University of Alberta

Modeling the Transmission of Tuberculosis in Long-Term Care Facilities using a Network Model

by

Alison Muscat

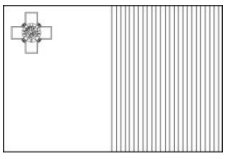
A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Applied Mathematics

Department of Mathematical and Statistical Sciences

Fall 2012
Edmonton, Alberta

The research work disclosed in this publication is partially funded by the Strategic Educational Pathways Scholarship (Malta). This scholarship is part-financed by the European Union - European Social Fund (ESF) under Operational Programme II - Cohesion Policy 2007 - 2013, "Empowering People for More Jobs and a Better Quality of Life".



Operational Programme II – Cohesion Policy 2007-2013
Empowering People for More Jobs and a Better Quality of Life
Scholarship Part-financed by the European Union
European Social Fund (ESF)
Co-financing rate: 85% EU Funds; 15% National Funds

Investing in your future

Abstract

Tuberculosis (TB) infection in the elderly is frequently misdiagnosed. The resulting treatment delay may increase TB transmission which is higher in long-term care (LTC) facilities. The CDC's recommendations to prevent and control TB in LTC facilities include TB education and better initial screening methods on entry into the facility. However, TB education programs might not always be given priority and comparing screening methods experimentally is often not feasible.

To address these problems, we develop a general conceptual SEIR network model for LTC facilities and present a case study of a specific outbreak that occurred in a nursing home in Arkansas. We investigate the impact of reducing diagnosis delay on the Arkansas outbreak and evaluate potential screening programs for that setting. Our results quantify the effectiveness of reducing diagnosis delay, justifying a good TB education program. We also suggest multiple screening programs that were found to produce equivalent results.

Acknowledgement

First and foremost, I would like to thank God, for none of this would have been possible without Him. He has given me the grace and the strength to pursue this journey and see it through.

I would like to thank my supervisor, Dr. Michael Li, for his support and guidance, and my colleagues and friends, Betsy Varughese and Peng Du, for their invaluable input to this project.

Another word of thanks goes to Dr. Philip Jacobs and Dr. Geetika Verma, who presented us with the literature that sparked this project, and to Michael Akinwumi, colleague and friend, off of whom I bounced various ideas.

Last, but not least, I would like to thank my family and friends who have supported me throughout this journey. I have been blessed with a great support network, for which I am immensely grateful.

Table of Contents

List of Tables

List of Figures

List of Symbols

Parameter Descriptions and Values for the Arkansas Model

Parameter	Definition	Units	Value
γ	Prop. of successfully treated cases of active TB	%	0 or 100
ω	Prop. of converters successfully treated for latent TB	%	80
ε	Rate of progression from latent to active TB	Day^{-1}	3.00×10^{-4}
λ	Transmission probability	$(Contact.Day)^{-1}$	9.90×10^{-2}
	Mean number of contacts within:		
μ_1	Wing 1	Day^{-1}	3
μ_2	Wing 2	Day^{-1}	14
μ_3	Wing 3	Day^{-1}	0
	Number of visits from:		
β_{12}	Wing 1 to Wing 2	Day^{-1}	86
β_{21}	Wing 2 to Wing 1	Day^{-1}	7
β_{13}	Wing 1 to Wing 3	Day^{-1}	38
β_{31}	Wing 3 to Wing 1	Day^{-1}	0
β_{23}	Wing 2 to Wing 3	Day^{-1}	12
β_{32}	Wing 3 to Wing 2	Day^{-1}	0

Glossary

active TB disease	A person who is experiencing TB symptoms and is able to transmit infection to others.
converter	A person who is known to have been tuberculin-negative but is now testing positive, indicating a new or recent latent TB infection.
latent TB infection	A person who has been exposed to TB but is not experiencing any symptoms and thus is not able to transmit infection.
Latin Hypercube Sampling	A statistical method used to generate a sample of parameter sets with more likely values according to the specified parameter distributions.
Monte Carlo simulation	A method to evaluate a model through a number of iterations, each of which is run with a random set of inputs. Individual outcomes are stored and used to calculate mean model outcomes.
nonlinear least squares	A method for fitting a nonlinear model to a dataset.
reactor	A person who has a positive reaction to the tuberculin skin test and thus has latent TB infection.
sensitivity analysis	A study to quantify the impact of uncertain factors of a model on the model's outcomes.
tuberculosis	An airborne infectious disease caused by the *Mycobacterium tuberculosis* complex.

1. Introduction

Tuberculosis in the elderly (>65 years old) is an important issue in Canada. In 2009, this section of the population accounted for 20% of all reported Tuberculosis (TB) cases [2]. Previous exposure to TB, a compromised immune system and the presence of chronic conditions, are some of the factors that put the elderly at a higher risk for TB infection [3]. This risk is further increased if the elderly are living in an enclosed space, sharing the same sources of air, food and water, such as in a nursing home [4]. The costs associated with an active TB case are also significant. The Canadian government spent $47,290 for each active TB case in 2004 and this includes the direct costs for the care of the patient with active TB, contact investigations and the treatment of any resulting latent TB infections (LTBI) [5].

In older adults the misdiagnosis of TB infection is a common problem since more frequently diagnosed illnesses, such as pneumonia and lung cancer, share a number of symptoms with TB [3,6,7]. This misdiagnosis results in treatment delay which may lead to an increase in disease transmission and could

potentially endanger the patient's recovery [8]. The presence of a good TB screening program is thus important. Continuous screening of the residents for latent TB infection would be highly beneficial but is not feasible as it is costly and puts a substantial burden on the nursing staff.

Quite often, for ethical reasons or otherwise, it is not possible to conduct experiments to assess the effectiveness of different intervention strategies. Mathematical modeling has proven to be a useful tool to investigate and compare such strategies [9,10]. In this thesis, we will develop a network model for TB transmission in long-term care facilities and use it to investigate different aspects of TB intervention strategies.

1.1 Objectives

The main objectives of this thesis are:

1. To develop a conceptual network model for Tuberculosis (TB) transmission in long-term care facilities.
2. As a case study, to adapt the conceptual model to a TB outbreak in an Arkansas nursing home and develop a specific network model that can recreate the Arkansas outbreak.
3. To incorporate different TB interventions into our Arkansas Model to investigate:
    - the impacts of reducing diagnosis delay, and
    - the effects of different screening programs.
4. To demonstrate applicability of network models to the study of control and prevention of disease transmission in long-term care facilities.

1.2 Some Background on Tuberculosis

Tuberculosis (TB) is an infectious disease that is caused by the *Mycobacterium tuberculosis* complex. It is transmitted from person to person through the air. Infection occurs primarily through inhalation [11] of droplet nuclei of this pathogen (1-5μm in diameter [12]) that are expelled from the infectious person's body into the air mainly through, but not limited to, sneezing or coughing [13]. Any susceptible individual that is close enough to inhale these airborne pathogens may become infected. The infection usually attacks the lungs, in which case it is called pulmonary TB. However, extra-pulmonary TB may develop if the infection spreads from the lungs to some other part of the body which may include, amongst others, the central nervous system, the bones and the genitourinary system. [11,13]

Individuals infected with the TB mycobacterium do not always develop active TB disease [11]. The bacteria may remain inactive inside the lungs for a period of time during which the individual is said to have latent TB infection (LTBI). At this stage one does not display any symptoms and is therefore not infectious [13]. If no preventive treatment is given for LTBI around 10% of the latently infected population will eventually develop active TB. These latently infected individuals have a higher probability of developing active TB in the first two years after infection [11]. Once the pathogens in the lungs become activated, the individual starts showing symptoms and is able to transmit the disease. Common symptoms of active TB disease include coughing, fever, night sweats, loss of appetite and weight loss [13].

Diagnosis of new disease usually involves skin testing for conversions, chest radiography for those that are tuberculin-positive and analysis of sputum for acid-fast bacilli to determine whether active TB is present [3]. The tuberculin skin test (TST) is used to diagnose TB infection for which a five tuberculin unit dose of purified protein derivative (PPD) is administered intradermally using the

Mantoux method. The reaction is read 48 to 72 hours later and a diagnosis is made according to the size of the measured induration whilst taking into account the person's risk category [6].

Tuberculosis in Long-Term Care Facilities

Tuberculosis is a nosocomial (healthcare-associated) infection for residents and employees of long-term care facilities [14]. Elderly persons residing in such long-term care facilities are at a higher risk for TB than other elderly persons living within the community [15-17]. In a survey conducted in the USA, the TB incidence rate for nursing home residents was found to be 1.8 times higher than that of elderly community dwellers [17]. A combination of personal and environmental factors make the elderly, especially those living in long-term care facilities, more susceptible to TB infection. Previous exposure to TB, a compromised immune system and the presence of chronic conditions are examples of such personal factors, whereas the concentration of M. tuberculosis in the air and congregate living within such facilities further increase the risk of infection [3].

The elderly who are susceptible to TB include those with no prior exposure to the bacterium, those with a dormant infection that is at risk of reactivation, as well as those who had previously recovered from infection but may now be susceptible to reinfection [15]. Reactivation of a latent infection is the reason behind the majority of active TB cases in the elderly [6,15,18,19]. The risk of an elderly person getting active TB disease from a new infection was estimated to be between 8% and 12% [19,20].

Identifying, diagnosing and treating TB infection and active disease in the elderly are more challenging [3,6]. Elderly people having active TB do not necessarily display the more common symptoms of TB [15]. For this reason, a set of recommendations for the prevention and control of TB in long-term care

facilities were published by the Advisory Committee for Elimination of TB and include the four principal elements of surveillance, containment, assessment and education [16].

Surveillance includes the identification and reporting of all cases of both TB infection and disease within the facility. Such measures include screening all new residents on admission and all employees on employment unless a previous positive reaction to the skin test is already documented. Containment deals with the prevention of further disease transmission within the facility. The advised treatment should be administered and completed and the proper ventilation system switched on. Assessment primarily includes the monitoring and evaluation of both surveillance and containment within the long-term care facility. Education includes preparing residents, families, visitors and employees to realize the importance of their cooperation with surveillance, containment and assessment activities. This is done by giving them appropriate information and by equipping them with the right skills. [16].

The nursing staff`s role in the control of TB includes the facilitation of early diagnosis of TB and the upholding of the required isolation practices for newly diagnosed patients, amongst others [3].

1.3 Mathematical Modeling of Infectious Disease Transmission

The role of mathematical modeling in epidemiology has become more prominent over the past century. Some diseases may be managed in different ways including, but not limited to, strategies of prevention, isolation of infectious patients, quarantine of susceptibles or treatment of symptoms [9]. The effectiveness of such strategies cannot usually be compared using experiments, for ethical reasons or otherwise, and the only possible solution is to develop a mathematical model to make predictions and compare the outcomes of the possible management strategies [9,10]. It was through mathematical modeling that Ross concluded that

the spread of malaria could be managed by controlling the number of mosquitoes [21]. A brief summary of some major accomplishments of mathematical modeling are described by Brauer in [9].

The suggested control strategies may be considered a side benefit of using modeling. A mathematical model mainly provides insight into what affects the spread of disease, which may not be evident from the data itself due to inaccuracies in measurement or otherwise [9,10].

Brauer [9,10] also briefly discusses the trade-off which is present in most areas of mathematical modeling, including modeling of disease transmission. This trade-off is between simple models and more detailed (and thus more complex) ones. A useful model is one that gives, with no excessive complexity, a plausible answer to the question for which it was set up.

We now give a brief introduction to compartmental models based on Brauer's more detailed discussion of compartmental modeling in chapter 2 of *Mathematical Epidemiology* [10]. In such models, the population is split into a number of compartments depending on the nature of the disease that is being considered. For diseases with immunity, one would have a class for susceptible individuals (S), one for the infectious individuals (I) and one for the recovered (R). A model for a population moving forward across these three compartments is called a SIR model [10]. A well-known example of such a model is the one introduced by Kermack and McKendrick in 1927 [22].

For some diseases, such as tuberculosis, one formulates SEIR models because of the existence of a latency period (E) where individuals have been infected but are thus not yet capable of transmitting the disease to others. Compartmental models for other different types of diseases include SIS and SIRS models, amongst others [10].

The variables in this type of model are the time *t* and the numbers of susceptibles, infectious and recovered individuals at time *t* are *S(t)*, *I(t)* and *R(t)*, respectively. Assumptions are made on the transfer rates across compartments and since the model is formulated based on the changes in the size of each compartment using differential equations. This kind of model is therefore deterministic with the disease progression within a population completely predictable from the modeling equations and initial conditions. Another important concept is the basic reproductive number, R_0, which measures the number of secondary infectious cases produced by one infectious case in an entirely susceptible population throughout its entire infectious period [10].

Deterministic models are suitable for large populations such that homogeneous mixing is justified. However, especially at the beginning of an outbreak, there may be many susceptibles and very few infectious individuals. The pattern of contacts is thus more important, suggesting that deterministic models should be replaced by stochastic models. Two possible options are complete stochastic models and network models. Stochastic models require many simulations and each simulation may have a different outcome, whereas one simulation is enough for deterministic models as the same output is obtained every time the model is run [10].

Network models are stochastic and are very useful in simulating the transmission of infectious diseases within relatively small populations. They take into account the individual contacts between members of the population. As such, they give a more realistic picture of how the disease is spreading in that particular community and allow one to extract more information from the model. When dealing with a sufficiently large population, one may switch to a deterministic model once the outbreak is past its initial stages [10]. A more detailed description of network models is provided by Brauer in chapter 4 of *Mathematical Epidemiology* [10], whereas a detailed introduction to complete stochastic models is provided by Allen in chapter 3 of the same book [10].

1.4 Literature Review on TB models

We here present a brief introduction to some of the TB models published in the literature. The majority of these are differential equations SEIR-type models and include different aspects of TB such as drug-resistant strains of TB, treatment and vaccination strategies, TB-HIV co-infection and TB amongst immigrants.

Blower et al. [23] present a theoretical framework for the investigation of TB transmission dynamics. This is done through two SEIR-type differential equation models, a simple one and a more detailed one that includes both active TB (infectious and non-infectious) and recovery. The very slow dynamics of TB epidemics are also discussed.

A model by Ziv et al. [24] focuses on the reduction of the incidence rate for TB through the treatment of early latent infection. They conclude that contact investigations, with the appropriate treatment when required, has a significant impact on reducing the spread of TB.

Castillo-Chavez and Feng [25] analyze an age-structure model for TB transmission in the presence of vaccination to identify an optimal vaccination strategy. They conclude that this optimal strategy can be either vaccination of the susceptible population at one particular age (one-age strategy) or vaccination of a proportion of susceptibles at an age and the rest at another later age (two-age strategy).

Another TB model that incorporates vaccination is presented by Lietman and Blower [26]. They develop two simple mathematical models investigating the impact on TB epidemics of pre-exposure vaccines in the first and the impact of post-exposure vaccines in the other. Their conclusion is that vaccines which are only moderately effective, may still have an impact on the reduction of TB

epidemics, as long as they are used in conjunction with treatment rates that are sufficiently high.

A *reinfection threshold* is detected by Gomes et al. [27] in a general TB epidemic model. Once this threshold is reached, disease transmission becomes mainly due to reinfection. At that stage, vaccination fails due to the resulting drastic increase in infection levels. They conclude their findings by showing how vaccination programmes can be used to manipulate these reinfection thresholds. Cohen and Murray [28] present another model that includes reinfection, this time amongst US immigrants.

Cohen et al. [29] develop a network model for TB transmission that takes into account reinfection. Other models, including Cohen and Murray's model for TB amongst immigrants [28], had stated that reinfection is a particularly major factor in high-incidence regions. Due to the absence of a homogenous mixing assumption in network models, Cohen et al. [29] suggest that reinfection may also be a major factor in moderate- or low-incidence regions.

Castillo-Chavez and Feng [30] present two models. First they develop a simple SEIR model for one-strain TB transmission and then they present a two-strain TB model that includes a resistant strain which is left untreated. They use the latter model first with a naturally-resistant strain and then with an antibiotic-resistant strain. Their conclusion is that although not frequent, co-existence of a typical strain and a naturally-resistant strain is possible. On the other hand, co-existence is standard when the resistant strain was resistant due to antibiotics.

Cohen et al. [31] developed a model for populations coinfected with TB and HIV. They use this model to analyze the effects on the dynamics of drug-resistant TB caused by the administration of isoniazid preventive therapy (IPT) to the coinfected population (community-wide IPT). Although drug-resistant TB strains would emerge more quickly on implementation of such a programme, they

conclude that this strategy is still beneficial as long as the necessary identification and effective-treatment policies for the increasing proportion of drug-resistant TB cases were implemented.

This brief literature review is not meant as a full review of TB models but rather it aims to give a brief overview of the various aspects that TB models try to address. A more detailed review on TB models is presented by Colijn et al. [32].

1.5 Overview of Our Results

1. We used the network modeling approach to develop a conceptual model for TB transmission in long-term care facilities.
2. Case study (the Arkansas Model): we adapted the conceptual model to a specific TB outbreak that occurred in a nursing home in Arkansas, USA, between 1977 and 1979, as described by Stead [1].
3. We used the Arkansas Model:
    a. to investigate the impact of reducing diagnosis delay, and
    b. to identify an optimal screening program for that nursing home's residents.

    For the Arkansas setting, reducing diagnosis delay to one month results in a reduction of 25% in the number of latent TB infections and a reduction of 47% in the number of active TB cases. Multiple screening programs for this Arkansas setting were found to produce the same results.

Our results demonstrate how a network model may be used to implement different aspects of TB intervention strategies and investigate their effectiveness leading to recommendations for policy making.

1.6 Thesis Outline

Chapter 2 describes the modeling process and introduces an example of a real TB outbreak that is described in published literature [1]. Details from this outbreak provide us with the data used for parameter estimation, the process for which is described in this section. We also include a description of the sensitivity analysis process.

Chapter 3 first describes the results of both the parameter estimation and sensitivity analysis processes. We then compare the results of our Arkansas Model with the outbreak described by Stead [1]. This Arkansas Model is then used to explore the impact of reducing diagnosis delay. We also assess potential screening programs that test for conversions and if appropriate, treat, a percentage of the susceptible population every number of months. The latter is aimed at reducing the burden on the nursing staff by finding a good balance between the testing frequency and the proportion of residents to test. We conclude this thesis in Chapter 4 by providing a brief summary of this work and a discussion of any limitations and potential future work.

2. The Modeling Process

Network models have been increasingly used to model infectious disease transmission [9]. Keeling and Eames [33] provide a review of basic aspects of network theory and epidemiology, linking the two together. Additional theoretical background on network models can be found in [10,34,35]. Others [34,36-38] apply network models to investigate various infectious diseases through the use of simulations. A network model for TB transmission is presented by Cohen et al. [29]. We have adapted the network modeling approach to develop a general conceptual model for TB transmission in long-term care facilities. This conceptual model can be adapted to specific long-term care facilities.

In this chapter we describe the modeling process. We first develop in Section 2.1, a general conceptual SEIR TB network model in the setting of a long-term care facility. We then adapt this modeling framework to a particular case of a TB outbreak that happened in a nursing home in Arkansas, USA, in the late 1970s, as described by Stead [1]. Using the data extracted from this paper, we perform parameter estimation and model fitting using random sampling methods. This process produces a network based stochastic model whose expected outcome

reproduces the Arkansas outbreak. Sensitivity analysis is then performed using two methods to determine the parameters that are the most sensitive with respect to model outcome. Results from the analysis will be presented in Chapter 3.

2.1 A General Conceptual SEIR Network Model for TB transmission in Nursing Homes and other Long-Term Care Facilities

A network (graph) is used to represent the contact among residents in a long-term care facility. The nodes in the network correspond to rooms of the facility whilst the edges between the nodes reflect the contacts between one or more residents of the corresponding rooms. The degree of a node is the number of edges at the node and describes the average number of contacts of that room.

To model the spread of TB in the facility, the resident population is compartmentalized into 4 categories (*S*, *E*, *I*, *R*) based on their TB health status: susceptible, exposed (latent TB infection), infectious (active TB disease) and recovered (received treatment and is therefore no longer susceptible to the same strain of TB).

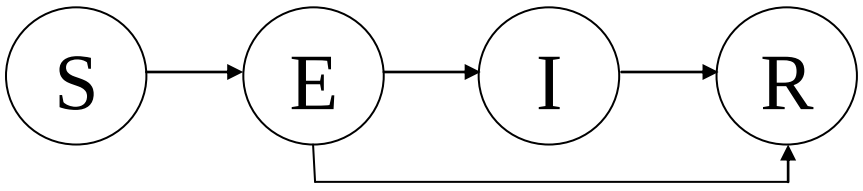


Figure 2.1. An SEIR process for the progression of Tuberculosis in elderly residents of a long-term care facility.
Stages of the process include susceptibility to TB (*S*), latent TB infection (*E*), active TB disease (*I*) and recovery (*R*) from either infection or disease.

At each node of the network, we defined a set of attributes that describe the corresponding room. This set contains the room number, the wing number and the health status of each of its residents. The health status of a resident may have only one of these values at any point in time: susceptible, reactor, converter, infectious, recovered converter and recovered infectious. As the disease progresses throughout the residence, this value will change based on the resident's contacts and the disease transmission rates, amongst others.

A reactor is a person who has a positive reaction to the tuberculin skin test (TST). A converter is a person who is known to have been tuberculin-negative but is now testing positive, indicating a new or recent infection [39]. This is measured as an increase of over 10mm in induration over a period of two years [40]. Both reactors and converters are latently infected (*E*) but their probability of developing active TB disease are different. They are thus assigned different rates of progression from *E* to *I*. The risk of developing active TB is higher in the first two years after infection [11] so the converters are given priority over the reactors for LTBI treatment [41].

Connections within a wing

Long-term care facilities are typically organized into different wings. The contacts between residents living within the same wing *i* are created first. We make the following assumptions on the number and distribution of contacts within a wing:

1. there is only a small number of rooms that either have few contacts or else many contacts; and
2. most rooms have an average of μ_i contacts.

Therefore, we let *X* be a discrete random variable that represents the number of contacts per day that a room has with other rooms located in the same wing. For wing *i*, preceding assumptions 1 and 2 are the same as assuming that *X* has a

Poisson distribution with mean μ_i. Thus, the probability *P* that a room has *k* contacts is given by

$$P(X = k) = \frac{\mu_i^k e^{-\mu_i}}{k!}$$

The following explains the process of creating these connections:

Step 1

Create a vector *P* such that *P(k)* contains the probability of a room having *k* contacts (as described above). If the number of rooms in the wing is *n*, then the maximum value of *k* is equal to *n - 1* since a room may not be connected to itself. Another vector *N* is created such that *N(k)* holds the number of rooms that have *k* contacts. This is obtained by multiplying each entry of *P* by *n* and then rounding its value since the number of rooms is an integer value.

Step 2

Summing the entries of *N* should result in *n*. However, due to the rounding performed in Step 1, this might not be the case.

- If *sum(N)* < *n*, then a number of rooms have not been assigned a degree. The difference is thus added to $N(\mu_i)$ to ensure that the mean of the distribution is not changed.
- Otherwise, if *sum(N)* > n, a number of rooms have been assigned multiple degrees. To ensure that all rooms are assigned only one degree, a random element of *N*, say *N(j)*, is selected and its degree is reduced by 1 as long as $j \neq \mu_i$ and $N(j) \neq 0$.

Step 3

A sampling vector *AS* of rooms is then created. *AS* contains the room numbers of all non-isolated rooms, with each room number repeated according to that room's number of contacts, given by *N(j)* for room *j*. A second sampling vector *AS1* is

also created and used later to make the connections between different wings. This second vector is created in the same way as *AS* with the main difference that each room number *j* is repeated *N(j)* times up to a maximum of μ_i. Capping at the mean μ_i the maximum number of times a room number may appear in this vector avoids the possibility of just selecting the most connected rooms as the ones also connected to the other wings.

Step 4

To establish the links between rooms in wing *i*, a vector *CM* of length *n* is created, where *CM(j)* is a list of all the rooms that are connected with room *j*. We create the links between the rooms and thus populate the vector *CM* as follows:

- select 2 rooms, *r1* and *r2* from *AS*, making sure $r1 \neq r2$ to avoid having self-connections;
- add *r1* to *CM(r2)*, which is the list of contacts of *r2*;
- similarly, add *r2* to *CM(r1)*. This is done based on the assumption that connections between rooms located in the same wing are bidirectional;
- remove *r1* and *r2* from *AS* by switching them with the last and next-to-last elements of *AS* respectively, and then truncate *AS* by two elements. The switching is done to further randomize this process.

Connections between wings

The connections between rooms located in different wings are established next. A connection from wing *i* to wing *j* represents a visit made by a resident living in wing *i* to a resident located in wing *j*. For such connections, we assume that a resident who is highly connected in their own wing, i.e. is very sociable, has a higher probability of having contact with residents of other wings. The total number of visits from wing *i* to wing *j* is given by β_{ij}.

To create these connections, *probability proportional to size (PPS) sampling* [42] was used to select 2 different samples, each from *AS1*. The likelihood of a room number being selected is proportional to the number of contacts the room has within its own wing.

- The first sample, S_{iX}, has length $\beta_{ij} + \beta_{ik}$ and it contains the room numbers of those rooms in wing *i* whose residents will *visit* residents in wing *j* and wing *k*.
- The second sample, S_{Xi}, contains the room numbers of those rooms in wing *i receiving* visits from residents living in wings *j* and *k* and it therefore has length $\beta_{ji} + \beta_{ki}$.

Another vector *WXtoWi* is then created, composed of 2 random samples selected as follows:

- β_{ji} rooms were selected from S_{jX}
- β_{ki} rooms were selected from S_{kX}

These are the rooms in wings *j* and *k* whose residents are visiting wing *i*. For each room *z* in S_{iX}, a random room from *WXtoWi* is selected and added to *CM(z)* to make the connection. *WXtoWi* is updated accordingly.

After the connections are set up, the index case of TB and the reactors are selected amongst the resident population. Letting the index case be a very sociable resident that is located in wing *i*, we select our index case to be that resident from wing *i* that has the most contacts. The *CM* vector is used to compare the lengths of the list of contacts of the rooms in wing *i* to obtain the most connected room in that wing. A resident from that room is then randomly selected to be the index case, by switching its health status to infectious. A number of reactors in each wing are selected randomly, allowing for the possibility of having multiple reactors in one room.

The network model set up using this process is stochastic. Each run of the model will result in a different network configuration. The number of nodes and the number of edges are fixed. However, the connections themselves are created randomly such that two connected nodes in one model evaluation are not necessarily connected in the next. Monte Carlo methods and mean model outcomes are thus used. A Monte Carlo simulation consisting of a number of model evaluations is performed, each time recording the output. The mean output is computed at the end. Most simulations in the studies of this thesis, including those for model validation, the investigation of the impact of reducing diagnosis delay and the assessment of different screening programs (Chapter 3), make use of 20 000 model runs to ensure that the mean model outcome is well-established.

The SEIR process is performed on the individuals of the population at every single time-step. The chosen time-step is one day and one evaluation of the model investigates TB progression in the long-term care facility for a specified duration.

At the program initialization, the network is configured and all room attributes are assigned. Then each individual's movement is followed across compartments from day 1 until the last day. Transitional probabilities are assigned to the relevant individuals to assess whether they should progress to another compartment. Flowcharts describe the general structure of the program (Figure A.1) and also how possible progression from one compartment to the next occurs (Figures A.2 - A.5).

2.2 Case Study: A TB outbreak in a nursing home in Arkansas, USA

Stead [1] describes an outbreak that occurred in a nursing home in Arkansas, USA, between 1977 and 1979. The first wave of the outbreak was caused by an index case that was misdiagnosed for more than a year. A second wave of active

TB cases was mainly due to treatment delay of the converters, which could not immediately be distinguished from reactors. We would like to adapt the conceptual model described in Section 2.1 to recreate this specific outbreak.

The nursing home consisted of 240 beds in 112 rooms, with 2 or 3 residents in each room, spread over six wings. Two of these wings were for skilled care so its residents very rarely left their rooms, whilst residents from the other four wings ate and played games together in the dining room. The resident population was mainly white of middle-class background with a mean age of 76.1 years.

At the time, the facility's TB control strategy consisted of a chest roentgenogram for residents close to admission to exclude active TB and skin tests for employees before starting employment at the facility. Chest roentgenograms were obtained from those employees whose reaction had an induration of 10mm or more within 48 hours. Employees were then tested annually and all converters, irrespective of age, would be treated with isoniazid.

In June 1978, a public health nurse noticed a number of conversions among the facility's employees, who were annually tested for latent TB infection. This prompted an investigation, which led to the discovery of the index case, a sociable 72-year old man who had been a resident of the nursing home for three years. On a visit to the hospital for minor surgery in June 1977, the index case had a routine chest roentgenogram which showed some abnormalities in the lungs. Diagnosed as a "probable bronchogenic carcinoma" by the radiologist, the patient turned down further investigations and was taken back to the nursing home. His symptoms, which included coughing and weight loss, were attributed to the carcinoma so he kept on interacting with the other residents.

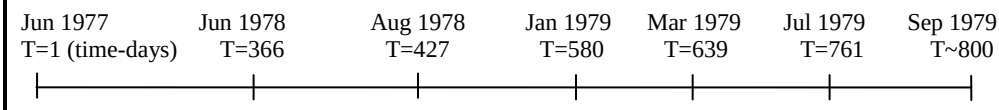
On discovery of the index case, all residents and employees with no known prior positive reaction to the skin test were skin tested and those that now

resulted positive were given chest x-rays. Tuberculin-negative residents and employees were tested again after two months. All employees that had converted were given prophylactic treatment with isoniazid and residents with active TB disease were also treated. Resident converters could not yet be identified and because of their age, preventive treatment was withheld.

In March 1979, an active TB case was confirmed in a 65-year old man with chronic lymphatic leukemia. He was found to be anergic through skin testing with different antigens since previous TST was negative. All previously tuberculin-negative residents and employees were tested once again and more converters were found among both groups. A 97-year old man was among the converters and, because of his age, it was decided that confirmation of active TB disease was needed before treatment was to be initiated. Before test results came back, he was admitted with abdominal pain to hospital where he was diagnosed with pneumonia. He did not respond to the corresponding treatment and died a week later. Confirmation of active TB disease arrived a few days later.

Realising that an outbreak of TB was occurring in the nursing home, more effort was put in to locate the residents' previous TST records. The majority of such records were finally obtained, allowing the distinction between reactors and converters to be made for the first time and thus the administration of the appropriate treatment.

Figure 2.2 outlines the major events of this outbreak. The outbreak data and the distribution of the reactors amongst the wings, as extracted from Stead [1], are described in Tables 2.1 and 2.2, respectively. The active TB cases are counted separately from the converters in Table 2.1.



June 1977: Misdiagnosis of the index case during hospital visit.

June 1978: Index case is diagnosed with active TB disease. All resident and employees with no known tuberculin-positive records are tested. Chest x-rays are given to those now resulting in a positive reaction.

Aug 1978: Tuberculin-negative residents and employees are tested again. 3 more cases of active TB disease and 3 new cases of LTBI are discovered among the residents. Residents with LTBI are **NOT** treated due to old age and the absence of certain medical records preventing the identification of converters from reactors. All active TB cases are treated, together with cases of LTBI amongst the employees.

Jan 1979: Misdiagnosis of another resident due to symptoms resulting from complications of leukemia.

Mar 1979: Resident is diagnosed with active TB disease. All tuberculin-negative residents and employees are tested once again, resulting in 20 new cases of LTBI and 1 case of active TB disease amongst the residents. Treatment is finally given to all converters as they could finally be distinguished from reactors.

July 1979: 1 more resident with active TB disease and another with LTBI are discovered and treated accordingly.

Figure 2.2. Timeline of the major events of a TB outbreak in a nursing home in Arkansas as described by [1].

2.3 Our Baseline Network Model for TB transmission in the Arkansas Nursing Home (the Arkansas Model)

Our Arkansas Model was built according to the description of the general conceptual model in Section 2.1. MATLAB R2007b is used to develop the program for model simulations.

Additional simplifying assumptions for our Arkansas Model in relation to the data include:

- 2 residents residing in each room (total of 224 residents vs. the 240 stated in the paper);
- the nursing home rooms are regrouped into 3 wings according to location and type;
- minimal transmission from employees and visitors to residents - model simulates TB transmission solely amongst the resident population;
- a constant resident population, no new admissions or deaths considered;
- no common areas - dining room is excluded;
- 2 periods of transmission: Jun 1977 – May 1978 and Oct 1978 – Feb 1979. Otherwise, all active TB cases are discovered and treated (treatment efficacy of 100%);
- rate of progression to active TB for reactors is chosen to be a *quarter* of the corresponding rate of progression of converters as none of the residents who were tuberculin-positive before June 1978 developed active TB disease [1];
- on average, frequency of contact between roommates is higher than that between residents of different rooms. For simplicity, we assume a contact ratio of 10:1;

- network configuration is determined at the beginning. New contacts cannot be created and existing ones cannot be removed during each model evaluation.

Table 2.1. Outbreak data, obtained from Stead [1]

Date	Jun 77	Jun 78	Aug 78	Mar 79	Jul 79
Time (days)	1	366	427	639	761
Accumulated number of converters	0	18	21	41	42
Accumulated number of active TB cases	1	3	6	7	8

Table 2.2. Data about reactors, obtained from Stead [1]

Wing Number	**1**	**2**	**3**
Number of Reactors	27	8	12

Table 2.3. Distribution of cases among different wings by the end of the outbreak [1]

Wing Number	**1**	**2**	**3**
Active TB Cases (actual)	8	0	0
Latent Cases/Converters (actual)	21	18	3

The timeline in Figure 2.3 describes the periods of transmission, which include the year taken for the index case to be discovered (June 77 - June 78) and the second transmission window that starts in October 78, based on our assumption that active TB cases were diagnosed and treated for the 4 months following the discovery of the index case. This second period of transmission ends in March 79, with the diagnosis of a new active TB case and the treatment of both converters (ω = 0.8) and active TB cases (γ = 1).

A schematic diagram of the network representing our nursing home, taking into account the above assumptions, is displayed in Figure 2.4.

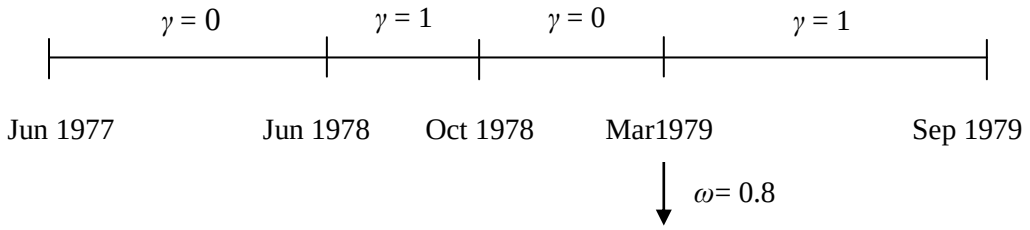


Figure 2.3. Timeline of the Arkansas outbreak including our assumptions.
These include assumptions about periods of diagnosis delay and the proportions of successfully treated converters (ω = 0.8) and active TB cases (γ = 1).

2.4 Parameter Estimation

The progression rate (per day) from latent TB infection to active TB disease is described by the parameter ε. The data in Table 2.1 was used to estimate its value as follows: over the course of one year, between June 1977 and June 1978, 2 new active TB cases developed from a total of 20 converters. An estimate for the value of ε was thus obtained by the following calculation:

$$\varepsilon = \frac{2}{23 \times 365} \approx 0.0003$$

The parameter ω describes the recovery rate of the converters. The Tuberculin Skin Test has a false negative rate of up to 20% [43], implying that only 80% of the converters are effectively diagnosed with LTBI. A treatment efficacy of 100% is assumed and thus ω is assigned the value 0.8.

The recovery rate from active TB disease is described by the parameter γ. During periods of transmission, γ takes on the value 0, otherwise it is assumed that all infectious cases are isolated and treated immediately. A treatment efficacy of 100% is also assumed and thus during these periods γ has a value of 1.

The description of the outbreak provided by Stead [1] does not include any details about either the contact rates or the transmission rate. These parameters are thus fitted using the data extracted from Stead [1], which includes the accumulated number of converters (five data points) and the accumulated number of active TB cases (five data points) at specific points in time (Table 2.1), as well as the distribution of LTBI (three data points) and active TB cases (three data points) amongst the different wings (Table 2.3).

The method of nonlinear least squares is used to fit these parameters:

- λ, which is the probability that a contact per day results in infection;
- the mean of the Poisson distribution of contacts of the rooms in wings 1 and 2, which are given by μ_1 and μ_2; and
- the number of visits originating from wings 1 & 2 (β_{12}, β_{13}, β_{21} and β_{23}). As wing 3 is for skilled care, no contacts occur amongst its residents ($\mu_3 = 0$) and the residents do not leave their rooms ($\beta_{31} = 0 = \beta_{31}$).

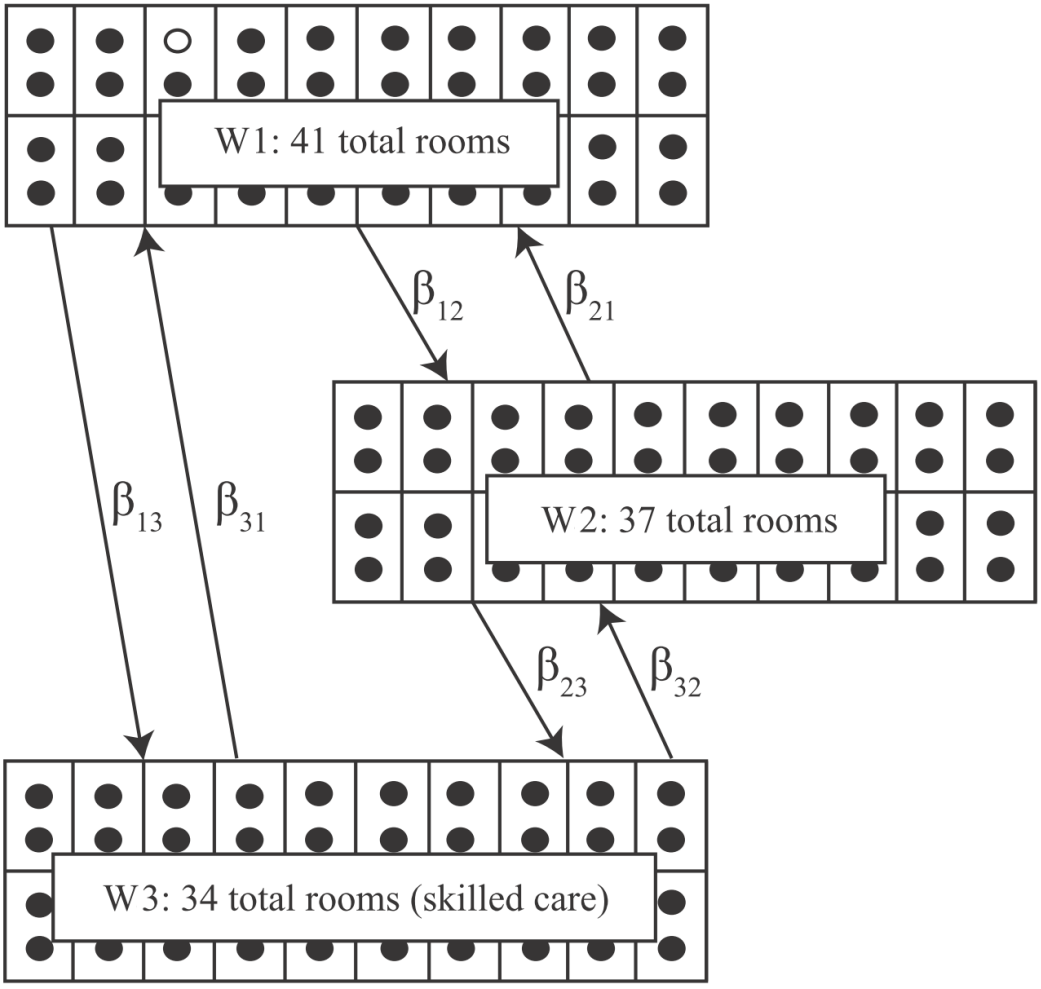


Figure 2.4. A schematic diagram of the network representing the nursing home of the Arkansas Model.
The parameters β_{ij} represent the number of visits per day from Wing *i* to Wing *j*.

Rooms in wings 1 and 2 are not generally isolated so the minimum value for both μ_1 and μ_2 is set to be equal to 1. The upper end of the range of values of μ_i is restricted to one less than the number of rooms in that wing *i* since a room can at most be connected to all other rooms in that wing but may not be connected to itself. The maximum values for μ_1 and μ_2 are thus equal to 40 and 36, respectively.

The assumption that, each day, residents living in the same room had 10 times as much contact with their roommate than they had with other residents, led to the restriction of the value of the parameter λ to a maximum of 0.1.

Details about the type of contact amongst wings were not provided and thus some further assumptions were made. Wings 1 and 2 have a total of 82 and 74 residents, respectively so the maximum number of visits from wing 1 to wing 2 (β_{12}) and vice versa (β_{21}) was set to 100. A lower maximum of 50 visits originating from each of wings 1 and 2 to wing 3 (β_{13} and β_{23}) was chosen so that each wing may now have a maximum total of 100 visits (Figure 2.5).

Table 2.4. Range of possible values for fitted parameters

Parameter	[min,max]
λ	[0,0.1]
μ_1	[1,40]
μ_2	[1,36]
β_{12}	[1,100]
β_{21}	[1,100]
β_{13}	[1,50]
β_{23}	[1,50]

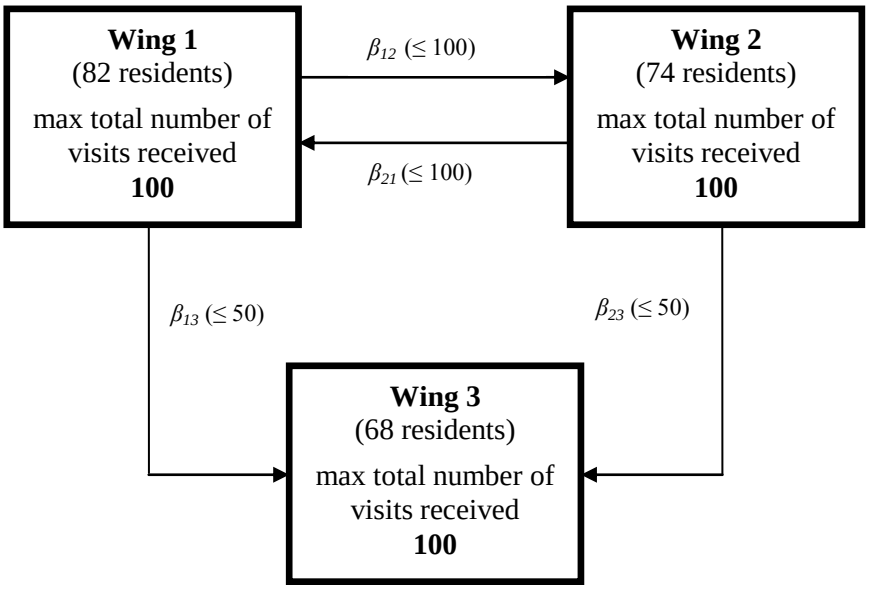


Figure 2.5. The maximum number of visits per day for each wing.

Parameter Fitting - Stage 1

To make the fitting tractable, a sample of 5000 parameter sets was selected. Each parameter set *p* contains a value for each of the 7 parameters to be fitted, which were selected randomly from the parameter ranges described in Table 2.4. Due to the stochasticity of the model, each time the model is evaluated, the network's edges are configured differently. Mean model outcomes thus needed to be used. For each of the 5000 parameter sets, a Monte Carlo simulation of 500 model evaluations is executed using the same parameter set *p*. The outcomes recorded after each model evaluation include the number of converters and the number of active TB cases at specific time intervals over the duration of the outbreak, as well as the distribution of both LTBI and active cases across the 3 different wings. A

vector of length 500 is thus created for each outcome x and the corresponding mean model outcome, d_x (Table 2.5), is calculated by taking that vector's mean.

Table 2.5. Model Outcomes

Outcome description	Outcome representation
# of active TB cases at specific points in time	d_1, d_2, d_3, d_4, d_5
# of converters at specific points in time	d_6, d_7, d_8, d_9, d_{10}
# of active TB cases across the 3 wings	d_{11}, d_{12}, d_{13}
# of converters across the 3 wings	d_{14}, d_{15}, d_{16}

The mean model outcome for a parameter set p is described by the vector *data*, where it is defined by:

$$data = \left(d_x\right)_{x=1}^{16}$$

The observed values, provided by Stead [1], are described by the vector *data**. For each set p, the error is calculated, squared and summed up to obtain the sum of squares, $SS(p)$, for that parameter set.

$$SS(p) = \sum_{i=1}^{16} |d_i^* - d_i|^2$$

The 5000 parameter sets are then sorted in ascending order according to their sum of squares.

Parameter Fitting - Stage 2

The next step makes use of the top 50 parameter sets obtained through the process described in Stage 1. The three most frequent values for each of the contact parameters (μ_1, μ_2, β_{12}, β_{13}, β_{21} and β_{23}) were selected, whilst the range for the parameter λ was linearized into 12 values. Another sample of parameter sets was

formed by obtaining all possible combinations of these selected values. The size of this sample is thus equal to 8748 (i.e. 3^6 x 12). The same procedure as in Stage 1 is followed for all parameter sets and again, they are ordered by the lowest sum of squares. The parameter set that has the lowest sum of squares is considered to be the best fit for our model.

2.5 Sensitivity Analysis

The purpose of performing sensitivity analysis is to determine which parameters have the largest effect on each of the outputs of interest. It is not simply a method of finding the most important factors since the definition of importance needs to first be specified [44]. In our case, the outputs of interest are the number of converters and the number of active TB cases occurring during a fixed period of time. We would like to determine which of the parameters has the most impact on these numbers.

Sensitivity analysis may be classified into either *local* or *global*. Local methods, sometimes also referred to as *One-Factor-at-a-Time (OFAT)* methods, perturb one parameter whilst fixing the rest [45]. Results may include graphs presenting the size of the output as it changes with variations in the investigated parameter. One major drawback of this kind of method is that the interactions of the investigated parameter with those parameters that are kept fixed are completely disregarded. To overcome this, global sensitivity analysis is opted for instead of local methods.

Marino [46] gives an overview of when it is appropriate to use some standard methods based on the type of relationship that exists between input and output. Scatter plots may be used initially to determine any nonlinearities between the parameters and the outcomes of interest. If a linear relationship exists, standard methods for its measurement include the Pearson correlation coefficient and the Partial Correlation Coefficients (PCC), amongst others. Other methods,

such as the Spearman Rank Correlation Coefficients (SRC) and Partial Rank Correlation Coefficients (PRCC), are useful in the presence of a monotonic nonlinear relationship, where parameters and outputs are first ranked. The third class of methods is used for non-monotonic nonlinear relationships and are based on variance decomposition of model output. These include the Sobol method and the Fourier Amplitude Sensitivity Test (FAST) [46].

In our sensitivity analysis, scatter plots of the fitted parameters against the outputs of interest were generated and based on the absence of a clearly defined linear relationship, PRCC was used for our global sensitivity analysis. The sample of parameter sets used was generated using Latin Hypercube Sampling [47]. We also investigate the local sensitivity of the estimated parameters ε and ω using the OFAT method.

Partial Rank Correlation Coefficient Method (PRCC)

We generate a sample of 3000 parameter sets using Latin Hypercube Sampling (LHS) [47] and run a Monte Carlo simulation of 500 model runs for each set. The corresponding mean model outcomes are recorded in Table 2.5. A matrix P of parameter sets is formed, with each column corresponding to a different parameter and the two outcomes of interest, d_5 and d_{10}, are attached to P as another two columns, forming the matrix D. The partial correlation coefficients for both outcomes are then calculated using MATLAB's *partialcorr* function.

For the LHS, a triangular distribution, with endpoints a, b and mode c, is used to generate the 3000 parameter sets. The endpoints for each parameter correspond to the minimum (a) and maximum (b) value that parameter can take (Table 2.4), whilst the mode (c) was chosen to be the respective previously fitted values. The interval [0,1] is split into 3000 intervals of equal length. A random value is chosen from each interval and the corresponding parameter value is then calculated (and rounded accordingly).

A vector of 3000 values is thus formed for each of the contact parameters μ_1, μ_2, β_{12}, β_{13}, β_{21} and β_{23}, and the transmission parameter λ. Each vector is randomized separately and the resulting vectors are now the columns of a 3000x7 matrix P describing our sample set. Each row of the matrix corresponds to a parameter set, which is used to run a Monte Carlo simulation of 500 model evaluations. The two mean model outcomes of interest are recorded for each sample set, forming another two vectors of length 3000. Together with the matrix P, they are used as input to the Matlab function *partialcorr* to calculate the partial correlation coefficients and determine an order of sensitivity for the parameters.

One-Factor-at-a-Time Method (OFAT)

We now investigate the effects the parameters ε and ω have on the size of the outbreak. A range of values (Table 2.6) is selected for both the rate of progression from LTBI to active TB disease, ε, and the recovery rate of the converters, ω. The values for ε are minor perturbations to either side of the previously chosen value of 0.0003, whilst the values for ω reflect the statement that the TST has a false negative of *up to* 20% [43]. Both ranges are linearized into 11 points. For each parameter that is being investigated, a Monte Carlo simulation of 2 000 model evaluations is performed for each of the parameter's 11 values, keeping all the other parameters fixed. The corresponding mean model outcomes representing the number of active TB cases and the number of converters after 800 days are then calculated.

Table 2.6. Range of values for ω and ε

Parameter	Range
ω	[0.8, 1.0]
ε	[0.0002, 0.0004]

2.6 Summary

In this chapter we have described how a general conceptual SEIR network model for TB transmission in a long-term care facility may be developed. We have adapted this conceptual model to a specific nursing home in Arkansas as described by Stead [1]. Using data from published literature [1] we have estimated some of the parameters and fit the rest with a two-stage fitting process using the method of Nonlinear Least Squares. Sensitivity analysis for the fitted parameters was done using the Partial Correlation Coefficient Method on a sample of 3000 parameter sets generated by Latin Hypercube Sampling, whilst the One-Factor-at-a-Time method was used for the estimated parameters ε and ω. Results are presented in the next chapter.

3. Results

In this chapter we present our results for the case study of the Arkansas nursing home using the modeling process described in Chapter 2. For each subsection, we first list the main results and then describe how these results are established. Section 3.1 describes the outcome of parameter fitting, which will be used in subsequent simulations. Section 3.2 highlights the results of sensitivity analysis, whilst in Section 3.3 we present the results of a Monte Carlo simulation of the outbreak that occurred in a nursing home in Arkansas, USA, in the late 1970s, as described by Stead [1].

We also investigate two aspects of TB intervention strategies using the fitted Arkansas Model in the setting of the Arkansas nursing home. In Section 3.4, we investigate the impact of reducing diagnosis delay. In Section 3.5, we investigate the effects of various screening programs.

3.1 Parameter Fitting

Using our two-stage fitting process, we have determined the set of parameter values for which the mean model outcomes best fit our data (Tables 2.1 and 2.3).

RESULT 1 - Best-fit parameter values

The parameter values that provide the best fit mean model outcome for our data are:

λ	μ_1	μ_2	β_{12}	β_{21}	β_{13}	β_{23}
0.0990	3	14	86	7	38	12

We now describe the outcomes of the intermediate steps taken in the parameter fitting process.

Stage 1

The five parameter sets with the lowest sum of squares out of the 5000 sampled sets are described in Table 3.1. To select the values for the next stage of the parameter fitting process, we considered the best 1% parameter sets out of the initial 5000. This is equivalent to selecting those 50 parameter sets that resulted in the lowest sums of squares. The three most frequently occurring values for each parameter were selected and whenever multiple values for a parameter had the same frequency across those 50 sets, three values were randomly selected from them. Figure 3.1 describes the frequency of each possible parameter value across these 50 parameter sets, whilst Table 3.2 summarizes the selected values.

Table 3.1. Top 5 Fitted Parameter Sets for Stage 1

	1	**2**	**3**	**4**	**5**
λ	0.0956	0.0959	0.0714	0.0709	0.0243
μ_1	3	3	3	3	3
μ_2	7	8	8	12	30
μ_3	0	0	0	0	0
β_{12}	63	42	91	10	92
β_{21}	30	13	11	9	8
β_{13}	7	23	31	18	28
β_{31}	0	0	0	0	0
β_{23}	10	29	17	18	14
β_{32}	0	0	0	0	0
Sum of squares	271.30	285.85	288.40	293.85	310.68

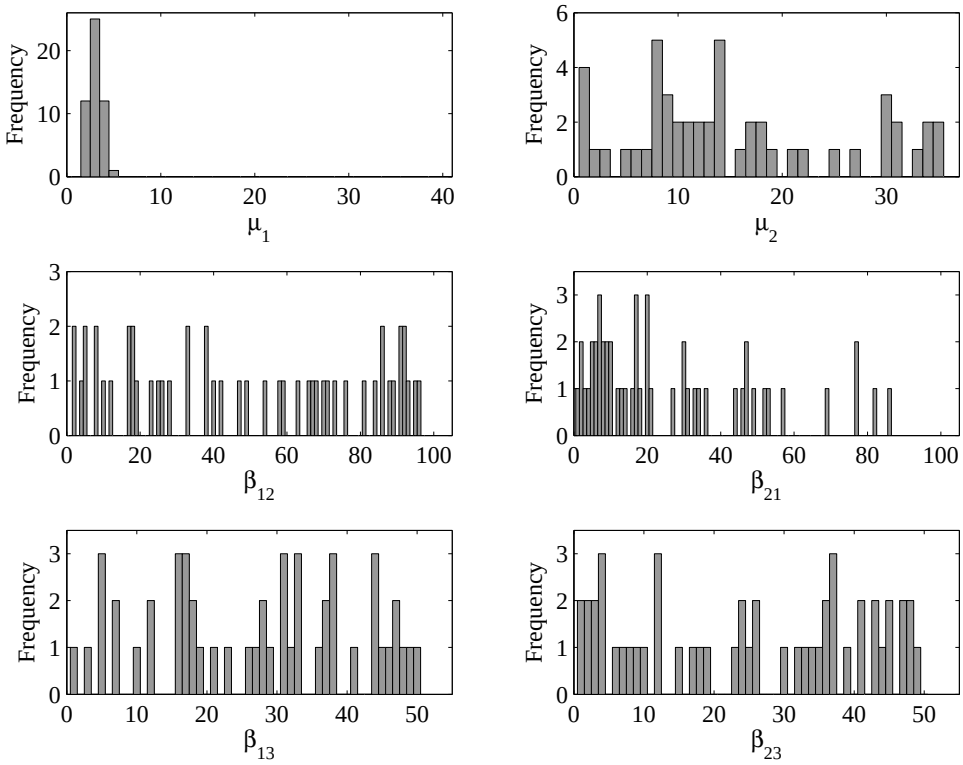


Figure 3.1. Frequency distribution of the parameter values.

Each graph shows the number of occurrences of the corresponding parameter values across the best 50 sets (ordered by least sum of squares) resulting from the random sample of 5000 parameter sets.

Table 3.2. Top 3 most frequent values for each contact parameter.

Values obtained after running the model with each of the 5000 randomly chosen parameter sets. More than 3 values shared the same highest frequency for both β_{12} and β_{13} so the 3 values were randomly selected amongst the most frequent values.

Parameter	Values
μ_1	{2, 3, 4}
μ_2	{1, 8, 14}
β_{12}	{8, 38, 86}
β_{13}	{5, 17, 38}
β_{21}	{7, 17, 47}
β_{23}	{4, 12, 37}

Stage 2

For the second stage of the parameter fitting process, we created a new sample of parameter sets by first linearizing the parameter λ into 12 points between 0.001 and 0.099. The resulting set of values, {0.0010, 0.0099, 0.0188, 0.0277, 0.0366, 0.0455, 0.0545, 0.0634, 0.0723, 0.0812, 0.0901, 0.0990}, together with the selected values for the other parameters as described in Table 3.2, are combined to create all possible combinations of these values. The resulting sample of 8748 parameter sets is then used to find the parameter set with the least sum of squares. Its resulting mean model outcomes are then considered to best fit our data.

We describe the best five parameter sets obtained through this second stage in Table 3.3 and we summarise all the parameters used by the model, estimated and fitted, in Table 3.4.

Table 3.3. Top 5 Fitted Parameter Sets for Stage 2

	1	**2**	**3**	**4**	**5**
λ	0.0990	0.0545	0.0455	0.0990	0.0990
μ_1	3	3	3	3	3
μ_2	14	14	14	8	14
μ_3	0	0	0	0	0
β_{12}	86	86	86	86	86
β_{21}	7	7	17	17	7
β_{13}	38	17	17	38	17
β_{31}	0	0	0	0	0
β_{23}	12	37	12	4	12
β_{32}	0	0	0	0	0
Sum of squares	238.55	253.36	255.71	259.29	259.85

Table 3.4. Parameter Descriptions and Values for the Arkansas Model

Parameter	Definition	Units	Value
γ	Prop. of successfully treated cases of active TB	%	0 or 100
ω	Prop. of converters successfully treated for latent TB	%	80
ε	Rate of progression from latent to active TB	Day^{-1}	3.00×10^{-4}
λ	Transmission probability	$(Contact.Day)^{-1}$	9.90×10^{-2}
	Mean number of contacts within:		
μ_1	Wing 1	Day^{-1}	3
μ_2	Wing 2	Day^{-1}	14
μ_3	Wing 3	Day^{-1}	0
	Number of visits from:		
β_{12}	Wing 1 to Wing 2	Day^{-1}	86
β_{21}	Wing 2 to Wing 1	Day^{-1}	7
β_{13}	Wing 1 to Wing 3	Day^{-1}	38
β_{31}	Wing 3 to Wing 1	Day^{-1}	0
β_{23}	Wing 2 to Wing 3	Day^{-1}	12
β_{32}	Wing 3 to Wing 2	Day^{-1}	0

3.2 Sensitivity Analysis

The main results of the sensitivity analysis performed on our Arkansas Model, with respect to the data set from the Arkansas nursing home, are displayed below. Result 2 describes which of the fitted parameters has the highest and lowest correlation to the size of the outbreak. We describe the ranking of the fitted parameters obtained by the Partial Rank Correlation Coefficient Method in Table 3.5. The two rankings related to the two outcomes, the number of active TB cases (I) and the number of converters (E), are listed separately. The rank is a measure of the correlation between the parameter and the size of the outbreak. The second result summarises the main observations of the local sensitivity analysis performed on ω, the proportion of successfully treated converters, and ε, the rate of progression from latent TB infection to active TB disease. We provide further details supporting these observations in Table 3.6.

RESULT 2 - Global Sensitivity Analysis of the fitted parameters

- Parameter $\boldsymbol{\mu_1}$, the mean number of contacts within Wing 1, has the **highest** correlation to both the number of active TB cases and the number of converters.
- Parameter $\boldsymbol{\beta_{23}}$, the mean number of contacts from Wing 2 to Wing 3**,** has the **lowest** correlation to the number of active TB cases.
- Parameter $\boldsymbol{\beta_{13}}$, the mean number of contacts from Wing 1 to Wing 3, has the **lowest** correlation to the number of converters.

Biological Interpretation

The most sensitive of the fitted parameters is μ_1, which represents the mean number of contacts in wing 1. The value of μ_1 is important to our model particularly because the index case resides in wing 1. The higher μ_1 is, the easier it is for the infection to spread as the index case would have a higher probability of being connected to residents with a larger number of connections.

The least sensitive parameters amongst the ones analysed are β_{13} and β_{23}, representing the total number of visits received by the residents of wing 3, originating from wings 1 and 2, respectively. The values of these parameters have relatively less impact on the size of the outbreak as the residents in wing 3 are mainly confined to their rooms.

Biological Implications

Our sensitivity results imply that, to reduce the potential number of LTBI or active TB cases, it is best to restrict the number of contacts within the index case's wing. A reduction in the number of visits to wing 3 is not strictly necessary as it does not have a major influence on the spread of the disease across the nursing home.

Table 3.5. Ranking of the correlations of the fitted parameters.

Rank from highest (1) to lowest (7) correlation to the size of the outbreak.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Number of Active TB cases (I)	μ_1	β_{21}	μ_2	λ	β_{12}	β_{13}	β_{23}
Number of Converters (E)	μ_1	β_{21}	μ_2	λ	β_{12}	β_{23}	β_{13}

RESULT 3 - Local Sensitivity Analysis of estimated parameters

- Perturbing ε, the progression rate from latent to active TB, has resulted in the most drastic changes to both the number of active TB cases and the number of converters.
- Perturbing ω, the fraction of converters successfully treated for latent TB, does not have any significant effect on either the number of active TB cases or the number of converters.

Biological Implications

Our local sensitivity analysis implies that among the parameters ε and ω, reducing **ε** will have a significant impact on the number of LTBI and active TB cases.

Table 3.6. Local sensitivity analysis of the estimated parameters ω and ε

Parameter	Observation
ω	Decreasing the rate of false negatives resulting from the TST kept the mean number of converters between 34 and 36, and the mean number of active TB cases between 7 and 7.6.
ε	Perturbing the rate of progression from LTBI to active TB disease, from a value of 0.0002 to 0.0004, led to the mean number of converters to increase from 26 to 44, and the mean number of active TB cases to increase from 4.8 to 11.

We now describe in more detail the results just described by including all the relevant values and figures that established our observations.

Partial Rank Correlation Coefficients (PRCC)

Using the mean model outcomes for each of the 3000 parameter sets generated by the Latin Hypercube Sampling method, scatter plots of each of the fitted parameters against each of the two outcomes of interest were plotted. An example of such a scatter plot is shown in Figure 3.2. These plots displayed the nonlinear relationships between the fitted parameters and the two outcomes. The PRCC method was thus used for the global sensitivity analysis of our fitted parameters.

Using MATLAB's *partialcorr* function, we calculated the partial correlation coefficients that measure how sensitive an outcome is to a particular parameter. We first ranked the individual vectors of both parameters and outcomes using MATLAB's *tiedrank* and then used these vectors as input to the *partialcorr* function, to obtain the partial rank correlation coefficients. The higher the absolute value of the coefficient associated with the parameter for some outcome, the higher the correlation between that parameter and that outcome. We present the results in Table 3.7, where both the coefficients and the ranking in terms of sensitivity (1 being the highest) are given.

Table 3.7. Partial Rank Correlation Coefficients for the Fitted Parameters

	μ_1	μ_2	β_{12}	β_{21}	β_{13}	β_{23}	λ
Active TB Cases (I)	0.9644	0.5844	0.0478	0.8104	0.0421	0.0376	0.4656
Rank (I)	1	3	5	2	6	7	4
Converters (E)	0.9138	0.6287	0.0555	0.8142	0.0331	0.0335	0.3132
Rank (E)	1	3	5	2	7	6	4

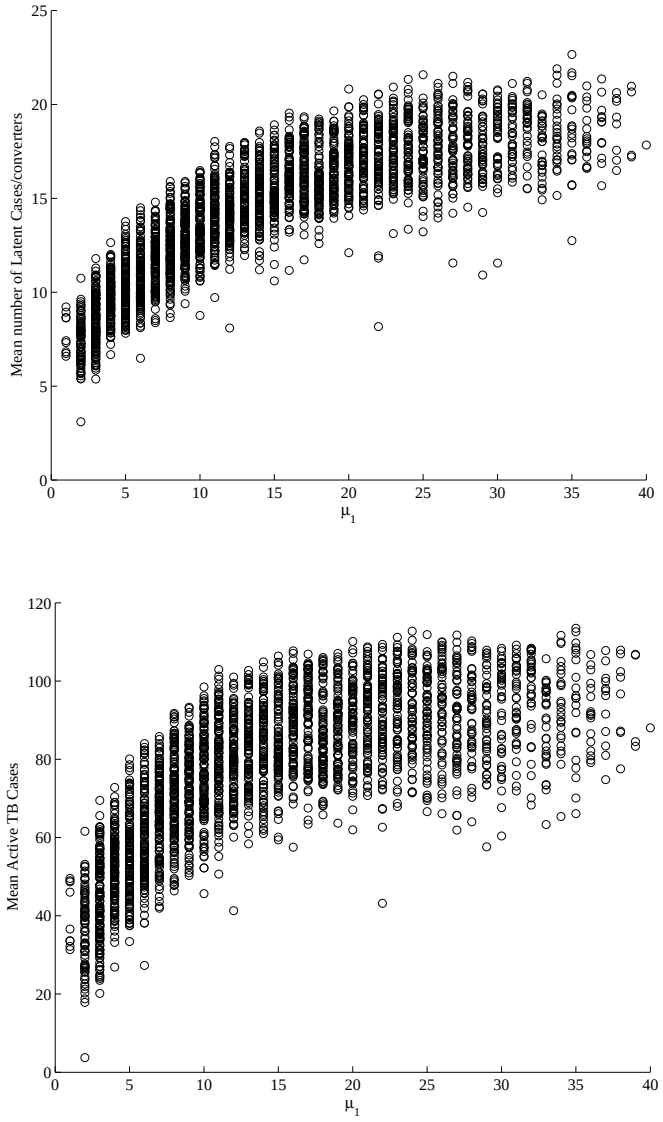


Figure 3.2 Scatter plot of the most sensitive contact parameter (μ_1).

The pattern in both plots shows the nonlinear relationship between (μ_1) and the size of the outbreak, using data obtained from a Monte Carlo simulation of 500 model evaluations performed for each of the 3000 parameter sets generated by Latin Hypercube Sampling.

One-Factor-at-a-Time Method (OFAT)

We show the effects of perturbing the value of ε by 0.00002, from 0.0002 to 0.0004, in Figure 3.3. The top graph describes the changes in the number of converters by the end of the 800 days. The bottom graph shows the changes in the number of active TB cases. We observe that small changes in the value of ε lead to an increase in both the number of converters and the number of active TB cases.

We have also perturbed the parameter ω to investigate its effects on the size of the outbreak. The values of ω reflect different percentages of the false negatives resulting from the Tuberculin Skin Test, which may be up to 20% [43]. Its lowest value is thus our chosen value of 0.8 with increments of 0.02 until the hypothetical perfect testing that can detect all latent TB infections ($\omega = 1$). The changes in the size of the outbreak by the end of the 800 days are shown in Figure 3.4. It is observed that the fluctuations in the value of ω lead to minimal changes to both the total number of converters and the total number of active TB cases.

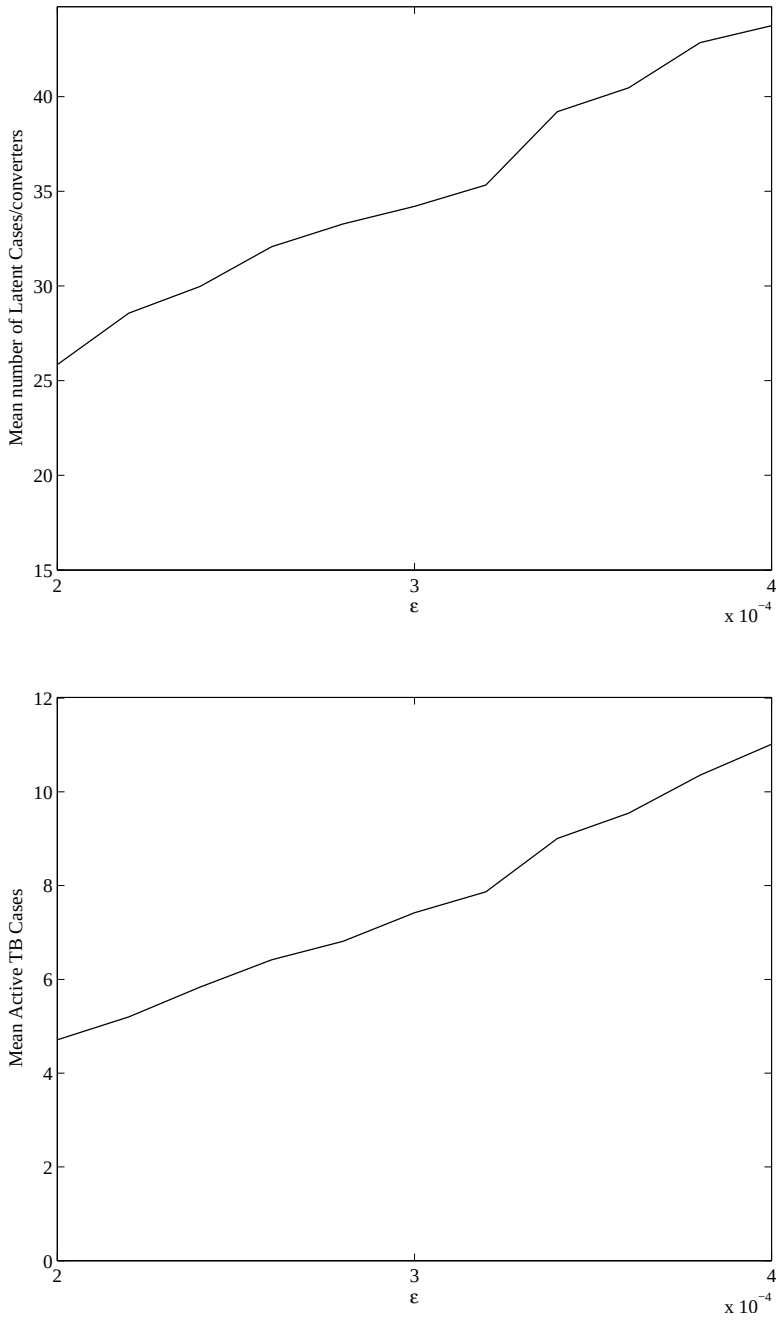


Figure 3.3 Local sensitivity analysis results for ε.

Investigating how the rate of progression (ε) from LTBI to active TB affects the mean number of converters (top) and the mean number of active TB cases (bottom) after a Monte Carlo simulation of 20 000 model evaluations.

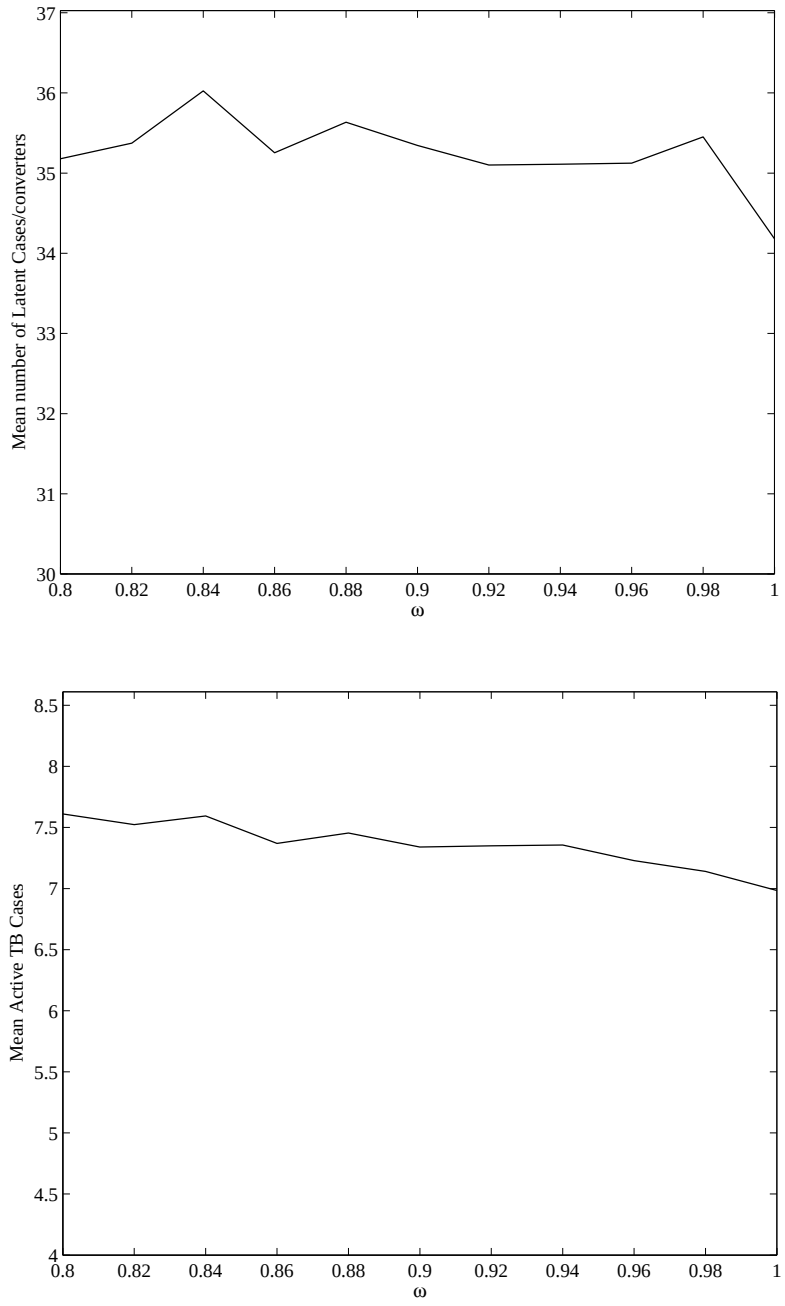


Figure 3.4 Local sensitivity analysis results for ω.

Investigating how the proportion of false negatives resulting from the TST ω affects the mean number of converters (top) and the mean number of active TB cases (bottom) after a Monte Carlo simulation of 20 000 model evaluations.

3.3 Simulating the Outbreak of the Nursing Home in Arkansas, USA

A Monte Carlo simulation consists of a number of model evaluations. For each evaluation, we randomly generate a set of parameter values from chosen parameter distributions, evaluate the model using that set of values and record the outcome. The result of a Monte Carlo simulation is then the mean outcome that is calculated from the outcomes of the individual evaluations of that simulation.

As described in Section 2.1, the general conceptual model is stochastic. Each time the model is evaluated, the connections between the nodes are created randomly. Two connected nodes in one model evaluation do not have to be connected in another and so we make use of Monte Carlo methods. The outcome of a simulation is thus not the result of a single model evaluation but the mean outcome of a series of model evaluations.

We ran a Monte Carlo simulation of 20 000 evaluations of the Arkansas Model. The simulation's mean model outcomes are then compared to the scenario as described by Stead [1].

RESULT 4

The mean outcomes of a Monte Carlo simulation of the Arkansas Model are comparable to Stead's description of the outbreak:

- 3.4% of the population developed active TB disease in our simulated outbreak compared to the 3.3% described by Stead.
- 15.7% of the population contracted LTBI in our simulation compared to the 17.5% described in Stead's outbreak.

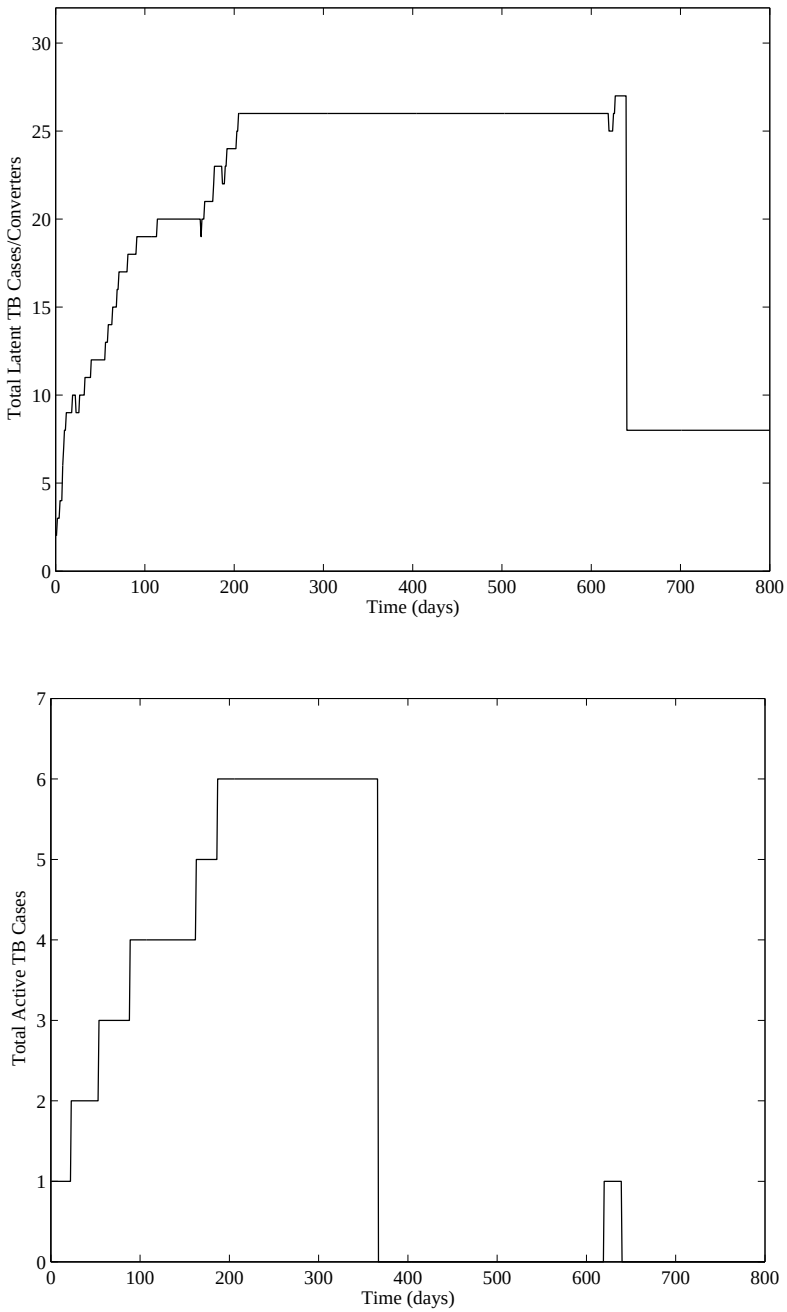


Figure 3.5. Output of the simulated Arkansas Model.

The total number of LTBI/converters (top) and active TB cases (bottom) at each day from one evaluation of the Arkansas Model.

A sample output of a single evaluation of the Arkansas Model includes the graphs shown in Figure 3.5. The top graph describes the number of converters and the bottom graph describes the number of active TB cases, both for each day of the simulated outbreak.

Starting with one active TB case (the index case), we have an increasing number of such infectious cases until around day 200. A corresponding increase can also be seen in the number of latent TB infections. All active TB cases are treated as soon as the index case is discovered at T = 365 days (June 77). However, none of the residents with LTBI is treated and thus that number of cases is steady until in this case, one of them develops active TB disease at around T = 620 days. We call this the second wave of active TB cases. This can be seen by the decrease of one converter and the corresponding increase by one in the number of active TB cases. The probability of our model producing this second wave of active TB cases is 72.1%.

The delay in diagnosis of this new infectious case leads to a few more infections until all known converters and active TB cases are treated at T = 639 (March 79). A number of converters were undiagnosed due to the 20% false negative rate of the Tuberculin skin test and this is reflected in the remaining number of converters left untreated.

The distribution of both the average number of active TB cases and the latent TB infections are described in Figure 3.6. The spread of TB infection across the nursing home is influenced by the index case's contacts. The more highly connected they are, the higher the probability of more residents becoming infected. With the index case being the most connected resident in Wing 1 together with such a low mean number of contacts in Wing 1, there's a higher probability that the index case ends up connected (randomly) to residents which are not very social. We thus see the distribution of both the active TB cases and the latent infections (Figure 3.6) peaking at lower values than the mean. However,

using the mean outcomes after evaluating the model for a sufficient number of times (in our case, 20 000 model evaluations), we obtain a good indication of how our simulated outbreak behaves in comparison to the scenario described by Stead [1].

The mean percentages of the population that were either infected with LTBI or else developed active TB disease throughout the duration of the outbreak, summarised in Result 4, are listed again in Table 3.8. For comparison purposes, we include values for both the actual outbreak as occurred in the original scenario as well as the outbreak simulated by our model.

Table 3.8. Population affected by the outbreak

	Actual (%)	**Simulated (%)**
Active TB Cases	3.3	3.4
Latent Cases/Converters	17.5	15.7

We also compare the distribution of cases amongst the three wings at the end of the outbreak. Both distributions of converters and active TB cases are listed in Table 3.9.

Table 3.9. Percentage distribution of Cases

Wing	**1**	**2**	**3**
Active TB Cases (actual)	**3.3**	**0**	**0**
Mean # of infectious cases (simulated)	2.09	0.95	0.15
Latent Cases/Converters (actual)	**8.75**	**7.5**	**1.25**
Mean # of latent cases/Converters (simulated)	8.83	6.76	0.15

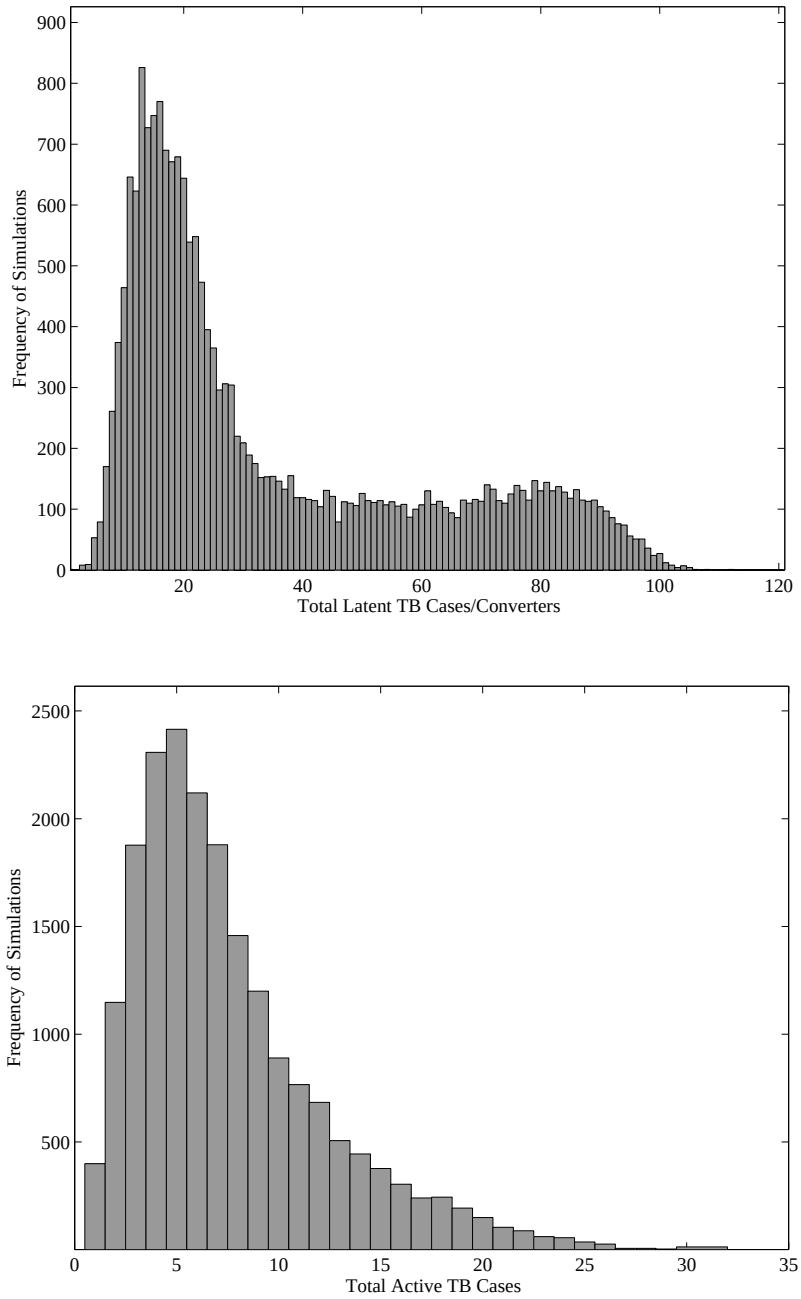


Figure 3.6. A frequency distribution of the total number of converters and active TB cases from one Monte Carlo simulation.

The number of LTBI/converters (top) and active TB cases (bottom) were obtained at the end of an 800-day period from a Monte Carlo simulation of 20 000 evaluations of the Arkansas Model.

3.4 The Impact of Reducing Diagnosis Delay

In the outbreak scenario described by Stead [1], the index case was diagnosed with active TB disease an entire year after the initial misdiagnosis. A second case, a number of months later, was also misdiagnosed, leading to a second wave of infections. We investigate the effect of diagnosis delay on the size of the outbreak and the mean time for the second wave to appear, to give some estimates as to when to be more alert than usual for active TB cases. We assume that active TB cases are successfully treated as soon as they are diagnosed. Therefore, reducing diagnosis delay is crucial for narrowing the window of transmission for undiagnosed cases, and for prevention of potential outbreaks.

RESULT 5

Reducing diagnosis delay to 1 month results in:

- a 47.1% reduction in the number of active TB cases;
- a 25.5% reduction in the number of converters; and
- a mean time T^* of 380 days for a second wave of active TB cases to occur.

We also list, in Table 3.10, the reduction in the number of cases, when compared to the simulated Arkansas outbreak, that resulted from different reduction of diagnosis delay. We also list the mean time it takes for the second wave of active TB cases to occur.

Table 3.10 Reduction in Outbreak Size and Mean Time (T^*) for second wave to occur.

Values resulting from the reduction in diagnosis delay.

X (months)	E (%)	I (%)	Mean time (T^*)
1	25.5	47.1	380
3	21	44.1	425
6	12.1	35.3	490
9	0.6	23.5	552

To obtain these results, we have incorporated into our baseline Arkansas Model both the standard intervention practices, such as immediate treatment of any active TB cases upon diagnosis, follow-up screening and treatment of diagnosed converters for 8 weeks, and a delay of diagnosis of the index case. More specifically, we have made the following modified assumptions:

- Index case is misdiagnosed for X months. We vary X to be equal to 1, 3, 6 and 9 months to investigate how the length of the transmission window ($\gamma = 0$) affects the size of the outbreak.
- Any converters are treated once, 8 weeks after the diagnosis of the index case ($\omega = 0.8$).
- All cases of active TB disease are diagnosed and treated right away ($\gamma = 1$) during the 4 months following the discovery of the first active TB case for that period.
- A resident develops active TB disease after the first wave of cases is considered to be over. We denote this time by T^*. This resident is again misdiagnosed for X months, leading to a second wave of infections, at which time, the same procedure as the one adopted during the first wave is followed.
- The period of time we consider is the same as for the original scenario: Jun 1977 until around Sep 1979 (800 days).

The timeline describing periods of transmission ($\gamma = 0$) and the recovery of both active TB cases ($\gamma = 1$) and converters ($\omega = 0.8$) is detailed in Figure 3.7.

We ran a Monte Carlo simulation of 20 000 model evaluations for each value of X and tabulated the results in Table 3.11, which describes the percentage of the population that has been affected by the outbreak by the end of the first 800 days. The values for the original scenario described in Section 3.3 are included for comparison purposes.

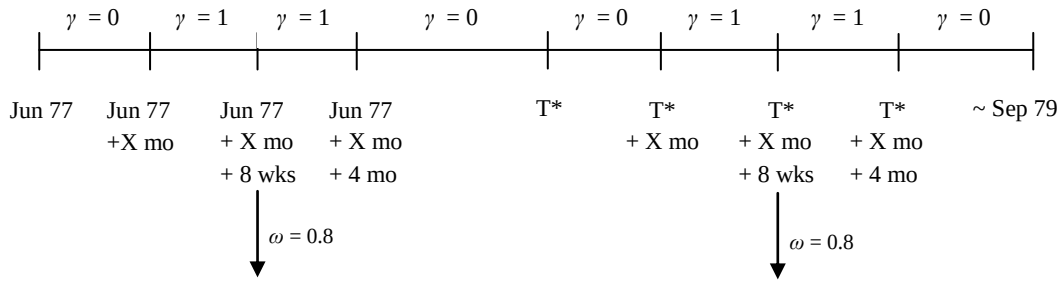


Figure 3.7. Timeline of events as we investigate the effects of diagnosis delay on the size of the outbreak after 800 days.
The length of the transmission window is denoted by X months, indicating the periods of disease transmission ($\gamma = 0$). All active TB cases are treated ($\gamma = 1$) during the 4 months after the discovery of the index case. Treatment of converters occurs once, 8 weeks after the first active TB case is diagnosed.

Table 3.11. Population affected by the outbreak when the transmission period is *X* months long

***X* (months)**	***E* (%)**	***I* (%)**
1	11.7	1.8
3	12.4	1.9
6	13.8	2.2
9	15.6	2.6
Arkansas Model	15.7	3.4

The second wave appears very frequently, independent of the length of the transmission window *X*. What is affected by *X* is the time at which this second wave of infections occurs. Having an idea of when it is more likely for new infections to develop will help the nurses be more alert than usual especially around that time. We describe the probability of having this second wave within the first 800 days and the mean time it takes for it to appear in Table 3.12.

Table 3.12. Probability of obtaining a second wave of infections within 800 days and the mean time for it to appear.

Values obtained when the transmission period is *X* months long.

***X* (months)**	**Probability (%)**	**Mean time (T*)**
1	83.3	380
3	82.3	425
6	77.7	490
9	73.1	552
Arkansas Model	72.1	545

A sample model output of a single evaluation of one of the simulations with the transmission window being one month long is shown in Figure 3.8. Translating one month into 30 days, the bottom graph clearly shows that starting with the index case, he is misdiagnosed for 30 days during which time, as shown in the top graph, he infected a number of residents. At the end of that month, he is diagnosed and thus treated, which is reflected in the drop of the number of active TB cases. This results in the number of converters being constant as none of them had yet become infectious.

After eight weeks from the diagnosis of the index case, all known converters are treated and thus we see a drop in their number. In this case, none of the converters became infectious until just before $T = 400$ days. This new active

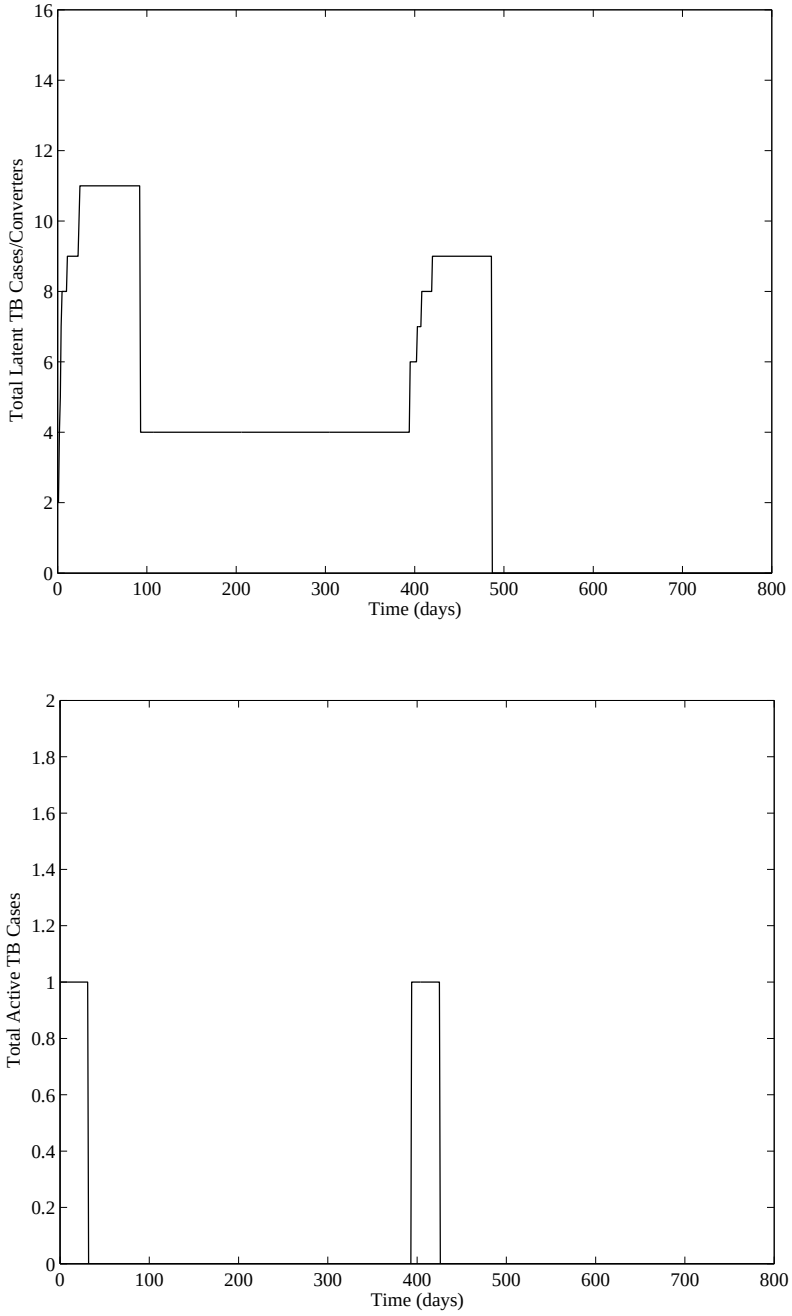


Figure 3.8. The daily total number of LTBI/converters (top) and active TB cases (bottom) from one model evaluation of the Arkansas model with reduced diagnosis delay.

TB case is again misdiagnosed for one month, during which a number of residents got infected. The same procedure is followed, where the active TB case is diagnosed and treated after one month and the diagnosed converters are treated after eight weeks (56 days).

3.5 Evaluation of Potential Screening Programs

During the Arkansas outbreak, the nursing home had no regular screening program for its residents. Employees were tested yearly, and the importance of having at least that type of screening program in place for the employees is shown by the fact that the investigations which eventually led to discovery of the index case were triggered by a public nurse's observation that a number of employees had converted over the previous year.

We would like to investigate the effects of implementing different screening programs for the residents. Testing all residents on a monthly basis would significantly decrease the possibility of a TB outbreak occurring in the nursing home. However, this is often not feasible, as it puts a large burden on the nurses and also inconveniences the residents. We look at the possibility of testing a percentage Y of the susceptible population every X number of months. If an active TB case is encountered, it is treated right away. Additionally, if a proportion Z of the *tested* population is found to have converted during the screening, we look for any active TB cases and treat them right away.

We compare the effectiveness of the different screening programs by measuring the total number of converters and the total number of active TB cases that occurred throughout the 800-day period. The most successful screening program is the one resulting in the least number of LTBI and active TB cases.

RESULT 6

Testing ***Y***% of the susceptible population for LTBI every ***X*** months produces a comparable reduction in both the number of active TB cases and the number of converters, to testing (***nY***)% every (***nX***) months. For example:

- testing 25% of the susceptible population every 3 months;
- testing 50% every 6 months; and
- testing 75% every 9 months,

all lead to approx. 2.6% of the resident population developing active TB disease and 18.2% contracting LTBI.

The flowcharts in Figures 3.9(a) and 3.9(b) describe the algorithm developed for the implementation of different screening programs.

Tables 3.14(a), 3.14(b) and 3.14(c) reflect the percentages (converters, active TB cases) of the population that were affected by the outbreak when a screening program with the appropriate parameters described above is included in our simulation. Once again, a Monte Carlo simulation of 20 000 model evaluations is performed and mean model outcomes are used.

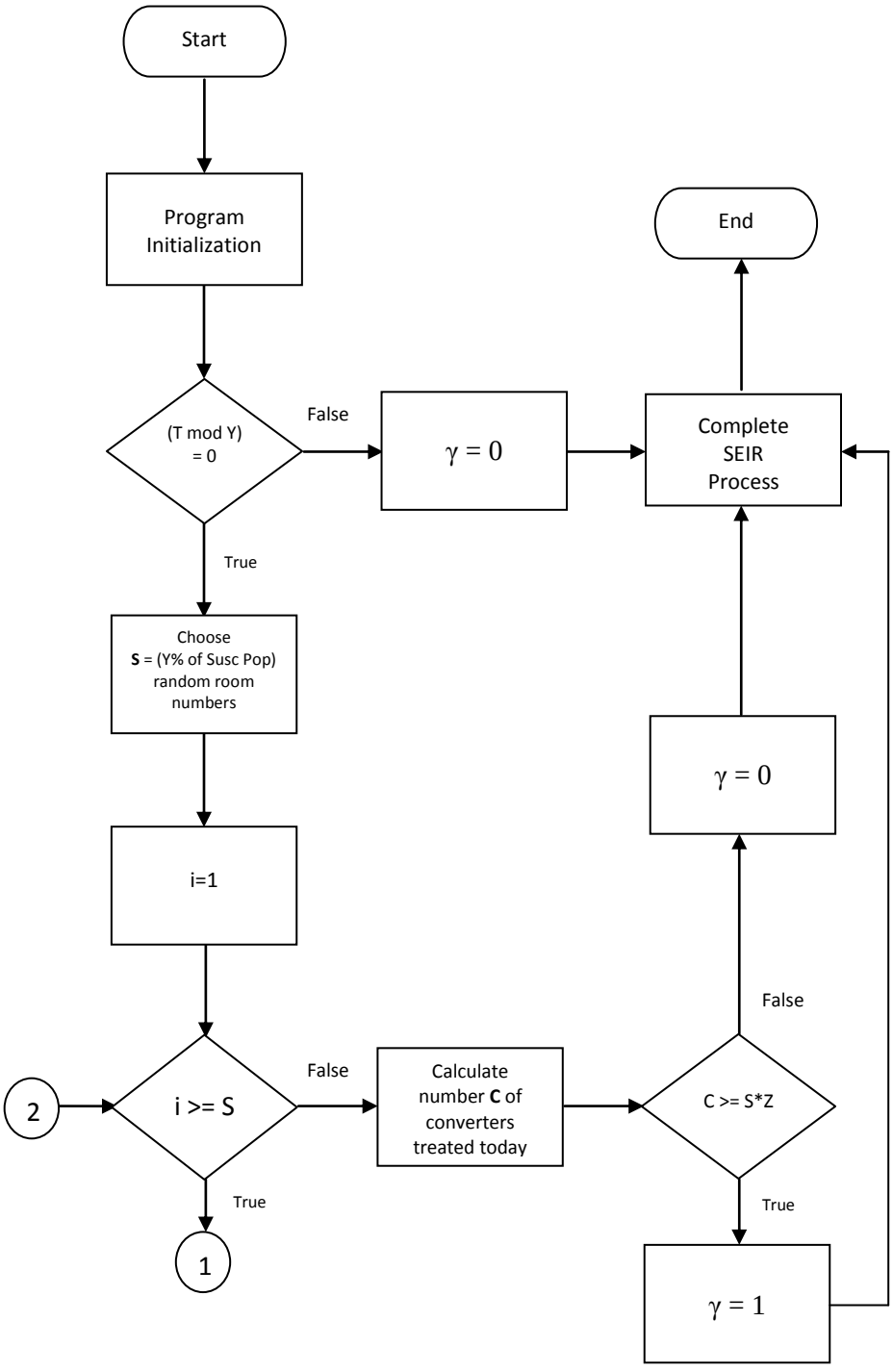


Figure 3.9 (a). Flow diagram describing the algorithm for a screening program.

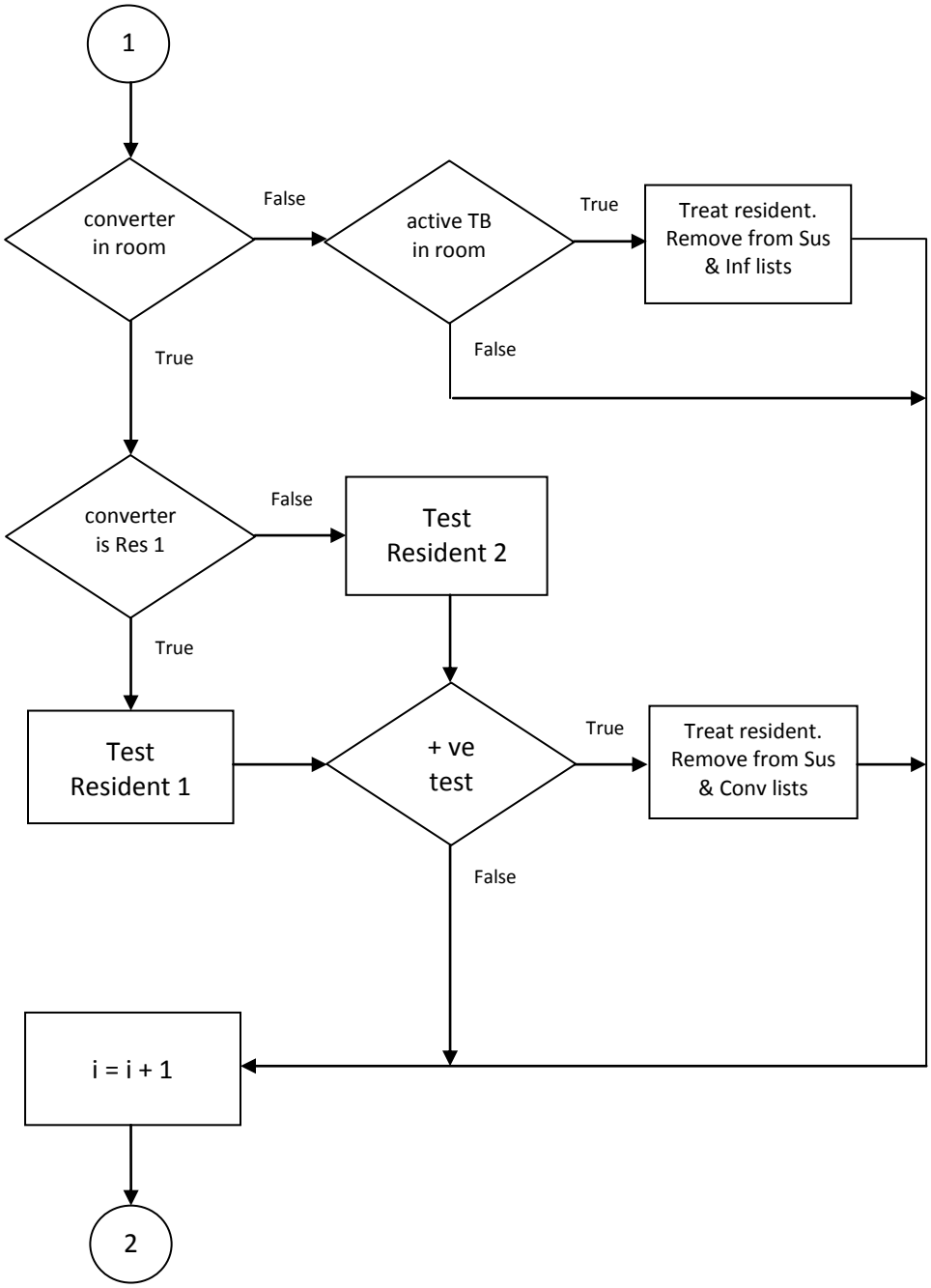


Figure 3.9 (b). Flow diagram describing the algorithm for a screening program (cont.).

Table 3.14 (a). Percentages (*E*, *I*) of population affected by outbreak with screening program in place (*Z* = 0.05)

Z = 0.05	*X* months			
Y (%)	1	3	6	9
25	(12.6, 1.7)	(17.8, 2.6)	(20.4, 3.3)	(21.8, 3.8)
50	(11.4, 1.4)	(15.1, 2.0)	(18.0, 2.6)	(20.3, 3.2)
75	(10.7, 1.3)	(13.5, 1.7)	(16.3, 2.2)	(18.5, 2.7)
100	(10.3, 1.3)	(12.5, 1.5)	(15.0, 1.9)	(16.9, 2.3)

Table 3.14 (b) Percentages (*E*, *I*) of population affected by outbreak with screening program in place (*Z* = 0.10)

Z = 0.10	*X* months			
Y (%)	1	3	6	9
25	(13.5, 1.7)	(17.9, 2.6)	(20.7, 3.3)	(21.8, 3.8)
50	(11.9, 1.4)	(15.2, 1.9)	(18.3, 2.6)	(20.1, 3.1)
75	(11.1, 1.3)	(13.7, 1.7)	(16.4, 2.2)	(18.3, 2.6)
100	(10.3, 1.3)	(12.7, 1.5)	(14.9, 1.9)	(16.8, 2.3)

Table 3.14 (c) Percentages (*E*, *I*) of population affected by outbreak with screening program in place (*Z* = 0.15)

Z = 0.15	*X* months			
Y (%)	1	3	6	9
25	(14.0, 1.7)	(18.4, 2.7)	(20.8, 3.4)	(22.0, 3.8)
50	(12.0, 1.4)	(15.4, 2.0)	(18.4, 2.6)	(20.2, 3.2)
75	(11.0, 1.3)	(13.8, 1.7)	(16.5, 2.2)	(18.7, 2.7)
100	(10.3, 1.3)	(12.8, 1.5)	(15.0, 1.9)	(17.0, 2.3)

3.6 Summary

In this section we have presented the results from our parameter fitting and sensitivity analysis processes. We have described the kind of output one may obtain from our model, by taking the outbreak described by Stead [1] as an example. Two aspects of TB intervention strategies are investigated using the fitted model. One aspect focuses on reducing diagnosis delay, and the other is the investigation of different screening programs.

4. Discussion

In this concluding chapter, we first give a summary of main results in the thesis in Section 4.1, and then, in Section 4.2, we discuss some limitations of our TB network model and suggest possible solutions that may be implemented in the future.

4.1 Summary

Our main objective is to model TB transmission in long-term care facilities such as nursing homes, for the purpose of assessing the effectiveness of TB intervention and management strategies at these facilities. We have developed a general conceptual SEIR network model for TB transmission in long-term care facilities. To demonstrate the applicability of the conceptual model, we have built a specific network model to simulate a TB outbreak that occurred in a nursing home in Arkansas, USA, between 1977 and 1979, as described by Stead [1]. We have implemented two aspects of different TB intervention strategies into the Arkansas Model and assessed their effectiveness measured by the reduction in the number of LTBI and active TB cases within a given period of time.

We have investigated the impact of reducing diagnosis delay of an infectious case using our Arkansas Model, by varying the duration of the transmission window of an infectious TB case while adopting the current practice of contact tracing and treatment of converters after 8 weeks following the diagnosis of an infectious case. Our results show that if the transmission window of an infectious TB case is cut to within a month, the number of converters at the Arkansas nursing home can be reduced by 25%, whilst the number of active TB cases will be reduced by 47%.

We have also implemented a hypothetical screening program into our Arkansas Model, to investigate the impact of testing different percentages of the susceptible population at different time intervals. We observed a useful linear relationship between the percentage of susceptibles screened and the frequency of screening. For instance, in the case of the Arkansas nursing home, screening 25% of the susceptible population every 3 months will achieve a similar result to screening 50% every 6 months. This suggests that, for the same number of staff nurses at the nursing home, to test a smaller percentage of residents more frequently is preferable than testing a larger population less frequently.

Results in this thesis have demonstrated feasibility and usefulness of using network models to study control and prevention of transmission of TB in long-term care facilities. We expect that the same modeling approach can be applied to the control and prevention of other disease transmissions in a similar setting.

4.2 Limitations and Future Work

Our network's static nature prevents the number of contacts and the actual connections themselves from changing over time. This closed and static system does not take into account any admissions to the nursing home. Any deaths and outside visitors are also excluded from the model. A dynamic network may be

created to reflect the nature of human relationships as well as the changing population of the nursing home. The extra stochasticity introduced by having such a dynamic network will increase the time required to run the model simulations, with the parameter fitting process taking considerably more time. Our Arkansas model produced results that were comparable to the Arkansas outbreak as described by Stead [1]. The introduction of a dynamic network for more accurate results will thus not result in a significant improvement.

The parameter fitting process did not identify a single prominent unique set of parameters as the data extracted from the description provided by Stead [1] was not enough. The top five parameter sets, shown in Table 3.3, produce similar outcomes. To overcome this, more data needs to be obtained. Another possibility is talking to a physician to acquire a better understanding of which parameter values make the most sense. This would lead us to restrict further the range of values each parameter may take, reducing the sample space of parameter values. This will result in a more effective parameter fitting process.

For our model, we have implemented two aspects of different TB intervention strategies. Additional strategies may be implemented and compared in order to determine whether there are better alternatives to the strategies that are currently in use. One such alternative would be the inclusion of staff members and their interactions with the resident population. This provides an opportunity for a different type of screening, where the impact of regularly screening a percentage of the staff rather than the residents may be investigated.

Costs associated with treating LTBI and active TB cases may also be included in the model to assess the cost-effectiveness of the different investigated strategies. Such assessments may further aid policy makers in decision-making. This network model may also be adapted to model transmission of other infectious diseases in similar settings, including detention centers, hospitals and aboriginal reserves.

Bibliography

[1] W.W. Stead, Tuberculosis among elderly persons: an outbreak in a nursing home, Ann. Intern. Med. 94 (1981) 606-610.

[2] E. Ellis, K. Dawson, V. Gallant, M. Phypers, D. Scholten, Tuberculosis in Canada 2009: Pre-release, HP37-5/2009E-PDF (2010) 1.

[3] M.L. Hopkins, L. Schoener, Tuberculosis and the elderly living in long-term care facilities, Geriatr. Nurs. 17 (1996) 27-32.

[4] L.J. Strausbaugh, S.R. Sukumar, C.L. Joseph, Infectious disease outbreaks in nursing homes: an unappreciated hazard for frail elderly persons, Clin. Infect. Dis. 36 (2003) 870-876.

[5] D. Menzies, M. Lewis, O. Oxlade, Costs for tuberculosis care in Canada, Can. J. Public Health. 99 (2008) 391-396.

[6] L. Thrupp, S. Bradley, P. Smith, A. Simor, N. Gantz, K. Crossley, M. Loeb, L. Strausbaugh, L. Nicolle, SHEA Long-Term-Care Committee, Tuberculosis prevention and control in long-term-care facilities for older adults, Infect. Control Hosp. Epidemiol. 25 (2004) 1097-1108.

[7] K. Ijaz, J.A. Dillaha, Z. Yang, M.D. Cave, J.H. Bates, Unrecognized tuberculosis in a nursing home causing death with spread of tuberculosis to the community, J. Am. Geriatr. Soc. 50 (2002) 1213-1218.

[8] S.K. Teo, Tuberculosis--barriers to early diagnosis, Singapore Med. J. 43 (2002) 169-171.

[9] F. Brauer, Mathematical epidemiology is not an oxymoron, BMC Public Health. 9 Suppl 1 (2009) S2.

[10] C. Castillo-Chávez, D. Earn, M. Martcheva, M. Nuno, J. Watmough, F. Brauer, Pauline van den Driessche, J. Wu, Mathematical Epidemiology (Lecture Notes in Mathematics / Mathematical Biosciences Subseries) Springer.

[11] R. Long, E. Ellis, K. Elwood, Canadian Tuberculosis Standards / Editors, Richard Long, Edward Ellis ; Associate Editors, Kevin Elwood ... [Et Al.], [Ottawa] : Public Health Agency of Canada/Agence de la santé publique du Canada, 2007; 6th ed, 2007.

[12] D. Edwards, C.H. Kirkpatrick, The immunology of mycobacterial diseases, Am. Rev. Respir. Dis. 134 (1986) 1062-1071.

[13] Diagnostic Standards and Classification of Tuberculosis in Adults and Children. This official statement of the American Thoracic Society and the Centers for Disease Control and Prevention was adopted by the ATS Board

of Directors, July 1999. This statement was endorsed by the Council of the Infectious Disease Society of America, September 1999, Am. J. Respir. Crit. Care Med. 161 (2000) 1376-1395.

[14] D.W. Bentley, Tuberculosis in long-term care facilities, Infect. Control Hosp. Epidemiol. 11 (1990) 42-46.

[15] S. Rajagopalan, T.T. Yoshikawa, Tuberculosis in long-term-care facilities, Infect. Control Hosp. Epidemiol. 21 (2000) 611-615.

[16] Centers for Disease Control and Prevention (CDC), Prevention and control of tuberculosis in facilities providing long-term care to the elderly. Recommendations of the Advisory Committee for Elimination of Tuberculosis, MMWR Recomm Rep. 39 (1990) 7-13.

[17] M.D. Hutton, G.M. Cauthen, A.B. Bloch, Results of a 29-state survey of tuberculosis in nursing homes and correctional facilities, Public Health Rep. 108 (1993) 305-314.

[18] T.T. Yoshikawa, Tuberculosis in aging adults, J. Am. Geriatr. Soc. 40 (1992) 178-187.

[19] W.W. Stead, J.P. Lofgren, E. Warren, C. Thomas, Tuberculosis as an endemic and nosocomial infection among the elderly in nursing homes, N. Engl. J. Med. 312 (1985) 1483-1487.

[20] E.H. Steimke, M.F. Tenholder, M.I. McCormick, J.P. Rissing, Tuberculosis surveillance: lessons from a cluster of skin test conversions, Am. J. Infect. Control. 22 (1994) 236-241.

[21] R. Ross , The Prevention of Malaria / by Ronald Ross, Murray, London, 1910.

[22] W.O. Kermack, A.G. McKendrick, A Contribution to the Mathematical Theory of Epidemics, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character. 115 (1927) 700-721.

[23] S. Blower, A. Mclean, T. Porco, P. Small, P. Hopewell, M. Sanchez, A. Moss, The intrinsic transmission dynamics of tuberculosis epidemics, Nat. Med. 1 (2006) 10-821.

[24] E. Ziv, C.L. Daley, S.M. Blower, Early therapy for latent tuberculosis infection, Am. J. Epidemiol. 153 (2001) 381-385.

[25] C. Castillo-Chavez, Z. Feng, Global stability of an age-structure model for TB and its applications to optimal vaccination strategies, Math. Biosci. 151 (1998) 135-154.

[26] T. Lietman, S.M. Blower, Potential Impact of Tuberculosis Vaccines as Epidemic Control Agents, Clinical Infectious Diseases. 30 (2000) S316.

[27] M.G.M. Gomes, G.F. Medley, A.O. Franco, M.C. Gomes, The Reinfection Threshold Promotes Variability in Tuberculosis Epidemiology and Vaccine Efficacy, Proceedings: Biological Sciences. 271 (2004) 617-623.

[28] T. Cohen, M. Murray, Incident Tuberculosis among Recent US Immigrants and Exogenous Reinfection, Emerging Infectious Diseases. 11 (2005) 725-728.

[29] T. Cohen, C. Colijn, B. Finklea, M. Murray, Exogenous re-infection and the dynamics of tuberculosis epidemics: local effects in a network model of transmission, Journal of the Royal Society Interface. 4 (2007) 523-531.

[30] C. Castillo-Chavez, Z. Feng, To treat or not to treat: the case of tuberculosis, J. Math. Biol. 35 (1997) 629-656.

[31] T. Cohen, M. Lipsitch, R.P. Walensky, M. Murray, Beneficial and perverse effects of isoniazid preventive therapy for latent tuberculosis infection in HIV-tuberculosis coinfected populations, Proceedings of the National Academy of Sciences. 103 (2006) 7042-7047.

[32] C. Colijn, M. Murray, T. Cohen, Mathematical Models of Tuberculosis: Accomplishments and Future Challenges, in: BIOMAT 2006, World Scientific, 2007, pp. 123-148.

[33] M.J. Keeling, K. Eames, Networks and epidemic models, Journal of the Royal Society Interface. 2 (2005) 295-307.

[34] M.E.J. Newman, Spread of epidemic disease on networks, Phys Rev E. 66 (2002) 016128.

[35] L.A. Meyers, M.E.J. Newman, B. Pourbohloul, Predicting epidemics on directed contact networks, J. Theor. Biol. 240 (2006) 400-418.

[36] N.M. Ferguson, D.A.T. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, D.S. Burke, Strategies for containing an emerging influenza pandemic in Southeast Asia, Nature. 437 (2005) 209-214.

[37] J.M. Read, M.J. Keeling, Disease Evolution on Networks: The Role of Contact Structure, Proceedings: Biological Sciences. 270 (2003) pp. 699-708.

[38] I.M. Longini, M.E. Halloran, A. Nizam, Y. Yang, Containing Pandemic Influenza with Antiviral Agents, American Journal of Epidemiology. 159 (2004) 623-633.

[39] Tuberculosis surveillance and screening for long term care facilities in Colorado, Colorado State Publications Library Digital Repository 2012.

[40] A. Ayub, S.H. Yale, K.D. Reed, R.M. Nasser, S.R. Gilbert, Testing for Latent Tuberculosis, Clinical Medicine & Research. 2 (2004) 191-194.

[41] Federal Bureau of Prisons, Washington D.C., Clinical Practice Guidelines: Management of Tuberculosis, (2010).

[42] P.S. Levy, S. Lemeshow, Sampling of populations: methods and applications, in: Hoboken, N.J. : Wiley, c2008; 4th ed, 2008, pp. 340.

[43] D.N. Rose, C.B. Schechter, J.J. Adler, Interpretation of the tuberculin skin test, J. Gen. Intern. Med. 10 (1995) 635-642.

[44] A. Saltelli, M. Ratto, S. Tarantola, F. Campolongo, Sensitivity analysis practices: Strategies for model-based inference, Reliab. Eng. Syst. Saf. 91 (2006) 1109-1125.

[45] C. Daniel, One-at-a-Time Plans, Journal of the American Statistical Association. 68 (1973) 353-360.

[46] S. Marino, I.B. Hogue, C.J. Ray, D.E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology, J. Theor. Biol. 254 (2008) 178-196.

[47] J.C. Helton, F.J. Davis, Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, Reliability Engineering and System Safety. 81 (2003) 23-69.

Appendix

Flowcharts

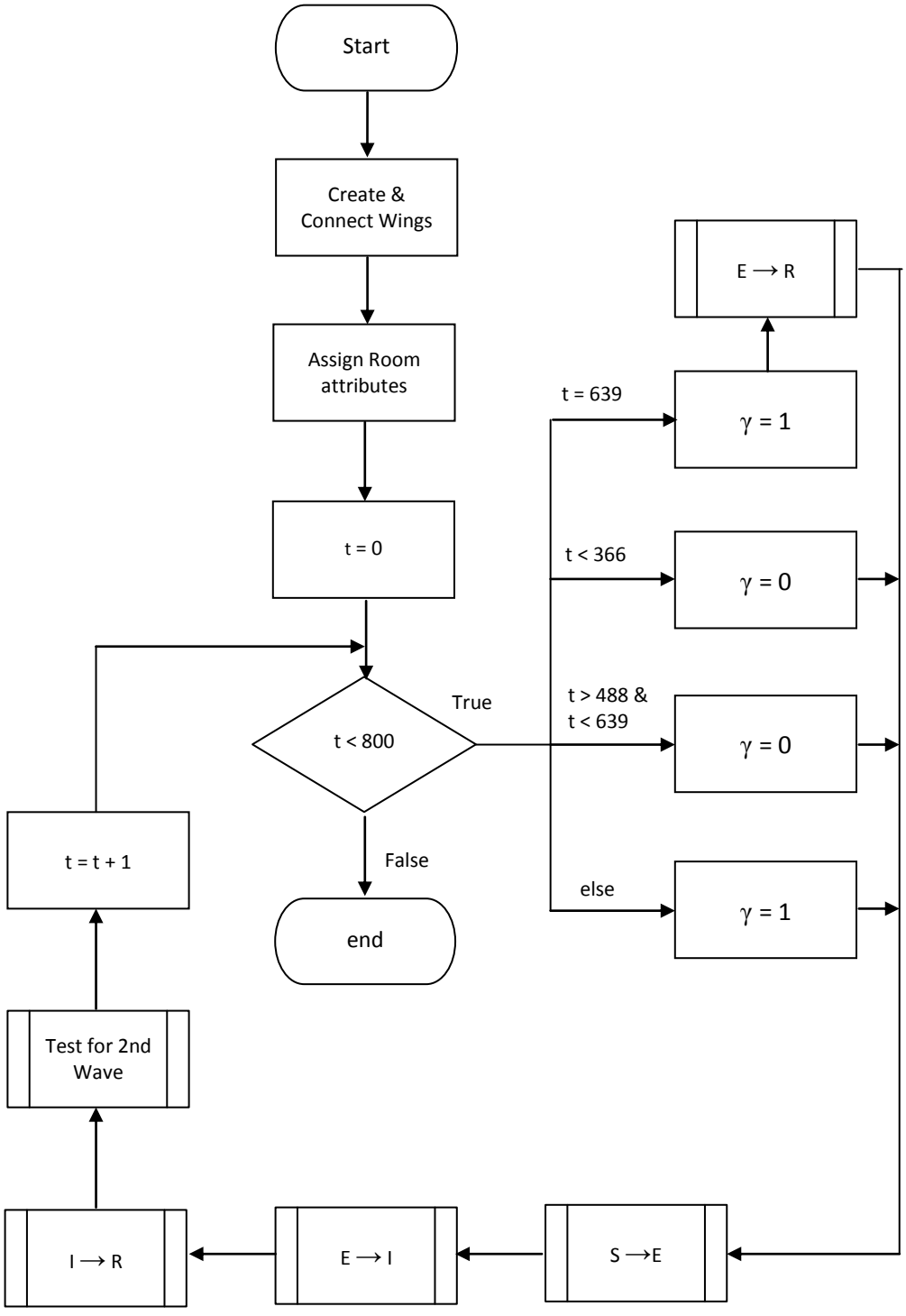


Figure A.1. Overall program structure

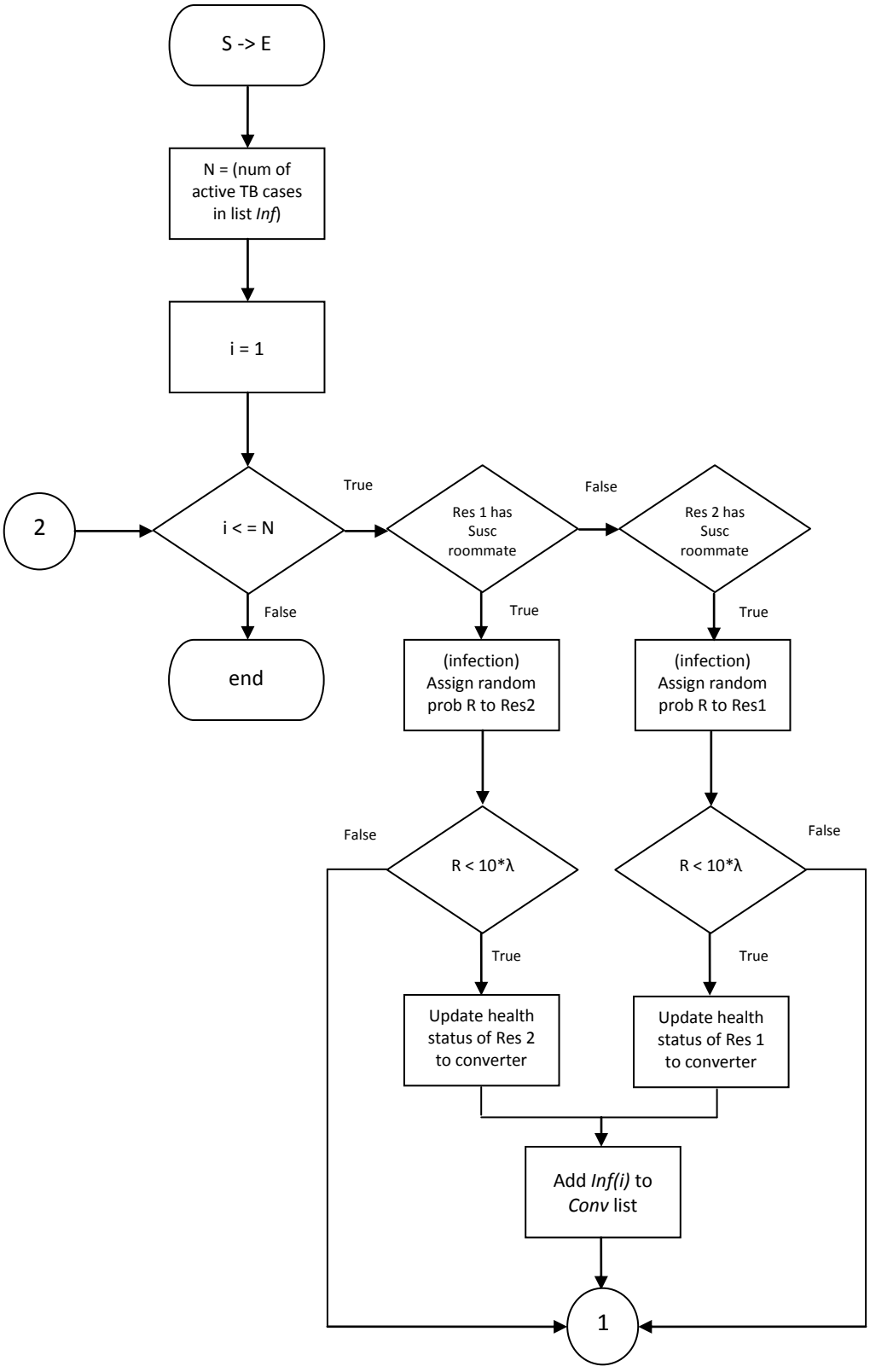


Figure A.2. Subroutine for the S → E process

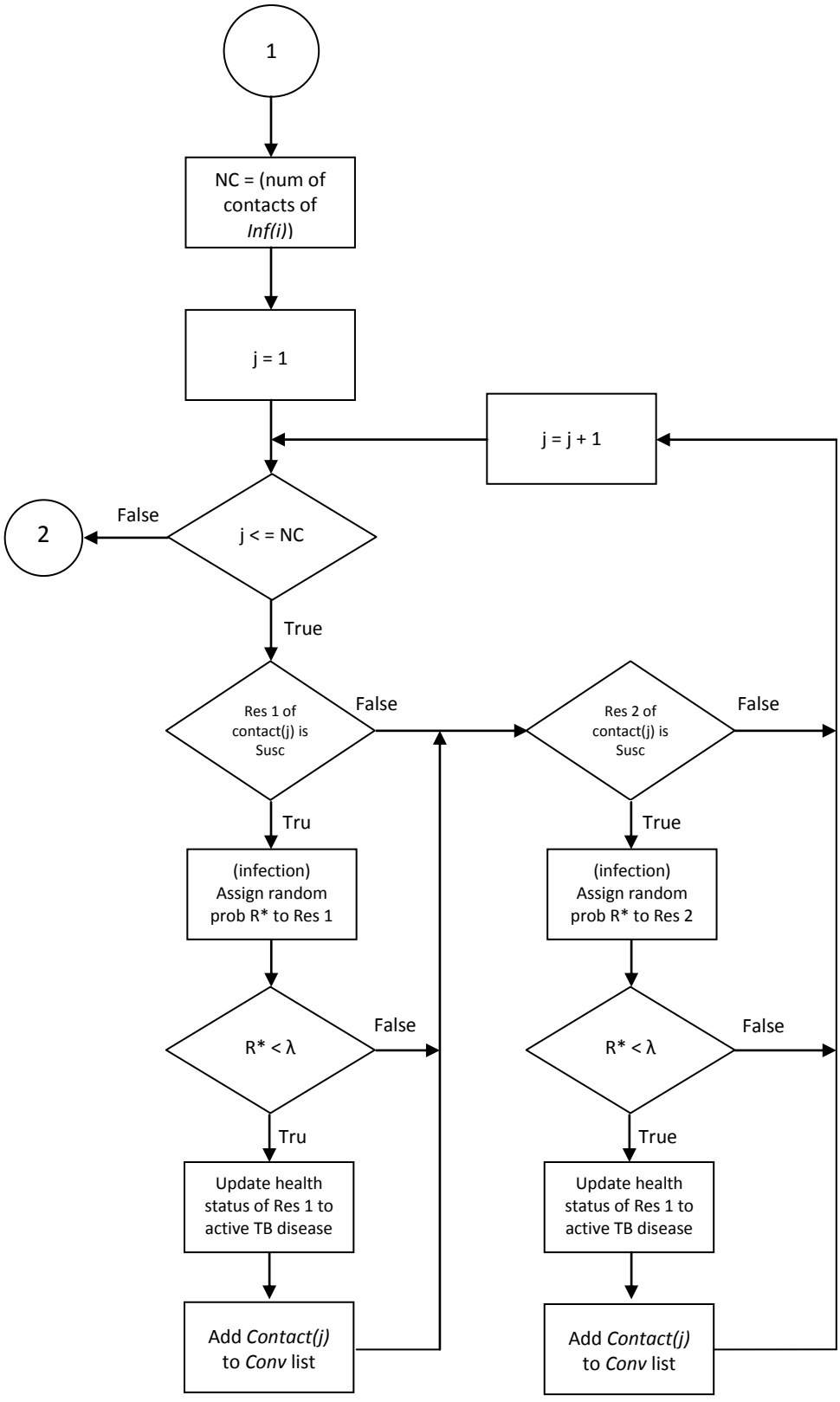


Figure A.2. Subroutine for the S → E process (cont.)

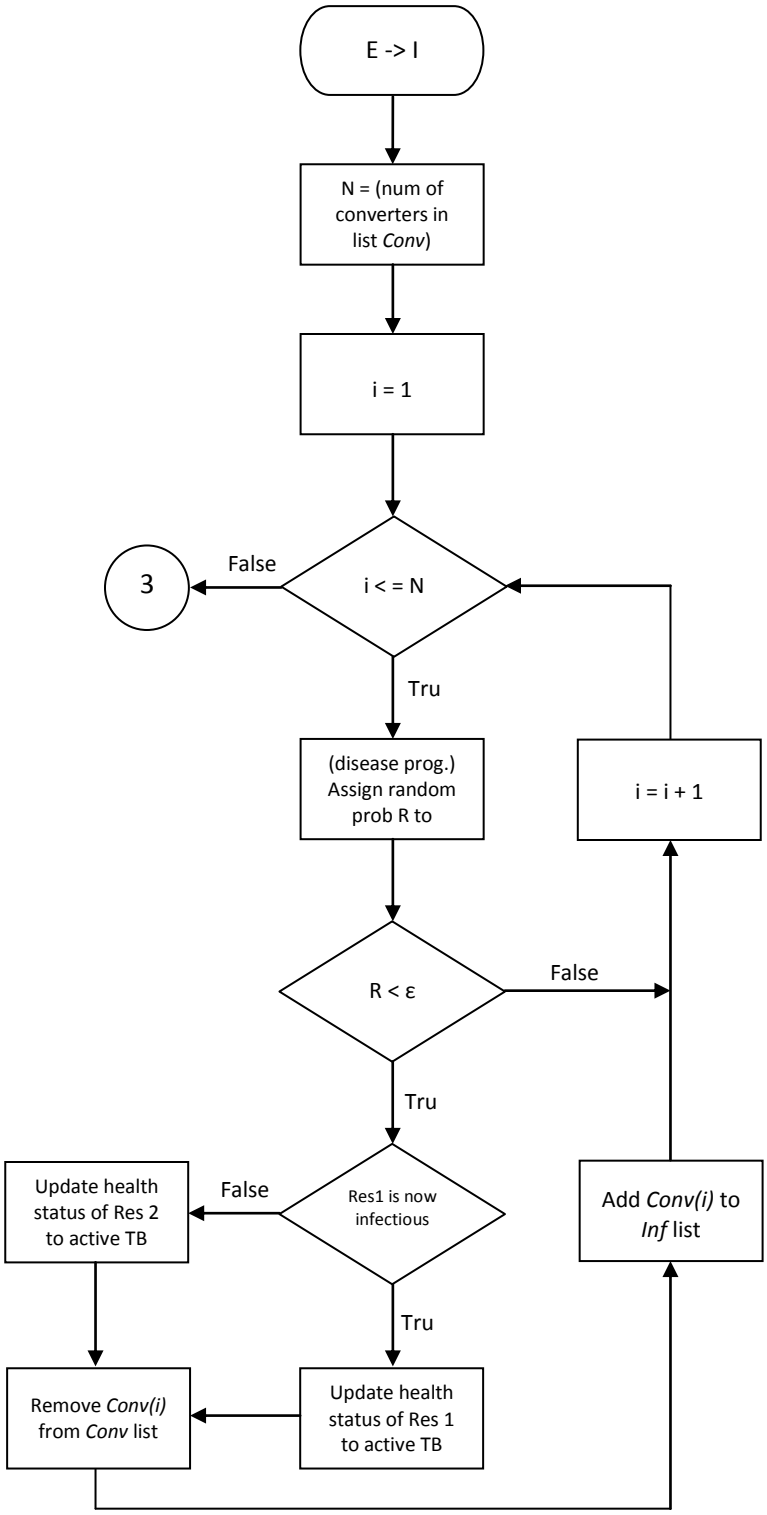


Figure A.3. Subroutine for the E → I process

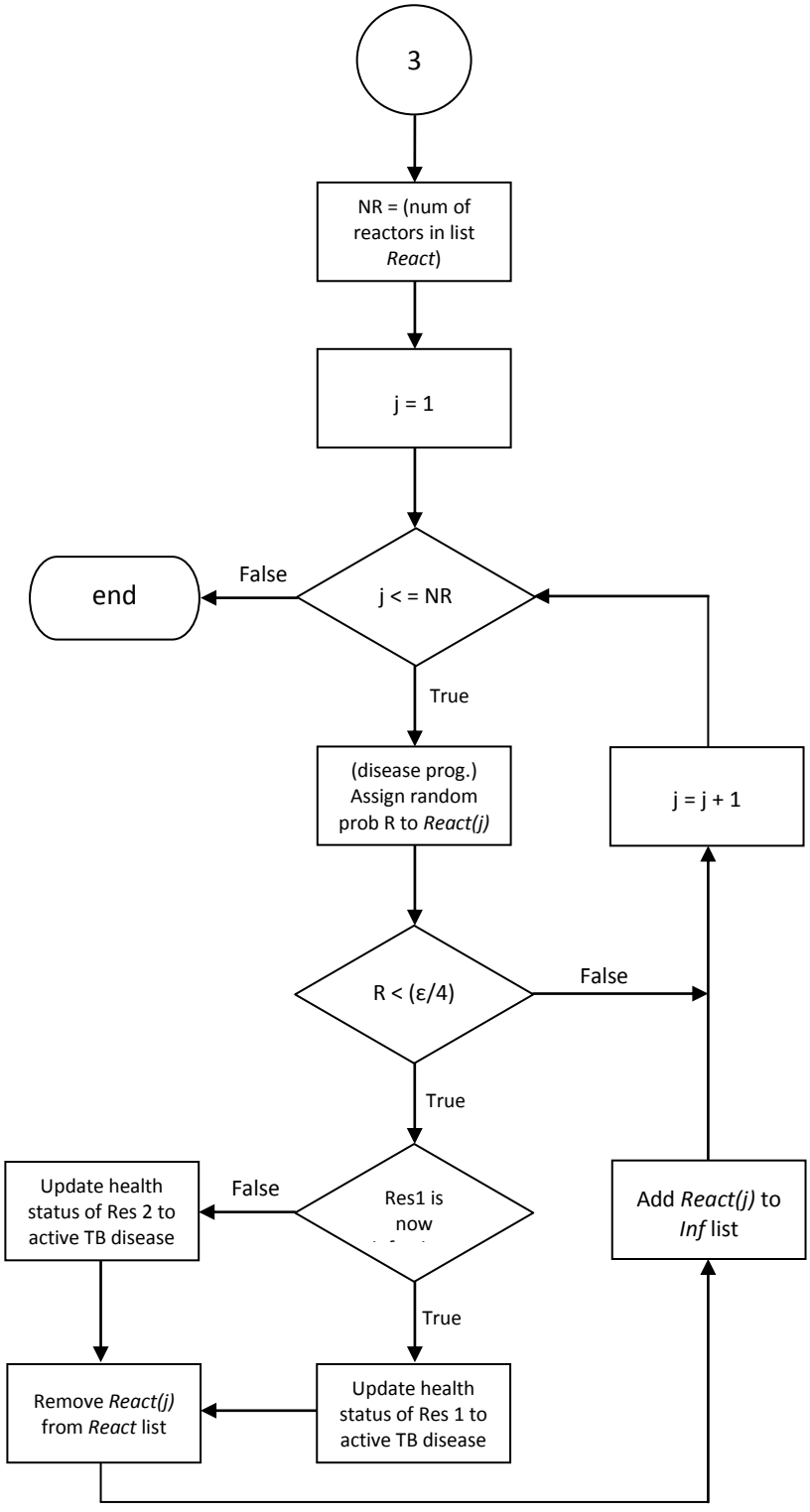


Figure A.3. Subroutine for the E → I process (cont.)

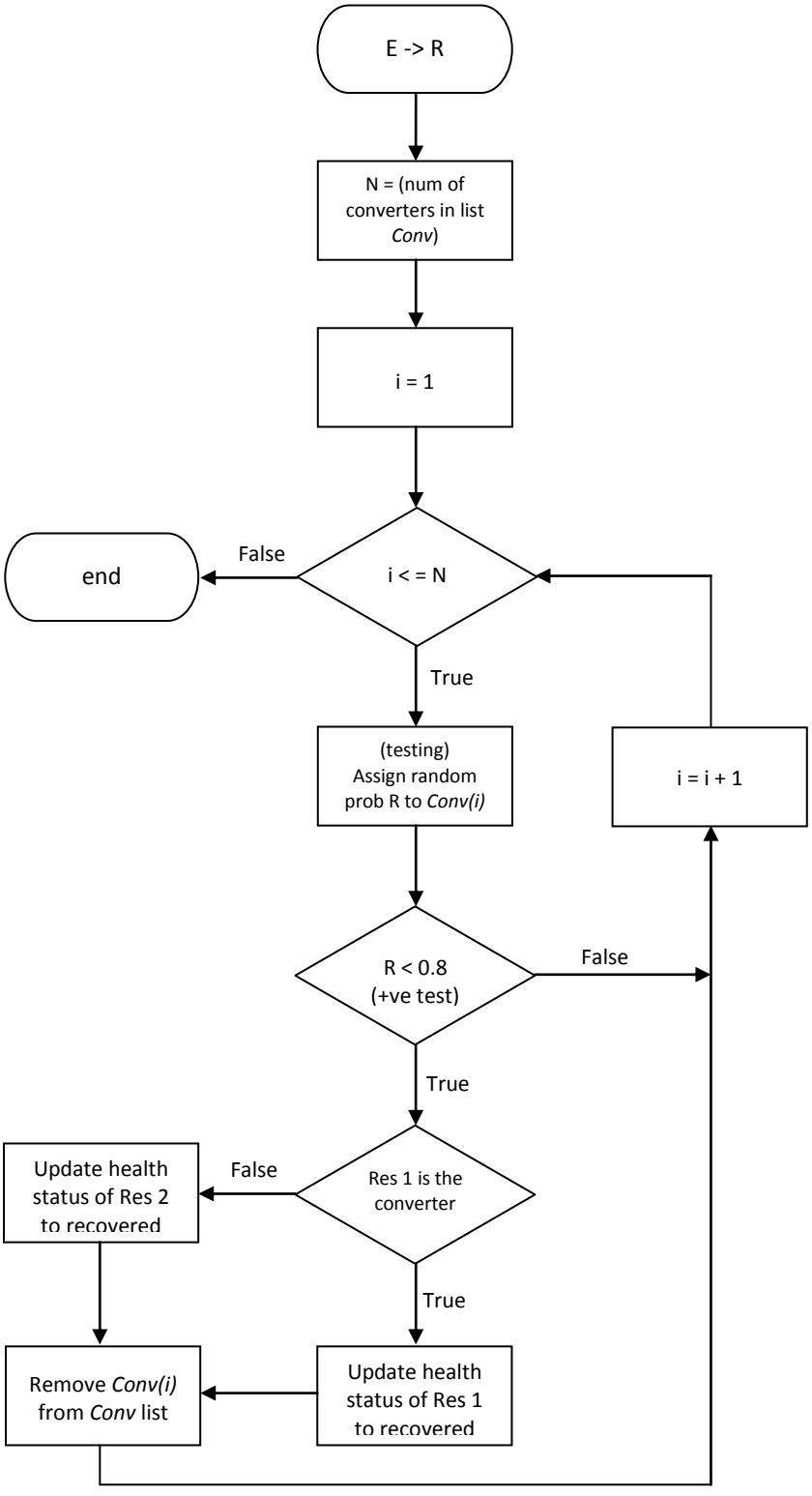


Figure A.4. Subroutine for the E → R process

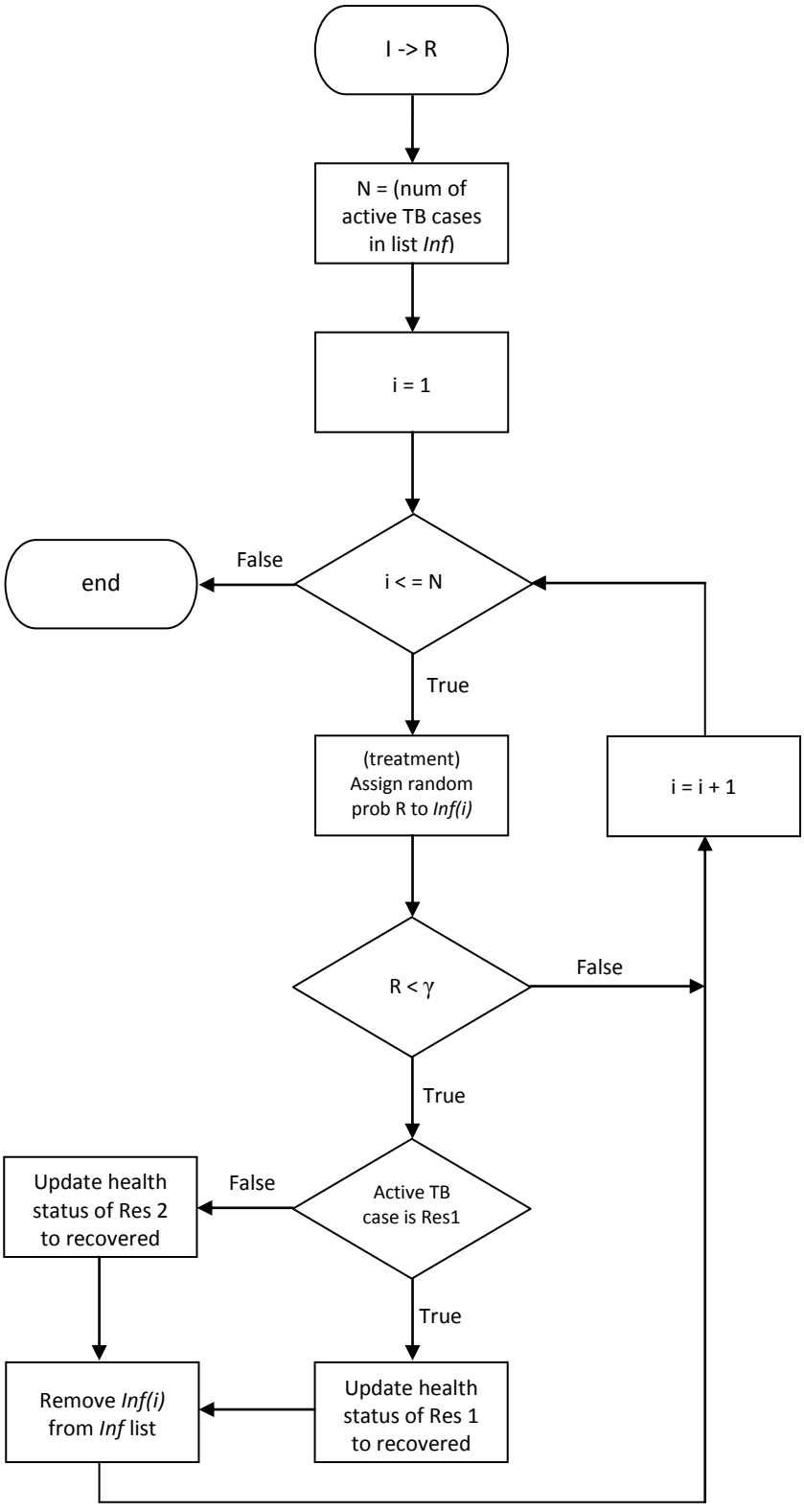


Figure A.5. Subroutine for the I → R process

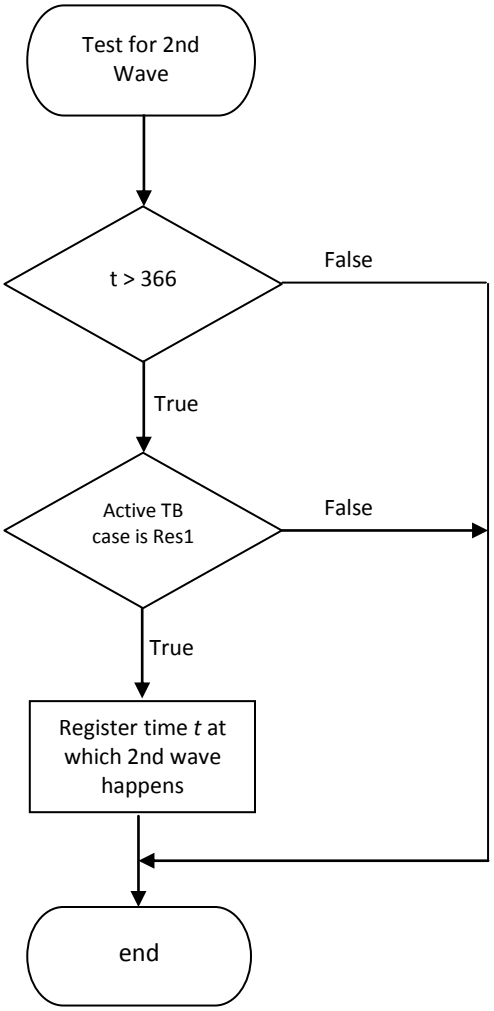


Figure A.6. Subroutine for the Test for 2nd wave process.

Checks if a second wave of infectious cases occurs, in which case it will register the time at which it occurs.