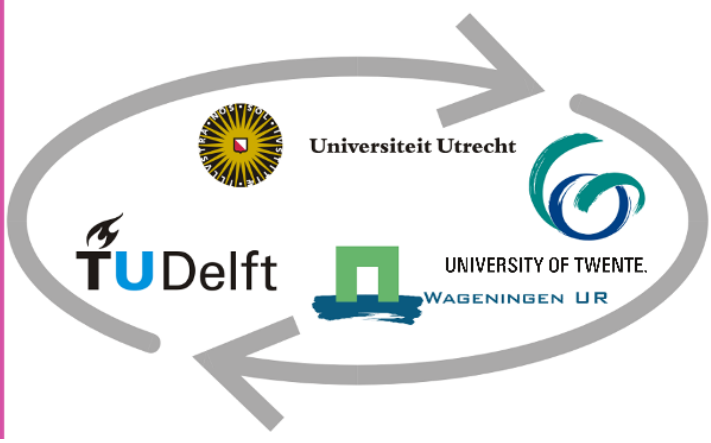
LEPRASIM: An Agent Based Model for the spatio-temporal diffusion dynamics of leprosy infections.

Stijn August van der Hoek



GIMA Master Thesis

Supervisor: Ir. P.W.M. Augustijn
Professor: prof. dr. M.J. Kraak

November 2016

Colophon

November 2016
GIMA MSc Thesis

Contact Information:
Stijn August van der Hoek
Dickenslaan 39-2
3533 BT Utrecht
s.vanderhoek@geogap.nl

UU student number: 3033325
ITC student number: s6020453

Preface

Now that this thesis has finally reached completion, I would like to thank those that have helped me to complete it.

First of all, I would like to extend my sincerest gratitude to Ir. Ellen-Wien Augustijn for her continuous guidance, criticism, enthusiasm and support. Your ability to keep me on the right track has been impressive. Furthermore, my gratitude goes out to Dr. Mirjam Bakker for her support. Her knowledge of leprosy, medication and the study area has been invaluable to this research. Finally, I would like to thank Suzanne for her motivation and love during this period of writing my thesis.

Summary

Leprosy is an infectious disease caused by the *Mycobacterium Leprae* bacterium. In the 1990s, the World Health Organization (WHO) launched a campaign to eliminate leprosy worldwide. Until today this goal has not been reached as the disease maintains itself in India, Brazil and Indonesia. As the worldwide number of new cases has been dropping in recent years, leprosy control programs have shifted their focus from active case detection programs to preventive strategies aimed at direct contacts of people infected with leprosy, as a global interruption of the transmission of the disease is required to fully eradicate the disease. By prescription of an antibiotic (rifampicin) to these contacts the leprosy prevention strategies aim to stop the transmission of leprosy.

In this thesis, an Agent Based Model (ABM) is developed to gain a deeper understanding of the spatio-temporal diffusion dynamics of leprosy infections and the long-term effectiveness of these prevention measures in reducing the occurrence of leprosy in a population (LEPRASIM). In an ABM the heterogeneity of population characteristics relevant to leprosy infection are modeled at an individual level in a spatially explicit manner. For leprosy, the most important characteristics are the susceptibility of individuals to the disease, the (great) variance in incubation time, the variability in infectiousness and the contact groups an individual has contact with.

The model is applied to a case study on a group of five islands in the Flores Sea, Indonesia. Empirical data on leprosy incidence and prevalence are available for these islands as a result of a set of intervention strategy applied by Mirjam Bakker during the period 2000-2010. Although the intervention strategy entailed an active case detection program in combination with a preventive proscription of rifampicin prophylaxis to those contact groups, with the highest probability of infection, leprosy re-emerged on the islands. Several hypotheses exist to explain this re-emergence that can be tested via LEPRASIM. Two sets of experiments are conducted.

In the first experiment a set of two possible explanations for the unexpected re-emergence of leprosy on one of the islands in the case study are considered: (1) a reintroduction of the disease through inter-island marriages; (2) a reintroduction of the disease through contact of fisherman from different islands on fishing boats. The likelihood of these explanations is explored by explicitly modeling them as activities within LEPRASIM. In the second experiment the long-term effectiveness (2000-2025) of a set of eight leprosy prevention strategies, aimed at different contact groups of symptomatic patients through contact tracing of these individuals, is measured. These prevention strategies vary in both the contact groups they are aimed at, as well as in the moment(s) of deployment.

Both the inter-island marriages and the contact of fishermen on fishing boats are proven to be valid explanations for the re-emergence of the disease. The unexpected re-emergence of leprosy can however best be explained by the inter-island marriages, as the effect of this activity on the spatio-temporal diffusion of leprosy in the study area is much larger and of a more deterministic nature.

The effect of the prevention strategies on the incidence rates over time, relative to the baseline model performance (control approach) is measured to examine the long-term effects of each of these prevention strategies. The effect is measured in two different ways: the effect per number of recipients of the medication (1) and the impact on the cumulative incidence rate (2). A prevention strategies aimed at household contacts has the largest and most stable effect when measuring the effectiveness per recipient. This is however only the case in an endemic situation (extremely high prevalence rate), as the effect per recipient declines rapidly with a decrease in prevalence rate. Within the prevention strategies aimed solely at household contacts, the number of people receiving the medication is too low to have a significant impact on the cumulative incidence rate over a longer time period. These prevention strategies are thus not suitable to reduce leprosy incidence over a longer time period in this high prevalence setting.

The Blanket Approach (providing medication to the complete island population) showed both a (relatively) high effectiveness per recipient and a significant reduction in cumulative incidence rates. The Contact Extended Blanket Approach (One time medication to the complete population followed by contact approach) shows the biggest decrease in cumulative incidence rates. In an area where leprosy is highly endemic, as is the case in the case study, the deployment of a prevention strategy which starts with a blanket approach thus is the most effective. As this blanket approach leads to a rather large immediate reduction in incidence rates, the approach can best be followed by a contact approach aimed at both household as well as neighbor contacts of newly detected patients, as this strategy yields the highest effect per recipient of the medication.

Table of Contents

Figures

Tables

Abbreviations

ABM	Agent Based Model
ABDM	Agent Based Disease Model
BCG	Bacille Calmette-Guérin
BB	Mid-Borderline
BI	Bacterial Index
BL	Borderline Lepromatous
BLA	Blanket Approach
BT	Borderline Tuberculoid
CA	Cellular Automata
CIR	Cumulative Incidence Rate
CNT	Contact Approach
CTR	Control Approach
EBB	Extended Blanket Approach
EBC	Contact Extended Blanket Approach
EBH	Household Contact Extended Blanket Approach
ECC	Extended Contact Approach
EHH	Extended Household Contact Approach
GIS	Geographical Information System
HHC	Household Contact Approach
ILI	Influenza-Like Infections
IR	Incidence Rate
KIT	Koninklijk Instituut voor de Tropen
LEPRASIM	The ABM for spatio-temporal diffusion of leprosy infections developed in this thesis.
LL	Lepromatous Leprosy
MB	Multibacillary
MDT	Multi Drug Therapy
NCDR	New Case Detection Rate
PB	Paucibacillary
PEP	Post Exposure Prophylaxis
PR	Prevalence Rate
R^2	Coefficient of determination
SEIR	Susceptible Exposed Infectious Recovered
SIMCOLEP	Leprosy Simulation Model developed by Fisher et al. (2010)
SIMLEP	Leprosy Simulation Model developed by Meima et al. (2004)
TT	Tuberculoid
WHO	World Health Organization

1 Introduction

1.1 Background

Leprosy (Hansen's disease) is an infectious disease caused by the *Mycobacterium Leprae* bacterium. Depending on the immunological response in the patient an infection with this bacterium can cause skin lesions, nerve function impairments and even chronic disability (Blok et al., 2015). In addition, a strong social stigma is related to the disease (Sermrittirong & van Brakel, 2014), leading to isolation and even banishment of leprosy patients over the years.

In 1982 the World Health Organization (WHO) introduced active case detection programs in combination with Multidrug Therapy (MDT) as the main leprosy control program (WHO, 1991). These programs proved to be very successful, as a global leprosy prevalence of less than 1 case per 10.000 people was achieved in 2000 (Meima et al., 2002). Leprosy however still remains a problem in specific parts of the world. The most recent data on the global leprosy situation indicates that in 2014 approximately two hundred thousand new cases were detected, of which 81% are located in only three countries in the world: Brazil (14,5%), India (58.8%) and Indonesia (7.9%) – the so-called endemic countries. In these countries the new case detection rates (NCDR) have remained at a constant level from 1995 onwards (WHO, 2015b, Blok et al., 2015), even after nearly 30 years of MDT (Blok et al., 2014; WHO, 2015b). In these areas, the transmission of the *Mycobacterium Leprae* bacterium has thus not been stopped by the active case detection programs (Blok et al., 2015; Bakker et al., 2006).

The persevering disease burden led to the current leprosy control strategy: the "*Enhanced global strategy for further reducing the disease burden due to leprosy 2011-2015"* (WHO, 2009a), which has been updated (and extended to 2020) by a roadmap target in 2012 stating that the aim is to reduce the number of leprosy patients with visible disabilities by at least 35% in 2015 (when compared to 2010) and the global prevalence rate to 1 per million in 2020 (WHO, 2012). To achieve this goal a global interruption of the transmission of the disease is required, as the transmission of the M*ycobacterium Leprae* bacterium has to be stopped to truly eliminate leprosy (Blok et al., 2015). Two main issues arise in achieving this global interruption of disease transmission. First of all, the actual mechanism of leprosy transmission is not known (WHO, 2015a). Secondly, as even in the endemic countries new cases have become relatively rare, the active case detection programs are no longer cost-effective (Blok et al., 2015; Smith & Smith, 2000). A more targeted approach is thus required to reduce the transmission of the bacterium. New tools and technologies aimed at pre-emptive treatment of direct contacts of leprosy patients are in development (Rodrigues & Lockwood, 2011; Duthie et al., 2012; Bakker, 2005a, Bakker et al., 2006), but their potential long-term impact on the control of leprosy is unknown (Fischer et al., 2011; Bakker, 2005b). To gain an insight into the long-term impact of these pre-emptive treatment strategies this impact either has to be monitored over a long time period, or examined using a computational model.

In the continuous search for improvements to disease control the World Health Organization (2009b) is actively promoting the use of Geographical Information Systems (GIS), as these systems can offer enormous benefits to the control programs. People involved in leprosy control, however, make only a very limited use of GIS in leprosy control. Those leprosy control programs using GIS often do so on an ad-hoc basis (Bakker et al., 2009). For example, by simply mapping the home addresses of patients, like was done first by John Snow in his famous cholera maps of London (Snow, 1854), spatial clusters of leprosy patients are revealed in the study by Guay et al. (2007). Using a spatial scan statistic (SatScan) significant clusters of patients were found on a municipality level in Brazil (Penna et al., 2009) and on a household level in Indonesia (Bakker et al., 2004), clearly showing the added benefit a GIS can have for leprosy control. Using a GIS various epidemiological and potential risk factors can be analyzed together (Bakker et al., 2009). Furthermore, using a spatial model high risk groups and/ or areas for leprosy infection can be predicted (Cromley & McLafferty, 2002).

In recent years, the use of GIS and computational models for simulating the spread of endemic diseases in general has rapidly advanced, as the potential of computational models for simulating the spread of diseases is enormous (Epstein, 2008). This advancement has however occurred in the study of other diseases like tuberculosis (Rodrigues et al., 2015), cholera (Crooks & Hailegiorgis, 2014) and malaria (Linard et al., 2009), but not in the study of leprosy. Epidemiological models aim to test whether mechanisms used to explain an observed phenomenon are sufficient to understand empirical evidence of this phenomenon, and to reveal gaps or inconsistencies in these mechanisms (Helbing, 2012). Most computer-simulation models follow an equation-based approach, but for simulating the spread of an infectious disease an Agent-Based approach is more suited (Helbing, 2012). The spread of a disease can often be considered to be a complex system – a system with many interacting entities and non-linear interactions (Helbing, 2012). Recent models for simulating the spread of endemic diseases (Rodrigues et al., 2015; Crooks & Hailegiorgis, 2014; Linard et al., 2009) therefor primarily use this Agent-Based approach.

The focus of an Agent Based Model (ABM) lies in explaining a certain phenomenon based on realistic assumptions about the elements producing this phenomenon. The goal of an ABM is to gain a greater understanding of the processes causing an observed phenomenon (Helbing, 2012). Furthermore, an ABM can be used to test certain hypotheses in a very detailed manner (Helbing, 2012). For simulating the spread of an infectious disease, like leprosy, an ABM is very suitable, as transmission of the *Mycobacterium Leprae* bacterium occurs at an individual level in a heterogeneous manner (Blok et al., 2015; Grimm et al., 2010).

1.2 Treating Leprosy

Leprosy can be (partially) prevented by either vaccinations or chemoprophylaxis (Blok et al., 2015). Vaccination is the administration of antigenic material to stimulate an individual's immune system to develop immunity to a pathogen and is administered on a whole population (Woolf et al., 2008). Currently, the vaccination against tuberculosis – Bacille Calmette-Guérin (BCG) - offers a

degree of protection against leprosy (Meima et al., 2004; Fisher et al., 2010), but no true leprosy vaccine has been developed yet (Meima et al., 2004; Blok et al., 2015).

Chemoprophylaxis – also known as chemoprevention or post-exposure prophylaxis (PEP) – means the administration of a medication (chemo) to prevent (prophylaxis) a disease or infection (Woolf et al., 2008). In the leprosy case this implies the administration of an antibiotic to potential leprosy patients. As direct contact with infectious individuals is considered to be the main risk factor for leprosy infection (Blok et al., 2015) and a cost-efficient approach is needed (Rodrigues & Lockwood, 2011), leprosy control activities are shifting from a population-based approach to a targeted approach aimed at high-risk groups (Bakker et al., 2005; WHO, 2009a). The cost-effectiveness of this chemoprophylaxis approach for reducing leprosy incidence has been proven by Smith & Smith (2000).

A major problem in leprosy control is that many leprosy cases remain undetected for a long time period as the incubation time of leprosy is very long (up to 20 years) (Blok et al., 2015; WHO, 2003). As these undetected patients are infectious, transmission of the bacterium continues. Therefore, active case detection programs aimed at household and neighbor contacts of newly detected leprosy patients have been employed (Bakker et al., 2006). However, infections do not arise from household and neighbor contacts alone, but from other contacts as well. For this reason Van Beers et al. (1999) have retrospectively researched from which groups of contacts leprosy infections occurred in a 25 year time-period. Besides household and neighbor contacts, social and family contacts were distinguished as separate contact groups. The results of the study are that of all incident leprosy cases only 28% could be classified as originating from a household contact. In the study, 36% of infections originated from a neighbor contact, and 15% from a social or family contact. The remaining 15% could not be classified (Van Beers et al., 1999).

1.3 Rifampicin Prophylaxis

Bakker et al. (2006) have researched the effectiveness of rifampicin (antibiotic) prophylaxis as a preventive measure for leprosy control on a group of five islands in the Flores Sea in Indonesia, highly endemic for leprosy: Sapuka, Sailus, Pelokang, Tampaang and Kembanglemari. Indonesia has a long history of leprosy control. As early as 1969 the government started to integrate leprosy control in the general health services (Peters et al., 2013). On the group of Islands in the Flores Sea leprosy control however has been very irregular (Bakker, 2005a). This is due to the geographical isolation of the islands (Figure 1.1) making healthcare facilities in general and thus leprosy control, very irregular. MDT was introduced on the islands as late as 1993 (Bakker, 2005a). Before this time little to no *dapsone* treatment was available (dapsone is a specific kind of antibiotic). This has resulted in an extremely high prevalence rate on the islands of 195 per 10.000 inhabitants in 2000 (Bakker et al., 2006). This in combination with the isolated location of the islands made the area very suitable for testing the effect of a leprosy prevention program. In addition, the total population of the islands consisted of only 4793 people in June 2000 (Bakker et al., 2006), making a yearly examination of (nearly) the entire population for clinical signs of leprosy possible.

Figure 1.1 Study Area (Google Maps, 2016)

The study by Bakker et al. (2006) is based on the premise (posed among others by van Beers et al. (1999) & Meima et al. (2004)) that not only leprosy patients – those showing the actual signs of leprosy – are infectious, but also sub clinically infected persons – those in the incubation period of the disease (Cree & Smith, 1998). In the study by Bakker et al. (2006), administration of rifampicin (a specific kind of antibiotic) to direct contacts of infectious patients is compared to administration of this drug to the whole population, in order to research at which categories of contacts rifampicin prophylaxis should be aimed to achieve an effective reduction of leprosy incidence in the population. Bakker et al. (2006) define direct contacts as being household, first and second neighbor contacts (both housed within fifty meters), adding a spatial factor to the study.

The antibiotics were administered in 2000 and data on leprosy prevalence and new case detection rates were collected on a yearly basis until 2010 (Bakker, 2016). In this time period the population of all islands was actively screened for signs of leprosy on a yearly basis. On all the islands diagnosed leprosy patients were directly treated with Multi-Drug Treatment (MDT) (Bakker, 2005a). On the first island – Sailus - no chemoprophylaxis was given to anyone – the control group. On the second island – Sapuka - chemoprophylaxis was given to all eligible persons – the blanket group. On the third (group of) islands – Kembanglemari, Pelokang & Tampaang- only eligible direct contacts of known leprosy patients (in 2000) were given chemoprophylaxis – the contact group. All three (groups of) islands have roughly the same population size.

Table 1.1 Study Bakker et al (2006)			
	Control Group	**Contact Groups**	**Blanket Group**
Islands	Sailus	Sapuka	Kembanglemari, Pelokang & Tampaang
Chemoprophylaxis for:	-	Eligible contacts of known leprosy patients in 2000	All Eligible Persons
Observed Incidence 2000-2003 (per 10.000 person-years)	39	35	11

The study by Bakker et al. (2006) showed that chemoprophylaxis had a detectable effect on leprosy incidence for the blanket group only (Table 1.1). No significant differences between the control and contact group emerged, although this was to be expected, given the results of the

study by van Beers et al. (1999) on the causes of leprosy infection. Bakker (2005b) blames the lack of a significant effect of the chemoprophylaxis on leprosy incidence in the contact group primarily to the spatial interpretation of the concept "contact", and feels the incorporation of social or other types of contacts in their study would have altered the results (Bakker, 2005b). Incorporation of social contacts in a prospective study is however hard to accomplish given the size of the islands, the size of the population and the limited funds.

Furthermore, the unexpected re-emergence of leprosy in both the blanket and the contact group of the study raises further questions about the possible influence of factors not incorporated in the study. Two possible explanations for the re-emergence of leprosy in these groups are identified by Bakker (2005b):

- Leprosy is reintroduced on the islands through marriages between the islands
- Leprosy is reintroduced on the islands through contact of fishermen on fishing boats

The cumulative incidence for the blanket group is significantly lower than the cumulative incidence of the control and contact groups (Table 1.1). This is an indication that the rifampicin chemoprophylaxis has had the desired effect, but a longer follow-up period is required to validate this indication (Bakker, 2005b). For the effect of rifampicin chemoprophylaxis three possibilities are identified by Bakker (2005b):

(1) Chemoprophylaxis only delays the development of leprosy;
(2) Chemoprophylaxis prevents leprosy, but only has a temporal effect on transmission of the disease;
(3) Chemoprophylaxis prevents leprosy and also reduces the transmission;

In the first case, the cumulative incidence rates in the blanket group are bound to rise to the same level as the control group over a longer time-period. In the second case the yearly incidence rates will rise to the same level after a shorter time-period, in the third case the two incidence rates will continue to diverge. Both the actual mechanism of leprosy transmittance as the potential long-term influence of rifampicin chemoprophylaxis on the prevalence of leprosy and incidence rates thus remain unknown. To gain a greater insight into these mechanisms and the potential influence of rifampicin chemoprophylaxis on these mechanisms in a cost-effective way a computational model is needed. To model the spatial characteristics of leprosy diffusion in combination with the heterogeneity of the population in their susceptibility to the disease a spatially explicit Agent-Based Model is required.

1.4 Research Objectives

Over the years a number of equation-based computational models for Leprosy have been developed. The first model was developed by Lechat et al. (1974). Extensions on this model were made by Meima et al. (2004) (SIMLEP-model) and Fisher et al. (2010; 2011) (SIMCOLEP-Model). In collaboration with Dr. Bakker a group of UT minor students have developed an initial prototype Agent Based Model (ABM) for the simulation of leprosy in the Flores Sea area (Holtrup et al., 2015). In this master thesis this prototype model shall be expanded into an operational ABM called LEPRASIM for simulating the spread of leprosy. The primary research objective is as follows:

To develop and validate an Agent-Based Model for simulating the spread of leprosy in order to gain a greater insight into the spatio-temporal diffusion dynamics of leprosy infections and the effect of different prevention strategies on this phenomenon.

The main inputs for LEPRASIM are the prototype model developed by Holtrup et al. (2015), the research by Bakker et al. (2002; 2004; 2006), van Beers et al. (1999), the SIMLEP model (Meima et al., 2004) and the SIMCOLEP model (Fischer et al., 2008; 2010; 2011). LEPRASIM will be used to answer the following three research questions:

1. Can the spatio-temporal diffusion dynamics of leprosy be modeled using an Agent-Based approach and does this approach provide a greater insight into these dynamics?
2. Which of the identified reasons for the unexpected re-emergence of leprosy in the blanket and contact group of the study by Bakker et al. (2006) best explains this phenomenon?
3. At which contact group(s) of infectious individuals should a leprosy prevention strategy using rifampicin prophylaxis be aimed to be most effective?

The first research question entails the retrospective part of the study. It is answered by developing the model, documenting it in a reproducible way using the ODD protocol (Grimm et al., 2010) and testing the model's sensitivity to its input parameters: effectively translating the existing leprosy models (SIMLEP/ SIMCOLEP) into an Agent-Based Model. By calibrating the model to the observations made in the study area by Bakker et al. (2002; 2004) on the prevalence rate on each of the islands in 2000 and the source of new infections (household or neighbor contacts) in the period 2000-2003 the spatio-temporal diffusion dynamics of leprosy in the study area are captured.

The second and third research questions entail the prospective part of the study. To answer the second research question LEPRASIM is extended to include the reasons for the unexpected re-emergence of leprosy by explicitly modeling these activities. The effect of the modeled activities on the leprosy prevalence and incidence rates and the spatial clustering of leprosy is examined.

As was shown by van Beers et al. (1999) the different contact groups of infected individuals should play a major role in a leprosy prevention strategy, as the different contact groups account for a

different percentage of the newly detected cases of leprosy. The third research question is answered by extending LEPRASIM with a set of eight leprosy prevention strategies aimed at these contact groups and an examination of their long-term effect on leprosy prevalence and incidence rates in the population.

1.5 Scope

This study will not include the effect of prevention strategies using vaccination as a treatment method. Only the effect of prevention strategies using rifampicin prophylaxis aimed at different contact group(s) of known patients will be examined, as this follows the existing research done by Bakker et al. (2006) most closely. Prevention strategies using dapsone prophylaxis will thus not be included. In addition, although a significant effect of the heritability of susceptibility to leprosy is strongly suspected (Bakker, 2005a; Meima et al., 2004), this is not part of this study, as this effect has been thoroughly researched by Fischer et al. (2011).

1.6 Outline of the thesis

In chapter 2 the theoretical background on leprosy, its diffusion, associated risk factors, disease modeling and agent-based modeling briefly presented in this introduction shall be expanded, resulting in an inventory of existing leprosy models. The third chapter deals with the methodology used for this thesis. In chapter 4 the LEPRASIM-model is presented using the ODD protocol (Grimm et al., 2010). In chapter 5 the sensitivity analysis, calibration and verification of LEPRASIM are presented. A sensitivity analysis on LEPRASIM is performed, consisting of an investigation of the model stability and a sensitivity analysis, using the OAT-method (Hassani-Mahmooei & Parris, 2013). The model is calibrated via a three-step global calibration approach, using the outcomes of the sensitivity analysis. The model is verified by an examination of both the disease as the population model's behavior and its spatial implications. In chapter 6 the results of both the validation of the model on observations made by Bakker et al. (2002; 2004; 2005b; 2006) in the study area, as well as the results of the two experiments are presented. This thesis is closed off with a conclusion and discussion, including suggestions for further research.

2. Theoretical Background

2.1 Introduction

In this chapter a framework for the agent based modeling of leprosy is presented. A review of literature on disease modeling, and agent-based disease modeling (ABDM), in particular, provides the background for this framework (section 2.2). Next, the dynamics particular to the diffusion of leprosy, are presented (section 2.3), leading to an identification of a set of leprosy-specific risk factors (2.3.8). Using the framework for disease models and this set of risk factors the existing leprosy simulation models are reviewed (section 2.4).

2.2 Disease Modeling

By translating a system or phenomenon into a simplified (computational) model, a greater understanding of this phenomenon can be obtained (Voinov, 1999). In general, models are used to test whether mechanisms used to explain an observed phenomenon are sufficient to understand empirical evidence of this phenomenon, and to reveal gaps or inconsistencies in these mechanisms (Helbing, 2012). In recent years, the use of GIS and computational models for simulating the spread of endemic diseases has rapidly advanced, as the potential of computational models for simulating the spread of diseases is enormous (Epstein, 2008). This advancement has occurred in the study of leprosy as well (Meima et al., 2004, Fischer et al., 2010). In other disease areas like tuberculosis (Rodrigues et al., 2015), cholera (Crooks & Hailegiorgis, 2014) and malaria (Linard et al., 2009) disease models have undergone a development from non-spatial equation-based approaches to spatially explicit individual-based (or agent-based) models, but not in the study of leprosy.

Disease models can (in general) be divided into two sub-models: a population-model and a disease-model. In the population-model the development of the population under study over time is modeled using birth, death and ageing mechanisms. The population is divided into compartments representing different disease stages. These compartments can be either groups of individuals or individuals themselves. In traditional disease modeling, transmission of disease occurs between these compartments based on mathematical equations. In more recent disease modeling efforts like individual- or agent-based models this transmission occurs between individuals based on a certain topology of connectedness. This connectedness can be represented by a spatial, household, neighbor or network structure (Blok et al., 2015). In case no topology of connectedness is explicitly modeled this connectedness is implicitly modeled in the general population.

For the analysis of existing leprosy models (section 2.4) the following parameters are used: the type of model and the topology of connectedness employed in the model (Table 2.1). The type of model is compartmental, individual-based or agent-based. The topology of connectedness determines how disease transmissions are modeled.

Table 2.1 : Disease Model Parameters	
Type of Model	Compartmental, Individual-Based, Agent-Based
Topology of Connectedness	General Population, Spatial, Household, Neighbor, Network

First, the traditional disease modeling approach shall be explored (section 2.2.1). Next, the spatially explicit agent-based approach to modeling shall be presented (section 2.2.2). The disease model structure is coupled to the Agent-Based Model structure in section 2.2.3.

2.2.1 Traditional Disease Modeling

Classic disease models model epidemics of infectious diseases using a population-based, non-spatial approach (Perez & Dragicevic, 2009). In these traditional disease models the population is divided into different population compartments (Bian & Liebner, 2005), i.e. the models are compartmental models. These compartmental disease models have a deterministic nature in common, assuming that the entire population is equal on the characteristics relevant for the disease being modeled. Spatial effects in the spread of a disease, individual contact processes and the effects of individual behaviors are thus ignored (Perez & Dragicevic, 2009). The most basic compartmental model for epidemic spread divides a population in two compartments: *Susceptible* and *Infected*, giving the SI model structure. In the SI model structure, a susceptible individual can become infected, and remains infected until death.

A number of extensions on this basic structure can be made: the SIS- model structure, making a re-transition from infected to susceptible possible, the SIR-model-structure, adding a *Recovered* compartment, which is immune to infection, and the SER-structure, replacing the *Infected* with an *Exposed* compartment (Perra & Goncalves, 2015). Lastly, the SIR-model structure has been expanded into the SEIR-model by adding the *Exposed* state to the model, incorporating the incubation phase of a disease, i.e. an infected, but not yet infectious compartment (Perra & Goncalves, 2015), as is done for example in the Global Epidemic And Mobility Model (GLEAM) (Colizza et al., 2006). For bacterial infections - or Influenza Like Infections (ILI) - this SEIR-model-structure is most widely used (Figure 2.1) (Perra & Goncalves, 2015). Within the SEIR model structure, the compartments are not static: individuals can transition between the different compartments.

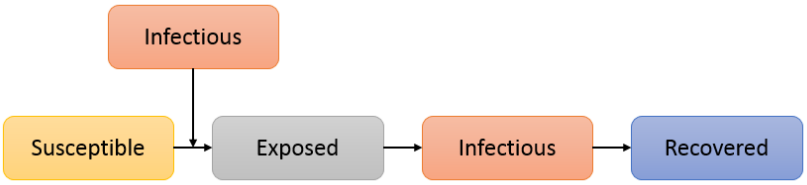


Figure 2.1 SEIR-Model Structure (adapted from Perra & Goncalves, 2015)

The transition of individuals between different stages of a disease can be of two types: spontaneous or interactive. Spontaneous transitions occur at a fixed rate, while interactive transitions require contact between two individuals in different states. The transition from *Infectious*

to *Recovered* can be considered to be spontaneous: some people naturally recover from the modeled disease at a given probability or rate. The transition from the *Susceptible* to the *Infectious* compartment depends on contact of a susceptible person with an infectious person: an interactive transition.

In the compartmental approach the probability of an infection occurring for the *Susceptible* compartment is dependent on the size of the *Infectious* compartment (Perra & Goncalves, 2015). To make the transition from the *Suceptible* to the *Infectious* compartment truly interactive, i.e. to incorporate (spatial) heterogeneity in the number of contacts of individuals within the *Susceptible* compartment and the effect of the behavior of these individuals, one has to move beyond the compartmental modeling approach to an Agent-Based modeling approach (Perez & Dragicevic, 2009).

2.2.2 Agent Based Modeling

Recent developments in disease modeling show the emergence of the Agent-Based modeling approach in this field (Rodrigues et al., 2015; Crooks & Hailegiorgis, 2014; Linard et al., 2009, Blok et al., 2015). Agent-Based Modeling is a specific form of object-based modeling in which individuals, or groups of individuals, are explicitly modeled as agents in a spatial environment. Agent-based simulations are used to study the dynamics of complex systems, i.e. systems with many non-linearly interacting entities (Helbing, 2012). An Agent-Based Model (ABM) typically consists of a representation of time and the following three elements (Macal & North, 2010):

(1) Agents, representing individuals (or groups), their attributes and behaviors. These behaviors can be mathematically formalized, or (more generally) dependent on decision-rules.
(2) Agent relationships and methods of interaction. A certain topology of connectedness must be defined to determine how and with whom agents interact.
(3) The environment, as agents interact with their environment.

By modeling individuals and their relationships in a rule-based way, an ABM can produce features of a system as a whole as an emergent phenomenon, without making assumptions at this aggregate level. It is in the replication of these emergent phenomena, using a minimum set off assumptions that the power of ABMs lies. Agent Based simulations are based on the local interaction among agents, as no central authority exists within the model structure to govern the models behavior as a whole (Macal & North, 2006). The behavior of the system as a whole is thus dependent on the interaction of the individual agents. For simulating the spread of an infectious disease an ABM is very suitable, as transmission of infectious diseases occurs at an individual level in a heterogeneous manner. In addition, ABMs make the study of specific spatial aspects of the spread of a disease possible (Perez & Dragicevic, 2009).

Agents

Within an ABM the population consists of (groups of) individual actors called agents. These agents are simple, self-contained programs (Perez & Dragicevic, 2009). Each agent is a discrete entity with its own goals and behaviors, operating autonomous, having the capability to modify its own

behavior (Macal & North, 2006). Assumptions about key aspects of this behavior and the mechanisms of interaction form the basis of an ABM. From these assumptions the features of a complex system as a whole are modeled as an emergent phenomenon (Epstein et al., 2004). In epidemiology the progression of a disease in a population is modelled through individual agents, making the population heterogeneous in variables relevant to the progression of this disease (for example disease history, health status or susceptibility).

Relationships and rule-based method of interaction

Interactive disease transmissions are explicitly modeled through the contacts of each agent with other agents (Epstein et al., 2004), i.e. via the agent's relationships. Agents can be connected via various topologies, ranging from local neighborhoods in Cellular Automata (CA) to static and/or dynamics networks in social studies (Macal & North, 2006). Within this set of agent relationships rule-based methods of interaction are employed.

Environment

The environment in an ABM consists of the spatial environment and the driving factors of the model. The spatial environment can be explicitly modeled as a Euclidean Space, a network, a raster (Cellular Automata) or more advanced tessellations of space using a GIS (Macal & North, 2006). By modeling the spatial environment explicitly the role of spatial heterogeneity in spreading dynamics of a disease can be understood (Perez & Dragicevic, 2009). The driving factors are mathematical equations determining the conditions by/ in which the agents operate. These driving forces are often based on input data (Fischer et al., 2010).

2.2.3 Agent-Based Disease Modeling

Most ABMs for disease modeling use the SEIR model-structure (Rodrigues et al., 2015; Crooks & Hailegiorgis, 2014; Linard et al., 2009). At initialization of the Agent-Based Disease Model (ABDM) the total population of agents is divided into a susceptible, immune/ recovered and infected sub-set of agents. The infected sub-set of agents is in turn divided into an exposed and infectious sub-set. The period after which an individual agent moves from the exposed to the infectious state is determined by the latency period. The period after which an individual agent moves from the infectious to the recovered state is determined by the infectious period (see Figure 2.2). In ABMs, infections occur in a heterogeneous manner based on individual infectiousness of agents in the Infectious stage of the disease.

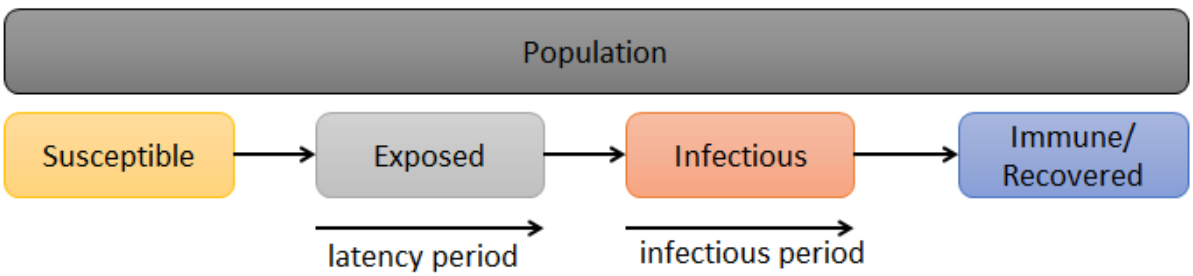


Figure 2.2 ABDM Structure

2.3 Leprosy

2.3.1 Clinical Signs

Leprosy, also known as Hansen's disease, is an infectious disease caused by the *Mycobacterium Leprae* bacterium. Those infected by the bacterium can develop a wide range of clinical signs. The first symptoms usually are skin lesions and enlarged nerves (Bakker, 2005a). The eyes and the mucosa of the upper respiratory tract can be affected as well (WHO, 2003). With progression of the disease nerve function impairments, anesthesia, i.e. the loss of sensitivity, and permanent disabilities can develop (WHO, 2003). The degree in which these symptoms emerge depends on the response of the immune system of infected individuals (Blok et al., 2015). The permanent disabilities are caused by acute increased immunological responses to the bacterium (Lienhardt & Fine, 1994), also known as "reactions" (Bakker, 2005a). In addition to these clinical signs, people affected by leprosy suffer from the social stigma related to the disease (Sermrittirong & van Brakel, 2014).

2.3.2 Diagnosis

The detection of leprosy is dependent on the detection of the clinical signs (2.3.1) (Bakker, 2005a), which means that leprosy can only be detected by means of a physical examination (Blok et al., 2015). After clinical diagnosis, the diagnosis can be complemented by the demonstration of acid-fast *Mycobacterium Leprae* bacteria in slit-skin smears via microscopic examination (Bakker, 2005a). In this way the bacterial load can be determined, giving an indication of the degree of infection. The bacterial load is expressed in the Bacterial Index (BI) (Rees & Young, 1985).

Table 2.2 : Classification of Leprosy Patients

Source	Classes				
Ridley & Jopling (1966)	Tuberculoid (TT)	Borderline Tuberculoid (BT)	Mid-Borderline (BB)	Boderline Lepromatous (BL)	Lepromatous Leprosy (LL)
WHO (1982)	Paucibacillary (PB)		Multibacillary (MB)		
WHO (1998)	PB1	PB2 - PB5		MB	
Meima et al (2004)	Self-Healing			Chronic	

2.3.3 Classification

The first classification system for leprosy patients was introduced by Ridley and Jopling (1966). This system ranged from tuberculoid (TT) to lepromatous (LL) leprosy patients, with three intermediate stages: borderline tuberculoid (BT), mid-borderline (BB) and borderline lepromatous (BL). In 1982 this classification was simplified by the WHO for treatment purposes. All BB, BL and LL patients were classified as multibacillary (MB), all TT and BT patients as paucibacillary (PB). In 1998 this classification was extended based on the number of skin lesions in patients and the bacterial load (BI). In this new classification MB patients show more than five skin lesions, PB1 patients show 1 skin lesion, while PB2-5 show 2 to 5 skin lesions. For modeling purposes Meima et al. (1999) use a different terminology: self-healing and chronic. The self-healing stage of the disease corresponds to the PB classes as defined by the WHO in 1998, while the chronic cases correspond to the MB cases (Meima et al., 1999; 2004) (Table 2.2).

2.3.4 Treatment and self-healing

As leprosy is a bacterial infection, it can be treated with antibiotics. Until 1982 leprosy was treated with a specific antibiotic called *dapsone* (Bakker, 2005a). Over time resistance in patients against this antibiotic grew, leading to the recommendation of the WHO in 1982 to treat leprosy with Multi Drug Therapy (MDT), which involves a combination of two different antibiotics for PB leprosy, and three antibiotics for MB leprosy (WHO, 1997; WHO, 1998; Meima et al., 2004). For MB patients the duration of this MDT is twelve months, for PB patients 6 months (WHO, 1997). PB patients are known to heal spontaneously, but an exact percentage is hard to determine as all detected patients receive treatment (Bakker, 2005a). The percentage of leprosy patients spontaneously healing over time ranges from 15% to 80% (Fine, 1982).

2.3.5 Leprosy control programs and their effect

Since 1982, the main leprosy control program has been an active case detection programs in combination with MDT. As this approach proved to be effective, the World Health Organization stated the goal to eliminate leprosy globally as a public health problem at the 44th World Health Assembly in 1991 (WHO, 1991), defined by a prevalence rate of less than 1 case per 10.000 people. This goal was achieved by active detection programs and MDT at the end of 2000 (Meima et al., 2004). The problem has however not been totally eliminated, as the most recent data on the global leprosy situation (WHO, 2015b) shows that in 2014 a total of 213.899 new leprosy cases were detected. Of the new cases 81% are located in only three countries in the world: Brazil (31,064 = 14.5%), India (125,785 = 58.8%) and Indonesia (17,025 = 7.9%) – the so-called endemic countries.

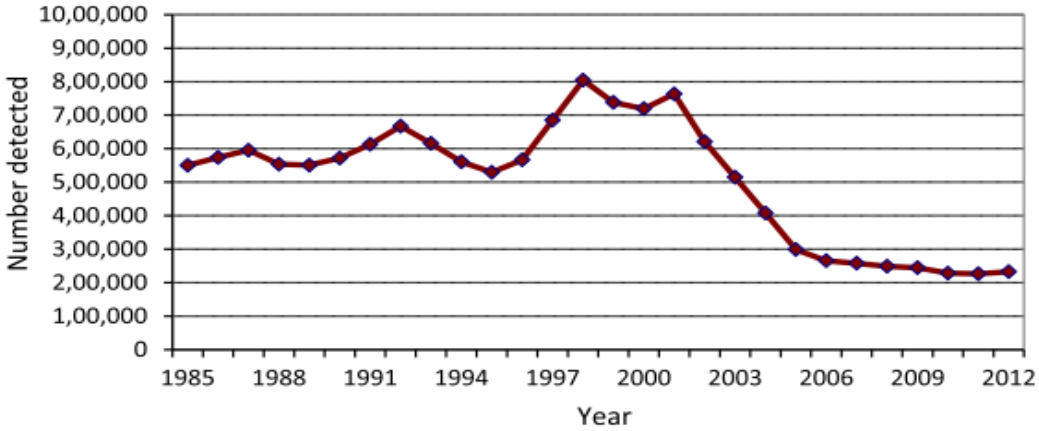


Figure 2.3: Global leprosy new case detection 1985 – 2012 (Blok et al., 2015)

This persevering disease burden has led to the current leprosy control strategy: the "*Enhanced global strategy for further reducing the disease burden due to leprosy 2011-2015*" (WHO, 2009a), which aims to reduce the number of leprosy patients with visible disabilities (PB2 or higher) by at least 35% in 2015 (when compared to 2010). This strategy puts an emphasis on early case detection (Blok et al., 2015). In 2012 the WHO has formulated a roadmap target, extending the duration of the leprosy control strategy, to reduce the number of leprosy patients with visible disabilities to below 1 per million population globally in 2020 (WHO, 2012), and to achieve a

global interruption of the transmission of the disease, as the transmission of the *Mycobacterium Leprae* bacterium has to be stopped to truly eliminate leprosy (Blok et al., 2015).

The effect of the leprosy control approach on the global new case detection rates is shown in Figure 2.3. The increase in new case detections in the period 1994-1997 can be explained by the intensification of the active case detection programs in this period. The New Case Detection Rate (NCDR) (i.e. the number of new cases) has been relatively stable from 2006 onwards (WHO, 2015b), indicating that transmission of the *Mycobacterium Leprae* bacterium has not been stopped (Blok et al., 2015).

To truly eliminate leprosy, as is the goal defined by the WHO (2012), this transmission has to be interrupted. Even in the endemic countries, new cases have however become relatively rare and the classic population-based case detection approaches, which were so successful between 1990 and 2000, are no longer cost-effective (Blok et al., 2015). To eliminate leprosy new tools and technologies aimed at pre-emptive treatment of high risk groups in the population are needed (Rodrigues and Lockwood, 2011): a shift from intervention to prevention. At the moment of diagnosis, contact tracing is performed to locate these high risk groups (Bakker, 2005a, Fischer et al., 2010). A number of new leprosy prevention strategies aimed at household contacts of infectious individuals are in development (Bakker, 2005a, Bakker et al., 2006, Duthie et al., 2012), but their potential long-term impact on the control of leprosy is unknown (Fischer et al., 2011; Blok et al., 2015).

2.3.6 Transmission and sources of infection

Consensus exists about leprosy being a directly-transmitted bacterial infection, thus requiring direct contact between an infected and susceptible person for transmission (Blok et al., 2015). The nose is considered to be the main port of exit and entry of the *Mycobacterium Leprae* bacterium (Shephard, 1962), although skin-to-skin transmissions are considered to be viable routes of transmission as well (Hatta et al., 1995). Indirect routes of transmission cannot be excluded entirely, as *Mycobacterium Leprae* remains viable outside of the human body for some time (Desikan, 1995). The most probable route of transmission is however via the nose.

As untreated multibacillary (MB) leprosy patients have the highest number of bacteria (measured as the Bacterial Index (BI)), they are the main source of infection (Bakker, 2005a). After the first dose of MDT, MB patients are considered to be no longer infectious (WHO, 2003). Over the last 15 years the stability of new case detection rates has shown that the transmission of the *Mycobacterium Leprae* has not been stopped, despite the intensification of the leprosy control programs. This leads to the general acknowledgement that sub clinically infected persons, i.e. those persons in the incubation period of the disease, are sources of infection as well (Meima et al., 2004; Bakker, 2005a; Cree & Smith, 1998). The infectiousness of PB leprosy patients has not been proven nor disproven (Bakker, 2005a), but is dependent on the BI of these patients, which is relatively low.

2.3.7 Incubation Periods

The exact time of infection cannot be determined in most cases, leading to an approximation of the incubation time from war veterans who served in endemic areas (Fine, 1982; Noordeen, 1985). For tuberculoid/ PB1 leprosy, this incubation period is between 2 and 5 years, for lepromatous/ MB leprosy between 8 and 12 years (Bakker, 2005a). These incubation periods are known to vary widely, ranging from just a few to up to 30 years (Bakker 2005a, Noordeen, 1985). Possible explanations for the extremely long incubation periods might be re-infection or reactivation of dormant bacilli (Noordeen, 1985; Meima et al., 2002), but the exact explanation is unknown (Bakker, 2005a).

2.3.8 Risk factors

Differences between individuals in the amount of exposure to infectious leprosy patients (contacts) and the development of leprosy after exposure (susceptibility) cause heterogeneity in leprosy infections (Blok et al., 2015). The factors determining a vulnerability to progression to a disease are known as risk factors (Bakker, 2005a). Bakker (2005a) has performed a review of 41 available cohort studies to identify these risk factors, using a general set of criteria. For leprosy the risk factors for – or determinants of – leprosy can be categorized into three levels: the individual, contact and macro level (Bakker, 2005a). The risk factors for leprosy shall be discussed following this categorization (Table 2.3).

Table 2.3 : Leprosy Risk Factors

Level	Risk Factor
Individual	Susceptibility
	Genetic Factors
	Health Factors
	Serological Status
	BCG Vaccination
	Chemoprophylaxis
Contact	Type of Contact
	Infectiousness of Contact
	Number of Contacts
Macro	Leprosy Prevalence in Population
	Socio-economic Factors
	Social Stigma
	Leprosy Control Programs

2.3.8.1 Individual level

Susceptibility

Different people react differently to an exposure to *Mycobacterium Leprae*: their susceptibility to leprosy differs. In his pioneer research on leprosy transmittance, Fine (1982) discovered that approximately 5-20% of the population is susceptible to the development of leprosy after exposure. Differences in susceptibility are caused by a combination of genetic and health factors (Blok et al., 2015).

Genetic Factors

Genetic factors explain up to 57% of the total variance in susceptibility in a study performed by Bakker et al. (2006). These genetic factors also play a role in the development of the disease from PB to MB leprosy (Bakker et al., 2006; Mira et al., 2004). It is unclear whether susceptibility to leprosy is inherited via a recessive or a dominant genome (Fischer et al., 2010), or how much of the variability in leprosy susceptibility can exactly be explained by genetic factors (Bakker, 2005a). The literature review performed by Bakker (2005a) does however strongly suggest that there is a relationship.

Health Factors

Health factors relate to the housing conditions, sex, age and nutritional status, influencing the overall health of people and thus their susceptibility to any bacterial infection (Bakker et al., 2006). The study by Pönnighaus et al. (1994) shows a clear relationship between housing conditions and the risk of leprosy, as this risk increases with falling housing standards. In her literature review Bakker (2005a) shows that the relationship between sex and developing leprosy is not that straightforward. In general it can however be stated that men have a higher risk of developing leprosy (Bakker, 2005a). The relationship between age and the development of leprosy is not clear, as most studies reported the highest incident rates among children (aged 5-14) (Bakker, 2005a), while several studies reported the highest incident rates among the elderly (Bakker, 2005a). The relationship between nutritional status and the risk of developing leprosy has not been researched (as this is very hard to do). It can however be logically assumed that people with a bad nutritional status or people in contact with a lot of other mycobacteria have a weaker immune system and are thus more susceptible to leprosy.

Serological Status

When a person is seropositive to a disease it means that the person shows a "*presence of antibodies or other immune markers in serum that indicate prior exposure to a particular organism or antigen*" (O'Toole, 1997). In her review of studies on leprosy risk factors, Bakker (2005a) shows that different studies show different relationships between seropositivity and the risk of leprosy infection, ranging from a slightly negative to a strong positive relationship. In their own study on clustering of seropositive individuals Bakker et al. (2004) conclude that living in the vicinity of

seropositive patients increases the risk of getting infected. The serological status of a person (whether he/she is seropositive or seronegative to leprosy) appears to be a good indicator for transmission potential of patients. Being seropositive to leprosy, i.e. having been exposed to the disease at an earlier moment in time, thus highly increases the risk of developing the disease.

BCG Vaccination

Leprosy can be (partially) prevented by either vaccinations or chemoprophylaxis. Vaccination is the administration of antigenic material to stimulate an individual's immune system to develop immunity to a pathogen and is administered on a whole population (Woolf et al., 2008). Currently, the vaccination against tuberculosis –Bacille Calmette-Guérin (BCG) - offers a degree of protection against leprosy (Fine, 1982), but no true leprosy vaccine has been developed (Meima et al., 2004). The effect of the BCG vaccination on the prevention of leprosy has been researched by a number of authors, (Fine, 1982; 1988; Stanley et al., 1981) leading to widely varying results (Bakker, 2005a). The efficacy of the BCG vaccination against leprosy is estimated at 20-80% (Bakker, 2005a). The degree of protection to leprosy the BCG vaccination gives is thus unclear, but it is clear that it offers some protection.

Chemoprophylaxis

Chemoprophylaxis – also known as chemoprevention – means the administration of a medication (chemo) to prevent (prophylaxis) a disease or infection (Woolf et al., 2008), in this case the administration of an antibiotic to potential leprosy patients. This serves two separate control goals. Firstly, the development of leprosy in infected, but not detected, individuals is prevented. Secondly, the transmission of the *Mycobacterium Leprae* bacterium from infected individuals to susceptible individuals is prevented (Bakker, 2005a). The cost-efficiency of (dapsone) chemoprophylaxis for household contacts of infected individuals was proven by Smith & Smith (2000) and the effect of rifampicin chemoprophylaxis examined by Bakker et al. (2006).

2.3.8.2 Contact level

Type of contact

As leprosy is a bacterial infection, contact between an infected and non-infected person is needed for the disease to transmit. The amount of contact a person has with infected people thus is a major risk factor (for infection with any directly-transmitted disease) (Wallinga et al., 1999). A number of studies have shown this effect to be true for the spread of leprosy in countries all over the world (Blok et al., 2015; Moet et al., 2004; Bakker et al., 2004): contacts of leprosy patients have a higher risk of getting infected. As leprosy is transmitted through direct contact, and the total number of leprosy patients is decreasing, an increasing percentage of new cases are resulting from contacts within households (Richardus et al., 2005). In her literature review, Bakker (2005a) shows that all studies show an increased risk for household contacts, although the magnitude of this increase in risk is unclear (ranging from a factor 2 to 9).

In their retrospective study on determinants for incident leprosy van Beers et al. (1999) also showed that contact is the major determinant for new infections (80% of all new cases can be explained by a specific infectious contact), but their definition of a contact is broader. They define a contact as a household member, a neighbor or a social contact, explaining 28%, 36% and 15% of new infections respectively. In general, clustering of leprosy patients within households, families and neighborhoods has been reported many times (Fischer et al., 2010). As family members often live together in a household, it is however hard to distinguish between the effect of susceptibility and contact risk factors (Bakker, 2005a; Bakker et al., 2006; Fischer et al., 2010).

In an effort to make this distinction Bakker et al. (2002) expanded the concept of contacts to include neighborhood contacts as well. A distinction was made between direct and next neighbor contacts (neighbor 1 and neighbor 2), both living within 50 meters distance of a leprosy patient (Bakker et al., 2002). An analysis of the clustering of leprosy patients in three types of clusters – within a house, a house and its direct neighbors or a house, its direct and next neighbors – showed that patients indeed tend to cluster (Bakker et al., 2002), giving a 2.07, a 1.75 and a 1.34 times higher risk for the respective clusters. This means that spatial clustering of leprosy primarily occurs within households, but within neighborhoods as well. Having infectious neighbor contacts thus increases the risk of a leprosy infection. In the follow-up study Bakker et al. (2004) further investigated the spatial clustering of leprosy prevalence using GIS, indicating that leprosy infections are significantly higher among persons living in close proximity (up to 75 meters) to leprosy patients. This clustering of leprosy patients is very context dependent, as no clear relationship between population density and leprosy incidence in general has been found (Pönnighaus et al., 1994).

The retrospective study performed by van Beers et al. (1999) shows that not only household contacts and neighbors have an increased risk of developing leprosy, but also social contacts. No further research into the magnitude of this risk has been done.

Infectiousness of contact

Untreated multibacillary (MB) leprosy patients are more infectious than PB patients (Bakker, 2005a). Furthermore, sub clinically infected persons are infectious as well, but less infectious than symptomatic patients (Meima et al., 2004; Bakker, 2005a; Cree & Smith, 1998). In their cohort study on risk factors for developing leprosy Bakker et al. (2006) show that household contacts of MB patients have an approximately five times higher risk of developing leprosy than household contacts of PB patients.

Number of contacts

The stage of leprosy a contact has, i.e. PB or MB, has a bigger influence on the risk of developing leprosy than the number of infectious contacts an uninfected individual comes into contact with (Meima et al., 2004). The risk does however increase with the number of infectious contacts (Bakker, 2005a). Furthermore Bakker (2005b) indicates that the bigger the household one lives in is, the higher the risk of developing leprosy is which is logical given the transmittance of leprosy

through direct contact. Household contacts of more than one leprosy patient have an approximately three times higher risk than household contacts of one leprosy patient in the study by Bakker et al. (2006).

2.3.8.3 Macro level

Leprosy Prevalence in population

In general it can be stated (for any bacterial infection) that the higher the prevalence of a disease in a population is, the higher the risk of getting infected is for any individual in that population (Perra & Goncalves, 2015). This is also true for leprosy, as the probability of contact with infectious patients, and thus the probability of infection through this contact, increases with the amount of infectious patients in the population.

Socio-economic factors

An improvement of the socio-economic situation has a lowering effect on leprosy incidence in a population (Bakker, 2005). The clearest evidence of this is the absence of leprosy in the wealthier countries of the world (WHO, 2015b). A clear relationship between education levels and leprosy incidence is shown by Pönnighaus et al. (1994). This relationship can likely be explained by the relationship between the education level and wealth, and thus the housing conditions of an individual. Better housing conditions, i.e. better hygienic facilities, improve the overall health of individuals, and thus lower the chance of infection with any bacterium.

Social Stigma

The stigmatization as a result of leprosy has existed since ancient times. The causes for this stigma lie in the chronic disability caused by leprosy, religious/ cultural believes, the fear of transmission and public-health related interventions (Sermrittirong & van Brakel, 2014). For example, in Indonesia leprosy patients were isolated compulsory from 1655 until 1932 in leprosy asylums (Peters et al., 2013). Similar segregation policies were used in the end of the nineteenth century in many countries of the world (Sermrittirong & van Brakel, 2014), increasing the fear of leprosy and thus the stigma related to the disease. The effect of this stigma on leprosy incidence and prevalence rates is hard to quantify (Bakker, 2005a).

Leprosy control programs

The efforts to control leprosy are aimed at preventing nerve damage in leprosy patients and an interruption of the transmittance of the disease. A study on the effects of leprosy control measures by Leiker & Fisher (1976) showed that segregation policies, i.e. the compulsory isolation of identified leprosy patients, has had no effect on leprosy incidence, while active case detection programs, dapsone treatment and MDT do cause a decline in incidence.

2.4 Existing Leprosy Models

Over the years four models simulating leprosy diffusion dynamics have been developed. Lechat et al. (1974) developed the first model, which Meima et al. (2004) extended into the SIMLEP model. In 2010 Fischer et al. presented a micro-simulation (or individual-based) model called SIMCOLEP. In 2015 Holtrup et al. developed a prototype Agent-Based Model for leprosy diffusion on the islands in the Flores Sea. Over the years a development from deterministic compartmental leprosy models (Lechat et al., 1974, Meima et al., 2004) to an individual-based (Fischer et al., 2010) and Agent-Based (Holtrup et al., 2015) approach can be observed. The four identified models are presented and the different disease model structures are made comparable using the SEIR-model structure. Each model is analyzed on the general disease model parameters and inclusion of the leprosy risk factors (Table 2.4).

Table 2.4 : Framework Analysis Existing Leprosy Models

Disease Model Parameters	
Type of Model	Compartmental, Individual-Based, Agent-Based
Topology of Connectedness	General Population, Spatial, Household, Neighbor, Network
Leprosy Risk Factors	
Level	**Risk Factor**
Individual	Susceptibility
	Genetic Factors
	Health Factors
	Serological Status
	BCG Vaccination
	Chemoprophylaxis
Contact	Type of Contact
	Infectiousness of Contact
	Number of Contacts
Macro	Leprosy Prevalence in Population
	Socio-economic Factors
	Social Stigma
	Leprosy Control Programs

2.4.1 Lechat's model

Lechat et al. (1974) developed the first compartmental leprosy model, enabling the investigation of leprosy development in populations and the long-term effect of different leprosy control strategies. In the model the whole population is considered to be susceptible to leprosy (Lechat et al., 1985). New infections are modeled as a function of the number of infected persons in the whole population (Blok et al., 2015). At infection, an individual enters the latent stage, where a distinction between PB and MB is made. After a latency period, the individual enters the disease stage (Lechat et al., 1974). As in the SEIR-model, a distinction between uninfected, latent, infective and non-infective compartments is thus made (Figure 2.4).

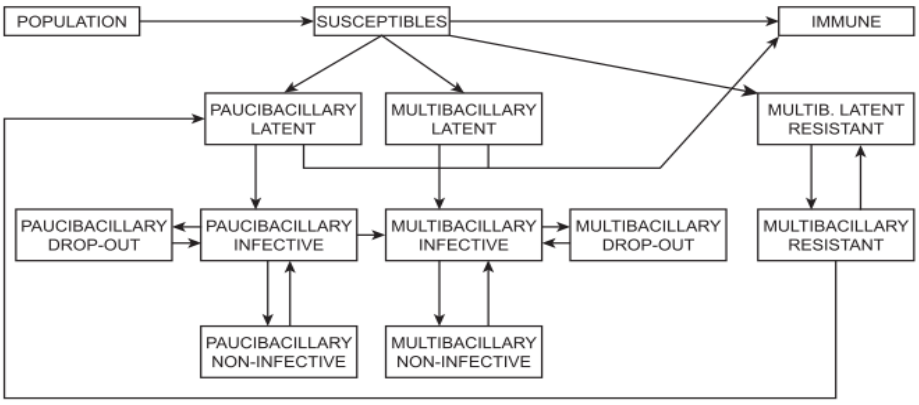


Figure 2.4: Lechat's Leprosy Model (Lechat et al., 1985)

In their expansion on the model Lechat et al. (1985) made the incidence rates age- and sex-specific. Whether a person develops the disease and becomes infective or becomes immune depends on the reaction of his immune system to the disease, caused by a combination of health and genetic factors. An infected person can either drop-out (i.e. die or move out of the study area) or self-heal. An infected individual can also be treated and become non-infective (Recovered). In Lechat's model, the capacity of a person to transmit leprosy is thus a function of the stage of disease and the treatment-status (Lechat, 1992). By adding a "resistant" stage to the model structure Lechat et al. (1985) have taken the serological status of patients into account. It should be noted that in the model structure, patients can develop leprosy from PB to MB, only when they are infective, as a person infected with PB will already be infective before he can reach the MB-stage.

Translated to the SEIR-model, Lechat's leprosy model looks like the structure depicted in Figure 2.5. The stages PB/ MB latent and PB/MB infective in Lechat's model are seen as different classes of the exposed and infectious stage in the SEIR model respectively. The Recovered stage of the SEIR model has been modeled via four different stages in Lechat's model, as the serological, treatment and disease-stage status of individuals were taken into account.

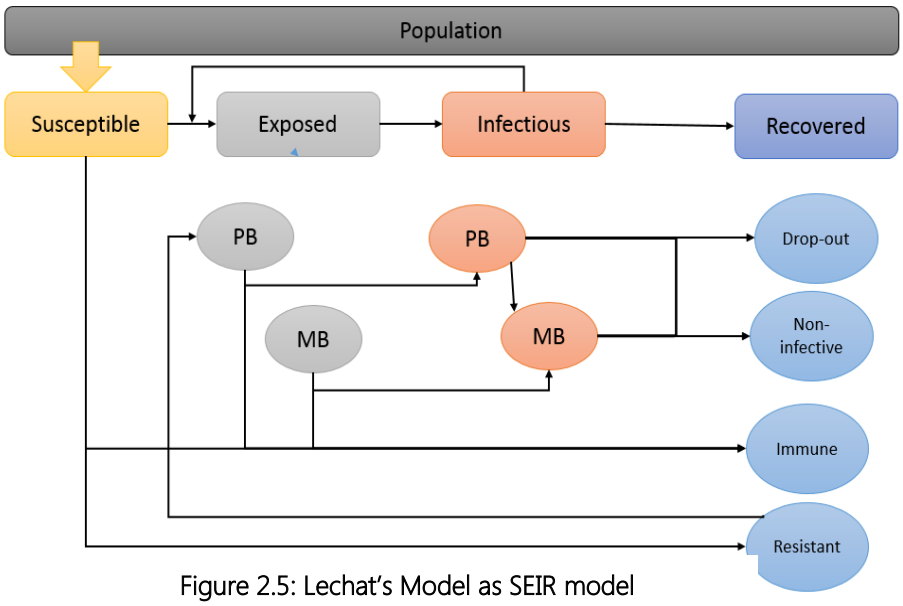


Figure 2.5: Lechat's Model as SEIR model

In Table 2.5 Lechat's model is analyzed on the general disease model parameters and incorporation of identified risk factors. The model is compartmental and models infections only in the general population. The model does account for heterogeneity in infectiousness, the serological status of individuals and the effect of leprosy prevalence in the population on the probability of infection. The model is used to test the effect of a number of leprosy control programs (Lechat, 1992). None of the other identified risk factors are incorporated in the model.

Table 2.5 Lechat's Model (1974)			
General Disease Model parameters			
Type of model		Compartmental	
Topology of connectedness		General Population	
Level	**Risk Factor**	**Incorporation?**	**Description**
Individual	Susceptibility	-	The whole population is equally susceptible to leprosy
	Genetic factors	-	-
	Health factors	-	-
	Serological status	x	A compartment of the population can become resistant to leprosy
	BCG vaccination	x	The effect of BCG vaccination on leprosy prevalence is tested
	Chemoprophylaxis	-	-
Contact	Type of contact	-	
	Infectiousness of contact	x	Different stages of the disease (PB/ MB) have different levels of infectiousness
	Number of contacts	x	As the number of infected individuals in the population increases, the probability of infection does as well.
Macro	Leprosy Prevalence in population	x	Infections are modeled as a function of the leprosy prevalence in the population
	Socio-economic factors	-	-
	Social Stigma	-	-
	Leprosy control programs	x	The effect of dapsone monotherapy, MDT treatment and BCG vaccination on leprosy prevalence is explored

2.4.2 SIMLEP model

The SIMLEP model (Meima et al., 1999) expanded on Lechat's model in order to make better predictions of future trends in leprosy (Meima et al., 2004). In the SIMLEP model variations in natural immunity, incubation periods and asymptomatic infection (latent in Lechat's model structure), delays in awareness and thus delays in treatments are taken into account (Blok et al., 2015). The level of infectiousness differs per type of infection, and assumptions are made about these levels (Meima et al., 2004). The most important assumption is that only MB patients are considered to be infectious, but are so during the incubation phase of the disease as well (Meima et al., 2004). All PB patients are considered to be non-infectious and self-healing over time (Blok et al., 2015). The infectiousness of MB patients increases during the latency period (asymptomatic stage) and remains stable during the symptomatic stage (Figure 2.6).

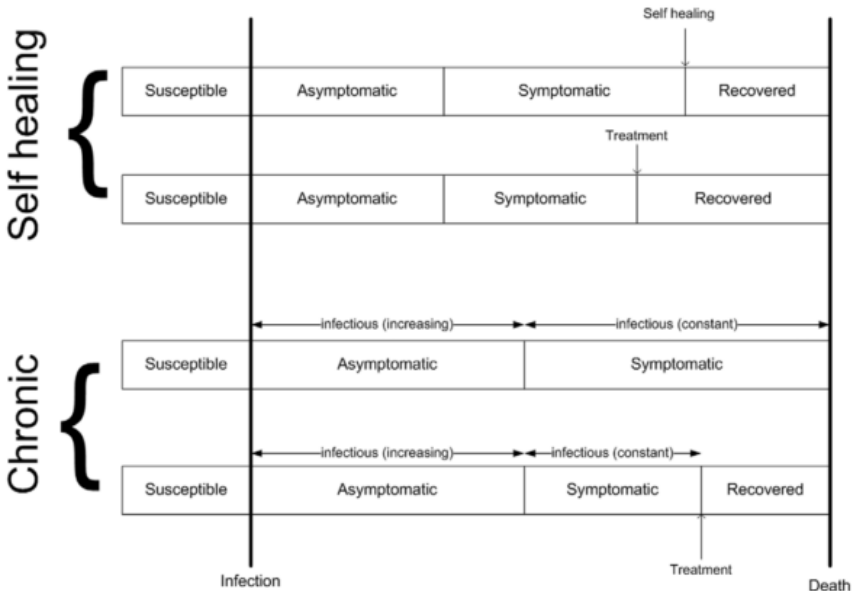


Figure 2.6: SIMLEP Model (Meima et al., 2004)

In contrast to Lechat's model, not the whole population is considered to be susceptible to leprosy in the SIMLEP model: a non-susceptible (or immune) compartment is added. In the SIMLEP model an asymptomatic infection can develop over time (incorporating the incubation period) into one of two types of symptomatic infections: self-healing (PB) or chronic (MB).

The number of new infections for each time-step in the SIMLEP model is a function of the number of infectious individuals in the population, with the infectiousness being stage-dependent (Blok et al., 2015; Meima et al., 2004). The effect of treatments (and relapses on these treatments) are modeled, including time-delays, by moving a fraction of the self-healing symptomatic compartment to either the immune (Treated) or susceptible compartment in each time-step. The chronic patients do not self-heal over time, but can only recover through treatment, incorporating a time-delay before this treatment (Meima et al., 2004). The model structure, translated to the SEIR terminology is depicted in Figure 2.7.

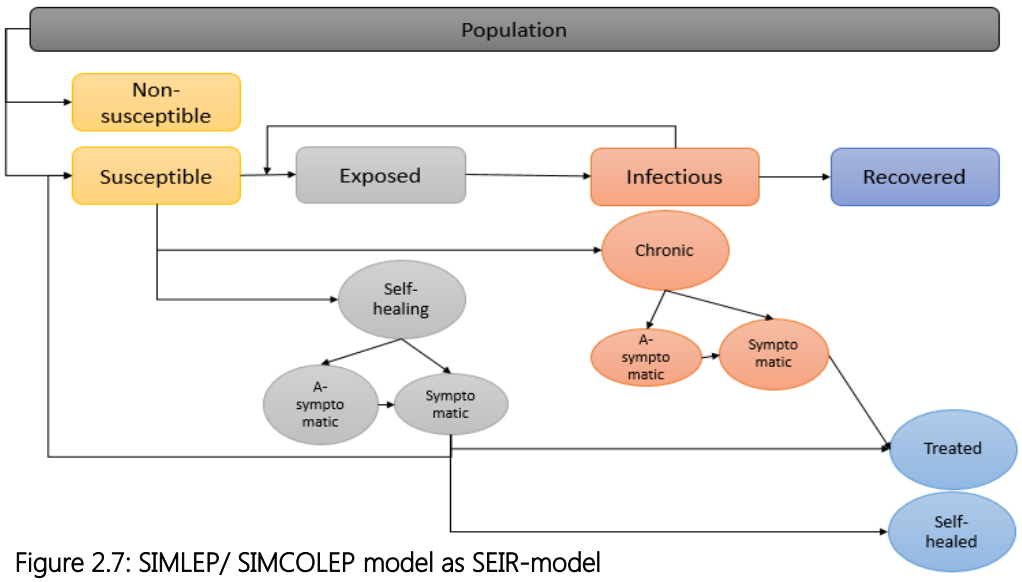


Figure 2.7: SIMLEP/ SIMCOLEP model as SEIR-model

The SIMLEP model was used to test the impact of an active case detection program, or "early case detection strategy" as Meima et al. (2004) call it, on leprosy prevalence in a population over a longer time period. Furthermore, the consequences of not sustaining this strategy on this prevalence are tested, under a variation of assumptions. Scenario's with and without BCG vaccination are employed and varied on different coverage and protection grades from this vaccination (Meima et al., 2004). The study concluded that not continuing the active case detection program will lead to an increase in leprosy prevalence. When continuing the early case detection strategy the decline in leprosy incidence will slowly continue. The implementation of a global long-term strategy to control leprosy is highly recommended (Meima et al., 2004), as uncertainty about the adverse effects of a longer delay in detection are ultimately unknown.

Table 2.6 SIMLEP (Meima et al, 2004)			
General Disease Model parameters			
Type of model		Compartmental	
Topology of connectedness		General Population	
Level	**Risk Factor**	**Incorporation?**	**Description**
Individual	Susceptibility	x	A fixed percentage of the population is susceptible to leprosy
	Genetic factors	-	-
	Health factors	-	-
	Serological status	x	Self-healed patients are immune to further infection
	BCG vaccination	x	The model is used to test the long-term effect of BCG
	Chemoprophylaxis	-	
Contact	Type of contact	-	
	Infectiousness of contact	x	Different stages of the disease (PB/ MB) have different levels of infectiousness
	Number of contacts	x	As the number of infected individuals in the population increases, the probability of infection does as well.
Macro	Leprosy Prevalence in population	x	Infections are modeled as a function of the leprosy prevalence in the population
	Socio-economic factors	-	-
	Social Stigma	-	-
	Leprosy control programs	x	The long-term effect of BCG vaccinations in combination with various degrees of early case detection (two to four years detection delay) are tested

Analysis of the SIMLEP shows that the SIMLEP model differs from Lechat's model on a number of points (Table 2.6). The most important point is that the SIMLEP model accounts for susceptibility to leprosy, while Lechat's model does not. The model is however still a compartmental model and thus does not account for heterogeneity in this susceptibility or the disease dynamics within households. To test the effect of interventions aimed at household contacts of infected individuals (like chemoprophylaxis) this was needed. Therefore the SIMCOLEP model was developed (Blok et al., 2015).

2.4.3 SIMCOLEP model

In order to take heterogeneity in susceptibility into account the SIMCOLEP model was developed by Fischer et al. (2010). In this model population processes are split from disease and transmission processes, and households are introduced as simulation units (Figure 2.8). The model represents a shift from a deterministic compartmental modeling approach to a stochastic individual-based modeling approach in the study of leprosy. Within SIMCOLEP disease dynamics at the population level are procured via aggregation of the disease dynamics on an individual level (Fischer et al., 2010). The delays in awareness and treatments, as introduced by Meima et al. (1999), are incorporated into the SIMCOLEP model through the disease processes.

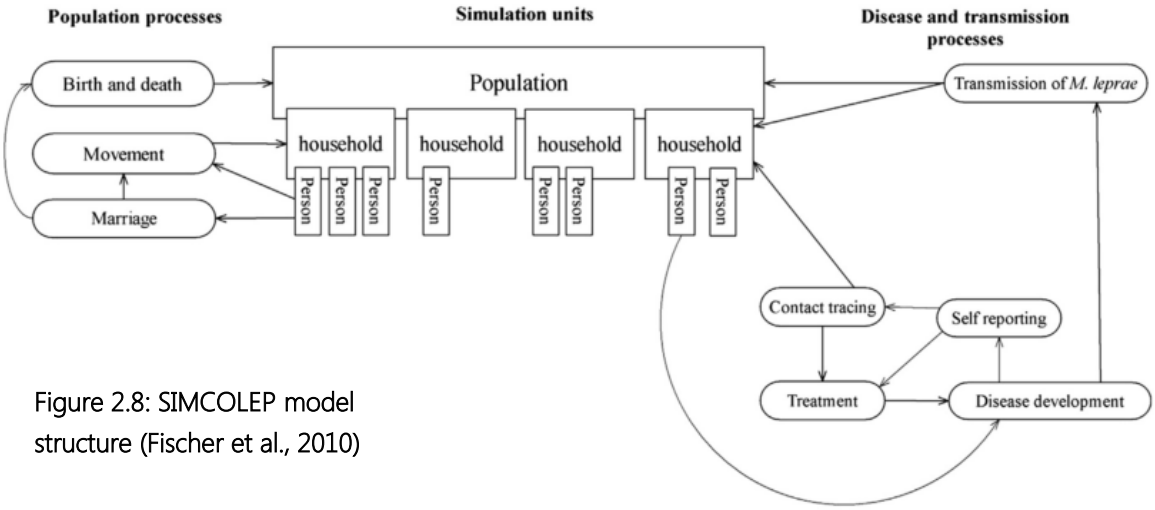


Figure 2.8: SIMCOLEP model structure (Fischer et al., 2010)

The SIMCOLEP model was developed (among other reasons) to explore the long-term effectiveness of different intervention strategies (Fischer et al., 2011) aimed at household contacts of infected individuals. Social and neighborhood (spatial) contacts are however not taken into account. The SIMCOLEP model uses the same disease model structure as the SIMLEP model (Figure 2.7), but implements it at an individual level. In SIMCOLEP the life history and natural history of infection with the *Mycobacterium Leprae* bacterium of individuals is simulated. Demographic (or population) processes in the model are defined by birth-, death- and movement-rates via a so-called life-Table (Fischer et al., 2010), giving a growing population distributed over a dynamic household structure during one model run. This makes the evaluation of intervention strategies aimed at household contacts of infected individuals possible.

Following the SIMLEP model, the SIMCOLEP model assumes a fixed percentage of individuals to be susceptible to leprosy. Infectiousness in the SIMCOLEP model is dependent on the type of disease: self-healing or chronic. As in the SIMLEP model, only the "chronic" stage is infectious. For this subset, the infectiousness increases during the asymptomatic stage, and remains constant during the symptomatic stage (Fischer et al., 2010).

Infections in SIMCOLEP occur via two separate processes. First of all, transmission of leprosy from an Infectious individual to random individuals in the population occurs at a fixed rate of contacts per year multiplied by a factor representing the infectiousness of the individual (an infection probability). Secondly, additional transmission of leprosy within households containing an infectious individual is explicitly modeled by applying the infection probability (infectiousness) of this individual on each household member. Only symptomatic patients are treated in SIMCOLEP-simulations (Fischer et al., 2010).

Table 2.7 SIMCOLEP (Fischer et al, 2010)			
General Disease Model parameters			
Type of model		Individual-Based	
Topology of connectedness		General Population, Household Structure	
Level	**Risk Factor**	**Incorporation?**	**Description**
Individual	Susceptibility	x	A fixed percentage of the population is susceptible to leprosy
	Genetic factors	x	The model is used to test the influence of heritability of susceptibility to leprosy on leprosy prevalence
	Health factors	-	
	Serological status	x	Self-healed patients are immune to further infection, given a relapse rate
	BCG vaccination	x	BCG coverage and protection are incorporated in the model
	Chemoprophylaxis	x	One of the intervention strategies tested with the model was chemoprophylaxis aimed at household contacts of infected individuals
Contact	Type of contact	x	Infections of household contacts and contacts in the general population are seperately modeled.
	Infectiousness of contact	x	Different stages of the disease (PB/ MB) have different levels of infectiousness
	Number of contacts	x	As the number of infected individuals in the population/ household increases, the probability of infection does as well.
Macro	Leprosy Prevalence in population	-	The global infection rate in the model is independent of the leprosy prevalence in the population
	Socio-economic factors	-	
	Social Stigma	-	
	Leprosy control programs	x	Seven different intervention strategies are tested: baseline scenario, no contact tracing, chemoprophylaxis to all household contacts, early case detection, BCG vaccination to all new-borns, BCG + chemoprophylaxis, BCG + early case detection

In the first study by Fischer et al. (2010) the SIMCOLEP-model is used to test for the effect of different genetic inheritance of susceptibility to leprosy on leprosy prevalence in a population, by adding different inheritance-mechanisms to the individual-based model. The result of this study was that none of the tested mechanisms for leprosy susceptibility allocation could be ruled out (Fischer et al., 2010). In another study (Fischer et al., 2011) the same model was used to test the effect of different intervention strategies, of which active case detection programs in combination with chemoprophylaxis aimed at household contacts proved to be most effective.

The SIMCOLEP model represents the shift in leprosy modeling from a compartmental to an individual-based approach (Table 2.7). Of the identified risk factors, only the individual health factors, socio-economic factors and the role of social stigma are not taken into account within SIMCOLEP. The model is thus a rather complete model. It is however not spatial, making the effect of infections to neighbor contacts (as observed by Bakker, 2005a) on leprosy prevalence a gap in the model's representation of the dispersion dynamics of leprosy.

2.4.4 Prototype Agent-Based Model

The prototype agent-based model developed by Holtrup et al. (2015) represents the first effort of an Agent-Based approach to simulating the spread of leprosy. In their prototype model Holtrup et al. (2015) have split up their model into four components: a world, a population, an activity and a disease model. In the world model the spatial characteristics of the islands are represented in a NetLogo environment by representing the study area - a group of islands in the Flores Sea - on a map. In the population model a population, or agent-set, is "created" with an appropriate age and sex distribution taken from the work by Bakker (2005a). The population is distributed over a number of randomly distributed houses on the islands by allocating one man, woman and a random number of children to each house (Holtrup et al., 2015), effectively distributing the population over a number of households. The activity model represents the population processes as described in the SIMCOLEP model (Figure 2.8) and the explicit modeling of men going fishing on boats, where their chance of leprosy infection changes. All other activities of agents are not modeled as such, but as infection chances. The model assumes that the population size remains the same, and implements this assumption by replacing every dying agent with a new one (Holtrup et al., 2015).

The disease model consists of four infection chances: the world, island, household and boat infection chances (Holtrup et al., 2015). The world infection chance represents the chance of infection by contact of an agent with any of the other agents in the simulation world: a representation of leprosy prevalence in the population. The island infection chance represents the chance of infection by contact of an agent with any other agents on the same island in the simulation world. For the household infection chance this is contact within a household, for the boat infection chance the chance of infection within a boat. A boat consists of a random number of men from any of the islands, making cross-infection between the islands a part of the model (Holtrup et al., 2015). It should be noted that Holtrup et al. (2015) already incorporate this effect by

using the leprosy prevalence in the population as an infection chance within the world infection chance, and are thus modelling the same effect twice.

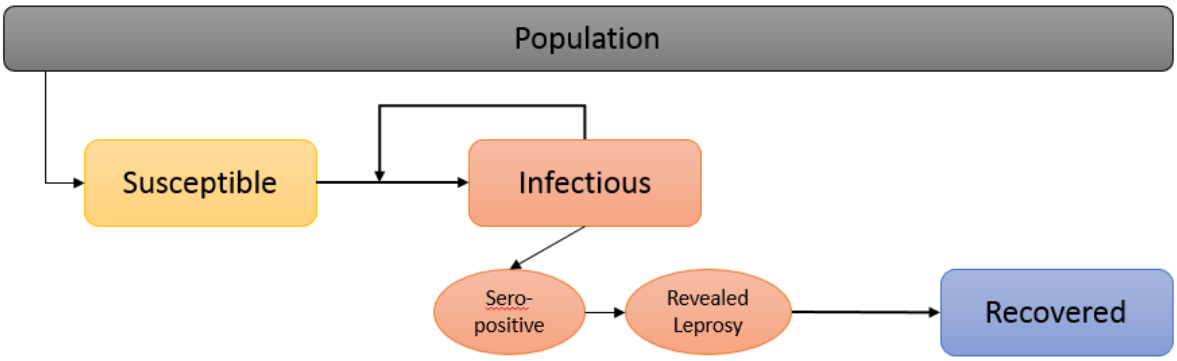


Figure 2.9: Prototype ABM (Holtrup et al., 2015) as SIR-model

Each infection chance consists of a multiplication of a risk and an intimacy-factor. The risk is calculated by counting the number of seropositive agents in each of the groups divided by the number of agents in this group. The intimacy-factor is dependent on the type of chance, gradually increasing from the world infection up unto the household and boat infection chances. By adding up the four infection chances at each time-step for each agent the total infection chance is obtained for each agent. At each time-step each individual has this chance to get infected with leprosy. When infected the agent becomes seropositive (Exposed), after which he has a certain probability to "reveal Leprosy". This is not a distinction between Exposed and Infectious as in the SEIR model, as all seropositive agents are considered to be equally infectious in the prototype ABM (Holtrup et al., 2015). As no distinction is made between Exposed and Infectious individuals in the prototype ABM, the SEIR-model is not applicable. The disease model used in the prototype ABM can, however, be represented as a SIR model (Figure 2.9). It should be noted that the effect of treatment strategies has not been incorporated in the model, but it is assumed that the idea was to make *Infectious* agents able to become *Recovered* in one way or the other.

The prototype model represents a first effort at an agent-based approach to modeling leprosy and its transmission in a population (Table 2.8). The model, however, does not incorporate interventions, or any treatment for that matter, yet. Furthermore, the explicitly spatial structure of the ABM, which is one of its great advantages over traditional modelling approaches, is not used to its full potential by Holtrup et al. (2015), as no neighbor relationships, and accompanying infection probabilities, are modeled.

Table 2.8 Prototype ABM (Holtrup et al, 2015)			
General Disease Model parameters			
Type of model		Agent-Based	
Topology of connectedness		Household structure, Social structure	
Level	**Risk Factor**	**Incorporation?**	**Description**
Individual	Susceptibility	-	The whole population is equally susceptible to leprosy
	Genetic factors	-	-
	Health factors	-	-
	Serological status	x	Only seropositive agents are infectious. Recovered agents are no longer infectious.
	BCG vaccination	-	-
	Chemoprophylaxis	-	-
Contact	Type of contact	x	World, Island, Household and Fishing contacts each acount for different probabilities of infection
	Infectiousness of contact	-	All seropositive agents are equally infectious
	Number of contacts	x	The probability of infection increases with the number of infected contacts
Macro	Leprosy Prevalence in population	x	The World Infection Chance is dependent on the leprosy prevalence in the study population
	Socio-economic factors	-	-
	Social Stigma	-	-
	Leprosy control programs	-	-

2.4.5 Inventory existing leprosy models

Over the years, models aimed at modeling leprosy diffusion dynamics have developed from deterministic compartmental disease models (Lechat et al., 1974, Meima et al., 2004) to an individual-based (Fischer et al., 2010) approach. The prototype ABM developed by Holtrup et al. (2015) represents a first effort at an Agent-Based model for leprosy, but does not use the full potential an ABM can offer. Furthermore, the prototype ABM does not incorporate heterogeneity in susceptibility or heterogeneity in infectiousness, which are considered to be essential factors in leprosy diffusion dynamics (Table 2.9) (Meima et al., 2004; Bakker, 2005a, Fischer et al., 2010). None of the models incorporate neighbor relationships, although research by van Beers et al. (1999) and the findings by Bakker et al. (2006) clearly indicates that neighbors of infectious leprosy patients have a significantly higher risk of contracting leprosy than non-contacts.

Table 2.9: Inventory Existing Leprosy Models

		Lechat's Model	SIMLEP	SIMCOLEP	Prototype ABM
General Disease Model Parameters					
Type of model		Compartmental	Compartmental	Individual-Based	Agent-Based
Topology of connectedness		General Population	General Population	General Population, Household Stucture	Household Structure, Social Structure
Level	**Leprosy Risk Factors**	**Incorporation in Model?**			
Individual	Susceptibility	-	x	x	-
	Genetic factors	-	-	x	-
	Health factors	-	-	-	-
	Serological status	x	x	x	x
	BCG vaccination	x	x	x	-
	Chemoprophylaxis	-	-	x	-
Contact	Type of contact	-	-	household contacts	world, island, household & fishing contacts
	Infectiousness of contact	x	x	x	-
	Number of contacts	x	x	x	x
Macro	Leprosy Prevalence in population	x	x	-	x
	Socio-economic factors	-	-	-	-
	Social Stigma	-	-	-	-
	Leprosy control programs	x	x	x	-

Of the existing leprosy models the SIMCOLEP-model (Fischer et al., 2010) offers the greatest potential, as it is the most complete model, i.e. as it incorporates the most known risk factors for infection with leprosy. It is however not an Agent-Based Approach, which is needed to include heterogeneity in susceptibility and contacts. LEPRASIM will therefore be a translation of the SIMCOLEP-model structure to an Agent-Based environment, with incorporation of different contact groups as in the prototype ABM (Holtrup et al., 2015).

3. Methodology

The methodology of this study is divided into six steps (Figure 3.1). The first step consists of the model development in the NetLogo environment based on the literature study presented in chapter 2 and a number of interviews with Dr. Mirjam Bakker (KIT). The resulting model "LEPRASIM" is presented using the ODD protocol (Grimm et al., 2010; chapter 4). The second step consists of an assessment of the model stability and a sensitivity analysis on this model. Using the outcomes of this sensitivity analysis the model is calibrated on observations made in the study area by Bakker et al. (2002; 2004; 2005b) in the third step. The behavior of the calibrated model is verified in the fourth step, and validated in the fifth step of this study. The sixth and final step consists of two experiments designed to assess the influence of the modeled activities on the re-emergence of leprosy in the study area and the long-term effects of the prevention strategies.

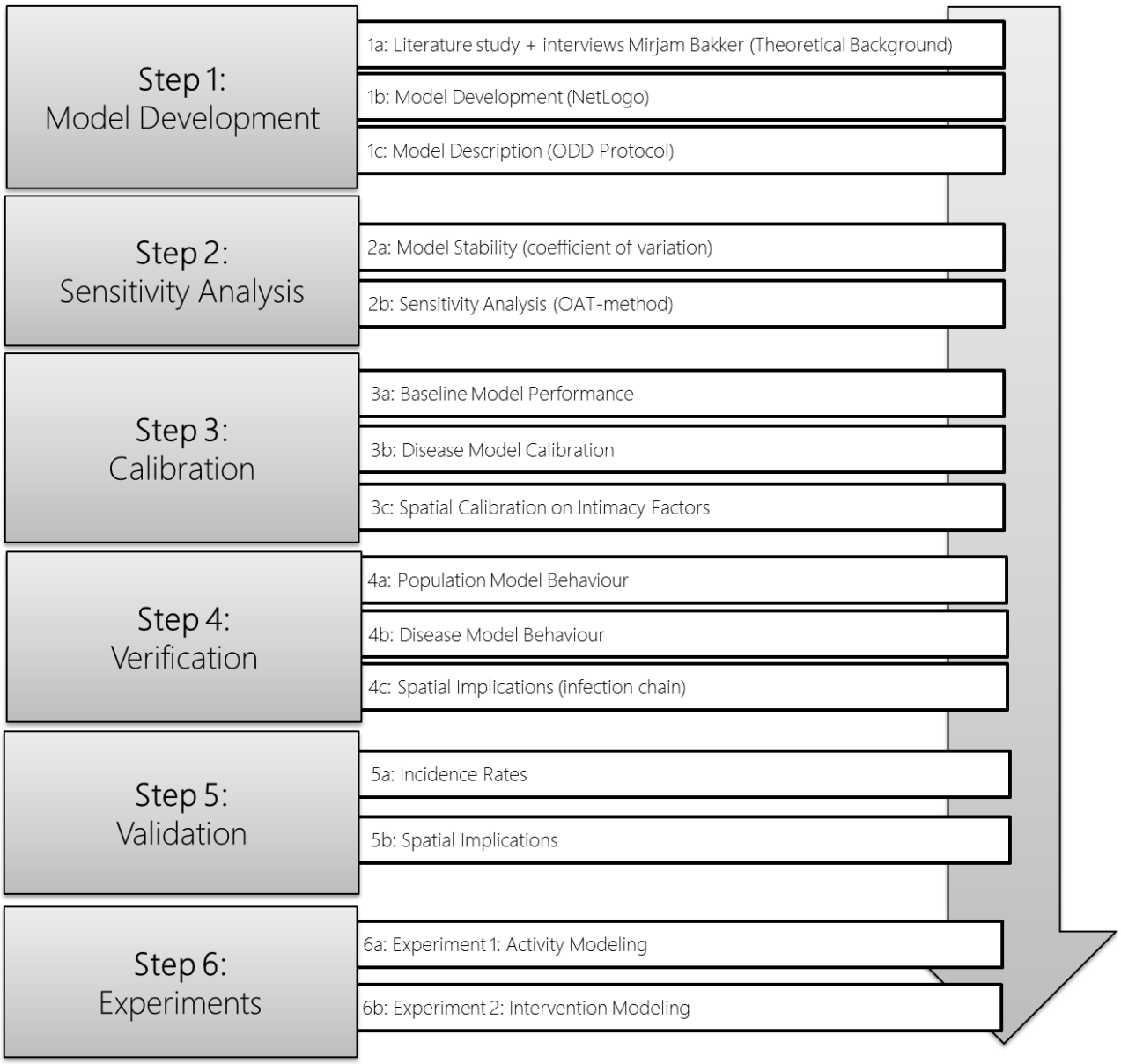


Figure 3.1: Methodology

3.1 Step 1: Model development

Through a literature review and a number of interviews with Dr. Mirjam Bakker (KIT) an insight into the known mechanisms causing the spread of leprosy was obtained. An inventory of risk factors for leprosy infection identified in other studies has been made, to provide an insight into individual characteristics relevant for the spread of leprosy, and thus for an ABM of this phenomenon. An exploration and comparison of existing leprosy simulation models was made. Lechat's leprosy model (Lechat, 1992), the SIMLEP model (Meima et al., 2004), the SIMCOLEP model (Fischer et al., 2010) and the prototype ABM (Holtrup et al., 2015) were compared on model structure and incorporation of the relevant risk factors, extending the work of Blok et al. (2015) in their review of the first three of these models. It should be noted that these models (except for the prototype ABM) are not ABMs. The identified common parameters/ assumptions thus need to be translated into this environment.

From this comparison of existing leprosy models a number of essential factors to be included in LEPRASIM are distilled and implemented in the NetLogo environment. Special attention is given to modelling the different contact groups (agent collectives) in the population as this is the essence of leprosy diffusion and the experiments performed in this study.

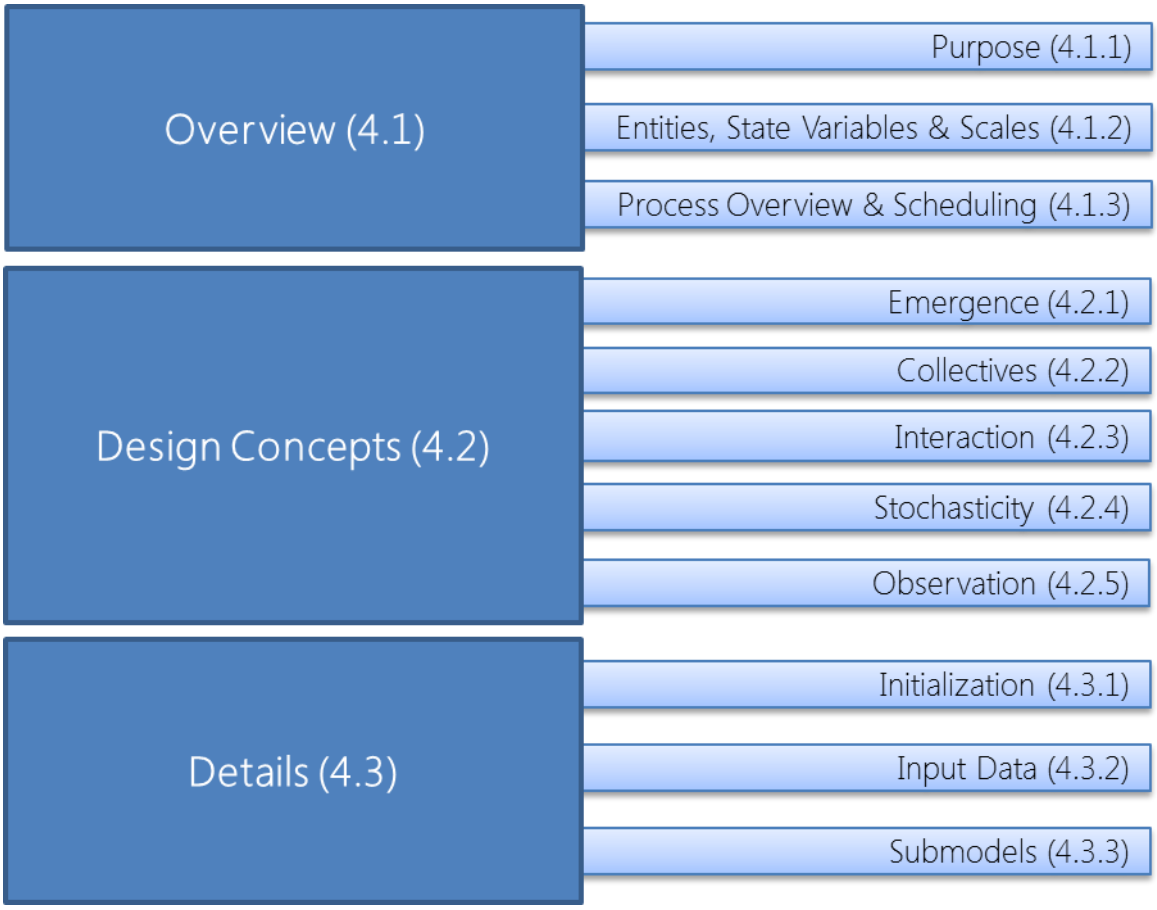


Figure 3.2: ODD-Protocol (Grimm et al., 2010)

The LEPRASIM model is described (in chapter 4) following the updated Overview, Design concepts & Details (ODD) Protocol (Grimm et al., 2010). This protocol has been developed as a response to criticism on ABM's being so poorly documented, that they could not be evaluated. The ODD protocol aims to provide a generic structure and standard format for documentation of ABMs, making these models easier to replicate and understand (Grimm et al., 2010). The ODD protocol consists of seven elements grouped in three categories (Figure 3.2). In the Overview section (4.2) the general purpose of the model, the entities used in the model, their variables and the scales of these variables are presented. In addition an overview of the model entities (UML2 class diagram), the modeled processes and the scheduling of these events is provided. The Design Concepts section (4.3) gives an overview of the emerging phenomena, agent collectives and interactions, the degree of stochasticity of the model and the method of observation. In the Details section (4.4) the initialization of the model and the input data used to drive the model are presented. In addition, the sub-models are explained in detail.

3.2 Step 2: Sensitivity Analysis

The second step of the methodology is the sensitivity analysis of LEPRASIM. The aim of is this step is to assess the internal validity of the model: does the model operate as intended? The internal validation of the model therefor starts with a "stability check" to determine at which point LEPRASIM becomes stable, i.e. to determine the minimum sample size of model-runs needed to produce a stable result. The methodology developed by Lorscheid et al. (2012) is employed. In this method a coefficient of variation is used to express the model's stability (Lee et al., 2015). The next step is the sensitivity analysis itself. This sensitivity analysis is performed for two reasons: to check whether the model shows the expected behavior and to determine the model's sensitivity to the different input parameters. The sensitivity analysis is performed using the one-parameter-at-a-time (OAT) –method (Hassani-Mahmooei & Parris, 2013). In this method, each input parameter is examined over a predefined range of values in isolation, as the remaining parameters are set at a constant baseline (Lee et al., 2015). As LEPRASIM is divided into a population, activity and a disease model (see section 4.1.2) the sensitivity analysis comprises of several steps. In the first step the sensitivity of the population model to the driving factors of the activity model is assessed. In the second step the internal sensitivity of the disease model is examined. In the third step the sensitivity of the disease model to variations in the population model's output parameters is assessed. In this way a complete sensitivity analysis off the LEPRASIM model is constructed.

3.3 Step 3: Calibration

In the third step (Figure 3.1) LEPRASIM is calibrated using a global calibration process. The goal is to assess the external validity of the model. By variation on a set of parameters identified in the sensitivity analysis the model's output is calibrated to match the observation's made by Bakker et al. (2002; 2006) in the study area. The calibration process consists of three steps. In the first step the baseline model performance is compared to the observations. In the second step the disease

model is calibrated to match the cumulative incidence and prevalence rate observed in the study area. In the third step the distribution of these infections over households and islands is calibrated using the "intimacy factors". In other words, a spatial calibration of LEPRASIM is performed.

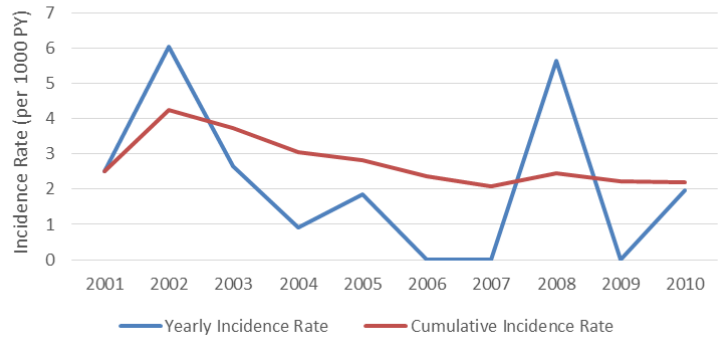
3.4 Step 4: Verification

By verifying the behavior of LEPRASIM a greater insight into the spatio-temporal diffusion dynamics of leprosy is obtained. The verification of the calibrated model consists of a number of steps. First, an assessment of both the population as the disease model's behavior is made by examination of a single model run. Next, an insight into the spatial implications of the disease model is obtained via an "infection-chain".

3.5 Step 5: Validation

By validating LEPRASIM the agent-based approach to the modeling of leprosy is proven to be possible. The model is validated via two sets of validation parameters. As the model has been calibrated on the situation in 2000, the simulation of the spread of leprosy will be validated firstly by comparing the model output cumulative incidence rates in the period after the intervention by Bakker et al. started (2000-2010) to the observations made by Bakker (2016) in the study area. The figures on the yearly incidence rates (per 1000 PY) for the control group in the period 2000-2010 were provided by Bakker (2016). From these figures a cumulative incidence rate (per 1000 PY) is calculated (Table 3.1). As can be seen from Figure 3.3 the yearly incidence rates show a rather large variation. This is due to the relatively small size of the population participating in the study and the effects of the performed intervention in the control group: the yearly screening of the population for signs of leprosy infection and subsequent treatment of detected patients with MDT.

Figure 3.3: Incidence Rate Control Group 2000-2010 (Bakker, 2016)



In addition, a number of spatial measures, not used in the calibration process, will be used for the validation of the spatial implications of the model. The first of these spatial measures are the Hazard Ratios for the period 2000-2003 on two conditions (Bakker et al., 2004): the household size and the contact group. In the study by Bakker et al. (2004) the status of participants on these two measures was identified in 2000. In the period 2000-2003 the cumulative incidence rates for each of these groups were measured, resulting in a relative risk per status in 2000, also known as a

hazard ratio. For the household size the categories "1-4", "5-7" and "greater than 7" are distinguished, for the contact status "no contact", "household contact" and "neighbour contact" are distinguished (Table 3.1).

The second spatial measure is the clustering of patients in clusters of contact groups in 2000. This is identified (Bakker et al., 2002) by the percentage of patients (symptomatic agents) and controls (non-symptomatic/ non-infected agents) living in a cluster with at least one other patient. These clusters are the "households" and the "households + direct neighbours". Where the hazard ratios are an indicator of the models behaviour after the intervention is started (in 2000), this measure is an indication of the models behaviour before it is.

The last measure of the spatial implications of the model is the clustering on the islands in 2000 (Bakker et al., 2002), which is measured by the minimum and maximum prevalence rate per island in comparison to the total prevalence rate (Table 3.1). In other words: the degree of variation in clustering of patients on islands.

Table 3.1 Validation Parameters

Cumulative Incidence Rate Control Group 2000-2010 (Bakker, 2016)										
Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
CIR (per 1000 PY)	2.52	4.25	3.73	3.05	2.82	2.38	2.07	2.44	2.21	2.2
Clustering in Contact Groups: Hazard Ratio 2000-2003 per status in 2000 on (Bakker et al., 2004):										
	Household Size				Contact Group					
Status	1-4	5-7	>7		no contact		neighbor		household	
CIR (2000-2003)	188 (104 - 3.40)	3.09 (2.02 - 4.74)	5.61(3.19 - 9.88)		2.88 (2.00 - 4.15)		3.31(149 - 7.38)		6.67 (3.00 - 14.9)	
Hazard Ratio	1	171(0.82 - 3.56)	3.47 (151 - 7.98)		1		152 (0.50 - 4.59)		3.29 (111 - 9.77)	
Clustering in Contact Groups: percentage of patients/ controls living in cluster with at least one other patient 2000 (Bakker et al., 2002)										
Cluster	Patients			Controls			Patients / Controls			
Household	32%			16%			2.07			
Household + Neighbors	52%			30%			175			
Clustering on Islands: Prevelance Rate 2000 (Bakker et al., 2002)										
Island	Tampaang	Sapuka	Sailus	Pelokang	Kembanglemari		All		Minimum	Maximum
Absolute PR	89	145	179	181	440		195		89	440
PR as percentage of total	45.6%	74.4%	918%	92.8%	225.6%		100.0%		45.6%	225.6%

3.6 Step 6: Experiments

To answer the second and third research questions two separate experiments are conducted. To answer the second research question - *which of the identified reasons for the unexpected re-emergence of leprosy in the blanket and contact group of the study by Bakker et al. (2006) best explains this phenomenon within LEPRASIM? -,* these reasons are explicitly modeled within LEPRASIM by adding activities to the model. The third research question - *at which contact group(s) of infectious individuals should a leprosy prevention strategy using rifampicin prophylaxis be aimed to be most effective? -* is answered by modeling these leprosy prevention strategies explicitly within LEPRASIM and examining their long-term effect on leprosy prevalence in the population.

3.6.1 Experiment 1: Activity Modeling

An experiment is conducted to explore the possible explanations identified by Bakker (2005b) for the unexpected re-emergence of leprosy in the contact and blanket group after the intervention in 2000. The possible explanations are:

- Leprosy is reintroduced on the islands through marriages between the islands
- Leprosy is reintroduced on the islands through contact of fishermen on fishing boats

To explore the feasibility of these potential explanations the activities have been explicitly modeled in LEPRASIM. The marriage activity is modeled by "turning on" marriage between the islands, meaning that in the altered model male agents can choose a potential partner from any of the islands, instead of only their own island. The chosen partner (and her children) then moves to the island of the male agent, leading to a flow of people between the islands.

Approximately 90% of the male population in the study area is a fisherman (Bakker, 2005a). The fishing activity is modeled by placing a random selection of 90% of the male agents on fishing boats twice a year for a period ranging from 1 to 3 months. In this period the agent can only effect other agents on the boat (when he is infectious himself), leading to an additional source of infections. When on a fishing trip, the male agent cannot infect household or neighbor contacts. 60 boats, each carrying 10 agents are deployed (leading to a total of 600 fishermen per fishing event), as not all agents go on "long" fishing trips on which they have contact with other agents twice a year, but do so irregularly (Bakker, 2005a).

Experiment 1 consists of two parts. In the first part the effect of the addition of the activities to LEPRASIM on the occurrence of leprosy in the entire study area is explored. In the second part of the experiment the magnitude of this effect is explored.

Experiment 1a: Effect of activity modeling on spatio-temporal diffusion of leprosy

In the first part of this experiment (1a) LEPRASIM is run for 50 iterations (for the period 1960-2010). In these model runs the entire population is treated as the control group, which means that no intervention is employed other than the yearly examination of the entire population on leprosy and treating those that are symptomatic. The effect of the two additional activities and the combination of the two on the cumulative incidence rates (2000-2010) and the development of the prevalence rate per island over time (2000-2010) is compared to the performance of LEPRASIM. The hypotheses are that each of the modeled activities:

1. Slows the decline of the total cumulative incidence rate in the period 2000-2010.
2. Increases the minimum prevalence rate per island in the period 2000-2010.
3. Increases the maximum prevalence rate per island in the period 2000-2010.

The first part of this experiment is thus designed to establish the influence the addition of the activities has on the development of leprosy in the total study area.

Experiment 1b: Reintroduction of leprosy through activity modeling

In the second part of the experiment (1b) the magnitude of this influence is explored by a close examination of one of the islands in the study area. In order to do so a hypothetical intervention in the year 2000 is modeled within LEPRASIM. This hypothetical intervention consists of changing the state of all agents on one of the islands (Sapuka Besar) to "not infected" in 2000, effectively

eradicating leprosy on that particular island. Hereafter, the model is run for 10 more years (2000-2010), in which leprosy can only be reintroduced on the islands via the modeled activities. By tracing the cumulative incidence rate on the island over this time period the effect of the addition of the activities is measured.

3.6.2 Experiment 2: Prevention Strategy Modeling

A second set of experiments is conducted to examine the long-term effect of different leprosy prevention strategies on the cumulative incidence and prevalence rates in the study area. The best fitting model identified in experiment 1, is adopted to include a set off prevention strategies. Within this adapted model, the prevention strategies are applied to the entire population of all islands. The modeled prevention strategies apply rifampicin prophylaxis to different contact groups of infected individuals. For the effect of rifampicin chemoprophylaxis on an individual level three possibilities are identified by Bakker (2005b): (1) it only delays the development of leprosy, (2) it prevents leprosy, but only has a temporal effect on transmission of the disease and (3) it prevents leprosy and also reduces the transmission of the disease. As the exact effect of the chemoprophylaxis on leprosy transmission is not known, the effect of the pre-emptive medication within the experiment is limited to the leprosy prevention effect (2) of the medicine. Within the experiment this implies that at the moment(s) off intervention all "asymptomatic" leprosy patients are cured ("not infected"). Each prevention strategy is aimed at a different sub-set of the population. The effect of the prevention strategies is measured via the reduction in incidence rates over time, relative to the control approach (the baseline model performance).

Experiment 2a: Effect of prevention strategies on spatio-temporal diffusion of leprosy

Within the first part of experiment 2 (2a) the prevention strategies of the study by Bakker et al. (2006) (see introduction), with an addition of the "household contact approach", are employed within LEPRASIM:

- Control Approach (CTR): no prevention strategy is employed
- Blanket Approach (BLA): prevention strategy aimed at entire population
- Contact Approach (CNT): prevention strategy aimed at household and neighbor contacts of symptomatic patients
- Household Contact Approach (HHC): prevention strategy aimed at household contacts of symptomatic patients

In each of these four approaches the prevention strategy is employed one time in the year 2000 (tick 485). The effect on the incidence rates over time, relative to the baseline model performance (control approach) is measured to examine the long-term effects of each of these prevention strategies.

Experiment 2b: Effectiveness of prevention strategies

In the second part of the experiment this relative reduction in incidence rate is related to the number of agents receiving the pre-emptive medication, in order to deduce the relative

effectiveness of each of the prevention strategies. As the aim of the experiment is to find the most effective prevention strategy, the prevention strategies are extended by addition of the following strategies:

- Extended Blanket Approach (EBB): Blanket approach, followed by a five-yearly repetition of this blanket approach
- Household Contact Extended Blanket Approach (EBH): Blanket approach, followed by the household contact approach at each detection
- Contact Extended Blanket Approach (EBC): Blanket approach, followed by the contact approach at each detection
- Extended Household Contact Approach (EHH): Household contact approach, followed by the household contact approach at each detection
- Extended Contact Approach (ECC): Contact approach, followed by the contact approach at each detection

The effectiveness of a prevention strategy is determined by two measures: (1) the reduction in cumulative incidence rate for the period 2000-2025 and (2) the relative reduction in the incidence rate over time (measured via a 5-yearly cumulative incidence rate) per recipient of the medication.

4. Model Description (ODD-Protocol)

4.1 Overview

4.1.1 Purpose

LEPRASIM is an agent-based disease model (ABDM) serving two purposes. Firstly, the model is designed to explore the spatio-temporal diffusion dynamics of leprosy in a population on a group of islands in the Flores Sea of Indonesia. The main purpose of the model is to test the effect of (inter-island) marriages and fishing activities on this phenomenon. Secondly, the model is used to test the effects of leprosy prevention strategies.

4.1.2 State variables and scales

LEPRASIM is divided into three sub-models: the population model, the activity model and the disease model. LEPRASIM contains one type of agent – a person, from now on referred to as an "Agent". The state variable of this agent is its "Leprosy Health Status". Agents can be susceptible, infected, or recovered from Leprosy, and when infected, agents can be in the symptomatic or a-symptomatic stage of either an MB or PB type of leprosy infection. This gives a set of seven possible values for the state variable: immune, susceptible, a-symptomatic PB, symptomatic PB, a-symptomatic MB, symptomatic MB and recovered (see section 4.3.3.3)

One time-step in the model represent a single month. The maximum length of simulations is 65 years (or 782 time-steps) ranging from 1960 to 2025. Within the model one cell represents an area of approximately 200 x 200 meters. The entities of the model are presented in Figure 4.1 using a UML2 Class Diagram.

Agents are grouped in four aggregations representing the set of other agents they have (close) interaction (contact) with. These groups are referred to as contact groups and include: household contacts, neighborhood contacts, fishing contacts and island contacts.

Household contacts are agents that belong to the same household and have the closest form of contact. Neighborhood contacts are considered to be less intensive. They are defined at the household level and all individuals belonging to the household inherit the same neighborhood contacts. Neighborhood contacts can only be formed between households on the same island. Both the household and the neighborhood contacts change over the run of a simulation due to the birth and death of agents (population model) and marriage and movement activities (activity model).

Fishing contacts are non-permanent, and short in duration. They are formed at the level of the individual agent (not the household) and can be inter-island contacts (involving agents from different islands). Island contacts are the least intensive kind of contacts. They are formed at the individual level. They can change over the run of the simulation due to inter-island marriages. The contacts groups of each agent are determined at each time-step. This process is described in section 4.1.3. A complete overview of all agent variables is provided in Table 4.1.

Table 4.1 LEPRASIM Agents: state variables		
Agent		
Sub-model	Variable	Description
Population Model	ID	The ID of the agent.
	age	The age of the agent.
	sex	The sex of the agent.
	life-expectancy	The life-expectancy of the agent. Determined by global variable *Life-expectancy at birth*.
Activity Model	age_to_marry	The age a male agent will start looking for a female partner. Applicability is determined by global variable *Marry Rate*.
	partner_ID	The ID of the partner of the agent.
	household_ID	The ID of the household the agent is residing in.
	island_ID	The ID of the island that the household of the agents is located on.
	fisherman?	A boolean indicating whether the male agent is a fisherman embarking on fishing trips. Determined by global variable *Percentage Fisherman*.
	boat_ID	The ID of the boat the agent is located on.
Disease Model	leprosy_health_status	The state variable of the agent.
	suceptible?	A boolean indicating whether the agent is susceptible to leprosy. Determined by global variable *Percentage Susceptible*.
	probability_of_infection (P_i)	The probability an uninfected agent is infected with leprosy, which is calculated each time-step based on interactions, the global variable *Infection Rate* and the various *intimacy factors*.
	genetic_leprosy_type	The type of leprosy the agent will develop when infected. The options are "PB" or "MB". Determined by global variable *Percentage genetic type MB*
	infected?	A boolean indicating whether the agent is infected with leprosy.
	stage_of_disease	The stage of disease an infected agent has. The options are "asymptomatic" or "symptomatic".
	time_since_infection	The number of ticks the agent has been infected.
	incubation_time	The time in ticks between the moment of infection and the transition from the asymptomatic to the symptomatic stage of leprosy. Determined by global variables *MB/ PB Incubation Time*.
	infectiousness (I)	The infectiousness of an infected agent, with genetic_leprosy_type MB. Ranging from 0 to 1
	time_until_selfhealing	The time in ticks between the transition to the symptomatic stage of disease and the transition to the recovered stage for symptomatic PB patients.
	recovered?	A boolean indicating whether the agent is recovered from leprosy. To these agents the global variable *Relapse Rate* applies.
Contact Groups: agent collectives		
Activity Model	Household Contacts (C_h)	Collective of agents which are part of the same household as an agent
	Neighbor Contacts (C_n)	Collective of agents which are part of households neighboring the household of an agent
	Fishing Contacts (C_f)	Collective of agents which are part of the same boat as an agent
	Island Contacts (C_i)	Collective of agents which are part of the same island as an agent

Besides the Agent described above, the model contains four environments: the islands, the sea, the households and the boats (Figure 4.1). The island environment consists of five separate islands, representing the study area. Each island has a number of inhabitants and infectious agents, changing over time. On each of the islands a number of "houses" are located that can be occupied by "household contact groups" consisting of individual agents. Houses are randomly distributed over the island space and can be occupied or empty. Each house can only provide a location to a single household.

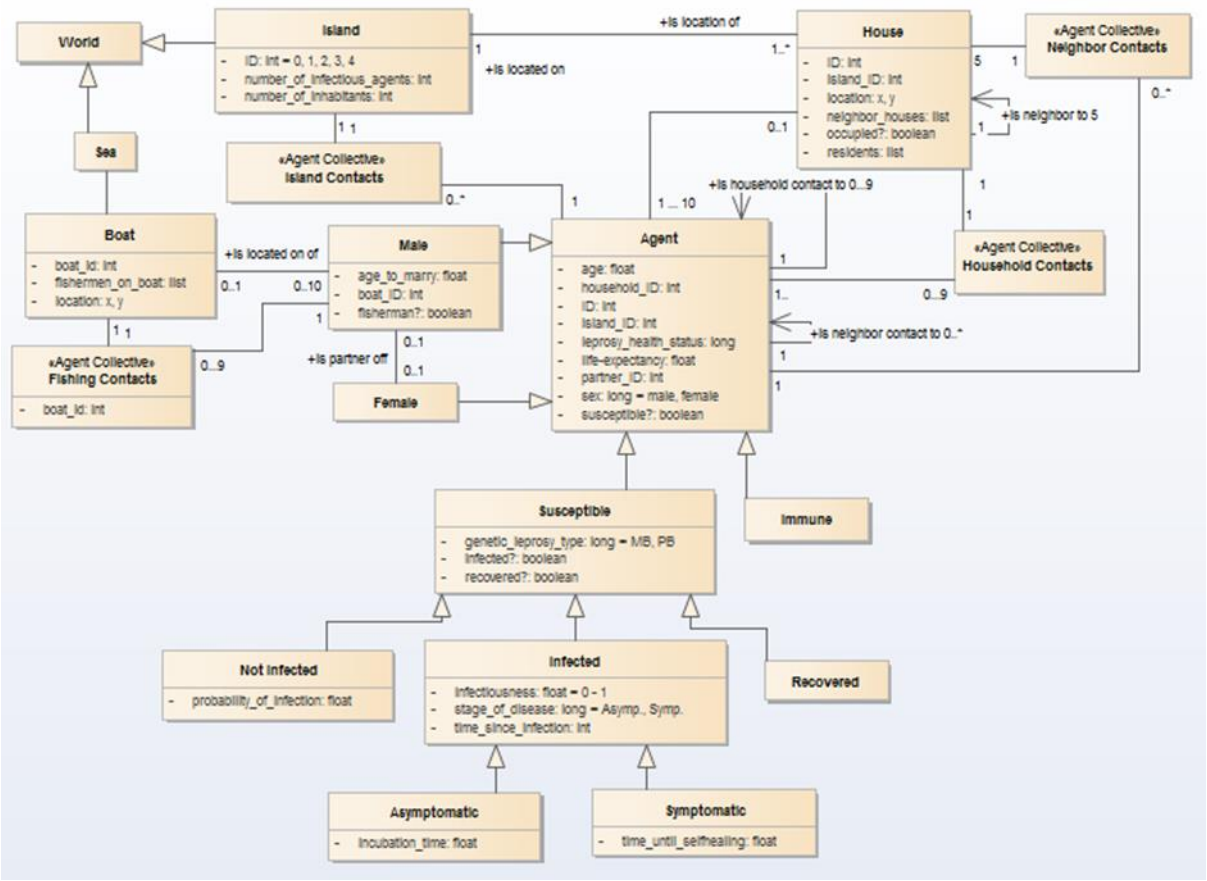


Figure 4.1: UML2 Class Diagram LEPRASIM

Besides the island environment, the model also contains a "sea" environment that represents the area between the islands. This sea environment defines the space in which the "boats" can be located. Boats are distributed randomly over this space at the beginning of the simulation and remain permanent during the simulation. Boats provide a location to fishermen and can be empty or contain many different fishermen during a simulation run. The exact location of the boats is irrelevant. A complete overview off all environment variables is provided in Table 4.2.

Table 4.2 LEPRASIM Environment: state variables		
Environment		
Environment	Variable	Description
House	ID	The ID of the house
	location	The location of the house (x, y).
	residents	A list of the agents currently residing in the household.
	island_ID	The ID of the island the house is located on.
	neighbor_houses	A list of the neighbor houses of the house, determining the neighbor contacts of the agents residing in the household.
	occupied?	A boolean indicating whether the house is occupied by agents, i.e. whether the list of residents is not empty
Island	ID	The ID of the island
	number_of_inhabitants	The number of agents residing in a house located on the island
	number_of_infectious_agents	The number of infectious agents residing in a house located on the island multiplied by their infectiousness
Boat	ID	The ID of the boat
	fishermen_on_boat	A list of the agents currently located on the boat
	location	The location of the boat (x, y).

The population model (through births and deaths) and activity model change (through marriage, movement and fishing activities) influence the distribution of agents over the households, islands and boats over time, and thus the collectives the agents are part of. In the disease model infections are modeled as a result of contact of an Agent with other agents in the collectives it is part of. The size of this population and its spatial distribution over the study area develops over time as described in section 4.1.3 (process overview). The method through which the spatial environment is initialized, the initial population is generated and the distribution off this population over the spatial environment at initialization is described in section 4.3.1 (initialization). The initialization of the disease model is described in section 4.3.1.4. The disease model itself is described in section 4.3.3.3.

4.1.3 Process Overview and Scheduling

LEPRASIM is divided into three sub-models: the population model, the activity model and the disease model (Figure 4.2). Within LEPRASIM a given population of individual agents grows over time. The population growth is controlled by the population model through births, death and aging events and is driven by external data (see section 4.3.2: Input). A detailed description of the population sub model is provided in section 4.3.3.1.

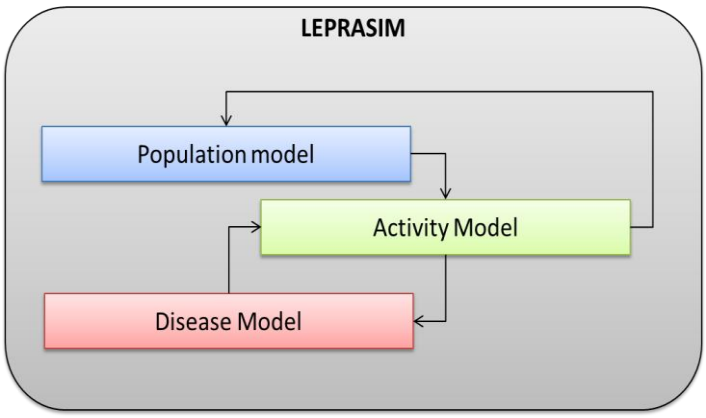


Figure 4.2: Abstract Process Overview

The activity model controls what activities are carried out by the agents. No behavior is defined at the level of the agent collectives: all activities are modelled at the level of the Agent. The activity model contains the following activities: marriage, movement and fishing (see section 4.3.3.2). Activities as controlled by the activity model affect the distribution of the population over the spatial environment. This changing distribution over the spatial environment gives each agent a changing set of contacts (agent collectives): household, neighbor, island and fishing contacts. These different contact groups serve as input for the disease model.

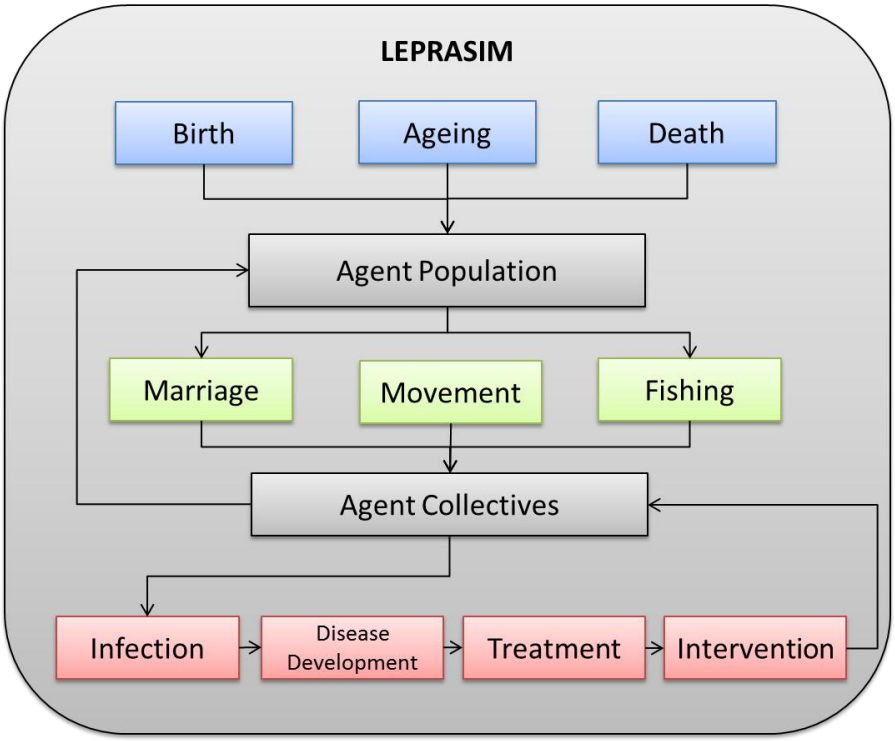


Figure 4.3: Detailed Process Overview

Within the disease model a probability of infection is calculated for and applied to each individual agent at each time-step. Each time-step disease development and treatment events are applied. An intervention event is applied yearly, starting in time-step 485 (the year 2000). The intervention event uses the spatial environment to determine the contact groups a prevention strategy is aimed at. The relationships between the events in the three sub-models are schematically shown in figure 4.3.

In LEPRASIM time is modeled using discrete one month time-steps. Each time-step event is scheduled in the order shown in Figure 4.4. The events attributed to the population model are run first, followed by the events attributed to the activity model, causing changes in the spatial distribution off the population. Next, the events of the disease model are run, starting with the infection event. In this infection event a probability of infection is calculated for and applied to each agent. The disease development and treatment events are deployed next. From time-step 485 this is followed by the intervention event. The last event in each time-step is the ageing of the population. For a detailed description of each of the events, see section 4.4.3.

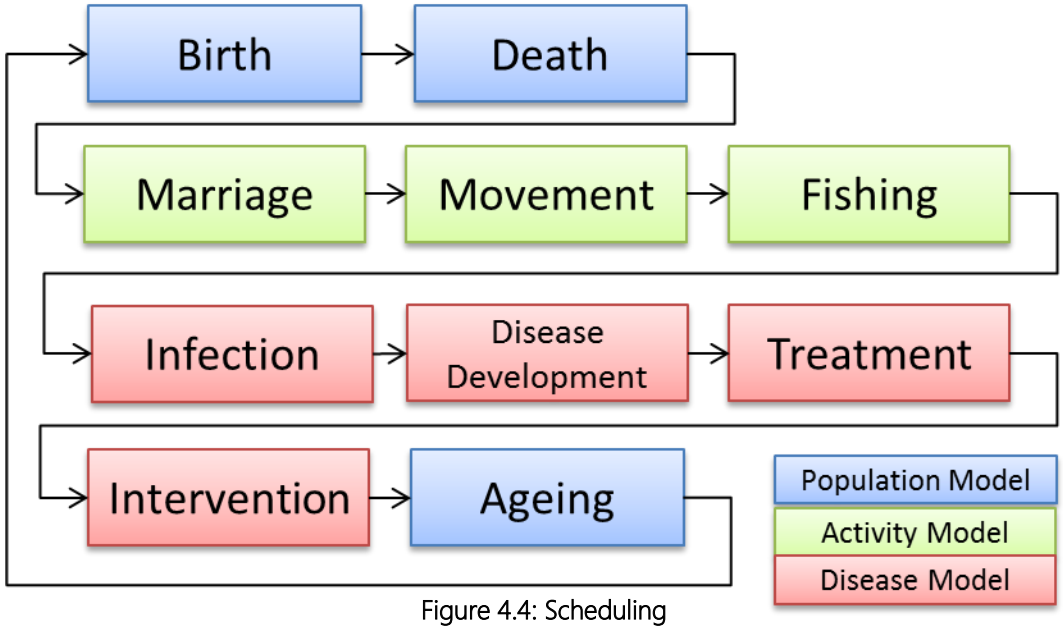


Figure 4.4: Scheduling

4.2 Design Concepts

4.2.1 Emergence

In addition two extensions on LEPRASIM are modeled, to explain two separate emergent phenomena. For the first experiment (section 3.5.1) the emergent phenomenon is the unexpected re-emergence of leprosy on an island, cleared of infection, as a result of a specific change in the activity model (2000-2010). For the second experiment (section 3.5.2) the emergent phenomenon is the change in the total cumulative incidence rate per recipient of rifampicin prophylaxis (2000-2025) as a result of a number of leprosy prevention strategies aimed at different agent collectives.

4.2.2 Collectives

Although the agent collectives in LEPRASIM do not show any behavior different from that of the individual agents, and are thus not agent collectives in the strict sense (Bonabeau, 2002), they are presented as such in this model description. Within LEPRASIM these collectives are modeled as separate entities, which means that the model can be easily extended to include the behavior of these entities in the future. Four types of agent collectives exist in LEPRASIM for each individual agent:

- Household contacts: Agents located in the same house.
- Neighborhood contacts: Agents located in a neighbor house.
- Fishing contacts: Agents located on the same boat.
- Island contacts: Agents located on the same island.

On each agent collective a specific *intimacy factor* (F) applies, which is a measure of the frequency of this contact. The intimacy factor thus represents the attribution of this agent collective to the agent's probability of infection. The island contacts have the lowest, neighborhood contacts a

slightly higher and household/ fishing contacts the highest intimacy factor (van Beers et al., 1999). Within LEPRASIM agents and agent collectives show no adaptive behavior, sensing or learning.

4.2.3 Interaction

Agents interact indirectly with each other in each of the sub-models of LEPRASIM. In the population model birth events occur "to" a female agent implying an interaction between this female agent and her child: the female agent gives birth to a child agent. Within the activity model, the marriage and movement event involve an interaction of male agents that, at a predefined age, "make" one of the female agents their partner. The consequence is that this female agent "moves" to the house of the male agent, becoming a part of his household contact group. In the disease model infections occur within the agent collectives, mimicking interactions, and resulting infections, within these contact groups.

Agents interact with the environment in a number of ways. Through the marriage and movement events (in the activity model) new households are formed and subsequently allocated to an unoccupied house (*occupied? = false)*. This changes the status of the house to true (*occupied?* = true). In the fishing event ten randomly selected male agents (with *fisherman?* = true) are allocated to each boat for a set number of ticks (three). This changes the location of these agents and the agents in their agent collectives. As the agent is now located on a boat, his household and neighbor collectives are emptied. The fishing contacts are those agents allocated to the same boat. At the end of the fishing event, the agents move back to the location of their house.

4.2.4 Stochasticity

As the actual route of leprosy transmissions (WHO, 2015a), the length of incubation times (Fine, 1982; Nordeen, 1985) and the population growth on the islands in the future are unknown LEPRASIM is a highly stochastic model, as is common in many Agent-Based Disease Models (Rodrigues et al., 2015; Crooks & Hailegiorgis, 2014; Linard et al., 2009) and is also the case in the SIMCOLEP-model (Fisher et al., 2011). LEPRASIM uses stochasticity to model birth events, marriage events, fishing events, infection events, incubation times, the probability of treatment and the occurrence off relapses. Birth events are modeled via a monthly chance to give birth (for women aged 18-40). Within the marriage event, a male agent chooses a female agent randomly from the population of his own island (or the entire population in experiment 1). In the fishing event a fixed number of fishermen are randomly selected from the population of fishermen to embark on this event. Each infection that occurs in LEPRASIM is a stochastic event, as for each agent the number of contacts and the degree/level of infectiousness of these contacts in each of the contact groups form an agent-specific probability of infection ranging from zero to one. The incubation times (for both PB and MB patients) are stochastically drawn from a normal distribution, as is the time until self-healing for PB patients. The probability of treatment and the relapse rate are determined in the environment, but applied to the agents in a stochastic manner.

4.2.5 Observation

The key observations are the prevalence rate in 2000 and the yearly (cumulative) incidence rates for the period 2000-2025. In addition the distribution of patients over households and

neighborhoods is observed via the cumulative incidence rate 2000-2003 per status (in 2000) on household size and contact with infected agents. The distribution of patients over the different islands is observed via the prevalence rate per island.

4.3 Details

4.3.1 Initialization

4.3.1.1 Population size

To initialize the population model an approximation of the population on the five islands in 1960, and an approximation of the distribution of this population over the households is needed. In order to generate the initial population the following assumption was made: *the percentage of Indonesians living on the five islands has not changed in the period 1960-2000.* As the population of the whole of Indonesia developed from approximately 89 to 209 million over this time-period (World Bank, 2015), while the population on the islands was only 4774 in 2000 (Bakker et al., 2002), this is considered a save assumption.

Table 4.3: Approximation initial population

Year	Population Indonesia (World Bank, 2015)	Population Islands Flores Sea (Bakker et al, 2005)
2000	208938698	4774
1960	88692697	**2027**

Using demographic data of the World Bank (2015) on the population size of Indonesia and the population count done on the five islands in 2000 by Bakker et al. (2002), an approximation of the initial population was obtained (Table 4.3): 2027 inhabitants. Using this approximation as an initial input the population sub-model is calibrated to match the number of inhabitants in 2000. This calibration process resulted in a lower initial population of 1779 inhabitants as an optimal input, as the input data used caused a slight overestimation (approximately 14%) of the population development on the islands (see section 4.3.2: input). The population is distributed over the five islands, based on the assumption that *the percentage of people living on each of the islands has remained the same over time* (Table 4.4).

Table 4.4: Initial Population per Island

	2000 (Bakker et al, 2005)	Percentage	Initialization / 1960
Total Inhabitants	**4774**	**100%**	**1779**
Sapuka	2068	43.32%	**771**
Sailus	1451	30.39%	**541**
Kembanglemari	637	13.34%	**237**
Pelokang	393	8.23%	**146**
Tampaang	225	4.71%	**84**

Within the population sub-model the agents have three attributes: age, sex and life-expectancy. World Bank data (2015) on the spread of the Indonesian population across sex and age in 1960 showed that 48% of all Indonesians were female, and 40% were aged 15 or younger. The start

population in the model is generated using these figures. Approximately 48% of all agents are assigned a "female" sex; the rest is assigned the "male" sex. Approximately 40% of agents receive a random age between 0 and 15; the rest receives a random age between 16 and 59 (the life-expectancy at that time, see section 4.4.2: Input).

4.3.1.2 Spatial environment

A spatial environment is an essential part of any Agent-Based Model (Macal & North, 2010), as this environment shapes the conditions within which the agents can interact. In LEPRASIM the spatial environment consists of a representation of the group of five islands in the Flores Sea within a NetLogo environment. As the exact geographical location of the islands relative to each other has no impact on the model's behavior, the location of the islands has been altered to be more compact (using QGIS). In this way, the spatial resolution of the model is increased, as less "empty" area (sea) needs to be part of the spatial structure (Figure 4.5; actual location of the islands in Figure 1.1).

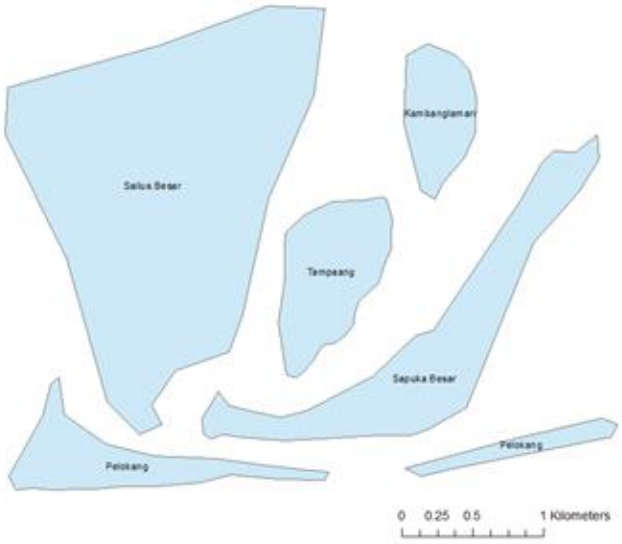


Figure 4.5: Model representation of study area

Per island the location of the houses was randomly generated using QGIS. The absolute location of the houses was not used, as no data on these locations was available. As no population is present (yet), all houses are empty (*occupied?* = false). The next step is to assign the five nearest houses of each house as neighbor houses to this house (*neighbor_houses)* within QGIS via a k-nearest-neighbor algorithm. By measuring the distance of each house to each other house on the same island, and subsequently assigning the five nearest houses to the house a list of (five) neighbor houses is created. This list is implemented in the NetLogo environment via "links", which are separate entities connecting two agents (NetLogo, 2016). As the neighbor relationships are defined using a k-nearest-neighbor algorithm, the absolute location of the houses has no influence on the dynamics of LEPRASIM. As the absolute location of the boats in the model

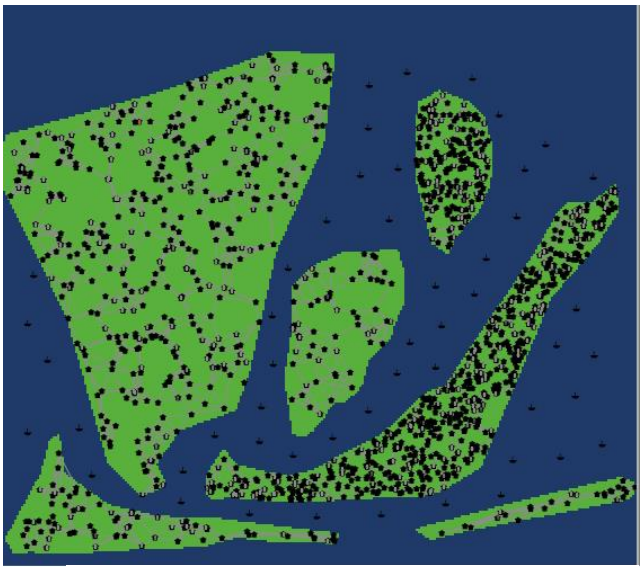


Figure 4.6: NetLogo representation of study

does not influence the model's behavior a fixed number of boats (50) were randomly distributed over the sea between the islands.

The model's appearance within the NetLogo environment is shown in Figure 4.6. The bottom two islands are considered to be one island in the model, as in the study by Bakker et al. (2006). Grey houses are occupied by agents, black houses are not. Neighbor relations between the houses are determined at the start of the model and are depicted with grey links between the houses (Figure 4.6). The neighbor houses of a house are thus not necessarily occupied. As the model develops over time, an increasing amount of the houses becomes occupied (by marrying individuals), mimicking the process of new houses being built on the islands and the intensification of contact on the islands.

4.3.1.3 Distribution of population over households

For the initialization of the household structure a seven-step-algorithm was developed, based on the algorithm developed by Gargiulo et al. (2010) (Figure 4.7). The main assumption used for this algorithm is a *maximum household size*. By making this assumption, an algorithm can be employed which ensures a constant initial distribution of the population over the household sizes for each model run. The maximum household size was set at ten. Although bigger household sizes do exist in the study area, the additional size (above a size of ten) of these households has no significant effect on the chance of leprosy infection (Bakker et al., 2006).

The first step of the algorithm is to approximate the future number of households per island in 2025 based on an approximation of the population in 2025 and an approximation of the average household size in 2025. The population size in 2025 was approximated by extrapolating the trend in population growth for Indonesia in the period 1960-2015 (World Bank, 2015) until 2025. This was done by performing 100 iterations of the population growth sub-model, without a distribution over the households, resulting in an expected population of approximately 5050 people on the islands in 2025.

The average household size in the study area was 4.29 in 2000 (Bakker et al., 2004). An approximation of the average household size in 1960 was obtained by taking the average household size of Bangladesh in 1960: 5.21 (Fisher et al., 2010). As Bangladesh had a similar economic status as Indonesia in 1960, and no data on the average household size in Indonesia in 1960 is available, this was used as an initial approximation. This gives a yearly decline of 0.023 in average household size for the period 1960-2000. Extending this trend to 2025 gives an expected average household size of 3.71 in 2025. By dividing the expected population by this expected average household size the needed number of houses in 2025 was obtained: 1361. This number of houses was divided over the islands based on the number of houses observed by Bakker et al. (2002) on each of the island in 2000. Of the houses in 2000 a certain percentage was located on each of the islands. In LEPRASIM this percentage of the 1361 houses is allocated to each of the islands.

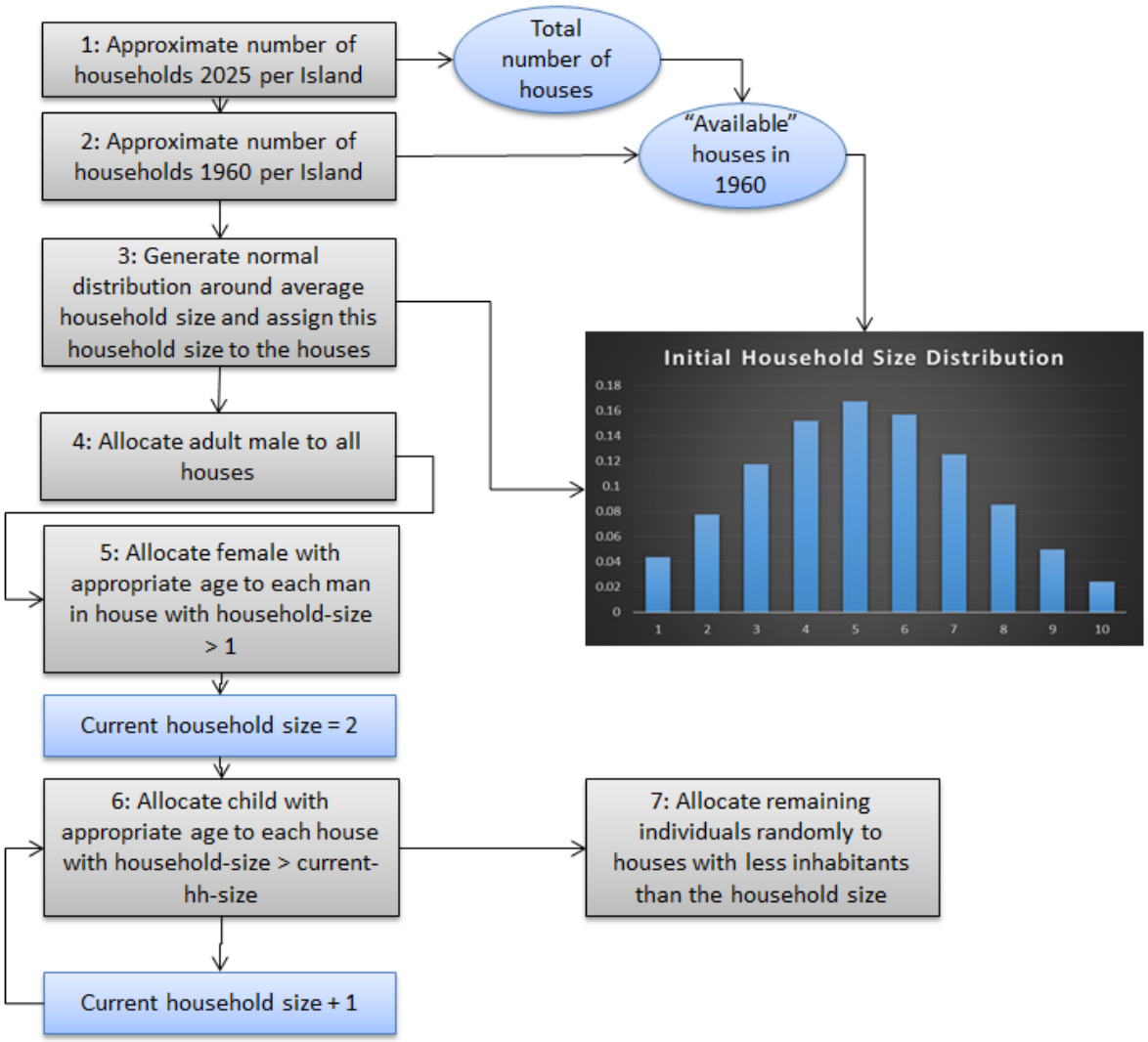


Figure 4.7: Algorithm for initial household size distribution

For the second step of the algorithm the needed number of household in 1960 is approximated. This was done by dividing the initial number of people on each of the islands by the average household size in 1960 (Fischer et al., 2010). Per island this number of houses was assigned "available" for allocation of agents. In the third step of the algorithm a normal distribution with a standard deviation of 2 around the average household size (5.21) was created, as was done in the initialization process of the SIMCOLEP model (Fischer et al., 2010). This normal distribution was applied to the number of houses on each of the islands, which means that for each model run the initial distribution off the houses over the household sizes is identical. The input table for the household size distribution is depicted (for all of the islands) in Table 4.5.

Table 4.5: Initial Household Size Distribution

Household Size	Count
1	15
2	26
3	40
4	52
5	57
6	52
7	44
8	27
9	17
10	11

With an approximation of the initial number of houses and people per island and a distribution off the houses over the household sizes the fourth step op the algorithm commences: allocation of

one adult male agent (age > 18 and sex = male) to each of the houses. This means all one-person households are "filled" with an adult male agent (*occupied?* = true). The fifth step of the algorithm allocates one adult female (age > 18 and sex = female) of an appropriate age (age of male +/- 5) to all male agents not residing in a one-person household, and sets the agents as each other's "partner". In the sixth step of the algorithm children (age < 18) with an appropriate age (age <= age of female in house - 15) are iteratively allocated to the houses per household size. In the first round one child is searched for all households with a household size greater than 2. In the next round a child is searched for all households with a household size greater than 3. In case the algorithm cannot find a child with an appropriate age, no individual is allocated to the house. This process is repeated until the household size is 9. In the final step of the algorithm all remaining individuals are allocated one by one at random to those houses, whose household size exceeds the number of individuals currently allocated to the house. This process is repeated until all 1779 people are allocated and thus all houses are "filled" according to the initial household size distribution.

4.3.1.4 Initialization of Disease Model

For an overview of the global variables used in LEPRASIM, see table (4.X). At initialization (1960, tick = 0) the global variable *percentage susceptible* is used to determine which agents are susceptible to leprosy infections (*susceptible?* = true). For these agents the global variable *percentage genetic type MB* is used to assign that percentage of susceptible agents the *genetic_leprosy_type* "MB". The rest of the susceptible agents receive *genetic_leprosy_type* "PB". Furthermore, the "Leprosy Health Status" of one Agent with *genetic_leprosy_type* = MB and an *age* of 20 is set to "symptomatic MB", representing the introduction of leprosy on the islands in 1960, as identified by Bakker (2005a). This agent is thus infected (*infected?* = true), has *genetic_leprosy_type* "MB", and has an *infectiousness* of 1 (the maximum). Furthermore, for this agent the *time_until_detection* is set, using the *probability of treatment* and *detection delay* variables (for more information: section 4.3.3.3).

4.3.2 Input

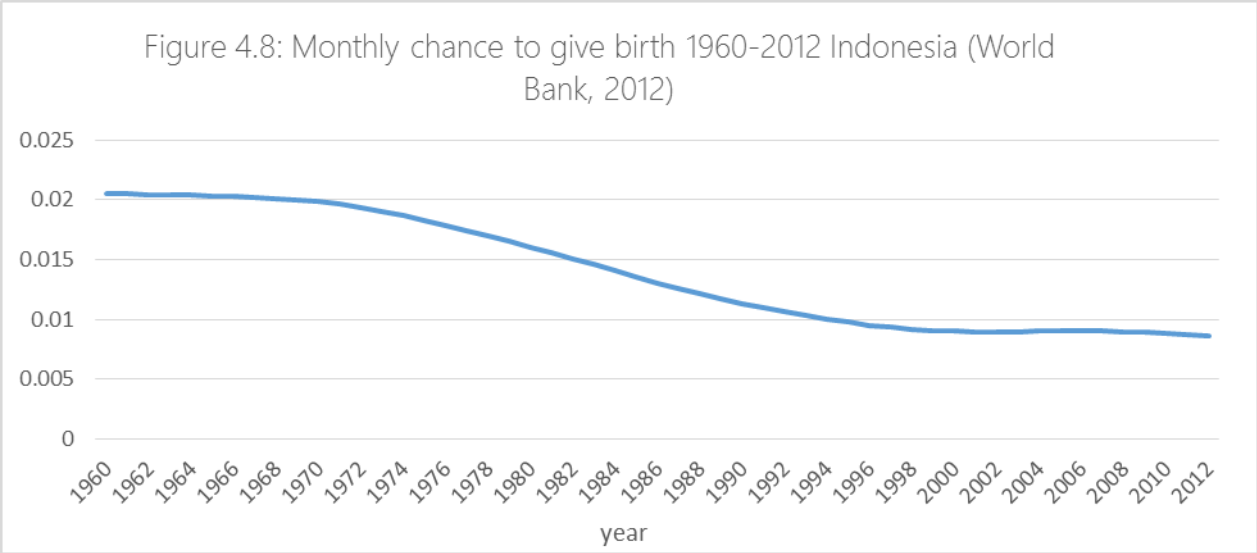
Within LEPRASIM only the population sub-model is driven by input data. The general equation used to model the population growth is the following:

(1) Population (t+1) = Population (t) + Births – Deaths

This equation means that during each time-step the change in population size is determined by death and birth events applied to each agent. The input data for these events is World Bank data for Indonesia in the period 1960-2014 (World Bank, 2015) on the fertility rate, child mortality rate and life-expectancy at birth. It is assumed that the population of the islands has followed these Indonesian demographics over time. The following sections describe the method used to employ this data in the LEPRASIM model.

4.3.2.1 Birth rate

A fertility rate represents the average number of children a woman gets in her lifetime. Yearly data on fertility rates for Indonesia (World Bank, 2015) provide the input data for the birth events. In order to translate this fertility rate figures to a "monthly chance to give birth", which is needed for the model structure, it is assumed that only women aged 18-40 give birth. By dividing the yearly fertility rates by the fertile months of these women (264 months) a monthly chance to give birth, changing over time, was obtained. A plot of the data shows that this chance has decreased over time, but not in a linear fashion (Figure 4.8).



Using the least square method a trend line was fitted through the data (2nd order polynomial, R^2= 0.9976), which was used as the birth rate for the model. The chance a woman aged between 18 and 40 gets a child is calculated at each time-step using the following equation:

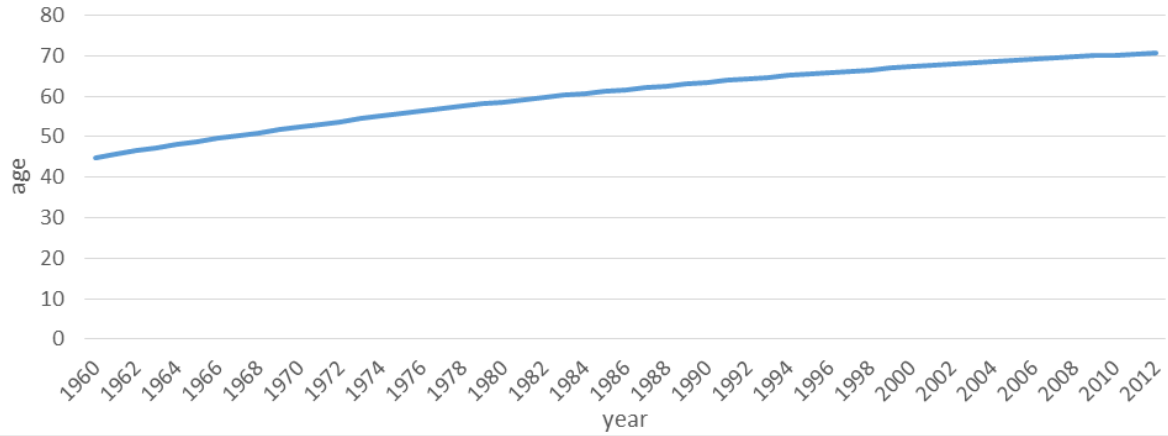
$$\mathbf{(2)\ B = 3 * 10^{-6} * x^2 - 4 * 10^{-4} * x + 0.0227}$$

Where B is fertility rate; x is the year the model is currently in.

4.3.2.2 Death rate

A combination of data on the average life-expectancy at birth and the child mortality rate for Indonesia for the same time-period (World Bank, 2012) was used to model deaths in LEPRASIM. The data on life-expectancy at birth (World Bank, 2012) is shown in Figure 4.9.

Figure 4.9 Average life-expectancy at birth Indonesia 1960-2012 (World Bank, 2012)



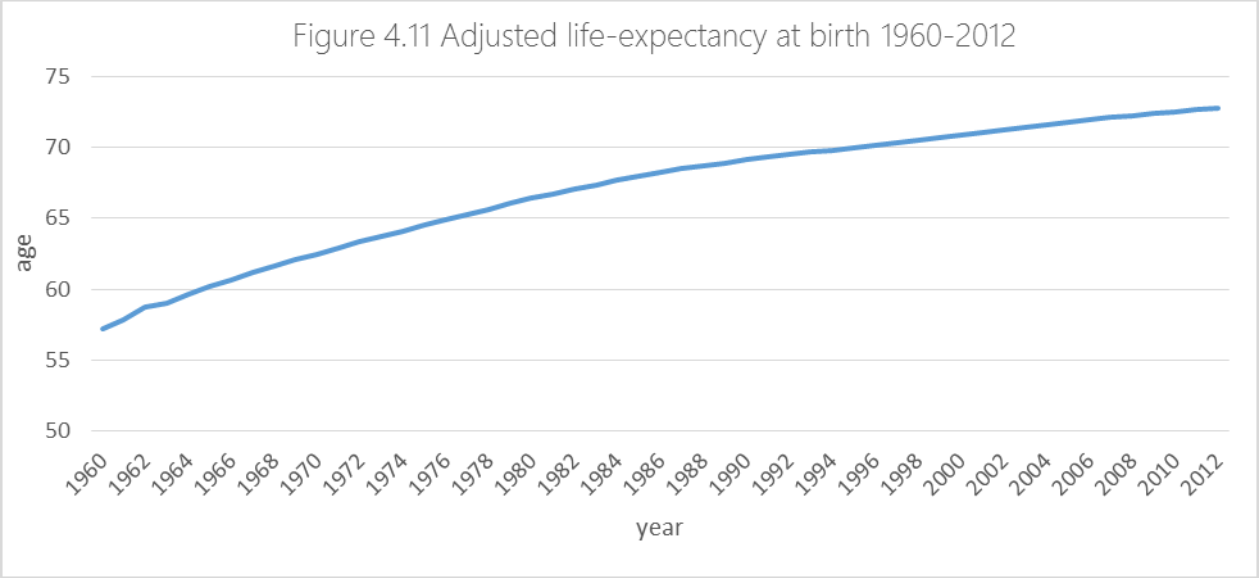
Dependent on the year the model is in during a run, new-born agents are assigned the life-expectancy valid at that moment in time. As the model progresses an agent dies (and is removed from the simulation) when his age exceeds this life-expectancy. At the start of the model all agents are given the life-expectancy valid in 1960.

Besides the life expectancy at birth another factor needs to be taken into account: the child-mortality rate. A high child mortality rate has a significant effect on population growth, as these children will not "produce" any children as time passes by in the case that they die as an infant. The child mortality rate for children aged below 5 for Indonesia is shown in Figure 4.10 (World Bank, 2012). Again, a 2nd order polynomial was fitted through the data using the least square method.

Figure 4.10 Child Mortality Rate Indonesia 1960-2012 (World Bank, 2012)

The mortality rate for children below the age of 5 was implemented in the model at the moment of birth. When a new agent is born, the child mortality rate valid at that point in time determines whether the child survives or not. Incorporation of the child mortality however has an effect on the life expectancy for agents passing the age of five. Of the child mortality approximately half of the infants dies before the age of one (World Bank, 2015), making the approximate life-expectancy for the children, dying before the age of five, two. The life-expectancy for agents, not dying as a consequence of child mortality, was adjusted for this child mortality by using the following equation:

(3) Adjusted life expectancy = (life expectancy – (child mortality rate * 2) / (1- child mortality rate)



The resulting *adjusted life expectancy* was plotted, and a second order polynomial was fitted through this data using the least square method. The resulting graph can be seen in Figure 4.911. This gives the following two equations for the adjusted life-expectancy at birth and child mortality rate respectively:

(4) $\boldsymbol{L = -4.6 * 10^{-3} * x^2 + 0.5203 * x + 57.631}$

(5) $\boldsymbol{C = 4 * 10^{-5} * x^2 - 6 * 10^{-3} * x + 0.2241}$

Where L is (adjusted) Life-expectancy at birth; C is child mortality rate

4.3.3 Submodels

An overview of the global variables used in each of LEPRASIM's sub-models is provided in table 4.6. Per global variable the symbol used in this model description and a short description of the global variable is provided. In this section the inner workings of each of the sub-models is explained.

Table 4.6 LEPRASIM: global variables			
Sub-model	Symbol	Variable	Description
Population Model	B	Fertility Rate	The monthly chance women aged 18-40 give birth
	C	Child Mortality Rate	The chance a child dies before the age of 5
	L	Life-expectancy at birth	The life-expectancy at birth
Activity Model	M_r	Marry Rate	The percentage of male agents which will search a female partner
	M_o	Percentage Own House at Marriage	The percentage of marrying agents which will form a new household collective at marriage
	F	Percentage Fishermen	The percentage of male agents who are a fisherman, embarking on fishing events.
Disease Model	G	Global Infection Rate	The global infection rate, indicating the chance of infeciton from an infectious agent to a susceptible agent at interaction.
	S	Percentage Suceptible	The percentage of the population susceptible to leprosy infections.
	M_b	Percentage genetic type MB	The percentage of susceptible agents which will develop MB leprosy after infection. The rest develops PB leprosy.
	P_t	Probability of Treatment	The probability a symptomatic agent transitions to the "recovered" stage. Applied at transition from asymptomatic to symptomatic stage.
	D	Detection Delay	The time-delay in ticks between the moment of transition to the symptomatic stage and the transition to the recoverd stage, in case probability of treatment applies.
	R	Relapse Rate	The probability an agent in the recovered stage transitions back to the symptomatic stage. Applied yearly to all recovered agents.
	T_{mh}	MB incubation Time	The time in ticks between the moment of infection and the transition from the asymptomatic to the symptomatic stage of leprosy for MB patients.
	T_{pb}	PB incubation time	The time in ticks between the moment of infection and the transition from the asymptomatic to the symptomatic stage of leprosy for PB patients.
	F_h	Household/ Boat Intimacy Factor	The relative risk of infection for household/ fishing contacts of an infected agent.
	F_n	Neighbor Intimacy Factor	The relative risk of infection for neighbor contacts of an infected agent.
	F_i	Island Intimacy Factor	The relative risk of infection per island. Determined by multiplication of the island intimacy factor with the percentage of the island population infected with leprosy.
	-	Prevention Strategy	The applied prevention strategy aimed at different contact groups of infected agents: CTR, CNT, BLA, HHC, EHH, ECC, EBB, EBH & EBC

4.3.3.1 Population model

The population growth is modeled using input data (World Bank, 2012, section 4.3.2), via equation (1). As no data on migration to, from and between the islands is available, the net migration has not been taken into account. The data driving the population growth within the population model has been described in section 4.4.2 (Input). In this section the events within the population model (Figure 4.4) will be described in detail.

Birth events

Within LEPRASIM, a birth event follows three steps (Figure 4.12). In the first step all "potential mothers" are located: agents with a "female" sex and an age between 18 and 40. In the second step the *fertility rate* (B; the monthly chance to give birth) is applied to these agents. If the probability applies the third step of the birth event takes effect: the application of the *child mortality rate* (C). If the child mortality rate does not apply, a birth event takes place: a new agent is placed in the household of the mother. The environmental variables *fertility rate* (B) and *child mortality rate* (C) are determined via equation (2) and (5). When an agent is born in the model, he first receives the life-expectancy valid at that moment in time. Next, he can die as a result of child mortality. If not, he lives until his age exceeds this adjusted life-expectancy.

Figure 4.12: Birth Event

Death events

Each time-step all agents are asked whether their age exceed their life-expectancy. This life-expectancy is set at initialization/ birth, and is determined by the environmental variable *life-expectancy at birth* (L), which increases over time via equation (4). If the age of an agent exceeds his life-expectancy a death event takes place: the agent is removed from the simulation (Figure 4.13).

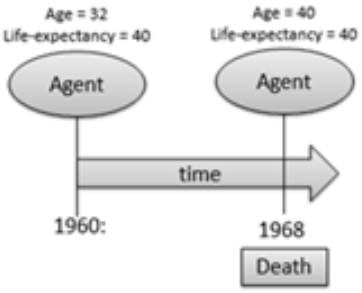


Figure 4.13: Death Event

Ageing event

At the end of each time-step the age of all agents in the model is increased by 1/12 (0.08333) years (or one month).

4.3.3.2 Activity model

Marriage and movement events

In LEPRASIM, as in the SIMCOLEP-model (Fisher et al., 2010), movement events are triggered by marriage events (Figure 4.14). To simulate marriage the environmental variable *marry-rate* (M_r) is used. At birth/ initialization this parameter determines which percentage of male agents will marry during the simulation. A marry-rate of 0.80 implies that 80% of male agents will search for a suitable partner during the simulation run. To these agents an *age-to-marry* (a random number between 18 and 23) is applied. When the age of the agent exceeds this *age-to-marry* he will search for a suitable partner. A suitable partner is a female partner located on the same island and with an age within a 5-year range of the age of the searching agent. In the extensions made on LEPRASIM male agents can marry female agents residing on any of the islands if inter-island marriages are part of the experiment.

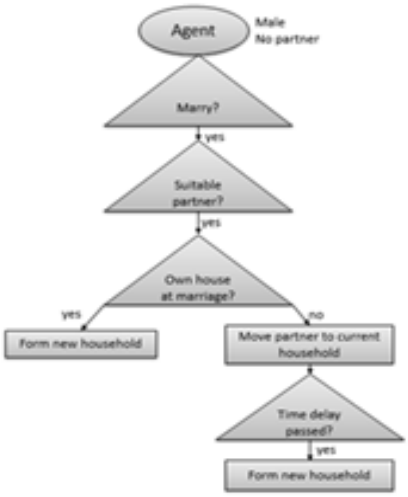


Figure 4.14: Marriage and Movement Events

In case a suitable partner is found, the environmental variable *percentage own house at marriage* (M_o)determines whether a new household is formed at marriage, or a time-delay between the marriage and movement event is set. This time-delay is normally distributed around an average of 10 years, with a standard deviation of 2 years, as is done in the SIMCOLEP-model (Fisher et al., 2010). After the time-delay has passed a new household is formed: a movement event is triggered. A new household is formed by moving the agent, his partner and their children to an unoccupied household on the same island where the agent resides (Figure 4.14).

Fishing events

In experiment 1 and 2 a fishing activity is added to LEPRASIM. The environmental parameter *percentage fisherman* (F) determines which percentage of the adult male population is a fisherman (*fisherman?* = true). Starting in 2000, each half year (or 6 ticks) a fishing event is triggered: a random selection of 10 fishermen is placed on each of the boats, leading to a total of 500 fishermen. At this fishing event the time the agent spends on the boat is set randomly between 1 and 3 time-steps. After this time has passed, the agent returns to his household. During a fishing event, the fishermen can only be infected by their fishing contacts: the agents located on the same boat.

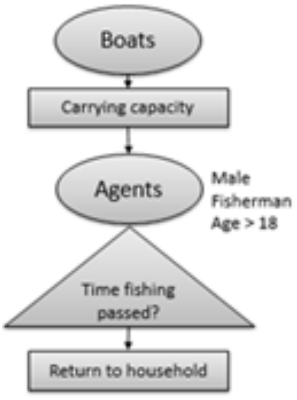


Figure 4.15: Fishing Event

4.3.3.3 Disease model

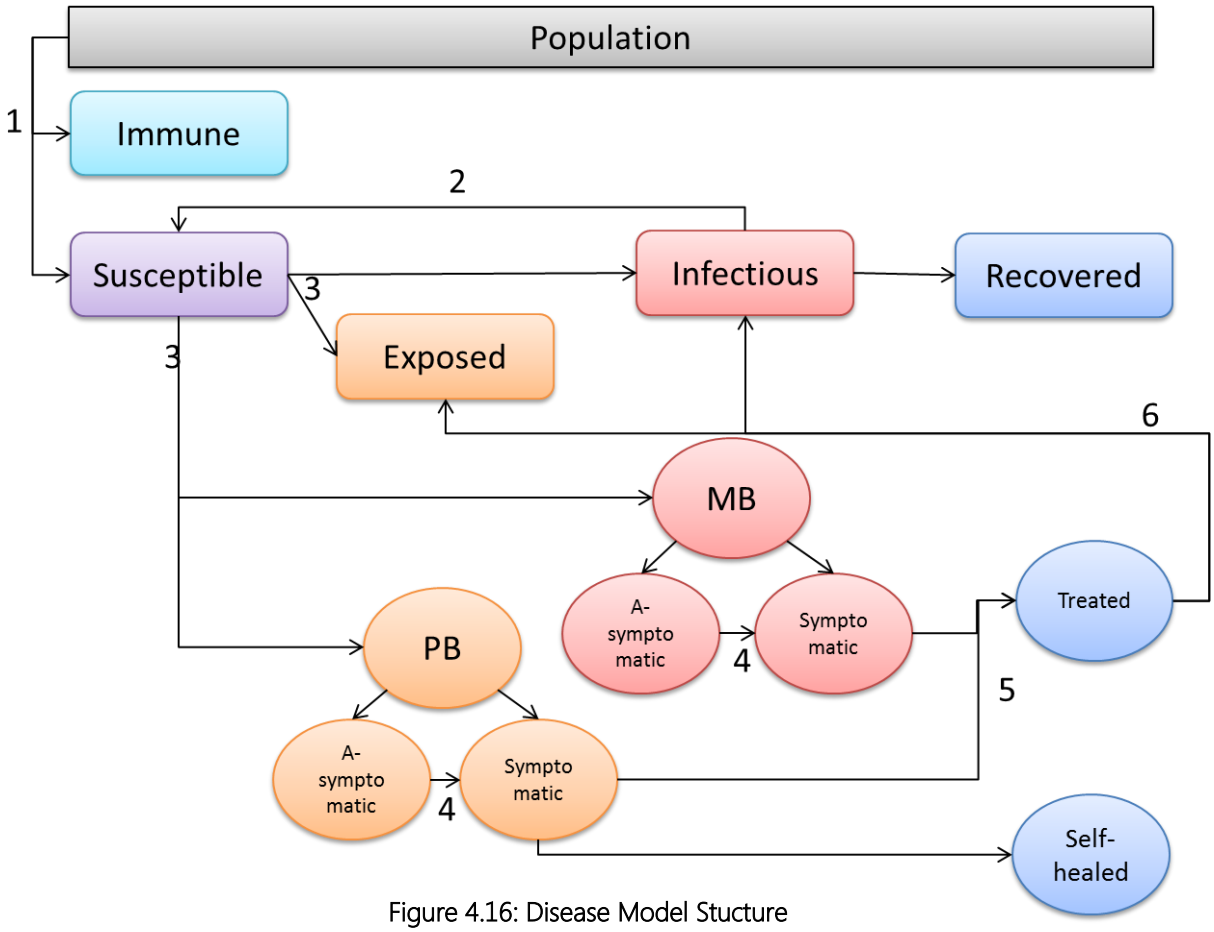


Figure 4.16: Disease Model Stucture

The disease model uses a SEIR model structure (Figure 4.16). The disease model operates via four events: infection, disease development, treatment and intervention events. The environmental variable *percentage susceptible* (S) determines the percentage of the population susceptible to leprosy (1 in Figure 4.16). The transition from the susceptible to the exposed/ infectious stage of leprosy is modeled via the infection events (2 in Figure 4.16), which are explained in more detail in the infection event. The environmental variable *percentage genetic type MB* (M_b) determines whether an individual enters the PB (exposed) or MB (infectious) disease track at the moment of infection (3). Separate incubation periods (*MB incubation time* (T_{mb}) *& PB incubation time* (T_{pb})), determining the latency period between the asymptomatic and symptomatic stage of the disease, apply to each track (4). After the incubation period, a self-healing mechanism is employed (agent-variable *time-until-self-healing)* within the PB track, while MB patients are considered to be chronic and thus not self-heal over time.

Via the treatment events the transition from the two symptomatic stages to the recovered (or treated) stage of the SEIR model structure is modeled. As in the SIMCOLEP model, the effects of the efforts made to treat leprosy patients on the islands in the period 1960-2000 are captured via a *probability of treatment* (P_t) and a *detection delay* (D), changing over time (5) (Fischer et al., 2010). Only symptomatic patients can be treated, as only these patients can be detected (Bakker,

2005a). In addition a *relapse rate* (R) determines which percentage of the treated compartment relapses (6) (Fischer et al., 2010). A predefined ratio (*percentage MB at relapse)* determines which percentage of these relapsing patients enters the PB/ MB track of disease, as this ratio differs significantly from the initial ratio at infection (Meima et al., 2004). This variable is set to 90%, as is done in the SIMLEP & SIMCOLEP models (Meima et al., 2004; Fisher et al., 2010).

In the disease model infections from infectious to susceptible individuals, disease development in infected individuals, relapses of recovered individuals, treatment of infected individuals and the intervention/ prevention strategies are modeled. In this section the separate events within the disease model are described in detail.

Infection event

Only MB patients are infectious (cause new infections), but are so during the asymptomatic state as well. The infectiousness of MB patients increases linearly (from 0 to 1) with the incubation time, and remains constant during the symptomatic stage of the disease (Figure 4.17).

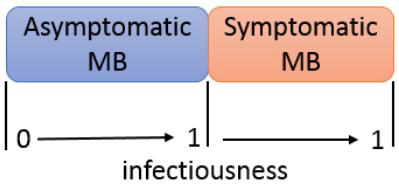


Figure 4.17: Infectiousness MB patients

Infection events are modeled by calculating the *probability of infection* for each susceptible agent at each time-step. This *probability of infection* is dependent on the number of infectious agents in the different contact groups of the agent and the infectiousness of these contacts. In order to calculate the infection probabilities the agent collectives for each agent are thus determined first. For each agent the household contacts (C_h) are those agents with the same *house_ID* as the agent. The neighbor contacts (C_n) are those agents with a *house_ID* equal to one of the *neighbor_houses* of the house of the agent. The island contacts (C_i) are all agents with same island_ID as the agent, while the fishing contacts (C_f) are all agents with the same boat_ID. To determine the probability of infection, the following equation applies:

$$(6)\ P_i = (P_s + \left(\frac{\sum(C_i * I)}{\sum C_i} * F_i\right) + \sum(C_n * I) * F_n + \sum(C_h * I) * F_h + \sum(C_f * I) * F_h) * G$$

$$Where\ P_i\ is\ probability\ of\ infection, P_s\ is\ start\ probability,\ C_i\ isIsland\ Contacts,$$
$$C_n is\ Neighbor\ Contacts, C_h is\ Household\ Contacts, C_f\ is\ Fishing\ Contacts,$$
$$F_i is\ Island\ Intimacy\ Factor, F_n\ is\ Neighbor\ Intimacy\ Factor, F_h is\ Household\ Intimacy\ Factor,$$
$$I\ is\ Infectiousness, G\ is\ Global\ Infection\ Rate$$

For the retrospective part of the simulation (1960-2000) the start probability (P_s) is set at 1 to simulate the increasing prevalence rate in the study area over this time period. This leads to a direct representing of the *global infection rate* (G) in the *probability of infection* (P_i).As the model starts with one infected agent in 1960 (section 4.3.1.4), an additional driving factor behind new infections needs to be applied to come to the correct prevalence rate in 2000. For the experiments (2000-2010/ 2000-2025) the start probability is set at zero, as the purpose of these experiments is

to gain an insight into the effect of these experiments on the incidence rates. Within the experiments, all new infections can thus be directly attributed to a contact within one of the contact groups.

For the island contacts the sum of the infectiousness (I) of each agent in the island contact-group(C_i), is divided by the total number of island contacts(C_i). This figure is multiplied by the *island intimacy factor*(F_i). In this way the influence of the island-specific prevalence rates on the *probability of infection* is captured in the model, as leprosy is known to cluster on the islands (Bakker et al., 2006). Furthermore, the degree to which the island contacts attribute to the probability of infection can be directly controlled via the *island intimacy factor*. As the study by Van Beers et al. (1999) could attribute approximately 64% of infections to household or neighbor contacts, the remaining 36% of infections occurred from another type of contact. Within LEPRASIM these contacts are covered (and controlled) via the island and fishing contacts.

For each agent the *probability of infection* (P_i) is increased by the sum of the infectiousness (I) of all contacts in each of the remaining contact groups (neighbor(C_n), household (C_h) and fishing contacts(C_f), multiplied by the appropriate *intimacy factor* (F_n & F_h) and the *global infection rate* (G). Boats are considered to be temporary households. If an agent is embarked on a fishing activity the *household intimacy factor* ($\mathrm{F_h}$) is thus applied to his fishing contacts(C_f). In this case, the number of household contacts, neighbor contacts and island contacts amount to zero, and the probability of infection is thus determined solely by the infectiousness (I) of the fishing contacts. In case the agent is not embarked on a fishing activity, the probability of infection is determined by the infectiousness of the household, neighbor and island contacts, as the number of fishing contacts amounts to zero.

The addition of the start probability and the four contact-group-probabilities multiplied by the *global infection rate* (G) gives the *probability of infection* ($\mathrm{P_i}$) for each agent. The *global infection rate* is set at $4.83 = 10^{-4}$, as this was the global infection rate, or "general probability of infection" found in the extensive retrospective study by Van Beers et al. (1999). The *probability of infection* ($\mathrm{P_i}$) is applied and resulting infections are implemented by changing the state of the *infected?*-variable to true for the newly infected.

Disease Development and Treatment Events

The environmental variable *percentage susceptible* (S) and *percentage genetic type MB* determine the percentage of population susceptible to leprosy infections and the type of disease these agents will develop after infection. The ratio between potential MB and PB patients is set for each agent at initialization (and at each birth) through the *genetic leprosy type* variable, and determines whether the agent will develop PB or MB leprosy after infection. At infection an individual receives the stage of disease "asymptomatic" and enters a latency period (or incubation period) dependent on the type of leprosy he will develop. This incubation period differs for PB and MB leprosy infections, and is applied by drawing a value from a normal distribution at the moment of infection. The values for these incubation times are drawn from the work by Meima et al. (2004)

and Fisher et al. (2010) on the SIMLEP-model and SIMCOLEP-model. For PB patients this average incubation time is 4.2, with a standard deviation of 1.9 years. For MB patients the average incubation time is 11.1, with a standard deviation of 5.0 years.

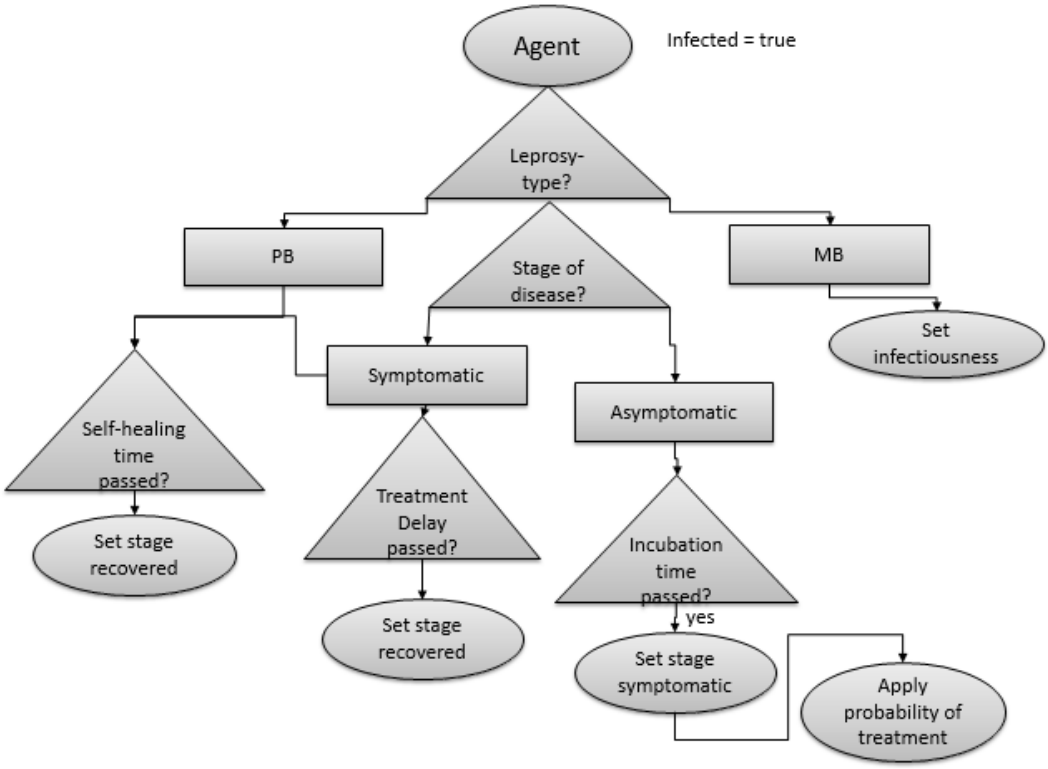


Figure 4.18: Disease Development and Treatment Events

Each time-step the schematic depicted in Figure 4.18 applies to all infected agents. For each infected agent the leprosy-type and stage of disease are determined. For all agents in the asymptomatic stage of the disease (PB or MB) each time-step the passing of the incubation period is checked. If so, they are moved to the symptomatic stage of infection (PB or MB, dependent on the *genetic leprosy type* variable) and the environmental variable *probability of treatment* (P_i) is applied to determine whether they will be treated. If the *probability of treatment* applies, a treatment delay is set, determined by the environmental variable *detection delay* (D). For all symptomatic patients each time-step the passing of this treatment delay is checked (treatment event). If so, the agent is moved to the recovered stage of the disease, and is no longer infected.

In addition, the level of infectiousness for all MB patients (symptomatic and asymptomatic) is set at each time-step. For symptomatic patients this infectiousness is 1, for a-symptomatic patients it is calculated by dividing the *time_since_infection* by the *incubation_time* giving a ratio from 0 to 1. Lastly, as PB patients are considered to be self-healing over time (Meima et al., 2004), for agents with a genetic-leprosy type PB a *time-until-self-healing* is set at the moment of transition from asymptomatic to symptomatic. After this time had passed, the agent transitions to the "recovered" stage, and is no longer infected.

In the efforts made to control leprosy on the islands in the Flores Sea in the period from 1960 to 2000 three periods have been distinguished by Bakker (2005a). In the period 1960-1993 health care on the islands was very limited and the treatment for leprosy occurred with the dapsone antibiotic. From 1993 onwards a passive case detection program with subsequent deployment of MDT has been deployed on the islands (Bakker, 2005a). From 2000 onwards an active case detection program has been employed on the islands, as Bakker et al. conducted their study in the area. These three different treatment programs have been represented in LEPRASIM using the *probability of treatment* (P_t) and a *detection delay* (D) variables. For the values the initial parameters used in SIMCOLEP for Bangladesh in 1960 (Fisher et al., 2010) have been used as initial values of these same parameters in LEPRASIM. The changes in these parameters over time have been determined based on an interview with Dr. Mirjam Bakker.

The *probability of treatment* (P_t) starts at 0.05 in 1960, increasing to 0.2 in 1993. Between 1993 and 2000 the probability increases from 0.2 to 0.4, after 2000 the probability is set at 1, as for the purpose of the experiment the entire population has been screened for signs of leprosy on a yearly basis (Figure 4.19).

The detection delay (D) starts at 12 years with a standard deviation of 3.5 years in 1960 (Fischer et al., 2010), decreasing to 5 years, with the same standard deviation in 1993 as the health care facilities on the islands gradually improve. In the period 1993 to 2000 the detection delay is 5 years, from 2000 onwards 1 year (Figure 4.20).

Figure 4.19: Probability of Treatment

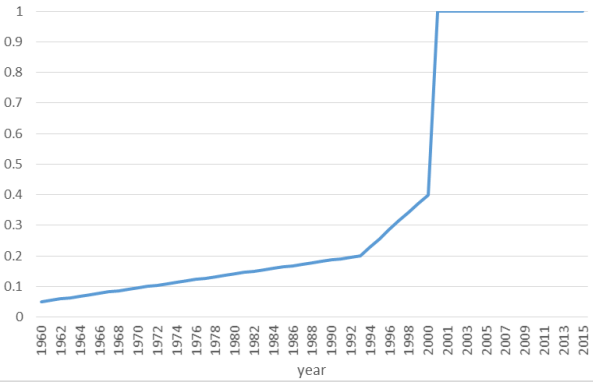
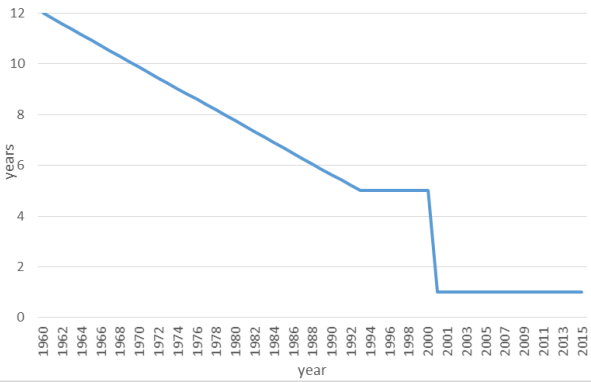


Figure 4.20: Detection Delay



Relapse events

A yearly *relapse rate* (R) applies to all recovered individuals, as in the SIMCOLEP model (Fischer et al., 2010). This implies that yearly (each twelve time steps) each recovered individual has a fixed probability of moving back to the infected stage of the disease. Which type of leprosy he will develop is no longer dependent on the initial ratio between PB and MB patients, but on the global variable *percentage MB at relapse*, which is set 90%, as is done in the SIMCOLEP-model by Fisher et al. (2010). The relapse rate decreases over time, as MDT has been refined, and thus has become "better" over the years, decreasing linearly from 0.015 in 1960 to 0.001 in 2000 via the following equation (taken from the work by Fisher et al., 2010):

$$(7)\ R = 0.015 - 0.0004 * x$$

Where R is relapse rate; x is model-year

Intervention events

With the arrival of Dr. Mirjam Bakker and her study in 2000 the passive case detection program on the islands has been replaced with an active case detection program. This means that in the model a different *probability of treatment* (P_t) and *detection delay* (D) apply (Figures 4.19 & 4.20): there is an absolute probability of treatment (P_t = 1), as the entire population off the experiment-population is screened for leprosy on a yearly basis (Bakker et al., 2006). The *detection delay* is thus set at one year (D = 1) In addition to the shift to an active case detection program, in experiment 2 chemoprophylaxis is employed via contact tracing of infected individuals.

The prevention strategies are aimed at the agent collectives. In order to apply the prevention strategy, these agent collectives are thus determined, as is done for an infection event. The difference is that the agent collectives are determined for each MB symptomatic and PB symptomatic patient, instead of each susceptible agent. The pre-emptive medication works by changing the state of all "asymptomatic" patients within the relevant agent collectives to "susceptible". In this way, a set of eight prevention strategies aimed at different agent collectives at different time-intervals are tested on their effect on the cumulative incidence rate in the period 2000-2025, compared to the cumulative incidence rate in the same time-period in the baseline model performance.

4.4 Overview of global variables and initial values

In Table 4.7 an overview of the environmental variables, with their values is presented. For values, that are part of the sensitivity analysis or the calibration process, the possible range of values is shown, as this is the range drawn from literature.

Table 4.7 Global variables: initial values, range and source			
Sub-model	Variable	Value/ Range	Source
Population Model	Fertility Rate	$3 * 10^{-6} * year^2 - 4 * 10^{-4} * year + 0.227$	World Bank (2012)
	Child Mortality Rate	$4 * 10^{-5} * year^2 - 6 * 10^{-3} * year + 0.2241$	World Bank (2012)
	Life-expectancy at birth	$-4.6 * 10^{-3} * year^2 + 0.5203 * year + 57.631$	World Bank (2012)
Activity Model	Marry Rate	50% - 100%	Fischer et al (2010)
	Percentage Own House at Marriage	15% - 50%	Fischer et al (2010)
	Percentage Fishermen	90%	Bakker (2005a)
Disease Model	Infection Rate	$3 * 10^{-4} - 15 * 10^{-4}$	van Beers et al (1999)
	Percentage Suceptible	10%-30%	Fischer et al (2010)/ Interview Mirjam Bakker
	Percentage genetic type MB	10% - 50%	Meima et al (2004)
	Probability of Treatment	1960 - 1993: 0.05 - 0.2 1993-2000: 0.2 - 0.4 2000-2010: 1	Fischer et al (2010)/ Interview Mirjam Bakker
	Detection Delay	1960 - 1993: 12 years - 5 years (standard deviation = 3.5) 1993-2000: 5 years (standard deviation = 3.5) 2000-2010: 1 year	Fischer et al (2010)/ Interview Mirjam Bakker
	Relapse Rate	$0.015 - 0.0004 * year$	Fisher et al (2010)
	MB incubation Time	5 - 15 years (standard deviation = 5.0)	Meima et al (2004)
	PB incubation time	4.2 years (standard deviation = 1.9)	Meima et al (2004)
	Household/ Boat Intimacy Factor	6-20	van Beers et al (1999)
	Neighbor Intimacy Factor	2-10	van Beers et al (1999)
	Island Intimacy Factor	1-8	Interview M. Bakker

5. Sensitivity Analysis, Calibration and Verification

This chapter consists of the sensitivity analysis, calibration process and verification activities performed on the LEPRASIM model. The sensitivity analysis (5.1) starts with a "stability check" (5.1.1) to determine the sensitivity of the average model output to the number of model iterations. The point at which LEPRASIM becomes stable, i.e. the minimum sample size of model-runs needed to produce a stable result, is found using the methodology developed by Lorscheid et al. (2012). The sensitivity analysis itself (using the one-parameter-at-a-time (OAT) –method (Hassani-Mahmooei & Parris, 2013)) is presented in section 5.1.2. The calibration (5.2) of the model in terms of prevalence rate, MB:PB ratio and spatial sources of new infections on observations made by Bakker et al. (2002; 2004) in the study area in the period 2000-2003 is presented next. In the verification section (5.3) the implications of the calibration process on the models behaviour are verified by presenting the population model's behaviour, the disease model's behaviour and the spatial implications of the disease model over time. The results of the validation process are presented in chapter 6 (Results).

5.1 Sensitivity Analysis

5.1.1 Stability Check

In order to test at which point LEPRASIM becomes stable, i.e. to determine the **minimum sample size** of model-runs needed to produce a stable result, the methodology developed by Lorscheid et al. (2012) is employed. In this method a **coefficient of variation** is used to express the model's stability (Lee et al., 2015). The coefficient of variation is defined as the ratio between the standard deviation of a sample and the mean of that sample:

$$Cv = \frac{\sigma}{\mu}$$

Where Cv = coefficient of variation, σ *= standard deviation of sample and* μ *= sample mean.*

For a fixed parameter set of runs this metric is calculated for two output-parameters of the disease sub-model: the leprosy prevalence per 10.000 people and the MB:PB ratio in 2000. Both output-parameters are used in the model calibration, as data on these parameters is available in the study area (Bakker et al., 2005). The model's stability should thus be analyzed in light of these parameters as well. As the model needs to be stable on both parameters, the minimum number of runs for the model is the maximum of these minimum sample sizes (Lee et al., 2015).

In Figure 5.1 the individual model outcomes are plotted against the average of these outcomes for both parameters of interest for 300 runs. As can be seen the average model outcomes tend to become stable rather quickly as the number of model runs is increased. The coefficient of variation is calculated for the following parameter-set of runs: n = {5, 10, 25, 50, 100, 200, 300}. The resulting coefficients of variation for both output-parameters are show in Table 5.1 and Figure 5.2.

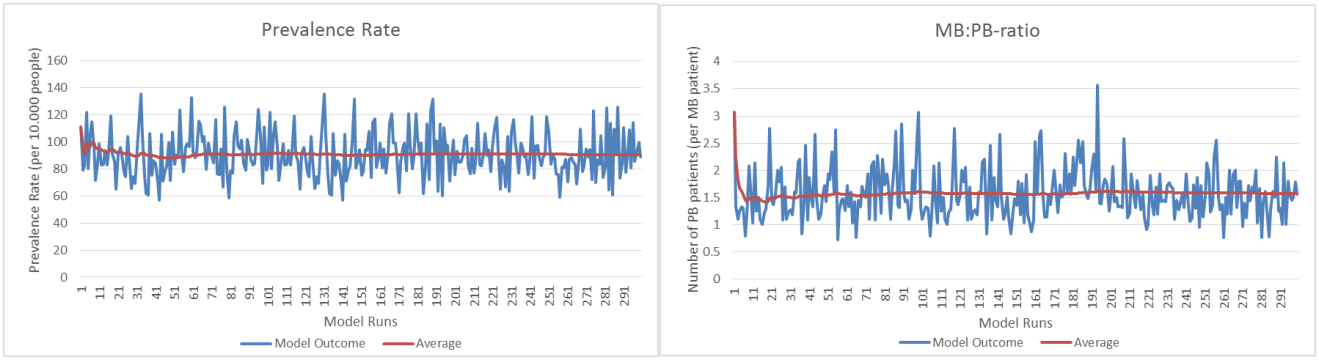


Figure 5.1 Outcome Disease Model Run vs. Average

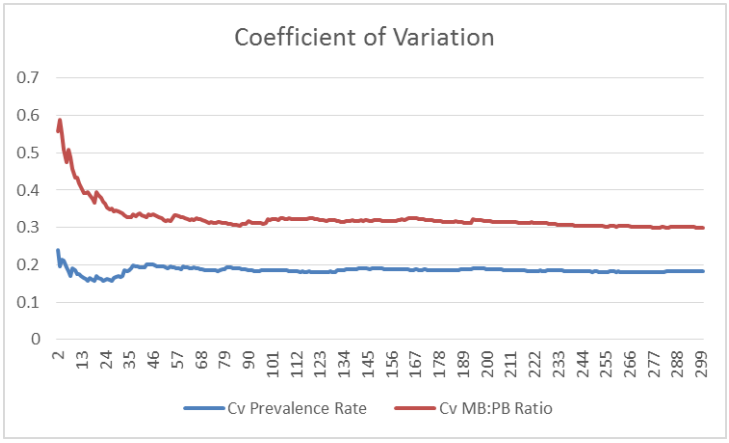


Table 5.1: Coefficient of Varation		
N	Cv	
	Prevalence Rate	MB:PB-Ratio
5	0.21	0.55
10	0.19	0.46
25	0.16	0.36
50	0.2	0.33
100	0.19	0.32
200	0.19	0.32
300	0.18	0.30

Figure 5.2 Coefficient of Variation Prevalence Rate and MB:PB ratio

The model output becomes stable rather quickly, as the coefficient of variation stabilizes after approximately fifty runs. Performing more than fifty model iterations does not add to the stability of the model outcomes in a significant manner, as the coefficient of variation does not alter more than 0.01. Fifty model iterations is thus the minimum number of model iterations required for the model to produce a reliable result. For the sensitivity analysis, calibration of and experiments performed with the model this is thus the number of runs performed.

5.1.2 Sensitivity Analysis

A sensitivity analysis on the model is performed to assess the sensitivity of the LEPRASIM-model to its input parameters. The sensitivity analysis is performed using the one-parameter-at-a-time (OAT) –method (Hassani-Mahmooei & Parris, 2013). In this method, each input parameter is examined over a predefined range of values in isolation, as the remaining parameters are set at a constant baseline (Lee et al., 2015). In order to make a complete assessment of LEPRASIM, the sensitivity off the sub-models to each other needs to be assessed. Therefore, the sensitivity analysis comprises of three steps (Figure 5.3). In the first step the sensitivity of the population sub-model to the driving factors of the activity sub-model is assessed (5.1.2.1). In the second step the internal sensitivity of the disease sub-model is examined, as the other sub-models serve as input for this disease sub-model (5.1.2.2). In the third step the sensitivity of the disease sub-model to variations in the population sub-model's output parameters is assessed, as the outcome off the population sub-model determines the influence of the activity sub-model on the disease sub-model (5.1.3.1).

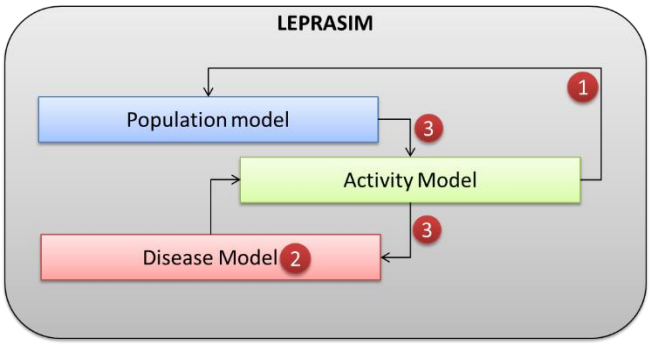


Figure 5.3: Steps Sensitivity Analysis

Within each of the three steps the effects of the input parameters on a selection of output-parameters are monitored (over fifty model iterations). During these iterations the value of the input parameter is randomly varied within the predefined range. The input parameter-output relationships are visualized using scatter-plots, revealing potential dependencies (Lee et al., 2015). In addition a standardized regression coefficient (SRC) is estimated for each relationship, giving an indication of the size of the relationship between the different input parameters and the specific outcome parameter. The SRC expresses the sensitivity of the independent variable to the dependent variable in terms of the effect one standard deviation in the dependent variable has on the independent variable as a percentage of its standard deviation (Lee et al., 2015). In this way the different scales of and variations in parameters are standardized, so that the relative effects of different independent variables are made comparable.

Per step of the sensitivity analysis, the in- and output parameters on which the sensitivity analysis is performed are shown in Table 5.2. The baseline input parameter values and the predefined range of variation for these input-parameters is shown as well. In addition, the source this range of variation is based on is provided.

In the first step the outcome parameters (Table 5.2) used are the population size, average household size, the percentage of male agents, the percentage of agents aged below 10 and the mean age. The sensitivity of these outcome parameters to the predefined variations in the *marry rate* and *percentage own house at marriage* input variables is assessed.

In the second step is the sensitivity of the disease model and its spatial representation to its input parameters is assessed. The prevalence rate and MB:PB ratio in 2000, as well as the cumulative

incidence rate in the period 2000-2003 are used as output parameters for the disease model. In addition, to assess the sensitivity of the models spatial distribution of infections a number of spatial measures are used. The percentage of houses with more than one patient is used as indicator for clustering within households. The percentage of new patients, which are a household or neighbor contact in the period 2000-2003, is used as an indicator of the spatial sources of new infections. The minimum and maximum prevalence rate per Island in 2000 is used as an indicator of the clustering of patients on Islands.

In the third step is the sensitivity of the disease model to the output of the population model is assessed, by examining the variation in prevalence rate, MB:PB ratio, incidence rates and percentages of houses with zero patients (as an indicator of clustering within households) in relation to variations in the population size and average household size.

Table 5.2: Sensitivity Analysis

Step 1: sensitivity of population model to activity model					
Input Parameters				**Output Parameters**	
Parameter	**Baseline Value**	**Range of Variation**	**Source**	**Parameter**	**Year**
Marry rate	80%	50% - 100%	Fischer et al (2010)	Population Size	2000
				Average Household Size	2000
Percentage own house at marriage	25%	15% - 50%	Fischer et al (2010)	Percentage Male	2000
				Percentage 0-10	2000
				Mean age	2000
Step 2: internal sensitivity of disease model					
Percentage Susceptible	20%	10%-30%	Fischer et al (2010)/ Interview Mirjam Bakker	Prevalence Rate	2000
				MB:PB-ratio	2000
Infection Rate	$4.83 * 10^{-4}$	$3 * 10^{-4} - 15 * 10^{-4}$	van Beers et al (1999)	Incidence Rate	2000-2004
				Percentage of houses with more than 1 patients	2000
MB Incubation Time	11.1 years, s = 5	5 - 15 years	Meima et al (2004)		
Percentage genetic type MB	20%	10% - 50%	Meima et al (2004)		
Household/ Boat Intimacy Factor	12.7	6-20	van Beers et al (1999)	Spatial source of new infections	2000-2004
Neighbor Intimacy Factor	5.2	2-10	van Beers et al (1999)	Minimum Prevalence Rate (per Island)	2000
Island Intimacy Factor	3	1-8	Interview M. Bakker	Maximum Prevalence Rate (per Island)	2000
Step 3: sensitivity of disease model to population model					
Population Size	2000	5441	Bakker (2005a)	Prevalence Rate	2000
				MB:PB-ratio	2000
Average Household Size	2000	4.29	Bakker (2005a)	Incidence Rate	2000-2004
				Percentage of houses with zero patients	2000

5.1.2.1 Step 1: sensitivity of population model to activity model

To test the sensitivity the population model to the activity model the *marry rate* and *percentage own house at marriage* variables were randomly varied within the predefined range (Table 5.2) using 50 model iterations. The effect on the outcome parameters of the population model is plotted per input variable in Appendix A1 (Figure A1.1 – A1.10) and the standard regression coefficients with associated R^2 values are calculated (Table 5.3).

Table 5.3 Step 1: sensitivity of population model to activity model						
		Population Size	Average Household Size	Percentage Male	Percentage Aged 0-10	Mean Age
Marry rate	SRC	0.065	**-0.924**	-0.167	0.189	-0.240
	R^2	0.004	**0.855**	0.028	0.036	0.057
Percentage own house at marriage	SRC	0.246	**-0.783**	0.097	0.230	-0.090
	R^2	0.060	**0.613**	0.009	0.053	0.008

As the population growth within the population model is driven by input data (World Bank, 2015) no correlation should occur between the input variables of the activity model and the output variables of the population model .As can be seen from the scatter-plots (appendix A1) and SRC's this indeed is not the case. The marry rate and the percentage of marrying couples forming a new household at marriage do show a strong correlation with the average household size (SRC of -0.924 and -0.783 respectively): an increase in the *marriage rate* or *percentage own house at marriage* parameters causes a decrease in the average household size. The *average household size* is however more sensitive to the *marry rate* than to the *percentage own house at marriage* variable (as can be seen by the difference in R^2 values). The population model thus behaves as was intended, with only the average household size being highly sensitive to the input parameters of the activity model.

5.1.2.2 Step 2: internal sensitivity of disease model

The influence of each of the disease model's input parameters on its output parameters was analyzed (using 50 model iterations). In Table 5.4 the SRC's with associated R^2 values for the disease model output parameters are shown. First, these relationships will be discussed for the three identified disease model output parameters (prevalence rate, MB:PB ratio and incidence rate). Hereafter, the sensitivity of the spatial distribution of these infections to the different input parameters will be discussed.

Table 5.4: Step 2: internal sensitivity of disease model				
		Prevalence Rate	MB:PB Ratio	Incidence Rate
Percentage Susceptible	SRC	0.883	0.249	0.871
	R^2	0.779	0.062	0.759
Infection Rate	SRC	0.042	0.141	0.958
	R^2	0.914	0.020	0.918
MB Incubation Time	SRC	-0.390	0.504	-0.303
	R^2	0.154	0.205	0.094
Percentage genetic type MB	SRC	0.868	-0.827	0.804
	R^2	0.753	0.683	0.646
Household Intimacy Factor	SRC	0.117	-0.099	0.116
	R^2	0.014	0.010	0.014
Neighbor Intimacy Factor	SRC	0.299	0.111	0.380
	R^2	0.089	0.012	0.144
Island Intimacy Factor	SRC	0.117	-0.053	0.178
	R^2	0.014	0.003	0.032

Prevalence Rate

Using the total prevalence rate in 2000 as an independent variable, the influence of the disease model input parameters was analyzed. As can be seen in Table 5.4 variations in the intimacy factors had no significant effect on the total prevalence rate. From the scatter plots (Appendix A2: Figure A2.1 – A2.4) the main determinants of the prevalence rate can be deduced. The *percentage susceptible* and *percentage genetic type MB* variables show the strongest positive correlation with the prevalence rate (SRC of 0.883 and 0.868 respectively). An inspection of the associated R^2 values shows a different image: the *infection rate* is the best determinant for the total prevalence rate. Although the SRC-value is lower, the relationship is more pronounced (R^2 of 0.914). The *MB incubation time* shows a negative correlation with the prevalence rate: as the incubation time for MB (infectious) patients increases, the total prevalence rate decreases.

MB:PB Ratio

The MB:PB ratio in 2000 is mainly dependent on the percentage of patients developing MB after infection, determined by the *percentage genetic type MB* variable (SRC = -0.827, R^2 = 0.683). As more patients develop MB leprosy at infection, the MB:PB ratio goes down. The MB:PB ratio is sensitive to the *MB incubation time* variable to a lesser extent (SRC = 0.504, R^2 = 0.205): as the incubation time for MB patients increases, the MB:PB ratio goes up, meaning there are relatively more PB patients in 2000. The MB:PB ratio is not sensitive to the *percentage susceptible* or *infection-rate* variables (see Appendix A2.2: Figure A2.5-A2.8).

Incidence Rate

The incidence rate (2000-2003) is most sensitive to the *infection rate* (SRC= 0.958, R^2 = 0.918). In addition the incidence rate is sensitive to the *percentage susceptible* and *percentage genetic type MB* variables as well, but to a lesser extent (Appendix A2.3: Figures A2.9-A2.12). To examine whether the incidence rate in the period 2000-2003 is truly sensitive to these parameters or to the Model's situation in 2000 the correlation between this model situation in 2000 in terms of the prevalence rate and the incidence rate was analyzed (Table 5.5 and Appendix A2.4: Figures A2.13-A2.14).

As can be seen from the scatter plots, SRC's and associated R^2 values (Table 5.5) the *incidence rate* in the period 2000-2003 is highly sensitive to the *prevalence rate* in 2000 (SRC =0.959, R^2 = 0.919). This means that in addition to the infection rate, the number of infected patients at that moment in time determines the incidence rate in the period 2000-2003 within LEPRASIM.

Table 5.5: Disease Model: sensitivity in time

		Incidence Rate (2000-2004)
Prevalence Rate	SRC	0.959
	R^2	0.919
MB:PB Ratio	SRC	-0.351
	R^2	0.380

Spatial Distribution of Infections

To assess the sensitivity of the spatial diffusion of leprosy infections to the disease model's input parameters three categories of output parameters are used as a dependent variable in the sensitivity analysis: an indicator of clustering within households, an indicator of clustering on islands and the spatial spread off new infections resulting from an infectious agent in different contact groups. The results can be seen in Table 5.6.

Table 5.6 Spatial distribution of infections

		Clustering within households	Clustering on Islands		Spatial sources of new infections (2000-2004)		
		% of houses with more than 1 patient	Minimum Prevalence Rate	Maximum Prevalence Rate	non-contact	household-contact	neighbor-contact
household intimacy factor	SRC	0.135	-0.032	0.051	-0.145	0.101	0.123
	R^2	0.018	0.001	0.003	0.021	0.010	0.015
neighbor intimacy factor	SRC	0.149	0.172	-0.154	-0.413	0.030	0.430
	R^2	0.022	0.030	0.024	0.171	0.001	0.185
island intimacy factor	SRC	0.135	0.020	-0.272	-0.134	-0.023	0.190
	R^2	0.018	0.000	0.074	0.018	0.001	0.036
percentage susceptible	SRC	0.611	0.030	-0.271	-0.559	0.447	0.444
	R^2	0.373	0.001	0.073	0.312	0.200	0.197
infection rate	SRC	0.838	0.183	-0.198	-0.827	0.362	0.777
	R^2	0.702	0.034	0.039	0.685	0.131	0.604
MB Incubation Time	SRC	-0.189	-0.303	-0.324	0.280	-0.189	-0.209
	R^2	0.036	0.092	0.105	0.081	0.037	0.045
Percentage genetic type MB	SRC	0.754	0.058	-0.317	-0.735	0.543	0.680
	R^2	0.568	0.003	0.100	0.540	0.295	0.462
Prevalence rate	SRC	0.872	0.403	0.101	-0.826	0.412	0.753
	R^2	0.761	0.162	0.010	0.683	0.170	0.567

Clustering within households

Using the percentage of households with more than 1 patient as a dependent variable the sensitivity of the degree of clustering of patients within households to the disease model input parameters was assessed. The clustering of patients within households is most sensitive to the infection rate and percentage of MB patients in the model (Table 5.9). As the household intimacy factor operates as a multiplier to the infection rate, and only MB patients are infectious, this is correct model behavior. This observation is supported by the sensitivity of the clustering within households to the prevalence rate: the higher the prevalence rate, and the higher the percentage of MB patients within this infected population, the higher the clustering within households becomes.

Clustering on islands

To assess the degree of clustering of patients on the islands the prevalence rate per island in 2000 was calculated in relation to the total prevalence rate in 2000 in that model iteration. From these five prevalence rates (as a percentage of the total prevalence rate) the minimum and maximum prevalence rates per model run were deduced. The effect of the input parameters on these minimum and maximum prevalence rates was expressed in a SRC, with associated R^2 value. The clustering of leprosy patients over the islands is most sensitive to the number of MB patients in a model run (*percentage genetic type MB)*. As the number of MB patients in the simulation increases, the clustering of patients within households, neighborhoods and thus islands becomes greater. As this clustering is modeled as being dependent on infections occurring within islands, household and neighbor contact groups, this is an expected dependence.

Spatial sources of new infections (2000-2004)

An increase in any of the *intimacy factors* decreases the percentage of new infections originating from non-contacts, as the relative weight of island, neighbor or household contacts with an infected patient on the infection probability is increased. The *island intimacy factor* shows a negative relation with the amount of infections originating from non-contacts as well, as the likelihood of an infected individual being a neighbor or household contact increases with the increase in the relative probability of infection from the island contacts. The *percentage susceptible*, *percentage genetic type MB* and *infection rate* variables show a strong negative correlation with the percentage of new infections originating from non-contacts, as all three variables increase the relative probability of transmission between household and neighbor-contacts.

5.1.2.3 Step 3: sensitivity of disease model to population model

To complete the sensitivity analysis of the LEPRASIM model the sensitivity of the disease model to the population model needs to be assessed, as the outputs of the population model might have an influence on the disease model output parameters. From the sensitivity analysis of the population model (5.1.2.1) the conclusion is drawn that only the average household size is sensitive to its input parameters. The input data used for the population model thus results in an approximation of the population size on the islands in 2000, with some variation caused by the stochastic nature of the population model. To assess the sensitivity of the disease model to the population model the correlation between the output parameters of the population model and the output parameters of the disease model is assessed.

The output parameters of the population model used for this assessment are the average household size and the population size. The disease model output parameters used are the prevalence rate (2000), the MB:PB ratio, the incidence rate (2000-2003) and the percentage of households with no patients in 2000. The correlations were assessed using 50 model iterations and are plotted in scatterplots (see appendix A3: Figures A3.1 – A3.8) for each of the disease model's output parameters. In these figures blue represents the population-size and red the average

household size as independent variable. As can be seen from the scatterplots the disease model is not sensitive to variations in the population model at all. These two sub-models of LEPRASIM thus act independently, as they were intended to do.

5.1.2.4 Model sensitivity: summary

The population model is primarily sensitive to the input data used to drive it. Within the population model the average household size is controlled by the *marry rate* and *percentage own house at marriage* parameters. The output average household size is most sensitive to the *marry rate*, and to a lesser extent to the *percentage own house at marriage* variable. The population model does not influence the dynamics within the disease model: the population size and average household size resulting from the population model do not influence the outputs of the disease model. Within the disease model the prevalence and incidence rate are most sensitive to the *percentage susceptible, infection rate* and *percentage genetic type MB* variables. The ratio between MB and PB patients is determined by the *percentage genetic type MB* variable, as was intended.

The clustering of infections within households can be explained to a large extent by the total prevalence rate occurring in a model run. As the prevalence rate is higher, the chance more agents within one household are infected increases. Although the household intimacy factor adds to this clustering, it does so to a smaller extent than the total prevalence. Looking at the sources of new infections (the contact groups), a more clear effect of the intimacy factors is noticeable: the amount of infections occurring from neighbor contacts is sensitive to the neighbor intimacy factor. The clustering on the islands is determined by a combination of the three intimacy factors and the prevalence rate.

5.2 Calibration

The external validity of the model is obtained via a global calibration process. By variation on a selection of the parameters identified in the sensitivity analysis the disease model's output is fit to the observation's made by Bakker et al. (2002; 2006) in the study area in 2000 and in the period 2000-2004. After these parameters are identified (5.2.1), the global calibration process can commence. In the first step of this process the baseline model performance is compared to the observations. In the second step the disease model is calibrated to match the cumulative incidence and prevalence rate observed in the study area. In the third step the distribution of these infections over households and islands is calibrated using the "intimacy factors". In other words, a spatial calibration of LEPRASIM is performed.

5.2.1 Calibration Parameters

The calibration of the disease model is done on both the general disease characteristics as well as its spatial implications. For the calibration off the disease model the *percentage of agents susceptible* to leprosy infection and *the percentage of infected agents developing MB leprosy* are used to calibrate the model on the Prevalence Rate and MB:PB ratio of infections

Table 5.7: Calibration Input Parameters

Parameter	
percentage susceptible	20%
percentage genetic type MB	20%
island intimacy factor	3
neighbor intimacy factor	5.2
household intimacy factor	12.7

observed in the study area in 2000 (Bakker et al., 2002). The three *intimacy factors* are used to control the degree of clustering within households, neighborhoods and islands. The spatial sources of new infections occurring in the model are calibrated on the observations made in the study area in the period 2000-2003 (Bakker et al., 2004) and the retrospective study by van Beers et al. (1999). The input parameters used for the calibration process are shown in Table 5.7, the output measures (with the accepted range) in Table 5.8. The accepted range for the prevalence rate is defined by the 95% Confidence Interval of the made observation in 2000 (Bakker et al., 2002). The range for the MB"PB ratio is defined by a maximum 5% difference in the percentage MB patients, giving an accepted ratio ranging from 1:1 to 1:33. The range for the spatial sources of new infections is defined by the results of two studies: the observations made in the study area in the period 2000-2003 (Bakker et al., 2004) and the retrospective study by Van Beers et al. (1999).

Table 5.8 Calibration Output Measures				
Measure	Value	Range	Year	Source
Prevalence Rate (per 10.000)	195	156-234	2000	Bakker et al (2002)
MB:PB ratio	1:1.22	1:1 - 1:33	2000	Bakker et al (2002)
Spatial Sources of Infection: New cases per contact group as % of all new cases				
non-contact	80%	36% - 80%	2000-2003	Bakker et al, 2004
	36%			van Beers et al, 1999
household contacts	9%	9% - 28%		Bakker et al, 2004
	28%			van Beers et al, 1999
neighbor contacts	11%	11% - 36%		Bakker et al, 2004
	36%			van Beers et al, 1999

5.2.2 Step 1: Baseline Model Performance

Using the initial values for the calibration input parameters (Table 5.7) the model was run for 50 iterations. This resulted in a rather large underestimation of both the prevalence (93) and MB:PB ratio (1:1.61) (Table 5.9). In other words: these model settings lead to an underestimation of the number of infectious patients, leading to an underestimation of the prevalence rate observed on the islands in 2000.

Table 5.9: Baseline Performance LEPRASIM		
Measure	Observation	Baseline Model Performance
Prevalence Rate (per 10.000)	195	93
MB:PB ratio	1:1.22	1:1.61
Spatial Sources of Infection: New cases per contact group as % of all new cases		
non-contact	80%	65%
household contacts	9%	12%
neighbor contacts	11%	23%

5.2.3 Step 2: Calibration of General Disease Model Parameters

As was shown in the sensitivity analysis, the prevalence rate is most sensitive to the *percentage susceptible* and *percentage genetic type MB* variables (Table 5.4). From the interviews with Mirjam Bakker the assumption that the percentage of people susceptible to leprosy infections living on the Islands is higher than elsewhere on the world was formed. Furthermore, as a random allocation of

susceptibility to the agents is used, and no genetic factors are used to steer this allocation, the research by Fisher et al. (2010) suggests that the percentage of susceptible individuals should likely be increased to make the model fit the observations. For this reason, as a first step in the global calibration process, the percentage of susceptible individuals was raised (from 20%) to 30%, and the model was run for 50 more iterations. As the MB:PB ratio was underestimated as well, the effect of an increase in the *percentage MB at infection* parameter (from 20%) to 30% was analyzed as well (Table 5.10). Each model setting was run for 50 iterations.

An increase in the *percentage susceptible* parameter leads to the expected increase in the prevalence rate (to 159). The increase in the *percentage susceptible* variable, and thus in the prevalence rate, however alters the MB:PB ratio in the wrong direction. As more people are susceptible to leprosy infections and the percentage of patients developing MB leprosy remains at a constant level, the ratio logically shifts towards more PB patients. Looking at the spatial sources of infections a slight overestimation of both the infections originating from household as well as neighbor contacts remains the case. Compared to the baseline model output this overestimation increases. To approximate the MB:PB ratio more closely and further increase the prevalence rate, the percentage of susceptible individuals developing MB leprosy was increased from 20 to 30% (fourth column Table 5.10). With these parameter settings the model slightly overestimates the prevalence rate. The MB:PB ratio is now slightly underestimated.

Table 5.10: Calibration of General Disease Model Parameters

Percentage Susceptible		**20%**	30%	30%	**30%**
Percentage MB at infection		**20%**	20%	30%	**25%**
Measure	Observation	Baseline Model Performance	Increased Susceptibility Performance	Increased Susceptibility and MB performance	Calibrated MB Performance
Prevalence Rate (per 10.000)	195	93	159	233	210
MB:PB ratio	1:1.22	1:1.61	1:1.72	1:1.12	1:1.15
Spatial Sources of Infection: New cases per contact group as % of all new cases					
non-contact	80%	65%	55%	42%	46%
household contacts	9%	12%	14%	19%	18%
neighbor contacts	11%	23%	32%	38%	37%

This first exploration (Table 5.10) shows that the model runs with a susceptibility percentage of 30% and a *percentage genetic type MB* of 20% and 30% resulted in an MB:PB ratio of 1:1.72 and 1:12 respectively. As an additional step the *percentage of MB at* infection was set at 25%. The model was run for another 50 iterations to verify the effect of this alteration (Table 5.10). The correct MB:PB ratio is approximated to a high degree (1:1.15, meaning a 5.74% difference with the observed ratio). As the observed ratio has some uncertainty of its own, this approximation is within the predefined range. The total prevalence rate is captured within acceptable boundaries as well. As was expected from the results of the sensitivity analysis, the amount of new infections occurring as a result of household and neighbor contacts increases with an increase in both the input parameters. To capture the spatial diffusion of leprosy observed in the study area, the intimacy factors are calibrated to match this diffusion, keeping in mind the relative uncertainty of these observations.

5.2.4 Step 3: Spatial calibration on intimacy factors

As the amount of infections occurring from household and neighbour contacts are overestimated, both the household and neighbour intimacy factors were lowered to more closely match the observations made in the study area. Although this led to a closer approximation of the distribution of spatial sources of infections over the three contact groups, the overall prevalence rate was greatly reduced (to 137), as the household and neighbour intimacy factors are strongly correlated with the prevalence rate. To compensate for this effect, the island intimacy factor is increased. As expected from the sensitivity analysis this led to an increase of the prevalence rate to an acceptable level (177), while preserving the distribution over the different contact groups.

Table 5.11: Spatial Calibration on intimacy factors				
household intimacy factor		**12.7**	7.0	7.0
neighbor intimacy factor		**5.2**	3.0	3.0
island intimacy factor		**3.0**	3.0	15.0
Measure	Observation	Calibrated MB Performance	Decreased HH & NB intimacy factor	Increased Island intimacy factor
Prevalence Rate (per 10.000)	195	210	137	177
MB:PB ratio	1:1.22	1:1.15	1:1.35	1:1.31
Spatial Sources of Infection: New cases per contact group as % of all new cases				
non-contact	80%	46%	63%	62%
household contacts	9%	18%	12%	11%
neighbor contacts	11%	37%	26%	28%

Although the distribution over the different contact groups is now more closely approximated, and within the accepted range, the prevalence rate is still slightly underestimated. One final calibration step increases the percentage of susceptible individuals from 30% to 32%, leading to an increase of the prevalence rate from 177 to 193. The final calibration parameters are shown in Table 5.12.

Table 5.12: Final Calibration		
model parameters		
percentage susceptible		32%
percentage genetic type MB		25%
island intimacy factor		15
neighbor intimacy factor		3
household intimacy factor		7
output measures		
	observation (range)	model
Prevalence Rate	195	193
MB: PB Ratio	1:1.22	1:1.29
Spatial Sources of New Infections		
non-contacts	36% - 80%	59%
hosuehold-contacts	9% - 28%	12%
neighbor-contacts	11% - 36%	29%

5.3 Verification

The calibrated model's behaviour is verified in a number of ways. First, an assessment of both the population model's behavior is made. Next, the disease model's behavior is examined via the output of a single model run. Next, an insight into the spatial implications of the disease model is obtained via an "infection-chain".

5.3.1 Population Model Behaviour

Using 1779 agents as an initial population, and spread over age and sex as described in chapter 4, the population sub-model was run for 1000 iterations. The resulting histogram of the outputs (Figure 5.4) and the main statistics (Table 5.13) are shown.

Figure 5.4: Histogram Population Size 2004 (1000 model iterations)

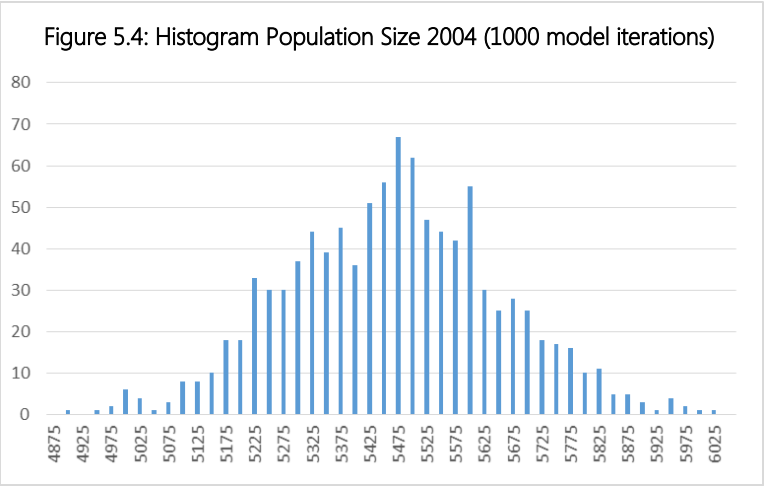


Table 5.13: Summary Statistics Population Model Output 2004	
Initial Population	1780
Iterations	1000
Mean	5454.41
Standard Error	5.85
Median	5458.5
Mode	5493
Standard Deviation	185.23
Sample Variance	34309
Kurtosis	-0.0725
Skewness	0.0275
Range	1136
Minimum	4889
Maximum	6025

The outputs of the model, in terms of total population, show a normal distribution. The mean of this distribution for the 1000 performed iterations is 5454, with a standard deviation of 185. The distribution was tested for normality using the chi-squared statistic. A p-value of 0.37 indicated that it is indeed a normal distribution given a 0.05 significance level. (Chi-square = 44.42, with a critical chi-value of 58.12). The standard deviation of 185, with a standard error of 5.85, indicates that the model is not very precise in predicting the "right" outcome: it shows a rather large variance. This variance is however not very skewed (0.02). 1.000 model iterations give a confidence interval of 5454 +/- 11.48 for a significance of 0.05. When averaging over a large number of runs (1000) the average of the model output can thus be considered to be a good approximation of reality in terms of total population size in 2004. To further investigate the population model its behaviour in relation to the observations made in the study area by Bakker et al. (2002; 2006) was tested (Table 5.14).

Table 5.14: Population Model Behavior compared to Observations Bakker et al. (2002; 2004)

Year	Parameter	Observed Value	Average Model output (1000 iterations, 0.05 significance)
2000	Total Population	4774	5009 +/- 10.20
	Percentage Male	48%	0.4799 +/- 0.00045
	Percentage 0-10	30%	0.298 +/- 0.000439
	Median Age	20	19.73 +/- 0.0228
	Mean Age	24	21.37 +/- 0.0127
2004	Total Population	5441	5456 +/- 11.48

The population model correctly simulates the total inhabitants for 2004, the percentage male and children and the median age in 2000. The mean age however is slightly underestimated. The biggest error occurs in predicting the number of inhabitants in 2000, as the model averagely predicts 5009 inhabitants, while in reality there were only 4774 (5% overestimation). The yearly population growth in the years 2000-2004 on the islands however has been 115 inhabitants, which is significantly larger than the average growth rate obtained from the population figures for whole Indonesia (56 new-borns per year). It thus appears that the growth rate on the islands has been lower than the average for Indonesia in the years prior to 2000, while it has been significantly higher in the years 2000-2004. Another possibility is that this growth rate is an effect of migration, which has not been taken into account.

The population model follows the population statistics for whole Indonesia. The demographics of the population on the two islands however differ (slightly) from these demographics, as the population has shown a higher birth-rate in the years 2000-2004. As no demographic data prior 2000 is available for the islands, and the exact causes of the higher population growth between 2000 and 2004 on the islands is unknown, the population model is the best approximation of the demographic processes possible. The population model accurately predicts the population in 2004 (by means of calibration), with a normal distribution, having acceptable characteristics (passed chi-square test with $p = 0.37 > 0.05$)).

5.3.2 Disease Model Behaviour

The disease model's behavior over one model run is shown in Figure 5.5 -5.7. At initialization one symptomatic MB patient is "placed" in the model, to mimic the introduction of leprosy on the islands. Starting from this symptomatic infectious MB patient and the global infection rate (*start infection probability * infection rate)* infections occur as the model progresses over time. The number of symptomatic and asymptomatic patients is roughly the same as the model progresses, while more agents have PB then MB leprosy. This is in line with findings by Bakker et al. (2005). The treatment of leprosy patients is virtually non-existent in the first 100 steps of the model, to represent the lack of treatment on the islands in this time (1960 – 1970). After this the number of recovered individuals gradually goes up. The strength of this process is increased as the last period of treatment (2000-2004) begins. From this point on we see the number of recovered agents increasing rapidly, as was the case from the moment Bakker et al. (2005) began the active case detection programs on the islands.

Figure 5.5: Disease Model Behavior: MP-PB ratio

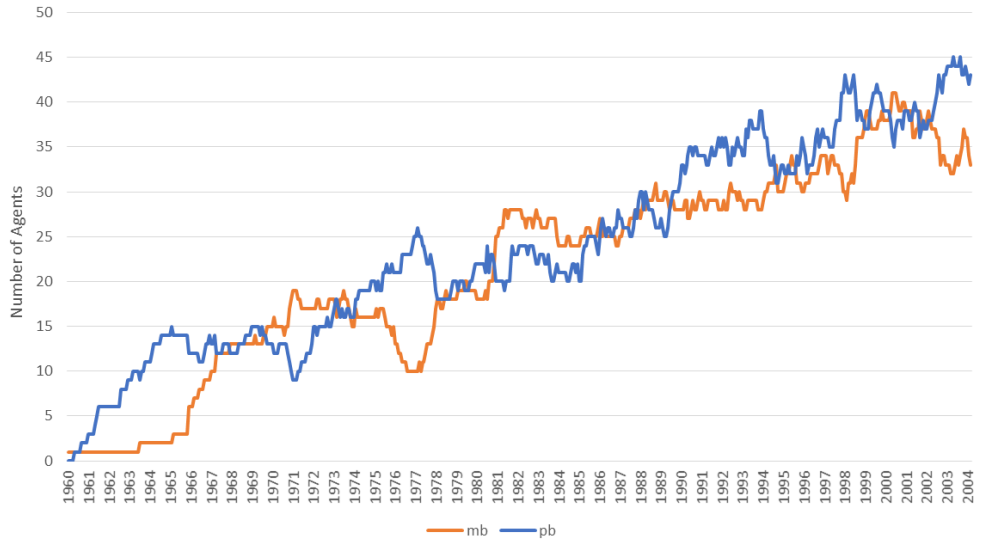


Figure 5.6: Disease Model Behavior: Infected & Recovered

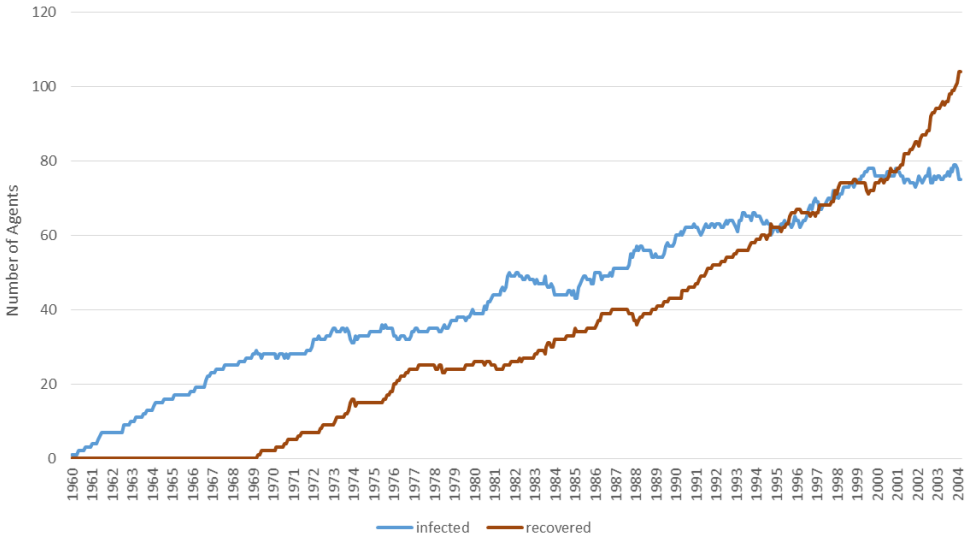
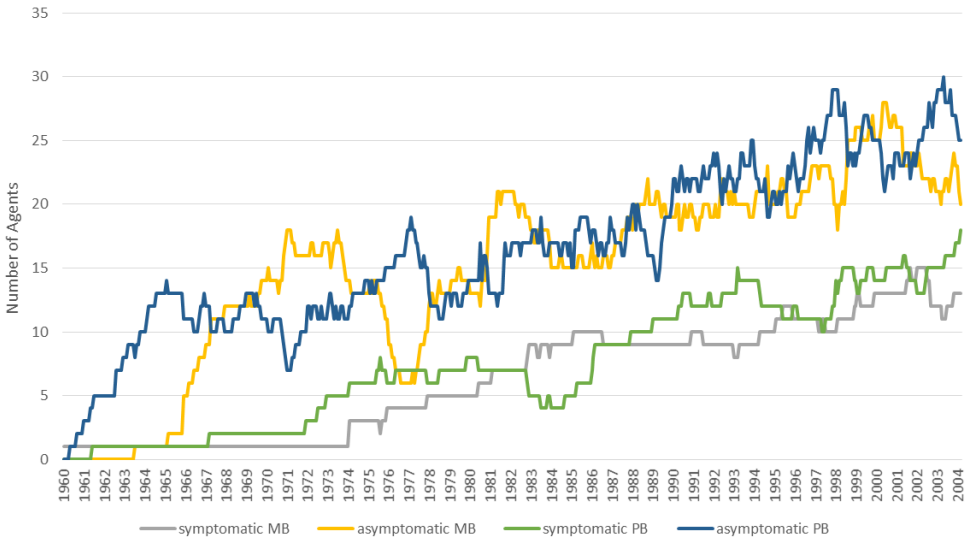


Figure 5.7: Disease Model Behavior: symptomatic/ asymptomatic



5.3.3 Spatial Implications Disease Model Structure

To provide an insight into the spatial implications of the combination of the population and disease model within LEPRASIM a small cluster of households has been "followed" during one model run (Figure 5.8 – 5.10) to visualize a chain of infections occurring within the model as a result of household and neighbour infections. The household, containing the first infected patient at initialization was selected, in combination with its first and second degree neighbouring houses. In Figure 5.8 -5.10 the colour of the houses indicates how much agents are residing in the household. The selected house is labelled X, its first order neighbours 1A:1E, its second order neighbours 2A:2D. The format of the label is as follows: *house-label: number of* ***susceptible agents****, number of* ***infected agents****, number of* ***infectious (MB) agents****.*

Red circles and arrows indicate a transmission of leprosy as a result of an increased probability of infection due to an infectious household or neighbour contact respectively. In Figure 5.11 the infections in the separate time-steps are added to form a cumulative infection chain.

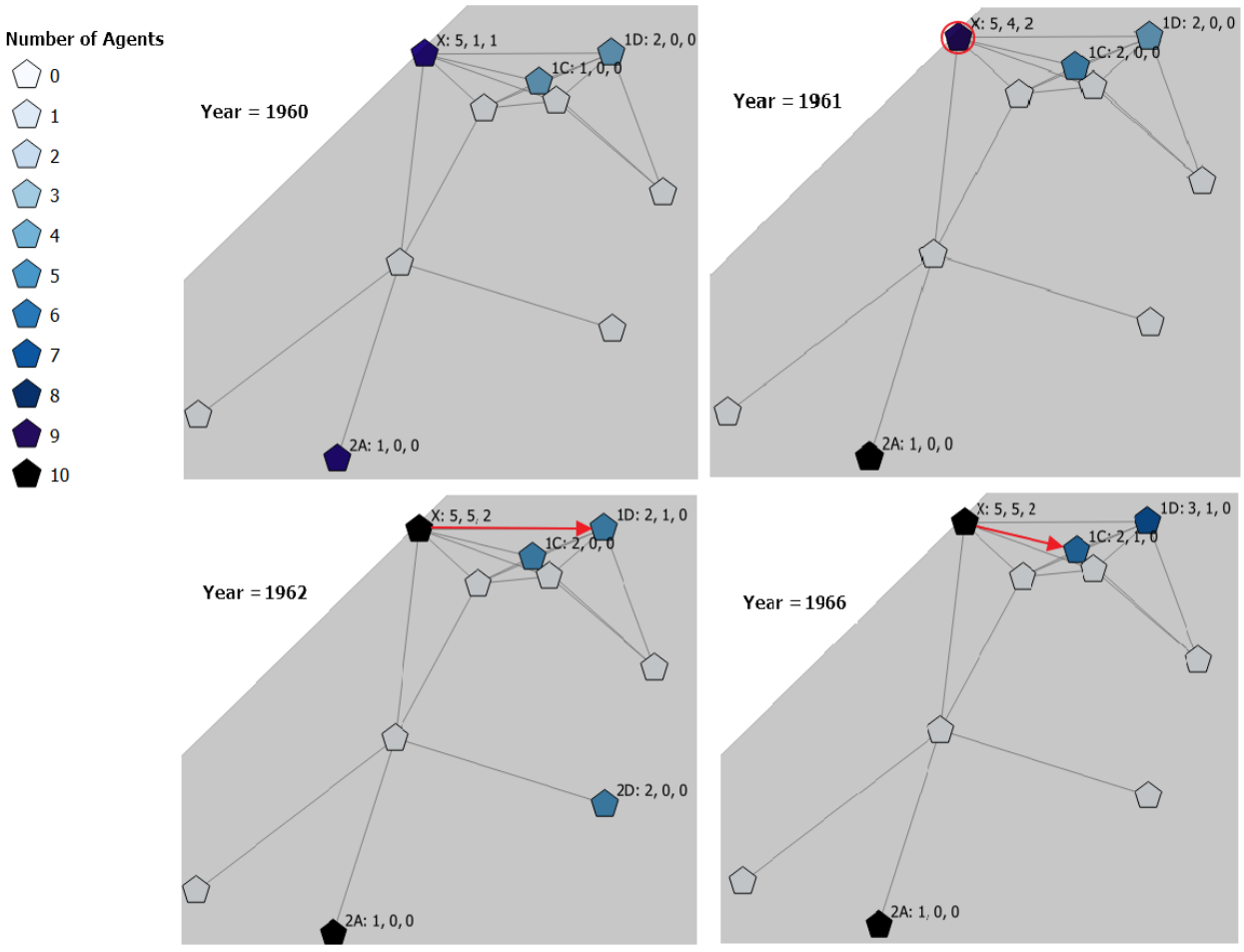


Figure 5.8: Infection Chain: Model Situation 1960, 1961, 1962 and 1966

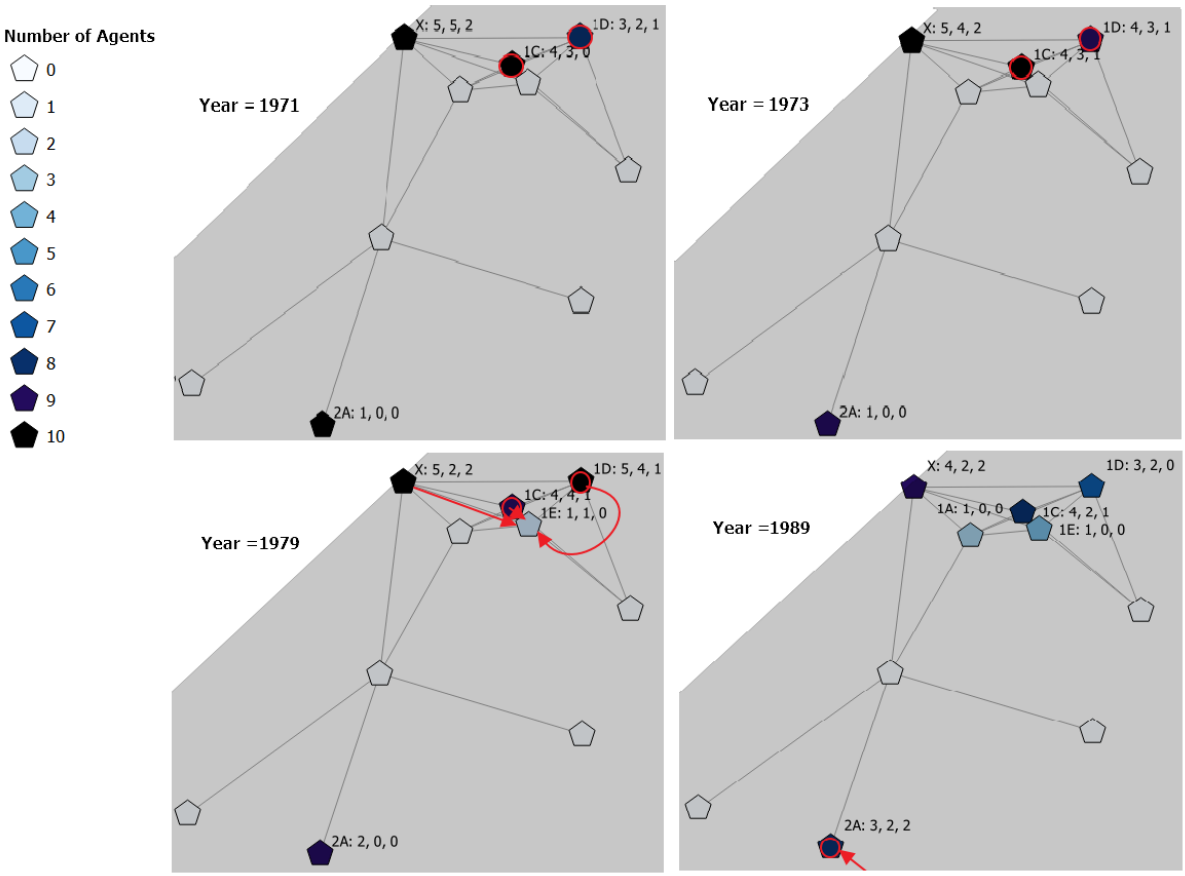


Figure 5.9: Infection Chain: Model Situation 1971, 1973, 1979 and 1989

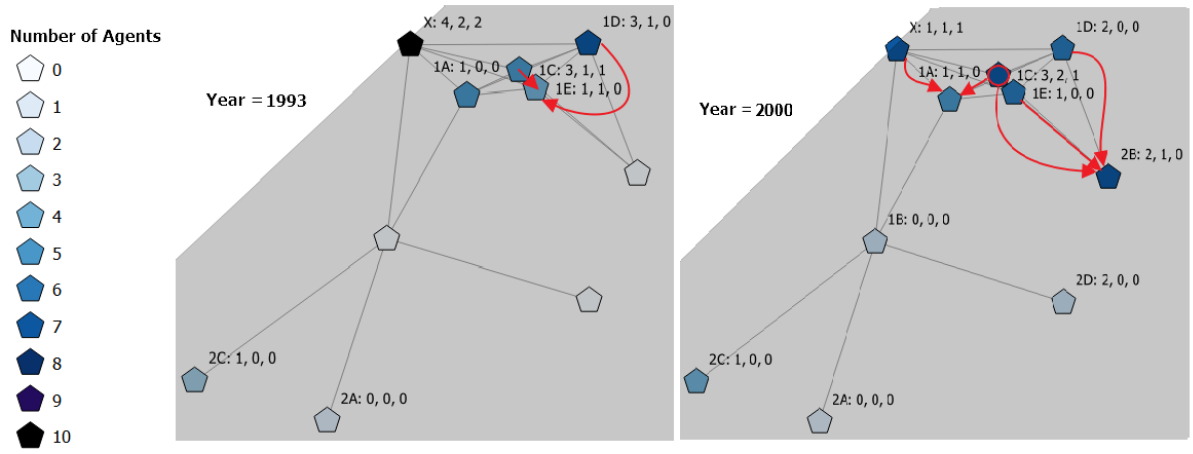


Figure 5.10: Infection Chain: Model Situation 1993 and 2000

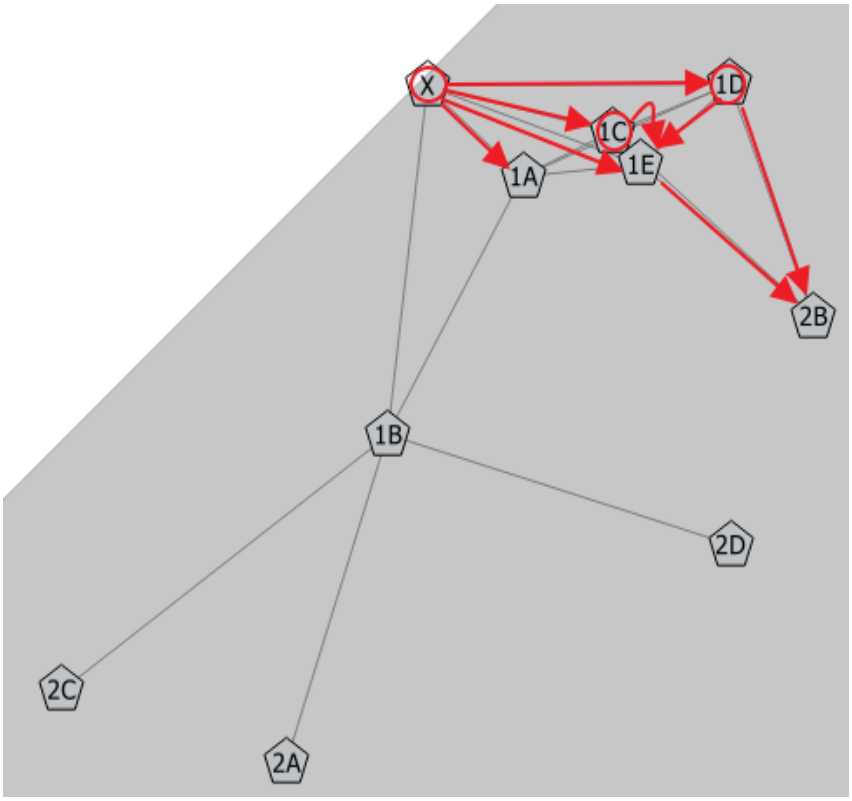


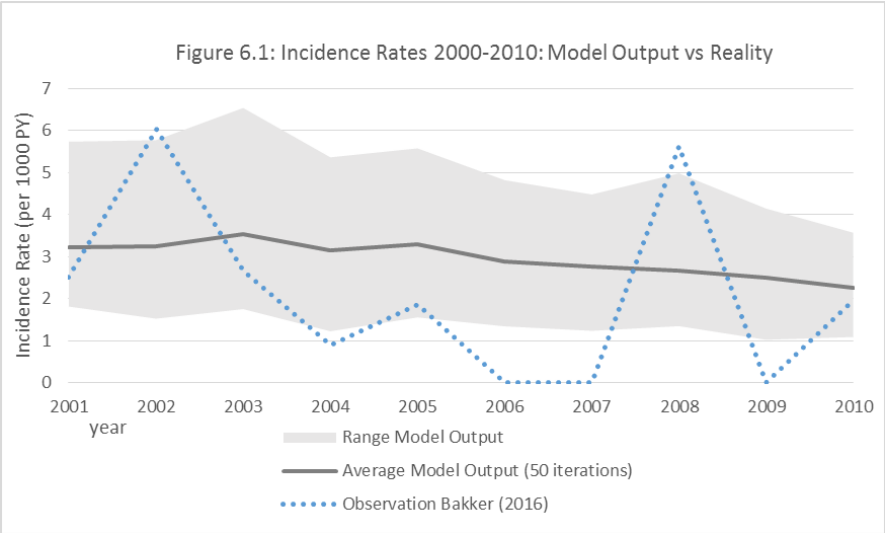
Figure 5.11: Cumulative Infection Chain

During the model run, the initial infection in 1960, located in household X resulted in a chain of infections leading to an infection in household 2B in 2000. As can be seen from the developments in the area over time, this infection chain is dependent on both the distribution off the population over the households as the increase in the *infection probability* of certain agents caused by their household/ neighbour and island contacts. Within houses X, 1C and 1D infections occurred as a result of an infectious household contact, for houses 1A, 1E and 2B the infection was the result of an infectious neighbour contact.

6. Results

6.1 Validation

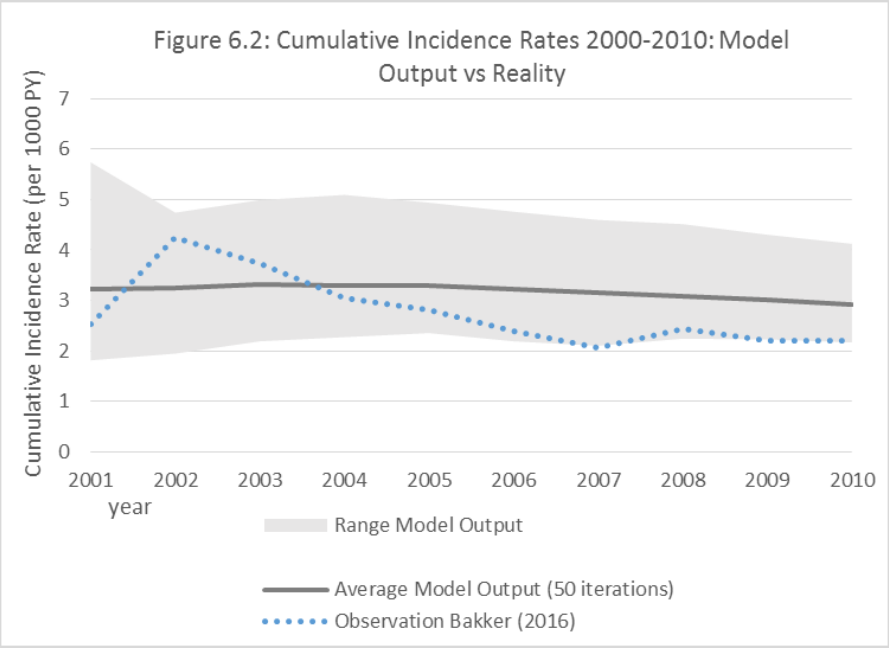
The model was run for 50 iterations using the settings presented in Table 5.15. The results are discussed below per validation parameter.



Incidence Rates Control Group 2000-2010

As can be seem from Figure 6.1, the yearly incidence rate per 1000 PY observed in the study area in the period 2000-2010 (Bakker, 2010) varies significantly. The rather large variation in yearly incidence rate (blue line in Figure 6.1) indicates an incidence rate based on a (very) small number of participants. The average model output over 50 model iterations (black line) and the range of variation in the model output (grey area) are provided as well (grey area). Although the range of variation in the observation does not fit within the range of variation in model output, this is explained by the relative uncertainty of the observations (Bakker, 2016).

In Figure 6.2 the cumulative incidence rate for the control group in the period 2000-2010 is shown. The cumulative incidence rate observed by Bakker (2016) in the study area fits within the range of the cumulative incidence rates produced by the model (grey area in Figure 6.2). On average, LEPRASIM however slightly overestimates this cumulative incidence rate. Where this Figure declined from 2.51 (2001) to 2.2 (2010) in reality (a decline of 0.31), in the simulated output it declines from 3.23 (2000) to 2.92 (2010) (which is also a decline of 0.31). Although the model thus slightly overestimates the height of the cumulative incidence rate, it does follow the trend of the observation made in reality: a decline in cumulative incidence rates as a result of the intervention done by Bakker et al. in 2000. The relative decrease in cumulative incidence rates ($1 - \frac{CIR_{2000-2010}}{CIR_{2000-2001}}$) in the model is 10%, where it was 13% in reality.



The difference between the model output and reality can be explained via three reasons: (1) the relative uncertainty of the observations made by Bakker (2016), as the incidence rates are based on a (very) small number of participants, and thus show a rather large variation (blue line in Figure 6.1). (2) The incidence rate (and cumulative incidence rate) produced by the model is the mean of 50 model-iterations, and is thus less likely to fluctuate. (3) The relatively low cumulative incidence rate in the observation for the period 2000-2001, compared to 2000-2002. This spike in new case detections in the observation in 2002, as a result of the yearly screening of the entire population, is less pronounced in the model output than it was in reality.

Clustering in Contact Groups: Hazard Ratio 2000-2003

Looking at the cumulative incidence rates (2000-2003) for the status on household size in 2000 (Table 6.1) the model underestimates the effect of household size on the risk of getting infected with leprosy. Where the hazard ratio was 3.47 for individuals living in a household with more than seven inhabitants in reality, in the model this hazard ratio is only 1.17. This is explained by both an overestimation of the cumulative incidence rate for the small households (1-4), as well as an underestimation of the cumulative incidence rate for the large households (>7). For the medium-

Table 6.1 Model Validation on clustering in contact groups: Hazard Ratio 2003

Status in 2000 on:		Cumulative Incidence Rate (2000-2003)		Hazard Ratio	
		Reality	Model	Reality	Model
household size	1-4	188 (104 – 3.40)	2.99	1	1
	5-7	3.09 (2.02 – 4.74)	3.32	171 (0.82 – 3.56)	111
	>7	5.61 (3.19 – 9.88)	3.51	3.47 (151 – 7.98)	117
contact status	no contact	2.88 (2.00 -4.15)	2.53	1	1
	neighbor contact	3.31 (149 - 7.38)	3.77	152 (0.50 - 4.59)	149
	household contact	6.67 (3.00 – 14.9)	6.18	3.29 (111 – 9.77)	2.44

sized households (5-7), the model correctly predicts the cumulative incidence rates. Although for all three groups the cumulative incidence rate is within the 95% confidence interval (Bakker et al., 2006), due to the underestimation of the influence of household size, the hazard ratio for the large households (>7) is not.

When looking at the cumulative incidence rates (2000-2003) for the contact status in 2000 we can see the model correctly predicting both the cumulative incidence rates, as well as the resulting hazard ratios. Within LEPRASIM, household contacts of infected individuals thus have the highest risk of getting infected, followed by neighbor contacts of infected individuals, as is the case in reality. In addition, the greater the household an agent lives in, the greater the risk of getting infected, although this effect is slightly underestimated.

Clustering in contact groups: situation in 2000

The results of the validation of the model output on clustering in contact groups are presented in Table 6.2 When comparing the observations made by Bakker et al. (2006) on the percentage of patients and controls (people not showing signs of leprosy) living in a cluster with at least one other patient in 2000 to the model output, we can again observe the same underestimation of the clustering in households. When looking at the relation between percentage of patients and controls living in a household with at least one other patient the chance a patient is living in such a household is approximately 1.75 times larger than a control living in such a household. In the observation by Bakker et al. (2006) this ratio is 2.07. The percentage of controls living in a neighbor-cluster on the other hand is slightly overestimated (40% compared to 30%), leading to an underestimation of the relative risk for patients living in such a cluster. All in all, the model correctly represents the relative probability of a patient living in a household with another patient and the percentage of patients living a neighbor-cluster.

Table 6.2 Model Validation on clustering in contact groups: situation 2000

Percentage of patients/ controls living in cluster with at least one other patient:				
	House		Household + Neighbors	
	reality	model	reality	model
Patients	32%	16%	52%	51%
Controls	16%	9%	30%	40%
Patients / controls	2.07	1.75	1.75	1.29

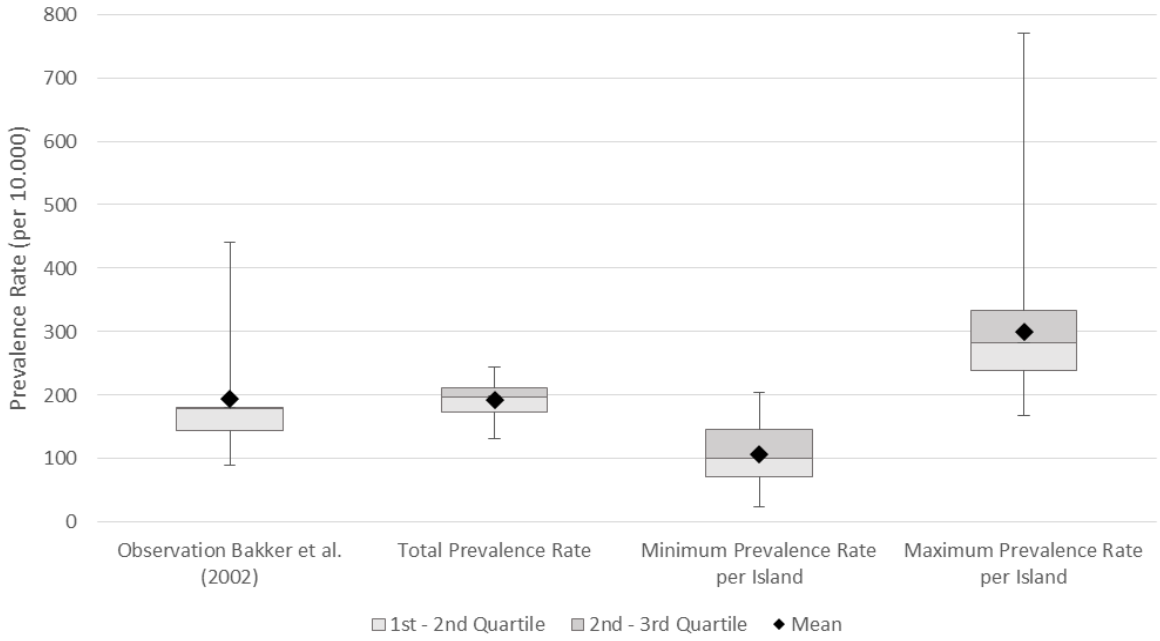
Clustering on Islands: prevalence rate per island 2000

Looking at the total prevalence rate in 2000 (Table 6.3), we can see that the model correctly simulates the number of infected individuals in the population (193 mean prevalence rate compared to 195 in observation). The clustering on islands is represented by the mean of the minimum and maximum prevalence rates measured over all five islands over 50 model iterations. In reality this minimum prevalence rate was 89, while the maximum was 440. The model replicates this clustering on islands, although to a less extreme extent: a mean minimum prevalence rate of 107 and a mean maximum prevalence rate of 299. Although this mean maximum prevalence rate is lower than the maximum prevalence rate in the observation (Bakker et al., 2006), it should be noted that this observation is located at a rather large distance from the 3rd quartile range (max=

440, 3rd Q = 181) and can thus be considered to be an outlier (see Figure 6.3). Looking at the variation in the maximum prevalence rate, it can be seen that the model replicates this phenomenon of strong variation in the maximum prevalence rate. The minimum prevalence rate, i.e. the clustering of non-patients on an island, is correctly represented by the model, as the mean minimum prevalence rate (107) approximates the minimum prevalence rate in the observation (89) and shows a relatively small variation around this point.

Table 6.3 Model Validation: Clustering on Islands 2000				
	Prevalence Rate		Percentage	
	reality	model	reality	model
Total	195	193	100.00%	100%
Minimum	89	107	45.64%	54.92%
Maximum	440	299	225.64%	122.53%

Figure 6.3: Model Output: Total PR, Min PR & Max PR 2000

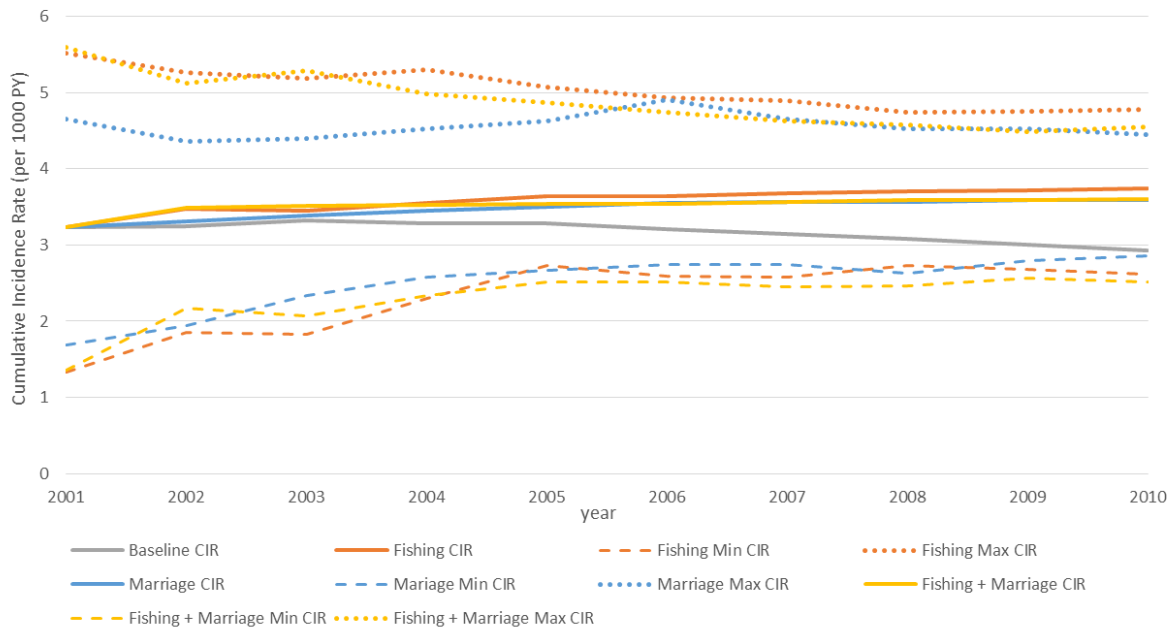


6.2 Experiment 1: Activity Modeling

6.2.1 Experiment 1a: Effect of activity modeling on spatio-temporal diffusion of leprosy

As can be seen from Figure 6.4, the addition of marriage and fishing activities to LEPRASIM leads to an increase of the cumulative incidence rates over time. Where this cumulative incidence rate decreases over time in the baseline model, as a consequence of the intervention employed in which all symptomatic agents are treated yearly (grey solid line), addition of the marriage and/ or fishing activities counteracts this effect (blue and orange solid lines).

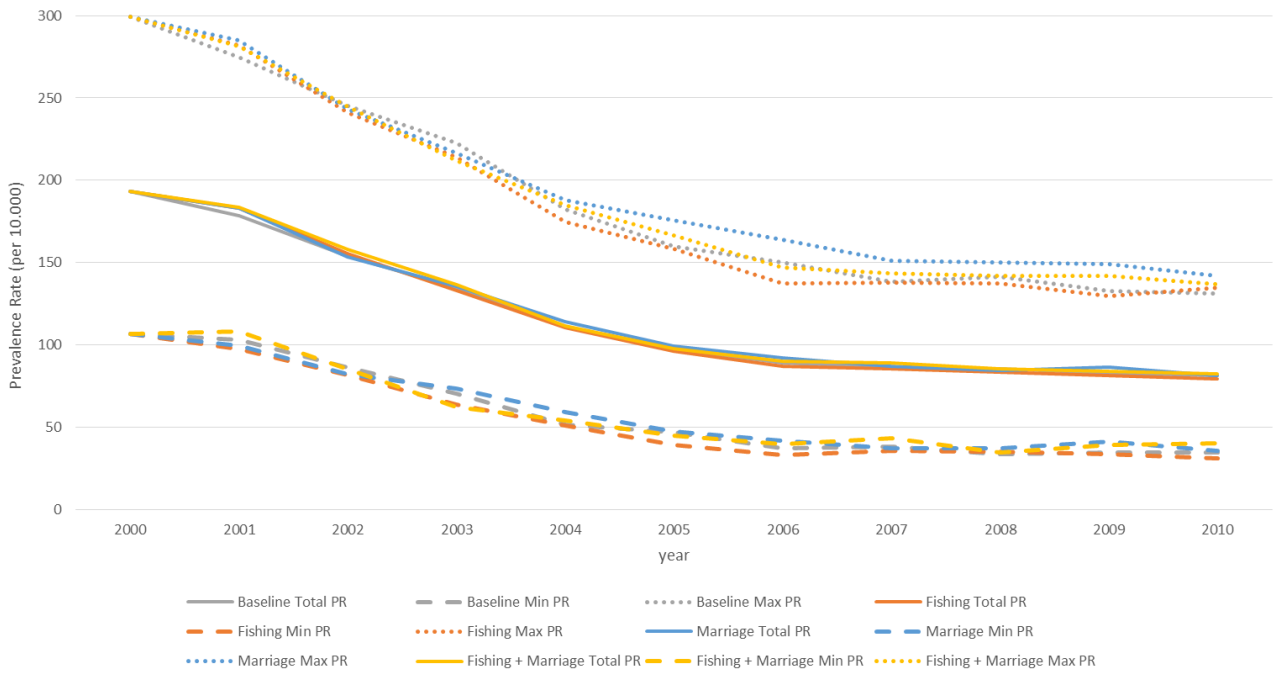
Figure 6.4: Effect of activity modeling on cumulative incidence rates 2000-2010



The reintroduction of leprosy in the control and blanket group of the study by Bakker et al. (2006) can thus be explained by a reintroduction of the disease via both marriage as well as fishing activities. The combination of both the fishing and the marriage activity (yellow solid line), however, does not lead to a larger increase in cumulative incidence rate over time, but to a similar outcome as of the separate modeling of the two activities. This effect can be explained by the intervention employed on the islands from the year 2000 on. As time goes by, all symptomatic individuals (infected and showing physical signs of leprosy) are treated each year. As the percentage of agents which is susceptible to the disease is limited (32.5%), the amount of susceptible untreated agents thus declines to such a low number, that the likelihood of an incidence rate above the shown average declines as well. This is confirmed by the convergence of the minimum and maximum cumulative incidence rates in the period 2006-2010 (Figure 6.4: dotted and dashed lines). The effect of the yearly treatment of infected symptomatic individuals on the total cumulative incidence rate thus exceeds the effect of the modeled activities. The first hypothesis can, however, be considered to be proven for both activities, as a significant difference in cumulative incidence rates with the baseline model can be observed.

When looking at the development of the prevalence rate in the same time-period (2000-2010) this observation is confirmed (Figure 6.5). Both the marriage and the fishing activities do not influence the average total prevalence rate in the study area over time. The influence of the active screening of the entire population for leprosy, and subsequently treating those infected, on this prevalence rate outweighs the influence of the modeled activities. The minimum and maximum prevalence rates are increased slightly by the marriage activity and the combination of the two activities. This means that the second and third hypotheses are only proven for the marriage activity, while they are false for the fishing activity.

Figure 6.5 Effect of activity modeling on total prevalence rate and minimum/ maximum prevalence rate per island 2000-2010

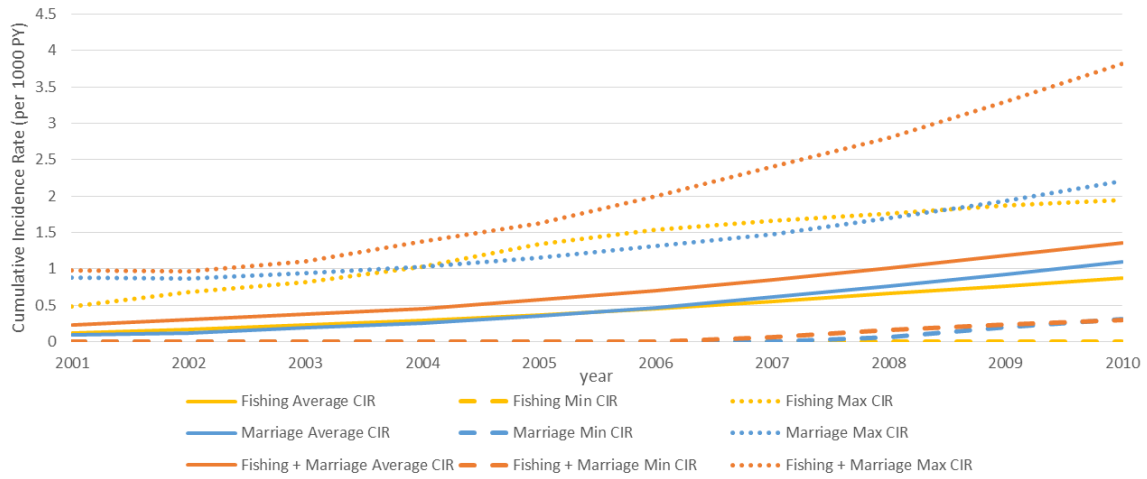


The addition of marriage/ fishing activities to LEPRASIM increases the likelihood of clustering of leprosy on the islands (higher max incidence rate) and the total incidence rate. The prevalence rate remains unaltered, as the yearly treatment of all symptomatic agents compensates for the effect of the increase in cumulative incidence rates. This means that the re-emergence of leprosy in the blanket and control group of the study by Bakker et al. (2006) can be explained by both the marriage as well as the fishing activity, although the magnitude of these activities remains unclear.

6.2.2 Experiment 1b: Reintroduction of leprosy through activity modelling

To gain an insight into the magnitude of the effect of the addition of inter-island activities on the reintroduction of leprosy within LEPRASIM a hypothetical intervention is modeled on one of the islands (Sapuka Besar), clearing all agents of infection in the year 2000. The cumulative incidence rate is measured for the period 2000-2010. The resulting figures are shown below.

Figure 6.6: Development of cumulative incidence rates Sapuka (2000-2010) as a result of activity modeling



The addition of the marriage activity and/or fishing activity leads to an increasing cumulative incidence rate on the test-island (Sapuka Besar) in the period 2000-2010 (Figure 6.6). The marriage activity has a slightly larger impact than the fishing activity as the average and maximum cumulative incidence rates for the period 2000-2010 are higher. In addition, the minimum cumulative incidence rate for the marriage activity (based on 50 model iterations) is larger than zero, while it remains zero for the fishing activity. This means that the addition of the marriage activity to LEPRASIM definitely leads to a reintroduction of the disease on the test-island, while this is not the case for the fishing activity. The combination of the fishing and marriage activity leads to a slightly larger cumulative incidence rate on Sapuka Besar. An inspection of the maximum cumulative incidence rate for the combination of the two activities (3.82) shows that the effect of the two activities adds up to a certain extent, as this value is 2.21 for the marriage activity and 1.95 for the fishing activity (which adds up to 4.16).

Table 6.4 Effect Activities on CIR 2000-2010

Activity	Cumulative Incidence Rate 2000-2010	Number of Events	Increase in CIR per event
Marriage	+109	28	+0.03892
Fishing	+0.87	4163	+0.00021

On average 28 marriages between male inhabitants of the island and a female agent from another island, leading to a movement of the female agent to the island, occur in the period 2000-2010 (based on 50 model iterations). The average total number of "fishing trips", leading to a potential contact with an infectious fisherman from another island, fishermen on the island make in this same period is 4163. Translated to an effect per event (see Table 6.4) we see that one marriage event has a much larger effect than one fishing event.

This difference is explained by the fact that a fishing activity leads to a "one-off" contact of one agent with 9 other agents on a fishing boat, in which an infection can occur, while a marriage activity leads to an introduction of an additional agent(s) on the island, which from that point on can lead to an infection, or be infected. As the percentage of the population that is susceptible to leprosy is relatively low (30%), the permanent redistribution of susceptible agents across the islands, caused by the marriage activity, has a greater influence on the cumulative incidence rates, than the contact between agents on fishing boats.

The reintroduction of leprosy in the control and blanket group of the study by Bakker et al. (2006) can thus be explained by both inter-island marriages and contact of fishermen on fishing boats. The effect of the inter-island marriages on this re-emergence of leprosy is, however, much larger and of a more deterministic nature. The unexpected re-emergence of leprosy can thus best be explained by a reintroduction of the disease through inter-island marriages.

6.3 Experiment 2: Intervention Modeling

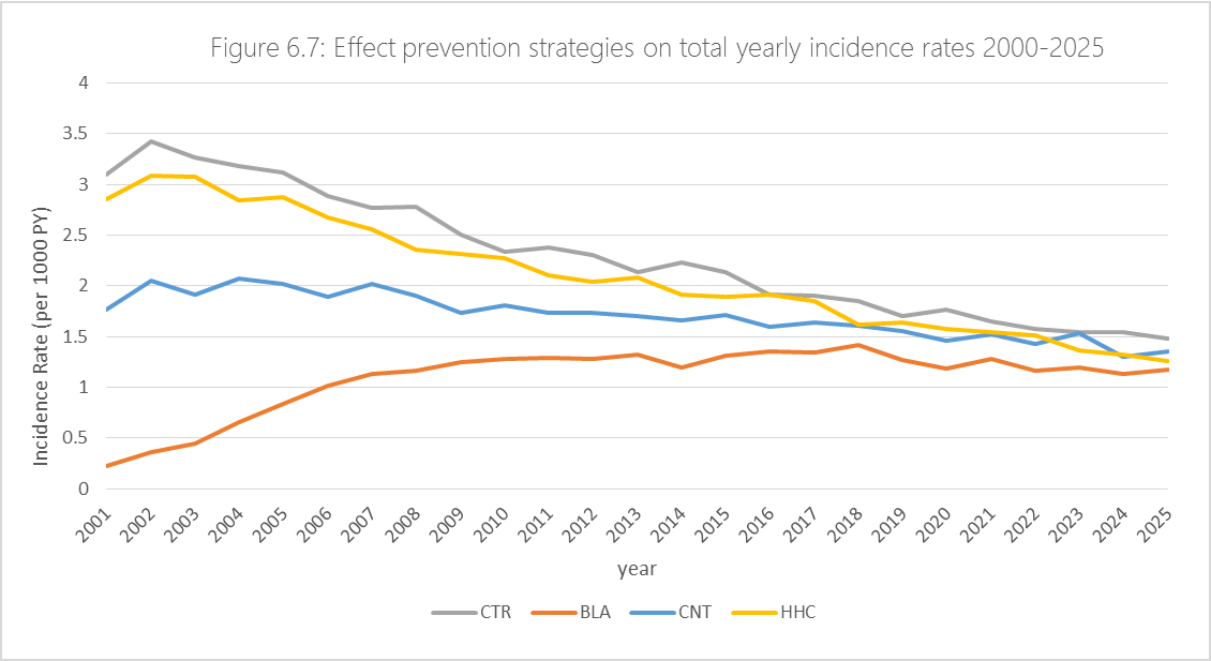
6.3.1 Experiment 2a: Effect of prevention strategies on spatio-temporal diffusion of leprosy

The LEPRASIM-variant from experiment 1 which includes both marriage as well as fishing activities is used in the second experiment, as addition of both these activities explained the re-emergence of leprosy in the blanket and contact group best.

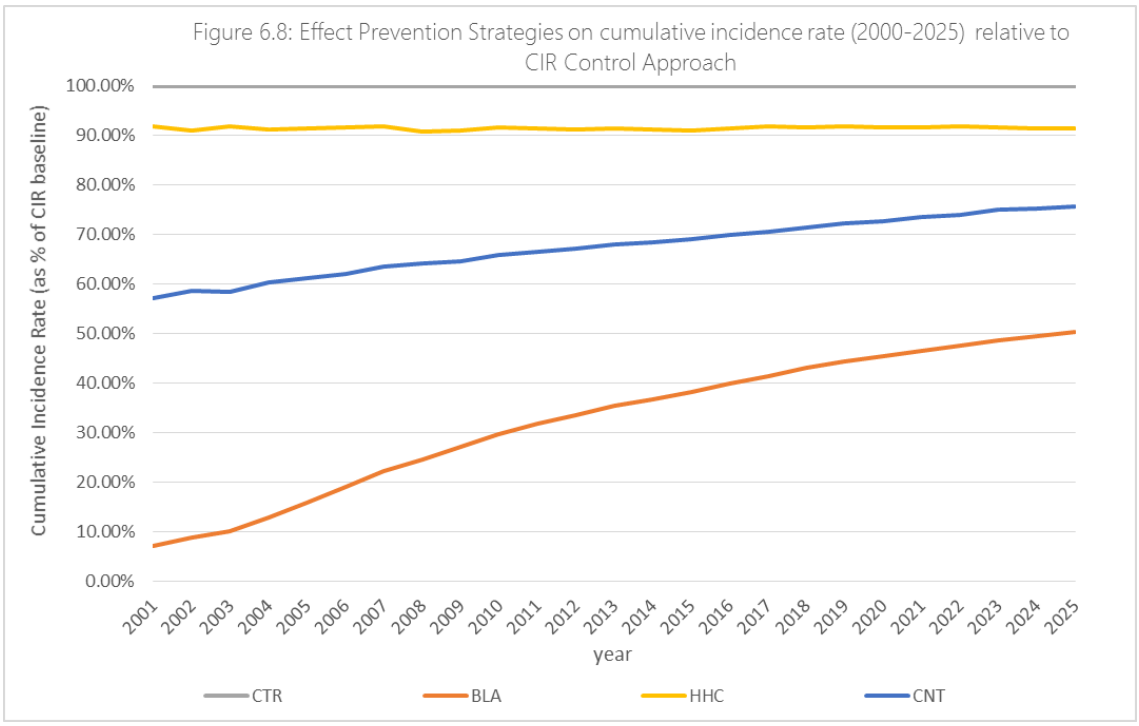
In the first part of this second experiment the effect of the following prevention strategies (as identified and researched by Bakker et al. (2006)) on the reduction in incidence rates in the period 2000-2025, relative to the baseline model performance, is measured:

- Control approach (CTR): no prevention strategy/ baseline model performance
- Blanket approach (BLA): prevention strategy aimed at entire population
- Contact approach (CNT): prevention strategy aimed at household and neighbor contacts of symptomatic patients
- Household Contact Approach (HHC): prevention strategy aimed at household contacts of symptomatic patients

The prevention strategies "work" by changing the state of all asymptomatic individuals, affected by the prevention strategy, to "not infected" at the moment the prevention strategy is deployed (tick 485/ the year 2000).



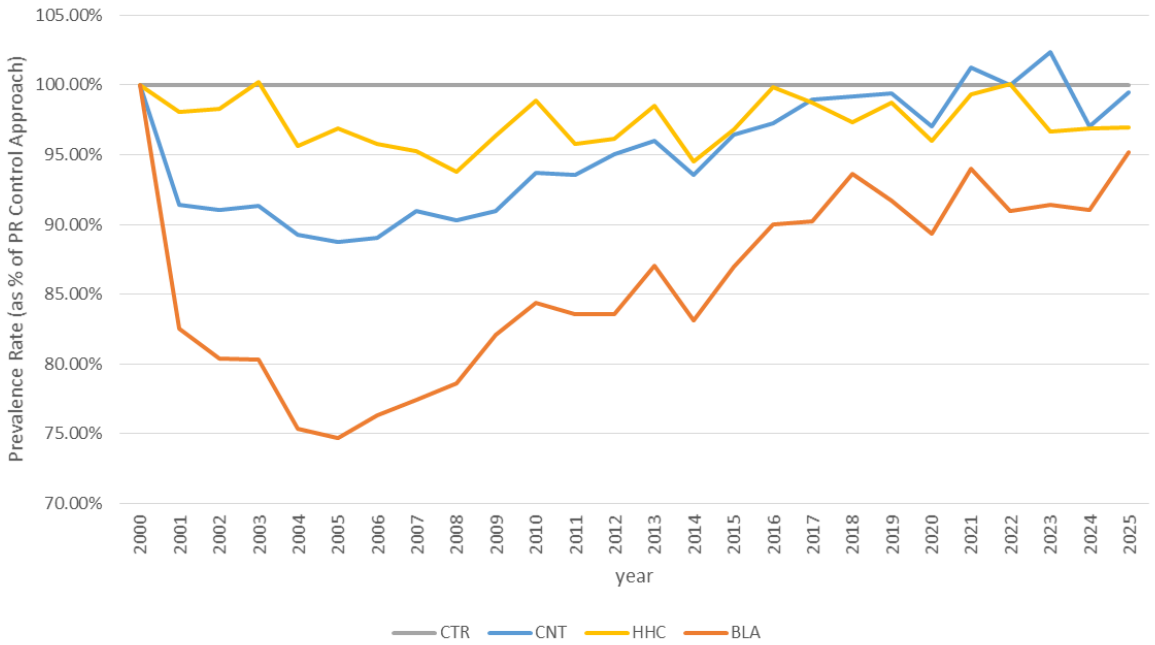
In Figure 6.7 the development in yearly incidence rates in the period 2000-2025 is shown. The blanket approach leads to the greatest immediate reduction in the yearly incidence rate (0.25 in 2000-2001), followed by the contact approach (1.75 in 2000-2001). Where this yearly incidence rate increases for the blanket approach as time goes by (period 2001-2013), the contact approach appears to have a more stable effect. Looking at the household approach, an even more stable, although small, effect can be observed, continuing onto 2025. A prevention strategy aimed at household contacts of symptomatic individuals thus has the most stable effect on the yearly incidence rates, followed by the contact and blanket approach respectively. In an absolute sense the blanket approach has the biggest effect, especially on the short term. This effect diminishes over time, where the effect of the other two prevention strategies remains more constant.



In Figure 6.8 the development of the cumulative incidence rates over the same time-period is shown, as a percentage of the cumulative incidence rate observed in the control group, providing an insight into the development of the effect off the different prevention strategies over time. Where the blanket approach (orange line) leads to an immediate reduction of the cumulative incidence rate to approximately 8% in the first year after the intervention is employed, this effect gradually decreases to approximately 50% in the year 2025. For the contact approach a similar development is observed, although less pronounced: an increase from 58% (2001) to 76% (2025). The household contact approach has a lasting effect on the incidence rate, as the cumulative incidence rate remains at 91% of the control approach over the whole time-period.

The effect of each of the prevention strategies on the total prevalence rates over the same time period (2000-2025) is shown in Figure 6.9. As can be seen, only the blanket approach (orange line) has a significant effect on the prevalence rate in 2025. The household contact approach has a very small diminishing effect on the total prevalence rate over time, while the contact approach has a larger effect on the short term, which decreases to (less than) zero as time goes on.

Figure 6.9: Effect Prevention Strategies on prevalence rate (2000-2025) relative to PR Control Approach



A one-time deployment of any of the prevention strategies does not lead to the desired reduction of the total prevalence rate in 2025, as all the lines (Figure 6.9) converge towards the prevalence rate of the control approach. As the prevalence rate in the control approach is itself rapidly declining as a result of the yearly treatment of all symptomatic people (see Figure 6.5), the prevention strategies no longer show a significant additional effect on this prevalence rate on the long term. The contact approach (on average) even leads to a higher prevalence rate in the period 2021-2023 than the control approach. This is due to the fact that asymptomatic patients are "treated" by the prevention strategy in 2000, effectively increasing the number of susceptible

people in the period there-after. On the short term (2000-2005) the prevalence rate is (greatly) reduced by the prevention strategies: a 25% and 9% decline for the blanket and control approach respectively. To investigate the possibility of deployment of the prevention strategies for the long-term, a number of extended strategies are developed and tested in the second part of this experiment.

6.3.2 Experiment 2b: Effectiveness of prevention strategies

As the blanket approach leads to the desired reduction in both incidence, as well as prevalence rates, the blanket approach is extended to attempt to achieve a more significant and permanent reduction of the incidence rate over time. After the blanket approach in 2000 a number of alternative extensions to the preventions strategy are examined:

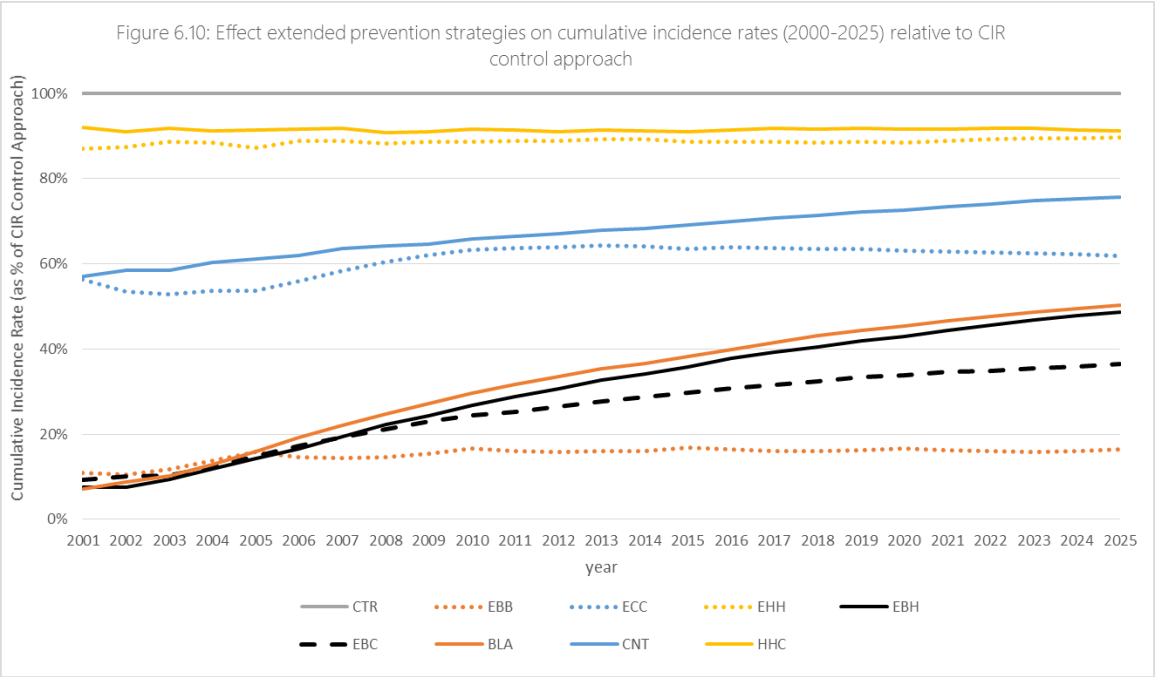
- Extended Blanket Approach (EBB): Blanket approach, followed by a five-yearly repetition of this blanket approach (2000, 2005, 2010, 2015, 2020).
- Household Contact Extend Blanket Approach (EBH): Blanket approach (2000), followed by the household contact approach at each detection (2000-2025).
- Contact Extended Blanket Approach (EBC): Blanket approach (2000), followed by the contact approach at each detection (2000-2025).

The first alternative is the repetition of this prevention strategy with a 5-year interval. The second alternative is the deployment of the household contact approach at the detection of each symptomatic leprosy patient through contact tracing of this patient in the period after 2000. In the third alternative not only the household contacts of these new patients are given the pre-emptive medication, but also the neighbor contacts.

In addition the potential of the household contact and contact approach on the long-term is examined by deploying these strategies at detection of a leprosy patient, through contact tracing of all new symptomatic patients in the period 2000-2005. This means the following two prevention strategies are deployed:

- Extended Household Contact Approach (EHH): Household contact approach (2000), followed by the household contact approach at each detection (2000-2025)
- Extended Contact Approach (ECC): Contact approach (2000), followed by the contact approach at each detection (2000-2025).

The effect of each of all eight strategies on the cumulative incidence rates over time, relative to the control approach, is shown in Figure 6.10.



Extended (Household) Contact Approach as Prevention Strategy

Looking at the extended variant of the household contact approach (Figure 6.10: EHH) a small stable reduction in cumulative incidence rates over the entire period can be observed, when compared to the household contact approach (HHC: yellow line). The total reduction in cumulative incidence rate is however very small and does not diverge from the variant without contact tracing. This is explained by a decrease in the influence of household contact on the incidence rate, given the declining prevalence rate (Figure 6.10), due to the intervention employed in the study area. As the number of infected people declines, so does the incidence rate. As the intimacy factors remain constant, this means that an increasing percentage of the new infections do not occur within households, but through contact with neighbors/ other island members or on fishing boats, as an individual agent has a higher number of contacts outside his household than within his household. Given the decreasing prevalence, the relative importance of the agent's household contacts to the probability the agent is infected thus decreases. This means that the household contact approach (with or without contact tracing) cannot be employed as a leprosy prevention strategy on the long-term, as it does not lead to a continuing reduction in incidence rates.

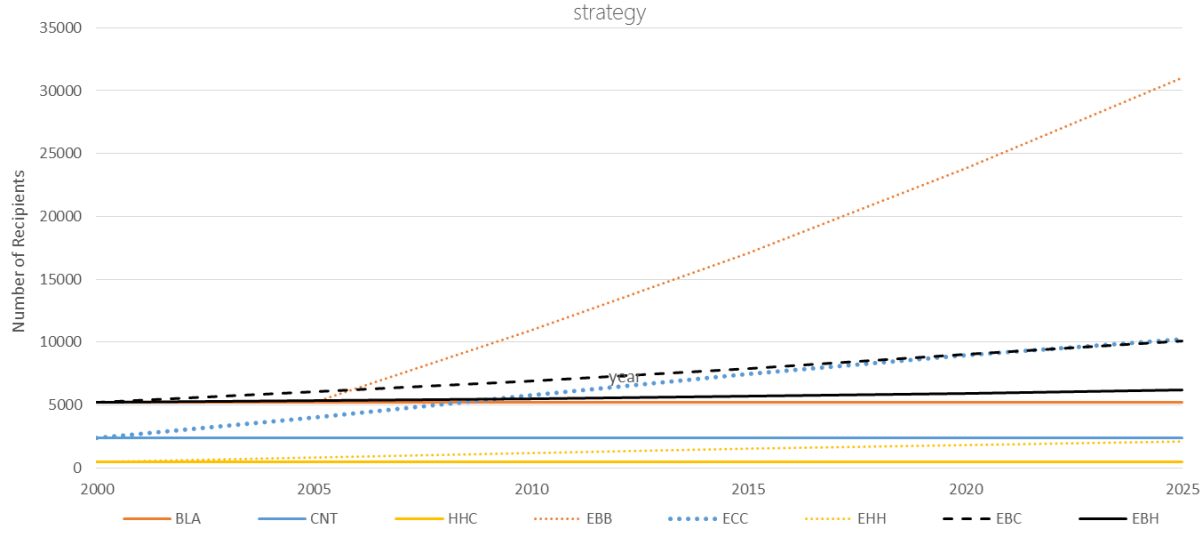
A continuation of the contact approach, aimed at both household and neighbor contacts of symptomatic individuals (Figure 6.10: ECC), leads to a stable reduction in cumulative incidence rates, when compared to one-of deployment of this contact approach (Figure 6.10: CNT). Where the cumulative incidence rate for the contact approach (CNT) converged towards the control approach (CTR) over time, a continuous deployment of this contact approach (ECC), through contact tracing of detected symptomatic individuals, stabilizes the reduction in cumulative incidence rate to around 60% of the control approach over time. This means that the contact approach, as defined by Bakker et al. (2006) is indeed a good strategy to reduce leprosy incidence rates over a longer time period.

Extended Blanket Approaches:

A comparison of the 5-yearly deployment of the blanket approach (Figure 6.10: EBB) with the one-of deployment of this strategy (BLA) shows that the reduction in cumulative incidence rate obtained by the deployment of the prevention strategy in 2000 (to approximately 18%) is pertained by the repetition of this strategy at five year intervals. This is logical, as the entire not infected population receives the rifampicin chemoprophylaxis once per 5 years. New cases that due arise, are due to relapses of recovered patients, and infections that occur within the period between an agent becoming symptomatic and his treatment (one year), as the population is screened for signs of leprosy on a yearly basis.

The deployment of the blanket approach in 2000, followed by the household contact approach through contact tracing of each newly detected patient (Figure 6.10: EBH) further confirms the observation made about this prevention strategy: the household contact approach does not lead to a significant reduction in cumulative incidence rates, but has a small stable effect. Employing the contact approach through contact tracing of each newly detected patient after the blanket approach is deployed in 2000 (Figure 6.10: EBC) does lead to a stabilization of the reduction of the cumulative incidence rate at approximately 40%, when compared to the control approach (CTR). This means that the effect of the blanket approach on the cumulative incidence rate is pertained (to a certain extent) by actively treating household and neighbor contacts of newly detected patients.

Figure 6.11: Cumulative number of recipients of rifampicin chemoprophylaxis (2000-2025) per prevention strategy



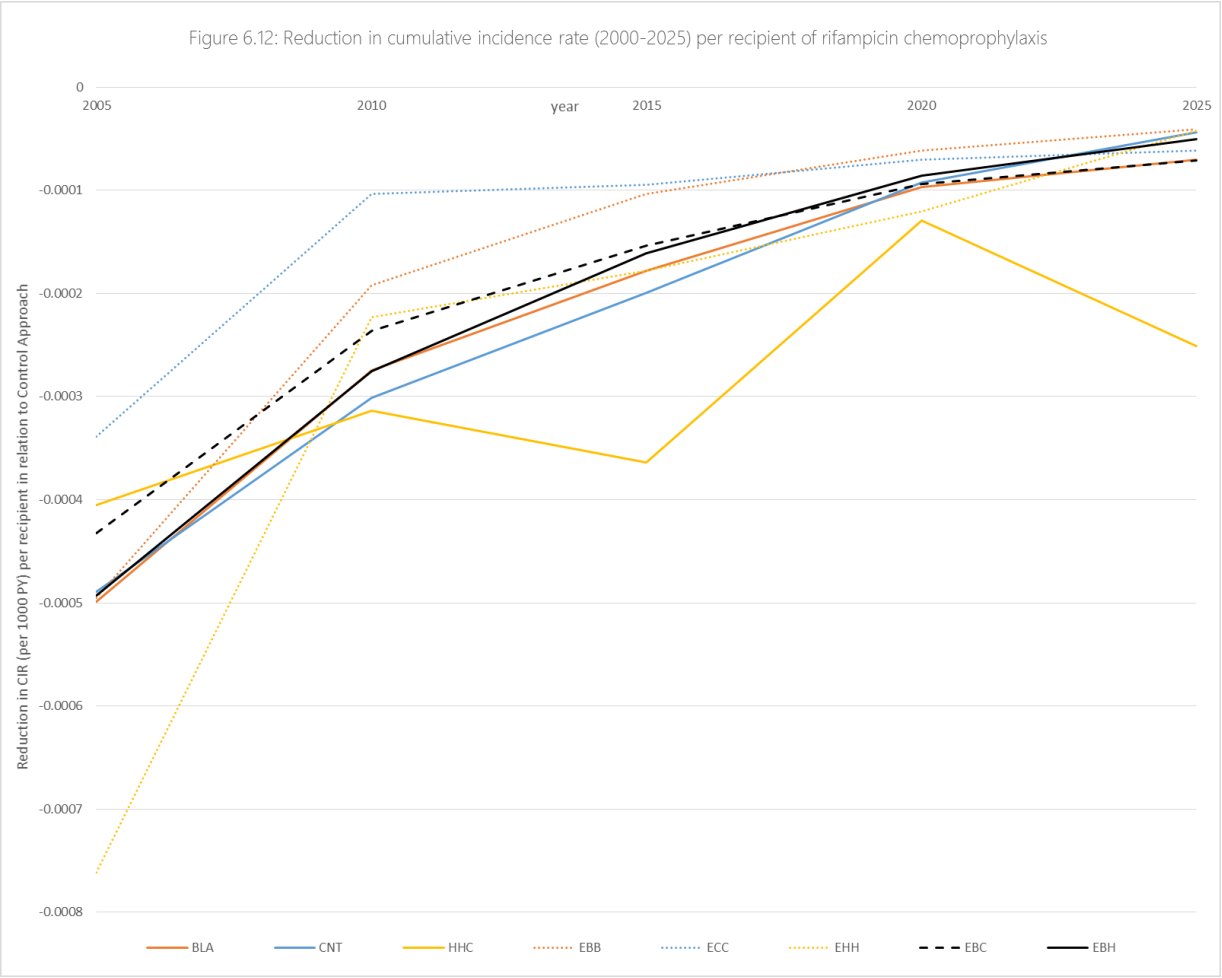
The effect of each of the prevention strategies should however be seen in the light of the number of patients receiving the pre-emptive medication. The number of recipients (of rifampicin chemoprophylaxis) for each of the prevention strategies is shown in Figure 6.11 (using 5-year intervals). This Figure shows that in the period 2000-2025 approximately the same number of people are given the pre-emptive medication in the contact approach (CNT) and the extended household contact approach (EHH). The same applies for the extended contact approach (ECC)

and the blanket approach, followed by this contact approach (EBC). Of the first pair, the contact approach leads to a greater reduction of cumulative incidence rate. For the second pair, the blanket approach followed by the contact approach (black dashed line) leads to a greater reduction of the cumulative incidence rate.

Table 6.5 Effectiveness Prevention Strategies				
Code	Prevention Strategy	No. of recipients (2000 - 2025)	Effect per recipient (reduction in CIR 2000-2025)	Reduction off CIR (2000-2025)
BLA	Blanket Approach	5236	$0.7 * 10^{-4}$	-11
HHC	Household Contact Approach	468	$2.5 * 10^{-4}$	-0.2
CNT	Contact Approach	2387	$0.4 * 10^{-4}$	-0.5
EHH	Extended Household Contact Approach	2133	$0.4 * 10^{-4}$	-0.2
ECC	Extended Contact Approach	10274	$0.6 * 10^{-4}$	-0.8
EBB	Extended Blanket Approach	31012	$0.4 * 10^{-4}$	-19
EBH	Household Contact Extended Blanket Approach	6193	$0.5 * 10^{-4}$	-11
EBC	Contact Extended Blanket Approach	10116	$0.7 * 10^{-4}$	-14

When the decrease in cumulative incidence rate is divided by the number of recipients of the pre-emptive medication the effectiveness of the different prevention strategies becomes apparent (Table 6.5; Figure 6.12). The one-of deployment of the household contact approach (HHC) has the largest effect per recipient of the medication (Figure 6.12: yellow line): a $\mathbf{2.5 * 10^{-4}}$ reduction in cumulative incidence rate for the period 2000-2025 (per recipient of the chemoprophylaxis in 2000), but has a very limited effect on the cumulative incidence rate itself (Figure 6.13) as only a low number of people receive the medication (Figure 6.11). A continuation of this strategy at each detection (EHH: yellow-dashed line in Figure 6.12) decreases the effect per recipient (to $\mathbf{0.4 * 10^{-4}}$), as more people receive the medication (Figure 6.11), while the relative influence of household contacts on the occurrence of new infections decreases with the declining prevalence rate. In the period 2000-2005 this prevalence is still rather high (see Figure 6.5), which is confirmed by the relatively large effect per recipient for the extended household contact approach (yellow dashed line) in this period.

The blanket approach (BLA) is more effective than the contact approach (CNT) ($\mathbf{0.7 * 10^{-4}}$ and $\mathbf{0.4 * 10^{-4}}$ effect per recipient respectively; Figure 6.12), which means that the blanket approach has a much larger effect on the cumulative incidence rate (mean cumulative incidence rate of 1.2 in comparison to 1.7) (Figure 6.13), especially as much more people receive the medication (Figure 6.11; Table 6.5).



The extended strategies (EBB, ECC, EHH, EBC & EBH) all have a lower effect per recipient than the blanket approach (BLA), as the effectiveness of the pre-emptive medication decreases with the prevalence rate. An extension off the blanket approach with a household contact or contact approach at detection (EBH, EBC) reduces the effect per recipient (to $\mathbf{0.5 * 10^{-4}}$ and $\mathbf{0.7 * 10^{-4}}$ respectively (Table 6.5). The Household Extended Blanket Approach (EBH) however has no effect on the cumulative incidence rate 2000-2025 (see Figure 6.13). This effect is noticeable in the Contact Extend Blanket Approach (EBC) (Figure 6.13; mean CIR blanket approach: 1.2, blanket approach + contact approach at detection: 0.8)

The Extended Contact Approach (ECC) leads to a (much) lower effect per recipient (Figure 6.12: blue lines), but to a significant decrease in the cumulative incidence rate when compared to the Contact Approach (CNT) (Table 6.5: -0.8 instead of -0.5). The same applies for the blanket approach: the Extended Blanket Approach (EBB) leads to a much lower effect per recipient (Figure 6.12; Table 6.5), but to a significant decrease in cumulative incidence rates (Figure 6.13)

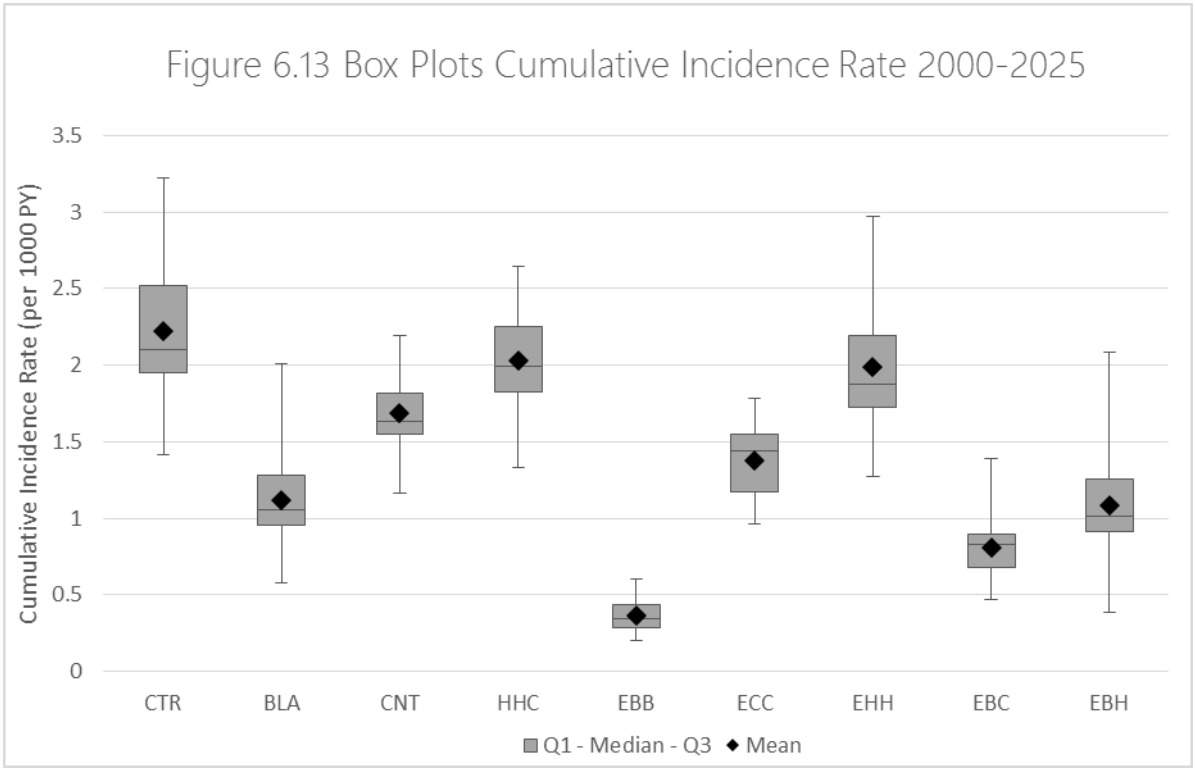


Figure 6.13 Box Plots Cumulative Incidence Rate 2000-2025
Cumulative Incidence Rate (per 1000 PY)
3.5
3
2.5
2
1.5
1
0.5
0
CTR
BLA
CNT
HHC
EBB
ECC
EHH
EBC
EBH
Q1 - Median - Q3
Mean

Conclusion

In this thesis research, the spatiotemporal dynamics of leprosy infections were modeled using an Agent-Based approach called LEPRASIM to gain a greater insight into these dynamics. In addition, the effect of explicitly modeled activities and prevention strategies (rifampicin prophylaxis) aimed at different contact groups of symptomatic patients on this phenomenon were explored. Three research questions were posed:

1 Can the spatio-temporal diffusion dynamics of leprosy be modeled using an Agent-Based approach and does this approach provide a greater insight into these dynamics?
2 Which of the identified reasons for the unexpected re-emergence of leprosy in the blanket and contact group of the study by Bakker et al. (2006) best explains this phenomenon?
3 At which contact group(s) of infectious individuals should a leprosy prevention strategy using rifampicin prophylaxis be aimed to be most effective?

LEPRASIM is tested on a case study for a group of Islands in the Flores Sea of Indonesia, on which Bakker et al. (2002, 2004, 2005a, 2006) performed an experiment using a number of leprosy prevention strategies. LEPRASIM uses a SEIR-model structure at an individual level, in which leprosy infections are modeled within agent-specific contact-groups: household, neighbor, island and fishing contacts. The hazard ratios for each of these contact groups, identified by Bakker et al. (2006) & van Beers et al. (1999), were translated to contact-group specific intimacy factors.

The model is presented using the extended ODD-protocol (Grimm et al., 2010), making the model verifiable and reproducible. Through an extensive sensitivity analysis a greater insight into the models behavior was obtained. The main conclusions of this sensitivity analysis are that for the diffusion of leprosy in a population the percentage of patients developing the multibacillary (MB) variant of the disease and the percentage of the population susceptible to the disease are the primary determinants off the prevalence and incidence rates. As only MB patients are considered to be infectious and only susceptible agents can be infected, this behavior was expected. Furthermore, the neighbor contacts of an individual have the strongest influence on the probability that this individual is infected. Although the relative influence of household contacts on this probability of infection is greater, the larger number of neighbor contacts leads to a bigger absolute influence.

After a three-step calibration process, LEPRASIM is validated using data on the incidence rate in the control group for the period 2000-2010 (Bakker, 2010). It is shown that LEPRASIM produces a cumulative incidence rate, following the trend in the observation made by Bakker (2010), LEPRASIM correctly simulates the clustering of patients in contact groups, but underestimates the influence of household size on the clustering of patients in households. This is due to the fact that susceptibility to leprosy is allocated randomly in LEPRASIM, while in reality a genetic factor is of influence. Furthermore, the clustering of patients on islands (measured by the average minimum and maximum prevalence rate per island over 50 model iterations) is very similar to the observation made by Bakker et al. (2002) in the study area.

Two possible explanations for the unexpected re-emergence of leprosy in the blanket and contact group of the study by Bakker et al. (2006) were examined by extending the LEPRASIM model (experiment 1) to include marriage and fishing activities: (1) a reintroduction of the disease through cross-island marriages and (2) a reintroduction of the disease through contact of fishermen from different islands on fishing-boats. Both explanations were proven to be valid, as both led to an increase in cumulative incidence rate (2000-2010). Of the two explanations, the inter-island marriage activity has the biggest impact on the cumulative incidence rate in an absolute sense (+1.09 compared to +0.87 for the addition of the marriage activity) as well as in an effect per activity (marriage: +0.039; fishing: +0.00021). Furthermore, the effect of the marriage activity is more deterministic, as the addition of the marriage activity leads to an increase of the minimum recorded incidence rate (from the year 2008) for all model iterations, while this is not the case for the fishing activity.

To determine at which contact group(s) a leprosy prevention strategy using rifampicin prophylaxis should be aimed to be most effective, LEPRASIM was extended to include a set of eight prevention strategies (Table 6.5). The effectiveness of a prevention strategy is two-fold: the reduction in cumulative incidence rates (2000-2025) and the reduction in this cumulative incidence rate (2000-2025) per recipient of the medication. In an absolute sense the Extended Blanket Approach (EBB) proved to be most effective: where the control approach (CTR) showed a cumulative incidence rate of 2.3 per 1000 person years for this time-period, EBB showed a CIR of only 0.4 (-1.9). The Contact Extended Blanket Approach (EBC) has a significantly larger effect than the Household Contact Extend Blanket Approach (EBH) (-1.4 compared to -1.1). Compared to the one-of deployment of the blanket approach in 2000 (BLA), the EBH-strategy does however not lead to an additional decrease in this CIR.

The effect per recipient of rifampicin prophylaxis on the CIR 2000-2025 shows that the Household Contact Approach (HHC) has the largest and most stable effect ($-2.5 * 10^{-4}$ per recipient), followed by strategy Blanket Approach (BLA), the Contact Extended Blanket Approach (EBC) ($-0.7 * 10^{-4}$per recipient) and the Extended Contact Approach (ECC) ($-0.6 * 10^{-4}$per recipient). A prescription of medication to household contacts of infected individuals thus leads to the greatest effect per recipient. This is however only the case when the prevalence rate is high, as the effect per recipient declines rapidly with a decrease in prevalence rate. Furthermore, the number of people receiving the medication is too low to have a significant impact on the cumulative incidence rate, as a large part of the infectious patients is in the asymptomatic stage of the disease and can thus not be detected by means of physical examination. The HHC or EHH-approaches are thus not suitable to reduce leprosy incidence over a longer time period.

Of the three prevention strategies (BLA, EBC & ECC), which showed the highest effectiveness per recipient (taking HHC out of the equation), the EBC-approach shows the biggest decrease in cumulative incidence rates (Table 6.5) (-1.4), followed by the BLA-approach (-1.1) and the implementation of the ECC-approach (-0.8). In an area where leprosy is highly endemic, as is the case in the study area in the year 2000, the deployment of a prevention strategy which starts with

a blanket approach thus is the most effective. As this blanket approach leads to a rather large immediate reduction in incidence rates, the approach can best be followed by a contact approach aimed at both household as well as neighbor contacts of newly detected patients, as this strategy yields the highest effect per recipient of the medication (Table 6.5: EBB, EBH & EBC).

Discussion

LEPRASIM simulates the transmission of leprosy at an individual level, effectively translating the work done in the SIMLEP-model (Meima et al., 2004) and SIMCOLEP-model (Fisher et al., 2011) to an Agent Based Model. By making the transition to an agent-based approach the heterogeneity of the population in the probability of infection, determined by the infectiousness of their household, neighbor and island contacts (Bakker, 2005a) has been captured within the model. In the development of LEPRASIM, or any other model, assumptions on factors determining the occurrence of a phenomenon were made with the greatest care. In this discussion the most important assumptions are discussed.

The model's behavior contradicts the assumption made by Richardus et al. (2005) that as leprosy is transmitted through direct contact, and the total number of leprosy patients is decreasing, an increasing percentage of new cases are resulting from contacts within households. LEPRASIM actually shows that with a decrease in the prevalence rate the effect of a prevention strategy aimed at these household contacts decreases as well. This means that the percentage of new cases resulting from these household contacts is also decreasing. LEPRASIM thus shows that the percentages of new cases resulting from neighbor or other contacts increase with a decreasing prevalence rate: the spatio-temporal diffusion dynamics of leprosy become more random. This means that the higher the prevalence of leprosy in a population is the more effective a strategy aimed at direct contacts of infected individuals will be.

Within LEPRASIM, neighbor contacts are modeled by assigning the five nearest agent-collectives (households) as a neighbor to each agent-collective. This means that, within LEPRASIM, each household has the same number of neighboring households in case all households are occupied. As the households get occupied over time, due to the increasing population size, the influence of this neighbor contacts on the probability of infection is underestimated for the period 1960-2000. In the period after 2000, in which the experiments take place, the model behaves correctly.

In reality, not all houses have the same number of neighbors within a predefined distance. Especially since not all houses are randomly distributed over the islands (as is the case in LEPRASIM), but rather tend to concentrate along the shores of the island. An addition of the spatial dimension to the definition of the neighbor contacts would greatly improve the insight into the influence of these neighbor contacts on the probability of a leprosy infection. In order to make LEPRASIM geographically explicit, data on the location of each house is required. This data is not (yet) available for the case study used. An application of the LEPRASIM model to a different study area, for which this data is available, would greatly improve LEPRASIM.

The focus of this research has not been on the factors determining heterogeneity in susceptibility to leprosy at an individual level, as this effect has been extensively researched, using the SIMCOLEP-model (Fisher et al., 2011). As genetic factors are known to explain up to 57% of the

variation in susceptibility to leprosy (Bakker et al., 2005) an incorporation of the heritability mechanisms, as identified by Fisher et al. (2011), into LEPRASIM is expected to further improve the model. This genetic factor is suspected to be the explanation for the underestimation of the influence of household size on the clustering of leprosy within these households in LEPRASIM, but further research on the matter is required. As the study by Pönnighaus et al. (1994) showed a negative correlation between housing conditions and the risk of leprosy infection, the heterogeneity in susceptibility to leprosy might be caused by differences in these housing conditions as well. LEPRASIM offers a platform to further investigate this matter.

The prevention strategies examined within LEPRASIM are all deployed in addition to the intervention performed in the study area, which entails a yearly screening of the entire population on signs of leprosy. When a symptomatic patient is detected in this screening, he is treated immediately (and transitions to the "recovered" stage). The effect of this intervention, which is necessary for the contact tracing of these infected individuals, obscures the effect of the prevention strategies on both the prevalence rate, as well as the incidence rate. A variation on the model, in which the population is screened at a less frequent interval, might thus lead to differences in the magnitude of the effects of the prevention strategies. Furthermore, LEPRASIM assumes a "perfect" situation, in which each infected individual is detected and treated and each contact is eligible to receive the pre-emptive antibiotics. Although this is not the case in reality, as people miss their doctors' appointments, hide their leprosy scars or are not eligible to receive antibiotics due to for example pregnancy, the distribution of these cases over the households and islands can be considered to be random.

So far, the effect of social stigma on leprosy incidence and prevalence rates has proven hard to quantify (Bakker, 2005a). As LEPRASIM models each individual agent, activities related to social stigma can be introduced to the model. The effect of segregation of leprosy patients can for example be modeled by an isolation of symptomatic agents in LEPRASIM in a separate household or island, as was done on a national scale from 1655 until 1932 in Indonesia (Peters et al., 2013).

LEPRASIM has been calibrated using data on the cumulative incidence rate for the control group in the study by Bakker et al. (2006) for the period 2000-2010. An extended calibration of the model including data on the cumulative incidence rates for the contact and blanket group of this same study (Bakker et al., 2006) for the same time-period would have led to a refinement of the intimacy factors applied in LEPRASIM. This data was however not available during this thesis research.

An important assumption has been made on the effect of rifampicin prophylaxis: within the experiment conducted in LEPRASIM this effect is limited to the treatment of infected asymptomatic patients. Bakker (2007) however identifies three possible effects of this medication: (1) Chemoprophylaxis only delays the development of leprosy; (2) Chemoprophylaxis prevents leprosy, but only has a temporal effect on transmission of the disease; (3) Chemoprophylaxis prevents leprosy and also reduces the transmission;

An examination of the data on the cumulative incidence rates in the blanket group of the study by Bakker et al. (2006) for the period 2000-2010 will provide a first insight into what the actual effect of the rifampicin prophylaxis is. In the first case, the cumulative incidence rates in the blanket

group will have risen to the same level as the control group. In the second case the yearly incidence rates will have risen to the same level after a shorter time-period, in the third case the two incidence rates will have continued to diverge. Once the effect has been determined, LEPRASIM offers a platform to further investigate the long-term effect of the different prevention strategies using this medication.

LEPRASIM is the first attempt at an Agent Based Model for modeling the spatio-temporal diffusion dynamics of leprosy. Hereby, the added value ABMs bring to disease modeling, so beneficiary to the study of other diseases (Rodrigues et al., 2015; Crooks & Hailegiorgis, 2014; Linard et al., 2009) ,has been brought to the study of leprosy. The added value of an ABM to disease modeling lies in its capability to model the heterogeneity of a population on characteristics and behavior relevant to the spread of that disease at an individual level, and its spatially explicit nature. By incorporating agent-specific contact-groups though multiple levels (household, neighbor and island) and activities (marriage and fishing), evolving over time, LEPRASIM is a robust agent-based structure for modeling infections, resulting from contacts within various sub-populations. The heterogeneity in the population, most relevant to leprosy infections, has been captured through the modeling of agent-specific contact groups.

Within LEPRASIM, the spatial dimension, so relevant to leprosy infections, has been captured implicitly via the modeling of households, within a neighbor (and island) structure. By making this spatial dimension explicit, the value of LEPRASIM for examining the long-term effects of leprosy prevention strategies can be further improved. When the model is made spatially explicit, it can be applied to other areas of the world, rather easily, as the parameters of the model can be fit to any situation. The applicability of the model to other disease areas is limited to Influenza-Like-Infections (ILI) with a rather long incubation time, as LEPRASIM is a model for bacterial infections and uses one-month time-steps.

References

Bakker, M.I., Hatta, M., Kwenang, A., Klatser, P.R. & Oskam, L. (2002). Epidemiology of leprosy on five isolated islands in the Flores Sea, Indonesia. *Tropical Medical International Health, 7,* pp. 780-787.

Bakker, M.I., Hatta, M. Kwenang, A. Faber, W.R., van Beers, S.M., Klatser, P.R. & Oskam, L. (2004). Population survey to determine risk factors for Mycobacterium leprae transmission and infection. *International Journal Epidemiology*, *33*, pp. 1329-1336

Bakker, M.I., Hatta, M., Kwenang, A., van Benthem, B.H.B., van Beers, S., Klatser, P.R. & Oskam, L. (2006). Prevention of leprosy using rifampicin as chemoprophylaxis. *American Journal of Tropical Medicine and Hygiene, 72(4),* pp. 443-448.

Bakker, M.I. (2005a). Chapter 1: Introduction *In: Bakker, M.I. (2005) Epidemiology and prevention of leprosy: a cohort study in Indonesia.* pp. 9-54.

Bakker, M.I. (2005b) Chapter 7: General Discussion *In: Bakker, M.I. (2005) Epidemiology and prevention of leprosy: a cohort study in Indonesia.* pp. 141-152.

Bakker, M.I., Scheelbeek, P.F.D. & van Beers, S.M. (2009). The use of GIS in leprosy control. *Leprosy Review, 80,* pp. 327-331.

Bakker, M.I. (2016), Excel-file containing data on cumulative incidence rates of control group in study Bakker et al. (2006).

Van Beers, S., Hatta, M. & Klatser, P.R. (1999). Patient contact is the major determinant in incident leprosy: implications for future control. *International Journal of Leprosy and other mycobacterial diseases, 67(2)*, pp. 119-128.

Bian, L. & Liebner, D. (2005). Simulating spatially explicit networks for dispersion of infectious diseases. *In: Goodchild, M.D. & Batty, M. (2005) GIS: Spatial analysis and Modeling. California: ESRI Press,* pp. 245-264.

Blok, D.J., de Vlas, S.J., Fischer, E.A.J. & Richardus, J.H. (2015). Chapter 2: Mathemetical Modeling of Leprosy and its Control. *Advances in Parasitology,* 87, pp. 33-51.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences, 99 (3)*, pp. 7280-7287.

Colizza, V., Barrat, A., Barthélemy, M., & Vespignani, A. (2006). The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences*, *103(7)*, pp. 2015–2020.

Cree, I.A. & Smith, W.C. (1998). Leprosy transmission and mucosal immunity: towards eradication? *Leprosy Review, 67*, pp. 112-121.

Cromley, E.K. & McLafferty, S.L. (2002). GIS and public health. Guilford Press: New York.

Crooks, A. T., & Hailegiorgis, A. B. (2014). An agent-based modeling approach applied to the spread of cholera. *Environmental Modelling & Software, 62*, pp. 164-177.

Desikan, K.V. (1995). Extended studies on the viability of Mycobacterium Leprae outside the human body. *Leprosy Review, 66,* pp. 287-295.

Duthie, M.S., Saunderson, P., & Reed, S.G. (2012). The potential for vaccination in leprosy elimination: new tools for targeted interventions. *Memórias do Instituto Oswaldo Cruz, 107,* pp. 190-196.

Epstein, J.M., Cummings, D., Chakravarty, S., Singa, R. & Burke, D. (2004). Toward a containment strategy for smallpox bioterror: an individual-based computational approach. Washington D.C.: Brookings Institution Press.

Epstein, J.M. (2008). Why Model? *Journal of Artificial Societies and Social Simulation,* 11(4), pp. 1-12

Fine, P.E. (1982). Leprosy: the epidemiology of a slow bacterium. *Epidemiol Review, 4,* pp. 161-188.

Fine, P.E. (1988). BCG vaccination against tuberculosis and leprosy. *British Medical Bulletin, 44,* pp. 691-703.

Fischer, E., Pahan, D., Chowdhury, S., & Richardus, J. (2008). The spatial distribution of leprosy cases during 15 years of a leprosy control program in Bangladesh: an observational study. *BMC Infectious Diseases*, *8(1)*, Available online at: <http://bmcinfectdis.biomedcentral.com /articles/ 10.1186/1471-2334-8-126>

Fischer, E., Vlas, D.S., Meima, A. Habbema, D. & Richardus, J. (2010). Different mechanisms for heterogeneity in leprosy susceptibility can explain disease clustering within households. *PloS One 5(11), e14061.* Available online at: <http://journals.plos.org/plosone/article?id=10.1371/ journal.pone.0014061>

Fischer, E.A.J., de Vlas, S.J., Habbema, D.F. & Richardus, J.H. (2011). The long term effect of current and new interventions on the new case detection of leprosy: a modeling study. *PLoS Neglected Tropical Diseases, 5(9), e1330.* Available online at: <http://journals.plos.org/plosntds/article?id= 10.1371/journal.pntd.0001330>

Guay, J.S., Hino, P. & dos Santos, C.B. (2007). Spatial distribution of leprosy cases in Ribeirao Preto, Brazil, 2004. *Revista latino-americana de enfermagem, 15 (3),* pp. 460-465. Available online at: < http://www.scielo.br/scielo.php?pid=S0104-11692007000300015&script=sci_arttext&tlng=es>

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J. & Railsback, S.F. (2010), The ODD protocol: a review and first update. *Ecological Modelling, 221,* pp. 2760-2768.

Hassanai-Mahmooei, B. & Parris, B. W. (2013). Resource scarcity, effort allocation and environmental security: An agent-based theoretical approach. *Economic Modelling, 30*, pp. 183–192.

Hatta, M., van Beers, S.M., Madjid, B., Djumadi, A., de Wit, M.Y. & Klatser, P.R. (1995). Distribution and persistence of Mycobacterium leprae nasal carriage among a population in which leprosy is endimic in Indonesia. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 89.4*, pp. 381-385.

Helbing, D. (2012). Social Self-Organization, Understanding Complex Systems. *Chapter 2: Agent-Based Modeling*. Berling: Springer-Verlag, pp. 25-70

Holtrup, R., de Jonge, M. & Peters, M. (2015). Leprosy in the Flores Sea: An agent based model. University of Twente: unpublished.

Lechat, M.F., Misson, J.Y. & Vellut, C.M. (1974). An epidemetric model of leprosy. *Bulletin World Health Organisation, 51,* pp. 361-373

Lechat, M. F., Misson, C. B., Lambert, A., Bouckaert, A., Vanderveken, M., & Vellut, C. (1985*).* Simulation of vaccination and resistance in leprosy using an epidemiometric model. *International Journal of Leprosy, 53,* pp. 461-467.

Lechat, M.F. (1992). Epidemiometric modeling in leprosy based on Indian data. *Leprosy Review, 63,* pp. S1-39.

Lee, J., Filatova, T., Ligmann-Zielinska, A., Hassani-Mahmooei, B., Stonedahl, Forrest, Lorscheid, I., Voinov, A., Polhill, G., Sun, Z. & and Parker, D.C. (2015). The Complexities of Agent-Based Modeling Output Analysis. *Journal of Artificial Societies and Social Simulation, 18 (4),* pp. 4-18, Available online at: <http://jasss.soc.surrey.ac.uk/18/4/4.html#lorscheid2012>

Leiker, D.L. & Fischer, P. (1976). The incidence of leprosy between 1943 and 1973 in a hyperendemic area, before and after the introduction of leprosy control measures, *Leprosy Review, 47,* pp. 115-125.

Lienhardt, C. & Fine, P.E. (1994). Type 1 reaction, neuritis and disability in leprosy. What is the current epidemiological status? *Leprosy Review, 65*, pp. 9-33.

Linard, C., Ponçon, N., Fontenille, D., & Lambin, E. F. (2009). A multi-agent simulation to assess the risk of malaria re-emergence in southern France. *Ecological Modelling, 220(2),* pp. 160-174.

Lorscheid, I., Heine, B.-O. & Meyer, M. (2012). Opening the 'black box of simulations: Increased transparency and effective communication through the systematic design of experiments. *Computational and Mathematical Organization Theory*, *18*, pp. 22-62.

Macal, C.M. & North, M.J. (2006). Introduction to Agent-based Modeling and Simulation. *MCS Lnd Informal Seminar, November 29 2006*. Available online at: < http://www.mcs.anl.gov/~leyffer/listn/slides-06/MacalNorth.pdf>

Macal, C.M. & North, M.J. (2010). Tutorial on agent-based modelling and simulation. *Journal of Simulation, 4*, pp. 151-162.

Meima, A., Gupte, M.D., van Oostmarssen, G.J. & Habbema, J.D.F. (1999). SIMLEP: a simulation model for leprosy transmission and control. *International Journal of Leprosy and Other Mycobacterial Diseases, 67,* pp. 213-236.

Meima, A., Irgens, L.M., van Oortmarssen, G.J., Richardus, J.H. & Habbema, J.D. (2002). Disappearance of leprosy from Norway: an exploration of critical factors using an epidemiological modelling approach. *International Journal of Epidemiology, 31,* pp. 991-1000.

Meima, A., Cairns, W., Smith, S. van Oostmarssen, G.J., Richardus, J.H. & Habbema, J.D.F. (2004) The future incidence of leprosy: a scenario analysis. *Bulletin of the World Health Organization 2004, 82,* pp. 373-380.

Mira, M.T., Alcais, A., Nguyen, V.T., Moraes, M.O., Di Flumeri, C., Vu, H.T., Mai, C.P., Nguyen, T.H., Nguyen, N.B., Pham, X.K., Sarna, E.N., Alter, A. Montpetit, A. Moraes, M.E., Moraes, J.R., Dore, C. Gallant, C.J., Lepage, P., Verner, A., van de Vosse, E., Hudson, T.J., Abel, L. & Schurr, E. (2004). Susceptibility to leprosy is associated with PARK2 and PACRG. *Nature*, *427 (6975),* pp. 636-640.

Moet, F.J., Meima, A., Oskam, L. & Richardus, J.H.. (2004). Risk factors for the development of clinical leprosy among contacts, and their relevance for targeted interventions. *Leprosy Review, 75,* pp. 310-326.

NetLogo (2016). NetLogo 5.3.1 User Manual: Links. Available online at: <http://ccl.northwestern.edu/netlogo/docs/programming.html#links>

Noordeen, S. K. (1985). Vaccination against leprosy; recent advances and practical implications. *Leprosy review, 56(1),* pp. 1-3.

O'Toole, M. (1997). Miller-Keane Exncyclopedia & Dictionary of Medicine, Nursing & Allies Health, 6th ed. Philadelphia: Saunders, Available online at <http://medical-dictionary.the freedictionary .com/seropositive>

Perez, L. & Dragicevic, S. (2009). An agent-based approach for modeling dynamics of contagious disease spread. *International journal of health geographics, 8 (1),* pp. 1-50.. Available online at < https://ij-healthgeographics.biomedcentral.com/articles/10.1186/1476-072X-8-50>

Peters, R.M.H., Dadun, Lusli, M., Miranda-Galarza, B. Van Brakel, W.H., Zweekhorst, M.B.M., Damayanti, R. Seda, F.S.S.E, Bunders, J.F.G & Irwanto (2013). The meaning of leprosy and everyday experiences: an exploration in Cirebon, Indonesia. *Journal of Tropical Medicine, vol. 2013, Article ID 507034,* pp. 1-10.

Perra, N. & Goncalves, B. (2015). Chapter 4 Modeling and Predicting Human Infectious Diseases. *In: Social Phenomena. From Data Analysis to Model*. London: Springer International Publishing, pp. 59-84.

Penna, M.L., Wand-Del-Rey-de-Oliveira, M.L., & Penna, G. (2009). Spatial distribution of leprosy in the Amazon region of Brazil. *Emergent Infectious Diseases*, *15,* pp. 650-652.

Pönnighaus, J.M., Fine, P.E.M., Sterne, J.A.C., Bliss, L. Wilson, R.J., Malema, S.S. & Kileta, S. (1994). Incidence rates of leprosy in Karonga District, Northern Malawi: patterns by age, sex, BCG status and classification. *International Journal for Leprosy and Other Mycobacterial Diseases, 62,* pp. 10-23.

Rees, R.J.W. & Young, D.B. (1985). The microbiology of leprosy. *In: Hastings, R.C. & Opromolla, D.V.A. (1985) Leprosy.* Edingburgh: Churchil Livingstone, pp. 31-52.

Richardus, J.H., Meima, A., van Marrewijk, C.J., Croft, R.P., & Smith, T.C. (2005). Close contacts with leprosy in newly diagnosed leprosy patients in a high and low endemic area: comparison between Bangladesh and Thailand. *International Journal of leprosy and other mycobacterial diseases, 73(4),* pp. 249-257.

Ridley, D.S. & Jopling, W.H. (1966). Classification of leprosy according to immunity. A five-group system. *International Journal of Leprosy and Other Mycobacterial Diseases, 34(3),* pp. 255–273.

Rodrigues, L. C. & Lockwood, D. N. (2011). Leprosy now: epidemiology, progress, challenges, and research gaps. *The Lancet infectious diseases, 11(6),* pp. 464-470

Rodrigues, C. G., Espíndola, A. L., & Penna, T. J. P. (2015). An agent-based computational model for tuberculosis spreading on age-structured populations. *Physica A: Statistical Mechanics and its Applications, 428,* pp. 52-59.

Sermrittirong, S. & van Brakel, W.H. (2014). Stigma in leprosy: concepts, causes and determinants. *Leprosy Review,* 85, pp. 36-47.

Shephard, C.C. (1962). The nasal excretion of Mycobacterium leprae outside the human body. *Leprosy Review, 48,* pp. 10-18.

Smith, C.M. & Smith, W.C.S. (2000). Chemoprophylaxis is effective in the prevention of leprosy in endemic countries: a systematic review and meta-analysis. *Journal of Infection, 41(2),* pp. 137-142

Snow, J. (1854). On the mode of communication of cholera. Second edition. John Churchill: London.

Stanley, S.J., Howland, C., Stone, M.M., & Sutherland, I. (1981). BCG vaccination of children against leprosy in Uganda: final results. *Journal of Hygiene, 87(2),* pp. 233-248.

Voinov, A. (1999) Simulation Modeling, Online Course. Available online at: < http://www.likbez.com/AV/Simmod.html>

Wallinga, J., Edmunds, W.J., & Kretzschmar, M. (1999). Perspective: human contact patterns and the spread of airborne infectious diseases. *Trends Microbiology*, *7*, pp 372-377.

Woolf, S.H., Jonas, S & Kaplan-Liss, E. (2008). Health Promotion and Disease Prevention in Clinical Practice. Lippincot: Williams & Wilkins.

World Health Organization (1991). Elimination of leprosy: Resolution of the 44th World Health Assembly (WHA 44.9). WHO: Geneva

World Health Organization (1997). Shortening duration of treatment of multibacillary leprosy. *Weekly Epidemic Review, 73,* pp. 169-175.

World Health Organization (1998). WHO expert committee on leprosy – seventh report. *World Health Organization Technical Report Series, 874,* pp. 1-43.

World Health Organization (2003). Leprosy. Factsheet N101. Available online at <http://www.who.int/mediacentre.factsheets/fs101/en/>

World Health Organization (2009a). Enhanced Global Strategy for further reducing the disease burden due to leprosy 2011-2015. WHO, regional office for South-East Asia: New Delhi.

World Health Organization (2009b). Geographic Information Systems (GIS). Available online at <http://www.who.int/lep/monitor/gis/en/>

World Health Organization (2012). Accelerating work to overcome the global impact of neglected tropical diseases – a roadmap for implementation. WHO: Geneva.

World Health Organization (2015a). Leprosy elimination. Available online at <http://www.who.int/lep/transmission/en/>

World Health Organization (2015b). Global Leprosy Update, 2014: need for early case detection. *Weekly epidemiological record, 90, no. 36,* pp. 461- 476.

Appendix A: Scatter Plots

A1 Step 1: sensitivity of population model to activity model

A1.1 Input: Marry Rate

Figure A1.1: Marry Rate - Populaton Size (2000)

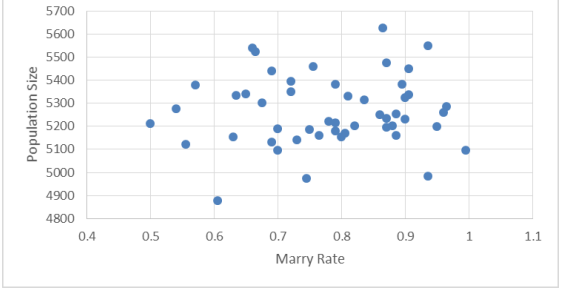


Figure A1.2: Marry Rate - Mean Hosuehold Size (2000)

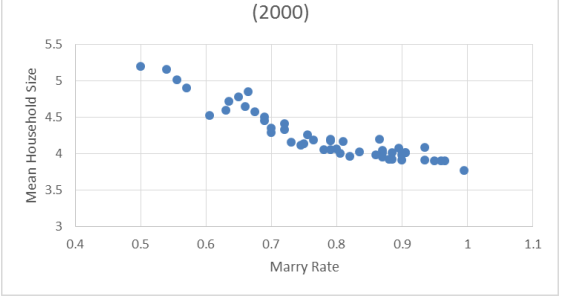


Figure A1.3: Marry Rate - Percentage under 10 (2000)

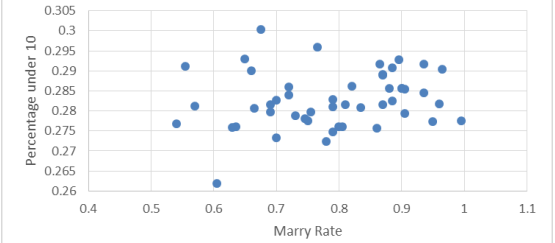


Figure A1.4: Marry Rate - Percentage male (2000)

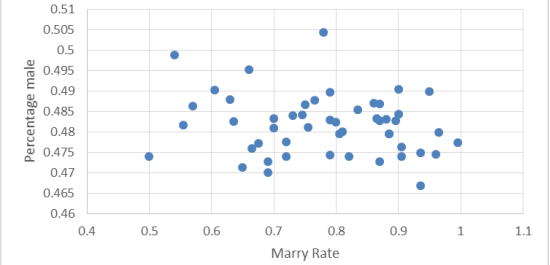
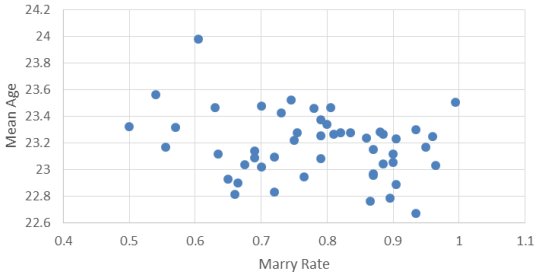


Figure A1.5: Marry Rate - Mean Age (2000)



A1.2 Input: Percentage Own House at Marriage

Figure A1.6 Percentage Own House at Marriage - Population Size (2000)

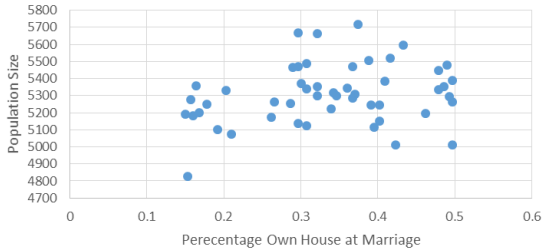


Figure A1.7 Percentage Own House at Marriage - Mean Household Size (2000)

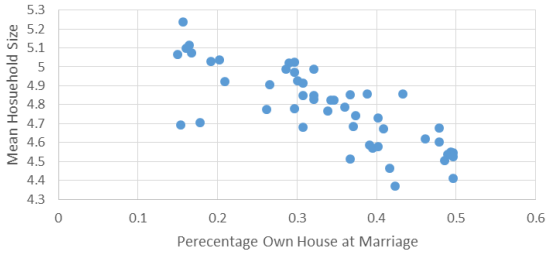


Figure A1.8 Percentage Own House at Marriage - Percentage under 10 (2000)

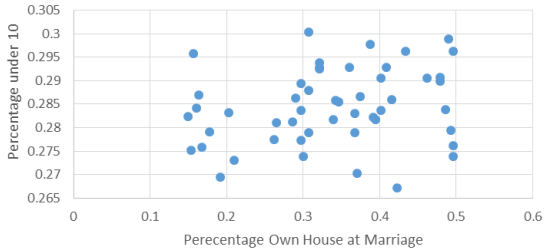


Figure A1.9 Percentage Own House at Marriage - Percentage Male(2000)

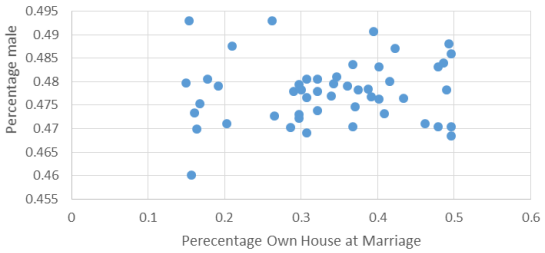
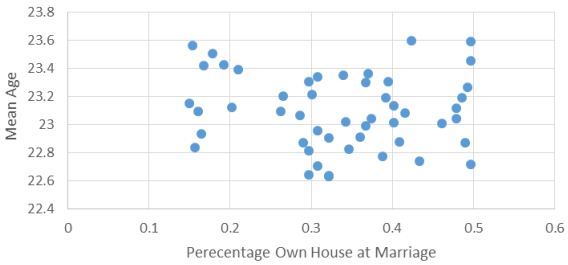
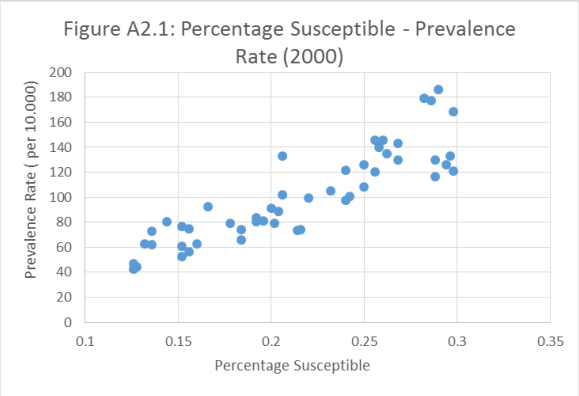


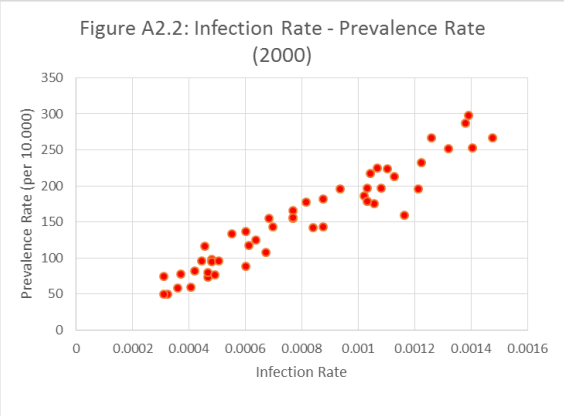
Figure A1.10 Percentage Own House at Marriage - Mean Age (2000)

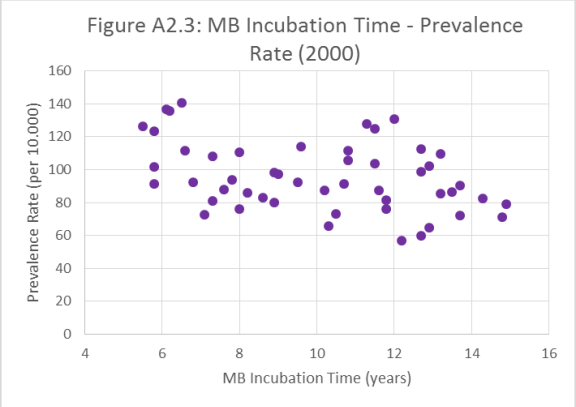


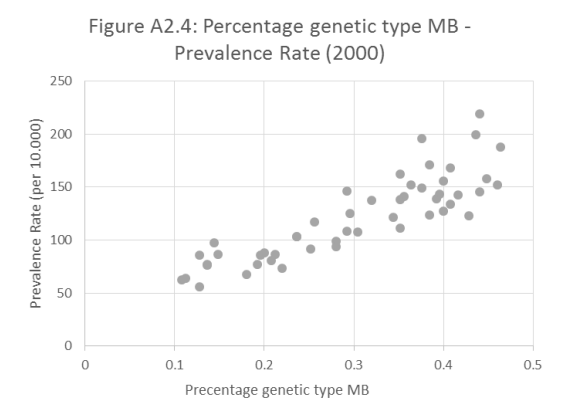
A2 Step 2: internal sensitivity of disease model A2.1 Output: Prevalence Rate

A2.1 Output: Prevalence Rate

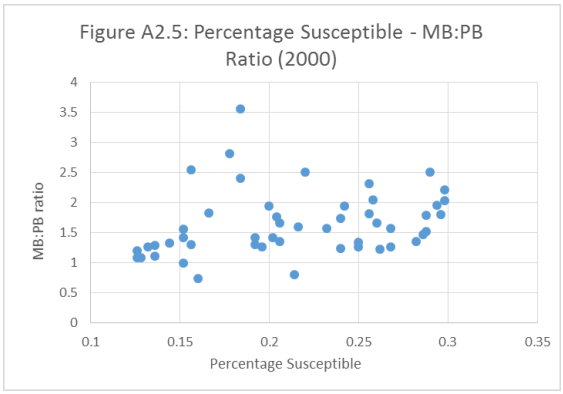


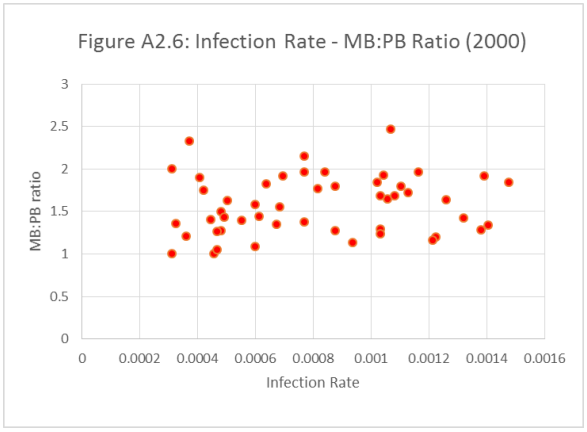


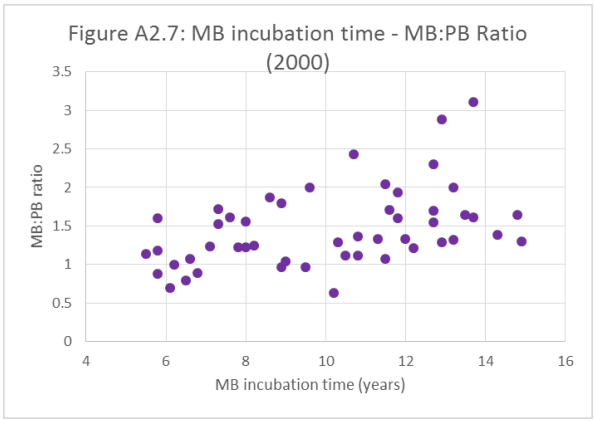


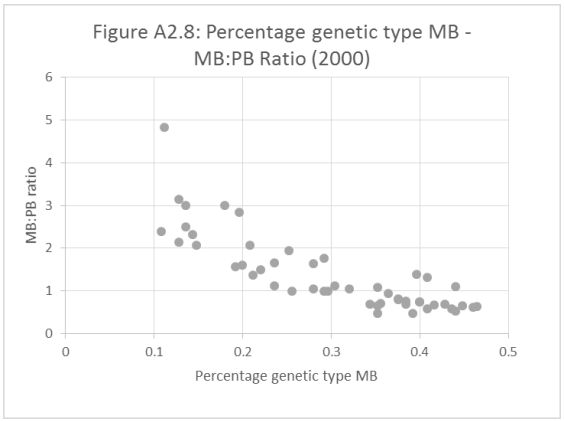


A2.2 Output: MB:PB-ratio









A2.3 Output: Incidence Rate

Figure A2.9: Percentage Susceptible - Incidence Rate (2000-2003)

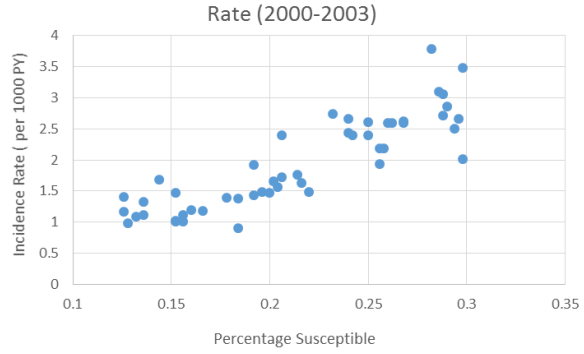


Figure A2.10: Infection Rate - Incidence Rate (2000-2003)

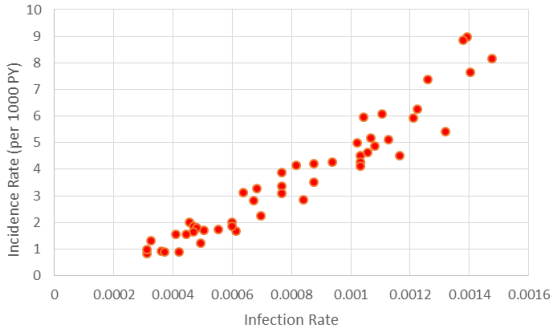


Figure A2.11: MB incubation time - Incidence Rate (2000 - 2003)

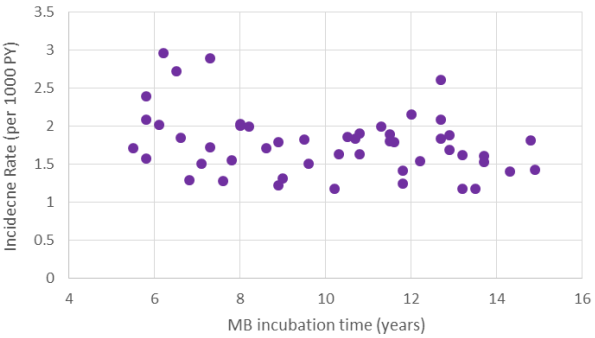
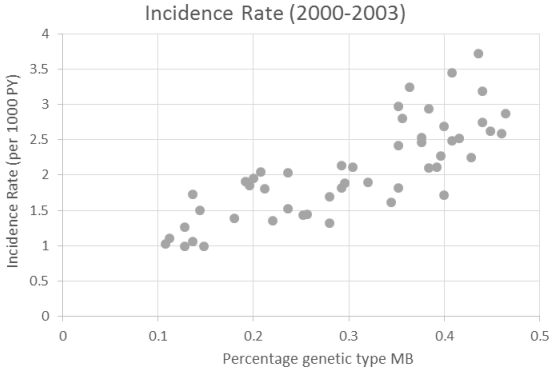


Figure A2.12: Percentage genetic type MB - Incidence Rate (2000-2003)



A2.4 Output: Incidence Rate: sensitivity in time

Figure A2.13: Prevalence Rate (2000) - Incidence Rate (2000-2003)

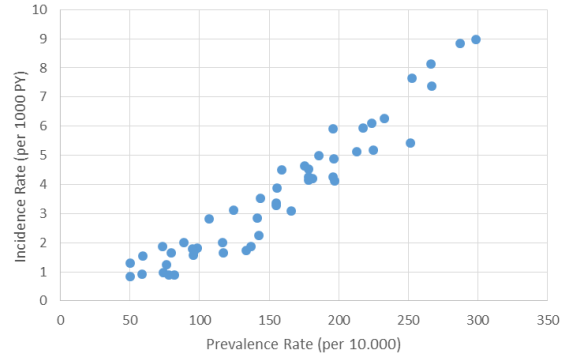
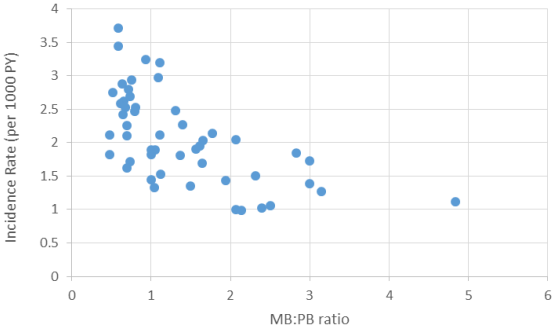


Figure A2.14: MB:PB Ratio (2000) - Incidence Rate (2000-2003)



A3 Step 3: sensitivity of disease model to population model

Figure A3.1: Population Size - Prevalence Rate (2000)

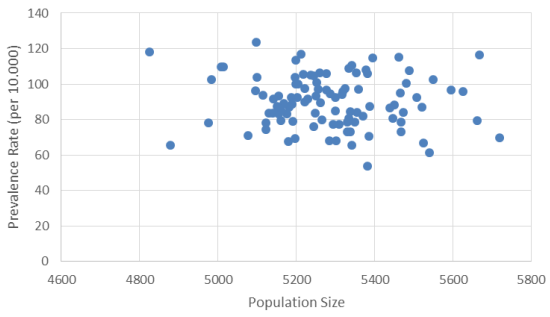


Figure A3.2: Mean Household Size - Prevalence Rate (2000)

Figure A3.3: Population Size - MB:PB Ratio (2000)

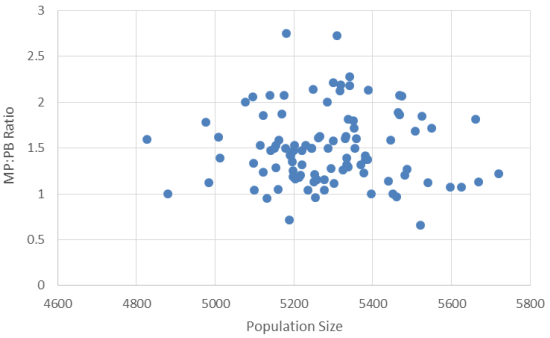


Figure A3.4: Mean Household Size - MB: PB Ratio (2000)

Figure A3.5: Population Size - Incidence Rate (2000 - 2003)

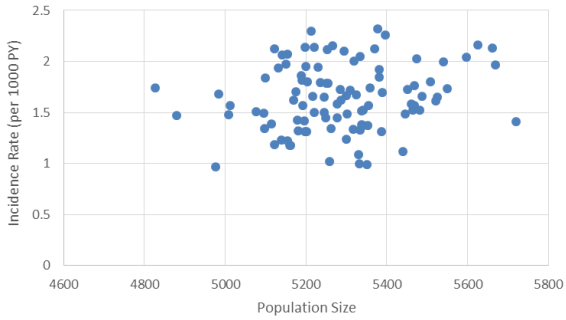


Figure A3.6: Mean Household Size - Incidence rate (2000-2003)

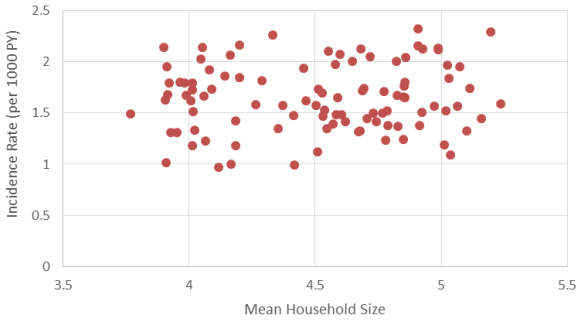


Figure A3.7: Population Size - Percentage of households with no patients (2000)

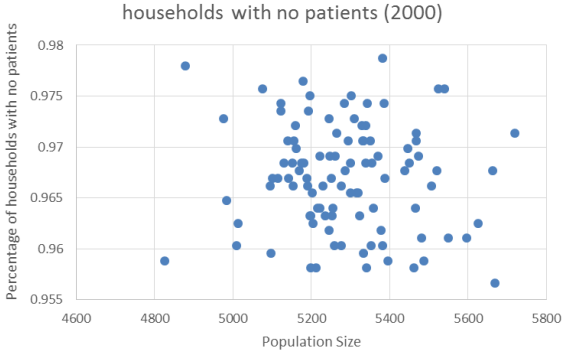


Figure A3.8: Mean Household Size - Percentage of households with no patients (2000)