicated and are more burdensome to implement than established conventional CRT analysis methods, which may inhibit the use of spatial methods in CRTs. There was a strong motivation to provide a method which could be used to assess the presence spatial spillover, irrespective of the type of outcome or the analysis model. The cluster reallocation method does not require any more statistical knowledge, than that required to analyse a CRT.

The main output of cluster reallocation is a graphical display showing the change in intervention estimates for different trial arm definitions. The plots provide an indication of the presence and range of spatial spillover effects. The analyst can assess whether the intervention effect decreases or increases when the cluster boundaries are varied. Thus, this approach provides a simple way to explore spillover in a wide range of settings. The method lacks maturity, and will need to be developed further to explore how sensitive it is to a wider range of conditions. However, a proof of concept has been presented that could allow others to explore spatial effects in CRTs, without requiring detailed knowledge of spatial statistics.

9.3. Strengths

In this section, I will discuss the strengths of the thesis.

9.3.1. Multi-disciplinary

A major strength of this thesis is that it draws upon several fields of research. The two overarching disciplines are CRTs and spatial statistics, but the thesis utilises knowledge from a range of other areas. Causal inference and economics provided

useful insights for how to think about and measure spillover effects [8–10]. While, machine learning and probabilistic graphical modeling literature was helpful in understanding Gaussian process models and Gaussian Markov random fields [11–13]. Vaccine trials provided a rich resource for considering indirect effects of interventions [14] and Bayesian methodology was necessary for fitting many of the spatial models [15, 16].

Making sense of literature that stemmed from diverse research areas was not without challenge. One of the main challenges was understanding the differing terminologies applied to the same or similar phenomena, see for example, the many related terms for spillover or contamination in chapter 2. In this thesis, I have tried to provide links between areas which at first may seem disparate, but without which it would not have been feasible to create this thesis. Several further strengths stem directly from this multidisciplinary approach

One strength of drawing from multiple disciplines is that methods can be applied in a novel setting. To the extent of my knowledge, this thesis applies several spatial methods to the CRT context for the first time. In chapter 5, GWR is used to explore the mechanisms of interventions effects and helped to reinforce the existence of a spatial pattern in the CRT. The join count statistic was also likely applied to a CRT for the first time to a CRT. This demonstrates that multiple spatial methods can be applied directly or adapted to a CRT setting, and that they complement the approaches of spatial variables and spatial models found in the systematic review. When discussing objective 3, I have made it clear that exploratory and visualisation based spatial approaches have been as instructive as formal models in understanding spatial spillovers in CRTs.

Although the collection of data is becoming cheaper in general, the cost of running trials is still considerable. A further strength of integrating methods from different disciplines into a new setting is in deriving additional value from existing datasets. This thesis is a good example of the utility of reanalysing secondary CRT datasets.

The trial in chapter 5 dates to 1993, and yet 25 years on provided a rich source of data for the application of many modern spatial methods. Several of the methods applied did not exist or were not widely available when the data was collected. In this case, applying modern spatial methods to existing data helped to deepen the understanding of the intervention effect. Furthermore, secondary analyses can help to inform future CRTs, and improve our understanding of spatial spillover without the need for further data collection.

9.3.2. Functional methodology

The spatial methods presented in this thesis require specialist knowledge and skills. In my opinion, for any trial methodology to be well used in practice it needs to be readily available, easy to understand, straightforward to implement, and not require a large amount of additional learning. Therefore, effort was made to develop a method that relates to the conventional and well understood approaches used in contemporary CRTs.

The cluster reallocation method does not require any statistical knowledge beyond that needed for a standard CRT analysis. In addition, the method is similar in ethos to a permutation test, which are sometimes used to analyse CRTs. This increases the chances of it being used and understood by statisticians who analyse CRTs. Furthermore, the spatial component for cluster reallocation is the expanding or contracting (dilation and erosion) of the cluster boundaries. It is hoped that this provides an intuitive approach that can be understood when presented to individuals without a formal statistical background. This ready interpretability could be beneficial in policy and planning contexts.

Unless trials are designed to test for spillover, any spatial spillover analysis will be a non-randomised comparison and subject to bias. To reflect the uncertainty about the presence of spillover, an exploratory visual method, rather than a hypothesis testing framework was used. It may be possible to test for spillover more formally

using a similar approach to normality testing or permutation tests. However, a visual method was chosen as it avoids a strict decision about whether spillover is present or absent. Furthermore, displaying a figure rather than a p-value may provide some transparency, making it easier for the results of a spillover analysis to be debated. This does not preclude the later development of a hypothesis testing approach that seeks to examine the null that no spatial spillover is present.

In terms of availability, work is planned to create an R package.

9.3.3. Reproducibility

A scientific finding is strengthened when it has been replicated by multiple independent researchers [17]. In medical research, full replication of a study is usually inhibited due to time and cost. In view of this, Peng et al. propose an attainable minimum standard, whereby the finding should be computationally reproducible [17]. Computational reproducibility refers to the ability to reproduce the findings of a study, based on the original data and/or the code used for an analysis. The gold standard is to make the data and code of the analysis available and to use open source software for the analyses [18]. This allows other researchers to check the veracity of the findings by rerunning the code, and independently reanalyse the data. Unfortunately, confidentiality may prohibit data sharing and a possible solution is the use of replication studies; where a team of independent researchers are provided with access to the data and code [19].

In this thesis, all analyses were conducted in R, which is free and open source. This means that, at least in theory, anyone can access the relevant software and libraries. Furthermore, the majority of the code written for the thesis is stored on private Github repositories, which can be made publicly available. Due to restrictions with collaborators, the code from chapter OPV was not stored on GitHub.

Steps have been taken to improve the computational reproducibility and veracity. Parts of the code from chapter 5 have been run independently by researchers in

Switzerland, and the results corroborated. In the absence of a further team of researchers, I have rerun the code on several different computers, and obtained consistent results. The process for this was: perform a fresh install of R; pull the relevant Github repository; load the relevant data to the computer; and then run a master file which calls all the relevant scripts. I have conducted this process multiple times for each analyses. For the simulation studies, a smaller number of iterations were performed to reduce unnecessary computation time.

It is intended that the simulation study chapters will be combined into a paper, and any data and code for these simulation studies will be made publicly available via Github. Due to the confidential nature of observation locations in the two applied papers, it is unlikely the data will be made publicly available. Where possible, I will attempt to make the code available, so that those who wish, can check the process used to reach the conclusions of the papers.

9.3.4. Analysis of real world data

Real data are messy, often containing dependencies, missing values, and structures that differ from those taught in traditional statistics classes [20]. Without the benefit of analysing real-world data, it is plausible to develop methods which only work well under a restricted set of conditions, that rarely occur in practice. The use of real data is invaluable in informing methodology development and helps to create methods which are useful and relevant for practitioners.

The methodology developed in this thesis stems from the analysis of real-world data. I was fortunate enough to have access to data from two real CRTs with spatial information. The trials had different aims, were of different types, and were set in different contexts. The analyses helped to inform the direction of the thesis and forced consideration of how well a single spatial method can be generalised to all CRTs. For instance, the spatial variable approach was of limited utility in the OPV study in chapter 4, but was useful in ITN trial in chapter 5. In this example

the usefulness was a function of the context of the CRTs, as opposed to differences in design. Access to real data also allowed for the cluster reallocation method to be tested outside of a simulated environment, thus demonstrating that the method could be applied in practice.

9.4. Limitations

Specific limitations of this thesis have been described at the end of each chapter. Here, limitations that affect the overall conclusions of the thesis will be discussed, and attempts have been made to minimise overlap.

9.4.1. Focus on conceptually simple CRTs

The CRTs in this thesis are typically parallel two-arm CRTs. This ignores the added complexity that can be involved in CRT design. I do not address the impact of multiple arms, multiple time points, or innovative designs such as stepped wedge trials. This stems from the PhD having a greater focus on spatial methods, rather than CRTs.

The greater focus on spatial methods was intentional, as it was perceived it would be more likely that a statistician on a CRT may want or need to conduct a spatial analysis, rather than a spatial statistician deciding to conduct a CRT. Despite this, it does mean that the impression of CRTs may be simplistic at times and may not easily generalise to more complex CRT design.

When describing how CRTs are analysed, considerable focus was placed on the one-stage method. The two-stage method was not explored, and the use of permutation tests ignored. The one-stage method was chosen as it had a clear spatial analogue. Furthermore, the two-stage method is more often used when the number of clusters is small; where it may be hard to detect spatial spillover, and difficult to fit spatial models. Additionally, permutation tests are computationally intensive, and when

combined with the complexity of the spatial models may have led to an unacceptable computational burden. For the reasons given, these two approaches were not considered, but they may be utilised in future work.

A further reason for the not considering more complex CRTs designs, is that given a lack of maturity of spatial analysis in CRTs, it was deemed more appropriate to start with the archetypal parallel CRT. Otherwise methods may have been applied in a multi-arm trial with multiple end points, before establishing whether they are appropriate in the simpler case. Therefore, spatial methods take priority in the thesis and consideration of further CRT designs, and methods combined with spatial analyses is left for future work.

9.4.2. Lack of temporal methods

CRT data are not collected instantaneously, and trials may measure multiple observations per participant at different time points. Additionally, the outcomes may incorporate time, such as in survival analysis where the outcome is typically time to an event. In these situations, the incorporation of time into the analysis through temporal methods may be important and ignoring time could lead to missing notable features in the data.

In this thesis, there was little consideration of the impact of multiple time points, and greater focus was given to purely spatial methods. This is mainly due to the immaturity of the use of spatial methods in CRTs. Temporal methods were used in chapter 4, where an autoregressive lag structure was included in the model to account for multiple time points. This approach could be used to incorporate time into the spatial methods used. Alternatively, an independent random effect could be included to represent time, which would make observations from the same time points related, but observations between timepoints unrelated. However, it could be argued that this form of analysis is a spatial and temporal approach, rather than a spatio-temporal method which allows space and time to interact.

The spatial aspects of a trial may also interact with the temporal aspects. For example, the introduction of ITN may affect the spatial distribution of mosquitoes initially, but over time as the insecticide wears off, the spatial distribution could change. In this instance, spatio-temporal methods would be required. Spatio-temporal methods have long been neglected, but their use is increasing, particularly in environmental epidemiology [21, 22]. The temporal approach used in chapter 4 treats time and space as separate parts of a model and does not allow for them to interact with each other. It would have been exciting to explore spatio-temporal methods within CRTs, but unfortunately, it was beyond the scope of the thesis. Therefore, the methods in this thesis may not generalise well to CRTs with multiple time points, or with spatio-temporal processes.

9.4.3. Simulating realistic spatial effects

A further limitation of the thesis relates to the difficulty of simulating realistic spatial effects. It is difficult to tell whether the spatial effects simulated in chapter 6 and 7 are representative of real spatial effects in CRTs. Due to a small amount of relevant data on spatial spillover, and spatial correlation in CRTs, it was difficult to establish plausible ranges and magnitudes for the effects. Attempts were made to simulate spillover effects similar to the ranges presented in ITN CRT literature, but those these ranged from a few hundred meters to distances of several kilometers [23, 24]. Presented with a lack of information, I decided to simulate a wide range of spatial effects. This means that the simulation study should include some scenarios which reflect real world data. However, the main drawback of this lack of knowledge is that the scenarios where spatial effects did impact of the results of CRT, may be unrealistic.

Conducting further spatial analyses of CRT will help to provide richer information on plausible ranges and magnitudes of spatial spillover. This can help to inform future simulation studies. An alternative approach could be to use mathematical

modelling to simulate spatial effects in the same way a disease outbreak is simulated. This may allow for more realistic scenarios to be created as there is a large literature on the dynamics of infectious diseases outbreaks .

9.5. Implications

9.5.1. Design

This thesis reinforces that careful thought and planning is required for designing CRTs. Spatial spillover effects may not be relevant in many contexts, but when they are, consideration of the type of spillover, and whether it is of interest is needed. The consideration of spatial effect is of particular importance when they stem from the intervention.

It is more likely that the trial design itself will motivate the types of spatial measures that are possible, rather than the desired spatial measure motivating the types of trial. This could result in a lack of meaningful spatial measurements. For example, a household CRT may preclude distinguishing between the location of the observations and the clusters. In the analysis of the OPV study, transmission dynamic were not patterned according to household distribution, and a different design approach may have helped to detect spatial effects.

The simulation studies showed that spatially correlated errors are unlikely to affect the intervention effect of a trial. This is reassuring in the sense that it confirms that randomised comparisons of a CRT are unbiased. This implies that only spatial effects that interact with the intervention need be considered during the design of a trial. Moreover, underlying spatial processes can be largely ignored unless they interact with the intervention.

Spatial analyses depend on the accuracy of the spatial data. Unfortunately, I was unable to make use of an additional dataset received for this PhD due to inaccura-

cies with the GPS recordings. There were observations that were not in the correct country, and points that were located far away from the study area. Whilst methods exist for interpolating missing attribute values at known spatial locations (e.g. kriging), I know of no methods that allow for the imputation of missing spatial location when attributes are known. Therefore, if the spatial data is collected during a CRT, then it should be checked early in the study, and individuals collecting it should be provided with training.

9.5.2. Ethics

When conducting spatial analyses of existing CRT data, the original purpose of the spatial data needs to be considered. If the data were collected for trial management, then participants may not have originally consented for it to be included in an analysis. It is not possible to inform individuals in a trial about all the possible uses of their data, particularly as methods which can now be applied may not have been developed at the time of the study. If consent was given for the data to be included in a map for publication, then this is potentially adequate for further analyses, as spatial visualisation could be considered as a form of exploratory spatial analysis.

When presenting spatial data, their granularity and identifiability should be assessed. In chapter 4, the maps were initially created using a background which identified roads and other features. It was determined during the analyses that this made the households relatively easy to identify, especially due to the small number of houses. It can be easy to disassociate data visualisations from the real life context of an analysis, and present maps of potentially identifiable and confidential data. However, just because data can be mapped does not mean it should be displayed. It may be useful for the researcher to consider whether they would be comfortable presenting the map if it was identifying their own location and medical history.

The use of tracking devices in future CRTs requires ethical considerations of the scale used to record and present spatial data. This scale will relate to the identi-

fiability of the study area. For instance, in rural settings a distance of 1km may enable identification of an individual, whereas in an urban setting 1km may provide anonymity. Furthermore, the time scale needs to be considered, a smaller time scale provides more reliable data, but also increases the identifiability of participants. A potential approach is to ask for consent to collect granular data which is useful for analysis, but ensure that only aggregated spatial data is presented.

One possible negative effect of collecting such detailed spatial data, is that it could deter some individuals from participating in the study. The knowledge that their location will be kept and analysed in the future may be unnerving. Therefore, a possible implication of incorporating spatial analyses in CRTs, and recording spatial data may be that fewer participants want to enrol in studies.

9.6. Future work

There are many ways in which the work presented in this thesis could be extended to provide further understanding of spatial analyses and spatial effects in CRTs.

9.6.1. Extend the cluster reallocation method

The cluster reallocation method provides an easy-to-understand approach to exploring spatial spillover in geographical CRTs. There are several aspects that could be explored further with the method.

Instead of spatial proximity based on cluster boundaries, the distance between observations could be used for reallocation. A network could be formed using the nearest neighbours of the participants. This would allow reallocation to be based on the status of the nearest neighbour. For instance, a control observation that is nearest to an intervention observation would be reallocated to intervention. Following this, a control observation who is a second neighbour with an intervention

observation would be reallocated. This approach would require extending or changing the code used for the method. This may help the method to be applicable beyond CRTs where the clusters are polygons and the observations are points. This approach would allow different measures of connectivity such as a social proximity, thus further generalising the method.

The output of cluster reallocation could also be extended. At present a single cluster reallocation plot is presented that represents the change in the effect estimate. There may be scope to present summary statistics alongside the visualisation. For instance, the average, or maximum estimate from increasing the intervention or control boundaries could be presented alongside the original cluster estimates. These values would complement the visual approach.

The models used calculated a relative effect size, and this was appropriate when assessing for positive spatial spillover. When assessing for negative spatial spillover, the cluster reallocation plots will likely look consistent with those of no spillover. Using an absolute effect estimate or presenting the absolute values in each trial arm is a possible way to assess for the effect of negative spatial spillover. Thus, different output metrics could be considered.

The use of the method will rely on accessible software. Therefore, future work will include developing an R package for the cluster reallocation method. Initially an attempt will be made to develop a function that can replicate the way the method has been used in the thesis. Following this, the extensions discussed will be developed and tested within the package.

9.6.2. Further simulations

The simulation study in this thesis has demonstrated that spatial effects can be a problem in CRTs. Building on this, there are several spatial and CRT characteristics that were fixed during the simulations that could be explored further.

The effect of different spatial structures could be explored by simulating non-random patterns. The study in the thesis simulated assuming spatial randomness. As discussed previously, this assumption is unlikely to reflect true neighbourhood structures, which are often clustered. Clustered patterns could be simulated, with few or many observations near the borders; road networks could be used for simulating clustered patterns; or real CRT data could be used and simulated values attached to the actual household locations. These approaches may increase the realism of the simulated spatial data.

The spatial structure of the observations was also kept fixed for all the simulations. In addition to using different types of structures, new point locations could be created for each iteration of the simulation study. This would remove the dependency of the results on any particular type of structure. However, these approaches may add considerably to the computational burden for the spatially correlated approach as a new spatial weight matrix would need to be calculated for each new structure. Changing the number of observations and the number of clusters could also be investigated further. Presumably, spatial spillover is more likely when the number of observations is high compared to the number of clusters. The effect of changing the number of observations for a fixed cluster size could be investigated to test this. Alternatively, a fixed sample size could be used, and the number of clusters could be varied. This might also help to determine whether there is a required number of clusters before spatial effects impact results and whether the spatial methods become unfeasible to use for large numbers of clusters and observations.

In combination with the quantity of the cluster and observations, the spatial extents could be varied. Simulations could be conducted where study areas are changed in size, or a minimum distance is established between observations. Beyond being interesting in itself, this type of simulation may allow for an estimation of an optimal study size, or number of clusters for a given area in a given context. When information is available about a spatial spillover effect, then a range of aggregations

(cluster definitions) could be simulated over the proposed study area. This could be combined with sample size calculations to determine the most effective number of clusters needed to reduce (or measure) spatial spillover, whilst also estimating the main effect of the trial.

A further aspect that was not explored was the impact of the intra-cluster correlation coefficient (ICC). I hypothesise that decreasing the ICC will reduce the impact of spatial spillover as the power of the trial will increase, and increasing the ICC will increase the impact of spatial effects as power is reduced. Alternatively, a greater ICC may result in more variability and thus magnitudes of the spatial effects may be smaller compared to the magnitudes of the simulated values.

The magnitude of the intervention is another factor that could be varied. A value of 2.0 was chosen in the study in chapter 7. Intuitively, it is likely that a smaller intervention effect will be more affected, and a larger intervention effect will be less affected. However, if the spatial effect is based on the magnitude of the intervention, then perhaps the impact is independent of the scale. This aspect is probably of less interest compared to the other study characteristics, but it could be varied within the other studies rather than being a specific study by itself.

9.6.3. Spatial models and spillover variables

The simulation study suggested that spatial models were also subject to bias. Although only two types of spatial models were assessed, they provide the foundation for a wide range of other spatial models. Spatial filtering and the use of copulas are approaches that are different from the considered methods and may be relevant[25, 26]. Testing a broader range of spatial models, would help to extend the generalisability of the results from the thesis.

In contrast to the spatial modelling approach, an alternative and perhaps more fruitful area of future research could be to develop spatial spillover measurements. The spatial variable approaches found in the literature made specific assumptions

about the spatial mechanism. Furthermore, they were context dependent. One possible solution is to develop a measure of proximity that makes fewer assumptions about the mechanism of the spillover effect. If such a general way of measuring proximity could be determined, that takes into account the proximity, density, and characteristics of nearby individuals, then it could be used in a variety of settings.

9.6.4. Application to further trials

In this thesis, spatial methods were applied to two existent trials. The systematic review also found few examples of spatial methods used in CRTs. Given the utility of spatial methods for secondary analysis, an obvious area of further research is to analyse more existing CRTs. A clear choice would be to conduct analyses on more ITN trials. A further area that received little attention in the PhD is primary care trials.

The systematic review in chapter 3 found several CRTs related to mosquitoes where spatial analyses were performed. Similar analyses to chapter 5 could be repeated to help provide a fuller picture of how often spatial effects are present in practice. Assessing the magnitude of spatial spillover in a range of trials could also be used to determine plausible ranges of spillover, providing context for the results from the simulation study in the thesis.

It would have been interesting to explore the impact of spatial dependence in primary care trials as per Silcocks and Kendrick [1]. This may have been an area where the spatial dependency models are more applicable. An additional interesting aspect of this trial design is that the space may be dynamic. If the definition of the primary care areas change during the course of the trial, then this will raise further implication for the analyses, such as how membership to a cluster is affected.

These two types of trials may make use of different types of spatial methods, furthering the range of spatial methods that are applicable to CRTs. In addition, reanalysis of existing CRT from a spatial perspective provide an exciting opportunity to gain

extra utility from already collected data. Applying spatial method to a broader range of CRT contexts could lead to further epidemiological and methodological understanding and development.

In this thesis, spatial methods were applied to two real trials. It would be interesting to apply spatial methods to a number of other CRTs, where spatial information is available. An obvious choice of CRT for further analyses would be ITN trials. An area of interest which received little attention in the PhD is primary care trials.

9.7. Concluding remarks

This PhD set out to explore the use of spatial methods within the analysis of CRTs, and provides one, if not the first, extensive consideration of the subject. The collective findings highlight that the use of spatial methods in CRTs is rare, and the presence of spatial effects can impact CRT results. However, the standard CRT one-stage approach is typically robust to small scale spillover effects. Consideration of the spatial dependence of observations appears to provide little extra utility in the main analysis of a trial. Despite this, spatial methods help provide additional insights into the mechanism of interventions, and are well suited to secondary analyses of CRTs, especially with the increasing collection of GPS data in CRTs.

9.8. Bibliography

- [1] Silcocks, P and Kendrick, D. Spatial effects should be allowed for in primary care and other community-based cluster RCTS. *Trials* Jan. 2010 11. (1) 55.
- [2] Chao, DL, Park, JK, Marks, F, Ochiai, RL, Longini, IM, and Halloran, ME. The contribution of neighbours to an individual's risk of typhoid outcome. *Epidemiology and infection* 2015 143. (16) 3520–7.
- [3] Global Polio Eradication Initiative and WHO. POLIO NOW. 2017. (Visited on 01/01/2017).
- [4] Morales, M, Tangermann, RH, and Wassilak, SG. Progress Toward Polio Eradication - Worldwide, 2015-2016. *MMWR. Morbidity and Mortality Weekly Report* May 2016 65. (18) 470–473.
- [5] Morales, M and Rudolf H. Wassilak, SGT. Polio Eradication & Endgame Strategic Plan 2013-2018. *Global Polio Eradication Initiative* 2013 134.
- [6] Lengeler, C. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database of Systematic Reviews* 2004. (2). arXiv: 14651858.CD000363.pub2. [Art. No.: CD000363. DOI: 10.1002].
- [7] Jarvis, C, Di Tanna, GL, Lewis, D, Alexander, N, and Edmunds, WJ. Spatial analysis of cluster randomised trials: A systematic review of analysis methods. *Emerging Themes in Epidemiology* 2017 14. (12).
- [8] Baylis, K and Ham, A. How important is spatial correlation in randomized controlled trials? 2015.
- [9] Benjamin-Chung, J, Arnold, BF, Berger, D, Luby, SP, Miguel, E, Colford Jr, JM, and Hubbard, AE. Spillover effects in epidemiology: parameters, study designs and methodological considerations. *International Journal of Epidemiology* 2017. (March) 332–347.

- [10] Benjamin-Chung, J, Abedin, J, Berger, D, Clark, A, Jimenez, V, Konagaya, E, Tran, D, Arnold, BF, Hubbard, AE, Luby, SP, Miguel, E, and Colford, JM. Spillover effects on health outcomes in low-and middle-income countries: A systematic review. *International Journal of Epidemiology* 2017 46. (4) 1251–1276.
- [11] Bishop, CM. Pattern Recognition and Machine Learning. Information Science and Statistics. Springer, 2006.
- [12] Koller, D and Friedman, N. Probabilistic Graphical Models: Principles and Techniques. Adaptive computation and machine learning. MIT Press, 2009.
- [13] Zheng, S, Jayasumana, S, Romera-Paredes, B, Vineet, V, Su, Z, Du, D, Huang, C, and Torr, PHS. Conditional Random Fields as Recurrent Neural Networks. 2015.
- [14] Halloran, ME, Longini, IM, and Struchiner, CJ. Design and Analysis of Vaccine Studies. Statistics for Biology and Health. Springer New York, 2009.
- [15] Gelman, A, Carlin, JB, Stern, HS, Dunson, DB, Vehtari, A, and Rubin, DB. Bayesian Data Analysis, Third Edition. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, 2013.
- [16] Geyer, CJ. Introduction to Markov Chain Monte Carlo. Ed. by Brooks, S, Gelman, A, Jones, G, and Meng, XL. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, 2011. Chap. 1.
- [17] Peng, RD, Dominici, F, and Zeger, SL. Reproducible epidemiologic research. *American Journal of Epidemiology* 2006 163. (9) 783–789.
- [18] Peng, RD. Reproducible research in computational science. *Science* 2011 334. (6060) 1226–1227. arXiv: 0901.4552.
- [19] Aiken, AM, Davey, C, Hargreaves, JR, and Hayes, RJ. Re-analysis of health and educational impacts of a school-based deworming programme in western

- Kenya: a pure replication. *International Journal of Epidemiology* 2015 44. (5) 1572–1580.
- [20] Legendre, P. Real data are messy. *Statistics and Computing* 1993 3. (4) 197–199.
- [21] Cressie, N and Wikle, CK. *Statistics for Spatio-Temporal Data*. Wiley, 2011.
- [22] Shaddick, G and Zidek, JV. *Spatio-Temporal Methods in Environmental Epidemiology*. CRC Texts in Statistical Science. Chapman Hall/CRC, June 2015. URL: <http://opus.bath.ac.uk/47731/>.
- [23] Lenhart, A, Orelus, N, Maskill, R, Alexander, N, Streit, T, and McCall, PJ. Insecticide-treated bednets to control dengue vectors: Preliminary evidence from a controlled trial in Haiti. *Tropical Medicine and International Health* 2008 13. (1) 56–67.
- [24] Howard, SC, Omumbo, J, Nevill, C, Some, ES, Donnelly, CA, and Snow, RW. Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2000 94. (4) 357–360.
- [25] Griffith, D. A linear regression solution to the spatial autocorrelation problem. *Journal of Geographical Systems* 2000 2 141–156.
- [26] Musgrove, DR, Hughes, J, and Eberly, LE. Hierarchical copula regression models for areal data. *Spatial Statistics* 2016 17 38–49.

A. LSHTM ethics approval

London School of Hygiene & Tropical Medicine

Keppel Street, London WC1E 7HT
 United Kingdom
 Switchboard: +44 (0)20 7636 8636

www.lshtm.ac.uk

LONDON
 SCHOOL of
 HYGIENE
 & TROPICAL
 MEDICINE



Observational / Interventions Research Ethics Committee

Mr Christopher Jarvis
 LSHTM

4 April 2017

Dear Christopher,

Study Title: Spatial Analysis of CRTs PhD

LSHTM Ethics Ref: 12260

Thank you for responding to the Observational Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

Approval is dependent on local ethical approval having been received, where relevant.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document Type	File Name	Date	Version
Local Approval	Binka et al. - 1996 - Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana a randomized contro	02/04/1996	1
Local Approval	Main LSHTM approval	28/05/2012	1
Investigator CV	Christopher_Jarvis_CV	31/01/2017	1
Protocol / Proposal	Spatial analyses of cluster randomised trials_ethics-protocol	31/01/2017	1
Covering Letter	Jarvis-12260-Response	29/03/2017	1
Local Approval	Binka, Indome, Smith - 1998 - Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Gha	29/03/2017	1
Local Approval	Re Spatial analysis Bednet data	29/03/2017	1
Local Approval	LaVIISWA protocol, binder	29/03/2017	1
Local Approval	FW data request form - Chris Jarvis	29/03/2017	1
Local Approval	Transfer of Data	29/03/2017	1
Local Approval	Re Permission to use OPV Study Data	29/03/2017	1
Local	Stanford IRB Application	29/03/2017	1

LSHTM ethics approval

Approval		
Local Approval	Stanford IRB Approval Letter	29/03/2017 1
Local Approval	forme1application_LaVIISWA	29/03/2017 1

After ethical review

The Chief Investigator (CI) or delegate is responsible for informing the ethics committee of any subsequent changes to the application. These must be submitted to the Committee for review using an Amendment form. Amendments must not be initiated before receipt of written favourable opinion from the committee.

The CI or delegate is also required to notify the ethics committee of any protocol violations and/or Suspected Unexpected Serious Adverse Reactions (SUSARs) which occur during the project by submitting a Serious Adverse Event form.

At the end of the study, the CI or delegate must notify the committee using an End of Study form.

All aforementioned forms are available on the ethics online applications website and can only be submitted to the committee via the website at: <http://leo.lshtm.ac.uk>

Additional information is available at: www.lshtm.ac.uk/ethics

Yours sincerely



Professor John DH Porter
Chair

ethics@lshtm.ac.uk
<http://www.lshtm.ac.uk/ethics/>

Improving health worldwide

B. License agreement for paper A

Christopher Jarvis

From: Murugappan, Magesh, Springer <magesh.murugappan@springer.com>
Sent: 13 March 2018 15:15
To: Christopher Jarvis
Subject: Permission to use research paper in PhD Thesis

Dear Dr. Jarvis,

Thank you for your email.

All authors of papers published by BioMed Central and SpringerOpen retain copyright of their manuscripts and can therefore use any part of it again as long as the original article is properly cited.
Our full copyright policy can be found here: <http://www.biomedcentral.com/getpublished/copyright-and-license> or <http://www.springeropen.com/get-published/copyright>

Please let us know if you have any further questions.

With best regards,
Magesh

Magesh Murugappan (Mr.)
Journals Editorial Office (JEO)

From: Christopher Jarvis [mailto:Christopher.Jarvis@lshtm.ac.uk]
Sent: Friday, March 09, 2018 10:59 PM
To: Murugappan, Magesh, Springer
Subject: Permission to use research paper in PhD Thesis

Dear Magesh,

I would like to request permission to include the following open access article that I published and authored with you: Jarvis C, Di Tanna GL, Lewis D, Alexander N, Edmunds JW. Spatial analysis of cluster randomised trials: A systematic review of analysis methods. Emerging themes in epidemiology. 2017;14.
in the digital copy of my thesis which will be made publically available through LSHTM Research Online <http://researchonline.lshtm.ac.uk/> . This is our institutional repository and is non commercial and openly available to all.

Yours sincerely
Christopher Jarvis

This e-mail is confidential and should not be used by anyone who is not the original intended recipient. If you have received this e-mail in error please inform the sender and delete it from your mailbox or any other storage mechanism. Scientific Publishing Services Private Limited does not accept liability for any statements made which are clearly the sender's own and not expressly made on behalf of Scientific Publishing Services Private Limited or one of their agents.
Please note that Scientific Publishing Services Private Limited and their agents and affiliates do not accept any responsibility for viruses or malware that may be contained in this e-mail or its attachments and it is your responsibility to scan the e-mail and attachments (if any).
Scientific Publishing Services Private Limited. Registered office: No. 6 & 7, 5th Street, R.K.Salai, Mylapore, Chennai, 600004, India.

C. Spatial modelling with random fields

C.1. Introduction

The conditional autoregressive (CAR) and simultaneous autoregressive (SAR) models are two of the most commonly used spatial models in the analysis of lattice or areal data (discrete spatial processes) [1]. The SAR model was introduced by Whittle in 1954 [2] and the CAR model introduced by Besag [3] some 20 years later. Due to the mathematical background of the authors and intended audience, the relation between Markov random fields (MRFs) and these models was clear at the time (See Comments [3]). Since then the use of CAR and SAR models has expanded beyond the Mathematical and Statistical communities into fields such as Geography, Spatial Econometrics, Epidemiology, and many others [4–6]. This has led to a blurring of the connection between MRFs and the SAR and CAR models.

CAR models were originally presented as a way to directly model spatial dependence in observed data [7]. When Besag presented the CAR models the unifying framework of generalised linear models (GLMs) were still in their infancy [8]. Since then spatial modelling has been incorporated into the GLM framework by use of generalised linear mixed models (GLMM) [9, 10]. This fusion allows spatial processes to be modelled separately from the outcome of interest [6, 10]. For instance, a binary outcome can be modelled with a binomial distribution, and the spatial process can be modelled as continuous assuming a Gaussian distribution. This subtle difference in model formulation provides great flexibility but leads to further difficulty by what is actually meant by a CAR model.

This chapter considers SAR and CAR models in the context of random fields, highlighting the connection between the spatial methods and the Statistical terminology. Starting with the origins of the SAR and CAR models to give context for their place in the modern approach of hierarchical models or GLMMs. I will outline what random fields are and how they relate to random variables. This provides a foundation from which to discuss spatial statistics from a more general platform of processes and fields.

The methods used for fitting the models is discussed, with particular focus on Integrated nested Laplace approximation (INLA) an approximate Bayesian fitting method. Finally, an exciting link between the geostatistical and areal data analysis methods is discussed.

C.2. Origins of SAR and CAR

The SAR and CAR models were developed as spatial extensions to time series methods [1, 7]. The conditional and joint specification of distributions are equivalent in time series analysis, because time has a specific direction from past to future. In space, the joint and conditional distributions are not equivalent, as space does not have a fixed direction.

The SAR model is the spatial analogue of the joint distribution approach of time series [1]. Introduced by Whittle [2], the model is equivalent to solving a series of simultaneous equations relating the outcome variable to observations that are nearby. Originally presented as a class of stationary processes on a plane, the SAR model has become ubiquitous in spatial econometrics but is used less frequently in Spatial statistics as it struggles to be extended to beyond Gaussian lattice data [11].

The CAR model is the spatial analogue of the conditional distribution approach of time series [1]. Introduced by Besag [3], some 20 years after Whittle's paper, the conditional approach was considered controversial at the time (See comments in [3]). The attractiveness of this approach is that specifying the conditional distribution is much easier than specifying the joint distribution. However this would not be sufficient without Besag's Lemma [11] [Often referred to as Brooks Lemma] which demonstrates that you can obtain the joint distribution from a carefully defined set of conditional distributions. The Intrinsic CAR (ICAR) model is a special case of the CAR model where the precision matrix is not of full rank [12].

Random fields models provide a framework for spatial modelling that includes the CAR and SAR approach as well as many of the modern-day approaches to spatial modelling [13]. To highlight the connection, we first need to define what is meant by a random field.

C.3. Random variables, processes and fields

The relation between random variables, stochastic processes, and Gaussian Markov random fields is presented visually in Figure C.1.

A stochastic process (SP) is a collection of random variables [14]. Typically, they are indexed by time, where each random variable relates to a realisation of an experiment at a specific time point [15]. A stochastic process is also called a random process, random function, and a stochastic function. An example of a stochastic process is the maximum temperature each day, the maximum temperature for a given day is a random variable, and the index is the day.

Further names are used to describe special types of stochastic process based on the properties of the random variables, two types we will describe are a Markov process and a Gaussian process (GP) . A Markov process is a stochastic process where the future values in the process depends only on the present value (Markov

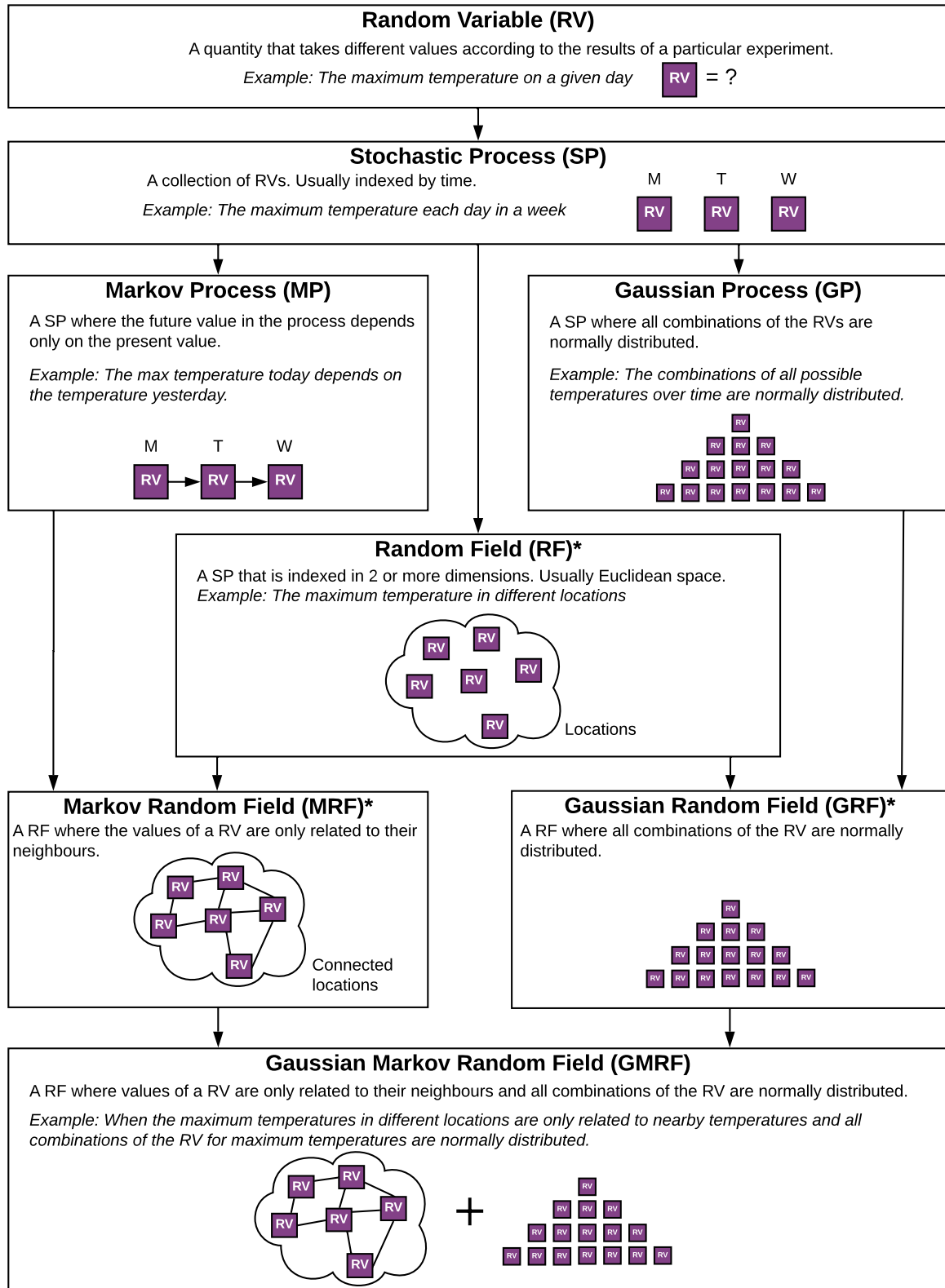
property) [16]. A GP is a stochastic process where all combinations of the random variables follow the multivariate normal distribution (MVN) [17, 18]. GPs are often defined as a Gaussian distribution with a mean and covariance function, as opposed to the constant mean and variance of a standard Gaussian distribution [13, 18]. Continuing our example if the maximum temperature today is only related to the maximum temperature yesterday then it would be a Markov Process. If all the possible combinations of the maximum temperatures in a stochastic process follow the MVN then it is a GP.

A stochastic process can also be indexed by space. When it is indexed by space in two or more dimensions it is also referred to as a random field [19]. Extending our example, a random field could be the maximum temperature for a given day in various locations. Similar to stochastic processes, a MRF is a random field with a Markov property (Markovian). A random field where all combinations of the random variables follow the MVN is a Gaussian random field [6]. In contrast to time, space does not typically have a direction from past to future and therefore the Markov property is extended to refer to neighbours, where random variables are related to their neighbours [3]. MRFs are also called undirected graphical models [20].

When a random field is Gaussian and Markovian then it is called a Gaussian Markov random field (GMRF) [19]. In our example a GMRF would be when the maximum temperature in a one location is related to locations that are nearby, and when all the combinations of the possible values that each location can take follow the MVN. When the precision matrices of a GMRF are not of full rank then they called Intrinsic GMRFs or improper GMRFs [19].

Although the term random field can be used to describe a stochastic process that is indexed in two or more dimensions, the distinction is not always used. A relevant example of this is the use of GPs to describe some spatial models [9]. In this context the term Gaussian field or GP could both be used but GP appears to be more prevalent [6, 13]. In this thesis GP will be used instead of Gaussian field.

Figure C.1. The connection between random variables and Gaussian Markov random fields



*Some literature refers to RFs as SPs. However explicitly stating RF makes clear that the SP is indexed in two or more dimensions. GP is often used to mean a of GF in spatial analysis of geostatistical data. MRF and GMRF is more commonly used for analysis of lattice data in spatial statistics.

C.4. Gaussian processes and Gaussian Markov fields

Geostatistical data can be modelled as GPs. When point data is collected, there will be locations between the points where no information is known. In this scenario a GP with a mean of zero, and a distance-based covariance structure can be used to model the data. For Areal data distance between polygon can be derived using centroids but this can be clumsy, and with irregular shaped polygons where the centroid may be outside the area of interest. A more natural way to model the data is to use adjacency between the areas of interest and this type of structure can be modelled using a MRF. For the continuous case the distribution with maximum entropy [21] is Gaussian [22] and therefore a GMRF is usually used. The GMRF is used when the spatial process is discrete, and the GP is used when the spatial process is a real valued and continuous. The CAR and SAR models are usually defined in the Gaussian case and are GMRFs and the ICAR is an intrinsic GMRF.

C.5. Non-Gaussian outcomes

The GP and GMRF approach can be used to model the outcome directly, however this would only naturally apply when the outcome is continuous. When Besag introduced his paper he presented further approaches of the auto-logistic, auto-Poisson approaches but noted that the auto-Poisson only works for negative spatial correlation and is therefore quite limited [3]. Cressie has extended this work to include auto-beta, auto-gamma and further classes of auto models [11] and the approach of extending auto models for further outcome is an area of active research [13].

An alternative approach is to use the methodology of GLM [8] and condition on the spatial process [9]. This allows the outcome to be modelled according to a distribution from the exponential family and the spatial process to be modelled in the linear predictor of the model. Here an outcome y_k is conditioned on a spatial process s_k where $k = \{1, \dots, K\}$ and represent the location where the observation was recorded for a binary outcome we can write the model as

$$\begin{aligned}y_k | s_k &\sim \text{Bin}(\pi, n) \\s_k &\sim \text{MVN}(\mu, \sigma^2 \Sigma)\end{aligned}$$

The spatial process is assumed to be MVN and is modelled separately from the outcome. Alternatively, we can write the model as

$$\begin{aligned}y_k &\sim \text{Bin}(\pi, n) \\ \log(\pi) &\sim \alpha + s_k \\ s_k &\sim \text{MVN}(\mu, \sigma^2 \Sigma)\end{aligned}$$

This approach is described as a Spatial GLM [13], but it is a GLMM as the spatial process is included as a random effect. More generally it is called a latent Gaussian model (LGM), as the unobserved (latent) process is assumed Gaussian.

This approach has allowed the auto-normal model to be applied to the error term of the linear predictor whilst modelling a non-normal outcome. This greatly extends the uses of the CAR model, and more specifically the Intrinsic CAR model. In a Bayesian setting the ICAR is no longer a model but is a type of prior distribution. Really it is a spatially interpretable form of the MVN. This approach allows the spatial effects to be defined a random effect, a famous example was presented by Besag, York, and Mollie [23] and is called the BYM model.

The BYM model includes a spatially structure random effect and a i.i.d random effect. For a count outcome the model can be written as

$$\begin{aligned}y_k &\sim \text{Pois}(\lambda_k) \\ \log(\lambda_k) &\sim \alpha + v_k + u_k\end{aligned}$$

Where the spatially structured random effect is $v_k \sim \text{MVN}(\mu, \sigma_v^2 \Sigma)$ and the i.i.d random effect is $u_k \sim \text{MVN}(\mu, \sigma_u^2)$. This model has been widely used in disease mapping [24, 25].

C.6. Model fitting

The SAR and CAR models can be fitted using maximum likelihood but the ICAR cannot. The SAR approach does not extend well to random effects and is therefore difficult in a hierarchical model setting [10]. The ICAR is difficult to fit with maximum likelihood methods and the CAR and ICAR approaches extends well to hierarchical models. These GLMM CAR types of spatial methods typically require a Bayesian methodology and perhaps this is why spatial statistics was an early adopter of the Markov chain Monte Carlo (MCMC) techniques [13]. MCMC is a simulation based approach used to sample from a probability distribution.

C.6.1. MCMC

MCMC are a group of algorithms that allow simulating from complex distributions using Markov chains [15]. The methods originate from Los Alamos and work conducted on the atomic bomb during World War II [26, 27]. The algorithms start with a distribution we wish to estimate (usually a posterior distribution) and then an algorithm is used to create a Markov chain. Running the Markov chain for a very long time results in the chain converging to the distribution we wish to estimate [28]. The Metropolis-Hastings algorithm is one of the MCMC algorithms that can be used [29, 30].

C.6.1.1. Metropolis-Hastings algorithm

Algorithm C.1 Metropolis-Hastings

- 1: Draw a random starting point θ^0 for the Markov chain
 - 2: Choose a proposal distribution $p(\theta^0|\theta^k)$ such that the probability of the current value given the previous is positive. (Typically, Gaussian centered around θ^k)
 - 3: **For each iteration** k :
 - 4: Generate a candidate θ^* from the proposal distribution
 - 5: Calculate the acceptance probability $r = \min(p(\theta^*/p(\theta^k)), 1)$
 - 6: Generate a uniform random number $u^* \in [0, 1]$
 - 7: **If** $r \geq u^*$:
 - 8: accept the candidate by setting $\theta^{(k+1)} = \theta^*$
 - 9: **Else:**
 - 10: set $\theta^{(k+1)} = \theta^k$
-

The beauty of MCMC is that we do not have to know the form of the distributions we are trying to estimate. This allows estimation of models with very complex distributions. The downside of MCMC is that it requires many computations, but this has become increasingly accessible and since the early 90s MCMC has become available to any researcher with a computer. It has revolutionised statistical computing and had a major impact for Bayesian statistics which was able to move on from simple conjugate prior models to more realistic and hopefully more insightful model forms [28].

Although MCMC has helped to make great strides in the types of models that statisticians are able to fit, work still continues on alternative approaches for statistical inference such as variational Bayes [20] and Laplace approximations.

C.6.2. INLA

Integrated nested Laplace approximation (INLA) [31] uses Laplace approximation to estimate the posterior distribution. The INLA method is based on three components. The Gaussian Markov random field (GMRF), latent gaussian models (LGMs), and Laplace approximation.

C.6.2.1. GMRFs

GMRFs have already been discussed in detail, in brief they are a collection of spatially indexed random variables that are distributed multivariate normal and have a Markov property [19]. One of the big advantages of GMRF is computational efficiency in model fitting. The Markovian property of the GMRF results in a sparse precision matrix, meaning that lots of the values of the matrix are zero and do not need to be involved in computations [31].

C.6.2.2. LGMS

LGMs have been mentioned briefly, and therefore will be described more thoroughly. LGMs are a subset of structured additive models [31, 32] and encompass a wide range of models. They can be thought of as a combination of GLMMs and Generalised additive models (GAMs). GAMs are GLMs which include unknown smooth functions of some of the predictor variables. LGMs can therefore contain fixed effects, random effects, and smoothed effects.

A Structured additive regression model can be represented as:

$$\eta = g(\mu) = \alpha + \gamma Z + \sum_{l=1}^L f_l(z_l) + \epsilon$$

Where γZ are the fixed effects and ϵ is a normally distributed error term. The term $f_l(\cdot)$ is function that can takes different forms such as smoothed effects of the covariates z_l , random intercept or slopes and temporal and seasonal effects. A latent field can be defined as $\theta = \{\alpha, \gamma, f_1(\cdot)\}$ (a collection of random variables that are being estimated in the model). Assuming the latent field can be described by a GMRF then model is a LGM, which includes a models such as [6]:

- Generalised linear (mixed) models.
- Generalised additive (mixed) models.
- Spline smoothing.

- Semi-parametric regression.
- Log-Gaussian Cox-processes.
- Spatio-temporal models.
- Survival analysis.

C.6.2.3. Laplace approximation

Laplace approximation [33] (also called Laplace’s method) is a technique based on Taylor series expansion used to approximate integrals of the form $\int_a^b e^{Mf(z)}$ where a , b could be infinite, $f(z)$ is twice differentiable and M is a large number. The method involves approximating the integral using a (multivariate) normal distribution. In Bayesian statistics calculating the Posterior distribution is equivalent to integration and thus the Laplace method is one way to calculate the posterior. It is typically fast as it only requires calculating the mode of the posterior then using as the mean of a MVN, then it consider the curvature (derivative) of the posterior at the mode to estimate the variance of the MVN [31]. With Laplace approximation the closer the posterior is too Gaussian the better the approximation works.

INLA uses Laplace approximation to estimate the marginal distribution of the latent field and the posterior distribution.

C.6.2.4. Stochastic Partial differential equations

GPs do not have a natural Markov property [6] and therefore INLA does not apply to these methods as clearly as it does to GMRFs. Fortunately a link between the continuously indexed GPs and the discrete indexed GMRFs has been proposed [34] which uses stochastic partial differential equations (SPDEs).

Lindgren et al [34] show that a Matern covariance model (A GP with a constant mean, and Matern covariance) is a solution to a SPDE. SPDEs can also be solved by using a finite element method. The finite element method is a process for solving a larger problem by breaking it down into simpler smaller parts [35]. In the spatial case this involves approximating an area with a finite number of triangles or a ‘mesh’, then the solution of the SPDE can be represented as the weighted sum of the vertices of the ‘mesh’. Assuming a Markov property the ‘mesh’ can then be modelled as a GMRF. Therefore, the GP is a solution of a SPDE, and the SPDE can be approximately solved by using a GMRF. Moreover, this means we can represent the GP in the form of a GMRF and use INLA to fit the model. Further details of the approach can be found in Lindgren et al. [34].

The computation of GPs can be intensive due to dense covariance matrices, therefore approximating the GP using a GMRF extends the computational advantages of the INLA approach to continuously indexed spatial process [36]. It also helps to link the methods of Geostatistics and GPs with GMRF and CAR models [34].

C.7. Summary

The CAR and SAR models were developed as extensions of the conditional and joint distribution specifications of time series analysis. The models are a subset of GMRF models which model discrete spatial processes. Although the terms RFs and SPs are used in some areas of spatial statistics often in more applied areas the CAR and SAR models are presented without the overarching theory and structures they relate to. With this in mind, RFs and SPs have been described in detail showing the link between random variables and GMRFs. This demonstrates that GMRFs are really an assumption about how the RVs we are attempting to model are connected and distributed. They are multi-dimensional equivalents for when the data is assumed to be normally distributed.

When the CAR was first proposed, the auto-normal, auto-logistic, and auto-Poisson were presented as separate methods to model continuous, binary, and count outcomes. This approach has been extended to include other forms of outcomes [11]. An alternative approach involves conditioning on a spatial process and modelling the outcome and process separately. GLMMs provide a framework for this approach allowing the auto-normal model to be applied to the spatial process and a range of distributions from the exponential family to model different types of outcomes. This greatly extends spatial analysis methods and provide a general formulation for geostatistical (GPs) and areal (GMRFs) models.

The SAR and CAR can be fitted using maximum likelihood, however the ICAR and Hierarchical models typically requires more complex fitting techniques such a MCMC or INLA. We briefly looked at MCMC and the three main components of INLA that are LGMs, GMRFs, and Laplace approximation.

The GMRF approach can be extended to point data through the use of distances but perhaps a more natural way is to model the point data as continuous using a GPs. GPs do not have a natural Markov property and requires estimations of a dense covariance function and therefore can be quite difficult to fit. As we have seen it turns out if you restrict the GP to have a Matern covariance function then it can be linked to a GMRF as they are both solutions of a SPDE. This helps to provide a bridge between geostatistics and GMRF models allowing users to model the data using a GP but benefit from the computational efficiencies of GMRFs.

In summary, this chapter has given an overview of modern day spatial modelling, showing the connections between the maximum likelihood methods of SAR and CAR through to the Spatial GLMs that require MCMC and INLA to fit. Although there is some confusing terminology once clearly defined, it reveals a reasonably unified approach that appears to be tending towards greater connections between areal and geostatistical data analysis. This chapter has provided a short review of spatial modelling approaches that were applied in the PhD.

C.8. Bibliography

- [1] Cressie, N. Statistics for Spatial Data. United States of America: John Wiley & Sons Ltd, 1991.
- [2] Whittle, P. On Stationary Processes in the Plane. *Biometrika* 1954 41. (3) 434–449.
- [3] Besag, J. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B* 1974 36. (2) 192–236.
- [4] Lichstein, JW, Simons, TR, Shriver, Sa, and Franzreb, KE. Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs* 2002 72. (3) 445–463.
- [5] Anselin, L. Thirty years of spatial econometrics. *Papers in Regional Science* 2010 89. (1) 3–25.
- [6] Shaddick, G and Zidek, JV. Spatio-Temporal Methods in Environmental Epidemiology. CRC Texts in Statistical Science. Chapman Hall/CRC, June 2015. URL: <http://opus.bath.ac.uk/47731/>.
- [7] Gelfand, A, Diggle, P, Fuentes, M, and Guttorp, P, eds. Handbook of Spatial Statistics. 2010.
- [8] Nelder, JA and Wedderburn, RWM. Generalized Linear Models. *Journal of the Royal Statistical Society. Series A* 1972 13517213. (3) 370–384.
- [9] Diggle, PJ, Tawn, JA, and Moyeed, RA. Model-based geostatistics. *Journal of the Royal Statistical Society: Series C* 1998 47. (3) 299–350. arXiv: MR1626544.
- [10] Banerjee, S, Carlin, BP, and Gelfrand, AE. Hierarchical Modeling and Analysis for Spatial Data. Boca Raton: Chapman & Hall/ CRC, 2015.
- [11] Cressie, N and Wikle, CK. Statistics for Spatio-Temporal Data. Wiley, 2011.
- [12] Besag, J and Kooperberg, C. On conditional and intrinsic autoregression. *Biometrika* 1995 82 733–746.
- [13] Haran, M. Gaussian random field models for spatial data. *Handbook of Markov Chain Monte Carlo*. 2011. Chap. 18449–478. arXiv: arXiv:1206.1901v1.
- [14] Clapham, C and Nicholson, J. The Concise Oxford Dictionary of Mathematics. Oxford Quick Reference. OUP Oxford, 2014.
- [15] Blitzstein, JK and Hwang, J. Introduction to Probability. Chapman & Hall/CRC Texts in Statistical Science. CRC Press/Taylor & Francis Group, 2014.
- [16] Kindermann, R and Snell, JL. Markov random fields and their applications. 1980.
- [17] Rasmussen, CE and Williams, CKI. Gaussian processes for machine learning. MIT Press, 2006. arXiv: 026218253X.
- [18] Gelfand, AE and Schliep, EM. Spatial statistics and Gaussian processes: A beautiful marriage. *Spatial Statistics* 2016 18 86–104.
- [19] Rue, H and Held, L. Gaussian Markov Random Fields: Theory and Applications. Monographs on Statistics and Applied Probability. London: Chapman & Hall, 2005.

- [20] Bishop, CM. Pattern Recognition and Machine Learning. Information Science and Statistics. Springer, 2006.
- [21] Mackay, DJC and Mackay, DJC. Information Theory , Inference , and Learning Algorithms. 2003.
- [22] McElreath, R. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2016.
- [23] Besag, J, York, J, and Mollié, A. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 1991 43. (1) 1–20.
- [24] Richardson, S and Guihenneuc-Jouyaux, C. Impact of Cliff and Ord (1969, 1981) on spatial epidemiology. *Geographical Analysis* 2009 41. (4) 444–451.
- [25] Lawson, A. Bayesian Disease Mapping. CRC Press, 2013.
- [26] Metropolis, N and Ulam, S. The Monte Carlo Method. *Journal of the American Statistical Association* 1949 44. (247) 335–341.
- [27] Metropolis, N. The beginning of the Monte Carlo method. *Los Alamos Science* 1987 15 125–130.
- [28] Geyer, CJ. Introduction to Markov Chain Monte Carlo. Ed. by Brooks, S, Gelman, A, Jones, G, and Meng, XL. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, 2011. Chap. 1.
- [29] Hastings, WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 1970 57. (1) 97–109.
- [30] Metropolis, N, Rosenbluth, AW, Rosenbluth, MN, Teller, AH, and Teller, E. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics* 1953 21. (6) 1087–1092. arXiv: 5744249209.
- [31] Rue, H, Martino, S, and Chopin, N. Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical Society Series B* 2009 71. (2) 319–392.
- [32] Fahrmeir, L and Tutz, G. Multivariate statistical modelling based on generalized linear models. Springer series in statistics. Springer-Verlag, 1994.
- [33] Laplace, PS. Memoir on the probability of causes of events. *Mémoires de Mathématique et de Physique, Tome Sixième*. Trans. by (English translation by S. M. Stigler 1986. *Statist. Sci.*, 1. 1776.
- [34] Lindgren, F, Rue, H, and Lindström, J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 2011 73 423–498.
- [35] Bathe, KJ. Finite element procedures. Second edition. 2nd. Klaus-Jürgen Bathe, 2006.
- [36] Blangiardo, M, Cameletti, M, Baio, G, and Rue, H. Spatial and spatio-temporal models with R-INLA. *Spatial and Spatio-temporal Epidemiology* 2013 7 39–55. arXiv: arXiv:1011.1669v3.