Design and analysis of trials for mosquito control interventions with contamination

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Lea Diana Multerer

Basel, 2021

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel, auf Antrag von

Prof. Dr. Thomas Smith, PD Dr. Tracy Glass und Prof. Dr. Bobby Reiner.

Basel, den 13. Oktober 2020

Der Dekan
Prof. Dr. Martin Spiess

SUMMARY

The movement of mosquitoes leads to a contamination of the effect of an intervention targeting mosquito-borne diseases because people living nearby might benefit from a reduced density of infectious mosquitoes. It is usually attempted to avoid these diffuse effects when testing new interventions. In this thesis however, we seek to understand these contamination effects and build the theoretical basis to estimate them since they may provide valuable information about the intervention effectiveness and should be considered.

We use partial differential equations to describe the dynamics of an *Aedes aegypti* population under a modified male mosquito release on an island to inform the trial design of such an intervention. By invoking optimal control theory, we assess the optimized release strategy under the constraint of a limited availability of modified males to achieve elimination as fast as possible. Our findings show that, to eliminate *Aedes aegypti* from a single location, the optimal release strategy is to initially release a high number of modified males and to subsequently release fewer mosquitoes proportionally to the decreasing female population. The best approach for elimination on the whole island is to target high mosquito density areas first and then move the focus in both directions along the periphery of the island until all areas have been covered. Sufficient release intensity has to be retained in the already targeted areas throughout this process to prevent reintroduction.

We then shift our focus to the analysis of trials targeting *Anopheles* mosquitoes. Informed by a model that describes the dispersion of mosquito with Gaussian kernels, we propose a nonlinear random effects model based on a sigmoid function of the distance to the nearest discordant household to analyze cluster randomized trials or stepped wedge cluster randomized trials of malaria interventions. This model approach leads to a closed-form contamination range that quantifies the measurable extent of the contamination. In a simulation study, we find that this approach indeed provides unbiased and precise estimates of effectiveness if an appropriate number of households is not affected by the contamination range. We extend this model to provide an estimate of the intervention effectiveness as a function of the intervention coverage at each household, that we define based on the estimated contamination range.

This methodological development is then applied to three trials of malaria interventions: the Navrongo trial on the use of insecticide treated nets, the SolarMal trial on the impact of mass trapping of mosquitoes with odor-baited traps and the AvecNet trial on the effect

of adding pyriproxyfen to long-lasting insecticidal nets. These three trials were conducted in different countries with different settlement patterns and were testing different malaria interventions. In our reanalyses we find that a sigmoid analysis yields a similar estimate of effectiveness compared to what was found in the original analysis. Furthermore, in all three trials the contamination range is around 100 to 200 meters, which is much less than the maximal distance *Anopheles* mosquitoes can fly.

For contamination effects to be estimable, the trial must be designed to collect information from zones where contamination is likely. We use the gained insights to give guidance on how to plan trials to allow for contamination and develop algorithmic approaches for cluster construction that allow for cluster boundaries to pass villages and hence enable the estimation of the contamination range. We conclude by connecting the work on the optimized release strategy for modified male mosquitoes with the analysis of trials with contamination by proposing a trial design that accounts for the dispersal of modified males and tests both the short-term effectiveness and the potential for elimination.

The work reported here creates a solid foundation for measuring and understanding the effects of contamination in trials of mosquito-borne diseases. Cluster size can be reduced to the minimum determined by operational factors or contamination effects, without the need for clusters to correspond to separated villages. This reduces the required number of participants in trials of malaria interventions and should lead to more cost-efficient trials and a better understanding of the indirect effects of interventions in protecting nearby nonusers.

ZUSAMMENFASSUNG

Die Bewegung von Mücken führt zu einer Kontamination der Wirksamkeit einer Intervention gegen mückenübertragbare Krankheiten, da benachbarte Personen von einer geringeren Dichte an infektiösen Mücken profitieren können. In der Regel wird versucht, diese diffusen Effekte bei der Validierung neuer Interventionen zu vermeiden. In dieser Arbeit versuchen wir jedoch, diese Kontaminationseffekte zu verstehen und die theoretische Grundlage zu schaffen, um diese Effekte zu messen, da sie wertvolle Informationen über die Wirksamkeit der Intervention liefern können und daher berücksichtigt werden sollten. Wir verwenden partielle Differentialgleichungen, um die Dynamik einer *Aedes aegypti*-Population unter der Freisetzung von modifizierten männlichen Mücken auf einer Insel zu beschreiben, mit dem Ziel, das Studiendesign einer solchen Intervention zu unterstützen. Indem wir auf die Theorie der optimalen Steuerung zurückgreifen, finden wir die optimierte Freisetzungsstrategie um eine schnellstmögliche Eliminierung zu erreichen, unter der Nebenbedingung einer begrenzten Verfügbarkeit von modifizierten männlichen Mücken pro Tag. Unsere Ergebnisse zeigen, dass die optimale Freisetzungsstrategie zur Eliminierung von *Aedes aegypti* an einem Ort darin besteht, zunächst eine hohe Anzahl modifizierter männlicher Mücken freizusetzen und anschliessend proportional zur abnehmenden weiblichen Population weniger männliche Mücken freizusetzen. Der beste Ansatz zur Eliminierung auf der gesamten Insel besteht darin, zuerst Gebiete mit hoher Mückendichte zu visieren und dann den Schwerpunkt in beide Richtungen entlang der Peripherie der Insel zu verlagern, bis alle Gebiete abgedeckt sind. Während dieses gesamten Prozesses muss in den bereits visierten Gebieten eine ausreichende Freisetzungsintensität beibehalten werden, um eine Wiederansiedlung zu verhindern.

Wir wenden uns dann der Analyse von Studien von Interventionen gegen *Anopheles*-Mücken zu. Gestützt auf ein Modell, das die Ausbreitung der Mücken mit Gausskernen beschreibt, schlagen wir ein nichtlineares gemischtes Modell basierend auf einer Sigmoidfunktion der Entfernung zum nächstgelegenen diskordanten Haushalt vor, um cluster randomisierte Studien zu Malaria-Interventionen zu analysieren. Dieser Modellansatz führt zu einem abgeschlossenen Kontaminationsbereich, der das messbare Ausmass der Kontamination quantifiziert. Durch eine Simulationsstudie finden wir heraus, dass dieser Ansatz tatsächlich unverzerrte und präzise Schätzer der Wirksamkeit liefert, wenn eine angemessene Zahl von Haushalten nicht vom Kontaminationsbereich betroffen ist. Wir erweitern dieses

Modell, um die Wirksamkeit der Intervention in Abhängigkeit ihrer Abdeckung in jedem Haushalt zu erhalten, die wir auf der Grundlage des geschätzten Kontaminationsbereichs definieren.
Diese methodologische Entwicklung wenden wir auf drei Studien zu Malaria-Interventionen an: der Navrongo-Studie zur Wirksamkeit von insektizidbehandelten Bettnetzen, der SolarMal-Studie zur Wirksamkeit von Mückenfallen mit Geruchsködern und der AvecNet-Studie zur Wirksamkeit von Pyriproxyfen in insektizidbehandelten Bettnetzen. Diese drei Studien wurden in verschiedenen Ländern, mit unterschiedlichen Besiedlungsmustern durchgeführt und es wurden unterschiedliche Malaria-Interventionen erprobt. In unseren Analysen finden wir, dass eine sigmoidale Analyse eine ähnliche Schätzung der Wirksamkeit ergibt, wie sie in der ursprünglichen Analyse gefunden wurde. Darüber hinaus liegt der Kontaminationsbereich in allen drei Studien bei etwa 100 bis 200 Meter, was wesentlich geringer ist als die maximale Entfernung, die *Anopheles*-Mücken fliegen können.
Damit Kontaminationseffekte schätzbar sind, muss die Studie so konzipiert sein, dass Information gesammelt wird aus Zonen, in denen eine Kontamination wahrscheinlich ist. Wir nutzen die gewonnenen Erkenntnisse, um Anleitung zur Planung von Studien zu geben und entwickeln algorithmische Ansätze für die Konstruktion von Clustern, deren Grenzen Dörfer durchlaufen und somit die Schätzung des Kontaminationsbereichs ermöglichen. Abschliessend verbinden wir die Resultate der optimierten Freisetzungsstrategie für modifizierte männlichen Mücken mit der Analyse von Studien mit Kontamination durch den Vorschlag eines Studiendesigns, das die Bewegung modifizierter Männchen einbindet und sowohl die kurzfristige Wirksamkeit als auch das Eliminationspotential testet.
Diese Arbeit schafft die Grundlage für die Schätzung und das Verständnis der Auswirkungen von Kontamination in Studien von mückenübertragbaren Krankheiten. Die Clustergrösse kann auf ein Minimum reduziert werden, bestimmt durch operationelle Faktoren oder Kontaminationseffekte, und Cluster müssen nicht klar separierten Dörfern entsprechen. Dadurch verringert sich die erforderliche Anzahl von Probanden in Studien zu Malaria-Interventionen, was zu kosteneffizienteren Versuchen und einem besseren Verständnis der indirekten Auswirkungen von Interventionen führt.

ACKNOWLEDGEMENTS

I wish to express my sincerest thanks to my supervisors Tom Smith and Tracy Glass for their scientific and personal support. They entrusted me with this fascinating topic, gave me the liberty to form this PhD and were always there to provide input and advice.

I thank Bobby Reiner for acting as external examiner in the PhD committee. I appreciate that he has taken the time and effort to evaluate my work.

I am grateful for having worked with so many inspiring colleagues. Above all, I would like to thank Fiona Vanobberghen and Nakul Chitnis for their guidance and scientific support. I thank the whole Infectious Disease Modelling unit for their encouragement and input. Special thanks go to my former office and PhD mates for being on this journey with me.

I express my gratitude to the Swiss National Science Foundation for making this work possible and to the get on track program for supporting me.

My deepest thanks goes to my family for all their love and for always trusting in me. I thank my husband Michael, who has always believed in me and supported me. Without them on my side I could not have achieved this. I am most grateful for my wonderful daughter Maria, for every laughter that made me forget everything else.

CONTENTS

Chapter I

INTRODUCTION

Mosquito-borne diseases remain a leading cause of morbidity and mortality in tropical and sub-tropical regions of the world. In 2018, an estimated 228 million cases of malaria and 96 million cases of dengue occurred worldwide [WHO, 2019, Bhatt et al., 2013, WHO, 2017a]. Reducing the burden of these diseases is of major public health concern. For most mosquito-borne diseases, vaccines are not available and mosquito control remains the primary means of reducing transmission. Effective mosquito control relies heavily on insecticidal interventions and in the last years, insecticide resistance of mosquitoes has been rising [WHO, 2018]. We are hence in urgent need of new mosquito control interventions to continue the success of the last decades.

New interventions require trials to confirm their effectiveness. The design and analysis of mosquito control trials imposes many challenges. Often, interventions have diffuse effects since individuals not receiving an intervention may still benefit from it because of mosquito movement. Other difficulties add to this. The impact of combinations of interventions might be unclear and operational constraints of the intervention deployment influence the trial design. A fundamental understanding of the targeted mosquito species and the intervention is needed in order to plan and analyze trials [WHO, 2017b]. Mathematical models and simulations can aid in this [Halloran et al., 2017]. They can assess the likely outcome of an intervention, trials can be simulated with different designs, analyses can be tested, and the magnitude of the simulated intervention effect and its variance can be predicted. Models can also be used to better understand contamination effects in interventions and in trials testing them by including mosquito movement in simulations and by allowing for it in the analysis.

1. Mosquito dynamics and how to simulate them

1.1 Mosquito biology and control interventions

The mosquito life cycle is divided into four stages, as indicated in figure I.1. Female

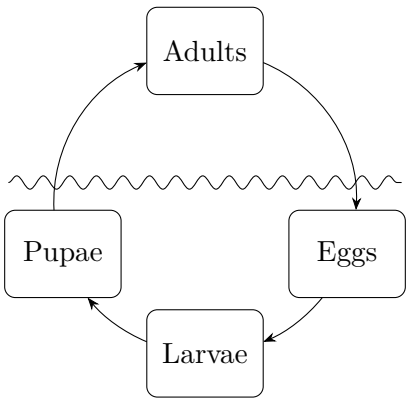


Figure I.1: Description of the mosquito life cycle. It consists of three aquatic stages and one adult stage.

adults lay eggs on stagnant water. These eggs pass three aquatic stages before an adult emerges from the water. While male mosquitoes feed primarily on plant nectar, female mosquitoes need blood meals for ovipositing. The development of each stage is temperature and resource dependent, leading to high seasonality in mosquito distribution and density. Mosquito control entails many different approaches, such as environmental management, chemical control by larvicides or adulticides and biological control [WHO, 2017a]. In this work, we focus on *Aedes* and *Anopheles* mosquitoes and their control. These species differ in many factors such as preferences of breeding sites, biting hours, average flight distance and in the diseases they carry.

Aedes mosquitoes exist almost everywhere in the world [Bhatt et al., 2013, WHO, 2009]. The species *Aedes aegypti* is the main vector of many viral diseases such as dengue, chikungunya and zika. They develop in densely populated areas and prefer to breed in containers, for example in discarded tires. *Aedes aegypti* are known to fly around 100 meters over their lifespan [Verdonschot and Besse-Lototskaya, 2014] and the biting of females happens predominantly during the day. Lately, the use of modified *Aedes* mosquitoes has gained a lot of attention [Dyck et al., 2005, Alphey et al., 2010, Harris et al., 2011, Carvalho et al., 2015, Zheng et al., 2019]. Male mosquitoes are modified and then released into the environment in order to mate with the native females. A female mosquito that mates with a modified male produces offspring that will die before reaching adulthood and the natural reproductive process is disrupted. Modification is usually achieved either by irradiation, by an infection with Wolbachia bacteria or by inserting a lethal gene [Flores and O'Neill, 2018]. Several trials testing the feasibility of suppressing *Aedes* mosquitoes are ongoing or have recently finished [Flores and O'Neill, 2018, Zheng et al., 2019].

Anopheles mosquitoes are the only vectors of human *Plasmodium*, the cause of human malaria. In constrast to *Aedes* mosquitoes, *Anopheles* prefer to breed in natural wa-

ter bodies and are hence mainly found in rural areas. They bite in the night hours and fly several hundred meters over their lifespan [Service, 1997, Verdonschot and Besse-Lototskaya, 2014, Guerra et al., 2014]. Between 2010 and 2018, the estimated number of malaria deaths worldwide decreased from 585 000 to 405 000 [WHO, 2019]. The continuation of this success is at risk due to increase in insecticide and drug resistance of *Anopheles* mosquitoes and *Plasmodium* parasites [WHO, 2018]. Vaccines are under development, the most advanced candidate being *RTS,S/AS01*, with a recently completed phase III trial [RTS,S Clinical Trials Partnership, 2015]. But for now, malaria control is mostly achieved with mosquito control, by using insecticide treated nets, indoor residual spraying or larval control.

1.2 Models of mosquito dynamics and movement

A model describes an observed process with a mathematical formulation. The theoretical and numerical analysis of this model can then provide insights on the observed process and help to generalize it beyond the instances that can be measured. The model approach chosen to describe a mosquito life cycle, the effect of an intervention, or the dispersal of mosquitoes depends on the goal, the previous knowledge of the process and the required level of detail the researcher has [Halloran et al., 2017].

There is a difference between statistics and mathematics in the relation of a model and observed data. In statistics, a model has the scope to characterize data and to estimate the likely future behavior based on past behavior. In mathematics, a model has the scope to describe a system and its changes and to assess the likely future behavior based on the model that was calibrated by the data. The use of data hence works in two ways. For a statistician, the data is used to understand the model and for a mathematician, the model is used to understand the data. In this thesis, we will use both concepts of a model, either to simulate interventions or to analyze them and we will use the word model for both statistical and mathematical models.

A common approach to describe mosquito dynamics is through dynamical systems with so-called ordinary differential equations (ODEs). The foundation was laid by Ross [Ross, 1911] and Macdonald [Macdonald, 1957] with a predator-prey model describing the interaction between female mosquitoes and human hosts. This model was extended and used extensively to gain insights on mosquito-borne disease transmission [Smith et al., 2012, Reiner et al., 2013, Smith et al., 2014]. It offers high flexibility and can be extended to describe the interaction between female and sterile male mosquitoes [Barclay and Mackauer, 1980, Dye, 1984] or to consider more life stages [Esteva and Yang, 2005] for instance. ODEs can appropriately capture interaction dynamics between different stages or species but have no allowance for spatial interaction. They can be extended to partial differential

equations (PDEs) by including a diffusion term. This term couples the interaction dynamics in time with space. This extension is crucial to understanding the spatial transmission of mosquito-borne diseases at a small scale [Dufourd and Dumont, 2013]. The theory of both ODEs and PDEs provides the right framework to answer the following question: given a dynamical model with a variable that can be controlled for, what control leads to the best outcome, based on a predefined goal? This question is answered with optimal control theory [Evans, 1983, Tröltzsch, 2010]. The desired outcome is described by an objective functional that is then maximized under the constraint of the dynamical system and possibly other constraints. In recent years, optimal control found many interesting applications in biology and epidemiology [Lenhart and Workman, 2007], especially to better understand interactions between different interventions [Thomé et al., 2010, Fister et al., 2013].

Although dynamical systems provide an attractive approach for the analysis of the effects of mosquito dynamics and movement, their use is limited if a high level of detail is required or if very heterogeneous settings are studied due to computational complexity and limited analytical knowledge of the solution. Because of this, other approaches are often chosen to simulate mosquito movement. PDEs can be simplified to a discrete space continuous time model, where the spatial component is replaced by patches interacting with each other, coupled with an ODE model [Lutambi et al., 2013]. A more data driven approach without considering the mosquito life cycle and other interactions is given through kernel density estimation [Kelsall and Diggle, 1995, Waller, 2010, Hazelton, 2016]. By approximating the likely movement of female mosquitoes, and hence the blood meals taken, around a household with a density function [Malinga et al., 2019a, Malinga et al., 2019b], the mosquito population and its changes under an intervention can be simulated.

2. Trial design for mosquito control interventions

2.1 Basic concepts of trial design

An individually randomized controlled trial is a type of experiment in which individuals are randomly assigned to either intervention or control group. The randomization aims to reduce selection bias [Wilson et al., 2015, Hayes and Moulton, 2009], a systematic error that arises with human choices. Sample size calculations to determine the required number of individuals per trial arm to achieve the desired power can be performed. An adequately powered individually randomized controlled trial is the gold standard for evaluating the effectiveness of an intervention [Wilson et al., 2015, WHO, 2017b].

In cluster randomized trials (CRTs), groups of individuals are randomized to the intervention or control group. This study design is appropriate if individual randomization is

not possible or not desirable [Hayes et al., 2000, Hayes and Moulton, 2009, Hussey and Hughes, 2007, Reich et al., 2012]. Conventional sample size calculations for CRTs inflate the required sample size for an individually randomized trial by the design effect, a measure for the increase in variance for using a more complex design [Kish, 1965, Donner et al., 1982, Hayes and Bennett, 1999, Hayes and Moulton, 2009]. This is required to correctly estimate the effect of interest because the clustering reduces the variance due to homogeneity within clusters. For a constant trial population, it is often attempted to subdivide the population into as many clusters as possible in the context of logistic constraints that impose a minimum cluster size, with the objective being to maximize the statistical efficiency conditional on some assumed or previously estimated intra-cluster correlation coefficient (ICC), a measure of variation of the outcome within clusters. Since mosquito control interventions are often allocated to households or entire areas and may provide both individual protection to the immediate recipients and also induce community effects by reducing onward transmission, their effectiveness is usually tested with CRTs.

Stepped wedge cluster randomized trials (SWCRTs) [Hussey and Hughes, 2007, Hayes and Moulton, 2009, Mdege et al., 2011] are a type of crossover CRT where all clusters first form the control group and are then sequentially and randomly assigned to the intervention group. SWCRTs have been criticized because they can be inferior in terms of power or bias compared to parallel designs [Wolbers et al., 2012, Kotz et al., 2013, Mdege et al., 2012] and are vulnerable to imprecision caused by temporal trends in underlying disease rates. Under certain logistical, practical or financial circumstances however, the use of SWCRTs is proposed [Wilson et al., 2015, Silkey et al., 2016]. They may for example provide an appropriate design if a trial of a mosquito control intervention aims to evaluate the potential of interrupting disease transmission and evaluation at maximum scale-up is needed.

2.2 Contamination effects in trials of mosquito control interventions

In both CRTs and SWCRTs of interventions targeting mosquito densities, the intervention effects can be contaminated due to interaction between clusters [Halloran and Struchiner, 1991, Halloran et al., 2010, Halloran et al., 2017]. There is no unified term for this phenomenon in the literature, it is either referred to as contamination, spill-over, indirect, mass or community effect. For mosquito control interventions, this contamination effect can occur in two ways, either due to human or mosquito movement [Hayes and Moulton, 2009, Wilson et al., 2015]. For *Aedes* transmitted diseases, where people get infected during the day, clusters are often chosen as schools or workplaces to minimize the human movement across cluster boundaries. With *Anopheles* mosquitoes, where transmission happens mostly indoors at night, human movement is not the main driver for contami-

nation. Clusters are therefore chosen as geographically contiguous areas of households to minimize contamination between them. This contamination can either be protective for nearby users, in that it lowers infectious mosquito densities, or result in an anti-protective effect, for example because the intervention only diverts mosquitoes without an added killing effect [Maia et al., 2013]. This reversed effect is not further considered here.
Although appropriate cluster choices reduce bias, the cluster boundaries are still subject to contamination arising from mosquito movement, unless they are completely separated. If clusters are too small then the effect of the interventions may be propagated beyond the cluster boundary via contamination throughout the whole population, biasing the difference between the trials arms towards zero. On the contrary, if clusters are very large, the contamination effect is negligible, but the geographical size of the trial drastically increases.
With the current standard methods, it is not possible to determine whether a low estimate of effectiveness is a result of bias caused by contamination due to inappropriate cluster size or of a disappointing effect of the intervention. In addition, estimating the spatial extent of contamination is of interest in itself [Halloran et al., 2017, Baird et al., 2018, Anaya-Izquierdo and Alexander, 2020]. The demonstration of measurable contamination effects in several CRTs of insecticide treated nets against malaria [Binka et al., 1998, Howard et al., 2000, Hawley et al., 2003] fed into the rationale for the massive distribution of nets across Africa. The range over which contamination is relevant does not necessarily equate with the maximal distance mosquitoes can fly. New statistical models are therefore needed that analyze this range of spatial contamination as trial outcome.

2.3 Analysis of trials with contamination

In the past, CRTs of mosquito control interventions were often designed so as to prevent contamination between clusters. This can be achieved by choosing well separated clusters or by enforcing the separation via a buffer zone for instance, in a so-called “fried egg” design [Hayes and Moulton, 2009, Wolbers et al., 2012, Protopopoff et al., 2015, Delrieu et al., 2015, Eisele et al., 2016, Protopopoff et al., 2018]. The whole cluster receives the intervention, but only the core (the “egg yolk”) is included in the analysis. If the per capita intervention costs are high, this may substantially increase the cost compared to a trial where only the “egg yolk” is intervened. An alternative is to fully exclude some clusters from the analysis [McCann et al., 2018] based on a criterium of closeness between households to attain a better separation of the intervention and control arms. This results in a trial design in which discordant clusters are separated by at least the contamination range but clusters in the same trial arm may be closer to each other. For both trial designs, a standard analysis [Hayes and Moulton, 2009] with an individual-based approach and a

random effect for the clustering is possible but comes at the cost of a large geographical trial or possible heterogeneity between clusters due to the spatial separation. Both designs also require an estimate of the measurable contamination range to quantify the buffer zone that can only be based on expert opinion.

If the contamination is not prevented in the design, it must be accounted for in the analysis. This might allow for trials with smaller areas and adds information to the trial outcome because of the possibility of estimating the contamination. On the other side, the analysis gets more complicated and there is potential for underestimating the contamination, leading to a biased estimate of effectiveness. A recent systematic review of spatial analysis methods for CRTs [Jarvis et al., 2017] reviewed CRTs that contain a method that accounts for the structure, location or relative distances between observations. Two categories of analyses were distinguished: analyses that assume that the spatial effect can be measured and include a spatial variable that relates the outcome to where the observations are located or analyses that include a spatially structured random effect.

In all analyses listed in the systematic review [Jarvis et al., 2017] that included a spatial variable, a linear model was extended with a term of the straight-line distance to the nearest discordant observation or a term of the density of households within a range that receive the intervention. This is quite simple to implement with existing software and the interpretation of the coefficients is straightforward. Nevertheless, although interaction terms and a link function can be used to adequately adapt a linear model to the problem, the contribution of each estimated coefficient remains linear and it is not possible to obtain a closed-form contamination range that specifies the maximal measurable extent of contamination. It is also not possible to obtain this information from a model with a spatially structured random effect. Another approach is to construct a nonlinear model, informed by the underlying dynamics of the mosquito movement. A nonlinear model can yield a closed-form estimate of the range of the contamination and hence adjust the estimate of intervention effectiveness. Simple random effects with an exchangeable correlation structure can be added to both linear and nonlinear model approaches for the clusters and, if needed, for survey rounds or hierarchical structures of the observations such as households or health facilities.

For any analysis accounting for contamination, the spatial distribution of the population in each arm, relative to the boundary is relevant. As cluster size becomes smaller, the contamination will tend to become more important and it might be difficult to attain an accurate and precise estimate of effectiveness. This represents a conventional problem in sample size determination and the relationship between cluster size, the level of contamination and the spatial distribution of the population has to be considered when applying sample size formulae. The connections between the design and analysis of a CRT

of mosquito control interventions with potential for contamination as described here can be found in figure I.2.

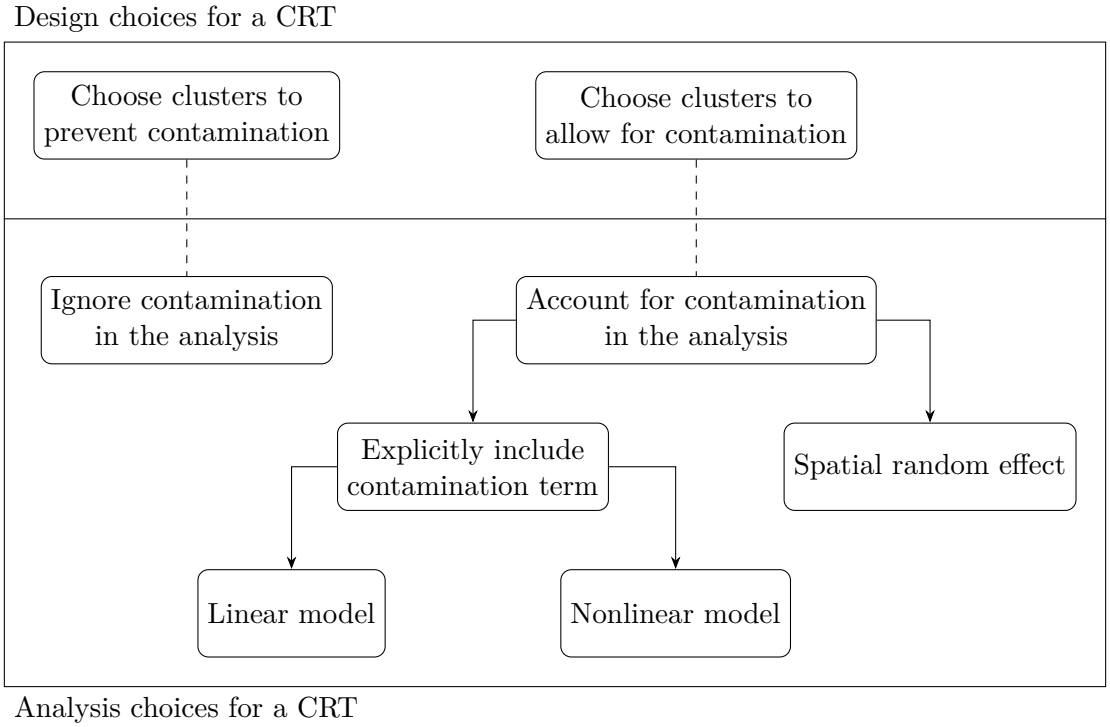


Figure I.2: Schematic summary of the design and analysis for a CRT of mosquito control interventions with potential for contamination. The blocks represent choices the researcher has to take, possibly leading to further choices indicated with arrows. The dashed lines show the counterparts of the design choices in the analysis choices.

3. Objectives and outline

The goal of this thesis is to use models to better understand contamination in new interventions of mosquito control and elimination and to provide the necessary analytical foundation to account for it in the trial design and analysis of such interventions. This overall goal breaks down into the following four objectives:

(i) Use a model of mosquito movement to describe an *Aedes aegypti* population under a modified male release and find the optimal strategy to eliminate them from an island.

(ii) Develop a model for the analysis of CRTs of malaria interventions that can adjust the intervention effectiveness for the contamination arising from mosquito movement and can estimate the contamination in closed-form.

(iii) Extend this model to SWCRTs and show the implications from estimating the contamination on the relation between the intervention coverage and the intervention effectiveness.

(iv) Show how trials can be designed for malaria control interventions in the presence of contamination.

Each of the chapters II,III,IV and V corresponds to one objective. In chapter II, we use a PDE model to simulate mosquito movement and apply optimal control theory to find the optimal strategy of elimination for *Aedes aegypti*. In chapter III we develop a nonlinear model for the analysis of CRTs with contamination effects and test the performance in a simulation study, where movement of *Anopheles* mosquitoes is simulated with a kernel density approach. The model is applied to a CRT of insecticide treated nets in northern Ghana. Chapter IV extends this work to SWCRTs and applies the model to two SWCRTs of different malaria interventions in Kenya and Burkina Faso. In this chapter we also show how the proposed analysis extends to estimating the intervention effectiveness at an arbitrary intervention coverage level. In chapter V, we answer remaining open questions from chapters III and IV on how to design such trials and develop several algorithmic approaches to cluster assignment. We conclude with a redesign of a previously analyzed trial. In the last chapter, we summarize all the results and put them in perspective. We also connect objective (i) with objectives (ii), (iii) and (iv) by outlining a potential plan for a modified male mosquito intervention, informed by the other results.

Chapter II

MODELING THE IMPACT OF STERILE MALES ON AN *AEDES AEGYPTI* POPULATION WITH OPTIMAL CONTROL

Lea Multerer[1,2] , Thomas Smith[1,2] , Nakul Chitnis[1,2]

[1] Swiss Tropical and Public Health Institute, Basel, Switzerland
[2] University of Basel, Basel, Switzerland

Published in *Mathematical Biosciences* in March 2019. (DOI: 10.1016/j.mbs.2019.03.003)

1. Abstract

We use partial differential equations to describe the dynamics of an *Aedes aegypti* mosquito population on an island, and the effects of a sterile male release. The model includes mosquito movement and an Allee effect to capture extinction events. We apply optimal control theory to identify the release strategy that eliminates the mosquitoes most rapidly, conditional on a limited availability of sterile males. The optimal solution for a single location is to initially release a substantial number of mosquitoes and to subsequently release fewer sterile males proportionally to the decreasing female population. The optimal solution for the whole island is intractable given a constraint on the total daily release of sterile males. The best approximation to the spatial optimal control strategy is to focus on the high mosquito density areas first and then move outwards (in both directions along the periphery of the island), until all areas have been covered, retaining throughout sufficient release intensity to prevent reintroduction in the already cleared areas.

Highlights

- Spatial models are needed to evaluate the impact of sterile male mosquito releases
- The application of optimal control gives insights for an optimal release strategy
- In every area, an initially high amount of sterile males has to be released
- An intervention is most effective when high density areas are targeted first

2. Introduction

The sterile insect technique (SIT) is a method of biological control, that entails releasing high numbers of sterile male organisms into the environment. The natural reproductive process of the population is disrupted by competition with fertile males, and the insect population may be eliminated if the number of released sterile males is high enough and is repeated over a sufficient duration of time.
This method was first introduced by Knipling [Knipling, 1955] in 1955 and was successfully applied to eliminate the Screwworm fly from North and Central America in the late 1950s and early 1960s [Dyck et al., 2005]. More recently, transgenic strategies have been developed to improve SIT [Alphey et al., 2010]. The RIDL (release of insects with dominant lethality) technique inserts a lethal gene into mosquitoes. A potential application of such strategies is to achieve elimination of a disease vector population on an island.
The development and analysis of mathematical models for SIT in mosquitoes goes back to the simple logistic model of Barclay and Mackauer [Barclay and Mackauer, 1980] and the similar model of Dye [Dye, 1984], specific for *Aedes aegypti* mosquitoes. Various simple models, describing the dynamics of an arbitrary mosquito population have been analyzed [Li, 2017, Li and Yuan, 2015, Cai et al., 2014]. Li [Li, 2017] compared release strategies for a simple model: constant release, release proportional to the population and proportional release with saturation. Esteva and Yang [Esteva and Yang, 2005] focused on a more extensive model, incorporating different life stages of *Aedes aegypti*.
Various researchers have applied optimal control theory to identify the optimal release strategy when SIT is combined with insecticide application [Thomé et al., 2010, Fister et al., 2013, Rafikov et al., 2015]. Kim *et al.* [Kim et al., 2017] applied an optimal control approach to determine the optimal ordering of an arbitrary intervention in a two patch dengue transmission model. A few analyses have considered spatial effects of SIT. Discrete models on cells have been considered for *Aedes aegypti* [Ferreira et al., 2008], coupled with a pulsed release of sterile males [Oléron Evans and Bishop, 2014]. Partial differential equations (PDEs) provide an attractive analytical approach for the analysis

of effects of mosquito movement. Dufourd and Dumont [Dufourd and Dumont, 2013] illustrated this by using advection-diffusion-reaction equations to simulate *Aedes albopictus* dispersal, taking into account environmental parameters. Seirin Lee *et al.* considered a PDE model to determine the optimal barrier size where sterile males are released in order to prevent immigration into an area [Seirin Lee et al., 2013b], and a control strategy for endemic and emerging outbreaks [Seirin Lee et al., 2013a]. Maidana and Yang [Maidana and Yang, 2008] described the geographic spread of Dengue by a system of partial differential reaction-diffusion equations and its travelling wave solution.
However, there are no results available on the optimal spatial strategy for a SIT on a domain, in particular one corresponding to an island. This paper aims to fill this gap by applying optimal control methods to a spatial PDE model for the dynamics of a mosquito population under a SIT. The model is parameterized to represent an *Aedes aegypti* population on a small island with a predominantly coastal human population, as is the case for e.g. Réunion in the Indian Ocean, Guam or Palau in the Pacific or Grenada, St. Vincent and St. Lucia in the Caribbean.

3. Model formulation

3.1 Equations for the single-site model

Let $u(t)$ denote the female, and $\psi(t)$ the effective sterile male mosquito population, equal to the surviving released mosquitoes, adjusted for their mating efficiency relative to the wild fertile males at time t. Under the assumption that fertile males and females have the same population dynamics we only consider wild females and sterile males. The per capita emergence rate for wild females is given by

$$\phi\left(1-\frac{u}{\kappa}\right),$$

where κ is the carrying capacity and ϕ the intrinsic oviposition rate of one female mosquito, divided by two (since only female mosquitoes are counted). The probability that a female mates with a fertile male is described by the term

$$\frac{u}{u+\psi},$$

where $u+\psi$ is the effective population of males, including the sterile males. Since we assume equal numbers of fertile males and females, u denotes the number of fertile male mosquitoes (as well as the number of females).
It is reasonable to assume that if the population u is very small, it can no longer sustain itself, since the mosquitoes rarely encounter mates. To include this effect into the model, we choose a function that has no effect if the population is big and leads the population

to zero if u falls under a certain threshold value. There is a variety of functions to model this Allee effect [Boukal and Berec, 2002]. Following Li [Li, 2017], we choose a function of Holling-II type [Holling, 1959]

$$\frac{\alpha u}{\alpha u + 1}, \tag{3.1}$$

where the constant α is the searching efficiency representing the encounter rate between two mosquitoes at very low density. The population ψ is described by the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi = \xi - \mu\psi,$$

where ξ is the number of effective sterile mosquitoes released every day and μ is the natural mortality rate, which we assume to be the same for all mosquitoes. The term ξ has to be adjusted by a constant if the mating efficiency of the sterile males is known to be less than for the wild males.
In conclusion, we require $u(t)$ and $\psi(t)$ to satisfy the following initial value problem:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}u &= \phi\left(1 - \frac{u}{\kappa}\right)\left(\frac{\alpha u}{\alpha u + 1}\right)\left(\frac{u}{u + \psi}\right)u - \mu u =: f(u, \psi), \\ \frac{\mathrm{d}}{\mathrm{d}t}\psi &= \xi - \mu\psi =: g(\psi), \end{aligned} \tag{3.2}$$

for $t > 0$, together with initial conditions $u(0) = u_0 \in [0, \kappa]$ and $\psi(0) = 0$. The terms ξ and κ can either be constant or depend on t. For simplicity, the explicit notation of this is dropped unless it is considered important. If no mosquitoes are released, this simplifies to the standard logistic model with an Allee effect.

3.2 The PDE model

We simplify the island by treating it as a circular boundary, corresponding to the coastline of a small island with a predominantly coastal human population hosting anthropophagic mosquitoes (such as *Aedes aegypti*).
The island is modeled by an interval $\Omega = [0, L]$ with periodic boundary conditions on its boundary $\partial\Omega = \{0, L\}$ (figure II.1). The model derived in the previous subsection

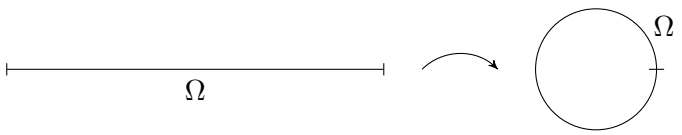


Figure II.1: The effect of periodic boundary conditions.

is easily extended to capture the spatial behavior of the population by adding diffusion

terms, resulting in the following initial value problem:

$$
\begin{aligned}
\frac{\partial u}{\partial t} &= D\Delta u + f(u,\psi), && \text{for } x \in \Omega,\ t > 0,\\
\frac{\partial \psi}{\partial t} &= D\Delta \psi + g(\psi), && \text{for } x \in \Omega,\ t > 0,\\
u(0,x) &= u_0(x),\ \psi(0,x) = 0, && \text{for } x \in \Omega,\\
u(t,0) &= u(t,L),\ \psi(t,0) = \psi(t,L), && \text{for } t > 0,\\
u_x(t,0) &= u_x(t,L),\ \psi_x(t,0) = \psi_x(t,L), && \text{for } t > 0,
\end{aligned}
\tag{3.3}
$$

where D is a diffusion coefficient. Apart from u and ψ, we allow ξ and κ to depend on x and t in $g(\psi)$ and $f(u,\psi)$. We assume that $u_0(x) > 0$ is well behaved, such that the problem above is feasible. This model thus consists of the variables and parameters listed in table III.1, where all the parameters are constrained to be positive with $\phi > \mu$ to hold.

Variables and parameters		Values	Units
u	Female mosquito population		mosq [km^{-1}]
ψ	Effective sterile male population		mosq [km^{-1}]
ϕ	Intrinsic oviposition rate	5	days^{-1}
μ	Natural mortality rate	0.15	days^{-1}
κ	Carrying capacity	1000	mosq [km^{-1}]
α	Searching efficiency	0.05	dimensionless
ξ	Release rate of sterile mosquitoes		mosq$\times$days^{-1}[km^{-1}]
D	Diffusion coefficient	5×10^{-5}	km^2 $\times$ days^{-1}
L	Length of domain	10	km

Table II.1: Summary of the variables and parameters used in the model. The extended units for the PDE model are written in brackets; mosq is an abbreviation for mosquitoes.

4. Parametrization for *Aedes aegypti*

Values for $\phi, \mu, \kappa, \alpha, D$ and L were chosen to represent the dynamics of a SIT for *Aedes aegypti* (see table III.1), drawing upon various parameter sets used in the literature [Esteva and Yang, 2005, Yang et al., 2016, Oléron Evans and Bishop, 2014, Seirin Lee et al., 2013b, Harris et al., 2011, Carvalho et al., 2015].

Estimates of the intrinsic oviposition rate range from 5 eggs per day [Esteva and Yang, 2005] to 16 eggs per day [Oléron Evans and Bishop, 2014]. We only consider female offspring and take the mean value; hence female *Aedes aegypti* lay an average of 5 female

eggs per day. For the mortality rate, Yang *et al.* [Yang et al., 2016] state mortality rates for larvae, pupae and adults. Combining these values with

$$(1-\mu) = (1-\mu_{larva})(1-\mu_{pupa})(1-\mu_{adult}),$$

we get $\mu = 0.15$. We have not found any values for the searching efficiency in the literature and assign a value of $\alpha = 0.05$, for when the population u is small [P. Müller, *personal communication*]. Estimates of the carrying capacity κ in the literature are highly variable. For the single-site model we assign $\kappa = 1000$ and set $u_0 = \kappa$.
When simulating the PDE, the units of u, ψ, ξ and κ change to mosquitoes per length unit. Following Oléron *et al.* [Oléron Evans and Bishop, 2014] the values for κ in the PDE model are chosen such that the mean over the island is 100 females per hectare, and a spatial variation is captured with a sinusoidal function representing four villages at equal distances around the island. Oléron *et al.* [Oléron Evans and Bishop, 2014] assume average movement of 30 meters per day. Approximately corresponding to this we set the diffusion coefficient D to be 50 square meters per day. We assume that the island has a circumference of $L = 10$km and hence $\Omega = [0, 10]$.

5. Analysis of the single-site model

5.1 Existence and uniqueness

(5.1) **Theorem.** There is a global unique solution w to the initial value problem (3.2) that is nonnegative and bounded.

Proof. The proof of this theorem uses standard arguments from the theory of differential equations. Recall that we assume all the parameters to be positive and $\phi > \mu$ to hold. Clearly, the functions $f(u, \psi)$ and $g(\psi)$ are globally Lipschitz continuous on $[0, \kappa]$ and hence bounded. The global existence and uniqueness of a solution $w := [u, \psi]^T$ for every $w_0 := [u_0, 0]^T$ directly follows from the Picard-Lindelöf Theorem.
It remains to show that if $u_0 \in [0, \kappa]$ is fulfilled, then the solution w is nonnegative and bounded. The boundedness is clear. To show that w is nonnegative, we define a sub solution v as

$$\frac{\mathrm{d}}{\mathrm{d}t} v := -c \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} < \begin{bmatrix} f(u,\psi) \\ g(\psi) \end{bmatrix} = \frac{\mathrm{d}}{\mathrm{d}t} w,$$

with a constant c. The existence of such a constant is ensured by the boundedness of $f(u, \psi)$ and $g(\psi)$. By a standard comparison result it follows that $v(t) \leqslant w(t)$ for $0 < t < \infty$ whenever $v(0) \leqslant w(0)$. We choose $v(0) = [0, 0]^T$, and hence, $v(t) \equiv 0$ holds for $0 < t < \infty$ and the claim follows. □

5.2 Equilibrium points

It is clear that ξ/μ is the only solution for $g(\psi) = 0$. A trivial equilibrium of system (3.2) is then given by $E_0 = [0, \xi/\mu]^T$, corresponding to the state when natural insects are absent and there is only a constant population of sterile males. Solving the equation $f(u, \psi) = 0$ for u, we find three more equilibria, $E_{1,2,3} = [e_{1,2,3}, \xi/\mu]^T$, where $e_{1,2,3}$ are given given by the roots of the polynomial

$$p(z) = \phi\alpha z^3 + \kappa\alpha(\mu - \phi)z^2 + \kappa(\mu + \alpha\xi)z + \kappa\xi.$$

By Descartes rule of signs, we easily see that $\phi > \mu$ is a necessary condition for two roots to be positive and real. It can easily be shown that e_3 is always negative and is hence disregarded. There remain three equilibria, $E_0 < E_2 \leqslant E_1$. We are ultimately interested in the optimal number of sterile insects to release in order to achieve extinction, and the minimal time required. The expressions for $E_{1,2}$ are real if the discriminant of the cubic polynomial is non negative. When the discriminant is less than zero, $E_{1,2}$ collapse and there is only one non-negative equilibrium point left, E_0. In this case, regardless of the initial conditions, the population can be driven to zero. To get a threshold value for the number of mosquitoes that must be released we set the discriminant to zero and rewrite the expression in terms of ξ. We find ξ° such that: if $\xi > \xi^\circ$ sterile mosquitoes are released, E_0 is the only equilibrium point and extinction will occur for any initial condition. The stability of the equilibria $E_{0,1,2}$ is determined by evaluating the sign of the eigenvalues of the Jacobian of $[f(u, \psi), g(\psi)]^T$ in the points $E_{0,1,2}$. A negative eigenvalue indicates a stable equilibrium and vice versa. We get the following relations:

$$\xi \begin{cases} < \xi^\circ, & E_0, E_1 \text{ are stable and } E_2 \text{ is unstable,} \\ = \xi^\circ, & E_{1,2} \text{ collapse to an unstable equilibrium and } E_0 \text{ is stable,} \\ > \xi^\circ, & E_{1,2} \text{ are infeasible and } E_0 \text{ is globally stable,} \end{cases}$$

leading to the bifurcation diagram in figure II.2.

5.3 Time to elimination

To assess the threshold mosquito population at which a SIT can be stopped, we assess the equilibria in the ordinary differential equations (ODE) (3.2) corresponding to the situation with no sterile males, $\xi = 0$. There are three equilibria, $\tilde{E}_0 = 0$ and

$$\tilde{E}_{1,2} = \frac{(\phi\kappa\alpha - \mu\kappa\alpha) \pm \sqrt{(\mu\kappa\alpha - \phi\kappa\alpha)^2 - 4\phi\mu\kappa\alpha}}{2\phi\alpha},$$

where $\tilde{E}_0$ and $\tilde{E}_1$ are stable, $\tilde{E}_2$ is unstable and $\tilde{E}_0 < \tilde{E}_2 \leqslant \tilde{E}_1$. Assume now that sterile mosquitoes are released until some time t is reached. If $u(t) < \tilde{E}_2$, the population will

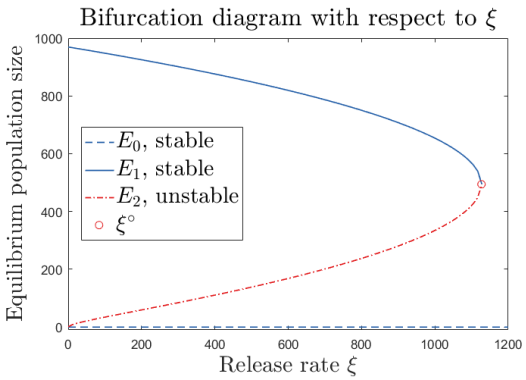


Figure II.2: Bifurcation diagram with respect to ξ, where the parameters are chosen as described in section 4.

reach the stable state $\tilde{E}_0 = 0$ and if $u(t) > \tilde{E}_2$, the population will increase again to $\tilde{E}_1$. Let t^* denote the threshold time that sterile males have to be released to bring the population below $\tilde{E}_2$, meaning

$$u(t) < \tilde{E}_2 \quad \text{for all } t \geqslant t^* > 0.$$

For the set of parameter values from section 4 it holds that $\tilde{E}_0 = 0$, $\tilde{E}_1 \simeq 969.4$, $\tilde{E}_2 \simeq 0.62$, implying that only when the population is below one female mosquito can releases of sterile males be discontinued.

5.4 Seasonal variations in the carrying capacity

The carrying capacity varies in time with external factors including temperature and rainfall. To explore the impact of seasonal variations in the carrying capacity we replace the constant κ with the periodic function

$$\kappa(t) = 1000 - \exp\left(2\pi \cos\left(\frac{2\pi}{T}t - \frac{\pi}{2}\right)\right),$$

where T is one year. When the release is barely sufficient to eliminate the population, (ξ is very close to the threshold value ξ°), then the transient drop in carrying capacity substantially shortens the time to extinction (figure II.3a). When the release is substantially more than the threshold, this effect is minimal, since the elimination is much faster than the effect of the change in carrying capacity (figure II.3b). It follows that temporal variation in carrying capacity should be taken into account in planning when to start a SIT intervention, on the basis of this model a good time to start an intervention would be when the carrying capacity is lowest. Periodic changes due to seasonality are less important because with a reasonable number of sterile males, the main impact of a SIT intervention happens in a very short time period. However, the model does not allow for dormancy of

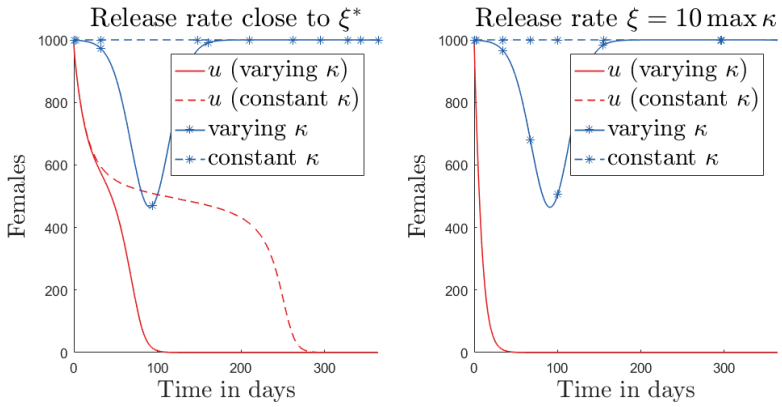


Figure II.3: Difference in the population dynamics with a varying and a constant carrying capacity κ for a) a release rate close to ξ° and b) a substantially higher release rate.

egg or larval stages which would need to be included in an extended model for situations where this is of frequent occurrence (such as where there is considerable seasonality driven by temperature or rainfall).

6. Optimal control solutions

For a constant release rate, ξ° denotes the minimal number of sterile males that must be released every day to eliminate the population. Yet ξ° alone does not account for the time needed for the population to fall below the threshold value $\tilde{E}_2$. Moreover, a constant release rate is not optimal, since there are obviously better strategies that entail adjusting the release to the number of females.

6.1 Problem formulation for the single-site model

The framework for determining the optimal release rate for the single-site model (3.2) is for example formulated in Evans [Evans, 1983]. We consider the system of ordinary differential equations

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}u(t) = f(u(t), \psi(t)), & \text{for } t > 0, \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\psi(t) = g(\psi(t), \xi(t)), & \text{for } t > 0, \\ u(0) = u_0 \in [0, \kappa], \\ \psi(0) = 0. \end{cases}$$

The goal is to minimize the population u on the time interval $[0, T]$ and to find the optimal

release rate of sterile mosquitoes $\xi^\star$ under certain aspects. To this end, we consider the payoff functional

$$J[\xi] := \frac{1}{2}\int_0^T Cu(t)^2 + \xi(t)^2 \ \mathrm{d}t,$$

and seek a function $\xi^\star$ such that $J[\xi^\star] \leqslant J[\xi]$ for all controls ξ with $\xi \in \mathcal{X}$ and

$$\mathcal{X} := \{\xi\colon [0,\infty) \to [0,M] \mid \xi \text{ is measurable}\}.$$

The payoff functional $J[\xi]$ is chosen following the examples in the book by Lenhart and Workman [Lenhart and Workman, 2007] (who provide an extensive list of applications of optimal control theory to biological problems). Since the ODE is nonlinear, we consider a quadratic functional. Other choices are discussed in A.1. The constant C weighs the importance of minimizing u against minimizing ξ. We now seek a function $\xi^\star$ that is the minimal solution for $J[\xi]$ among all functions that fulfill

$$0 \leqslant \xi \leqslant M,$$

where M is the maximal available number of sterile males per day. The function $\xi^\star$ is called the control and the corresponding $u^\star$ and $\psi^\star$ the response. Necessary conditions for optimality of a solution for the optimal control problem

$$\min_{\xi(t)} \left\{ \frac{1}{2}\int_0^T Cu(t)^2 + \xi(t)^2 \ \mathrm{d}t \right\} \tag{6.1}$$

subject to

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}u(t) = f(u(t),\psi(t)), & \text{for } t > 0, \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\psi(t) = g(\psi(t),\xi(t)), & \text{for } t > 0, \\ u(0) = u_0 \in [0,\kappa], \\ \psi(0) = 0, \\ 0 \leqslant \xi(t) \leqslant M, \end{cases} \tag{6.2}$$

are derived in the Pontryagin maximum principle [Evans, 1983, Theorem 4.3]. In this setting it can be stated as follows, omitting the dependence on t for all the functions:

(6.3) **Theorem.** Assume that $\xi^\star$ is optimal for problem (6.1) subject to (6.2). Let $u^\star$ and $\psi^\star$ be the corresponding responses and define

$$H(u,\psi,\lambda_1,\lambda_2,\xi) := f(u,\psi)\lambda_1 + g(\psi,\xi)\lambda_2 + \frac{1}{2}(Cu^2 + \xi^2)$$

for $u, \psi, \lambda_1, \lambda_2 \in \mathbb{R}$, $\xi \in \mathcal{X}$. Then there exist functions $\lambda_1^\star \colon [0,T] \to \mathbb{R}$ and $\lambda_2^\star \colon [0,T] \to \mathbb{R}$ such that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} u^\star &= \nabla_{\lambda_1} H(u^\star, \psi^\star, \lambda_1^\star, \lambda_2^\star, \xi^\star) = f(u^\star, \psi^\star), \\ \frac{\mathrm{d}}{\mathrm{d}t} \psi^\star &= \nabla_{\lambda_2} H(u^\star, \psi^\star, \lambda_1^\star, \lambda_2^\star, \xi^\star) = g(\psi^\star, \xi^\star), \end{aligned} \tag{6.4}$$

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \lambda_1^\star &= -\nabla_u H(u^\star, \psi^\star, \lambda_1^\star, \lambda_2^\star, \xi^\star) = -\nabla_u f(u^\star, \psi^\star)\lambda_1^\star - C u^\star, \\ \frac{\mathrm{d}}{\mathrm{d}t} \lambda_2^\star &= -\nabla_\psi H(u^\star, \psi^\star, \lambda_1^\star, \lambda_2^\star, \xi^\star) = -\nabla_\psi g(\psi^\star, \xi^\star)\lambda_2^\star, \end{aligned} \tag{6.5}$$

$$H(u^\star, \psi^\star, \lambda_1^\star, \lambda_2^\star, \xi^\star) = \min_{\xi \in \mathcal{X}} H(u^\star, \psi^\star, \lambda_1^\star, \lambda_2^\star, \xi), \quad \text{for } 0 \leqslant t \leqslant T. \tag{6.6}$$

Finally, there are the terminal conditions

$$\lambda_1^\star(T) = 0, \quad \lambda_2^\star(T) = 0.$$

Proof. See Evans [Evans, 1983, Theorem 4.3]. □

The identities (6.5) are called the adjoint equations and (6.6) is the minimization principle. Further information is deduced from equation (6.6) by solving the necessary condition for a local minimum,

$$\nabla_\xi H(u^\star, \psi^\star, \lambda_1^\star, \lambda_2^\star, \xi) = \nabla_\xi g(\psi^\star, \xi)\lambda_2^\star + \xi = \lambda_2^\star + \xi = 0.$$

This results in $\xi^\star = -\lambda_2^\star$, that can be explicitly computed after solving the equations (6.4) and (6.5)

6.2 Numerical results for the single-site model

One way of implementing theorem (6.3) is the Forward-Backward Sweep method [Lenhart and Workman, 2007]. The basic algorithm roughly consists of five steps, as shown in table II.2, dropping the * notation. Details for the implementation are provided by Lenhart and Workman [Lenhart and Workman, 2007]. We implement the algorithm in MATLAB R2016b, using a fourth-order Runge-Kutta method. We choose $\xi = 0$ as an initial guess in Step 1. To speed the convergence, we use a convex combination of the previously computed ξ and the current characterization in Step 4 and use a standard convergence criterion [Lenhart and Workman, 2007, Chapter 4].
We choose $C = 15/\mu^2$, since the elimination of the population is weighted much higher

Step 1	Make an initial guess for ξ over the time interval $[0, T]$.
Step 2	Solve the equations (6.4) forwards in time, using ξ and the initial conditions u_0 and ψ_0.
Step 3	Solve the equations (6.5) backwards in time, using ξ, u, ψ and the terminal conditions $\lambda_1(T) = 0$ and $\lambda_2(T) = 0$.
Step 4	Update ξ according to (6.6) by using the results from Steps 2-3 for u, ψ, λ_1 and λ_2 and enforce the condition $0 \leqslant \xi \leqslant M$ to hold.
Step 5	Repeat the Steps 2-4 until a convergence criteria is met.

Table II.2: Forward-Backward Sweep algorithm.

(15 times) than the number of released sterile males. The factor μ^2 has a scaling purpose, such that u^2 and $\xi(t)^2$ are of the same magnitude in the payoff functional. Other values for C are possible, the higher, the cheaper are high values of ξ. The choice $C = 15/\mu^2$ yielded a good trade off between a fast elimination and a reasonable amount of sterile males. If $C = 1$, no elimination is observed, since the easiest way to minimize the payoff functional is to release no sterile males. We simulate the population for $T = 80$ days and restrict the daily available mosquitoes to $M = 10\,000$. The result is shown in figure II.4. The optimal

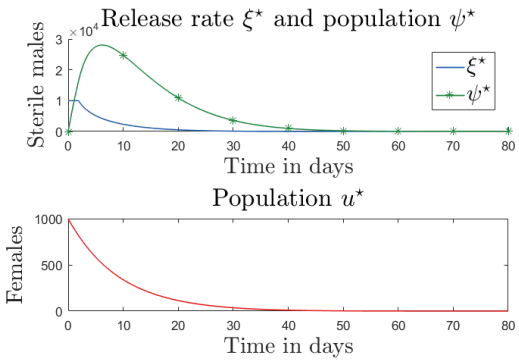


Figure II.4: Optimal control solution for the single-site model. The upper graph illustrates the release rate of sterile males together with the population of sterile males arising from the release. The lower graph shows the corresponding population of wild females.

way of eliminating the population u is to release a large number of sterile males at the beginning and as the population decreases, make smaller releases almost proportionally to the population. The same dynamics are observed for other choices of κ.
In subsection 5.3 it was stated that the population $u(t)$ has to fall below a threshold value $\tilde{E}_2$ before releases can stop. Here, we find that it takes $t^* \approx 60$ days until $u^\star(t^*) < \tilde{E}_2$. A comparison shows that if

$$\max_{0 \leqslant t \leqslant T} \xi^\star(t)$$

are released, elimination is achieved ~ 7 days earlier, as shown in figure II.5. However, many more mosquitoes must be released in this case.

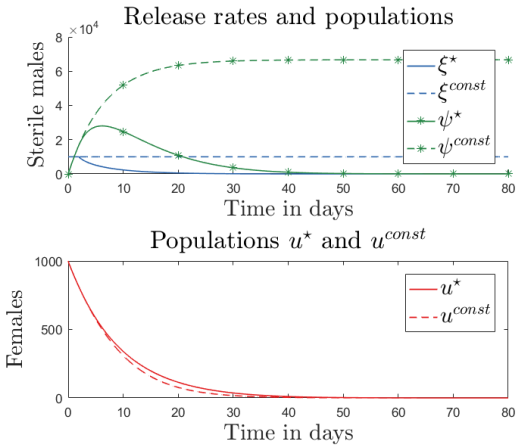


Figure II.5: Comparison of the optimal control solution with a constant release. The graphs are built in the same way as explained in figure II.4.

6.3 Optimal control problem formulation for the PDE model

In practice, there is a fixed number of sterile males available per day to allocate across the whole island requiring the application of optimal control theory to the spatial model (3.3). The Pontryagin maximum principle also holds for PDEs [Tröltzsch, 2010] and we can compute the solution with the Forward-Backward Sweep described in table II.2. However, this requires solution of a PDE forwards and backwards in time, which fundamentally complicates the numerics. We again consider the payoff functional

$$J[\xi] = \frac{1}{2}\int_0^T \int_\Omega C u(t,x)^2 + \xi(t,x)^2 \ \mathrm{d}x \ \mathrm{d}t,$$

and seek a function $\xi^\star(t,x)$ such that $J[\xi^\star] \leqslant J[\xi]$ for all controls ξ with $\xi \in \mathcal{X}$ and

$$\mathcal{X} := \{\xi \colon [0,\infty) \times \Omega \to [0,M] \mid \xi \text{ is measurable}\}.$$

The optimal control problem therefore reads

$$\min_{\xi(t,x)} \left\{ \frac{1}{2}\int_0^T \int_\Omega C u(t,x)^2 + \xi(t,x)^2 \ \mathrm{d}x \ \mathrm{d}t \right\} \tag{6.7}$$

subject to

$$
\text{(6.8)} \qquad \begin{cases} \dfrac{\partial u}{\partial t} = D\Delta u + f(u,\psi), & \text{for } x \in \Omega,\ t > 0, \\ \dfrac{\partial \psi}{\partial t} = D\Delta \psi + g(\psi), & \text{for } x \in \Omega,\ t > 0, \\ u(0,x) = u_0(x),\ \psi(0,x) = 0, & \text{for } x \in \Omega, \\ \text{Periodic boundary conditions}, & \text{for } x \in \partial\Omega, \\ 0 \leqslant \xi(t,x) \leqslant M, & \text{for } x \in \Omega,\ t > 0. \end{cases}
$$

The pointwise condition $0 \leqslant \xi(t,x) \leqslant M$ states that every point can only use a certain number M of sterile males. Optimally, the problem should be formulated with an isoperimetric constraint of the form

$$\int_\Omega \xi(t,x)\ \mathrm{d}x \leqslant M,$$

stressing the fact that there is a fixed number M for the whole island to allocate. However, this constraint is beyond the scope of this analysis, as it is not clear whether the Pontryagin maximum principle holds. Ding *et al.* [Ding et al., 2010] show that an optimal control problem for an elliptic PDE with such a constraint can, by a transformation, be led back to a form where the Pontryagin maximum principle holds, but the time dependency in this model makes this approach infeasible.
We state the Pontryagin maximum principle following Trölzsch [Tröltzsch, 2010, Section 5.6], but keeping our introduced notation.

(6.9) **Theorem.** Assume that $\xi^\star(t,x)$ is optimal for problem (6.7) subject to (6.8). Let $u^\star(t,x)$ and $\psi^\star(t,x)$ be the corresponding responses and again define

$$H(u,\psi,\lambda_1,\lambda_2,\xi) := f(u,\psi)\lambda_1 + g(\psi,\xi)\lambda_2 + \frac{1}{2}(Cu^2 + \xi^2)$$

for $u,\psi,\lambda_1,\lambda_2 \in \mathbb{R},\ \xi \in \mathcal{X}$.
Then there exist functions $\lambda_1^\star\colon [0,T]\times\Omega \to \mathbb{R}$ and $\lambda_2^\star\colon [0,T]\times\Omega \to \mathbb{R}$ such that

$$
\text{(6.10)} \qquad \begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}u^\star &= D\Delta u^\star + \nabla_{\lambda_1} H(u^\star,\psi^\star,\lambda_1^\star,\lambda_2^\star,\xi^\star) = D\Delta u^\star + f(u^\star,\psi^\star), \\ \frac{\mathrm{d}}{\mathrm{d}t}\psi^\star &= D\Delta \psi^\star + \nabla_{\lambda_2} H(u^\star,\psi^\star,\lambda_1^\star,\lambda_2^\star,\xi^\star) = D\Delta \psi^\star + g(\psi^\star,\xi^\star), \end{aligned}
$$

$$
\text{(6.11)} \qquad \begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\lambda_1^\star &= -D\Delta\lambda_1^\star - \nabla_u H(u^\star,\psi^\star,\lambda_1^\star,\lambda_2^\star,\xi^\star) = -D\Delta\lambda_1^\star - \nabla_u f(u^\star,\psi^\star)\lambda_1^\star - Cu^\star, \\ \frac{\mathrm{d}}{\mathrm{d}t}\lambda_2^\star &= -D\Delta\lambda_2^\star - \nabla_\psi H(u^\star,\psi^\star\lambda_1^\star,\lambda_2^\star,\xi^\star) = -D\Delta\lambda_2^\star - \nabla_\psi g(\psi^\star\xi^\star)\lambda_2^\star, \end{aligned}
$$

$$H(u^\star, \psi^\star, \lambda_1^\star, \lambda_2^\star, \xi^\star) = \min_{\xi \in \mathcal{X}} H(u^\star, \psi^\star, \lambda_1^\star, \lambda_2^\star, \xi), \quad \text{for } 0 \leqslant t \leqslant T. \tag{6.12}$$

Finally, there are terminal conditions

$$\lambda_1^\star(T) = 0, \quad \lambda_2^\star(T) = 0$$

and periodic boundary conditions on Ω.

Proof. See Trölzsch [Tröltzsch, 2010, Section 5.6]. □

Note that the definition of the Hamiltonian $H(u, \psi, \lambda_1, \lambda_2, \xi)$ is the same, but the negative sign extends to the stationary part of the adjoint equations.

7. Numerical results

We consider various release strategies for a SIT on an island, ranging from a very basic strategy to optimal control strategies and compare them in number of released mosquitoes and time needed for elimination.

The PDE (3.3) is solved with a Finite Element Method (FEM) approach in space and a θ-scheme in time. The results are computed with $\theta = 0.5$, the Crank-Nicolson method. To deal with the dependency of $f(u, \psi)$ and $g(\psi)$ on u and ψ a fixed point iteration is used. The parameters are chosen as described in subsection 4 and the island Ω is simulated for $T = 80$ days. To avoid instability caused by the large size of the modeled island relative to the rate of *Aedes aegypti* movement the problem is rescaled as described in A.2. The implementation of the optimal control solution is again based on the Forward-Backward Sweep method introduced in subsection 6.2 (table II.3). Note that in Step 3,

Step 1	Make an initial guess for $\xi(t, x)$ over the time interval $[0, T]$ and the domain Ω.
Step 2	Solve the equations (6.10) forwards in time, using $\xi(t, x)$ and the initial conditions $u_0(x)$ and $\psi_0(x)$. This is accomplished with FEM in space and a θ-scheme in time.
Step 3	Solve the equations (6.11) backwards in time, using $\xi(t, x)$, $u(t, x)$, $\psi(t, x)$ and the terminal conditions $\lambda_1(T, x) = 0$ and $\lambda_2(T, x) = 0$.
Step 4	Update $\xi(t, x)$ according to (6.12) by using the results from Steps 2-3 for $u(t, x)$, $\psi(t, x)$, $\lambda_1(t, x)$ and $\lambda_2(t, x)$ and enforce the condition $0 \leqslant \xi(t, x) \leqslant M$ to hold.
Step 5	Repeat the Steps 2-4 until a convergence criteria is met.

Table II.3: Forward-Backward Sweep algorithm for the PDE model.

a parabolic equation has to be solved backwards in time. This is only well defined for

a small neighborhood around $t \in [0, T]$. By taking small time steps and by using the fixed point iteration we can get a solution on the whole domain $[0, T]$ backwards in time. Another possibility would be to use an operator splitting method [Quiroga et al., 2015]. We experimentally choose $C = 10^5$. Apart from weighting the relative importance of each control, C is a term for the regularization of the problem. If C is chosen too small, convergence is not observed. The same happens if M is chosen too small. When the release is constrained to less than 5 times the maximum of the carrying capacity, convergence is no longer observed.

Constant release in time

An initial simulation considers a constant release in time of

$$\xi(x) = 10\kappa(x), \quad \text{for } x \in \Omega,$$

sterile males over the domain of the island. The carrying capacity $\kappa(x)$ includes a spatial variation of the female population representing four villages at equal distances around the island. The intervention is discarded when the population has fallen under the threshold value $\tilde{E}_2$ determined in the previous subsection (figure II.6).

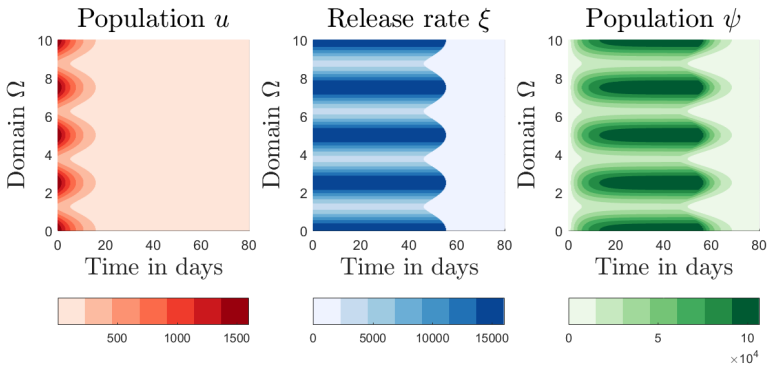


Figure II.6: Constant release of sterile males in time; Displayed on the y-axis of every figure is the island Ω, connected at its endpoints. Figure a) visualizes the population of female mosquitoes u with the initial condition of four villages. In figure b) and c) the release rate of sterile males, $\xi(x)$, and the population of sterile males, $\psi(x)$, are visualized. The shift of the colors indicates the accumulation of sterile males as the colorbar indicates the number of mosquitoes.

Optimal control strategy

The release strategy analyzed in the previous section is clearly not optimal, as it is constrained to be constant in time. For the optimal control solution allowing temporal variation, we choose

$$M = 10 \max_{x \in \Omega} \kappa(x),$$

such that $\xi(t, x)$ is limited to M for every $x \in \Omega$ and $t \in [0, T]$. The solution is then to initially release a substantial number of mosquitoes so that the number of sterile males accumulates, and then to release a smaller number to sustain the impact (figure II.7). This

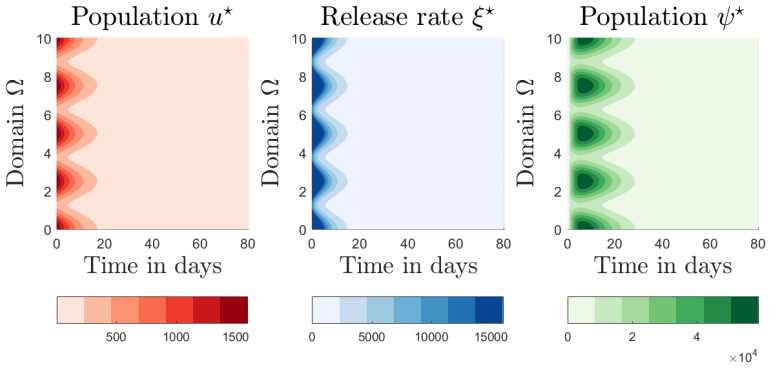


Figure II.7: Optimal control solution for the PDE, where $\xi(t, x)$ now depends on t and x. The construction of the graphs is the same as in figure II.6; On the left, the population of wild females is illustrated, together with the optimal control release of sterile males and the corresponding population.

result is very similar to the optimal control solution for the single-site model and optimal if there are sufficient resources.

Optimized strategies for a limited daily number of sterile males

In practice, the total number of sterile males available at any one time will be constrained, so there are unlikely to be enough to implement the optimal control strategy everywhere at once. We propose a variation of $\xi^\star(t, x)$, in which the decrease over time in release rate is constrained to be equal to that from the optimal control solution and the starting time of releases in every location is optimized. Once the program has started, the releases in new locations are delayed until there are enough sterile males available to introduce the intervention to other places.
The possible strategies within this constraint can be classified according to

(i) whether the zones between the initial locations are included incrementally by spreading out in one direction around the island, or in both directions from the initial location(s),

(ii) how many locations are initially included,

(iii) whether the initial locations should be high or low mosquito density areas.

An immediate result is that with a strategy of spreading out in only one direction the effect of the intervention disappears in a very short time (figure II.8). Once introduced

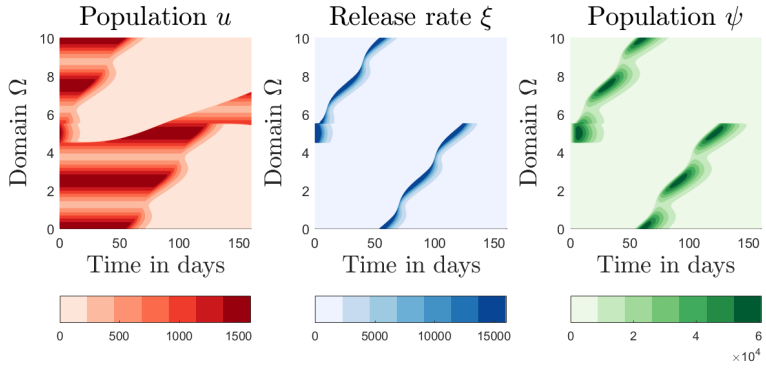


Figure II.8: Optimized release strategy for $\xi(t, x)$ that does not prevent reintroduction. The figure is constructed in the same way as figure II.6.

to a location, a strategy has to be spread out in both directions. Figures II.9 and II.10 illustrate the effect of the number of initially included locations. In each case the initial interval is divided into one, two or four locations starting either in the villages (i.e. the locations with higher densities) (figure II.9) or in the places with lower mosquito densities between the villages (figure II.10). The more sterile males are available per day, the more the plots resemble the findings from the spatial optimal control solution.
A good indicator to compare these strategies is the number of days needed to stop the intervention because the population has fallen below the threshold value $\tilde{E}_2$ (table II.4). The number of released sterile males before the intervention can be stopped (Σ) remains similar in the optimal control solution and the optimized strategies, variations arise due to diffusion. In this highly symmetrical situation, the strategies that release in one or two initial intervals are performing similarly and are overall faster (δ_2) if they start in low density areas. The strategy that releases in every village is faster when the intervention starts at high density areas. The common denominator of all the well performing strategies is that high density areas are dealt with earlier, making the intervention overall faster. This conclusion is underlined by simulations with different initial conditions and more

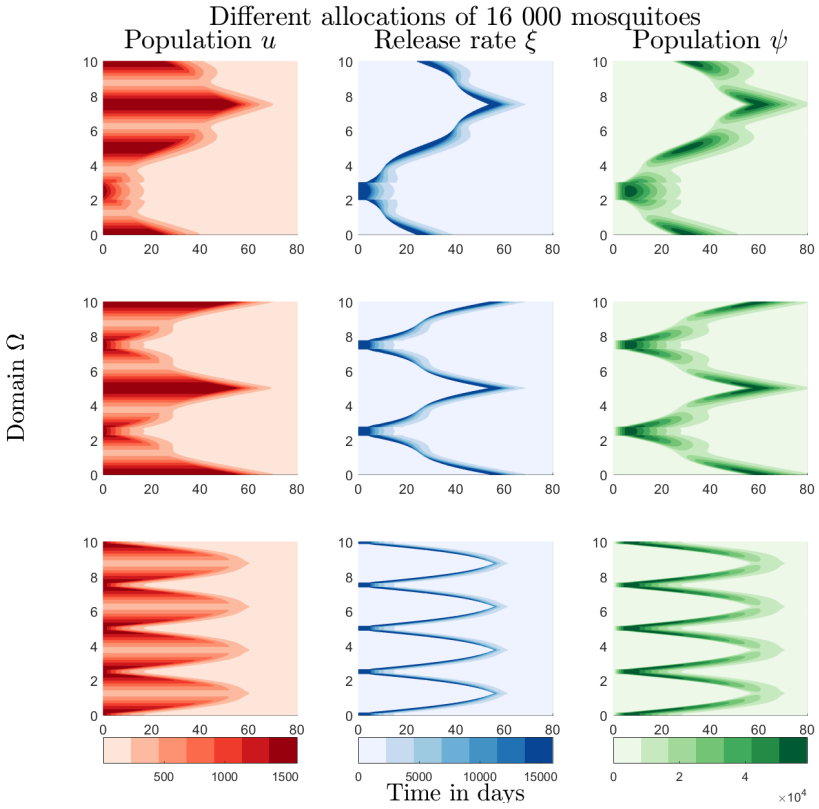


Figure II.9: Optimized release strategy for an initial release interval of 1km, resulting in an allocation of 16 000 mosquitoes. The release starts in high density areas. For a detailed description of the reading of the figure types, see figure II.6.

heterogeneous densities of female mosquitoes (figure II.11). Again comparing the number of days needed to stop the intervention shows that for this initial condition, a jump to a lower density area slows the intervention. The fastest strategy is to target the two higher density areas together (bottom of table II.4).

8. Discussion

This paper focuses on finding an optimal strategy for the allocation of a SIT intervention on an island, exemplified by its perimeter (on the assumption that the population of the island is coastal) under the constraint of a limited daily number of sterile males available. We constructed a PDE model for the dynamics of an *Aedes aegypti* population and applied optimal control theory to determine an optimal strategy of release of sterile males. The optimal control solution for the island (figure II.7) is to release mosquitoes everywhere, proportionately to the size of the endemic population if the resources are available. This result is intuitive and consistent with the results of the two-patch Dengue model of Kim *et*

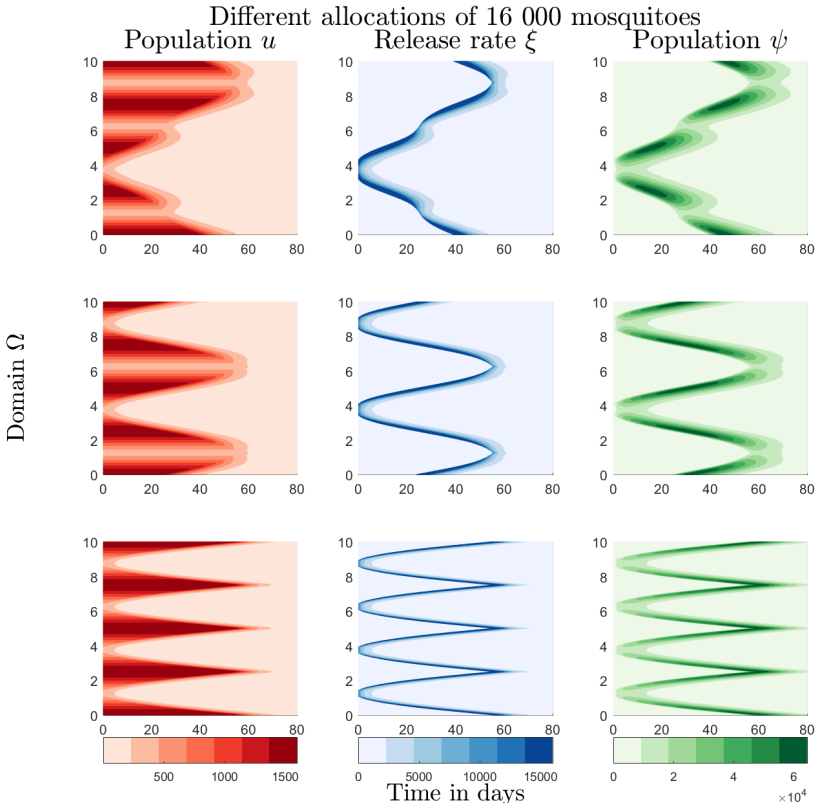


Figure II.10: Optimized release strategy starting in lower density areas, again with an allocation of 16 000 mosquitoes. For a detailed description of the reading of the figure types, see figure II.6.

al. [Kim et al., 2017]. It also coincides with the conclusion of Ferreira *et al.* [Ferreira et al., 2008], that spatially heterogeneous mosquito populations make it challenging to design a SIT intervention. The spatial optimal control solution follows the same procedure as the single-site solution in time with an initial high ratio of sterile males to females and then a decrease, proportional to the endemic population and is hence independent of the particular choice of the domain length L. Real programs face a constraint in the numbers of mosquitoes available to be released, and this constraint makes the general solution of the optimal control problem intractable.

We therefore propose an optimized strategy that builds on this result but includes a delay of the intervention in some areas to account for a limited daily number of sterile males (also exemplified in figure II.12):

(i) Identify high mosquito density areas on the island, together with their width and the number of sterile males available (M). For an optimal intervention effect, every area should be targeted with approximately 10 times the number

Strategy	Σ	δ_1 days	δ_2 days
Proportional, figure II.6	5 286 000	0	55.6
Optimal control, figure II.7	958 420	0	60.4
Optimized strategies:			
Daily limit 16 000, start at high density, figure II.9			
One initial interval	958 640	54.1	112.7
Two initial intervals	957 600	54.1	113.7
Four initial intervals	955 670	56.2	103.0
Daily limit 16 000, start at low density, figure II.10			
One initial interval	958 080	55.2	105.6
Two initial intervals	957 920	55.7	103.0
Four initial intervals	961 070	55.1	114.5
Heterogeneous initial distribution, daily limit 16 000, figure II.11			
Optimal control	608 350	0	60.1
One initial interval	607 760	33.0	85.1
Two initial intervals	606 750	33.3	83.5
Jump of intervention	607 020	34.3	85.9

Table II.4: Comparison of different strategies, based on the total number of released sterile males before the intervention can be stopped (Σ), the time until the intervention reaches every point (δ_1) and the time to elimination in every point (δ_2).

of females present. This proposed initial ratio of release (10:1) assumes perfect mating efficiency of the sterile males as well as perfect separating of sterile males and females. If this is not the case, this ratio has to be adjusted by a constant factor. Should the number of available sterile males exceed the number needed to target the highest density areas, more areas can be included. Else, the width of the areas can be made smaller.

(ii) In the identified areas, sterile males should initially be released with a high ratio of effective sterile males to females (we propose a ratio of 10:1), and then be withdrawn while maintaining approximately the same ratio to the decreasing endemic population (see result in figure II.4).

(iii) As soon as less sterile males are used in the targeted areas, move outwards (in both directions along the periphery of the island), until all areas have been covered, retaining throughout sufficient release intensity to prevent reintroduction in the already cleared areas. Hence, the last areas to be reached should be of low density.

The simulations have shown that more high density areas to start with are beneficial, compared to one (wider) high density area. This result is in line with the reported findings

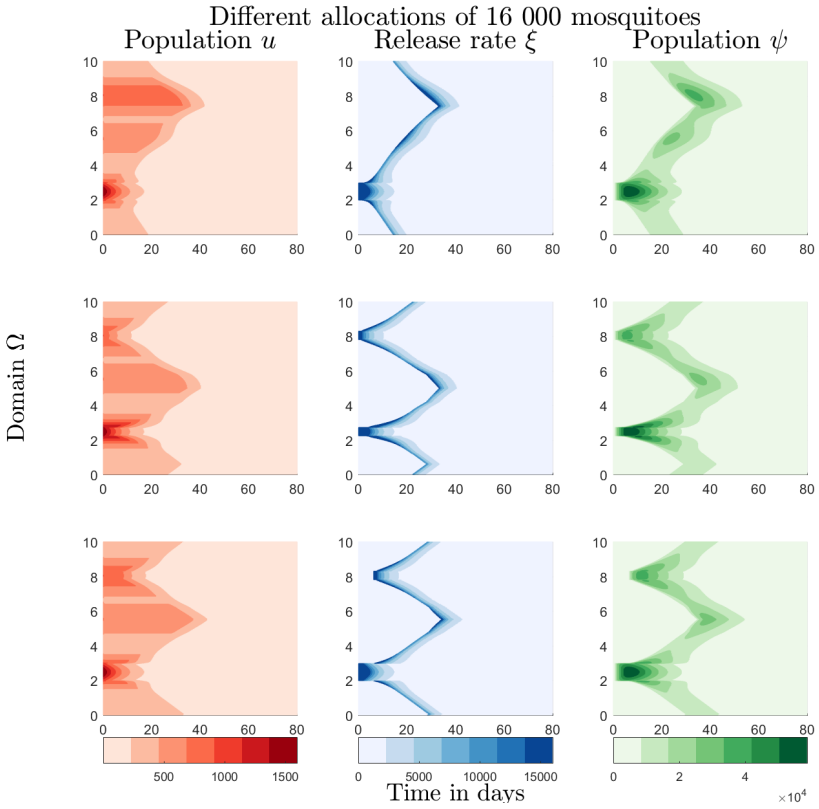


Figure II.11: Optimized release strategy with a more heterogeneous initial condition, representing three villages of different size; a) initial release in the highest density area, b) initial release in two higher density areas, c) initial release in the highest density area with a jump to a lower density area.

of Oléron *et al.* [Oléron Evans and Bishop, 2014] who use a spatial model and couple it with pulsed releases. They point out that, given a fixed quantity of sterile males available for release per unit of time for a given area, increasing the frequency of releases and the number of release sites improves the effectiveness of the intervention. A drop in the carrying capacity due to temporal variations can be used as an optimal start for a program, for if there are fewer female mosquitoes, fewer have to be eliminated [Ferreira et al., 2008]. Apart from that, simulations showed that the dynamics of seasonality did not influence the intervention if the number of released sterile males was high enough. This is not surprising, considering that the optimal control solution showed that the main impact of the intervention happened in the first three weeks. Seasonal variations act similar to time limited vector control interventions, and hence, a drop in the carrying capacity due to insecticide control can also be considered for a start of a SIT intervention. Thomé *et al.* [Thomé et al., 2010] also come to this conclusion, applying optimal control theory to find the best trade off between chemical control and a SIT intervention. They conclude

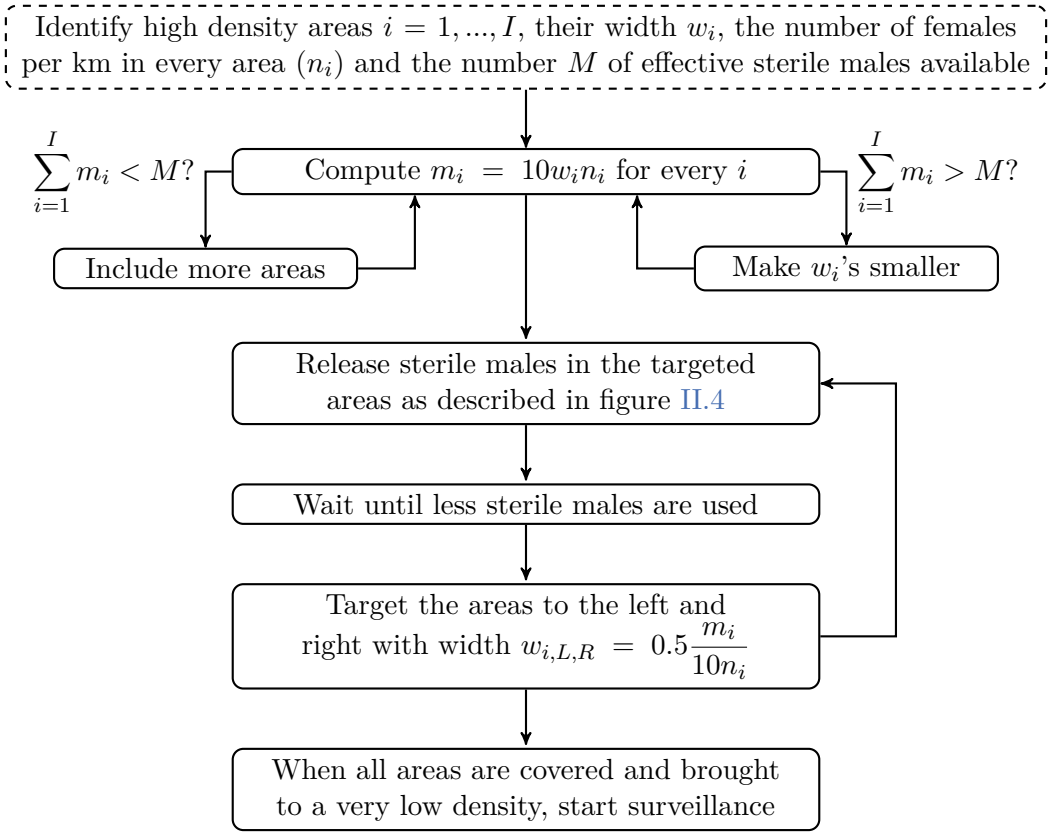


Figure II.12: Diagram visualizing the proposed optimized strategy.

that a high application of insecticide is needed at the beginning, leading to a drop in the carrying capacity that is beneficial to the intervention because the number of mosquitoes has decreased.

To model the stop of a SIT intervention without the population artificially jumping back to the carrying capacity, we incorporated an Allee effect of Holling-II type. It was then possible to state a threshold value $\tilde{E}_2$, given in terms of the chosen parameters, that is an indicator of how small the population has to become before the release can successfully be stopped. The constant $\tilde{E}_2$ attained a value below one for the chosen parameter set, implying that the population has to be completely eliminated before the release can stop. Varying the value of the parameter responsible for the Allee effect changes the required duration of the elimination efforts and total number of mosquitoes needed but does not affect the general conclusions. Regardless, this suggests that population replacement with the sterile males would need to be close to complete for the intervention to be successful, and a successful program will also have to ensure that sterile males are released near enough to all breeding sites so that they have a chance of mating with wild females. The model considers the effect on numbers of winged stage of mosquitoes, reasoning that

once these are eliminated no more larval stages can arise. A possible complication is the survival of dormant eggs from before the introduction of the program. This might be particularly important in settings with seasonal variation in temperatures or rainfall, where the capacity for dormancy is essential for mosquito survival. This, together with the need to guard against reintroductions and operational imperfections, (such as incomplete separation of females from sterile males in the production site), implies that effective surveillance would continue to be needed after the adult mosquito population has been removed. A response capacity for implementing further releases, maybe in form of a permanent barrier of sterile males as was done for the Screwworm fly, and/or other control measures must be held in place.

The proposed model is a simple predator-prey model with a diffusion term. It can easily be parameterized for other species and other domains than an island. The simplicity of the model makes the analysis tractable and possible, at least for the ODE system. A simple model cannot do justice to every complexity, but can highlight one particular aspect of a problem, in this case the question of an effective and successful release strategy of sterile males. Many factors influence the outcome of such a program: the dispersal ability of the female mosquitoes and the resulting rapid reimmigration are amongst the major obstacles to be faced. Nevertheless, if carefully planned and aligned with the geography of the targeted area, a SIT intervention may contribute substantially to eliminating a mosquito population, especially on closed domains like an island.

Chapter III

ANALYSIS OF CONTAMINATION IN CLUSTER RANDOMIZED TRIALS OF MALARIA INTERVENTIONS

Lea Multerer[1,2] , Tracy R. Glass[1,2] , Fiona Vanobberghen[1,2] , Thomas Smith[1,2]

[1] Swiss Tropical and Public Health Institute, Basel, Switzerland
[2] University of Basel, Basel, Switzerland

Published in *Trials* in September 2021. (DOI: 10.1186/s13063-021-05543-8)

1. Abstract

In cluster randomized trials (CRTs) of interventions against malaria, mosquito movement between households ultimately leads to contamination between intervention and control arms, unless they are separated by wide buffer zones.
This paper proposes a method for adjusting estimates of intervention effectiveness for contamination and for estimating a contamination range between intervention arms, the distance over which contamination measurably biases the estimate of effectiveness. A sigmoid function is fitted to malaria prevalence or incidence data as a function of the distance of households to the intervention boundary, stratified by intervention status and including a random effect for the clustering. The method is evaluated in a simulation study, corresponding to a range of rural settings with varying intervention effectiveness and contamination range, and applied to a CRT of insecticide treated nets in Ghana.
The simulations indicate that the method leads to approximately unbiased estimates of effectiveness. Precision decreases with increasing mosquito movement, but the contamination range is much smaller than the maximum distance traveled by mosquitoes. For

the method to provide precise and approximately unbiased estimates, at least 50% of the households should be at distances greater than the estimated contamination range from the discordant intervention arm.
A sigmoid approach provides an appropriate analysis for a CRT in the presence of contamination. Outcome data from boundary zones should not be discarded but used to provide estimates of the contamination range. This gives an alternative to "fried egg" designs, which use large clusters (increasing costs) and exclude buffer zones to avoid bias.

2. Background

Cluster randomized trials (CRTs) are often used in public health research to avoid contamination effects (also called indirect effects or spill-over effects) leading to averaging of estimates of effectiveness across the arms of a trial population in an individual-level randomized trial. The full effect of the intervention is only observed in comparisons of distinct clusters of individuals, but it may be difficult to ensure full separation between the intervention and control arms of the trial. This problem has long been recognized in the design of CRTs, especially for vaccine studies [Halloran and Struchiner, 1991, Halloran et al., 2010, Hussey and Hughes, 2007, Jarvis et al., 2017]. With directly transmitted diseases, dynamic models of the transmission across contact networks can provide an efficient, though technically challenging approach to optimizing trial design and estimating effects of contamination on effectiveness estimates [Staples et al., 2015].
With diseases transmitted by vectors, construction of contact networks is usually impossible and clusters are defined to correspond to the places where people get infected. In the case of *Aedes* transmitted diseases like dengue or zika, these may be schools or workplaces, since biting happens during the day. However, *Anopheles* mosquitoes transmitting malaria bite in the early night and early morning. Hence, most transmission of malaria is indoors or peri-domestic and can be geolocated to the host's primary residence. In trials of interventions, such as the deployment of insecticides or distribution of bed nets, clusters are therefore defined as geographically congruent areas, with contamination effects mainly induced by mosquito movement because people living nearby might benefit from a reduced density of infectious mosquitoes. Other contamination effects that are unrelated to geographical distance, such as relocation of human hosts, are relatively unimportant. The maximum effect of intervention is then observed only where high coverage is achieved throughout a substantial group of neighboring individuals. Since *Anopheles* mosquitoes can fly several kilometers [Verdonschot and Besse-Lototskaya, 2014], trial arms need to be separated by large distances if contamination at cluster boundaries is to be avoided. This has led to CRTs with clusters of much larger geographical size than are required to

estimate the effect of the intervention with wide buffer zones around each cluster where the intervention is introduced but excluded from data collection and analysis (a so-called "fried egg" design [Hayes and Moulton, 2009, Wolbers et al., 2012, Protopopoff et al., 2015, Protopopoff et al., 2018, Eisele et al., 2016, Delrieu et al., 2015]).

With the fried egg design, a simple mixed effects model provides a valid analysis [Hayes and Moulton, 2009], providing the buffer zone is large enough. But because the intervention must be introduced in the buffer zone, the trial may be very expensive if there are high per capita intervention costs. Since the buffer zone is excluded from data collection, there are usually no data on whether the buffer is large enough to avoid contamination effects, and an unexpectedly large contamination leading to substantial bias in the estimate of effect would go undetected. These considerations challenge the rationale for fried egg designs. Recently, an alternative was proposed to a simple fried egg design by either fully including or excluding clusters from both the intervention assignment and the analysis based on a criterium of closeness between households to attain a better separation between intervention and control arms [McCann et al., 2018]. This approach leads to smaller trials and a conventional analysis can be carried out, but, depending on the proximity of clusters, is very computationally expensive and information on the contamination range is still needed to design such a trial.

There are reasons why contamination effects should be measured [Halloran et al., 2017, Baird et al., 2018, McCann et al., 2018]. Evidence on contamination effects supports inference about indirect effects of the intervention, thus analyses of contamination in CRTs of Insecticide Treated Nets (ITNs) against malaria [Binka et al., 1998, Howard et al., 2000, Hawley et al., 2003] fed into the rationale for massive distribution of the nets across Africa. In the largest trial in Asembo, Kenya [Hawley et al., 2003], significant protective effects of ITNs were found for distances of up to 300 m from cluster boundaries, while on the coast effects persisted for distances of up to 1.5 km [Howard et al., 2000]. In these analyses, a linear model was extended to include a term of the distance to the nearest discordant observation. Nevertheless, it is not possible to obtain a closed-form range that specifies the maximal measurable extent of contamination from a linear model. Methods that jointly estimate contamination effects and adjust the estimate of effectiveness accordingly are needed. Neither the maximum distance that mosquitoes can fly, nor the distance over which contamination effects can be measured, necessarily equates with the distance over which contamination between trial arms is statistically relevant, and if contamination only biases the intervention effects over short distances then clusters could be smaller. This could lead to more cost-efficient, smaller trials while adding a new outcome measure to the analysis of CRTs.

This work proposes an approach for simultaneously estimating the intervention effective-

ness and the contamination range, defined as the extent of measurable contamination across the intervention boundaries in CRTs of malaria interventions. Simulations of CRTs of malaria interventions targeting mosquito densities and measuring prevalence as the outcome, for example with a rapid diagnostic test (RDT), were used to assess the model performance. Simulated mosquito movement leads to correlations between households and hence to contamination between intervention arms. The degree of mosquito movement, intervention efficacy, numbers of clusters, households per cluster and the pattern of spatial clustering in both the human and vector populations rates were varied. A reanalysis of a CRT for assessing the effects of ITNs on child mortality in the Kassena-Nankana district in northern Ghana (Navrongo trial) with the proposed method illustrates the findings [Binka et al., 1998, Binka et al., 1996].

3. Methods

3.1 Simulation of CRTs with contaminated intervention effects

The simulations of cluster randomized trials entailed generating simulated human populations at the household level, assigning disease distribution in the absence of intervention and implementation of intervention effects as follows:

Human populations and disease distribution in the absence of intervention

To approximate patterns of heterogeneous human dispersion, simulated human populations living in N households on a domain of $\eta \times \eta\,\mathrm{km}^2$ were generated via a (modified) Thomas cluster process [Thomas, 1949] (table III.1). This algorithm generates a uniform Poisson point process of parent points with intensity α_1, the cluster centers, and then replaces each parent point with a bivariate normally distributed cluster of offspring points, the households, with a mean number of points per cluster α_2 and a standard deviation of random displacement of a point from its parent α_3. This algorithm is implemented in the R-package spatstat [Baddeley and Turner, 2005] via the function rThomas.
The simulated transmission potential was a smooth function in space with local maxima at a simple random sample of τ index households. Each household j with coordinates x_j and y_j was assigned a local infection rate or vectorial capacity C_j, represented as a function of its location (see upper left part of figure III.1), generated as a sum of bivariate normal kernels centered on the index households with a bandwidth π and scaled to lie in $[\zeta_1, \zeta_2]$.
Mosquito movement was simulated by further smoothing these infection rates C_j via a simple diffusion process [Malinga et al., 2019a]. The acquisition of infection from mosquitoes

Fixed parameters used to generate the data sets			Values
N	Number of households in trial		2500
η	Domain size		5 km
α_1, α_2, α_3	Parameters for the Thomas process [Baddeley and Turner, 2005]		4, 50, 0.25 km
ξ, π	Index households and bandwidth for the KDE		200, 0.5 km
ζ_1, ζ_2	Scaling of C_i's		0.2, 0.6
Varying parameters used to generate the data sets		**Levels**	**Values**
E_s	Efficacy	2	0.2, 0.4
σ	Standard deviation	5	0.042, 0.106, 0.170, 0.234, 0.298 km
$\theta = 1.64\sqrt{2}\sigma$	$\Rightarrow$ Contamination range		0.10, 0.25, 0.40, 0.55, 0.70 km
c	Number of clusters		20, 25, 30, 35
h	Households per cluster		30, 40, 50, 60, 70
$c\&h$	Only combinations with $0.6N < 2hc \leqslant N$ households	8	
	Seed, sampled out of $[1, 100\,000]$ with seed(-1)	100	

Table III.1: Summary of the parameters used to generate the data sets.

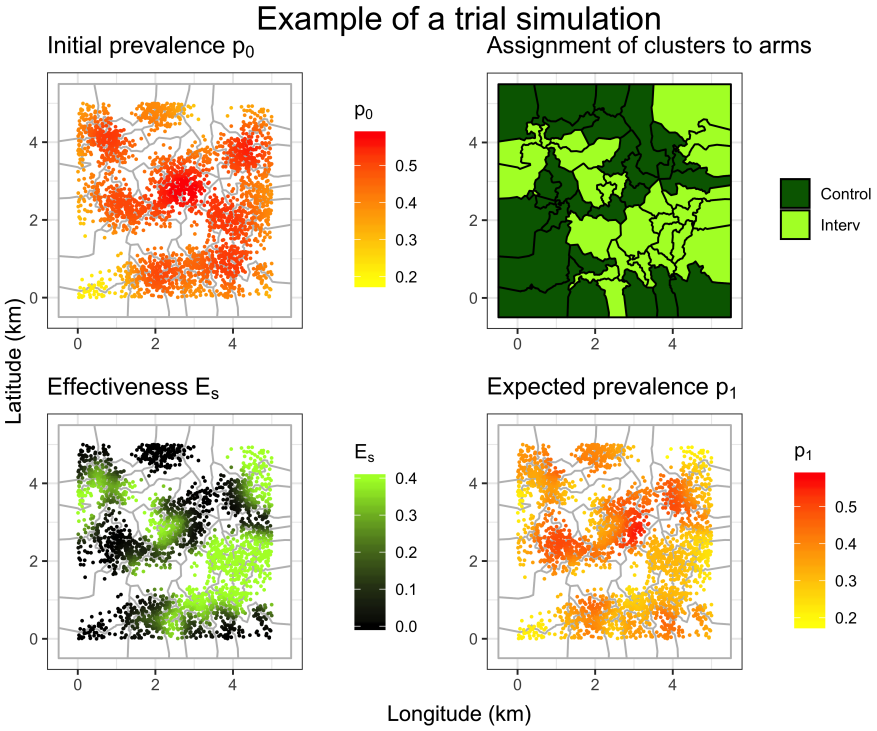


Figure III.1: Trial simulation and intervention assignment, visualizing the assigned initial malaria prevalence to a distribution of households (upper left), the division of the households into clusters based on a travelling salesman algorithm together with the cluster assignment (upper right), the assigned effectiveness varying at the cluster boundaries due to the mosquito movement (lower left) and the resulting expected prevalence (lower right). The parameters used to generate this data set can be found in table III.1, with the varying parameters chosen as follows: households were assigned to 50 clusters, each consisting of 50 households. The intervention was assumed to be 40% effective and the assigned contamination range was 0.4 km.

at each location is then proportional to C_j. In the absence of intervention, these infections are distributed to other locations i proportionately to a bivariate normal kernel, where

$$\Sigma := \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

is the diagonal covariance matrix and σ is the standard deviation of the distance moved by mosquitoes during the extrinsic cycle of the parasite, i.e. the time it takes for a malaria parasite to become transmissible. Equivalently, the numbers of infections distributed to house j from house i is a Gaussian function of distance between the houses,

$$f_{i-j} = f(x_i - x_j, y_i - y_j) := \frac{1}{2\pi\sigma^2} \exp\Big(- \frac{(x_i - x_j)^2 + (y_i - y_j)^2}{2\sigma^2} \Big) = f_{j-i}.$$

This two-dimensional function results in a total dispersion of infections quantifiable by the trace of Σ, that is $2\sigma^2$. For each household j, this means that 95% of the dispersion of infections happens within a radius $\theta := \Phi^{-1}(0.95) \times \sqrt{2\sigma^2}$ km where $\Phi^{-1}(p)$ denotes the quantile function of a standard normal distribution with $p \in [0, 1]$. Hereafter, θ is called the contamination range that quantifies the significant dispersion of infections (and hence the mosquito movement) in one direction.
The exposure to infection at location j in the absence of intervention is thus

$$z_{j,0} := \sum_i \Big(C_i \frac{f_{i-j}}{\sum_k f_{j-k}} \Big),$$

where the normalizing term $\sum_k f_{j-k}$ is required to ensure that the total vectorial capacity distributed from household j over all destination houses sums to C_j. The expected prevalence in household j in the absence of intervention, $p_{j,0}$, is scaled so that the mean of $p_{j,0}$ corresponds to a pre-defined value, $\bar{p}_0$. i.e.

$$p_{j,0} := \bar{p}_0 \frac{N}{\sum_k z_{k,0}} z_{j,0}.$$

Determination of clusters and assignment of intervention effects

The locations were grouped into c clusters per arm, each consisting of h households, by defining an efficient path through them with a heuristic algorithm for the traveling salesman problem (TSP) using the TSP package [Hahsler and Hornik, 2007] in R, as proposed by Silkey *et al.* [Silkey et al., 2016] for a trial of mosquito traps in Kenya [Homan et al., 2016], the SolarMal trial. Equal numbers of households were then allocated to each cluster along the derived path and a simple random sample of half the clusters was assigned to each arm of the trial (see upper right part of figure III.1).
The presence of an intervention acting at source household j reduces the total number of infections acquired from mosquitoes in that household by some efficacy E_s so that $z_{j,1}$, the exposure to infection of household j in the presence of the intervention, is

$$z_{j,1} := \sum_i \Big(C_i \frac{f_{i-j}}{\sum_k f_{j-k}} (1 - E_s \chi_i) \Big),$$

where the indicator function χ_i takes the value 1 if household i is intervened, and the value 0 if it is in the control arm (see lower left part of figure III.1). The expected prevalence in household j in the presence of intervention (using the same scale factor as for $p_{j,0}$) is then

$$p_{j,1} := \bar{p}_0 \frac{N}{\sum_k z_{k,0}} z_{j,1},$$

as shown in the lower right part of figure III.1. For each household in the trial population, a single sample was drawn from a Bernoulli distribution with probability $p_{j,1}$, such that

$$\kappa_j := \begin{cases} 1 & \text{with probability } p_{j,1}, \\ 0 & \text{with probability } 1 - p_{j,1}, \end{cases}$$

representing a malaria prevalence survey testing one person per household with an RDT for simplicity. This could easily be extended to more individuals per household by including another level of clustering in the Thomas cluster process.

Trial parameterization

The parameters of the simulation study were chosen to resemble a trial of mosquito traps in Kenya, the SolarMal trial [Silkey et al., 2016, Homan et al., 2016]. In this trial, clusters were assigned with a TSP and hence there were households within the contamination range spanning the cluster boundaries. In the upper part of table III.1, all fixed parameters of the trial simulation are listed. A domain of $\eta \times \eta$, where $\eta = 5\,\text{km}$, was chosen. The three parameters required for the Thomas cluster process were chosen to be $\alpha_1 = 4$, $\alpha_2 = 50$ and $\alpha_3 = 0.25\,\text{km}$, resulting in an expected number of 5000 households per realization. Of these households, $N = 2500$ were chosen to represent the trial population (to have a constant trial population over different simulations). For the kernel density estimation (KDE) of a subsample of the households, $\xi = 200$ households were randomly chosen with a bandwidth of $\pi = 0.5\,\text{km}$ for the Gaussian kernel. The resulting pattern was then scaled to lay in between 0.2 and 0.6 for the initial prevalence.

In addition to the fixed parameters, four parameters of interest were varied, influenced by the values chosen or calculated for the SolarMal trial: the efficacy E_s (20% and 40%); the standard deviation of the Gaussian functions σ, resulting in a contamination range $\theta = 1.64\sqrt{2\sigma^2}\,\text{km}$ ($\theta = 0.1$, 0.25, 0.4, 0.55, 0.7 km); five levels of cluster size for h (30, 40, 50, 60, 70 households per cluster); four levels of c (20, 25, 30, 35 clusters per arm), and of these 20 configurations of h and c only the ones with $0.6N < 2hc \leqslant N$ were included to keep the number of observations stable (8 levels, $(c, h) = (20, 40)$, $(20, 50)$, $(20, 60)$, $(25, 40)$, $(25, 50)$, $(30, 30)$, $(30, 40)$, $(35, 30)$). The theoretical intra-cluster correlation coefficient (ICC), a measure of variation of the outcome within clusters that is usually obtained from previous studies, was calculated for each data set, resulting in a mean ICC of 0.0021. This leads to an adequately powered study for an efficacy of 20% and an overpowered study for an efficacy of 40%, based on sample size calculations for malaria prevalence [Hayes and Moulton, 2009, Donner et al., 1982, Hayes and Bennett, 1999]. One hundred replicate data sets were produced using different seeds (and hence different patterns of households and infections) for each of the $2 \times 5 \times 8$ parameter configurations. Following guidelines

on simulation studies [Burton et al., 2006, Morris et al., 2019], it was calculated that 100 replicate data sets were sufficient since initial simulation showed that the variance of the main parameter of interest, $\hat{E}_s$ is very low, together with high accuracy for moderate θ for the sigmoid random effects model introduced below. All fixed and varying parameters can be found in table III.1.

3.2 Analysis of intervention effects in CRTs

Conventional linear analysis ignoring contamination

The simplest analysis of a CRT of an intervention targeting mosquito densities and measuring malaria prevalence is a calculation of the risk ratio comparing prevalence in the two trial arms based on cluster level summaries. This leads to an estimate $\tilde{E}_s$ of effectiveness as

$$\tilde{E}_s = 1 - \frac{\tilde{p}_I}{\tilde{p}_C},$$

where $\tilde{p}_C$ is the proportion infected in the control arm and $\tilde{p}_I$ the proportion infected in the intervention arm. The more mosquito movement is introduced, the more the estimate $\tilde{E}_s$ is biased towards the mean between intervention and control arms as it does not adjust for the contamination.

Intervention estimates based on individual-level data and allowing for clustering can be obtained using generalized linear mixed effects models (GLMMs) with the trial arm as the dependent variable and a logistic link function. However, in the special case of binary data, fitting the logistic regression random effect models using Gaussian quadrature may not always provide an adequate model fit due to the failure of the numerical quadrature invoked. If this is the case, it is recommended in the literature to then fit the model with generalized estimating equations (GEEs) [Zeger and Liang, 1986], [Diggle et al., 2002, p. 139] and an exchangeable correlation structure [Hayes and Moulton, 2009, p. 220]. The estimated effectiveness $\tilde{E}_s$ is obtained as above, by comparing the model outputs $\tilde{p}_C$ and $\tilde{p}_I$. It is possible to extend these linear models with a term of the straight-line distance to the nearest discordant observation or a term of the density of households within a range that receive the intervention [Jarvis et al., 2017, Howard et al., 2000, Hawley et al., 2003]. However, the contribution of each estimated coefficient remains linear and it is not possible to obtain a closed-form contamination range that specifies the maximal measurable extent of contamination from a linear model. It is also not possible to obtain this information from a model with a spatially structured random effect [Jarvis et al., 2017].

For malaria, interventions such as ITNs or indoor residual spraying are usually allocated to a household. The endpoint is then either measured in all residents of an area (as in the SolarMal trial [Homan et al., 2016]) or in a subgroup, normally children (as in the

Navrongo trial [Binka et al., 1996]). If there is more than one observation per household, clustering within the household should also be allowed for in the analysis. If the trial outcome is malaria incidence instead of prevalence, the effectiveness can be calculated via a rate ratio including the time at risk for each group. Individual-level analysis can then use a logarithmic link function and an offset for the time at risk.

Proposal of a nonlinear analysis allowing for contamination

If an intervention lowers mosquito densities in intervention clusters, the intervention effects are contaminated between trial arms due to mosquito movement. This contamination depends on the distance of a household to the nearest discordant household and is expected to follow a symmetrical smooth gradient in the boundary area between intervention and control clusters. Let Δ_{ij} denote the distance of the jth household in the ith cluster to the nearest household in the other arm, endowed with a negative sign for the households in the control arm and a positive sign for households in the intervention arm (hereafter called nearest discordant household). This smooth gradient of intervention effectiveness across arms can then be modeled by a nonlinear sigmoidal function of Δ_{ij}, governed by three parameters, β_1 and β_2, determining its position and height, and a parameter of steepness (growth rate) β_3. A variety of functions can be used to model this sigmoidal shape, the most natural choice being the sigmoid or logistic function (hereafter called sigmoid model and abbreviated with $\mathcal{S}$):

$$\mathcal{S}(\Delta_{ij}) := g^{-1}\Big(\beta_1 + \frac{\beta_2}{1+\exp(-\beta_3\Delta_{ij})}\Big).$$

It is assumed that mosquito densities are proportional to the number of acquired infections, such that the analysis can be carried out with data on malaria prevalence. The function g^{-1} hence denotes a logit link function, adjusting $\mathcal{S}$ for binary outcome data. This model formulation can easily be extended for malaria incidence by using a log link function g^{-1} with a Poisson error function and an offset for the time at risk, and other functions than $\mathcal{S}$ are also possible.

The prevalences in the intervention and control arm are then defined as $\hat{p}_C = g^{-1}(\beta_1)$, $\hat{p}_I = g^{-1}(\beta_1+\beta_2)$ and the resulting effectiveness is $\hat{E}_s = 1-\hat{p}_I/\hat{p}_C$. The parameter β_3 can be transformed to a measure of contamination range in km, the distance over which the estimate of effectiveness is measurably biased. This is defined here as the value of Δ_{ij} where $\mathcal{S}$ attains 95% of its growth, i.e. $\mathcal{S}(\Delta_{ij}) = g^{-1}(\beta_1+0.95\beta_2)$. Solving this for Δ_{ij} results in an interpretable contamination range of $\hat{\theta} = \beta_3^{-1}\log(0.95/0.05) = 2.944\beta_3^{-1}$. An illustration of the sigmoid function as well as how it fits the expected prevalence of an example data set can be found in figure III.2.

This sigmoid model can also be extended to allow for within-cluster correlation. One way

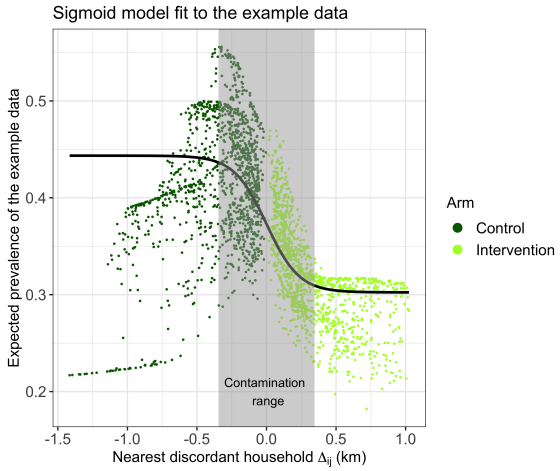


Figure III.2: Illustration of the sigmoid model function for an example data set of the subsequently described simulation study. Households are arranged based on their distance to the nearest discordant household stratified by intervention status on the x-axis (Δ_{ij}) and the expected prevalence is shown on the y-axis. The black line indicates the model fit and the grey rectangle the contamination range in both arms. The model is fitted to the same data set as is used in figure III.1, the detailed parameters are listed there. The patterns in the expected prevalence (such as the approximately linear grouping in the control arm with low expected prevalence) arise from the location of households.

to include a random effect for the clustering of the households is provided by Bayesian hierarchical models using Markov chain Monte Carlo (MCMC). A random effect $\beta_{1,i}$ is assigned to each cluster, with the random effects centered on the expected prevalence in the control arm on a logit scale. For malaria prevalence, the outcome Y_{ij} of the jth household in the ith cluster can then be described as follows:

$$\begin{aligned} Y_{ij} &\sim \text{Binomial}(p_{ij}), \\ \text{logit}(p_{ij}) &= \beta_{1,i} + \frac{\beta_2}{1+\exp(-\beta_3 \Delta_{ij})}, \\ \beta_{1,i} &\sim \text{Normal}(\mu, \tau). \end{aligned}$$

Again, for malaria incidence, a log link function and an offset for the time at risk must be used. The other parameters β_2, β_3, μ and τ are assigned non-informative priors. Hereafter, this model will be called sigmoid random effects model, abbreviated with $\mathcal{S}_{RE}$.
Opposed to a conventional, linear analysis ignoring contamination, zones where contamination is likely have to be included in a sigmoid analysis. The precision and accuracy of the estimate of effectiveness $\hat{E}_s$ and the contamination range $\hat{\theta}$ not only depend on the size and number of clusters but also on the geographical size relative to the contamination

range, the proximity of clusters in opposing arms and on the settlement distribution. This can be captured by considering the percentage of households unaffected by the contamination range $\hat{\theta}$ across the intervention boundary, namely the households whose distance to the discordant arm is greater than $\hat{\theta}$, hereafter called percentage of households in core, denoted by ω. To determine the premises under which a sigmoid analysis, either with $\mathcal{S}$ or $\mathcal{S}_{RE}$, yields precise and accurate estimates, the percentage of households in core ω will be used. A summary of all parameters introduced is listed in table III.2.

$\tilde{E}_s$, $\tilde{p}_I$, $\tilde{p}_C$	Effectiveness and prevalences in the intervention and control arm (ignoring contamination)
Δ_{ij}	Distance of jth household in the ith cluster to nearest discordant household
$\mathcal{S}$	Sigmoid model (allowing for contamination)
$\mathcal{S}_{RE}$	Sigmoid random effects model (allowing for contamination and including random effects)
β_1, β_2, β_3	Parameters for $\mathcal{S}$ and $\mathcal{S}_{RE}$, describing the position, height and steepness of the function
$\hat{E}_s$, $\hat{p}_I$, $\hat{p}_C$	Effectiveness and prevalences in the intervention and control arm for $\mathcal{S}$ or $\mathcal{S}_{RE}$ (allowing for contamination)
$\hat{\theta}$	Estimated contamination range from $\mathcal{S}$ or $\mathcal{S}_{RE}$
μ, τ	Hyperparameters for the Bayesian hierarchical model $\mathcal{S}_{RE}$
ω	Percentage of households in core, that is households whose distance to the discordant arm is greater than $\hat{\theta}$ (ignoring assignment to arms)

Table III.2: Summary of the important parameters and abbreviations defined.

Guide for the implementation in R

A trial can be analyzed with a sigmoid random effects model $\mathcal{S}_{RE}$ following the procedure outlined in table III.3. The R code as well as simulated datasets can be found in the

Input:	Geolocations of households; Cluster and intervention assignment of households; Trial outcome of interest (malaria prevalence or incidence); Technical parameters for the MCMC
Output:	Estimated effectiveness and contamination range with 95%CI
1:	Calculate distance to the nearest discordant household
2:	Set up the sigmoid random effects model in JAGS
3:	Fit the sigmoid random effects model
4:	Transform the model output for interpretation

Table III.3: Four steps to fit a sigmoid random effects model $\mathcal{S}_{RE}$ using MCMC.

additional files of the published version of this article. As input, data on the trial is needed, along with some technical parameters to fit the MCMC model. The output is the estimated

effectiveness and contamination range, with their 95% credible intervals (95%CI). In the first step, the distance to the nearest discordant household is calculated for each household, households in the control arm are additionally endowed with a minus sign. In the second step, the model is implemented as a Bayesian hierarchical model using MCMC, formulated in BUGS (Bayesian inference Using Gibbs Sampling) and fitted with JAGS [Plummer, 2019] (Just Another Gibbs Sampler). The parameter β_3 is constrained for the resulting contamination range to be interpretable, because $\hat{\theta}$ is calculated by taking the inverse of β_3. The model is then fitted (third step) and the parameters are transformed (fourth step) according to the chosen link function, to be interpretable. The back transformation for β_3 is independent of the link function, it holds that $\hat{\theta} = \log(0.95/0.05)\beta_3^{-1}$, as discussed above.

Both GEE and GLMMs are easily implemented in R with the packages geepack [Højsgaard et al., 2006] and lme4 [Bates et al., 2015] for instance. The function $\mathcal{S}(\Delta_{ij})$ can be fitted to prevalence data at the household level with a maximum likelihood method for Bernoulli data, assuming that households are independent of each other. The optimization can then be performed with a genetic algorithm (GA package [Scrucca, 2013]).

3.3 Analysis of simulations

Each data set was analyzed by: analysis allowing for within-cluster correlation (GEE); sigmoid model ($\mathcal{S}$); sigmoid model including a random effect ($\mathcal{S}_{RE}$). For each of the $2\times5\times8$ parameter configurations, the performance of the different models was assessed in terms of [Burton et al., 2006]: the relative bias with respect to the true value of the parameter of interest; the empirical standard error, that is the standard error of the parameter of interest; the average width of the 95% confidence intervals; the coverage probability, the proportion of the 95%CI that contain the true value of the parameter of interest. The first two performance measures are on the parameter of interest itself, measuring its accuracy and precision across replicate data sets, the third and fourth are on the 95%CI around the parameter of interest, quantifying the precision and accuracy of the 95%CI. A summary with the corresponding formulae can be found in table III.4.

Evaluation criteria	Abbreviation	Formula
Relative bias	relBias	$(\mathrm{E}[\hat{\Theta}] - \Theta)/\Theta$
Empirical standard error	EmpSE	$\mathrm{Var}[\hat{\Theta}]^{1/2}$
Width of 95%CI	Width	Average width of 95%CI
Coverage probability	CP	Proportion of 95%CI that contained Θ

Table III.4: Evaluation criteria for parameter estimations across replicate data sets. $\Theta \in \mathbb{R}$ denotes the true value of the parameter of interest, $\hat{\Theta} \in \mathbb{R}^m$ the parameter estimations for the m replicate data sets.

The 95%CI for the GEE and $\mathcal{S}$ analyses were calculated by parametric bootstrapping [Efron and Tibshirani, 1994], because this method is very generalizable (R package boot [Davison and Hinkley, 1997]). This step was repeated $R = 100$ times, leading to 10 000 resamples for each of the parameter configurations. For the fitting of $\mathcal{S}$, the parameter region for the genetic algorithm was chosen such that β_1, $\beta_2 \in [0, 1]$ and $\beta_3 > 0$. For the JAGS model, the 95% credible intervals were obtained from the 2.5 and 97.5 quantiles. Uninformative priors were chosen for β_1 and β_2 and a mildly informative prior for β_3 to constrain the resulting contamination range to be in $[0, 1.5]$ km. All simulations were performed at sciCORE scientific computing core facility at the University of Basel under R version 3.6.0 [R Core Team, 2017].

4. Results of the simulation study

The estimation of the two outcome parameters $\hat{E}_s = 1 - \hat{p}_I/\hat{p}_C$ and $\hat{\theta}$ for the sigmoid models are evaluated by four performance measures (relBias, EmpSE, Width, CP) in terms of the four parameters that were varied (E_s, θ, c, h) and compared against conventional methods for analysis. Both a GLMM and GEE showed very similar results and had acceptable model fit. However, a mixed effects model took slightly longer to fit. Simple cluster summaries also resulted in very similar results to a GEE or GLMM. Hence, only the results for a GEE analysis are used as comparison.
This section is divided into three parts: first an evaluation of the simulations for the parameters E_s and θ, followed by the evaluation for the parameters determining cluster size and number of households per cluster, c and h. Each of these two parts is further divided based on the different performance measurements. The third part is on the results in terms of the percentage of households in core, ω, and the difference between a GEE and a sigmoid analysis.

4.1 Varying the efficacy E_s and the contamination range θ

The results in this paragraph are averaged over all values of the number of clusters c and the number of households per cluster h.

Relative bias and empirical standard error

The relative bias and empirical standard error of the model fits are depicted in figure III.3. As the assigned contamination range θ increases, the effectiveness estimate of the GEE is biased towards zero. Results are similar across different levels of assigned efficacy E_s. The two sigmoid models also show greater bias towards zero with increasing contamination, but less so than with the GEE model. A similar pattern can be seen for the estimated

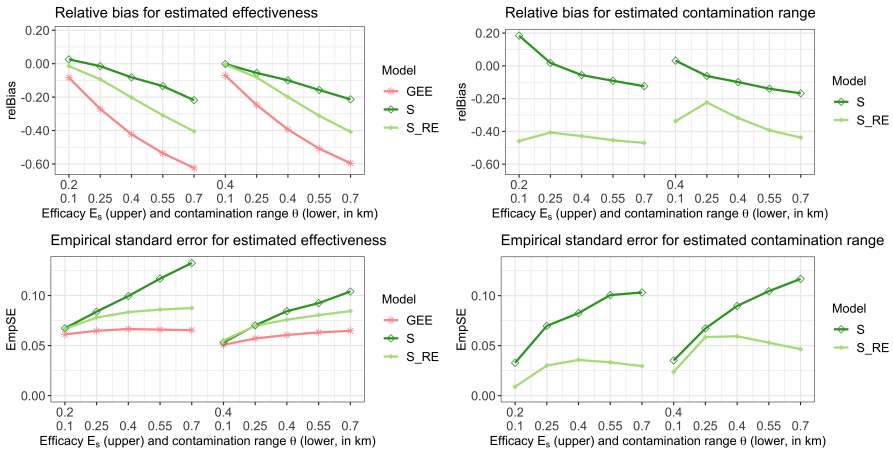


Figure III.3: Relative bias (upper plots) and empirical standard error (lower plots) for model fits. The graphs on the left illustrate the errors for the parameter estimation $\hat{E}_s$ for each of the three models and the graphs on the right for the contamination range $\hat{\theta}$ that only exists for the sigmoid models $\mathcal{S}$ and $\mathcal{S}_{RE}$. Depicted on the x-axis are 10 different levels of parameter variation for E_s and θ.

contamination range $\hat{\theta}$ for $\mathcal{S}$, with greater bias towards zero as $\hat{\theta}$ increases. In contrast, the $\mathcal{S}_{RE}$ model estimates a constant contamination range regardless of the value of θ, but is always substantially more biased towards zero than for $\mathcal{S}$. The empirical standard errors show the exact opposite trends. For both parameter estimations, the empirical standard error increases with greater contamination. The GEE analysis has the lowest variance for the estimated effectiveness, and the $\mathcal{S}$ model the highest. In conclusion, a GEE analysis shows lower accuracy but higher precision than the sigmoid models, $\mathcal{S}$ shows high accuracy together with low precision and the $\mathcal{S}_{RE}$ is in between.

Width of 95%CI and coverage probability

The coverage probability and the width of the 95%CI are highly correlated: a desirable result would be a narrow 95%CI together with a high coverage probability. The results are depicted in table III.5. GEE always has narrow 95%CI but shows a very bad coverage probability for increasing θ. The sigmoid models have much better coverage probabilities but wider confidence intervals, with the random effect model yielding even wider confidence intervals, as was expected because $\mathcal{S}$ does not account for the clustering, leading to incorrectly high precision. For increasing θ, the width of the confidence intervals for the sigmoid models increases and the coverage probability decreases. This decrease in coverage probability is higher for a higher assigned efficacy E_s. The width of the confidence intervals, however, is not altered by this parameter.

		GEE		$\mathcal{S}$		$\mathcal{S}_{RE}$		
E_s	θ	$\hat{E}_s$	$\hat{\theta}$	$\hat{E}_s$	$\hat{\theta}$	$\hat{E}_s$	$\hat{\theta}$	
0.2	0.1	82	-	85	100	96	100	CP (%)
		0.18	-	0.20	0.30	0.27	0.30	Width
	0.25	70	-	88	100	94	100	
		0.18	-	0.27	0.62	0.30	0.77	
	0.4	52	-	91	100	93	100	
		0.19	-	0.33	0.95	0.33	1.19	
	0.55	42	-	90	100	88	100	
		0.19	-	0.39	1.27	0.34	1.59	
	0.7	34	-	90	100	84	100	
		0.20	-	0.46	1.61	0.35	1.99	
0.4	0.1	80	-	88	100	96	100	
		0.15	-	0.18	0.22	0.23	0.30	
	0.25	37	-	90	100	94	100	
		0.17	-	0.25	0.45	0.29	0.78	
	0.4	11	-	90	99	85	100	
		0.18	-	0.30	0.70	0.32	1.18	
	0.55	2	-	92	100	71	100	
		0.19	-	0.36	0.98	0.33	1.54	
	0.7	0	-	91	99	54	100	
		0.20	-	0.42	1.29	0.35	1.91	

Table III.5: Coverage probability and the width of the confidence intervals for the two parameters of interest, E_s and θ (10 levels).

4.2 Varying the number of clusters c and the number of households in each cluster h

The results in this paragraph are averaged over all values of efficacy E_s and contamination range θ.

Relative bias and empirical standard error

The relative bias and empirical standard error of the model fits are visualized in figure III.4. All three models are quite robust with respect to the different variations of c and h. Again, as noted above, a GEE has the lowest accuracy and highest precision, while $\mathcal{S}$ shows the opposite. For all models, the empirical standard error is almost constant for both estimated parameters.

Width of 95%CI and coverage probability

The width of the confidence intervals and the coverage probability yield no new insights; the results are depicted in table III.6. Again, the sigmoid models have a very high coverage probability whereas this is very low in GEE. $\mathcal{S}_{RE}$ has wider credible intervals than $\mathcal{S}$, with comparable coverage probability.

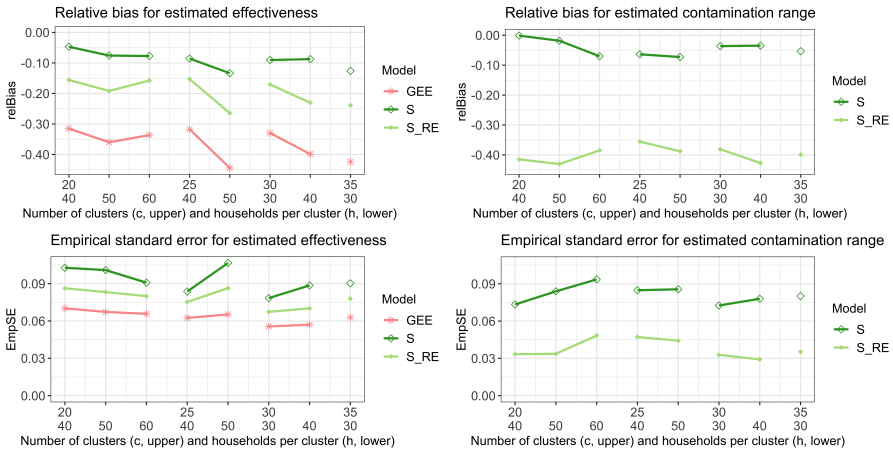


Figure III.4: Relative bias (upper plots) and empirical standard error (lower plots) for models fits when the parameters c (number of clusters in one arm, upper x-axis) and h (households per cluster, lower x-axis) are varied. The graphs on the left illustrate these errors for the parameter estimation $\hat{E}_s$ and the graphs on the right for $\hat{\theta}$ (only for sigmoid models).

		GEE		$\mathcal{S}$		$\mathcal{S}_{RE}$		
c	h	$\hat{E}_s$	$\hat{\theta}$	$\hat{E}_s$	$\hat{\theta}$	$\hat{E}_s$	$\hat{\theta}$	
20	40	59	-	89	100	91	100	CP (%)
		0.20	-	0.33	0.79	0.34	1.01	Width
	50	48	-	87	100	87	100	
		0.19	-	0.33	0.92	0.33	1.21	
	60	40	-	88	100	88	100	
		0.18	-	0.29	0.81	0.31	1.17	
25	40	41	-	86	99	88	100	
		0.17	-	0.27	0.73	0.30	1.11	
	50	30	-	86	100	81	100	
		0.19	-	0.35	0.95	0.32	1.30	
30	30	48	-	91	100	93	100	
		0.16	-	0.27	0.75	0.28	1.01	
	40	47	-	93	100	86	100	
		0.19	-	0.35	0.90	0.31	1.19	
35	30	32	-	90	100	80	100	
		0.19	-	0.33	0.83	0.31	1.16	

Table III.6: Coverage probability and the width of the confidence intervals for the two parameters of interest, c and h (8 levels).

4.3 Comparison between the analyses based on the percentage of households in core

The percentage of households in core, ω, is the key indicator for determining how small clusters can be with respect to the contamination range $\hat{\theta}$ for the sigmoid models. This measurement does not differentiate between households in the intervention or control arm, but if an equal number of same size clusters are allocated to both arms it is likely that there is a certain balance. For each of the simulations, ω was calculated and the relative bias

together with the width of the 95%CI was plotted with respect to ω, see figure III.5. This

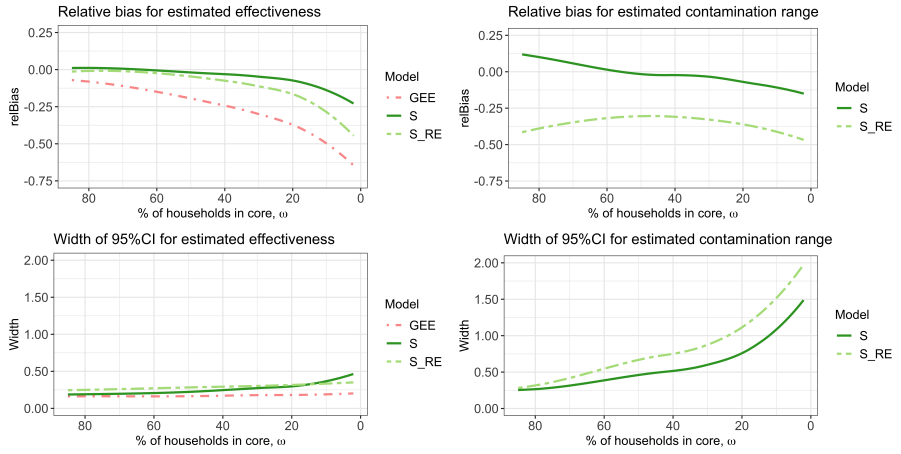


Figure III.5: Relative bias (upper plots) and width of the 95%CI (lower plots) for model fits in terms of the percentage of households in core ω. The graphs on the left illustrate these errors for the parameter estimation $\hat{E}_s$ and the graphs on the right for $\hat{\theta}$. The lines are smoothed over all parameters E_s, θ, c and h.

figure displays the same data as described in figures III.3 and III.4, but with respect to ω. From the graphs displaying the relative bias and width of the 95%CI for the estimated effectiveness $\hat{E}_s$, it becomes clear that at $\approx 20\%$ of households in core, the dynamics of the curves change. In all three models, there is considerably more bias moving towards less households in core. For a GEE analysis, it is clearly beneficial to have $\omega = 100\%$, all households in core. For a sigmoid analysis (with $\mathcal{S}$ or $\mathcal{S}_{RE}$) this is not the case. For the estimated effectiveness, if around 50% of the households are in core, the relative bias is approximately zero and the width of the confidence intervals is still small compared to the width if fewer households were in core. The relative bias for the estimated contamination range $\hat{\theta}$ is very flat and shows a nonlinear behaviour. The width of the 95%CI increases the fewer households there are in core, and once less than $\omega \approx 20\%$, the growth accelerates substantially.

5. Example: the Navrongo trial of ITNs

5.1 Study design

This large-scale CRT was conducted between July 1993 and June 1995 in the Kassena-Nankana districts of northern Ghana with the goal to assess the effect of ITNs compared to no ITNs on child mortality. The area was predominantly rural with people living in dispersed settlements, arranged in compounds. The study was a parallel CRT with 96

geographically contiguous clusters and an average of 120 compounds per cluster. Where possible, small paths or roads were used to delineate the clusters, but in most cases, the cluster boundaries did not correspond to natural barriers. The intervention of permethrin impregnated bed nets was allocated to 48 randomly chosen clusters and 31 000 ITNs were provided to intervention participants. A full description of the study design is reported elsewhere [Binka et al., 1996].

The outcome was all-cause mortality in children aged 6 months to 4 years, reported as a standardized mortality ratio (SMR). All children in the study area were included. The expected number of deaths for each cluster was computed by applying age-specific death rates derived from the pre-intervention population to the post-intervention time at risk and was treated as an offset for the regression models [Binka et al., 1998]. Data captured included the geographical coordinates of the household and the distance from each household to the nearest discordant household.

5.2 Published trial results

A total of 857 deaths occurred among children in the trial over the 2 years of follow-up. The original analysis found a 17% reduction in mortality (rate ratio (RR) comparing SMRs of 0.83, 95% CI $[0.69, 1.00]$) [Binka et al., 1996]. Subsequently, Binka *et al.* [Binka et al., 1998] graphed the ITN effect in relation to distance from the boundary. A regression approach incorporating this distance indicated that among children from clusters randomized to the control arm, the mortality risk increased by 6.7% with each additional shift of 100 m away from the nearest household in the intervention arm (95%CI $[1.8, 11.4]$%) [Binka et al., 1998]. Notably, due to the considerable spatial information available, the estimated confidence intervals (which did not allow for the spatial auto-correlation in the data) around the regression lines in this analysis were narrow, even though the overall estimate of effectiveness was imprecise [Binka et al., 1996]. This data set was recently reanalyzed [Jarvis et al., 2019] using multilevel models and geostatistical approaches to allow for spatial correlations and contamination effects. Including the distance to the nearest discordant household as a fixed effect in the multilevel model indicated an increase of the SMR with every additional 100 m away from the intervention arm of 1.7% (95% credible interval $[0.6, 2.6]$%) [Jarvis et al., 2019]. The main conclusion of the reanalysis was that, despite the evidence of a spatial contamination effect, the primary conclusions of the trial remain unaffected. The increase of the SMR with every additional 100 m was estimated to be less than was reported before, but the confidence intervals were similarly narrow. The confidence intervals around the main effect remained wide.

5.3 Methods

The Navrongo data was reanalyzed with the sigmoid models $\mathcal{S}$ and $\mathcal{S}_{RE}$ and the results were reported in terms of mortality incidence rates. Hence, a log link function g^{-1} with a Poisson error function was used. As in the original analysis, the expected number of deaths was treated as an offset. For comparison with the original spatial analysis, the increase in mortality rate with each additional 100 m away from the boundary was calculated by comparing the SMRs at the required distances. For fitting the JAGS models, the number of iterations was set at 20 000 with a burn-in period of 10 000 and uninformative priors were used for β_1 and β_2, together with a mildly informative prior for β_3, as for the simulation study. For fitting the sigmoid models without a random effect, R was set to 1000, and the valid parameter region for the search of the genetic algorithm for the contamination range was chosen to be $[0.05, 0.6]$ km. An extended reanalysis of the Navrongo data can be found in appendix A.3.

5.4 Results

Bed nets were associated with a 16.6% and 19.0% reduction in all-cause mortality in children aged 6 months to 4 years for $\mathcal{S}_{RE}$ and $\mathcal{S}$, the sigmoid models with and without a random effect, respectively. As in the original analysis, confidence intervals were wide. Contamination across the boundary was found to be around 0.2 km per arm, again with wide confidence intervals, especially for $\mathcal{S}_{RE}$. The parameter estimations for $\mathcal{S}_{RE}$ translate to an increase in mortality from the intervention boundary of 5.6%, 95%CI $[0.2, 15.5]$% up to 100 m and 2.4%, 95%CI $[0.1, 1.4]$% between 100 m and 200 m. After that the increase is very slow, since the contamination range is around 200 m. Since the model is symmetrical, the same numbers also hold for a decrease in mortality with each 100 m away from the nearest household without a bed net. All the estimates are displayed in table III.7. Figure III.6 illustrates the results for $\mathcal{S}_{RE}$, analogous to figure III.2.

6. Discussion

When contamination is anticipated in a cluster randomized trial of a malaria control intervention, a conventional analysis would lead to a biased estimate of effectiveness. To avoid this, a fried egg design is often recommended, attempting to separate the trial arms with buffer zones around each cluster [Hayes and Moulton, 2009]. This allows a conventional analysis, for instance with GEEs or GLMM, to be carried out, but leads to trials of much bigger geographical size than would be needed based on sample size formulae [Donner et al., 1982, Hayes and Bennett, 1999]. Further, an estimate of the measurable contamination range is needed to quantify the buffer zone, and in the absence of suitable data this

	Effectiveness:	Contamination range:	Increase mortality shift 100m; 200m:
$\mathcal{S}_{RE}$	16.6%, [2.2, 30.7]%	0.198 km, [0.092, 1.088] km	5.6%, [0.2, 15.5]%; 2.4%,[0.1, 1.4]%
$\mathcal{S}$	19.0%, [7.7, 28.1]%	0.170 km, [0.051, 0.495] km	-
Original [Binka et al., 1998]	17.0%, [0.0, 31.0]%	-	6.7%, [1.8, 11.4]%
Reanalysis [Jarvis et al., 2019]	18.0%, [5.0, 30.0]%	-	1.7%, [0.6, 2.6]%

Table III.7: Results for the parameter estimations for the Navrongo data for sigmoid models $\mathcal{S}$ and $\mathcal{S}_{RE}$, compared to the results of the original analysis [Binka et al., 1998] and a previous reanalysis [Jarvis et al., 2019]. In brackets, the 95%CI are given. The contamination range only exists for $\mathcal{S}$ and $\mathcal{S}_{RE}$. The last column indicates the increase in mortality after 100 m away from the nearest discordant household. For $\mathcal{S}_{RE}$ this is nonlinear, hence the increase after 100 m and 200 m is reported.

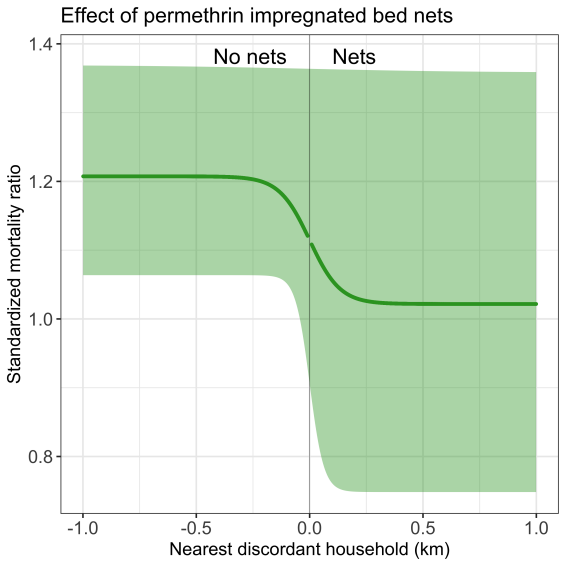


Figure III.6: Illustration of the results for the sigmoid model $\mathcal{S}_{RE}$. The thick line indicates the fitted sigmoid curve, together with the confidence intervals. On the x-axis, the distance to the nearest discordant household up to 1 km and on the y-axis, the standardized mortality ratio is plotted.

is typically based only on expert opinion. The contamination between arms in a cluster randomized trial contains information about the intervention *per se*. An analysis that takes this information into account as a trial outcome can lead to unbiased estimates of effectiveness, even when a substantial part of the data is affected by this contamination. This work proposes such an analysis for CRTs where contamination of intervention effects is introduced by mosquito movement and a nonlinear model is used to quantify the range of contamination across intervention arms.

The main strength of this approach lie in the adjustment of the estimate of intervention effectiveness to account for contamination, and yielding a closed-form estimate for the contamination range that can inform future analyses. Obtaining such a simple closed-form estimate would not be possible from a linear parametric or a nonparametric approach. The sigmoidal shape functions make the model nonlinear, complicating the analysis substantially. This also implies that an interpretation of the coefficients is not as straightforward as for a linear model and the contamination range needs a back transformation. Furthermore, the sigmoid function is symmetrical, which means that if there was an asymmetrical contamination, such as a protective effect of an intervention on nearby non-users, but no increase in risk for the intervention users associated with being near the boundary, this would not be captured with this proposed analysis. To capture this, an asymmetric function with another parameter for the asymmetry would be needed, further complicating

the analysis and interpretation.

The fitting of the random effects model was performed with an MCMC approach, where the binary structure of the outcome, representing a malaria prevalence survey, can easily be treated. The inclusion of random effects in a frequentist approach [Pinheiro and Bates, 2000] with the R package nlme [Pinheiro et al., 2018] proved unreliable, because nlme does not allow for binary outcome data structure. Representing the common practice in the field, an exchangeable correlation structure was chosen. For the Navrongo trial, a recent reanalysis [Jarvis et al., 2019] tested the impact of different spatial correlations and found minimal differences, supporting the choice of an exchangeable correlation structure.

The simulation study indicated that different cluster configurations (number of clusters and number of households per cluster) only slightly influenced the performance. This makes sense because rearrangement of households based on their distance to the intervention boundary does not take account of cluster assignment; the overall number of households and their spatial distribution relative to the boundary is more important. The results were not very sensitive to the assigned level of efficacy and a low efficacy of 20% did not impose fitting problems. Mosquito movement and hence contamination was simulated with varying widths of normal kernels centered at the households. An increase in mosquito movement biased the intervention effects towards the null in all of the analyses, but the bias was less extreme for the sigmoid functions.

It would be possible to extend the GEE model used for comparison with a term of the straight-line distance to the nearest discordant observation [Jarvis et al., 2017] and to include contamination as a parameter that quantifies the increase in effectiveness per distance unit away from the intervention boundary, as it has been done in the original analysis of the Navrongo trial [Binka et al., 1998]. In an initial analysis of the simulation study using this approach, it was found that the main parameter of effectiveness was not affected (results not shown). Hence, only the more basic GEE model is used for comparison to keep the focus of the simulation study on the sigmoid models.

The results of the simulation study were obtained by averaging over scenarios and varying parameters. This should be kept in mind when interpreting these results. Also, the variation arising due to a finite number of simulations, that could be assessed by Monte Carlo standard errors [Morris et al., 2019], was not addressed. More work is needed to better understand the differences between linear models incorporating contamination and nonlinear approaches and to better determine the premises under which a sigmoid model is suitable.

Since the performance of a sigmoid analysis, apart from the parameters that were varied in the simulation study, also depends on other factors such as geographical cluster size, we explored the simulations with respect to the percentage of households in core, i.e. house-

holds that are unaffected by the contamination range across the boundaries, which is a scale-free parameter. Usually, parameter configurations with a similar ratio between cluster size h and contamination range led to similar values of the percentage of households in core. When more than $\approx 50\%$ of the households are in core, the simulations indicated that it is possible to estimate the effectiveness without bias, irrespective of the cluster division. Since $\mathcal{S}_{RE}$ adjusts for the clustering, it is certainly to be preferred for primary analyses of efficacy over $\mathcal{S}$, although $\mathcal{S}$ yields better results for estimating the contamination range. Hence, these models are not a panacea for contamination in CRTs. As validation, the information gained from both models could be used to define buffer zones post hoc. A range of different buffer zones could be used, and the resulting estimates of effectiveness could be compared to the sigmoid model to check how the estimated contamination range relates to the size of buffer needed to avoid bias. Furthermore, it would be desirable to assemble estimates from multiple previous field studies, before an appropriate value of the contamination range can be assumed for use in designing a new trial for any specific site. This analysis raises the question of how best to divide populations into clusters. Many CRTs are designed with individual villages as clusters, which, depending on the settlement pattern, generally achieves spatial separation of trial arms by ensuring that cluster boundaries pass through unpopulated areas between villages. However, this approach leads to heterogeneity between clusters and varying cluster size (though a uniform number of households might be sampled in each cluster). If estimation of the contamination function is considered desirable it may be important for some cluster boundaries to pass through inhabited areas rather than avoiding them. This makes it feasible to define clusters with equal numbers of enrolled individuals, as was done in the simulation study. It is attractive to use an algorithmic approach to cluster assignment in a CRT, for instance using a travelling salesman algorithm [Silkey et al., 2016, Homan et al., 2016] as we did in the simulations. Further analysis would be needed to determine whether this is optimal in terms of maximizing trial efficiency.

A reanalysis of the Navrongo trial of the effect of ITNs on child mortality in northern Ghana with the proposed method yielded similar results to the original analysis [Binka et al., 1996]. For the sigmoid model $\mathcal{S}_{RE}$, bed nets were associated with a 16.6% reduction in all-cause mortality in children aged 6 months to 4 years (95%CI [2.2, 30.7]%) with a contamination range of 0.198 km per arm (95%CI [0.092, 1.088] km). Given that the outcome was all-cause mortality in children aged 6 months to 4 years and hence the data are rather sparse, it is unsurprising that the credible intervals for both the effectiveness and the contamination range estimate are wide. The result for the contamination is in line with what was found in a larger trial of ITNs in Asembo, Kenya [Hawley et al., 2003], where significant protective effects of ITNs were found for distances of up to 300 m from

cluster boundaries. In the original spatial analysis of the Navrongo data [Binka et al., 1998], an increase with each 100 m away from the nearest household with a bed net was reported to be 6.7% (and 1.7% for the spatial reanalysis, both with narrow confidence intervals). Since the sigmoid model $\mathcal{S}_{RE}$ is nonlinear, the increase with each unit is not a constant. The findings here of 5.6% increase in mortality for the first 100 meters and then 2.4% increase from $100-200$ m are similar to the previously reported results of 6.7%. The Navrongo trial had very large clusters and many households were unaffected by the estimated contamination range. It can be seen in this example that even when the contamination range is big enough to be estimable with such methods, this need not make much difference to the estimate of effectiveness.
An extension of the analyses in this paper would be to build on the results on the percentage of households in core by transforming the distance to the boundary into a measure of local coverage of the intervention, and hence estimating the effectiveness as a function of coverage. This framework could also easily be extended to account for another hierarchy of clustering at the household level or to trial designs with repeated sampling of individuals for either incidence or prevalence, using random effects terms to account for individual variation in addition to cluster effects. More estimates of the contamination range from other field studies are needed to design further trials. Guidelines for how to design CRTs for such an analysis as well as reanalyzes of other CRTs are planned.

7. Conclusions

Contamination measures are themselves valuable trial outcomes, providing information about the indirect effects of the intervention, and calculation of quantities derived from them might have several motivations. For some interventions, such as those intended to repel mosquitoes, the extent of contamination directly relates to the action of the intervention and will inform the density at which deployment is required. For any intervention, demonstration of significant contamination confirms that there is effectiveness: it is not possible for contamination to occur unless the two arms of the trial differ in the outcome. Estimates of this contamination range could be used to define buffer zones post hoc (using pre-specified criteria). But - more importantly - the possibility of statistically adjusting for contamination suggests not only that the size of buffer zones could be minimized, but that they could be completely avoided, leading to smaller and more cost-efficient trials of malaria interventions.

Chapter IV

ESTIMATING INTERVENTION EFFECTIVENESS IN TRIALS OF MALARIA INTERVENTIONS WITH CONTAMINATION

Lea Multerer[1,2] , Fiona Vanobberghen[1,2] , Tracy R. Glass[1,2] , Alexandra Hiscox[3,4] , Steven W. Lindsay[5] , Willem Takken[3] , Alfred Tiono[6] , Thomas Smith[1,2]

[1] Swiss Tropical and Public Health Institute, Basel, Switzerland
[2] University of Basel, Basel, Switzerland
[3] Laboratory of Entomology, Wageningen University and Research, Wageningen, The Netherlands
[4] ARCTEC, London School of Hygiene and Tropical Medicine, London, UK
[5] Department of Biosciences, Durham University, Durham, UK
[6] Centre National de Recherche et de Formation sur le Paludisme, Ouagadougou, Burkina Faso

Published in the *Malaria Journal* in October 2021. (DOI: 10.1186/s12936-021-03924-7)

1. Abstract

In cluster randomized trials (CRTs) or stepped wedge cluster randomized trials (SWCRTs) of malaria interventions, mosquito movement leads to contamination between trial arms unless buffer zones separate the clusters. Contamination can be accounted for in the analysis, yielding an estimate of the contamination range, the distance over which contamination measurably biases the effectiveness.

A previously described analysis for CRTs is extended to SWCRTs and estimates of effectiveness are provided as a function of intervention coverage. The methods are applied to two SWCRTs of malaria interventions, the SolarMal trial on the impact of mass trapping of mosquitoes with odor-baited traps and the AvecNet trial on the effect of adding pyriproxyfen to long-lasting insecticidal nets.
For the SolarMal trial, the contamination range was estimated to be 146 m (95% credible interval [0.052, 0.923] km), together with a 31.9% (95% credible interval [15.3, 45.8]%) reduction of *Plasmodium* infection, compared to the 30.0% reduction estimated without accounting for contamination. The estimated effectiveness had an approximately linear relationship with coverage. For the AvecNet trial, estimated contamination effects were minimal, with insufficient data from the cluster boundary regions to estimate the effectiveness as a function of coverage.
The contamination range in these trials of malaria interventions is much less than the distances *Anopheles* mosquitoes can fly. An appropriate analysis makes buffer zones unnecessary, enabling the design of more cost-efficient trials. Estimation of the contamination range requires information from the cluster boundary regions and trials should be designed to collect this.

2. Background

When testing new malaria control interventions, cluster randomized trials (CRTs) are often the study design of choice, because the intervention is either assigned at the household level, or contamination effects are anticipated between the households [Wilson et al., 2015]. With malaria, most transmission happens during the night when *Anopheles* mosquitoes bite and people are in their home. Movement of mosquitoes while searching for human hosts or oviposition sites is therefore the main cause of contamination in trials of mosquito control interventions against malaria. Whereas this is a challenge in field trials, the practical consequence is that intervention has a beneficial community effect on individuals living close by. To prevent this effect from biasing trial estimates of efficacy towards the null, clusters are usually chosen as geographically contiguous areas of households.
Since contamination may still arise at the cluster boundaries and hence bias trial results, malaria trials are often designed by choosing well separated clusters or enforcing separation by defining buffer zones around each cluster [Hayes and Moulton, 2009, Wolbers et al., 2012, Protopopoff et al., 2015, Protopopoff et al., 2018, Eisele et al., 2016, Delrieu et al., 2015]. Ideally, buffer size might be determined using estimates of the range of the contamination [Guerra et al., 2014], but very broad buffers are often used since other information is rarely available. In such trials, entire clusters receive the intervention, but only data

from cluster cores are analyzed. This allows a standard analytical approach [Hayes and Moulton, 2009] but at the cost of enrolling very large populations. Spatial separation may increase heterogeneity between clusters, and the cluster cores may be unrepresentative of the whole population if clusters correspond to natural units such as villages.
Estimating the spatial contamination is of scientific interest [Multerer et al., 2021, Jarvis et al., 2017, Halloran et al., 2017, Benjamin-Chung et al., 2017, Anaya-Izquierdo and Alexander, 2020] because protection of people living nearby is an important property of an intervention. Secondary analyses of several CRTs of malaria interventions have estimated contamination effects using linear models with terms measuring the distance between observations from one arm of the study to the nearest observation from the other study arm [Jarvis et al., 2017, Binka et al., 1998, Howard et al., 2000, Hawley et al., 2003]. These analyses all found evidence of spatial effects and tended to demonstrate the importance of accounting for contamination in estimating unbiased effects of the intervention. Nevertheless, these linear models cannot simultaneously provide closed-form estimates of the range over which the contamination is relevant while adjusting the estimate of effectiveness for the contamination. The authors recently demonstrated that this can be achieved with a sigmoid random effects model for the analysis of CRTs of malaria interventions with contamination arising from mosquito movement [Multerer et al., 2021].
Stepped wedge cluster randomized trials (SWCRTs) [Hayes and Moulton, 2009, Hussey and Hughes, 2007, Mdege et al., 2011] are a modification of CRTs in which the intervention is introduced progressively to all clusters in random order. To gain a better understanding of the effect of contamination in SWCRTs, the proposed model for CRTs [Multerer et al., 2021] is extended to analyze SWCRTs. It is then shown how the measurable contamination between trial arms (a quantity termed contamination range) leads to an estimate of the effective intervention coverage for each household and how this relates to the intervention effectiveness. These methods are applied to two SWCRTs of malaria interventions; the SolarMal trial assessing the effect of mass trapping with solar-powered odor-baited mosquito traps on Rusinga island, Kenya [Hiscox et al., 2016, Silkey et al., 2016, Homan et al., 2016], and the AvecNet trial investigating the effect of adding pyriproxyfen to long-lasting insecticidal nets in Burkina Faso [Tiono et al., 2015, Tiono et al., 2018].

3. Methods

3.1 The SolarMal trial of odor-baited mosquito traps

The SolarMal SWCRT aimed to reduce mosquito population size, reduce biting intensity and eliminate *Plasmodium falciparum* malaria on Rusinga island, Lake Victoria, Kenya [Hiscox et al., 2016, Silkey et al., 2016, Homan et al., 2016] using Solar-powered

Mosquito Trapping Systems [Hiscox et al., 2014] (SMoTS) which included odor-baited traps to lure and kill host-seeking mosquitoes. All households on Rusinga (area 44 km^2, mean population 24 879 [Homan et al., 2016]) were eligible to take part in the trial and were assigned to clusters using a travelling salesman algorithm, resulting in 81 geographically contiguous clusters of 50 - 51 households. Between June 2013 and May 2015, SMoTS were installed in one cluster per week, with a randomized order of clusters, until universal coverage of 4 358 households was achieved [Silkey et al., 2016].

The primary outcome was clinical malaria in individuals of any age, measured as fever plus a positive rapid diagnostic test (RDT) result and monitored through repeated household visits, secondary outcomes were malaria prevalence, measured by RDT, and mosquito densities. Data on malaria prevalence were collected at four-month intervals resulting in five survey rounds during rollout, at 22%, 46%, 63%, 76% and 95% intervention coverage. In each survey, malaria prevalence was recorded in a 10% random sample of households and clusters were excluded from analysis in the week during which the SMoTS were installed. Further details are given in the study protocol [Hiscox et al., 2016].

The clinical incidence of malaria episodes was unexpectedly low, hence the focus of the original analysis [Homan et al., 2016] shifted to the secondary outcome. Malaria prevalence was reported to be 31.4% (95% confidence intervals (CI): [27.5, 35.1]%) lower in intervention clusters (prevalence 23.7%, 1552/6550 people) than in control clusters (prevalence 34.5%, 2002/5795 people). Including random effects for clustering and survey round with generalized linear mixed models, the effectiveness of SMoTS on malaria prevalence was estimated to be 30.0% (95% CI: [20.9, 38.0]%).

3.2 The AvecNet trial of long-lasting insecticidal nets

The AvecNet trial assessed the effect of adding pyriproxyfen, an insect growth regulator, to long-lasting insecticidal nets (LLINs) in rural Burkina Faso, an area with intense malaria transmission and highly pyrethroid-resistant vectors. A baseline demographic survey enumerated 63 903 individuals living in 93 villages in an area of 1 250 km^2 [Tiono et al., 2015]. Over a two-year period, during high malaria transmission seasons, LLINs treated with permethrin were incrementally replaced by LLINs treated with permethrin and pyriproxyfen in a SWCRT with 40 clusters [Tiono et al., 2015, Tiono et al., 2018], with an overall 95% coverage of nets. Clusters were based on administrative units and an average of 50 children (aged 6 - 60 months) were selected in each cluster and followed up by passive case detection for clinical malaria at health centers. Each month from June to September in 2014 and 2015, five clusters switched from control to intervention arm.

The primary outcome was clinical malaria, measured as fever plus a positive RDT result for *Plasmodium falciparum*. The child-years at risk and the incidence rate ratio (IRR),

were calculated for each month in each group. Data were collected between June to December 2014 and May to December 2015, resulting in nine months with data from both the intervention and control arms. In these nine months, the mean intervention coverage was 17, 33, 44, 50, 51, 51, 56, 64 and 81%. Further details of trial design can be found in the study protocol [Tiono et al., 2015].
The original analysis [Tiono et al., 2018] estimated clinical malaria incidence of 2.0 episodes per child-year in the control group and 1.5 episodes per child-year in the intervention group (IRR 0.88, 95% CI: [0.77, 0.99], estimated from a Poisson model with offset for log-transformed exposure years, a random effect for cluster and fixed effects for month and health facility).

3.3 Analysis of SWCRTs allowing for contamination

Contamination between clusters because of mosquito movement between households might be expected to bias the intervention effects in both the SolarMal and AvecNet trials. This contamination is expected to follow a symmetrical smooth gradient in the boundary regions between intervention and control clusters and can hence be modeled by a sigmoid function of the distance of households to the nearest household in the discordant trial arm [Multerer et al., 2021] (figure IV.1). Analyses of simulated datasets [Multerer et al., 2021] found

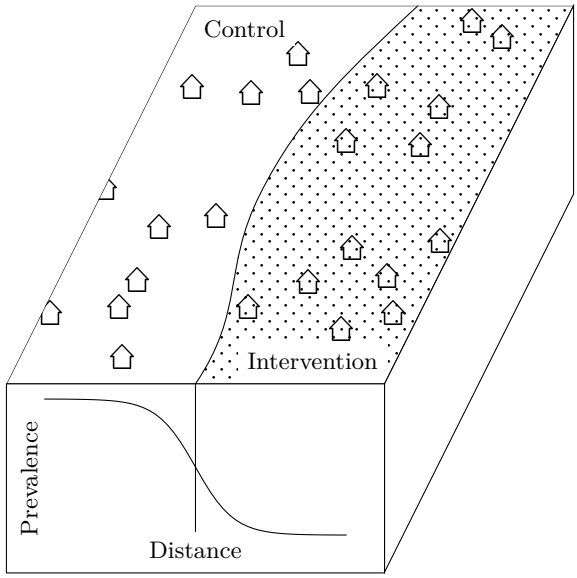


Figure IV.1: Schematic description of the effect of mosquito movement and the arising contamination in the boundary region of a CRT or SWCRT. On the front face of the rhomboid, the smooth decrease in prevalence between a control and intervention cluster (on the surface) based on the distance of a household to its nearest discordant household is visualized.

that this approach can provide unbiased and precise estimates of the contamination range and of the effectiveness, given that at least 50% of the households are at distances greater than the estimated contamination range from the nearest discordant household, hereafter called households in core.

With SWCRT designs, while the cluster size is constant, the assignment to arms changes, leading to variation in the distance to the nearest discordant household throughout the study. By accounting for this, and including a random effect for time, it is possible to extend the proposed sigmoid analysis to SWCRTs. Data with all clusters assigned to the intervention or control arm cannot be included, since the distance to the nearest discordant household is then not defined. Let Δ_{ijk} denote the distance of the jth household in the ith cluster at the kth time step to the nearest discordant household, endowed with a negative sign for the households in the control arm. For malaria prevalence, the outcome Y_{ijk} of the jth household in the ith cluster at the kth time step, with $i = 1, \ldots, c$, $j = 1, \ldots, h$ and $k = 1, \ldots, s$ (hereafter abbreviated with household ijk), can then be described by a Bayesian hierarchical model as follows [Multerer et al., 2021]:

$$\begin{aligned} Y_{ijk} &\sim \text{Binomial}(p_{ijk}), \\ \text{logit}(p_{ijk}) &= \beta_{1,ik} + \frac{\beta_2}{1 + \exp(-\beta_3 \Delta_{ijk})}, \\ \beta_{1,ik} &\sim \text{Normal}(\mu, \tau). \end{aligned} \tag{3.1}$$

For malaria incidence, a log link function together with an offset for the time at risk should be used. In this model formulation, $\beta_{1,ik}$ denotes a random effect parameter assigned to each cluster in each survey round [Hayes and Moulton, 2009], centered on the expected prevalence in the control arm, β_1. The other parameters β_2, β_3, μ and τ are assigned non-informative priors. The parameter β_2 denotes the intervention effect and β_3 can be transformed into the contamination range in km as $\hat{\theta} = \log(0.95/0.05)\beta_3^{-1}$.

This estimate $\hat{\theta}$ can be used to define the area around household ijk that influences the density of infectious mosquitoes, the effective intervention coverage $\mathcal{R}_{ijk}$. This is defined as the common density of the intervened households relative to the general density of households. The closer any other household m is to ijk, the greater is the contribution of m's intervention status to the effective intervention coverage (figure IV.2). This leads to a simple relationship between $\mathcal{R}_{ijk}$ and the distance to the nearest discordant household Δ_{ijk}. Households whose distance to the nearest discordant household is large are only surrounded by households with the same intervention status and hence $\mathcal{R}_{ijk}$ is either almost zero or one. By approximating this relationship with a sigmoid function whose

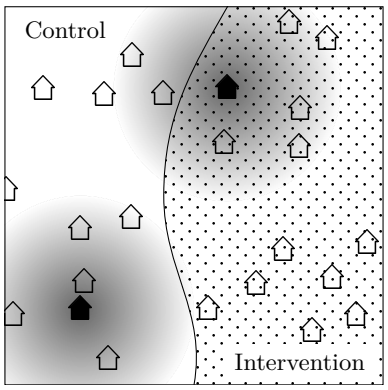


Figure IV.2: Schematic description of the effective coverage. For two households, one in the intervention arm (upper right corner) and one in the control arm (lower left corner), the area from which the effective coverage is calculated is shaded in grey. The closer a household is to one of these two households, the bigger its impact on the effective coverage (darker shade of grey). For the household in the control arm, the effective coverage is close to zero, since no intervention households are close. For the household in the intervention arm, the effective coverage is more than 50%.

growth rate depends on $\hat{\theta}$, and plugging it into the Bayesian hierarchical model (3.1), it holds that:

$$\text{logit}(p_{ijk}) \approx \beta_1 + \beta_2 \mathcal{R}_{ijk}.$$

With this formulation and after a back transformation, it is possible to describe the trial outcome as a linear function of the effective coverage, $\mathcal{R}_{ijk}$. This procedure is formally described in appendix A.4.

3.4 Analysis of the SolarMal and the AvecNet trials

Both SWCRTs were analyzed with the sigmoid random effects model from equation (3.1) and fitted using rjags [Plummer, 2019]. Uninformative priors were chosen for all parameters, and β_3 was constrained for the contamination range to be positive. All calculations were performed at sciCORE scientific computing core facility at the University of Basel under R version 4.0.0 [R Core Team, 2018].

For the SolarMal trial, the random effects parameter $\beta_{1,ik}$ was varied by survey round (as in the original publication [Homan et al., 2016]) and by survey round and households. The results are reported in terms of the reduction in odds ratio (OR), as well as the reduction in relative risk (RR). For the AvecNet trial, the random effects parameter was varied by survey round and health facility and the results are reported in terms of the IRR. Instead of the distance to the nearest discordant household, the distance to the household of the nearest discordant child enrolled in the trial was calculated, because only this data was

available. The intervention effectiveness is described in terms of the effective intervention coverage for both trials.

4. Results

4.1 SolarMal trial

For the five survey rounds, the mean distance to the nearest discordant household was 2.3, 0.9, 1.1, 1.6 and 2.7 km. For an assumed contamination range of 100 m, this results in 95, 91, 85, 93, and 98% of households in core, justifying a sigmoid random effects analysis [Multerer et al., 2021]. When including a random effect for survey round only, SMoTS were associated with a 31.9% reduction (95% credible interval (CrI): [15.3, 45.8]%) in odds ratio in the two arms, translating to a 25.2% reduction (95%CrI: [10.9, 39.0]%) in relative risk (table IV.1). The credible intervals were wider than the original confidence

	Sigmoid RE	Sigmoid RE + hh	GLMM
1−OR (in %):	31.9, [15.3, 45.8]	42.1, [32.2, 51.3]	30.0, [20.9, 38.0]
1−RR (in %):	25.2, [10.9, 39.0]	34.1, [24.4, 44.1]	-
$\hat{\theta}$ (in km):	0.146, [0.052, 0.923]	0.133, [0.052, 0.943]	-

Table IV.1: Results for the sigmoid random effects model (Sigmoid RE) and the sigmoid random effects model including a random effect for the households (Sigmoid RE + hh) for the SolarMal trial, compared to a generalized linear mixed effects model from the original analysis [Homan et al., 2016] (GLMM). The results are reported as the reduction in odds ratio 1−OR as well as the reduction in relative risk 1−RR. The contamination range $\hat{\theta}$ is only estimable for the two sigmoid random effects models.

intervals. The contamination range was estimated to be 146 m, also with a wide credible interval ([0.052, 0.923] km). With another random effect included for the household effects, SMoTS were associated with a 42.1% reduction (95%CrI: [32.2, 51.3]%) in odds ratio, and a 34.1% reduction (95%CrI: [24.4, 44.1]%) in relative risk. The contamination range was estimated to be 133 m ([0.052, 0.943] km). The effectiveness is almost linear in effective coverage (figure IV.3) rising from zero effectiveness at zero coverage to the maximal effectiveness of 34.1% when intervention households are only surrounded by other intervention households. The credible intervals increase with coverage.

4.2 AvecNet trial

The mean distance to the household of the nearest discordant child enrolled was high for all nine survey rounds, with a mean of 4.5 km. For an assumed contamination range of 100 m, this resulted in a mean of 98% of households in core. This indicates that the data

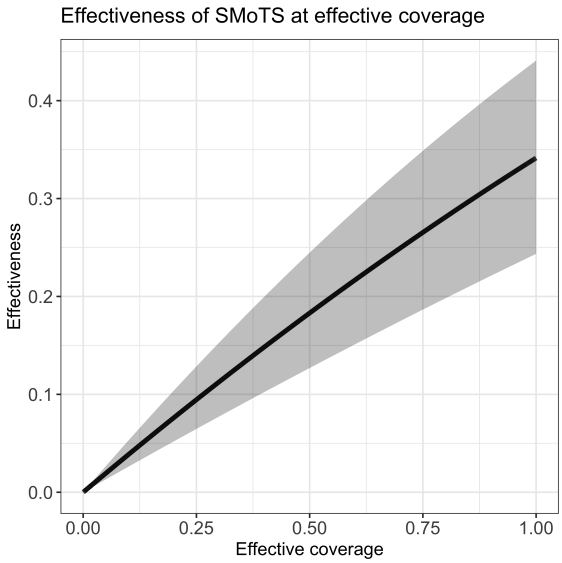


Figure IV.3: The effectiveness of Solar-powered Mosquito Trapping Systems (SMoTS) (on the y-axis) in terms of the effective coverage is visualized in black, with credible intervals in grey. The effectiveness was estimated with the model including a random effect for the household effects (Sigmoid RE + hh).

to estimate the contamination range are sparse, but a sigmoid random effects analysis can be carried out.

Adding pyriproxyfen to LLINs was associated with a reduction in incidence of clinical malaria in children of 17% (IRR 0.83, 95%CrI: [0.70, 1.00]), with credible intervals comparable to the confidence intervals from the original analysis (table IV.2). The contamination

	Sigmoid RE	GLMM
IRR:	0.83, [0.70, 1.00]	0.88, [0.77, 0.99]
$\hat{\theta}$ (in km):	0.101, [0.051, 0.745]	-

Table IV.2: Results for the sigmoid random effects model (Sigmoid RE) for the AvecNet trial, compared to a generalized linear mixed effects model from the original analysis [Tiono et al., 2018] (GLMM). The results are reported in terms of the incidence rate ratio (IRR). The contamination range $\hat{\theta}$ is only estimable for the sigmoid random effects model.

range was estimated to be 101 m, with a wide credible interval (95%CrI:[0.051, 0.745] km). The incidence rate ratio decreases as the effective coverage increases, in an almost linear fashion because of the inverse logarithmic transform, (figure IV.4). The credible intervals become wider with higher the coverage.

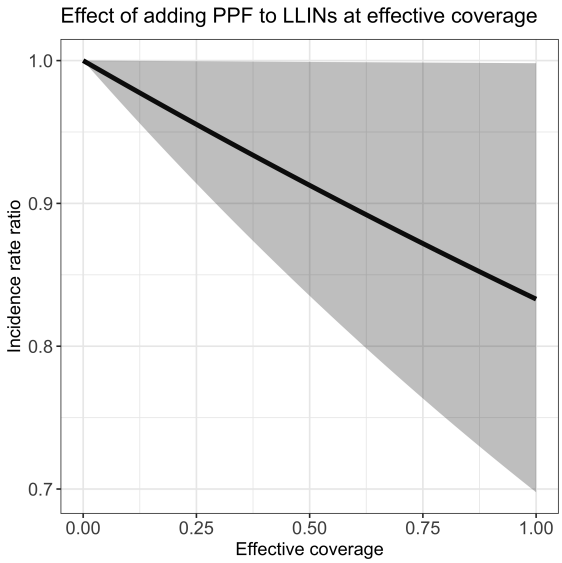


Figure IV.4: The effect of adding pyriproxyfen (PPF) to long-lasting insecticidal nets (LLINs) (on the y-axis) in terms of the effective coverage is visualized in black, with credible intervals in grey.

5. Discussion

In CRTs or SWCRTs of malaria interventions, contamination between the trial arms arises because of mosquito movement. In a conventional analysis this may bias effectiveness estimates, but this can be corrected with an appropriate analysis, such as a Bayesian hierarchical model with a sigmoid function for effectiveness as a function of distance to the nearest discordant household, that was recently proposed for CRTs [Multerer et al., 2021]. This model yields a closed-form contamination range that quantifies the contamination arising from mosquito movement between trial arms, and adjusts the main estimate of effectiveness for contamination, eliminating the need for buffer zones.

The proposed analytical approach is tailor-made for malaria interventions where transmission can be geolocated to the host's primary residence, and the main source of contamination between clusters arises from dispersal of adult female *Anopheles* mosquitoes, for which the proposed model, corresponding to mosquito dispersion by diffusion [Malinga et al., 2019a, Malinga et al., 2019b] is a reasonable approximation. In nature dispersal will vary between sites, within sites and by season, and depends on the extent and spatial distribution of aquatic habitats, households and alternative blood sources [Lutambi et al., 2013], as well as wind strength and direction and obstacles in the environment. For both interventions considered here, with effects mainly depending on mosquito densities, contamination was considered to be symmetrical. The intervention may protect nearby

nonusers, while users with many nearby nonusers have reduced intervention effects. With this model, a difference between homogeneously distributed and clustered interventions on the overall intervention effect is not distinguishable [Lutambi et al., 2013]. The same modelling approach might be applied where the intervention itself is designed to be dispersed by mosquitoes (for instance sterile insect techniques) and even with human-side interventions such as mass drug administration or mass vaccination, though in the latter cases contamination is less important relative to the overall efficacy since more of the impact is due to the direct effect of individual protection. Contamination also arises in CRTs of many other health interventions, but where transmission is not by night-biting mosquitoes the geometry is likely to be more complicated. For instance, where the intervention is behavioural and the primary source of contamination is social (and hence non-spatial), or with directly transmitted infections or those transmitted by less mobile and day biting *Aedes* mosquitoes (where infections often acquired at workplaces or schools, making geographically congruent clusters less desirable), different models of contamination are needed. In any given trial, the appropriateness and fit of the chosen contamination model should be carefully evaluated.

In this work, the sigmoid model is applied to two SWCRTs, the SolarMal and AvecNet trials. SWCRTs can be inferior in terms of power or bias compared to parallel designs and might be vulnerable to imprecision caused by temporal trends in underlying disease rates [Wolbers et al., 2012, Kotz et al., 2013] but may be required because of logistical, practical or financial constraints [Wilson et al., 2015, Mdege et al., 2012] (for example, in the SolarMal trial an objective was to assess whether interruption of transmission would occur at complete coverage [Silkey et al., 2016]). Because of the changing boundaries, the analysis of contamination effects in SWCRTs is more complicated, but in principle SWCRT data could be used to analyze changing patterns of contamination in time and place. At the same time, it is unclear how the imbalance between arms affects the precision and bias of the resulting estimates.

A reanalysis of the SolarMal trial yielded a slightly higher estimate of effectiveness than was reported in the original trial analysis [Homan et al., 2016], but with less precision. Adding a random effect for the households increased the estimate of effectiveness with reasonably wide credible intervals. Also for the AvecNet trial, the reduction in incidence of clinical malaria in children was higher than in the original analysis [Tiono et al., 2018], with only slightly less precision. The contamination range was consistently around 140 m in the SolarMal trial and around 100 m in the AvecNet trial, which is much less than the maximal distance *Anopheles* mosquitoes can fly [Guerra et al., 2014, Service, 1997].

The SolarMal trial was conducted in a small, densely populated area and had many small clusters. The AvecNet trial, in contrast, was conducted in a much larger area, with a

population density 10 times lower than that in the SolarMal trial (around 50 people per km^2 compared to more than 500 people per km^2). The settlement patterns where these trials were conducted are also different: in the region where the SolarMal trial took place around Lake Victoria, households are scattered, while the area where the AvecNet trial was conducted has villages with tight aggregations of houses, typical of the West African Sahel. These factors affect the percentage of households in core, the percentage of households unaffected by the contamination across cluster boundaries, where a balance is needed for the proposed analysis to yield unbiased and precise estimates. In the AvecNet trial, a subset of children was chosen from each village, to allow for clusters to be chosen as administrative units. This resulted in a high percentage of households in core, though this number is not comparable to the SolarMal trial, because only the distance to the household of the nearest discordant child was calculated. Informed by a previous simulation study [Multerer et al., 2021], it is assumed that with so little information from the boundary regions, the contamination range cannot be estimated reliably and the proposed model is not working.
Like AvecNet, many trials define clusters based on administrative units with cluster boundaries passing through uninhabited areas. However, for contamination effects to be estimable, the trial must be designed to collect information from the boundary zones where contamination is likely. If cluster boundaries can pass through inhabited areas, as in the SolarMal trial, equal-population clusters can be assigned giving a more balanced design with optimal power, therefore requiring fewer participants. When there is contamination there is also empirical information about every level of local coverage from within either a CRT or SWCRT, even without universal overall coverage. This enables extension of the analysis using kernel density estimation to infer from the contamination range how effectiveness depends on intervention coverage. These estimates could be used to support allocation decisions when interventions are deployed, but where resource constraints mean universal coverage is not achievable.

6. Conclusions

It was shown how trials with anticipated contamination effects arising from mosquito movement can be analyzed to give unbiased and precise estimates of effectiveness. Guidance is now needed on how to plan trials with adequate power and precision to allow for contamination. Without the need for buffer zones, or for clusters to correspond to villages, cluster size can be reduced to a minimum determined by operational factors or contamination effects, reducing the required numbers of participants in field trials of malaria

interventions. This should lead to more cost-efficient trials and a better understanding of the indirect effects of interventions in protecting nearby nonusers.

Chapter V

DESIGN OF TRIALS FOR MALARIA CONTROL INTERVENTIONS WITH CONTAMINATION

Working paper

1. Introduction

The previous two chapters III and IV on the analysis with a sigmoid random effects model for CRTs and SWCRTs focused on the analysis of already completed trials of malaria interventions. In this chapter, we use the insights from these chapters to outline how trials should be planned in the presence of contamination arising from mosquito movement. A crucial point in this is the choice of clusters. Usually, clusters correspond to administrative units or well-separated communities. This may lower logistical difficulties in distributing the intervention but often depends on an educated guess on how to draw cluster boundaries. In settings with no clear separation of communities it is proposed to rely on geospatial data combined with local geographical knowledge [Hayes and Moulton, 2009].

For the proposed sigmoid random effects model to result in unbiased and precise estimates of intervention effectiveness, it is necessary for cluster boundaries to pass inhabited areas. In this chapter, we indicate how clusters should be chosen and derive algorithmic approaches for assigning households to clusters. We give guidance for practitioners and then use the derived results to redesign the AvecNet trial, a trial that assessed the effect of adding pyriproxyfen to long-lasting insecticidal mosquito nets in rural Burkina Faso [Tiono et al., 2015, Tiono et al., 2018]. This trial was analyzed in chapter IV and it was concluded that contamination effects were minimal, with insufficient data from the cluster boundary regions to reliably estimate the contamination. Hence, we hypothesize

that this trial could have been smaller in terms of households enrolled in each cluster if clusters were not chosen as entire villages, but as arbitrary geographically contiguous areas.

2. Trial design

A malaria control intervention is to be tested in a CRT or SWCRT. The intervention is either assigned to a geolocated household, for example an intervention of insecticide treated nets, or to an area, such as the deployment of insecticide. To prevent contamination of both human and mosquito movement, clusters are to be defined as geographically contiguous areas corresponding to the places where people get infected, with possible operational constraints on the cluster size. The outcome of the trial is either epidemiological, such as disease prevalence or incidence, or entomological, for example adult mosquito or oviposition trap counts, and measured at geolocated households.
The analysis will be carried out with a sigmoid random effects model introduced in chapters III and IV, adapted to the outcome of choice via a link function and possible offsets. Hence, clusters need and should not be separated by natural or artificial buffer zones, because information on the contamination between clusters is needed to estimate and account for it.

2.1 Size of clusters

A crucial part in the designing of any CRT or SWCRT is to determine the size and the number of clusters needed to achieve the desired power. In practice, the size of clusters is often predefined based on operational or financial constraints and the other quantity is then computed via sample size formulae [Donner et al., 1982, Hayes and Bennett, 1999, Hayes and Moulton, 2009]. Constraints on the cluster size may arise in any of the following ways:

(i) clustering of the intervention is required only because of contamination effects, which therefore determine the size and configuration of clusters,

(ii) the size of clusters is fixed by operational factors, for instance clusters may be centered around schools or health facilities,

(iii) there are some operational constraints, for instance there may be a minimum cluster size corresponding to the work that a field team can conduct in one day, or week, but cluster configuration is not fixed.

If contamination due to mosquito movement is anticipated, clusters need to be big enough for the contamination effects to be measurable. In a previous simulation study (chapter III) it was found that the key factor for this is the percentage of households in core,

ω, the percentage of households whose distance to the nearest discordant arm is greater than the estimated contamination range $\hat{\theta}$. It was concluded that, for a sigmoid random effects model to result in unbiased and precise estimates of effectiveness, at least 50% of households should be in core. This is illustrated in figure V.1, for the special case that 50% of the households in each cluster are in core. If this is not feasible, either because the

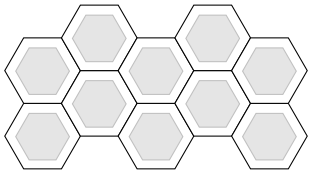


Figure V.1: Illustration of the requirement of 50% of households in core for designing a trial allowing for contamination. In each cluster (with boundaries delimited in black), 50% of the area is in core (in grey) and 50% is within the contamination range from the cluster boundary (in white).

assumed contamination range is too big or because of other constraints on cluster size, an iteration between an initial choice of cluster size h, calculation of sample size, cluster assignment and calculation of the percentage of households in core, ω is needed. If $\omega \approx 50\%$, no adjustment has to be made, if not, the initial proposal of h should be adjusted. If multiple candidate designs fulfil this requirement, one can randomly be chosen from them.

2.2 Number of clusters

Under the constraint of an appropriate number of households in core, a sigmoid random effects model results in unbiased and precise estimates of effectiveness. If this constraint is considered in the trial design, the overall size of the trial remains the same compared to a trial of an intervention with no contamination between clusters and standard sample size formulae to find an appropriate number of clusters can be used. This involves inflating the number of observations that would be needed in an individually randomized trial by the design effect, a measure for the increase in variance for using a more complex trial design [Kish, 1965, Hayes and Moulton, 2009]. For a binary outcome for example, this reads $DE = 1 + (h-1)\rho$, where h denotes the cluster size and ρ the intracluster correlation coefficient (ICC) [Hayes and Moulton, 2009, Donner et al., 1982, Hayes and Bennett, 1999] that is usually obtained from the study site or from similar areas. The design effect is calculated independently of the intervention and the contamination of an intervention hence cannot be considered. However, underlying the calculation is the assumption that clusters are uncorrelated, a theoretical requirement that will be violated at the cluster boundaries.

The required number of clusters c can then be calculated from standard sample size for-

mulae [Donner et al., 1982, Hayes and Bennett, 1999, Hayes and Moulton, 2009]. If the main outcome is a proportion, these read

$$c = 1 + n_{\text{ind}} \frac{DE}{h}, \tag{2.1}$$

where n_{ind} denotes the sample size for an individually randomized trial. For proportions p_C and $p_I = p_C(1 - E_s)$ with given effect size E_s, and standard normal distribution values for the upper tail probabilities $\alpha/2$ and β, $z_{\alpha/2}$ and z_β, it is defined as

$$n_{\text{ind}} = (z_{\alpha/2} + z_\beta)^2 \frac{p_C(1 - p_C) + p_I(1 - p_I)}{(p_C - p_I)^2}. \tag{2.2}$$

2.3 Algorithmic approach to cluster assignment

If there are no operational constraints on the cluster assignment, an algorithmic approach can be chosen to assign households to clusters. This leads to cluster boundaries passing through inhabited areas, which is needed to attain the minimum of $\omega \approx 50\%$. A controllable number of similar size clusters with high similarity between observations in the same cluster is desirable. Several algorithms exist that lead to an optimal allocation of the same number of observations to clusters [Bradley et al., 2000, Zhu et al., 2010], but they have not been used in trial design so far. Three computationally cheap and practicable algorithms are described below on how to divide a study area in an ad-hoc way into clusters.

(i) Run a travelling salesman problem (TSP, R package [Hahsler and Hornik, 2007]) heuristic starting at an arbitrary household and then group households along the TSP path into clusters of size k, as was proposed by Silkey *et al.* [Silkey et al., 2016, Hiscox et al., 2016] to construct clusters for the SolarMal trial.

(ii) Iteratively select one household and construct a cluster of size k with its $k - 1$ nearest neighbors (NN) before removing these points from the data set. Selecting the household furthest away from all the remaining others as cluster center will often lead to connected clusters, in a "fish scale" manner [Monlong, 2019].

(iii) Run a standard k-means algorithm, an iterative procedure that partitions observations based on their distance to cluster centroids.

The first two algorithms, TSP and NN, are constrained to clusters of same size. The k-means algorithm has more liberty and will result in slightly variable cluster size and more households in core, because cluster boundaries are less likely to pass densely inhabited

areas. We recommend simulating several cluster allocations and choose among the ones that are feasible [Halloran et al., 2017].

2.4 Practical guide for the trial design

Trials can be designed following the steps outlined in table V.1. Steps A–C represent

Step	Parameter		Choice of value
A	α	Two-sided confidence level	Conventionally, 0.05
B	$1-\beta$	Power	Conventionally, 0.8 or 0.9
C	E_s	Required effect size to detect	As proposed by investigator
D	ρ	ICC	Obtain from other studies in comparable settings
E	p_C	Baseline prevalence	Local data from study area
F	$\hat{\theta}$	Contamination range in km	Obtain from other studies in comparable settings
G		Geolocations of households in study area	Local data from study area (ground survey and/or satellite images)
H	h	Proposal for the number of households in cluster	Adjusted to meet operational constraints or maximize power
I	c	Required minimum number of clusters	Standard sample size formula for prevalence (2.1)
J		Cluster boundaries	Fixed, based on operational constraints, or derived using algorithmic approach
K		Randomization	Random assignment of clusters to arms
L	ω	% of households in core	Calculate from geolocations, randomization and $\hat{\theta}$

Table V.1: Summary of the steps A–L needed for designing a CRT for vector control interventions with contamination.

simple choices that are fixed by the investigator, and do not entail any calculations. Steps D–G require data, either from the study site or similar areas. If no good estimates can be obtained, steps D–G can be varied to assess the implications on sample size. The remaining parameters in steps H–L can be adjusted to optimize the design. Step H requires an initial proposal for the number of households in each cluster, h, as described in subsection 2.1. If there are no operational constraints on cluster size, the initial proposal for the cluster size h should be made such that $\approx 50\%$ of households in the trial are in core. The required number of clusters c (step I) can then be calculated from standard sample size formulae (2.1). For the determination of cluster boundaries, step J, an algorithm as described in subsection 2.3 can be used. A combination approach is also possible, in which different candidate assignments of cluster boundaries are drawn up and evaluated against

operational criteria. Conferring with local stakeholders at this stage does not compromise the randomization. The true percentage of households in core (step L) corresponding to the proposal of h, ω, is determined after randomization. The cluster size h should then iteratively be adjusted to respect the minimum of $\omega \approx 50\%$. Multiple candidate designs will fulfil this requirement, because ω is a function of the randomization, and an allocation scheme can randomly be chosen from them. Alternatively, steps J–L can be carried out for a range of different values of h and a design that fulfils $\omega \approx 50\%$ can be chosen.

3. Applications

3.1 Visualisation of the algorithmic cluster assignment

To illustrate the three different algorithms of cluster construction, we use simulations of CRTs. These simulations are designed in the same way as described in chapter III. In short, on a domain of size $\eta \times \eta$ km^2, inhomogeneous human populations of N households are generated with a (modified) Thomas cluster process [Thomas, 1949, Baddeley and Turner, 2005], that first generates a uniform Poisson point process of parent points with intensity α_1 that are then replaced with a cluster of offspring points. The offspring points are normally distributed around cluster centers and the clustering can be controlled by its mean α_2 and standard deviation α_3, that can be used to vary the settlement pattern.
The simulation of mosquito movement leads to a smoothing of the expected prevalence via a simple diffusion process, modeled with two dimensional Gaussian functions of distance between houses. The standard deviation σ describes the distance mosquitoes move during the extrinsic cycle of the parasite and translates to the contamination range θ, the range in km over which there is significant dispersion of infections as $\theta = \Phi^{-1}(0.95) \times \sqrt{2\sigma^2}$, where Φ^{-1} is the the quantile function of a standard normal distribution. The allocation of h households to c clusters is performed with an algorithmic approach, as described above. Half of the clusters are then assigned to the intervention arm. A summary of the trial parametrization is given in table V.2.
Three different CRT simulations, with three different settlement patterns, are generated with these values. For each of these three simulations, the three algorithms are run, resulting in nine different CRT simulations, illustrated in figure V.2. All simulations were performed at sciCORE scientific computing core facility at the University of Basel under R version 4.0.0 [R Core Team, 2018].
In this example, the k-means algorithm results in a higher percentage of households in core than the other two algorithms. This is unsurprising, since this algorithm has a certain flexibility to allow for varying cluster size. What is interesting is that the settlement pattern does not strongly influence the percentage of households in core. This is confirmed

Parameters used to generate the data sets		Values
η	Domain size	7.5 km
N	Number of households in trial	4000
α_1	Intensity of parent points	4
α_2	Mean number of offspring points	40
α_3	Standard deviation of offspring points	0.225, 0.175, 0.125 km
σ	Standard deviation of Gaussian function	0.106 km
θ	Contamination range for σ	0.25 km
c	Number of clusters	50
h	Households per cluster	80
	Seed, randomly sampled	5792
	Algorithm for cluster assignment	TSP, NN or k-means

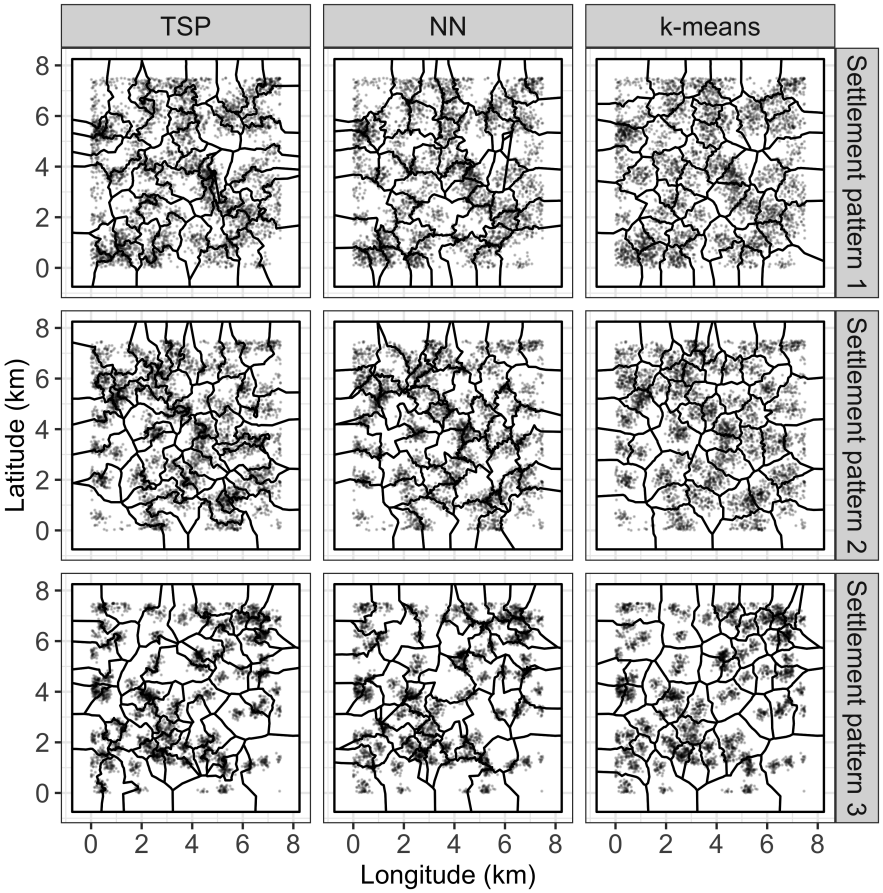
Table V.2: Summary of the parameters used to generate the three data sets.

in other simulations as well. For all three algorithms, the number of clusters is predefined, and a densely inhabited area might be divided into several clusters. In conclusion, the simulations show that no algorithm is strongly preferable. If equal cluster size is desired, a TSP or NN should be chosen, if not, then a very simple and fast k-means algorithm is best.

3.2 Redesign of the AvecNet trial

We redesign the AvecNet trial as a CRT following the Steps A–L from section 2.4 to answer the following question: what is the size of the smallest trial possible in terms of the number of households enrolled to attain the same power? This trial has been described in detail in chapter IV and in the original publications [Tiono et al., 2015, Tiono et al., 2018]. The parameter values for steps A–E can be found in the study protocol [Tiono et al., 2015]. The two-sided confidence level was chosen to be 5%, together with a power of 90% and a required effect size of 25%. The coefficient of variation was assumed to be 0.25, estimated from malaria incidence data from villages in the same region, and the baseline incidence was chosen as 1.5 malaria episodes per child per year, as was found in a previous study in Burkina Faso. In chapter IV, the contamination range for the AvecNet trial was found to be 0.101 km, with a credible interval of [0.051, 0.745] km. For step G, the geolocations of the households of the random sample of children enrolled in the study were collected, but we do not have the geolocations of all the households with children living in the study area.

In step H, a proposal for the number of households has to be made. In the original study, a mean of 50 children were monitored in each cluster, some of them living in the same household. For simplicity, it is assumed that each record of a child is a household



	TSP	NN	k-means
Settlement pattern 1	53%	57%	79%
Settlement pattern 2	56%	49%	79%
Settlement pattern 3	67%	48%	76%

Figure V.2: On top, the different algorithms for cluster assignment (travelling salesman heuristic (TSP), nearest neighbors (NN), and k-means) are illustrated for three different settlement patterns (from very homogeneous distribution of households on top to heterogeneous distribution on the bottom). On the bottom, the respective resulting percentages of households in core are listed. The parameters used to generate this data can be found in table V.2.

and the geolocations of the children living in the same household are slightly jiggled. As initial proposal, $h = 40$ is taken, the lowest number of children monitored in a cluster in the study [Tiono et al., 2015]. The required minimum number of clusters can then be calculated from standard sample size formulae [Hayes and Moulton, 2009], leading to a required minimum of 19.87 clusters per arm, hence an overall minimum of 40 clusters. The cluster boundaries are to be determined with an algorithmic approach. Every village in the study area is affiliated to one of six health centers. To prevent confounding, this

has to be considered, since each cluster should be affiliated with only one health center. A hierarchical approach with six metaclusters, the catchment areas of the six health centers, is chosen, and the greatest possible number of clusters within this metacluster is assigned. To correspond to the original study, the randomization is balanced to ensure that, within each metacluster, half of the clusters are intervened [Tiono et al., 2015]. The percentage of households in core can then be calculated from the geolocations, the randomization and the contamination range. Since the geolocations only consist of the geolocations of the enrolled children, only the distance to the household of the nearest discordant enrolled child can be calculated. It is assumed that in the mean, these two parameters are proportional.
We evaluated a series of candidate designs with the three different algorithms for cluster assignment. All led to a very high percentage of households in core, around 98%. Even with a more conservative estimate of the contamination range of 0.75 km, the upper credible interval value, around 93% of the households were in core. This result must be viewed with care, since it represents the percentage of households in core with respect to the nearest discordant household with an enrolled child. For a cluster size of $h = 40$ children per cluster, the overall study would consist of 1 600 children that have to be enrolled, instead of 1 980 that were enrolled at the beginning of the AvecNet trial, not counting all the children not enrolled but living in the study area.

4. Discussion

Preventing contamination in the trial design of a vector control intervention is possible, but often leads to very large trials. Clusters need to be separated by wide buffer zones, possibly leading to heterogeneity and the investigator needs information on the magnitude of contamination to choose the buffer zone. In contrast, the goal of this chapter was to give guidance on how to design trials that do not prevent contamination, but that allow for it, enabling the analysis to adjust for and estimate its magnitude. We showed how to perform sample size calculations and how to construct clusters with several algorithmic approaches. Table V.1 summarizes these findings and provides guidance for investigators. Corresponding R code is provided and used to redesign the AvecNet trial with smaller same size clusters.
The strength of our trial design proposal is that clusters need not be chosen as well-separated communities. We propose that at least half of the trial households ($\omega \approx 50\%$) are at distances greater than the assumed contamination range from the intervention boundary for the true intervention effect to be estimable. But the sigmoid random effects model requires data from the cluster boundary regions between intervention and control arm, allowing cluster boundaries to pass inhabited areas. Hence, clusters can be chosen of

equal size with an algorithmic approach. On the other hand, the theoretical assumption of independence between clusters is violated. By allowing cluster boundaries to pass through inhabited areas, there also might be an increase in human movement across cluster boundaries or a bias due to the social behavior of the participants, who might for example lend their net to their neighbor. This trial design is hence tailored for interventions where the major source of contamination is due to spatial diffusion on a small scale, as is the case with malaria interventions against *Anopheles* mosquitoes.

The redesigning of the AvecNet trial as a CRT showed that this trial could have been smaller in terms of the number of households enrolled. It would be possible to enroll all children of the study area in the trial, ensuring same size clusters with an algorithmic approach. However, several simplifications were made that would need consideration. We did not consider the age distribution of the children whereas in the original trial design, children were enrolled with equal numbers of children aged 6 to 35 months and 36 to 60 months. Furthermore, we assumed that from each household, there is only one child enrolled. Because we did not have the geolocations of all the households in the study area, we only calculated the distance to the nearest discordant household with a child enrolled. This does yield a biased percentage in core, which was extremely high, even after redesigning the study.

This chapter closes the methodological development of chapters III and IV by showing how to facilitate the proposed sigmoid random effects analysis and how algorithmic approaches and simulations from them can aid in trial design. But most of all, this chapter shows what has been hypothesized already in chapters III and IV: it is indeed possible to have cost efficient trials of malaria interventions with a reasonable number of participants and conducted in a relatively small area.

Chapter VI

DISCUSSION

There is increasing recognition of the need to use models and simulations in optimally designing trials of interventions against mosquito-borne diseases [Halloran et al., 2017]. Complex transmission and intervention dynamics can be captured, simulations of trials can aid in understanding the impact of a chosen design and models can be used to explore possible explanations for the observed effects. Models, simulations and their analysis hence should be in an interplay with the design and analysis of trials of mosquito-borne diseases. A model, deterministic or stochastic, parametric or non-parametric, that incorporates the disease, human or mosquito dynamics forms the basis for simulations and data that resembles a trial is randomly drawn from the model. This data is analyzed with an approach of choice, either simple summary statistics or with a more sophisticated statistical approach. Based on these three steps, a trial fulfilling the quality standards [Wilson et al., 2015, WHO, 2017b] is designed. After rollout, this trial is analyzed, leading to conclusions on the intervention effectiveness. At the same time, the trial analysis can inform the model choice with updated parameter estimations and model validation. This interplay can be visualized in a cycle, see figure VI.1.

In this thesis, we showed how the development of new models of intervention dynamics and their analysis aids in designing and analyzing trials of mosquito-borne diseases with contamination. We focused on several modeling approaches, developed a new approach to analyze cluster randomized trials and tested its robustness with simulations. These insights fed into our proposal for the design and analysis of new trials of mosquito-borne diseases, leading to a better understanding of interventions and their indirect effects. In the following sections, we will discuss our contributions to the interplay between models and trials visualized in figure VI.1 in more detail. We will touch upon points that were not discussed in the chapter discussions, give a new perspective of the obtained results and outline potential future work.

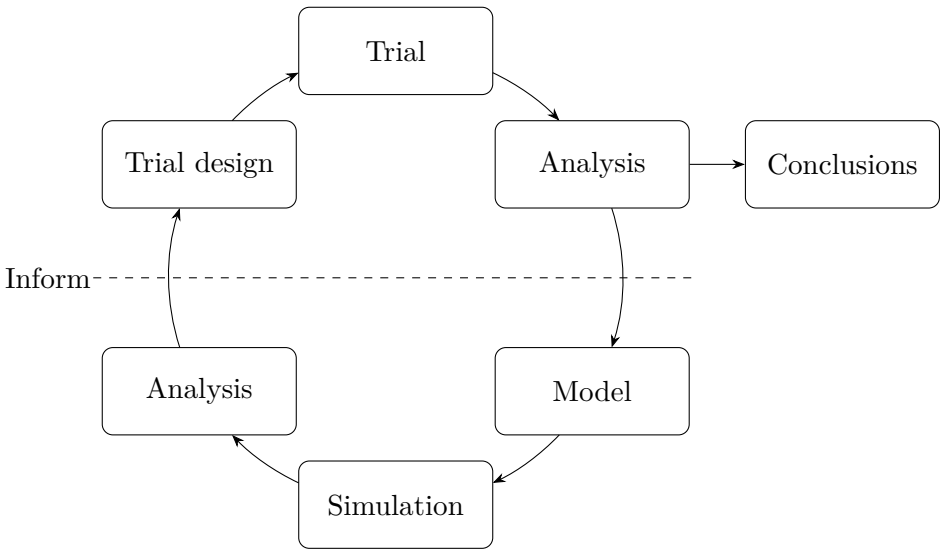


Figure VI.1: Cycle visualising the desirable interplay between trials and models. Models, simulations and their analysis can be used to inform the trial design and analysis, which in turn can inform the choice of the models. This is illustrated with a cycle, the dashed line denoting the boundary between simulations and real life.

1. Models and simulations of mosquito movement

We used two different approaches to model mosquito dynamics and mosquito movement in this thesis. In chapter II, we developed and calibrated a PDE model describing the dynamics of an *Aedes aegypti* population around the circumference of an island under a modified male mosquito release. Since any program testing the effectiveness of a modified male mosquito intervention is faced with limited production and release capacity of modified males [Bouyer et al., 2020], we used optimal control theory to identify an optimized release strategy with a limited availability of modified males per day. Because the optimal control solution with a daily limit of modified males for the island was intractable, we adapted the optimal control solution without this constraint by delaying the start of the intervention in some points, and retrospectively enforced the constraint. We concluded that, to eliminate *Aedes aegypti* from an island, the best approach is to target high mosquito density areas first and then move the focus in both directions along the periphery of the island. Initially, a high number of mosquitoes has to be released, followed by a release proportionally to the decreasing female population.

In chapter III, we implemented a more practicable model to describe mosquito movement that easily adapts to a given geography. We used kernel functions to describe the likely movement of *Anopheles* mosquitoes around a household [Malinga et al., 2019a, Malinga et al., 2019b] and calculated the common density of all kernel functions to draw inference

on the exposure to malaria infection at a household. In this model, the mosquito density at a household is influenced by the distance to surrounding households and an intervention effect can be simulated. We defined the significant dispersion of infection, the contamination range, as the radius of the circle around each household whose area covers 95% of the density function. Informed by this model, we also defined the effective intervention coverage at a given household in chapter IV. With this simple setup, it was possible to group households to clusters and simulate the contamination of intervention effects in a CRT. This formed the basis for a simulation study in which we simulated the human and mosquito populations, varied the mosquito movement, the effectiveness of intervention and the number and size of clusters per arm.

2. Development of a nonlinear model to analyze contamination in trials

The simulations of CRTs with mosquito movement motivated the proposed nonlinear model of the distance to the nearest discordant household for the analysis of CRTs, derived in chapter III. In the simulations, we chose bivariate Gaussian kernels to describe the mosquito movement. They can be approximated by logistic kernels, whose cumulative density function is the sigmoid function we chose as basis for our model, a very simple function often used in the modeling of biological processes. With such a nonlinear model type, it is possible to adjust the estimate of effectiveness for contamination caused by mosquito movement. This model type also yields a closed-form contamination range that quantifies the measurable distance over which contamination biases the estimate of effectiveness. We define this as a cut-off of a transformation of the growth rate from the sigmoid function when 95% of the growth are attained, in accordance with the mosquito model described above. The conducted simulation study showed that the model indeed yields unbiased and precise estimates of effectiveness, if more than 50% of the households are at distances greater than the contamination range from the discordant intervention arm. Chapter IV extended this model to analyze SWCRTs. We then introduced the estimation of the intervention coverage by means of the contamination range. This goes back to the kernel approach chosen to simulate the mosquito movement in CRTs. By assigning the coverage to each household, the output from the proposed sigmoid random effects model can be transformed to an estimate of the intervention effectiveness at any intervention coverage.

These two chapters complete figure I.2 by presenting a nonlinear model approach to analyze CRTs of mosquito control interventions. To our knowledge, this is the first time such an approach has been used. Since the use of nonlinear models in trial analysis is not

common, we provide R code for the user for the analysis of trials to facilitate the use. We propose to fit the model within the framework of a Bayesian hierarchical model. This is because standard frequentist libraries to fit nonlinear mixed effect models (for example the R package nlme [Pinheiro et al., 2018]) are not tailored to include other than normally distributed outcome variables. Extensions and adaptations of the model are possible. The dispersal of mosquitoes depends on the availability of resources and on other exogenous factors. The contamination range is thus likely to vary in time, space and with the intervention considered. It would be possible to consider it as a random effect in the proposed model, or, interpret it as being population density dependent. The starting point for this was made by considering the intervention effectiveness as a function of the effective intervention coverage. But to describe the contamination range as a function of the population density, another indicator than the nearest discordant household is needed as a basis for the model.

3. Implications for the design of trials for mosquito-borne diseases

The theoretical results we described so far on the development of models of mosquito movement and nonlinear models for the analysis of trials have implications on the design of trials. In chapter V, we used the insights from chapters III and IV to show the implications of conducting such a nonlinear analysis for the design of future CRTs and SWCRTs. We showed how sample size calculations are valid if the trial design fulfils that more than 50% of the households are at distances greater than the contamination range from the discordant intervention arm. This needs to be verified after clusters have been allocated and an iterative procedure may be necessary. It also implies that cluster boundaries should pass through inhabited areas. We showed that an algorithmic approach aids in the cluster allocation and presented three possibilities of simple algorithms.

We limited the work on the design and analysis of trials to malaria interventions, because of the biological properties of *Anopheles* mosquitoes that lead to trials where contamination mainly arises from mosquito movement between geographically contiguous areas. In general, these results can be extended to trials of vector-borne diseases with an entomological outcome. If the vector itself instead of the carried diseases is in focus, contamination sources such as human movement are less relevant, and any intervention where spatial contamination arises on a small scale could be analyzed with the proposed model. This might find applications in agriculture and with trials assessing the effectiveness of a modified male mosquito release in suppressing an endemic female *Aedes* population. We now connect the results of chapters III, IV and V with the results for the optimized release

strategy of modified males from chapter II to outline a potential plan for a trial of modified male mosquitoes.

The goal of a trial assessing the effectiveness and the potential for eliminating the local population of female mosquitoes with a modified male mosquito release is

(i) to assess the short-term effectiveness in comparison with random control units,

(ii) to determine the best deployment strategy,

(iii) to assess whether an area-wide program might eliminate the endemic female population.

To prevent immigration of females, the study region should be isolated, for example by choosing an island. The outcome would entail entomological and epidemiological measures, assessed with trapping devices [WHO and IAEA, 2020, Wilson et al., 2015, Bouyer et al., 2020]. Such a trial faces several particular challenges. The time frame between production and release of modified males is narrow, production capacity might be limited to a fixed total number of modified males per time unit and the release has to be recurrent to suppress the population (although the number of released modified males in one place can be controlled for). To test both the short-term effectiveness (i) and the potential for elimination (iii), two trial phases are needed. First an investigation phase where releases are randomized in space based on a predefined plan, allowing for the assessment of the optimal deployment strategy, and an adaptive control and elimination phase, with the objective to reduce the population to zero as quickly as possible.
In the investigation phase, a CRT or SWCRT is conducted. Modified male mosquitoes are released in the cluster areas and the entomological outcome can be measured with geolocated adult mosquito traps or with oviposition traps to assess the number of females or the number of viable eggs. If enough resources are available, a standard CRT can be carried out, allocating half of the clusters to the intervention arm and releasing modified males based on the endemic density of females. Based on the insights from chapter II, there is no need to keep the release constant, and after an initial high release, the release effort can be lowered. Should not enough resources be available to target all clusters at once, an SWCRT can be conducted, starting with all clusters in the control arm and allocating them to the intervention arm whenever sufficient resources are available. This is a randomized discrete analogy to the optimizied strategy from chapter II.
In accordance with chapter V, an initial proposal for the number of households in each cluster should be made, in correspondence with operational constraints. The number of clusters required can then be calculated and cluster boundaries can be drawn with an algorithmic approach, with the constraint that clusters should be of similar geographical

size if the household distribution is very heterogeneous. If the initial proposal for cluster size does not lead to more than 50% of households in core, another proposal should be made. An initial analysis of the investigation phase should then indicate whether such a release is justified and conditional on this outcome, the trial can move to the next phase [Bouyer et al., 2020]. The intervention effectiveness will be biased, because both the modified males and the endemic females can move across the cluster boundaries. With the sigmoid random effects model proposed in chapter III, this can be accounted for. The distance to the nearest discordant household should be replaced by the distance from a trap to the nearest discordant cluster area, because of the nature of the intervention.

In the adaptive control and elimination phase, the goal is to assess whether elimination of the female population is feasible. This is an all-or nothing outcome at the level of the whole area and requires scale-up to universal coverage. The work on the optimal release strategy on an island from chapter II supports this scale-up and rollout. Based on the density of the endemic female population (that will be varying depending on the chosen trial design of the first phase), high density areas can be targeted first and the focus can then be moved, preventing reintroduction in the already cleared areas, as it is described in chapter II.

This potential plan is intentionally vague, a more detailed plan will depend on the trial setting, the production and the timelines of the investigator. Nevertheless, this proposal answers questions an investigator might have: Bouyer *et al.* [Bouyer et al., 2020] formulated several outstanding questions arising from their work on a phased conditional approach for the use of modified males. One of them reads "How can we obtain reliable measures of female mosquitoes dispersal and of its epidemiological significance? How can the dispersal be accounted for when designing epidemiological trials? In other words, how can we measure the epidemiological impact of the SIT (sterile insect technique) against mosquitoes at a cluster level when this technique can only be successful when applied on an area-wide basis?". This question is answered with our proposed first investigation phase and its analysis. The second trial phase is similar to the proposed trial strategy in wave form by Bouyer *et al.* [Bouyer et al., 2020] where, starting from one point, local elimination is achieved and the intervention is moved, retaining temporary buffer zones. However, by proposing two trial stages, full randomization of the intervention assignment is achieved, allowing for a high-quality trial.

4. Analysis of trials of malaria interventions

We reanalyzed three trials of different malaria interventions in chapters III and IV. Here, we compare key points of these trials and the obtained results. The Navrongo trial,

assessing the effect of insecticide treated nets in a CRT in northern Ghana conducted 27 years ago [Binka et al., 1996], was recently reanalyzed with linear and geostatistical models to test for contamination effects [Jarvis et al., 2019]. In this work, which is not a core part of this thesis, we concluded that despite compelling evidence of contamination, the main conclusions of the trial remain unaffected by spatial model specifications because clusters were very big. This work provided us with important background information for our analysis, supporting the use of only simple correlation structures in the main analysis, hence considerably simplifying the statistical fitting and making the conclusions more robust. In chapter III we therefore reanalyzed the Navrongo trial with the sigmoid random effects model that was developed in parallel. In chapter IV, we reanalyzed the SolarMal [Homan et al., 2016] and the AvecNet [Tiono et al., 2018] SWCRTs.

The areas where the three trials were conducted differ in the settlement patterns and in the resulting population density. The Navrongo trial was conducted in an area with dispersed settlements consisting of small compounds, with a population density of approximately 104 people per km^2. This setting is similar to the AvecNet trial, conducted in an area with dispersed small villages, but with a population density of 51 people per km^2. The SolarMal trial, in contrast, was conducted on a small, densely populated island with a population density of 565 people per km^2. The percentage of households in core for the Navrongo trial was 82%, based on the distance to the nearest discordant compound. For the SolarMal, the percentage of households in core was 85% mid-rollout, bearing in mind that a hierarchical rollout strategy was employed, leading to a higher percentage of households in core than would be the case in a simple random assignment of clusters to arm. For the AvecNet trial, the available data only consists of the distance to the nearest discordant household with an enrolled child and the percentage of households in core is hence not comparable. Furthermore, the clusters were chosen as administrative units, and we assume that there is insufficient data from the cluster boundary regions to estimate the contamination. In chapter V, we conducted a redesign of this trial with an algorithmic approach to cluster allocation and showed that the clusters could have been smaller in terms of the households enrolled by not taking a subsample of all children to adhere to the administrative cluster boundaries.

The results for the Navrongo and the SolarMal trial are compared in table VI.1. In both trials, a sigmoid random effects analysis led to a comparable result in the respective estimate of effectiveness. For the SolarMal, there is an increase in effectiveness of 6%, probably because clusters were chosen much smaller in this trial. The contamination range was found to be around 200 m for the Navrongo trial and around 150 m for the SolarMal trial. This is much less than the maximum distances *Anopheles* mosquitoes fly [Guerra et al., 2014, Verdonschot and Besse-Lototskaya, 2014]. The difference between

	Navrongo trial	SolarMal trial
Original estimate:	17%, [0, 31]%	30.0%, [20.9, 38.0]%
Sigm RE estimate:	16.6%, [2.2, 30.7]%	31.9%, [15.3, 45.8]%
Difference:	−2.4%	6.0%
Contamination range:	0.198 km, [0.092, 1.088] km	0.146 km, [0.052, 0.923] km
Population density:	104 people per km^2	565 people per km^2
% in core:	82%	85% mid-rollout

Table VI.1: Comparison of the main outcomes reported for the Navrongo trial (reduction in all-cause mortality in children aged 6 months to 4 years) and the SolarMal trials (reduction in odds ratio of malaria prevalence in a random sample of 10% of all households). The primary outcome from the original analysis and from a sigmoid random effects model (Sigm RE) are listed and their difference is calculated.

the two estimates might be due to the difference of the interventions, or the differences in the population densities influencing the distance mosquitoes travel. These contamination ranges are a bit less than reported for a trial in Asembo, Kenya [Hawley et al., 2003], where significant protective effects of insecticide treated nets were found for distances of up to 300 m from cluster boundaries. The mean geographic distance between parent and offspring malaria infections estimated in Kilifi County, Kenya [Malinga et al., 2019b] was reported to be around 500 m. This does not directly relate to the contamination range but also supports our results that that the relevant contamination for malaria interventions arises within a small range from the intervention boundary and, if buffer zones are desired, they need not be wide.

5. Future work: maximize the information from the data for the model

In the previous sections of this chapter, we discussed the contributions of this thesis to the interplay between models and trials, visualized with a cycle in figure VI.1. In this section, we show how the circle can turn into a cycle. The theory of optimal experimental design can be used to maximize the information from the collected data in a CRT to fit a sigmoid random effects model. We emphasized the necessity to both have data from the cluster boundary regions where contamination is likely and the cluster core regions that are untouched by contamination to reliably estimate the true intervention effectiveness. This insight was gained from comparing the model performance in many simulations and by noting that the model results in unbiased and precise estimates of effectiveness if this is fulfilled. By invoking optimal experimental design theory, we might be able to precisely answer the question from what distance to the boundary the most data are needed to fit

the sigmoid random effects model with minimal variance in the estimated parameters in a mathematically optimal way.

The idea behind optimal experimental design is to design the best possible experiment such that model parameters can be estimated with the best possible statistical quality [Pronzato and Pázman, 2013, Kitsos, 2013]. We assume that we have a model that depends on p parameters $\theta = (\theta_1, ..., \theta_p)^T$ and an n-dimensional variable we can control, $\tau = (\tau_1, ..., \tau_n)^T$. We want to conduct an experiment to collect data that will be used to estimate the set of parameters θ. The aim is to choose the best configuration of τ to then estimate the parameters with low variance. Depending on the model and the goal, τ could represent the time points at which a drug concentration is measured, the number of modified males to be released at predefined time points or the best distance of N households to the discordant intervention boundary. To get the best configuration for τ, an optimization problem has to be solved. This is related to parameter estimation, that aims to determine parameter values for a model that give the best fit to experimental data, in the sense that both procedures minimize a related cost functional. A typical choice for the cost functional in parameter estimation is a maximum likelihood estimator. In optimal experimental design, the most intuitive choice for a cost functional involves the Fisher information matrix, the variance of the partial derivative in direction of the parameter of the log-likelihood function. By minimizing this type of cost functional we can determine the experimental setting that minimizes the variance, and hence maximizes the information, for the model under consideration [Pronzato and Pázman, 2013].

The theory of optimal experimental design has mainly found applications in pharmacodynamic modeling. It has been applied to CRTs in the context of assigning the best proportion of clusters to different intervention arms [Wu et al., 2017], but apart from this reference we are not aware of other connections between the fields of optimal experimental design and design of CRTs. This connection could benefit from the rich theory already developed in the last decades for optimal experimental design and might be able to substantially improve the design and analysis of CRTs of new interventions for mosquito-borne diseases.

6. Conclusions

In this thesis we developed the theoretical basis to allow for contamination in the design and analysis of trials of mosquito control interventions. We started this work by parametrizing a spatial model for the release of modified male mosquitoes with the goal of informing the trial design of such an intervention. By applying optimal control theory, we found an optimized strategy to release modified males on an island. In parallel we

developed a nonlinear model for the analysis of cluster randomized trials that can account for contamination arising from mosquito movement and leads to unbiased estimates of effectiveness. Major parts of this work were the development of simulations of cluster randomized trials with contamination to test the analysis in a simulation study and the application to a cluster randomized trial of insecticide treated nets. We then extended the developed sigmoid random effects model to analyze stepped wedge cluster randomized trials and to estimate the intervention effectiveness at any intervention coverage. This extension enabled us to analyze two stepped wedge cluster randomized trials of malaria interventions. We concluded this thesis with implications for the design of trials and connected the development on trial design and analysis with the results of the optimized release strategy of modified male mosquitoes.

Our results on the analysis of trials with contamination show how to account for and estimate mosquito dispersal in trials. The chosen sigmoid random effects model adjusts the estimate of effectiveness to be unbiased and the estimated contamination range is around 100 to 200 meters for malaria interventions, much less than the maximum distance *Anopheles* mosquitoes can fly. On the other hand, for such an analysis to add value it is crucial that data is collected from zones where contamination is likely when designing trials. Clusters need not be chosen as entire villages and the size can be reduced to the minimum determined by contamination effects or operational factors. Furthermore, buffer zones need not be excluded, or if a conventional approach is preferred, the buffer zone needs not be wide. This reduces the number of participants enrolled and hence the cost of the trial. Contamination effects are a valuable source of information on the effectiveness of an intervention. We hope that this work leads to a better understanding of the contamination arising from mosquito control interventions and to a paradigm shift in the way contamination is perceived by investigators.

Appendix

1. Other payoff functionals

Instead of the payoff functional (II.6.1), we can choose other functionals, depending on what property of $J[\xi]$ we lay the focus on. An intuitive choice would be

$$J[\xi] = \frac{1}{2}\int_0^T Cu(t)^2 + \xi(t)\ \mathrm{d}t,$$

considering a linear dependence of the control function ξ. Yet, with this choice, some analytical problems arise, since the function $Cu(t)^2 + \xi(t)$ is no longer convex. Another choice is

$$J[\xi] = \frac{1}{2}\int_0^T Ct^2u(t)^2 + \xi(t)^2\ \mathrm{d}t. \tag{1.1}$$

The factor t^2 leads to a higher penalty the more time passes and the algorithm has a motivation to minimize the population quickly.
When simulating the functional (1.1) with $C = 1/\mu^2$, we get a similar result as in figure II.4, though we get to elimination $\sim$ 7 days earlier, and the maximal number $M = 10\,000$ of sterile males is used for a longer time span (figure A.1).

2. Rescaling of the PDE model

Problem (II.3.3) can be rescaled such that it is independent of any unit. A detailed description of the steps taken, along with many examples, can be found in a recent book dealing with scaling of differential equations [Langtangen and Pedersen, 2016].
We rescale our problem and divide every variable and time dependent parameter by a convenient constant of the same units:

$$\bar{x} = \frac{x}{x_c},\ \bar{t} = \frac{t}{t_c},\ \bar{u} = \frac{u}{u_c},\ \bar{\psi} = \frac{\psi}{\psi_c},\ \bar{\xi} = \frac{\xi}{\xi_c},\ \bar{\kappa} = \frac{\kappa}{\kappa_c}.$$

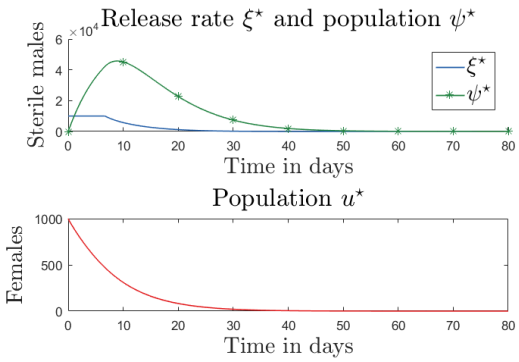


Figure A.1: Optimal control of the single-site model with a different functional. The graphs are built in the same way as explained in figure II.4.

We substitute the resulting variables and time dependent parameters back into the problem and use that

$$\frac{\partial u}{\partial t} = \frac{\partial (u_c \bar{u})}{\partial \bar{t}} \frac{\partial \bar{t}}{\partial t} = \frac{u_c}{t_c} \frac{\partial \bar{u}}{\partial \bar{t}},$$
$$\frac{\partial^2 u}{\partial x^2} = \frac{u_c}{x_c^2} \frac{\partial^2 \bar{u}}{\partial \bar{x}^2}.$$

From the resulting PDE problem we deduce that convenient constants to scale our problem are

$$x_c = \left(\frac{D}{\mu}\right)^{1/2},\ t_c = \frac{1}{\mu},\ k_c = \max_{0 \leqslant t \leqslant T} \kappa(t),\ u_c = \psi_c = k_c,\ \xi_c = \mu k_c.$$

This results in the following initial value problem:

$$\begin{aligned}
\frac{\partial \bar{u}}{\partial \bar{t}} &= \Delta \bar{u} + \bar{f}(\bar{u}, \bar{\psi}), && \text{for } \bar{x} \in \bar{\Omega},\ \bar{t} > 0,\\
\frac{\partial \bar{\psi}}{\partial \bar{t}} &= \Delta \bar{\psi} + \bar{g}(\bar{\psi}), && \text{for } \bar{x} \in \bar{\Omega},\ \bar{t} > 0,\\
\bar{u}(0, \bar{x}) &= \bar{u_0}(\bar{x}),\ \bar{\psi}(0, \bar{x}) = 0, && \text{for } \bar{x} \in \bar{\Omega},\\
&\text{Periodic boundary conditions}, && \text{for } \bar{t} > 0,
\end{aligned}$$

where $\bar{u_0}(\bar{x}) = \bar{\kappa}$, $\bar{\Omega} = [0, L/x_c]$ and $0 \leqslant \bar{t} \leqslant T/t_c$ with

$$\bar{f}(\bar{u}, \bar{\psi}) = \frac{\phi}{\mu}\left(1 - \frac{\bar{u}}{\bar{\kappa}}\right)\left(\frac{\alpha \kappa_c u}{\alpha \kappa_c u + 1}\right)\left(\frac{\bar{u}}{\bar{u} + \bar{\psi}}\right)\bar{u} - \bar{u},$$
$$\bar{g}(\bar{\psi}) = \bar{\xi} - \bar{\psi}.$$

Note that the diffusion D has disappeared from the spatial part Δ and now appears in the interval $\bar{\Omega}$.

3. Extension of the reanalysis for the Navrongo trial

The previous reanalysis [Jarvis et al., 2019] indicated that adjusting the main outcome of mortality for a contamination effect did not influence the results. At the same time, the confidence intervals around the contamination effect were confirmed to be narrow. It is hence assumed that clusters were chosen to be so large that even a contamination range of several hundred meters did not affect the main outcome, i.e. the percentage of households in core was very high. As shown above, for a sigmoid random effects analysis ($\mathcal{S}_{RE}$) to result in precise and accurate estimates of effectiveness, only around 50% of households need be in core. The trial cannot be redone with smaller clusters (and hence $\approx$ 50% of households in core), but simulations can show whether the estimate of effectiveness remains stable for a trial with smaller clusters. For each cluster, a subset of households far away from the discordant trial arm can be randomly excluded. This reduces the number of households per cluster without violating the cluster boundaries. Because only households far away from the discordant trial arm are chosen for exclusion, the percentage of households, ω, decreases. This could be seen as the opposite of a fried egg design, since in each cluster, households close to a discordant household are kept. It is hypothesized that a GEE analysis is more biased the more households in core are excluded and a sigmoid model analysis remains unaffected because the information of the contamination range remains the same, although confidence intervals will probably get wider.

If only households in core were eligible for exclusion, the resulting cluster sizes would be imbalanced, since the number of households in core in each cluster varies significantly. Hence, households lying further away from the nearest discordant household than the 20%-quantile were eligible for exclusion. Of these 80% of the households in each cluster, a percentage q was randomly selected and $0.8q$ households were randomly excluded. In total, 50 values for q were chosen and for each of those values, 50 replicate data sets were generated. The estimates were bootstrap corrected for 100 resamples, for the JAGS model the number of iterations was set to 2000 with a burn-in period of 500.

For a contamination range of 0.198 km, 82% of the households were in core and hence unaffected by the estimated contamination range. If only 50% of the households had been in core, the results for the sigmoid models remain unaffected, but with slightly wider confidence intervals. The result is displayed in figure A.2. For the GEE model, the estimated effectiveness decreases the more households are excluded (and hence the fewer households are in core) as expected. The estimated effectiveness for the sigmoid models $\mathcal{S}_{RE}$ and $\mathcal{S}$ remains constant, even when only $\approx$ 50% of households lay in core, corresponding to an exclusion of 64% of all households in the trial. The width of the 95%CI around the estimated effectiveness increases linearly for all three models, because

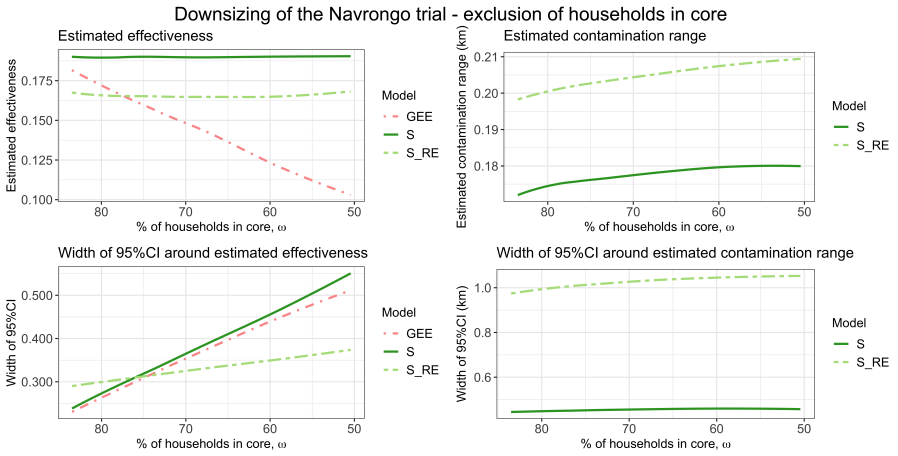


Figure A.2: Reanalysis of the Navrongo trial, where for each cluster, households far away from the intervention boundary were randomly excluded to vary the percentage of households in core ω. The estimated parameters (upper plots) and width of the 95%CI (lower plots) for a GEE, $\mathcal{S}$ and $\mathcal{S}_{RE}$ analysis in terms of the percentage of households in core are visualized.

fewer households (and hence smaller clusters) leads to a loss of power. This increase is slower for $\mathcal{S}_{RE}$. The estimated contamination range remains in the magnitude of 200 m, a slight increase for both models is noted as ω decreases. The width of the 95%CI for the estimated contamination range is constant for different ω (but quite large for $\mathcal{S}_{RE}$, around 1 km).

The conclusion that the Navrongo trial could have been much smaller without the results being affected hence holds. With only 36% of the original households included, the same parameter estimations are attained with a sigmoid random effects model, although the width of credible intervals increases. This underlines the findings from the simulation study that clusters could be much smaller in terms of number of households included and information from households close to the boundary should not be discarded.

4. Estimation of effectiveness as a function of intervention coverage

Effective intervention coverage

Assume that the N households in the study are ordered such that the first N_1 receive the intervention. This means that (x_m, y_m), $m = 1, \ldots, N_1$, denote all the coordinates of the N_1 intervened households and (x_m, y_m), $m = N_1 + 1, \ldots, N$, the coordinates of the control

households. For any household with coordinates (x, y) the effective intervention coverage $\mathcal{R}(x, y)$ is then defined as:

$$\mathcal{R}(x,y) = \frac{\sum_{m=1}^{N_1} K(x - x_m, y - y_m)}{\sum_{n=1}^{N} K(x - x_n, y - y_n)},$$

where $K(x, y)$ is an appropriate function.
If $K(x, y)$ is chosen as 1 whenever $x^2+y^2 \leqslant \hat{\theta}$, the expression $\mathcal{R}(x, y)$ compares the number of intervened households to the total number of households within the contamination range $\hat{\theta}$ and hence describes the percentage of households within an estimated contamination range $\hat{\theta}$ that receive the intervention, as it has been previously defined [Hawley et al., 2003, Silkey et al., 2016]. This discrete measurement is imprecise, since it does not consider the closeness between households within the contamination range. Instead, let $K(x, y)$ be a radially symmetric probability density function, for instance a two-dimensional Gaussian kernel. The expression $\mathcal{R}(x, y)$ then represents the common density of the intervened households relative to the general density of households and can be interpreted as a spatial relative risk function, as used in kernel density estimation [Hazelton, 2016, Kelsall and Diggle, 1995, Waller, 2010].

Choice of bandwidth

The bandwidth ε of the two-dimensional Gaussian kernel

$$K(x,y) = \frac{1}{2\pi\varepsilon^2} \exp\Big(-\frac{x^2+y^2}{2\varepsilon^2}\Big)$$

is chosen such that 95% of its distribution lies within a radius of the contamination range $\hat{\theta}$ around each household. Taking

$$\varepsilon = \frac{\pi}{\sqrt{6}\beta_3} (= 0.44\hat{\theta})$$

results in a bivariate normal distribution with 95% of the distribution laying within the contamination range $\hat{\theta}$. This is because the contamination range $\hat{\theta}$ was calculated from β_3, the growth rate of a sigmoid function. This sigmoid function is the cumulative density function of the one-dimensional logistic distribution with variance $\pi^2/(3\beta_3^2)$. On the other side, a covariance matrix of a two-dimensional Gaussian kernel gives rise to a variance in one direction quantified by $2\varepsilon^2$ [Multerer et al., 2021]. The effective intervention coverage of the ith household in the jth cluster in the kth survey round with coordinates (x_{ijk}, y_{ijk}) can then be calculated as $\mathcal{R}_{ijk} = \mathcal{R}(x_{ijk}, y_{ijk})$ and is defined analogously for CRTs, with less indexing.

Approximation of the effective intervention coverage

The effective intervention coverage $\mathcal{R}_{ijk}$ can be approximated by

$$\hat{\mathcal{R}}_{ijk} = \frac{1}{1 + \exp(-\beta_3 \Delta_{ijk})},$$

with the same β_3 as was fitted in the Bayesian hierarchical model

$$\begin{aligned} Y_{ijk} &\sim \text{Binomial}(p_{ijk}), \\ \text{logit}(p_{ijk}) &= \beta_{1,ik} + \frac{\beta_2}{1 + \exp(-\beta_3 \Delta_{ijk})}, \\ \beta_{1,ik} &\sim \text{Normal}(\mu, \tau). \end{aligned} \tag{4.1}$$

This relationship between the nearest discordant household Δ_{ijk} and $\mathcal{R}_{ijk}$ is illustrated in figure A.3 for the SolarMal trial. The percentage error

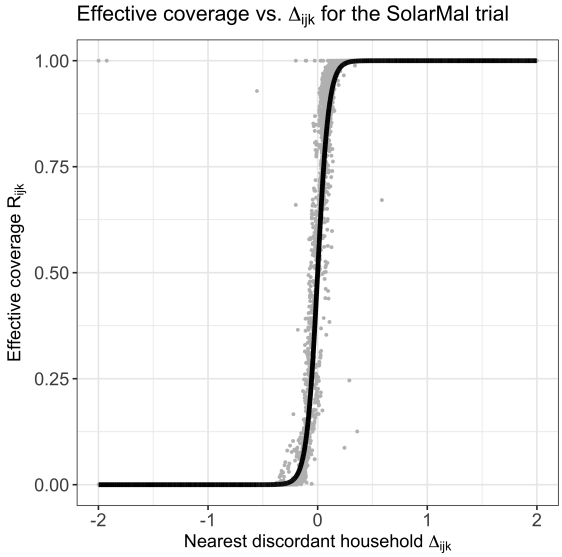


Figure A.3: Illustration of the sigmoid relationship between the nearest discordant household and the effective coverage for the SolarMal trial in grey. In black, the sigmoid curve for $\hat{\mathcal{R}}_{ijk}$ that approximates the effective coverage is illustrated. To increase the readability, only households within 2 km of the nearest discordant household are plotted on the x-axis.

$$\text{perc. err.} = 100\% \times \frac{\|R_{ijk} - \hat{R}_{ijk}\|_2}{\|R_{ijk}\|_2},$$

between the effective intervention coverage $\mathcal{R}_{ijk}$ and its approximation $\hat{\mathcal{R}}_{ijk}$ was estimated to be 6% for the SolarMal trial and 3% for the AvecNet trial. These small errors justify this approximation.

Estimation of the intervention effectiveness at the effective coverage

Plugging in the inverse relationship between Δ_{ijk} and $\mathcal{R}_{ijk} \approx \hat{\mathcal{R}}_{ijk}$ in equation (4.1) it holds that

$$\text{logit}(p_{ijk}) \approx \beta_1 + \beta_2 \mathcal{R}_{ijk}$$

and the trial outcome can be described as a function of the effective intervention coverage.

BIBLIOGRAPHY

[Alphey et al., 2010] Alphey, L., Benedict, M., Bellini, R., Clark, G., Dame, D., Service, M., and Dobson, S. (2010). Sterile-insect methods for control of mosquito-borne diseases: an analysis. *Vector-Borne and Zoonotic Diseases*, 10(3):295–311.

[Anaya-Izquierdo and Alexander, 2020] Anaya-Izquierdo, K. and Alexander, N. (2020). Spatial regression and spillover effects in cluster randomized trials with count outcomes. *Biometrics*, pages 1–16.

[Baddeley and Turner, 2005] Baddeley, A. and Turner, R. (2005). spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software*, 12(6):1–42.

[Baird et al., 2018] Baird, S., Bohren, J., McIntosh, C., and Özler, B. (2018). Optimal design of experiments in the presence of interference. *The Review of Economics and Statistics*, 100(5):844–860.

[Barclay and Mackauer, 1980] Barclay, H. and Mackauer, M. (1980). The sterile insect release method for pest control: a density dependent model. *Environmental Entomology*, 9:810–817.

[Bates et al., 2015] Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

[Benjamin-Chung et al., 2017] Benjamin-Chung, J., Abedin, J., Berger, D., Clark, A., Jimenez, V., Konagaya, E., Tran, D., Arnold, B., Hubbard, A., Luby, S., Miguel, E., and Colford, J. (2017). Spillover effects on health outcomes in low- and middle-income countries: a systematic review. *International Journal of Epidemiology*, 46(4):1251–1276.

[Bhatt et al., 2013] Bhatt, S., Gething, P., Brady, O., Messina, J., Farlow, A., Moyes, C., Drake, J., Brownstein, J., Hoen, A., Sankoh, O., Myers, M., George, D., Jaenisch, T., Wint, G., Simmons, C., Scott, T., Farrar, J., and Hay, S. (2013). The global distribution and burden of dengue. *Nature*, 496(7446):504 – 507.

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

[Binka et al., 1996] Binka, F., Kubaje, A., Adjuik, M., Williams, L., Lengeler, C., Maude, G., Armah, G., Kajihara, B., Adiamah, J., and Smith, P. (1996). Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Tropical Medicine & International Health*, 1(2):147–154.

[Boukal and Berec, 2002] Boukal, D. and Berec, L. (2002). Single-species models of the allee effect: extinction boundaries, sex ratios and mate encounters. *Journal of Theoretical Biology*, 218(3):375–394.

[Bouyer et al., 2020] Bouyer, J., Yamada, H., Pereira, R., Bourtzis, K., and Vreysen, M. (2020). Phased conditional approach for mosquito management using sterile insect technique. *Trends in Parasitology*, 36(4):325 – 336.

[Bradley et al., 2000] Bradley, P., Bennett, K., and Demiriz, A. (2000). Constrained k-means clustering. Technical report, MSR-TR-2000-65, Microsoft Research.

[Burton et al., 2006] Burton, A., Altman, D., Royston, P., and Holder, R. (2006). The design of simulation studies in medical statistics. *Statistics in Medicine*, 25:4279–4292.

[Cai et al., 2014] Cai, L., Ai, S., and Li, J. (2014). Dynamics of mosquitoes populations with different strategies for releasing sterile mosquitoes. *SIAM Journal on Applied Mathematics*, 74(6):1786–1809.

[Carvalho et al., 2015] Carvalho, D., McKemey, A., Garziera, L., Lacroix, R., Donnelly, C., Alphey, L., Malavasi, A., and Capurro, M. (2015). Suppression of a field population of *Aedes aegypti* in Brazil by sustained release of transgenic male mosquitoes. *PLoS Neglected Tropical Diseases*, 9(7):1–15.

[Davison and Hinkley, 1997] Davison, A. and Hinkley, D. (1997). *Bootstrap Methods and Their Applications*. Cambridge University Press.

[Delrieu et al., 2015] Delrieu, I., Leboulleux, D., Ivinson, K., Gessner, B., Chandramohan, D., Churcher, T., Drakeley, C., Halloran, E., Killeen, G., Kleinschmidt, I., Milligan, P., Robert, V., Rogier, C., Saul, A., Sinden, R., and Smith, T. (2015). Design of a phase III cluster randomized trial to assess the efficacy and safety of a malaria transmission blocking vaccine. *Vaccine*, 33(13):1518–1526.

[Diggle et al., 2002] Diggle, P., Heagerty, P., Liang, K., and Zeger, S. (2002). *Analysis of Longitudinal Data*. Oxford Statistical Science . Oxford University Press, 2 edition.

[Ding et al., 2010] Ding, W., Finotti, H., Lenhart, S., Lou, Y., and Ye, Q. (2010). Optimal control of growth coefficient on a steady-state population model. *Nonlinear Analysis: Real World Applications*, 11(2):688 – 704.

[Donner et al., 1982] Donner, A., Birkett, N., and Buck, C. (1982). Randomization by cluster: sample size requirements and analysis. *American Journal of Epidemiology*, 114:906–914.

[Dufourd and Dumont, 2013] Dufourd, C. and Dumont, Y. (2013). Impact of environmental factors on mosquito dispersal in the prospect of sterile insect technique control. *Computers and Mathematics with Applications*, 66(9):1695 – 1715.

[Dyck et al., 2005] Dyck, V., Hendrichs, J., and Robinson, A. (2005). *Sterile insect technique: principles and practice in area-wide integrated pest management*. Springer-Verlag, Dodrecht, The Netherlands.

[Dye, 1984] Dye, C. (1984). Models for the population dynamics of the yellow fever mosquito, *Aedes aegypti*. *Journal of Animal Ecology*, 53(1):247–268.

[Efron and Tibshirani, 1994] Efron, B. and Tibshirani, R. (1994). *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability . Chapman & Hall/CRC, Taylor & Francis Group.

[Eisele et al., 2016] Eisele, T., Bennett, A., Silumbe, K., Finn, T., Chalwe, V., Kamuliwo, M., Hamainza, B., Moonga, H., Kooma, E., Chizema Kawesha, E., Yukich, J., Keating, J., Porter, T., Conner, R., Earle, D., Steketee, R., and Miller, J. (2016). Short-term impact of mass drug administration with dihydroartemisinin plus piperaquine on malaria in southern province Zambia: a cluster-randomized controlled trial. *The Journal of Infectious Diseases*, 214(12):1831–1839.

[Esteva and Yang, 2005] Esteva, L. and Yang, H. (2005). Mathematical model to assess the control of *Aedes aegypti* mosquitoes by the sterile insect technique. *Mathematical Biosciences*, 198(2):132–147.

[Evans, 1983] Evans, L. (1983). Script: an introduction to mathematical optimal control theory, version 2.0. https://math.berkeley.edu/~evans/control.course.pdf. Accessed: 3.11.2017.

[Ferreira et al., 2008] Ferreira, C., Yang, H., and Esteva, L. (2008). Assessing the suitability of sterile insect technique applied to *Aedes aegypti*. *Journal of Biological Systems*, 16(4):565 – 577.

[Fister et al., 2013] Fister, K., McCarthy, M., Oppenheimer, S., and Collins, C. (2013). Optimal control of insects through sterile insect release and habitat modification. *Mathematical Biosciences*, 244(2):201–212.

[Flores and O'Neill, 2018] Flores, H. and O'Neill, S. (2018). Controlling vector-borne diseases by releasing modified mosquitoes. *Nature Reviews Microbiology*, 16(8):508 – 518.

[Guerra et al., 2014] Guerra, C., Reiner, R., Perkins, T., Lindsay, S., Midega, J., Brady, O., Barker, C., Reisen, W., Harrington, L., Takken, W., Kitron, U., Lloyd, A., Hay, S., Scott, T., and Smith, D. (2014). A global assembly of adult female mosquito mark-release-recapture data to inform the control of mosquito-borne pathogens. *Parasites & Vectors*, 7(1):1–15.

[Hahsler and Hornik, 2007] Hahsler, M. and Hornik, K. (2007). TSP – Infrastructure for the traveling salesperson problem. *Journal of Statistical Software*, 23(2):1–21.

[Halloran et al., 2017] Halloran, M., Auranen, K., Baird, S., Basta, N., Bellan, S., Brookmeyer, R., Cooper, B., DeGruttola, V., Hughes, J., Lessler, J., Lofgren, E., Longini, I., Onnela, J., Özler, B., Seage, G., Smith, T., Vespignani, A., Vynnycky, E., and Lipsitch, M. (2017). Simulations for designing and interpreting intervention trials in infectious diseases. *BMC Medicine*, 15(1):223–231.

[Halloran et al., 2010] Halloran, M., Longini Jr., I., and Struchiner, C. (2010). *Design and analysis of vaccine studies*. Statistics for Biology and Health. Springer-Verlag New York, 1 edition.

[Halloran and Struchiner, 1991] Halloran, M. and Struchiner, C. (1991). Study designs for dependent happenings. *Epidemiology*, 2(5):331–338.

[Harris et al., 2011] Harris, A., Nimmo, D., McKemey, A., Kelly, N., Scaife, S., Donnelly, C., Beech, C., Petrie, W., and Alphey, L. (2011). Field performance of engineered male mosquitoes. *Nature Biotechnology*, 29(11):1034–1037.

[Hawley et al., 2003] Hawley, W., Phillips-Howard, P., ter Kuile, F., Terlouw, D., Vulule, J., Ombok, M., Nahlen, B., Gimnig, J., Kariuki, S., Kolczak, M., and Hightower, A. (2003). Community-wide effects of Permethrin-treated bed nets on child mortality and malaria morbidiy in western Kenya. *American Journal of Tropical Medicine and Hygiene*, 68(4):121–127.

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

[Hayes and Bennett, 1999] Hayes, R. and Bennett, S. (1999). Simple sample size calculation for cluster-randomized trials. *International Journal of Epidemiology*, 28(2):319–326.

[Hayes and Moulton, 2009] Hayes, R. and Moulton, L. (2009). *Cluster Randomised Trials*. Biostatistics Series. Chapman & Hall/CRC, Taylor & Francis Group.

[Hazelton, 2016] Hazelton, M. (2016). Kernel smoothing methods. In Lawson, A., Banerjee, S., Haining, R., and Ugarte, M., editors, *Handbook of spatial epidemiology*, Handbooks of Modern Statistical Methods, chapter 10, pages 195–207. Chapman & Hall/CRC, Taylor & Francis Group.

[Hiscox et al., 2016] Hiscox, A., Homan, T., Mweresa, C., Maire, N., Di Pasquale, A., Masiga, D., Oria, P., Alaii, J., Leeuwis, C., Mukabana, W., Takken, W., and Smith, T. (2016). Mass mosquito trapping for malaria control in western Kenya: study protocol for a stepped wedge cluster-randomised trial. *Trials*, 17(1):356–368.

[Hiscox et al., 2014] Hiscox, A., Otieno, B., Kibet, A., Mweresa, C., Omusula, P., Geier, M., Rose, A., Mukabana, W., and Takken, W. (2014). Development and optimization of the Suna trap as a tool for mosquito monitoring and control. *Malaria Journal*, 13(1):257–271.

[Højsgaard et al., 2006] Højsgaard, S., Halekoh, U., and Yan, J. (2006). The R package geepack for generalized estimating equations. *Journal of Statistical Software*, 15(2):1–11.

[Holling, 1959] Holling, C. (1959). The components of predation as revealed by a study of small-mammal predation of the European pine sawfly. *Mathematical Biosciences*, 91(5):293–320.

[Homan et al., 2016] Homan, T., Hiscox, A., Mweresa, C., Masiga, D., Mukabana, W., Oria, P., Maire, N., Di Pasquale, A., Silkey, M., Alaii, J., Bousema, T., Leeuwis, C., Smith, T., and Takken, W. (2016). The effect of mass mosquito trapping on malaria transmission and disease burden (SolarMal): a stepped-wedge cluster-randomised trial. *The Lancet*, 388(10050):1193–1201.

[Howard et al., 2000] Howard, S., Omumbo, J., Nevill, C., Some, E., Donnelly, C., and Snow, R. (2000). 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*, 94(4):357–360.

[Hussey and Hughes, 2007] Hussey, M. and Hughes, J. (2007). Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials*, 28(2):182–191.

[Jarvis et al., 2017] Jarvis, C., Di Tanna, G., Lewis, D., Alexander, N., and Edmunds, W. (2017). Spatial analysis of cluster randomised trials: a systematic review of analysis methods. *Emerging Themes in Epidemiology*, 14(1):1–12.

[Jarvis et al., 2019] Jarvis, C., Multerer, L., Lewis, D., Binka, F., Edmunds, W., Alexander, N., and Smith, T. (2019). Spatial effects of permethrin-impregnated bed nets on child mortality: 26 years on, a spatial reanalysis of a cluster randomized trial. *American Journal of Tropical Medicine and Hygiene*, 101:1–8.

[Kelsall and Diggle, 1995] Kelsall, J. and Diggle, P. (1995). Kernel estimation of relative risk. *Bernoulli*, 1:3–16.

[Kim et al., 2017] Kim, J., Lee, H., Lee, C., and Lee, S. (2017). Assessment of optimal strategies in a two-patch dengue transmission model with seasonality. *PLoS ONE*, 12(3):1 – 21.

[Kish, 1965] Kish, L. (1965). *Survey sampling*. John Wiley & Sons.

[Kitsos, 2013] Kitsos, C. (2013). *Optimal Experimental Design for Non-Linear Models*. SpringerBriefs in Statistics . Springer-Verlag Berlin Heidelberg.

[Knipling, 1955] Knipling, E. (1955). Possibilities of insect control or eradication through the use of sexually sterile males. *Journal of Economic Entomology*, 48(4):459–462.

[Kotz et al., 2013] Kotz, D., Spigt, M., Arts, I., Crutzen, R., and Viechtbauer, W. (2013). The stepped wedge design does not inherently have more power than a cluster randomized controlled trial. *Journal of Clinical Epidemiology*, 66(9):1059–1060.

[Langtangen and Pedersen, 2016] Langtangen, H. and Pedersen, G. (2016). *Scaling of differential equations*, volume 2 of *Simula SpringerBriefs on Computing*. Springer International Publishing.

[Lenhart and Workman, 2007] Lenhart, S. and Workman, J. (2007). *Optimal control applied to biological models*. Mathematical and Computational Biology Series. Chapman & Hall/CRC, Taylor & Francis Group.

[Li, 2017] Li, J. (2017). New revised simple models for interactive wild and sterile mosquito populations and their dynamics. *Journal of Biological Dynamics*, 11:316–333.

[Li and Yuan, 2015] Li, J. and Yuan, Z. (2015). Modelling releases of sterile mosquitoes with different strategies. *Journal of Biological Dynamics*, 9(1):1–14.

[Lutambi et al., 2013] Lutambi, A., Penny, M., Smith, T., and Chitnis, N. (2013). Mathematical modelling of mosquito dispersal in a heterogeneous environment. *Mathematical Biosciences*, 241(2):198–216.

[Macdonald, 1957] Macdonald, G. (1957). *The epidemiology and control of malaria*. Oxford University Press, London.

[Maia et al., 2013] Maia, M., Onyango, S., Thele, M., Simfukwe, E., Turner, E., and Moore, S. (2013). Do topical repellents divert mosquitoes within a community? – health equity implications of topical repellents as a mosquito bite prevention tool. *PLoS ONE*, 8(12):1–7.

[Maidana and Yang, 2008] Maidana, N. and Yang, H. (2008). Describing the geographic spread of dengue disease by traveling waves. *Mathematical Biosciences*, 215(1):64 – 77.

[Malinga et al., 2019a] Malinga, J., Maia, M., Moore, S., and Ross, A. (2019a). Can trials of spatial repellents be used to estimate mosquito movement? *Parasites & Vectors*, 12(421):1–12.

[Malinga et al., 2019b] Malinga, J., Mogeni, P., Omedo, I., Rockett, K., Hubbart, C., Jeffreys, A., Williams, T., Kwiatkowski, D., Bejon, P., and Ross, A. (2019b). Investigating the drivers of the spatio-temporal patterns of genetic differences between plasmodium falciparum malaria infections in Kilifi County, Kenya. *Scientific Reports*, 9(1):1–13.

[McCann et al., 2018] McCann, R., van den Berg, H., Takken, W., Chetwynd, A., Giorgi, E., Terlouw, D., and Diggle, P. (2018). Reducing contamination risk in cluster-randomized infectious disease-intervention trials. *International Journal of Epidemiology*, 47(6):2015–2024.

[Mdege et al., 2011] Mdege, N., Man, M., Taylor (nee Brown), C., and Torgerson, D. (2011). Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *Journal of Clinical Epidemiology*, 64(9):936–948.

[Mdege et al., 2012] Mdege, N., Man, M., Taylor (nee Brown), C., and Torgerson, D. (2012). There are some circumstances where the stepped-wedge cluster randomized trial is preferable to the alternative: no randomized trial at all. Response to the commentary by Kotz and colleagues. *Journal of Clinical Epidemiology*, 65(12):1253–1254.

[Monlong, 2019] Monlong, J. (2019). Clustering into same size clusters. http://jmonlong.github.io/Hippocamplus/2018/06/09/cluster-same-size/#methods.

[Morris et al., 2019] Morris, T., White, I., and Crowther, M. (2019). Using simulation studies to evaluate statistical methods. *Statistics in medicine*, 38:2074–2102.

[Multerer et al., 2021] Multerer, L., Glass, T., Vanobberghen, F., and Smith, T. (2021). Analysis of contamination in cluster randomized trials of malaria interventions. *Trials*, 22(1):1–17.

[Oléron Evans and Bishop, 2014] Oléron Evans, T. and Bishop, S. (2014). A spatial model with pulsed releases to compare strategies for the sterile insect technique applied to the mosquito *Aedes aegypti*. *Mathematical Biosciences*, 254:6–27.

[Pinheiro and Bates, 2000] Pinheiro, J. and Bates, D. (2000). *Mixed-effects models in S and S-PLUS*. Statistics and Computing . Springer-Verlag New York.

[Pinheiro et al., 2018] Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2018). *nlme: linear and nonlinear mixed effects models*. R Foundation for Statistical Computing, Vienna, Austria. R package version 3.1-137.

[Plummer, 2019] Plummer, M. (2019). *rjags: Bayesian Graphical Models using MCMC*. R Foundation for Statistical Computing, Vienna, Austria. R package version 4-10.

[Pronzato and Pázman, 2013] Pronzato, L. and Pázman, A. (2013). *Design of Experiments in Nonlinear Models*. Lecture Notes in Statistics . Springer-Verlag New York.

[Protopopoff et al., 2018] Protopopoff, N., Mosha, J., Lukole, E., Charlwood, J., Wright, A., Mwalimu, C., Manjurano, A., Mosha, F., Kisinza, W., Kleinschmidt, I., and Rowland, M. (2018). Effectiveness of a long-lasting piperonyl butoxide-treated insecticidal net and indoor residual spray interventions, separately and together, against malaria transmitted by pyrethroid-resistant mosquitoes: a cluster, randomised controlled, two-by-two factorial design trial. *The Lancet*, 391:1577–1588.

[Protopopoff et al., 2015] Protopopoff, N., Wright, A., West, P., Tigererwa, R., Mosha, F., Kisinza, W., Kleinschmidt, I., and Rowland, M. (2015). Combination of insecticide treated nets and indoor residual spraying in northern Tanzania provides additional reduction in vector population density and malaria transmission rates compared to insecticide treated nets alone: a randomised control trial. *PLoS ONE*, 10(11):1–11.

[Quiroga et al., 2015] Quiroga, A., Fernández, D., Torres, G., and Turner, C. (2015). Adjoint method for a tumor invasion PDE-constrained optimization problem in 2D using adaptive finite element method. *Applied Mathematics and Computation*, 270:358 – 368.

[R Core Team, 2017] R Core Team (2017). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

[R Core Team, 2018] R Core Team (2018). *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[Rafikov et al., 2015] Rafikov, M., Rafikova, E., and Yang, H. (2015). Optimization of the *Aedes aegypti* control strategies for integrated vector management. *Journal of Applied Mathematics*, 9(7):1–8.

[Reich et al., 2012] Reich, N., Myers, J., Obeng, D., Milstone, A., and Perl, T. (2012). Empirical power and sample size calculations for cluster-randomized and cluster-randomized crossover studies. *PLoS ONE*, 7(4):1–7.

[Reiner et al., 2013] Reiner, R., Perkins, T., Barker, C., Niu, T., Chaves, L., Ellis, A., George, D., Le Menach, A., Pulliam, J., Bisanzio, D., Buckee, C., Chiyaka, C., Cummings, D., Garcia, A., Gatton, M., Gething, P., Hartley, D., Johnston, G., Klein, E., Michael, E., Lindsay, S., Lloyd, A., Pigott, D., Reisen, W., Ruktanonchai, N., Singh, B., Tatem, A., Kitron, U., Hay, S., Scott, T., and Smith, D. (2013). A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970-2010. *Journal of the Royal Society, Interface*, 10(81):1 – 13.

[Ross, 1911] Ross, R. (1911). *The prevention of malaria.* John Murray, London, 2 edition.

[RTS,S Clinical Trials Partnership, 2015] RTS,S Clinical Trials Partnership (2015). Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in africa: final results of a phase 3, individually randomised, controlled trial. *The Lancet*, 386(9988):31–45.

[Scrucca, 2013] Scrucca, L. (2013). GA: a package for genetic algorithms in R. *Journal of Statistical Software*, 53(4):1–37.

[Seirin Lee et al., 2013a] Seirin Lee, S., Baker, R., Gaffney, E., and White, S. (2013a). Modelling *Aedes aegypti* mosquito control via transgenic and sterile insect techniques: endemics and emerging outbreaks. *Journal of Theoretical Biology*, 331:78–90.

[Seirin Lee et al., 2013b] Seirin Lee, S., Baker, R., Gaffney, E., and White, S. (2013b). Optimal barrier zones for stopping the invasion of *Aedes aegypti* mosquitoes via transgenic or sterile insect techniques. *Theoretical Ecology*, 6(4):427–442.

[Service, 1997] Service, M. (1997). Mosquito (diptera: Culicidae) dispersal–the long and short of it. *Journal of Medical Entomology*, 34(6):579–588.

[Silkey et al., 2016] Silkey, M., Homan, T., Maire, N., Hiscox, A., Mukabana, R., Takken, W., and Smith, T. (2016). Design of trials for interrupting the transmission of endemic pathogens. *Trials*, 17(1):278–294.

[Smith et al., 2012] Smith, D., Battle, K., Hay, S., Barker, C., Scott, T., and McKenzie, F. (2012). Ross, macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLOS Pathogens*, 4(12):1–13.

[Smith et al., 2014] Smith, D., Perkins, T., Reiner, R., Barker, C., Niu, T., Chaves, L., Ellis, A., George, D., Le Menach, A., Pulliam, J., Bisanzio, D., Buckee, C., Chiyaka, C., Cummings, D., Garcia, A., Gatton, M., Gething, P., Hartley, D., Johnston, G., Klein, E., Michael, E., Lloyd, A., Pigott, D., Reisen, W., Ruktanonchai, N., Singh, B., Stoller, J., Tatem, A., Kitron, U., Godfray, H., Cohen, J., Hay, S., and Scott, T. (2014). Recasting the theory of mosquito-borne pathogen transmission dynamics and control. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 108(4):185 – 197.

[Staples et al., 2015] Staples, P., Ogburn, E., and Onnella, J. (2015). Incorporating contact network structure in cluster randomized trials. *Scientific Reports*, 5:1–12.

[Thomas, 1949] Thomas, M. (1949). A generalization of Poisson's binomial limit for use in ecology. *Biometrika*, 36(1/2):18–25.

[Thomé et al., 2010] Thomé, R., Yang, H., and Esteva, L. (2010). Optimal control of *Aedes aegypti* mosquitoes by the sterile insect technique and insecticide. *Mathematical Biosciences*, 223(1):12–23.

[Tiono et al., 2018] Tiono, A., Ouédraogo, A., Ouattara, D., Bougouma, E., Coulibaly, S., Diarra, A., Faragher, B., Guelbeogo, M., Grisales, N., Ouédraogo, I., Ouédraogo, Z., Pinder, M., Sanon, S., Smith, T., Vanobberghen, F., Sagnon, N., Ranson, H., and Lindsay, S. (2018). Efficacy of Olyset Duo, a bednet containing pyriproxyfen and permethrin, versus a permethrin-only net against clinical malaria in an area with highly pyrethroid-resistant vectors in rural Burkina Faso: a cluster-randomised controlled trial. *The Lancet*, 392(10147):569–580.

[Tiono et al., 2015] Tiono, A., Pinder, M., N'Fale, S., Faragher, B., Smith, T., Silkey, M., Ranson, H., and Lindsay, S. (2015). The AvecNet trial to assess whether addition of Pyriproxyfen, an insect juvenile hormone mimic, to long-lasting insecticidal mosquito nets provides additional protection against clinical malaria over current best practice in an area with pyrethroid-resistant vectors in rural Burkina Faso: study protocol for a randomised controlled trial. *Trials*, 16(1):113–125.

[Tröltzsch, 2010] Tröltzsch, F. (2010). *Optimal control of partial differential equations: theory, methods and applications*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island.

[Verdonschot and Besse-Lototskaya, 2014] Verdonschot, P. and Besse-Lototskaya, A. (2014). Flight distance of mosquitoes (Culicidae): a metadata analysis to support the management of barrier zones around rewetted and newly constructed wetlands. *Limnologica*, 45:69–79.

[Waller, 2010] Waller, L. (2010). Point process models and methods in spatial epidemiology. In Gelfand, A., Diggle, P., Fuentes, M., and Guttorp, P., editors, *Handbook of spatial statistics*, Handbooks of Modern Statistical Methods, chapter 22, pages 403–423. Chapman & Hall/CRC, Taylor & Francis Group.

[WHO, 2009] WHO (2009). *Dengue guidelines for diagnosis, treatment, prevention and control: new edition.* World Health Organization, Geneva. WHO/HTM/NTD/DEN/2009.1.

[WHO, 2017a] WHO (2017a). *Global vector control response 2017-2030.* World Health Organization, Geneva. Licence: CC BY-NC-SA 3.0 IGO.

[WHO, 2017b] WHO (2017b). *How to design vector control efficacy trials, guidance on phase III vector control field trial design.* World Health Organization, Geneva. Licence: CC BY-NC-SA 3.0 IGO.

[WHO, 2018] WHO (2018). *Global report on insecticide resistance in malaria vectors: 2010-2016.* World Health Organization, Geneva. Licence: CC BY-NC-SA 3.0 IGO.

[WHO, 2019] WHO (2019). *World malaria report 2019.* World Health Organization, Geneva. Licence: CC BY-NC-SA 3.0 IGO.

[WHO and IAEA, 2020] WHO and IAEA (2020). *Guidance framework for testing the sterile insect technique as a vector control tool against Aedes-borne diseases.* World Health Organization and the International Atomic Energy Agency, Geneva. Licence: CC BY-NC SA 3.0 IGO.

[Wilson et al., 2015] Wilson, A., Boelaert, M., Kleinschmidt, I., Pinder, M., Scott, T., Tusting, L., and Lindsay, S. (2015). Evidence-based vector control? improving the quality of vector control trials. *Trends in Parasitology*, 31(8):380 – 390.

[Wolbers et al., 2012] Wolbers, M., Kleinschmidt, I., Simmons, C., and Donnelly, C. (2012). Considerations in the design of clinical trials to test novel entomological approaches to dengue control. *PLoS Neglected Tropical Diseases*, 6(11):1–3.

[Wu et al., 2017] Wu, S., Wong, W., and Crespi, C. (2017). Maximin optimal designs for cluster randomized trials. *Biometrics*, 73(3):916–926.

[Yang et al., 2016] Yang, H., Boldrini, J., Fassoni, A., Freitas, L., Gomez, M., Lima, K. d., Andrade, V., and Freitas, A. (2016). Fitting the incidence data from the city of Campinas, Brazil, based on dengue transmission modellings considering time-dependent entomological parameters. *PLoS ONE*, 11(3):1–41.

[Zeger and Liang, 1986] Zeger, S. and Liang, K. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 42(1):121–130.

[Zheng et al., 2019] Zheng, X., Zhang, D., Li, Y., Yang, C., Wu, Y., Liang, X., Liang, Y., Pan, X., Hu, L., Sun, Q., Wang, X., Wei, Y., Zhu, J., Qian, W., Yan, Z., Parker, A., Gilles, J., Bourtzis, K., Bouyer, J., Tang, M., Zheng, B., Yu, J., Liu, J., Zhuang, J., Hu, Z., Zhang, M., Gong, J., Hong, X., Zhang, Z., Lin, L., Liu, Q., Hu, Z., Wu, Z., Baton, L., Hoffmann, A., and Xi, Z. (2019). Incompatible and sterile insect techniques combined eliminate mosquitoes. *Nature*, 572(7767):56 – 61.

[Zhu et al., 2010] Zhu, S., Dingding, W., and Li, T. (2010). Data clustering with size constraints. *Knowledge-Based Systems*, 23(8):883–889.

LIST OF FIGURES

LIST OF TABLES