



Faculty of Sciences

Department of Mathematics and Statistical Sciences

hellen.machingauta@studentmail.biust.ac.bw

**MATHEMATICAL MODELING OF INFECTIOUS AND NON
COMMUNICABLE DISEASES: EXPLORING PUBLIC HEALTH
INTERVENTION STRATEGIES**

by

MANDIDAYINGEYI HELLEN MACHINGAUTA

Student ID number: 17100008

BSc (Mathematics) (MSU, ZW), MSc (Applied Mathematical Modeling) (NUST, ZW)

A Dissertation/Thesis Submitted to the Faculty of Sciences in Partial Fulfilment of the
Requirements for the Award of the Degree of Master of Science/Doctor of Philosophy in Pure
and Applied Mathematics of BIUST

Supervisor(s): Professor Edward Lungu

Department of Mathematics
and Statistical Sciences

Faculty of Sciences, BIUST

lungue@biust.ac.bw

Dr Barbara Szomolay

Systems Immunity Research Unit

Cardiff University School of Medicine

Cardiff, United Kingdom

SzomolayB@cardiff.ac.uk

February, 2022

DECLARATION REGARDING THE WORK AND COPYRIGHT

Candidate (please write in caps or type) MANDIBAYINGEYI HELLEN MACHINGAUTA

Student ID: 17100008

Thesis Titled: MATHEMATICAL MODELING OF INFECTIOUS AND NON COMMUNICABLE DISEASES: EXPLORING PUBLIC HEALTH INTERVENTION STRATEGIES

I, the **Candidate**, certify that the Thesis is all my own original work and that I have not obtained a degree in this University or elsewhere on the basis of any of this work.

(If the thesis is based on a group project, then the student must indicate the extent of her / his contribution, with reference to any other theses submitted or published by each collaborator in the project, and a declaration to this effect must be included in the thesis)

This dissertation/thesis is copyright material protected under the Berne Convention, the Copyright and Neighbouring Rights Act, Act. No. 8 of 2000 and other international and national enactments, in that behalf, on intellectual property. It must not be reproduced by any means, in full or in part, except for short extracts in fair dealing; for researcher private study, critical scholarly review or discourse with an acknowledgement, without the written permission of the office of the Postgraduate School, on behalf of both the author and the BIUST.

Signed: HP Pupajena Date: 24/02/2022

Primary Supervisor (please write in caps or type) PROF EDWARD LUNGU

I, the Candidate's **Primary Supervisor**, hereby confirm that I have inspected the above titled thesis and, to the best of my knowledge, it is based on the original work of the candidate.

Signed: [Signature] Date: 25/2/2022

Declaration and Copyright

The work presented in Chapter 4 was previously published in *Mathematics* as *A mathematical model of contact tracing during the 2014-2026 West African Ebola outbreak* by Burton D, Lenhart S, Edholm C, Levy B, Washington M L, Greening B R, White K A J, **Lungu E**, Chimbola O, Kgosimore M, Chirove F, Ronoh M, **Machingauta M H**. This study was conceived by all of the authors. I contributed to the methodology, formal analysis, writing, review, editing and agreed to the published version of the manuscript

Dedication

For my dream keepers (Mutsa Patrick Pfupajena, Cathrine Mlambo) and my son Ethan.

Our deepest fear is not that we are inadequate. Our deepest fear is that we are powerful beyond measure. It is our light, not our darkness that most frightens us. We ask ourselves, Who am I to be brilliant, gorgeous, talented, fabulous? Actually, who are you not to be?

Marianne Williamson

Acknowledgements

Professor Edward Lungu, your guidance, dedication, understanding and time all meant so much to me. When I did not believe in my own capabilities and when the journey was tough both academically and personally you helped me keep the faith that I will get through. Not only did I attain a great mentor but I felt like part of the family. Thank you and may God bless you always.

Dr Barbara Szomolay, you were there to help guide me through, always providing encouragement and time to listen whenever I had questions or doubts. Thank you and may God bless you always.

The Department of Mathematics and Statistical Sciences, for providing a home away from home. Thank you to all the members of the department for providing an environment in which I could thrive and work towards my degree.

The Simons Foundation through The Research and Graduate Studies in Mathematics (RGSMA) project at Botswana International University of Science and Technology (BIUST) for funding my studies.

Special thanks to my family for all the love and support. I especially thank my late grandfather Barnes Pfupajena and my uncle Itai Pfupajena for starting me on this journey and teaching me that only I can limit myself.

Special thanks to Dr F. Chipepa, M. Gwazane, C. Mashiri, G. Chingwe, L. Gonzo, Dr C. Zidana and many more for providing the support and friendship that I needed.

Above all I owe everything to the Almighty God.

Contents

List of Figures	ix
List of Tables	xii
Abstract	xiii
1 Introduction	1
1.1 Introduction	1
1.2 Statement of the problem	6
1.3 Objectives of the study	6
1.4 Organization of the work	7
2 Literature Review	8
2.1 Summary	19
3 Mathematical tools	20
3.1 Autonomous dynamical systems	20
3.2 Hartman-Grobman theorem	23
3.3 Stability and Lyapunov functions	24
3.4 Definition and computation of R_0	25
3.5 Kamgang-Sallet stability theorem	27
3.6 Sensitivity analysis	27
3.7 Optimal Control	28
3.8 Brownian motion	30
3.8.1 Itô Process	30
3.8.2 The general Itô formula	31

4	A Model for Contact Tracing of an Infectious Disease: The West African Ebola Outbreak As An Example	32
4.1	Introduction	32
4.2	Model Description	33
4.3	The Reproductive number, \mathfrak{R}_0	37
4.4	Numerical Simulations	39
4.4.1	Parameter estimation	39
4.4.2	Importance of contact tracing	46
5	COVID-19 changing the face of the world. Can sub-Sahara Africa cope?	48
5.1	Introduction	48
5.2	Do people learn from previous epidemics: Ebola	49
5.3	Quality of the Lockdown	51
5.4	Model Description	52
5.5	Model Analysis	55
5.5.1	Positivity of solutions	55
5.5.2	Invariance	57
5.5.3	Disease Free Equilibrium (DFE)	57
5.5.3.1	Disease Free Equilibrium for $\pi = 0$	57
5.5.3.2	Disease Free Equilibrium for $\pi \neq 0$	58
5.5.3.3	Global stability of the Disease Free Equilibrium	60
5.5.4	Endemic Equilibrium Point (EEP)	62
5.5.4.1	Global stability of the Endemic Equilibrium Point	63
5.6	Numerical Simulations	65
5.6.0.1	Sensitivity Analysis	65
5.6.1	Italy	66
5.6.2	Optimal Control	71
5.6.2.1	Optimal control without incorporating household and community exposure	71
5.6.2.2	Optimal control incorporating household and community exposure.	72
5.6.3	Sub-Sahara Africa: South Africa as an example.	73

5.6.3.1	Scenario with no lockdown measures	76
5.6.3.2	Effect of early lockdown	79
6	The effect of smoking habits on lung cancer incidence.	81
6.1	Introduction	81
6.2	The Stochastic Model	83
6.3	The Deterministic Model	88
6.3.1	Model Analysis	92
6.3.1.1	Positivity and Boundedness of Solutions	92
6.3.2	The Reproduction Number	92
6.3.2.1	Global Stability of the DFE	94
6.3.3	Existence of the Endemic Equilibrium Point (EEP)	95
6.3.3.1	Bifurcation Analysis	97
6.4	Numerical Simulations	100
6.4.1	Sensitivity Analysis	101
6.4.2	Determination of lung cancer cases and deaths	102
6.4.3	Optimal Control	104
7	Discussion	109
	Bibliography	114
	Appendices	134

List of Figures

1.1	Daily new confirmed COVID-19 cases per million people. Source: Johns Hopkins University CSSE COVID-19 Data.	4
4.1	Flow diagram of population interactions.	34
4.2	The value of the objective for this simulation was $J = 0.0423$	39
4.3	Dynamics of class F in the upper left, class E_F in the upper right, their sum on the bottom left, and the integral of those leaving E_F to be hospitalized on the bottom right. These classes correspond to the parameters from Table 4.2 and the data simulations from Figure 4.2.	43
4.4	The graphs above correspond to the parameters from Table 4.2 and the data simulations from Figure 4.2. Note that the scales are all different.	44
4.5	Effect of varying the number of contact tracers available from 0 to 2000, with 1200 as the corresponding number in our model.	46
4.6	Effect of varying contact tracing parameters κ_1 and κ_2 on the total number of deaths by day 504 of the epidemic.	47
5.1	Ebola outbreaks plots for (a) Sudan and (b) the Democratic Republic of Congo.	50
5.2	Democratic Republic of Congo plots for (a) new Ebola cases (b) new Ebola deaths.	50
5.3	Time course for the environment	51
5.4	Model flow diagram	54
5.5	A diagram showing the sensitivity of R_0 to various model parameters.	66

5.6	Population level plots for hospitalized individuals for varying values of γ_1	67
5.7	Population level plots for hospitalized individuals for varying values of γ_1	68
5.8	Population level plots for hospitalized individuals for varying values of γ_2	69
5.9	Population level plots for hospitalized individuals for varying values of γ_2	69
5.10	Population level plots for hospitalized individuals for varying values of α_3	70
5.11	A comparison of state variables with and without controls	74
5.12	Total covid cases, UK, Spain, Italy, South Africa	76
5.13	Population level plots for hospitalized individuals for varying values of γ_1	77
5.14	Population level plots for hospitalized individuals for varying values of α_3	78
5.15	A comparison of state variables with and without controls	79
6.1	Fitting the stochastic model to proportions of (a) moderate smokers, (b) heavy smokers, (c) new cases for limited stage lung cancer, (d) new cases for extensive stage lung cancer and (e) lung cancer deaths	86
6.2	Flowchart of population interactions for the deterministic model.	88
6.3	(a) moderate smokers, (b) heavy smokers and (c) R_0 for varying values of Λ	101
6.4	PRCC showing the linear relationship between R_0 and certain model parameters, when the effect of the other parameters is discounted.	102
6.5	Simulations showing the comparison of actual and simulated cases of (a) moderate smokers, (b) heavy smokers (c) lung cancer sufferers and (d) lung cancer deaths using the deterministic model (6.6)-(6.12).	103

6.6 Longterm model predictions of actual and simulated cases of (a) moderate smokers, (b) heavy smokers (c) lung cancer sufferers and (d) lung cancer deaths. 104

6.7 Optimal control diagrams showing the effect of control measures on (a) moderate smokers, (b) heavy smokers population (c) lung cancer cases, (d) lung cancer deaths and (e) susceptible population. 108

6.8 Profiles for controls representing (a) social and (b) health policy measures. 108

7.1 First attempt match to the data of cumulative cases and cumulative deaths with all parameters constant. The value of J is 0.1963. 134

List of Tables

4.1	The parameters and compartment names in our model with their interpretations and units.	36
4.2	Values for parameters, with five parameters having early and late values. Parameters with * were taken from the data or the literature. Others were estimated.	45
5.1	Numerical values for the parameters of the Italian case	65
5.2	Comparison of deaths at different optimal cases	71
5.3	Numerical values for the parameters of the South African case	75
5.4	Number of people who must be hospitalized.	77
5.5	Comparison of deaths at different optimal cases	80
6.1	Model variables and parameters	90
7.1	Parameters and their ranges used for the sensitivity analysis	142

Abstract

As new viruses and new pandemics emerge we face the question as to whether our global health systems are well prepared to deal with them. Non pharmaceutical measures are a key control measure in the battle against infectious diseases especially in the absence of vaccines or when available vaccine quantities are not sufficient. The 2014-2016 West African outbreak of Ebola Virus Disease (EVD) was the largest and most deadly to date. Contact tracing, following up those who may have been infected through contact with an infected individual to prevent secondary spread, plays a vital role in controlling such outbreaks. Our aim in this work was to mechanistically represent the contact tracing process to illustrate potential areas of improvement in managing contact tracing efforts. We also explored the role contact tracing played in eventually ending the outbreak. We presented a system of ordinary differential equations to model contact tracing in Sierra Leone during the outbreak. We included the novel features of counting the total number of people being traced and tying this directly to the number of tracers doing this work. Our work highlighted the importance of incorporating changing behavior into one's model as needed when indicated by the data and reported trends. Our results showed that a larger contact tracing program would have reduced the death toll of the outbreak. Counting the total number of people being traced and including changes in behavior in our model led to better understanding of disease management.

Viral outbreaks differ in many ways, despite these differences policy responses used to tackle viral epidemics tend to be similar across time and countries. Substantial progress has been made since the 2014-2016 Ebola outbreak with lessons learnt from previous and ongoing outbreaks followed by significant investments into surveillance and preparedness and this has been of help in dealing with the COVID-19 pandemic. We formulated a mathematical model

for the spread of the coronavirus which incorporated adherence to disease prevention. The major results of this study were: first, we determined optimal infection coefficients such that high levels of coronavirus transmission were prevented. Secondly, we found that there existed several optimal pairs of removal rates, from the general population of asymptomatic and symptomatic infectives respectively that could protect hospital bed capacity and flatten the hospital admission curve. Of the many optimal strategies, this study recommended the pair that yielded the least number of coronavirus related deaths. The results for South Africa, which is better placed than the other sub-Saharan African countries, showed that failure to address hygiene and adherence issues will preclude the existence of an optimal strategy and could result in a more severe epidemic than the Italian COVID-19 epidemic. Relaxing lockdown measures to allow individuals to attend to vital needs such as food replenishment increases household and community infection rates and the severity of the overall infection.

Although the tobacco epidemic is one of the biggest health threats, responsible for more than 8 million deaths annually with 15% of these caused by second hand smoke , only a few mathematical models have addressed smoking in the context of lung cancer. In our work we present two models, a stochastic model and a deterministic model both of which are fitted to actual smoking data. The expected solution of the stochastic model predicts a steady state solution in the long run for the moderate and heavy smokers with proportions of these populations remaining to sustain the habit contrary to the trend in the actual data which suggests extinction of these populations. The deterministic model, revealed that the presence of highly quantifiable efficacious control measures can reduce the lung cancer load by 50% although the number of lung cancer deaths would remain the same for sometime. These results confirm the conclusions of the stochastic model and reveal further that these control measures can reduce the lung cancer load and lung cancer deaths by about 50% if there is a reduction of at least 20% in the population of susceptible individuals taking up smoking. Specifically, if the number of new potential (susceptible) smokers ex-

ceeds a quantifiable threshold, Λ^* , then even if $R_0 < 1$ there is persistence of the epidemic.

Keywords: *Contact tracing; Hospital capacity; Lung cancer; Optimal control; Smoking; Bifurcation analysis*

Chapter 1

Introduction

1.1 Introduction

Epidemiological studies in the twentieth century were dominated by studies of infectious diseases but as lifestyles have changed due to improved standards of living which ushered in processed food stuffs, laid back lifestyles etc., it has become apparent that non-communicable diseases (NCDs) have overtaken infectious diseases in terms of number infected and mortality [1, 2]. The dominance of NCDs is being challenged by the SARS-2 viral infections currently known as COVID-19 and its variants which is highly infectious and spreads in a variety of ways including close contact with contaminated surfaces, air, etc [3]. Luckily, humanity has been spared of the worst effects of the current pandemic due to tremendous scientific improvements in pharmacology, virology amongst other things. The rapid development of the vaccines (developed within 1 year) reduced the full effect of the SARS-2 virus in human beings [4].

The burden of NCDs on healthcare systems and generally economies, especially chronic diseases such as diabetes, cardiovascular diseases, cancers, asthma etc., is a growing concern for all nations. What is more worrying is how individuals infected with these diseases will fare if they are co-infected with SARS-2 virus (COVID-19) [5].

The World Health Organization (WHO), report of 2017 on NCDs shows that these diseases kill about 40 million people each year, equivalent to 70% of all deaths globally. The report shows that of these deaths, cardiovascular diseases

account for the most with 17.7 million deaths annually, followed by cancers (8.8 million), respiratory diseases (3.9 million) and diabetes (1.6 million) [6]. It is important to point out that in the 18 months of SARS-2, a respiratory disease, over 4 million individuals have lost their lives [7]. This is a warning to humanity as this respiratory disease (COVID-19) has contributed over 50% per annum to the 2017 mortality total of 3.9 million (for respiratory diseases).

The four diseases, cardiovascular, cancers, asthma and diabetes, account for over 80% of all premature NCDs deaths and constitute a threat to missing the 2030 targets of the Sustainable Development Goals (SDGs) of reducing premature NCDs deaths by one third [8]. Achieving the SDGs is dependent on reducing some of the most prominent risk factors such as tobacco use, physical inactivity, harmful use of alcohol and unhealthy diets.

Tobacco use is linked to severe COVID-19 but not COVID-19 deaths, an assertion which is disputed [9]. In any case, tobacco use is a preventable habit. However, in developing countries the argument against tobacco use appears to be a lost cause as advertising of tobacco products is associated with the high standard of living in the Western countries [10]. The use of alcohol is another habit which is popularized by lucrative advertizing and sponsorship of popular activities such as sport by breweries and the movie industry [11, 12].

Ozturk and Fidanci [13] have given the statistics on tobacco use which show that more than one billion people and a quarter of the adult population globally use cigarettes and other tobacco products. Consumption is on the rise especially in low-middle income countries (LMICs) as shown by the WHO 1980-2016 statistics [14], thereby presenting a possible double burden in terms of NCDs and infectious diseases for these countries that already have compromised health systems. The challenge for most LMIC's are their growing economies which are still largely dependent on developed countries, therefore their tax laws with regards to the tobacco industry are still more favorable to the industry and hence tobacco consumption is still on the rise [15].

Tobacco use causes cancer in multiple organs and is the main cause of lung cancer, responsible for at least 82% of the cases [16]. Landmark reports by the Royal College of Physicians in 1962 and the Advisory Commission to the Surgeon

General of the United States in 1964 provided indisputable proof that cigarette smoking lay at the heart of the worldwide increase in deaths from lung cancer. Environmental tobacco smoke (ETS) known as second hand smoke (SHS) or passive smoking is an established risk factor for lung cancer, it is a mixture of exhaled mainstream smoke and side stream smoke diluted with ambient air. According to the Global Cancer Observatory (GLOBOCAN) estimates for 2020 [17], more than 2 million people received a new diagnosis of lung cancer and 1.8 million died of lung cancer, comprising of 11.4% and 18% of all cancer cases respectively.

The year 2020 showed the world the crucial link between communicable diseases and NCDs, and that the management of communicable diseases can have a direct impact on the control strategies for NCDs. WHO completed a rapid assessment survey in May 2020 and found that 75% of countries reported interruptions to NCD services, amongst the most hard hit were public health campaigns and NCD surveillance efforts [18]. COVID-19 has seen countries go into strict lockdown situations, where day to day activities that were considered a norm were prohibited. Movement between states or countries faced restrictions, the education system encountered a complete shutdown especially in LMICs as they could not cope with the evolution from face-to-face learning to online learning, in hospitals COVID-19 cases took priority over other ailments, giving rise to questions on whether our health systems can stand up to pandemics and how sustainable these systems can be in the face of a major global health threat.

With the introduction of lockdown measures, only the most essential aspects of day to day life remained in place. Movement to and from shopping malls to obtain necessities for the home though at a decreased frequency was allowed, mandatory curfews with set times also meant individuals could still move around within reason and all essential workers were provided with special permits to be able to move to and from their workplaces during the strictest lockdown levels. In view of these national lockdowns there is possibility of overlooking households and communities as reservoirs that can worsen COVID-19 transmission as there is no foolproof way of restricting contact between individuals sharing the same house or that live within the same community. The

COVID-19 virus spreads mainly between people who are in close contact with each other when aerosols or droplets containing the virus are inhaled or come directly into contact with the eyes, nose or mouth [19], indoor environments such as homes, offices, malls, and public transport vehicles pose a particular challenge as these help fuel household and community transmission. As a result there is need to develop mathematical models that include household and community transmission when exploring the strategies that policy makers can use to contain COVID-19 transmission to combat the disease. In this thesis we develop a mathematical model for the spread of COVID-19 which incorporates adherence to disease prevention and the model also highlights how household and community infection rates can alter the severity of the overall infection.

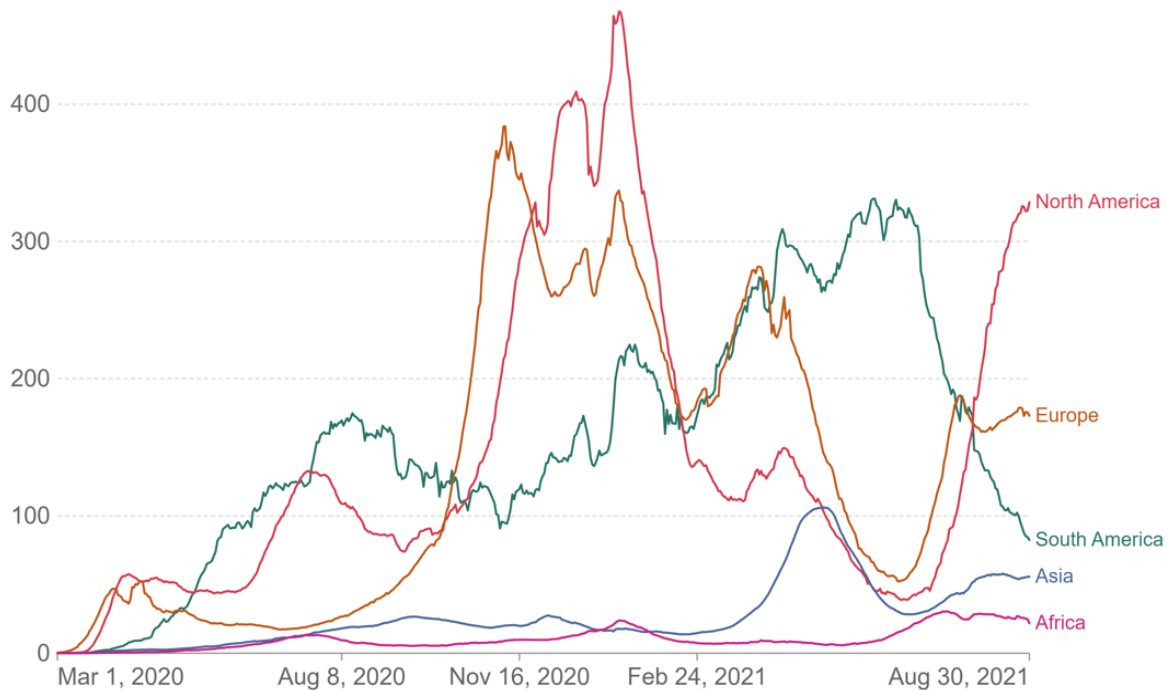


Figure 1.1: Daily new confirmed COVID-19 cases per million people.
Source: Johns Hopkins University CSSE COVID-19 Data.

The future is shaped by past experiences, in the approach to dealing with disease outbreak instead of relying only on new measures or new methods it is important to understand how in the past health systems withstood these sudden outbreaks. Africa carries the greatest health burden in the world in any given year, given its ongoing fight against recurring and new infectious diseases. On

average, 21 countries (45% of the continent) experience at least one epidemic annually [20]. In the last 10 years sub-Sahara Africa has experienced a number of epidemics including cholera, the Ebola virus disease (EVD), Lassa fever, monkeypox, Chikungunya fever and more recently COVID-19 [21]. Africa's history with epidemics seems to have worked to its advantage in the COVID-19 pandemic [22], fewer cases of COVID-19 have been reported in Africa as compared to the rest of the world (Figure 1.1). It has been suggested that this is a result of the lessons learned from the fight against EVD and cholera. For example the 2014-2016 outbreak (predominantly in West Africa) [23] and the 2018 cholera outbreak in Zimbabwe [24]. Many African countries responded swiftly to the WHO recommendation to ramp up preparedness even before the pandemic was declared a public health emergency of international concern, for example Sierra Leone a country devastated by the 2014 EVD outbreak closed its borders before recording a single COVID-19 case [22].

Until 2013, EVD outbreaks consisted of small numbers of cases that were contained by basic public health and containment measures. The largest EVD epidemic occurred in West Africa between 2013 and 2016 and detection of EVD cases in the United Kingdom (UK), Sardinia, Spain and the United States of America (USA) focused global attention on the epidemic. Like COVID-19, the rapid emergence of EVD required rapid response and targeted control measures. Africa with its poor health systems has had to rely largely on non-pharmaceutical health interventions to suppress epidemic curves. Contact tracing was used during the 2014 EVD outbreak with notable success in lowering transmission rates, this has helped inform and improve contact tracing efforts for COVID-19. Public health officials learnt to contact trace from these past outbreaks by learning that trust and communication between governments and members of the public are vital elements for contact tracing. The challenges of accessibility of some areas and manpower that were experienced in West Africa were improved upon as governments embraced modern day technology for contact tracing and the quick dissemination of information on health measures and steps that can be taken to prevent infection. Focusing on the challenges regarding resource allocation during the West African EVD outbreak, this thesis

presents a system of ordinary differential equations to model contact tracing in Sierra Leone. The model illustrates potential areas of improvement in managing contact tracing efforts by including new features like counting total number of people being traced and tying this directly to the number of contact tracers doing the work.

1.2 Statement of the problem

Emerging infectious diseases due to globalization can spread rapidly across national boundaries and communities, challenging the ability of public health systems to prevent and control the spread of the diseases, especially in resource-limited countries and regions. Contact tracing is an effective method to control emerging infectious diseases like EVD and COVID-19. Better understanding of the mechanisms of contact tracing is important for disease management as this plays a key role in controlling communicable diseases by seeking to break the chain of transmission between individuals.

To stop lung cancer being a major public health problem in the future, tobacco control is still needed to bring the rates down. It is imperative therefore to develop models incorporating characteristics that link smoking habits and lung cancer, as valuable tools to predict and understand the efficacy of smoking cessation interventions. To date there is no known level of safe exposure for non smokers to SHS resulting in non smokers acquiring lung cancer, there is need to study the effects secondary smoking has on lung cancer and quantify the risk that SHS confers on the development of lung cancer.

1.3 Objectives of the study

NCDs have received plenty of attention from governments but action to deal with combating them clearly remains inadequate. Programs to reduce NCDs remain chronically underfunded, only 2% of development funding goes toward their prevention [25]. Cigarette smoking prevalence can be a deceiving indicator of public health success since there is under reporting of socially disapproved

behaviors [26]. Much of the interpretation in changing lung cancer incidence (or mortality) trends must take into account patterns of tobacco consumption and related to this governmental efforts to reduce smoking prevalence [27]. For a better understanding of smoking prevalence trends, lung cancer incidence and mortality trends as well as strategies to help in the management of EVD and COVID-19 we conduct a study whose objectives are as follows:

- To apply contact tracing techniques to increase understanding of disease management.
- To determine isolation and quarantine strategies that will ensure that health systems are not overwhelmed by COVID-19 infections.
- To contribute towards policy recommendations on the management of COVID-19.
- To understand the dynamics between tobacco smoking and the reduction of lung cancer.
- To investigate efficacy levels of control measures capable of clearing the lung cancer caseload.

1.4 Organization of the work

This work is organized as follows: Chapter 2 reviews various studies about EVD, COVID-19, tobacco smoking and lung cancer. Chapter 3 will discuss mathematical and numerical tools used in this study. In Chapter 4 we present a mathematical model for EVD which focuses on contact tracing and the role it could possibly play in ending the EVD outbreaks. In Chapter 5 we present a mathematical model for COVID-19 which explores various strategies for controlling the disease. In Chapter 6 we present a mathematical model that explores the effect of smoking habits on lung cancer incidence. Chapter 7 discusses the results, recommendations and possibilities of future work.

Chapter 2

Literature Review

Throughout history, the spread of infectious diseases has been linked to human trade and travel, but in recent times, modern globalization and increasing populations have increased the probability of occurrence and the impact of outbreaks and pandemics. As new viruses and new pandemics emerge we face the question as to whether our global health systems are well prepared to deal with them. In the absence of vaccines or when vaccine quantities required are not sufficient an alternative approach is to stop infections through non pharmaceutical control measures. One such control measure is contact tracing and it has been shown to be a key control measure in the battle against infectious diseases [28].

Contact tracing involves identifying and monitoring individuals who have been in close contact with an infected person. In its simplest form contact tracing and isolation can be traced as far back as the medieval period in Europe, whilst the Black death (plague) raged through towns and cities across the continent, sufferers were not only impelled to quarantine in their homes, but their houses were marked with crosses on the door to alert others to avoid attempting contact with the infected individuals [29]. One of the most successful examples of contact tracing is probably the eradication of smallpox, due to a process known as “ring vaccination” which consisted of monitoring smallpox cases in a community and intervening for those who were at risk of infection. Small pox was officially declared eradicated in 1979 [30]. Early forms of contact tracing were applied to other infectious diseases for example tuberculosis, cholera, sex-

ually transmitted infections like HIV (human immunodeficiency virus) as the late nineteenth century progressed [31].

Between 2014 and 2016, when the Ebola virus disease (EVD) was ravaging communities in West Africa, healthcare workers extinguished the epidemic by finding and quarantining contacts of anyone who caught the disease. To date the 2014-2016 outbreak in West Africa has been the largest outbreak, 28 615 cases of EVD and 11 310 deaths were reported in Guinea, Liberia and Sierra Leone [32]. There were an additional 36 cases and 15 deaths that occurred when the outbreak spread outside of these three countries [32].

Transmission of EVD to humans can occur via two routes. Primary infection, via infected animals and non-human primates, and secondary infection, via human-to-human direct contact. In the latter case, transmission occurs via blood or body fluids of an infected/dead person due to EVD or due to contaminated objects with body fluids containing the virus [33]. The incubation period for EVD is between 2-21 days. EVD is not transmitted until the EVD-infected individual develops symptoms such as fever, weakness, fatigue, gastrointestinal symptoms and others [34].

Mathematical models are helpful in the process of linking biological processes of transmission and the dynamics of infection at the population level, allowing for optimal use of limited resources and efficient targeting of control measures in public health systems. Since the 1980's, modellers have developed a consistent theory for contact tracing, with the aim to find effective and efficient implementations and to assess the effects of contact tracing on the spread of infectious diseases [35]. A general modeling framework for measuring the population-level effect of contact tracing with epidemic data is needed for Ebola and other emerging pathogens. Chowell and Nishiura [36] illustrated the insights for disease management that can come from modeling, connected with Ebola epidemiological data and discussed the need for understanding the effectiveness of contact tracing. They show that mathematical modeling can be a valuable tool for assessing the risk of EVD epidemic and for predicting the impact of health measures on disease spread.

Rivers et al [37] built an SEIHR (Susceptible-Exposed-Infectious-

Hospitalized-Funeral-Recovered/removed) model of the epidemic in Sierra Leone and Liberia while it was ongoing and before it had reached a peak. A deterministic version of the model was fit and validated to the current outbreak data, whilst forecasting into the future was done using a stochastic version of the model. By increasing the proportion of diagnosed and hospitalized infected cases and by decreasing the time for hospitalization of infected individuals, their work showed that improved contact tracing can have a substantial impact on the number of Ebola cases, but is not sufficient to control the EVD epidemic. However since their model was calibrated using data before the peak of the epidemic in Liberia it is possible that the results could be biased as they could not capture contact tracing when resources were overwhelmed and what may be deemed a success during a period with lower reported number of cases may differ when the numbers are increased. In resource strapped countries it is important to take cognizance of such limitations and although they increased contact tracing they did not take into consideration the number of contact tracers needed in such a scenario.

Browne et al [38] in their SEIR model incorporated contact tracing by building separate compartments for exposed and infectious individuals being traced. They considered key features of contact tracing and characterized the impact of contact tracing on the effective reproduction number R_e of Ebola. Browne et al [38] determined how incubation period, infectious period and case reporting, along with varying monitoring protocols, affect the efficacy of contact tracing. They found that increasing the fraction of cases reported and increasing the fraction of reported contacts that were traced could bring R_0 below 1. They also provided weekly point estimates for the effective reproduction number for Guinea and Sierra Leone. However their model did not include spread from contact with deceased individuals yet they are a significant source of infection mainly due to traditional funeral practices where the washing or handling of the dead body is done by relatives and before the body is laid to rest family members or friends can kiss or embrace the body as a final farewell.

Webb et al [39] objective was to develop a mathematical model of the 2014 EVD epidemic in West Africa. Based on numerical simulations, they obtained

a reasonable visual fit and identified rough rules of thumbs for containing the epidemic. Webb et al [39] built an SEIR model form as well as a corresponding stochastic model implementation, which included a compartment for contaminated deceased (improperly handled corpses of infected) individuals, but did not include a hospitalized compartment. Projections by [39] indicated that the containment of the epidemic required a high level of both general identification and isolation process and contact process for removing infectious individuals from the susceptible population. The model included the main features of contact tracing, namely, the number of contacts per identified infectious case, the likelihood that a traced contact is infectious and the efficiency of the contact tracing process, but did not include the workload of tracing persons who do not become infected. More questions still need to be answered regarding the efficacy of tracing during times of heightened transmission as contact tracers could have been overwhelmed with high numbers of contacts that needed tracing.

Viral outbreaks differ in many ways, despite these differences policy responses used to tackle viral epidemics tend to be similar across time and countries. Substantial progress has been made since the 2014-2016 EVD outbreak with lessons learnt from previous and ongoing outbreaks followed by significant investments into surveillance and preparedness [40, 41]. Recent history of a similar event plays a role in altering the perceived threat from the disease as the loss and pain suffered during the event would still be fresh on people's minds. The World Health Organization was more timely in declaring COVID-19 as a Public Health Emergency of International Concern (PHEIC) as compared to the EVD outbreak. The first case of EVD was reported on 13 December 2013 and it was declared a PHEIC on 8 August 2014 (238 days) whereas the first case of COVID-19 was reported on 31 December 2019 and it was declared a PHEIC on 30 January 2020 (30 days) implying that having seen the devastating consequences of delaying considering the EVD outbreak as a PHEIC the organization could not afford to make the same mistake.

The EVD outbreak taught us that educating the public through the use of visual aids and messages that highlight the signs and symptoms of a disease as well as identifying risky behaviors has to be done in a timely manner, to allow

for early case identification, isolation and reduced transmission in the community [42]. The COVID-19 pandemic saw high media coverage as the messages on COVID-19 were put out on television, radio, print media as well as social media platforms to enable people to access the necessary information needed for self-preservation.

Community engagement must form the cornerstone of any COVID-19 response as it was key in ending the previous EVD outbreak in Africa [43, 44]. Swanson et al. [45] in their review of the usefulness of contact tracing during the 2014-2016 EVD outbreak recommend the implementation of intensified contact tracing efforts early in the outbreak and focusing on strengthening surveillance efforts on all fronts and electronic data systems, decentralizing management of multidisciplinary teams for better harmonization and oversight, developing and maintaining clear and well-detailed protocols and adapting community-led solutions to foster cooperation, trust and ownership. COVID-19 has seen contact tracing being a community led event, it was not restricted to qualified health personnel only but in various countries daily lists were compiled of people visiting various shopping malls, taxi drivers kept a record of passengers they would have ferried, alerts on phones in some countries would notify a person's contacts if they tested positive, to mention a few. This gave everyone a sense of purpose and created better contact tracing efforts that had a far greater outreach.

A united front goes a long way in helping achieving desired goals or attaining desired results in most situations. Outbreak response requires concerted effort and should not be left for any one government or organization. During the 2014-2016 EVD outbreak the international community committed USD100 million to fight the epidemic. This enabled the deployment of more medical personnel to assist in combating the outbreak [46]. The same efforts have been noted for COVID-19 with governments collaborating and coordinating events such as lockdown measures, the development of vaccines as well as their donation to other countries with less access to these vaccines. Due to these efforts the resilience of the human race amidst crisis has remained intact.

Since the first case of COVID-19, various mathematicians around the world have developed different mathematical models to understand the transmission

dynamics of the virus and have investigated the effects and impact of different intervention strategies on the disease.

Deressa et al [47] developed an SEIAHR mathematical model to study the transmission dynamics of COVID-19 in Ethiopia. Calibration of the model was done using the actual data of confirmed COVID-19 cases in Ethiopia. From their results intervention strategies greatly reduced the number of individuals in the compartments of exposed, symptomatic, asymptomatic and hospitalized cases for COVID-19. The much needed decreases in the compartments was achieved through the proper usage of personal protective measures but this had to be maintained for a long period of time. However, they only considered infection from symptomatic and asymptomatic infected individuals but neglected infection from the environment which is applicable in the African context as good sanitation and hygiene measures are difficult to maintain in countries with poor economies and poor health systems.

Garba et al [48] used a compartmental model to analyze the transmission dynamics of COVID-19 in South Africa. They used an SEIR model modified by incorporating classes that accounted for asymptomatic transmission and isolation/hospitalization. The model incorporated the role of environmental contamination by COVID-infected individuals. Garba et al [48] assessed the impact of various control and mitigation strategies and established that monitoring, testing and following up on the contacts and confirmed COVID-19 cases was important in curtailing the pandemic in South Africa. Garba et al [48] found the reproduction number for South Africa R_c , to be 2.9562, suggesting an outbreak which was out of control in South Africa. Their study suggested that care should be taken before reducing COVID-19 alerts so as to avoid loss of the gains made in managing the pandemic up to that moment.

Kassa et al [49] formulated and analyzed a mathematical model for the transmission of COVID-19. The model incorporated a behavior change function to account for the proportion of individuals who decided to use any of the self-protective measures and adhered to them, as well as the concentration of the pathogen in the environment. The study by [49] also considered a proportion of individuals with a history/knowledge of similar infections from the past and

practiced necessary protective measures right from the onset of the epidemic. Failure to develop permanent immunity against the disease for recovered individuals resulted in the model by [49] exhibiting a backward bifurcation. Kassa et al [49] suggested that if countries were unable to carry out extensive vaccination they needed to be able to identify and isolate at least 30% of the asymptomatic infectious group of individuals while treating in isolation at least 50% of symptomatic patients to control the disease. Although [49] consider transmission from the environment they did not consider household and community transmission as health personnel and key workers are possible infection agents in the COVID-19 context.

Mukandavire et al [50] adapted an SEIR compartmental model to fit to the observed cases of COVID-19 in South Africa and explored how the use of vaccines for different hypothetical vaccine efficacy scenarios could help in containing the disease. Early model estimates showed South Africa had a basic reproductive number of 2.95, confirming COVID-19 was there to stay in South Africa and the outbreak had the potential to crossover to other geographic locations. A vaccine with more than 70% efficacy as well as a high vaccine coverage rate of 94.44% would have enabled the containment of the COVID-19 pandemic as findings by [50] suggest. [50] recommended that maintaining social distancing measures at current levels will remain essential in controlling the infection in the absence of widespread vaccines or therapeutic options. However, since the study used hypothetical scenarios, the conclusions of this study could not be validated and further studies are required using real time data after the COVID-19 vaccination roll-out program has been done. Furthermore, focus still needed to be maintained to consider the efficacy of non-pharmaceutical control measures for COVID-19.

Ngonghala et al [51] developed a Kermack-McKendric SEIR type compartmental deterministic model incorporating features pertinent to COVID-19 transmission dynamics and control such as the quarantine of suspected cases and the isolation/hospitalization of confirmed COVID-19 cases. The model by [51] showed that the worst-case scenario projections for cumulative community mortality (based on the baseline levels of anti-COVID non-pharmaceutical in-

terventions considered in the study) decreased dramatically if the strict social distancing measures implemented are maintained until the study ended in May or June 2020. They suggested that using face masks in public is very useful in minimizing community transmission and burden of COVID-19 provided they are used community-wide.

COVID-19 is primarily transmitted through the respiratory tract (saliva). It is believed that smokers may be at increased risk of contracting the virus due to reduced lung function, impaired immune systems, cross-infection and susceptible hygiene habits [52]. For example, smokers tend to touch their lips with their fingers regularly when smoking, a habit that increases the probability of transmission of COVID-19 from hand to mouth [53]. On the other hand, smokers may have lung disease (or reduced lung capacity) which would greatly increase the risk of serious illnesses like COVID-19 [53]. The relationship between smoking and COVID-19 infection is contradictory. In particular, smoking prevalence among patients hospitalized with COVID-19 has been reported to be lower than the smoking prevalence in the general population [9]. Williamson et al [54] and Miyara et al [55] reported a slight protective effect against death from COVID-19 in current smokers. Although there may be contradictions on smoking and COVID-19, what remains unchallenged is the fact that smoking is a major risk factor for common chronic diseases, especially those closely related to the occurrence and development of respiratory disease [56]. The question is how COVID-19 and its management affects NCDs and their risk factors in the long run.

Since the tobacco epidemic is a major health challenge worldwide, numerous mathematical studies have been developed to explain the transmission dynamics and possible control measures to curb the epidemic. One of the earliest mathematical models on smoking dynamics was presented by Castillo-Garsow et al. [57]. In this model the population was sub-divided into 3 classes, potential smokers, smokers and quitters. This model was later extended and modified by other researchers [58, 59, 60, 61] by adding an extra class differentiating those who quit smoking as either temporary or permanent quitters. Other researchers [62, 63, 64] have classified smokers as occasional smokers or

heavy (chain) smokers and have explored the impact this classification has on the smoking epidemic.

Medical intervention such as behavioral guidance can be critical for recovery from tobacco addiction. Several studies [65, 66, 67, 68, 69] have explored control measures to understand smoking cessation or other impediments to stop smoking as a habit.

Cancer is a generic term for a large group of diseases that can affect any part of the body, it is the rapid creation of abnormal cells that grow beyond their usual boundaries to affect other parts of the body. Smoking is known to cause cancer and to block the body from fighting it in two different ways [70], toxins in cigarette smoke:

(i) can weaken the body's immune system resulting in the failure by the body to kill cancer cells and as a result the cells will grow without end.

(ii) can alter a cell's DNA, which affects a cell's normal growth and function.

When DNA is altered a cell can grow uncontrollably and create a cancer tumor.

Tobacco supplies nicotine, an addictive substance in cigarette smoke and smokeless tobacco products, which leads to long term dependence on tobacco use and continued exposure to these toxicants. More than 60 known carcinogens have been detected in cigarette smoke, notably smoking confers a higher risk for lung cancer, an average of between 5 and 10 fold [71].

Darby and Pike [72] proposed a multistage model to provide an accurate summary of the patterns of lung cancer risk among active smokers, ex-smokers and non-smokers. They looked at the stages of the carcinogenic process as affected by cigarettes smoked per day. They used data on numbers of lung cancers diagnosed and the distribution of man-years from a 20-year prospective study of male British doctors for people who had a record of unchanging smoking habits. Attention was restricted to smokers of cigarettes only. The results showed that the relation between lung cancer incidence and number of cigarettes smoked per day was greater than linear. The multistage model provided an accurate coherent summary of the patterns of lung cancer risk among active smokers. Epi-

demiological studies of active smokers have shown that the duration of smoking has much greater effect on lung cancer risk than the number of cigarettes smoked, suggesting that passive smoking is much more harmful than is predicted from levels of exposure alone, as it is often of very long duration usually beginning in early childhood [72].

Acevedo-Estefania et al. [73] investigated the long term impact of educating susceptible smokers and the effect it had on the epidemic (getting lung cancer). The average number of individuals who became smokers and the reduction of this average by an education program are determined. Their model focused on the impact peer pressure has on non smokers and the progression to lung cancer via first and secondhand smoke (SHS). They concluded that the best way to lower the number of smokers and individuals who developed lung cancer was by increasing awareness among individuals of the effects of smoking. SHS is now widely accepted as the third leading preventable health hazard after active smoking and alcohol [74], and although Acevedo-Estefania et al. [73] looked at the progression to lung cancer via SHS they considered its impact as negligible yet failed to suggest a safe level of exposure for SHS.

The study by Andest [75] focused on the relationship between cigarette smoking and the presence of nicotine in the lung. The analysis established that cigarette smoking is a great contributory factor to the amount of nicotine accumulation in the human lung, and concluded that reduction or stoppage of cigarette smoking can reduce the risk of lung cancer.

Sebastian and Victor [76] formulated and analyzed a discrete mathematical model for lung cancer involving smokers and non smokers. They considered an exponential rate for production of cancer cells to illustrate the most striking features of lung cancer cells growth and the genetic changes present in them. They concluded that having more smokers in the population resulted in more lung cancer cases. In their numerical simulations [76] looked only at how increasing the number of non smokers in the population would impact the number of active smokers and those who suffered from lung cancer. It is important though to consider also how increasing the number of active smokers can affect the dynamics explained by their model.

Trisilowati [77] discussed the influence of smoking behavior on the growth of the lung cancer population. The model [77] developed had four populations, susceptibles, active smokers, passive smokers and lung cancer patients. Education on smoking and its hazards was used as a control which was applied to the susceptible population. The purpose of optimal control was to minimize the population of smokers and lung cancer patients as well as the cost of educating individuals. The optimal control measure was effective in controlling the growth of the populations of passive smokers, active smokers and lung cancer patients. However the model by [77] focused on preventing the susceptible population from taking up smoking as a habit but did not address the need to also encourage active smokers to quit smoking through the use of smoking cessation interventions as a control in the model.

An SEIRS model with five classes namely susceptibles, exposed (smokers, victims of smoking), individuals infected with lung cancer and individuals recovered from lung cancer was developed by Ahmed and Biswas [78] to explore the dynamics of lung cancer. Their aim was to analyze the model of lung cancer to show the effect of smoking on lung cancer, to determine the basic reproduction number and to investigate the existence and stability of the disease free and endemic equilibrium points of the model. From numerical simulations it was observed that increasing the number of active smokers resulted in increasing the number of individuals in the victim population group as well as an increase in the number of individuals with lung cancer. They concluded that this can be minimized by controlling the number of active smokers, ensuring that the victim group is safe from smokers as well as improving the recovery rate for lung cancer. [78] state that improving recovery rate for lung cancer can only be possible when the disease is identified at a lower stage because as the cancer reaches a higher stage, it becomes more complicated and it is hard to recover and survive. However, they do not separate the compartment of individuals infected with lung cancer into stages yet they acknowledge that late detection of lung cancer is complicated and affects recovery.

2.1 Summary

The effectiveness and optimal levels of investment for contact tracing, particularly for emerging diseases and for acute epidemics are subjects of ongoing research and debate [45]. Contact tracing requires well organized databases, good surveillance management systems as well as community led tracing efforts tailored for the prevalent conditions [44, 45]. To contain outbreaks successfully there is need to efficiently report and isolate infectious cases as well as efficient monitoring of traced reported contacts [38, 39]. However as we experience more viral outbreaks there is room for improvement in the contact tracing process, thus inspiring further research in this area.

The emergence of COVID-19 required a rapid response and various mathematical models were developed to understand the transmission dynamics as well as mitigation strategies necessary to control the outbreak. It is important to maintain high standards of good sanitary and hygiene measure to present an environment that prohibits virus spread [47, 49, 51]. As the vaccines for COVID-19 were being developed hypothetical vaccine efficacy studies were already underway and showed how highly efficacious vaccines could control COVID-19 [50]. As an emerging infectious disease that affected the whole world, the need to understand strategies needed to mitigate the disease motivated this study.

The lung cancer epidemic is largely driven by smoking and it is important to both stop smoking initiation as well as encourage smoking cessation efforts [77, 78]. Though SHS exposure has received less attention when it comes to mathematical modeling of smoking and lung cancer it is important to consider it as there is no known safe level of exposure [72]. Smoking has such devastating consequences which are being felt world-wide, more so in LMIC's and yet there are few mathematical models on smoking and lung cancer, thus motivating this study to present insight into this area [73, 75, 76, 77, 78].

Chapter 3

Mathematical tools

In this chapter basic mathematical concepts and results that are used in this thesis are introduced.

3.1 Autonomous dynamical systems

A system of non-linear ordinary differential equations (ODEs) is said to be a non-linear autonomous dynamical system if it does not explicitly involve time. Consider a non-linear autonomous dynamical system:

$$\dot{x} = f(x). \quad (3.1)$$

Definition 3.1. *The vector or point $x = x^* \in \mathbb{R}^n$ will be an equilibrium point (critical point) of (3.1) if $f(x^*) = 0$, i.e., no change will appear in the system.*

Definition 3.2. *The corresponding Jacobian matrix of $f = (f_1, f_2, \dots, f_n)$ evaluated at x^* is given by:*

$$J|_{(x^*)} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \quad (3.2)$$

where f_i denotes the i^{th} component of f . The Jacobian matrix gives the best linear approximation of f near the point x^* . Therefore, it can be used to ensure stability near the equilibrium point.

The major question is whether the trajectories $x(t)$ for system (3.1) with ini-

tial condition x_0 will settle at x^* as t tends to infinity or will move away from the equilibrium point x^* as t tends to infinity.

Definition 3.3. *An equilibrium point x^* of (3.1) will be stable, if for each $\varepsilon > 0$ there exists a positive δ_ε such that for any arbitrary solution $x(t)$ the condition*

$$\|x(t_0) - x^*\| < \delta \Rightarrow \|x(t) - x^*\| < \varepsilon, \quad \text{for all } t > t_0, t_0 \in \mathfrak{R} \quad (3.3)$$

holds.

In Definition 3.3 $\|\cdot\|$ is the Euclidean vector norm. Furthermore, the equilibrium point x^* will be locally asymptotically stable (LAS) if it is stable, and if each solution around the equilibrium point x^* also converges to that equilibrium point. Mathematically, it can be expressed as

$$\|x(t_0) - x^*\| < \delta \text{ then } \lim_{t \rightarrow \infty} \|x(t) - x^*\| = 0. \quad (3.4)$$

The point x^* will be unstable if condition (3.3) is not satisfied. An equilibrium point x^* will be globally asymptotically stable (GAS) if it is stable and the conclusion of (3.4) holds for all solutions of the system (3.1).

The linearized system for (3.2) is given by

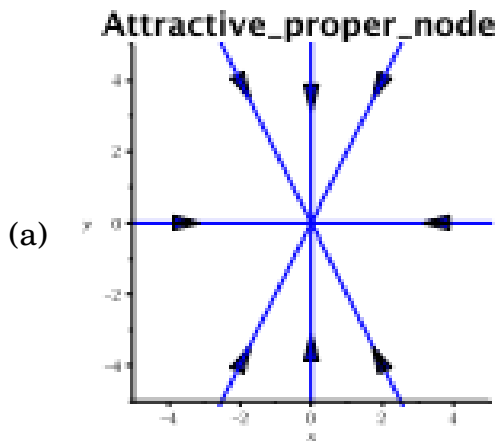
$$\dot{u} = J|_{(x^*)} u, \quad u = x - x^*.$$

Then

- (a) x^* is asymptotically stable if $\mathcal{R}_e(\lambda_i(J|_{(x^*)})) < 0$ for $i = 1, \dots, n$.
- (b) x^* is locally unstable if $\mathcal{R}_e(\lambda_i(J|_{(x^*)})) > 0$ for at least one i .

$\mathcal{R}_e(\lambda_i(J|_{(x^*)}))$ describes the real part of the i^{th} eigenvalue of the linearized system $J|_{(x^*)}$.

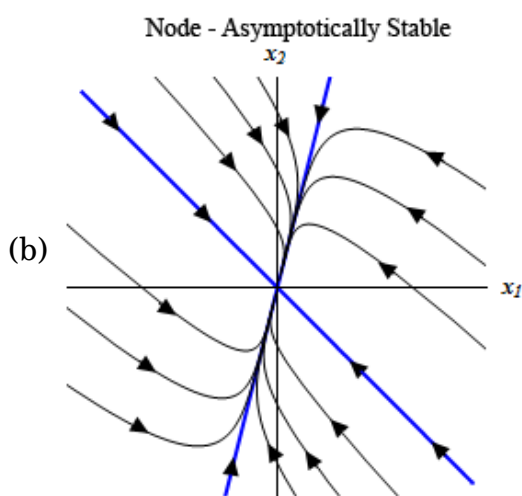
For linear systems like a direction field, a phase portrait can be a tool to predict the behaviors of a system's solutions. We have the following phase portrait classifications in two dimensions [85]:



Repeated real eigenvalue, two eigenvectors.

UNSTABLE if the eigenvalue is positive.

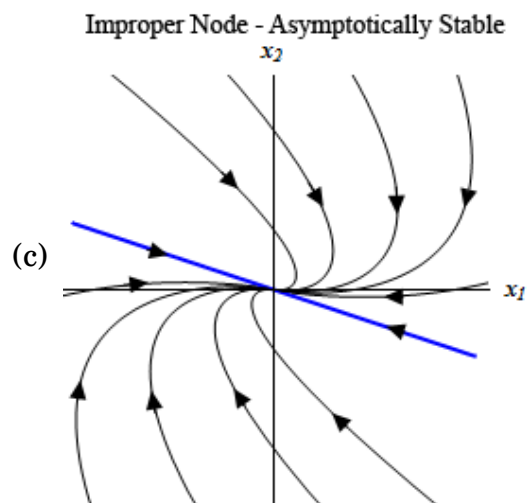
ASYMPTOTICALLY STABLE if the eigenvalue is negative.



Two distinct real eigenvalues of the same sign.

UNSTABLE if both eigenvalues are positive.

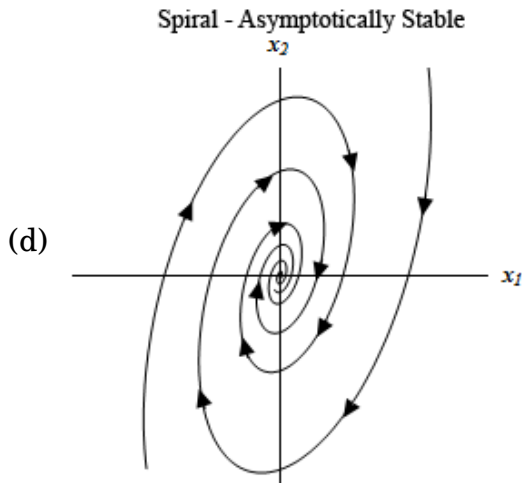
ASYMPTOTICALLY STABLE if both eigenvalues are negative.



Repeated real eigenvalue, one eigenvector only.

UNSTABLE if the eigenvalue is positive.

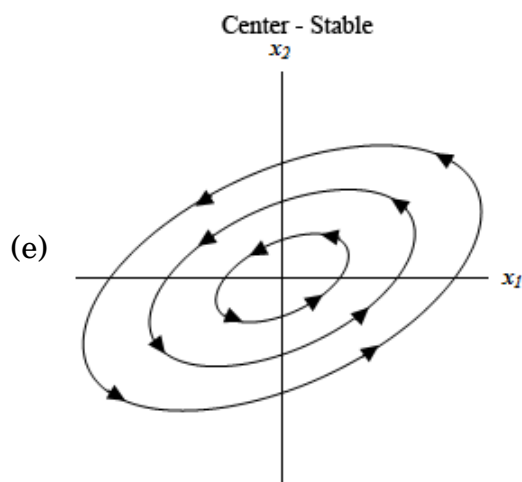
ASYMPTOTICALLY STABLE if the eigenvalue is negative.



Complex eigenvalues, with non-zero real part.

UNSTABLE if real part is positive.

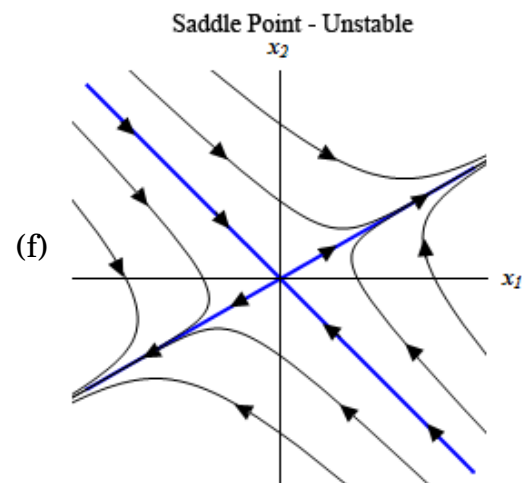
ASYMPTOTICALLY STABLE if real part is negative.



Complex eigenvalues, purely imaginary (real part = 0).

STABLE, or “NEUTRALLY STABLE”.

Not asymptotically stable.



Two distinct real eigenvalues with opposite signs.

Always UNSTABLE.

3.2 Hartman-Grobman theorem

Theorem 3.4. Let $J|_{(x^*)}$ denote the linearization of f at point x^* . Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a smooth map with a hyperbolic fixed point x^* (the eigenvalues of $J|_{(x^*)}$ evaluated at x^* all have non-zero real parts). Then there exists a neighborhood

U of x^* and a homeomorphism

$$h : U \rightarrow \mathbb{R}^n$$

such that

$$f_U = h^{-1} \circ A \circ h, \quad \text{where } A = J|_{x^*}$$

that is, in the neighborhood U of x^* , f is topologically conjugate to its linearization [86].

Theorem 3.4 justifies drawing conclusions about a non-linear system from studying the linearized system. The theorem roughly says that the solutions of an $n \times n$ autonomous system of ODEs in a neighborhood of a steady state behaves qualitatively just like the solutions of the linearized system near that point.

3.3 Stability and Lyapunov functions

Consider system (3.1) and let $V : D \rightarrow \mathbb{R}$ be a continuously differentiable function defined on the domain $D \subset \mathbb{R}^n$ containing a fixed point x^* . The derivative of the function $V(x_0) = V(x(t))$, where $x(t)$ is a solution of (3.1) with initial value x_0 along the trajectories of (3.1) is defined as:

$$\begin{aligned} \dot{V}(x(t)) &= \frac{dV(x(t))}{dt}, \\ &= \left[\frac{\partial V(x(t))}{\partial x_1}, \frac{\partial V(x(t))}{\partial x_2}, \dots, \frac{\partial V(x(t))}{\partial x_n} \right]^T \dot{x}, \\ &= \nabla V(x(t)) f(x), \end{aligned}$$

where $\nabla V(x)$ is the Jacobian of V with respect to x [82]. The necessary condition of the Lyapunov stability theorem is that all trajectories of the system decrease along the graph of $V(x(t))$ toward x^* , i.e $\dot{V}(x(t)) < 0, \forall x(t)$.

Theorem 3.5. Consider system (3.1). Let D be an open subset of \mathbb{R}^n containing x^* , where $f(x^*) = 0$. Furthermore, suppose that $V : \rightarrow \mathbb{R}$ is a real valued positive-definite function. Then

(a) if $\dot{V}(x(t)) \leq 0$ for all $x(t) \in D$, then x^* is stable.

(b) if $\dot{V}(x(t)) < 0$ for all $x(t) \in D - \{x^*\}$, then x^* is asymptotically stable.

Proof. For the proof of Theorem 3.5, see [83]. □

Definition 3.6. A continuously differentiable real-valued function $V : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ satisfying the conditions in Theorem 3.5 is called a Lyapunov function.

3.4 Definition and computation of R_0

The basic reproduction number is used to measure the ability of a disease to reproduce, and is denoted by R_0 .

Definition 3.7. The basic reproduction number, R_0 , of an infection is the average number of secondary cases caused by an infected individual in a wholly susceptible population [79].

This important epidemiological parameter can be determined by assessing the stability of equilibrium points using the Routh-Hurwitz criteria to analyze the characteristic equation, a method sometimes referred to as the Jacobian method (see [79] for examples). The Jacobian method gives a biologically reasonable R_0 , but for more complex compartmental models, especially those with more infected compartments, the method is more difficult to apply as it works well for models in which the necessary and sufficient conditions for stability of the Jacobian can be reduced to a single condition. An alternative approach, the next generation method as discussed by Van den Driessche and Watmough [80] is a general method for deriving R_0 in cases where one or more classes of infectives are involved.

Suppose we have a heterogeneous population whose individuals can be grouped into n compartments. Let $x = (x_1, \dots, x_n)^t$ with each $x_i \geq 0$, be the number of individuals in each compartment. We sort the compartments so that the first m compartments correspond to infected individuals. We define X_s to be the set of all disease free states, that is, $X_s = \{x \geq 0 \mid x_i = 0, i = 1, \dots, m\}$. In order to compute R_0 , it is important to distinguish new infections from all other

changes in the population. Let $\mathcal{F}_i(x)$ be the rate of appearance of new infections in compartment i , $\mathcal{V}_i^+(x)$ be the rate of transfer of individuals into compartment i by all other means, and $\mathcal{V}_i^-(x)$ be the rate of transfer of individuals out of compartment i . $\mathcal{F}_i(x)$, $\mathcal{V}_i^+(x)$ and $\mathcal{V}_i^-(x)$ are differentiable continuous functions. The disease transmission model consists of non negative initial conditions together with the following system of equations:

$$\begin{aligned}\dot{x}_i &= f_i(x) \\ &= \mathcal{F}_i(x) - \mathcal{V}_i(x), \quad i = 1, \dots, n,\end{aligned}$$

where $\mathcal{V}_i = \mathcal{V}_i^- - \mathcal{V}_i^+$ and the functions satisfy the following:

(A1) : If $x \geq 0$, then $\mathcal{F}_i(x)$, $\mathcal{V}_i^+(x)$, $\mathcal{V}_i^-(x) \geq 0$ for $i = 1, \dots, n$.

(A2) : If $x = 0$ then $\mathcal{V}_i^- = 0$. In particular, if $x \in X_s$ then $\mathcal{V}_i^- = 0$ for $i = 1, \dots, m$.

(A3) : If $i > m$, $\mathcal{F}_i = 0$.

(A4) : If $x \in X_s$, then $\mathcal{F}_i(x) = 0$, and $\mathcal{V}_i^+ = 0$ for $i = 1, \dots, m$.

(A5) : If $\mathcal{F}(x) = 0$, then the eigenvalues of $Df(x_0)$ have negative real parts and x_0 is the disease free equilibrium.

The conditions (A1) – (A5) allow us to partition the matrix $D\mathcal{F}(x)$ as shown:

Remark 3.8. $D\mathcal{F}(x_0)$ is the derivative $\left[\frac{\partial f_i}{\partial x_j}\right]$ evaluated at the disease free equilibrium. The conditions (A1) – (A5) allow us to partition the matrix $D\mathcal{F}(x_0)$ and $D\mathcal{V}(x_0)$ as :

$$D\mathcal{F}_i(x_0) = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix}, \text{ and } D\mathcal{V}_i(x_0) = \begin{bmatrix} V & 0 \\ J_3 & J_4 \end{bmatrix}.$$

where F and V are $m \times m$ matrices defined as $F = \left[\frac{\partial \mathcal{F}_i}{\partial x_j}(x_0)\right]$ and $V = \left[\frac{\partial \mathcal{V}_i}{\partial x_j}(x_0)\right]$ with $1, \leq i, j \leq m$. The non-negativity of F follows from (A1) and (A4), V is a non singular matrix by (A2) and (A4) and all eigenvalues of J_4 have positive real parts by (A5).

If $f(x)$ satisfies (A1) – (A5), then the reproduction number is defined as $R_0 = \rho(\mathcal{F}\mathcal{V}^{-1})$ by Diekmann et al [81], where ρ is the spectral radius .

3.5 Kamgang-Sallet stability theorem

Consider the system:

$$\dot{x}_1 = A_1(x)(x_1 - x_1^*) + A_{12}(x)x_2 \quad (3.5)$$

$$\dot{x}_2 = A_2(x)(x_2) \quad (3.6)$$

on the positively invariant set $\Omega \in \mathfrak{R}_+^{n_1+n_2}$. The model (3.5)-(3.6) is studied and analyzed under the following assumptions:

- (H1) The system is defined on a positively invariant set Ω of the non negative orthant. The system is dissipative on Ω .
- (H2) The sub-system $\dot{x}_1 = A_1(x_1, 0)(x_1 - x_1^*)$ is globally asymptotically stable at the equilibrium x_1^* on the canonical projection of Ω on $\mathfrak{R}_+^{n_1}$.
- (H3) The matrix $A_2(x)$ is a Metzler matrix, i.e. a matrix such that off diagonal terms are non negative and is irreducible for any given $x \in \Omega$.
- (H4) There exists a matrix \bar{A}_2 , which is an upper bound for the set $M = A_2(x) : x \in \Omega$.
- (H5) $\alpha(\bar{A}_2) \leq 0$

Theorem 3.9. *If the hypothesis H1-H5 are satisfied the disease free equilibrium is globally asymptotically stable for system (3.5)-(3.6).*

For proof and discussion on H1-H5, see [87].

3.6 Sensitivity analysis

For every mathematical model, input factors such as parameters are not always known with a sufficient degree of certainty, they are subject to changes and errors. Owing to uncertainty on parameter values, it is important to correctly understand the possible effects of changes in parameter values to the anticipated model output [88, 89]. Sensitivity analysis can aid in identifying influential model parameters and optimizing model structure. The Latin Hypercube

Sampling-Partial Rank Correlation Coefficient (LHS-PRCC), a sensitivity analysis method, can be used to explore the uncertainty of parameters . LHS-PRCC sensitivity analysis determines the full parameter space of a model with an optimal number of computer simulations [88]. Within a given range of parameter values, LHS samples them to generate different values at each simulation and PRCC uses those values to describe the relation of parameters with the output of a particular mathematical model [90]. The goal of LHS-PRCC sensitivity analysis is to identify significant parameters which have great impact for model prediction and to rank these parameters depending on their contribution for a precise model prediction [91]. The combined LHS-PRCC procedure generally involves [92]:

- (i) LHS of the parameter space,
- (ii) obtaining model output for each set of sampled parameters
- (iii) ranking parameter and output values and replacing their original values with their ranks, and
- (iv) calculating the PRCC for each input parameter.

3.7 Optimal Control

Consider the following general control problem:

$$\min \left\{ \Upsilon(t_1, w(t_1)) + \int_0^{t_1} P(t, w(t), u(t)) dt \right\},$$

where, $w = [w_1, w_2, \dots, w_n]$ and $u = [u_1, u_2, \dots, u_n]$ are respectively the state and control variables. Further, w satisfies the following differential equation:

$$\frac{dw}{dt} = f(t, w, u), \quad w_0 = w(0), \quad t \in [0, t_1]. \quad (3.7)$$

Pontryagin's Maximum Principle transforms the maximization (or minimization) of an objective functional along with the state variables into maximizing (or minimizing) a point-wise Hamiltonian denoted by \mathcal{H} with respect to $u(t)$. This principle can be stated as follows:

Theorem 3.10. *If $u^*(t)$ and $w^*(t)$ are respectively the optimal control and state variables for (3.7) then there exists adjoint variables $\lambda(t)$ such that,*

$$\mathcal{H}(t, w, u, \lambda) \leq \mathcal{H}(t, w^*, u^*, \lambda),$$

where \mathcal{H} is given by,

$$\mathcal{H} = P(t, w(t), u(t)) + \lambda(t) f(t, w(t), u(t))$$

with,

$$\frac{d\lambda_i}{dt} = - \frac{\partial \mathcal{H}(t, w^*, u^*, \lambda)}{\partial w}, \quad \lambda(t_1) = 0,$$

where t_1 denotes the final time.

Proof. For the proof of Theorem 3.10, see [93]. □

The necessary conditions for the optimal state and control variables denoted by w^* and u^* are given as follows:

$$\left. \begin{aligned} \frac{d\lambda(t)}{dt} &= - \frac{\partial \mathcal{H}(t, w^*, u^*, \lambda)}{\partial w}, \\ \lambda(t_1) &= 0, \\ \frac{\partial \mathcal{H}}{\partial u} &= 0. \end{aligned} \right\} \quad (3.8)$$

If w^* , u^* and λ satisfy (3.8), then the corresponding control and the state variables are optimal.

To solve the optimality system, and thus, the optimal control problem we use the Forward-Backward Sweep method [94]. A general outline of the algorithm is given below. In the algorithm we take $w^* = (w_1, \dots, w_{N+1})$ and $\lambda^* = (\lambda_1, \dots, \lambda_{N+1})$ as the vector approximations (where the indices refer to time points) for the state and adjoint.

Step 1. Make an initial guess for u^* over the interval.

Step 2. Using the initial condition $w_1 = w(t_0) = a$ and the values for u^* , solve for w^* forward in time according to its differential equation in the optimality system.

Step 3. Using the transversality condition $\lambda_{N+1} = \lambda(t_1) = 0$ and the values for u^* and w^* , solve λ^* backward in time according to its differential equation in the optimality system.

Step 4. Update u^* by entering the new w^* and λ^* values into the characterization of the optimal control.

Step 5. Check convergence. If values of the variables in this iteration and the last iteration are negligibly close, output the current values as solutions. If values are not close, return to **Step 2**.

3.8 Brownian motion

Definition 3.11. A *Brownian motion* is a continuous time stochastic process $\{B(t), t \geq 0\}$ that satisfies the following conditions [95]:

(i) $B(0) = 0$,

(ii) the paths $t \rightarrow B(t)$ are continuous,

(iii) for $0 \leq s < t < \infty$, the increment $B(t) - B(s)$ is independent of $B(s)$ and

(iv) for $0 \leq s < t < \infty$ the increment $B(t) - B(s)$ has the normal distribution with mean 0 and variance $t - s$.

3.8.1 Itô Process

An Itô process $X(t)$ is adapted (for every realisation and every n , $X_n(t)$ is known at time n) and of the form:

$$X(t) = \underbrace{X(0)}_{\text{Initial value}} + \underbrace{\int_0^t \mu(s)ds}_{\text{drift}} + \underbrace{\int_0^t \sigma(s)dB(s)}_{\text{martingale noise (volatility)}} \quad (3.9)$$

Equation 3.9 can be written in differential form as:

$$dX(t) = \mu dt + \sigma(t)dB(t); \quad X(0) = x \quad (3.10)$$

3.8.2 The general Itô formula

Let $dX_t = U_t dt + V_{nm} dB_t$ be an m -dimensional Itô process and let $g(t, x) = (g_1(t, x), \dots, g_p(t, x))$ be a C^2 map from $[0, \infty) \times \mathbb{R}^m$ into \mathbb{R}^p , Then the process $Y(t, \omega) = g(t, X(t))$ is an Itô process whose component number k is given by:

$$dY_k = \frac{\partial g_k(t, x)}{\partial t} dt + \sum_i \frac{\partial g_k(t, x)}{\partial x_i} dx_i + \frac{1}{2} \sum_{ij} \frac{\partial^2 g_k(t, x)}{\partial x_i \partial x_j} dx_i dx_j \quad (3.11)$$

Example 3.12. Let $B = (B_1, \dots, B_n)$ be a Brownian motion in \mathbb{R}^n , $n \geq 2$, and consider

$$\begin{aligned} R(t, w) &= |B(t, w)| \\ &= (B_1^2(t, w) + \dots + B_n^2(t, w))^{\frac{1}{2}}, \end{aligned}$$

that is, the distance to the origin of $B(t, w)$. The function $g(t, x) = |x|$ is not C^2 at the origin, but since B_t hits the origin with probability 0 when $n \geq 2$ Itô's formula still works and we get

$$dR = \sum_{i=1}^n \frac{B_i dB_i}{R} + \frac{n-1}{2R} dt.$$

The process R is called the n -dimensional Bessel process because its generator is the Bessel differential operator $Af(x) = \frac{1}{2} f''(x) + \frac{n-1}{2x} f'(x)$ [95].

Chapter 4

A Model for Contact Tracing of an Infectious Disease: The West African Ebola Outbreak As An Example

In this chapter we present work that is part of a jointly-authored publication in *Mathematics* [96].

4.1 Introduction

In March 2014, the most deadly outbreak to date of the Ebola virus disease (EVD), a hemorrhagic fever, began in Guinea and rapidly spread to Liberia, Nigeria, Senegal, and Sierra Leone [97]. In October 2014, the World Health Organization (WHO) Ebola Response Team estimated an overall case fatality rate of 70.8% and basic reproduction numbers (\mathcal{R}_0) of 1.71 for Guinea, 1.83 for Liberia and 1.38 for Sierra Leone [98]. Concern that Ebola might spread globally via airline travel led to recommendations for health assessments at airports in the affected countries [99]. A review and meta-analysis of 31 reports found that the main methods of spread were direct contact with an infected individual and contact with deceased loved ones during traditional funeral practices [100]. In the 2014-2016 outbreak in Sierra Leone, amongst individuals confirmed to

have EVD, 47.9% reported that they had had contact with someone suspected of having EVD and 25.5% reported having attended a funeral [101]. These transmission pathways are further indicated as important by both mathematical and statistical models [102, 103, 104, 105].

Contact tracing for Ebola focuses primarily on people who have been in some kind of close contact with the infected or deceased individual. Throughout the outbreak, the Centers for Disease Control and Prevention’s Morbidity and Mortality Weekly Report detailed the progress of the disease as well as some information about contact tracing efforts. We formulate a model to represent the contact tracing process. The model uses a novel feature, which is explicitly counting the people being traced and connecting the total persons traced with the workload of contact tracer workers. Data used is from the Sierra Leone Ministry of Health [106, 107].

In this chapter, we develop a model for contact tracing using Sierra Leone data to fit the model. The model developed here can be applied to any infectious disease such as the coronavirus with similar spread characteristics.

4.2 Model Description

We develop a compartmental model following an SEIR approach, similar to [37, 38, 39, 44, 108]. In addition to the Susceptible, Exposed, Infected, and Recovered classes, we also include a class (D) to account for individuals who have died from Ebola in the community (i.e., having not been effectively isolated in a hospital or by other means), because they are a significant source of infection due to traditional funeral practices such as hugging and kissing the body of a deceased loved one. We also include a Hospitalized class (H), in which individuals are assumed to be isolated and not contributing to infection, and if they die their bodies are assumed to be disposed of safely. We place no upper limit on the size of class H , which does not reflect the situation during the outbreak where insufficient beds and staffing were a major limiting factor in controlling the outbreak [109], but it allows us to examine the operation of a contact tracing system assuming hospital resources are readily available.

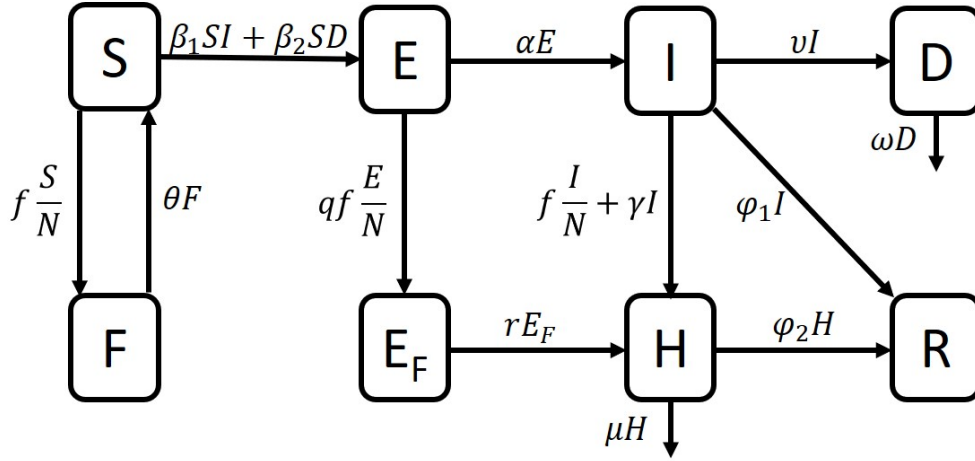


Figure 4.1: Flow diagram of population interactions.

Our investigation of contact tracing begins with the addition of two new classes of individuals being traced. Since exposure is a hidden trait, individuals being traced are either susceptible or exposed. We created a class F of susceptible individuals who are being traced because they came into contact with a suspected infected person or attended a funeral of an individual who died of Ebola. These individuals are observed for 21 days. If they do not become ill they return to S . A second class, E_F , for individuals being traced who are exposed and will become infectious, but these individuals could have moved elsewhere. Two events can lead to initiation of contact tracing:

- (i) either an individual is in contact with the D class or
- (ii) an individual enters the H class.

The contacts connected to the individuals involved in either of these two events will be contacted each day for 21 days by a contact tracer. We assume that individuals in the F class being traced will follow isolation guidelines to prevent them from becoming exposed. Individuals in E_F are moved to the hospital when they present symptoms. The function f alters the completion rate of key contact tracing steps based on the amount of work to be done along with the number of available contact tracing staff. There is a limited number of contact tracers, and each contact tracer is able to trace a limited number of individuals at a time. To account for this, we place a threshold on the total number of contacts that can be traced at a time. Part of the work done by contact tracers is moving individuals

to the hospital, and the remaining effort is dedicated to visiting contacts who haven't (yet) displayed any symptoms of Ebola. In Figure 4.1 one can see the terms with coefficients f representing the effects of contact tracing. We consider the following model:

$$S' = -\beta_1 SI - \beta_2 SD - f \frac{S}{N} + \theta F \quad (4.1)$$

$$F' = f \frac{S}{N} - \theta F \quad (4.2)$$

$$E' = \beta_1 SI + \beta_2 SD - qf \frac{E}{N} - \alpha E \quad (4.3)$$

$$E'_F = qf \frac{E}{N} - r E_F \quad (4.4)$$

$$I' = \alpha E - f \frac{I}{N} - \gamma I - \phi_1 I - \nu I \quad (4.5)$$

$$H' = r E_F + f \frac{I}{N} + \gamma I - \phi_2 H - \mu H \quad (4.6)$$

$$R' = \phi_1 I + \phi_2 H \quad (4.7)$$

$$D' = \nu I - \omega D \quad (4.8)$$

where $N = S + E + I$. The function f depends on F , E_F , and I and gives the rate of finding new contacts,

$$f = \begin{cases} \kappa_1 \gamma I + \kappa_2 \nu I & \text{if } F + E_F < (15)(1200)p \\ 0 & \text{else} \end{cases} \quad (4.9)$$

Here $1 - p$ is the proportion of the total available contact tracing effort dedicated to hospitalizing individuals identified as symptomatic. Note that the two events (movement into H and D) can be seen in the function f with the rates γI and νI . In the cutoff for f , the number 15 is how many contacts on average one contact tracer can trace and the number 1200 is the maximum number of contact tracers that were employed in the Western Area, Sierra Leone (containing the capital city of Freetown) during the 2014-2016 epidemic [110]. Although the total number of contact tracers varied throughout the outbreak, we decided to assume that the maximum of 1200 were available throughout the outbreak. The units and interpretation of each parameter are listed in Table 4.1. Note that we

Table 4.1: The parameters and compartment names in our model with their interpretations and units.

Variable or Parameter	Interpretation	Units
β_1	transmission from interactions between I and S	per person per time
β_2	transmission from interactions between D and S	per person per time
$1/\theta$	number of days a person is traced	time
$1/\alpha$	length of the exposed period	time
r	rate of hospitalization for traced individuals	per time
γ	rate of hospitalization for untraced individuals	per time
ϕ_1	recovery rate for untreated	per time
ϕ_2	recovery rate for treated	per time
ν	death rate for untreated	per time
μ	death rate for treated	per time
ω	rate at which dead bodies become non-infectious	per time
κ_1	contacts recruited from hospitalization of one person	unitless
κ_2	contacts recruited from funeral of one person	unitless
q	scaling factor for exposed contacts	unitless
S	susceptibles	individuals
F	susceptibles being traced	individuals
E	exposed	individuals
E_F	exposed being traced	individuals
I	infectious	individuals
H	hospitalized	individuals
D	dead bodies	individuals
R	recovered	individuals

do not account for births or for deaths from any other cause than Ebola.

People can move from the Susceptible class to the Exposed class by coming into contact with a member of the Infectious class (term $\beta_1 SI$) or by coming into contact with an infectious dead body (term $\beta_2 DS$). People who are being traced move from the Susceptible class to F or from the Exposed class to E_F by coming into contact with a person who has just been hospitalized or attending a funeral for somebody who has just died of Ebola (term $f \frac{S}{N} = (\kappa_1 \gamma I + \kappa_2 \nu I) \frac{S}{N}$). This term is scaled by N because the persons moving in tracing are moved proportionally to the ratio of persons in their current class. For example, a person being traced from S moves to F at a rate proportional to $\frac{S}{N} = \frac{S}{S+E+I}$. A person is more likely to be in E_F while being traced than to be in F because of the contact they had with either an infected person or a dead body. To account for this, we multiply the term $f \frac{E}{N}$ by a number $q > 1$, a scaling factor to increase the likelihood of individuals in E_F being traced relative to individuals in F being traced. People who have completed their time being traced and haven't developed symptoms move back into S (term θF). Once a person has been in the Exposed class for an average of 10 days, they move to the Infectious class (term αE). A person in the class E_F is moved to the hospital once they develop symptoms (term $r E_F$). If an individual being traced shows symptoms the first time they are contacted, they

are immediately moved to the hospital (term $f \frac{I}{N}$). Some Infectious people decide to go to the hospital on their own (term γI). Some Infectious people manage to survive Ebola and move to R (term $\phi_1 I$) but others die of the disease and we assume they are not safely buried and contribute to the class D (term νI). This is a simplifying assumption, because as the epidemic drew on many people who died in the community were safely buried. Some Hospitalized individuals will recover (term $\phi_2 H$) but others will die and be safely buried (term μH). After some time has passed, an unsafely buried dead body is no longer able to infect people (with decay term ωD).

4.3 The Reproductive number, \mathcal{R}_0 .

We derive the basic reproductive number, \mathcal{R}_0 , using the next generation matrix method [80, 111, 112, 113]. We expect that near the disease free equilibrium (DFE), the number of infections will be small but nonzero. The population affected by the outbreak consisted entirely of susceptibles at the beginning of the outbreak. Therefore for this analysis we assume that $f = \kappa_1 \gamma I + \kappa_2 \nu I$.

Equation (4.7) implies that

$$\phi_1 I^* = -\phi_2 H^*, \quad (4.10)$$

giving $I^* = H^* = 0$.

From equation (4.8), we get $D^* = 0$. Since $I^* = 0$, equation (4.2) gives $F^* = 0$ and equation (4.5) gives $E^* = 0$. Since $I^* = H^* = 0$, equation (4.6) gives $E_F^* = 0$. Since $E^* = I^* = 0$, we conclude that $S^* = N(0)$. We have the DFE: $(S^*, 0, 0, 0, 0, 0, R^*, 0)$, but we take $R^* = 0$ for computation of the next generation matrix. The diseased classes are: E , E_F , I , H , and D , with corresponding vectors $\mathcal{F} - \mathcal{V}$ forming the right hand side of the system with only diseased classes,

$$\mathcal{F} = \begin{pmatrix} \beta_1 SI + \beta_2 SD \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathcal{V} = \begin{pmatrix} \alpha E + q(\kappa_1 \gamma I + \kappa_2 \nu I) \frac{E}{S+I+E} \\ r E_F - q(\kappa_1 \gamma I + \kappa_2 \nu I) \frac{E}{S+I+E} \\ (\phi_1 + \nu + \gamma) I + (\kappa_1 \gamma I + \kappa_2 \nu I) \frac{I}{S+I+E} - \alpha E \\ (\phi_2 + \mu) H - r E_F - \gamma I - \frac{(\kappa_1 \gamma + \kappa_2 \nu) I^2}{S+I+E} \\ \omega D - \nu I \end{pmatrix}.$$

Note that our DFE is not unique and this is not required. We get the Jacobian matrices $D\mathcal{F}(E, E_F, I, H, D)$ and $D\mathcal{V}(E, E_F, I, H, D)$ at the DFE, $(0, 0, 0, 0, 0)$ with $S = S^*$,

$$D\mathcal{F}(0, 0, 0, 0, 0) = \begin{pmatrix} 0 & 0 & \beta_1 S^* & 0 & \beta_2 S^* \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$D\mathcal{V}(0, 0, 0, 0, 0) = \begin{pmatrix} \alpha & 0 & 0 & 0 & 0 \\ 0 & r & 0 & 0 & 0 \\ -\alpha & 0 & \phi_1 + \nu + \gamma & 0 & 0 \\ 0 & -r & -\gamma & \phi_2 + \mu & 0 \\ 0 & 0 & -\nu & 0 & \omega \end{pmatrix}.$$

Thus the basic reproductive number we obtain as the spectral radius of the matrix $D\mathcal{F}(0, 0, 0, 0, 0)(D\mathcal{V})^{-1}(0, 0, 0, 0, 0)$ is

$$\mathfrak{R}_0 = \frac{\beta_1 S^*}{\phi_1 + \nu + \gamma} + \frac{\nu \beta_2 S^*}{\omega(\phi_1 + \nu + \gamma)}. \quad (4.11)$$

The first term describes the number of new infections that we expect per individual from the I class, and the second term describes the number of new infections that we expect per body in the D class.

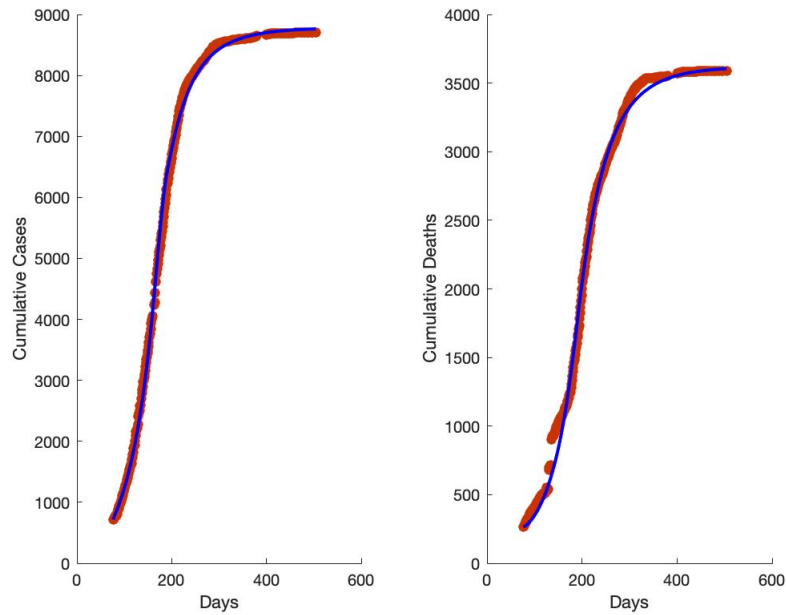


Figure 4.2: The value of the objective for this simulation was $J = 0.0423$.

4.4 Numerical Simulations

4.4.1 Parameter estimation

Our data are taken from the Sierra Leone Ministry of Health daily situation reports, published on their website during the epidemic. We accessed these old web sites via the Wayback Machine at:

https://web.archive.org/web/20150314233800/http://health.gov.sl/?page_id=583
(28/02/2020).

Data are listed in Appendix B. Situation reports were available beginning at day 77 with the final day being day 504, but not every intermediate day had a report. The missing data are for days when the Ministry of Health situation report was unavailable. There were 343 total reports available for usage. From these reports we used confirmed cases and deaths. There was one report we chose to exclude because it listed more confirmed deaths than subsequent reports, making the total number of data points 342. In Figure 4.2 the days with missing data are represented by gaps in the red dots.

We chose some parameters from the literature and estimated others using

data. The parameters

$$\alpha = 0.1, \frac{1}{\omega} = 4.5, \frac{1}{\theta} = 21,$$

were taken from the literature [37, 114, 44, 108]. Data indicated that the initial condition for the H class was $H(0) = 94$ individuals. We assumed the initial condition for the recovered class was $R(0) = 0$ individuals, and that the initial condition for S was roughly equivalent to the population of Sierra Leone at the time, $S(0) = 6,348,350$ people. We estimated the following parameters:

$$\beta_1, \beta_2, \gamma, \kappa_1, \kappa_2, r, p, \nu, \mu, \phi_1, \phi_2,$$

and the following initial conditions:

$$F(0), E(0), E_F(0), I(0), D(0).$$

The parameters were estimated in MATLAB using multistart to generate many vectors of starting parameter estimates. Each vector was used to initialize a search in `fmincon` (a local minimizer) and `ode45` served as our ODE solver. Parameter upper and lower bounds were based on ranges of parameters from the literature [44, 37] and from the Sierra Leone data. For example, parameters comparable to our $\beta_1, \beta_2, \phi_1,$ and ϕ_2 were found in Rivers [37]. Our lower limit for r was based on both papers [44, 37]. There is also a parameter comparable to our parameter κ_1 in [44]. The upper bound for F_0 was taken as 2500 because our data indicated that in the early days this was roughly the number of contacts being traced. To estimate the cumulative simulated cases, we summed over the entries into the H class, assuming that cases for people in the community were unconfirmed. To estimate the cumulative simulated deaths, we summed over the deaths from H and I together. The data to be compared with simulation results are cumulative confirmed cases and cumulative confirmed deaths. We minimized the following

$$J = \sum_{i=77}^{504} \left(\frac{(\text{Cases}_{Estimated}(i) - \text{Cases}_{Data}(i))^2}{(\text{Cases}_{Data}(i))^2} + \frac{(\text{Deaths}_{Estimated}(i) - \text{Deaths}_{Data}(i))^2}{(\text{Deaths}_{Data}(i))^2} \right), \quad (4.12)$$

which is a type of sum of least squares for our model.

We had two primary goals during the process of parameter estimation:

- (i) fit the data with a low J value and
- (ii) in each class we wanted reasonable dynamics, meaning approximately the correct magnitude in the size of each compartment.

We tried several ways of fitting the data. First, we estimated all the parameters listed above, holding them all constant throughout the epidemic. This resulted in simulated epidemic curves that did not flatten at the end, indicating the epidemic would have kept going (see Appendix A). In order to achieve a good simulation of the data with reasonable compartments, we modified the model by inserting the parameter q . In order to achieve a simulated fit of the data which would include a flattening of the cumulative cases and cumulative deaths curves, rather than simulations which indicated the epidemic wouldn't have ended, we allowed some parameters (specifically $\beta_1, \beta_2, \gamma, \kappa_1$, and κ_2) to vary over the course of the epidemic. These parameters were chosen because people's behavior change during an epidemic. We smoothed the transition from the first value of each of these parameters to the second value using piecewise functions such as the one below for each of the parameters,

$$\beta_1(t) = \begin{cases} 6.68e^{-8} & t < 160 \\ 6.68e^{-8}(1 - \frac{t-160}{30}) + 3.94e^{-8}(\frac{t-160}{30}) & 160 \leq t \leq 190 \\ 3.94e^{-8} & t > 190. \end{cases} \quad (4.13)$$

Chowell et al. [36] built a system of ODEs representing Ebola outbreaks in Congo and Uganda and used a smooth transition between two transmission rates due to control interventions (like education and contact tracing followed by quarantine).

The literature supports the decision to allow $\beta_1, \beta_2, \gamma, \kappa_1$ and κ_2 to change over the course of the epidemic. Senga et al. [108] analyzed data on probable and confirmed cases of EVD and their contacts in Kenema district, Sierra Leone taken from the national database. They found that the number of contacts per case increased over time. The low number of contacts per case reported early

in the epidemic was much lower than those reported in other countries, which they concluded meant that the contact listings were incomplete. Olu et al. found that during the months of June 2014 to November 2014 the average number of contacts per case was 9 and that during the months of December 2014 to May 2015 the average number of contacts per case increased to 16 [44]. Lokuge et al. reported that later in the epidemic people were more likely to come to the hospital of their own volition, less likely to report funeral contact, and that contact tracing increased in efficacy [109]. These findings from the literature indicate it is reasonable to conclude that values for $\beta_1, \beta_2, \gamma, \kappa_1$ and κ_2 changed during the course of the epidemic due to changes in behavior and level of education in the population about EVD.

However, we were unable to generate reasonable sizes for compartment E_F . Numerical simulations were showing very few people passing through E_F , which is not reflective of the success contact tracing achieved in locating exposed individuals during the outbreak. We decided to modify the model by adding a multiplier, q , in front of the $f \frac{E}{N}$ term. We tried several values and found that a value of $q = 100$ generated reasonable sizes for compartment E_F . This multiplier indicates that people who were being traced had had contact with an individual who was infectious or with a dead body, so they were more likely to have been exposed to Ebola than a member of the population who hadn't had such contact. These changes resulted in the simulations shown in Figures 4.2, 4.3, and 4.4 which were generated using the parameters found in Table 4.2.

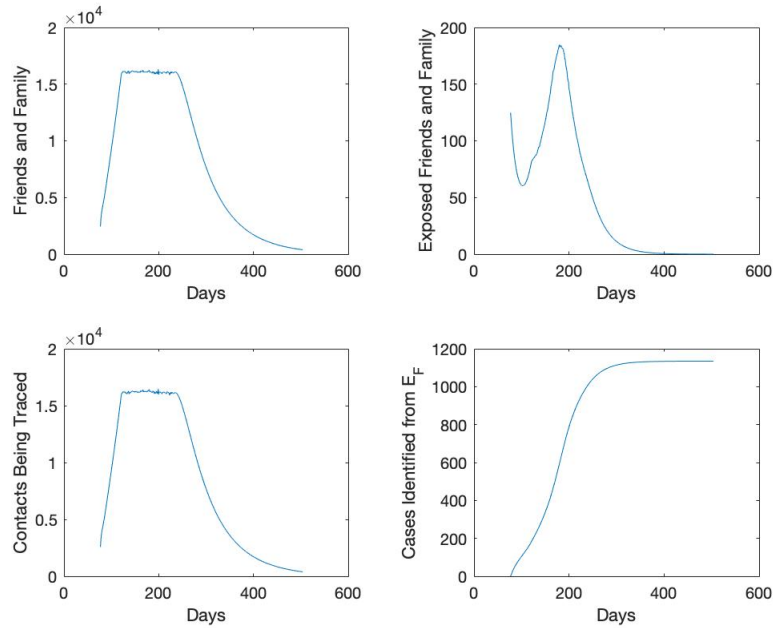


Figure 4.3: Dynamics of class F in the upper left, class E_F in the upper right, their sum on the bottom left, and the integral of those leaving E_F to be hospitalized on the bottom right. These classes correspond to the parameters from Table 4.2 and the data simulations from Figure 4.2.

Figure 4.3 shows the total number of cases that were identified as part of contact tracing efforts. Near the end of the outbreak this number reaches about 1100, which represents more than a tenth of all confirmed cases. This demonstrates the importance of successful contact tracing. The peak of contact tracing numbers corresponds to the slowing of the increase in cumulative cases, around day 200. This indicates that contact tracing efforts contributed to ending the epidemic.

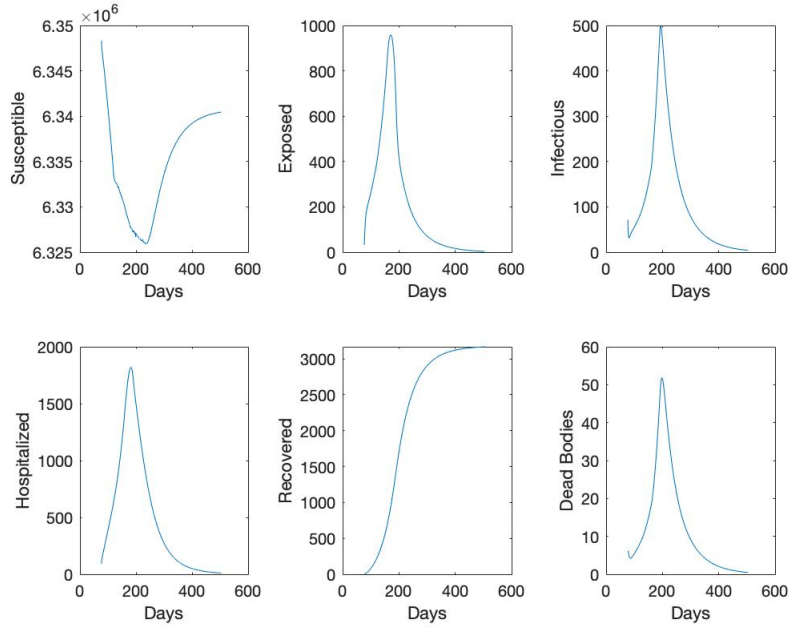


Figure 4.4: The graphs above correspond to the parameters from Table 4.2 and the data simulations from Figure 4.2. Note that the scales are all different.

Note that in Figure 4.4 the increase later in the epidemic of S results from people returning to S from F after being traced for 21 days and showing no symptoms. In Figure 4.4, the peak in E occurs at day 164, the peak in H about two weeks later on day 176, the peak in I about two weeks after that on day 192, and then the peak in D on day 197. It is not surprising that the peak in E precedes the other peaks, but it is surprising that the peak in D is the last peak to occur. This indicates that there may have been unsafely buried bodies later in the epidemic, but that fewer people were catching Ebola from funeral interactions despite this increase in funerals.

In Table 4.2, there is no difference between β_1 early and β_1 late. However, β_2 changes from an early value of $1.00 * 10^{-6}$ to a much lower later value of $1.00 * 10^{-7}$. These parameter values indicate that while the rate of transmission from interactions between S and I remained about the same throughout the epidemic, the rate of transmission from D to S decreased dramatically as people became more educated about Ebola. Oddly, $\gamma = 0.41$ decreases to a later value of $\gamma = 0.062$, which does not agree with accounts from the literature that people were more likely to come to the hospital once they developed symptoms later in

Table 4.2: Values for parameters, with five parameters having early and late values. Parameters with * were taken from the data or the literature. Others were estimated.

Parameter	Value	Parameter	Value
β_1 early	$1.00e^{-9}$	r	0.056
β_1 late	$1.00e^{-9}$	p	0.90
β_2 early	$1.00e^{-6}$	ν	0.024
β_2 late	$1.00e^{-7}$	μ	0.010
γ early	0.41	ϕ_1	0.020
γ late	0.062	ϕ_2	0.028
κ_1 early	29.74	$F(0)$	2451.10
κ_1 late	44.93	$E(0)$	32.04
κ_2 early	44.62	$E_F(0)$	124.88
κ_2 late	16.61	$I(0)$	71.76
		$D(0)$	6.09
α^*	0.1	$1/\omega^*$	4.5
$H(0)^*$	94	$S(0)^*$	6, 348, 350
$R(0)^*$	0	$1/\theta^*$	21

the epidemic than they were earlier in the epidemic. The value of $\kappa_1 = 29.74$ early increases to $\kappa_1 = 44.93$ late, corresponding to reports from the literature that people were more likely to report more complete lists of contacts later in the epidemic. However, $\kappa_2 = 44.62$ early decreased to $\kappa_2 = 16.61$ late, adding to the conclusion that people were less likely to attend traditional funerals later in the epidemic. The changes in these parameters during the outbreak might be caused by a combination of factors including educating the public about Ebola [115], increases in available beds at Ebola Treatment Centers, and more effective implementation of contact tracing.

The value of $r = 0.056$ means that contacts who were infected took an average of 18 days to show symptoms. This value for r is probably unrealistically small, as it should likely be closer to $\alpha = 0.1$. The parameter ν was slightly larger than μ , since those who were treated had slightly lower chance of dying from Ebola. Similarly, ϕ_2 was larger than ϕ_1 because those who were treated were more likely to recover from the disease.

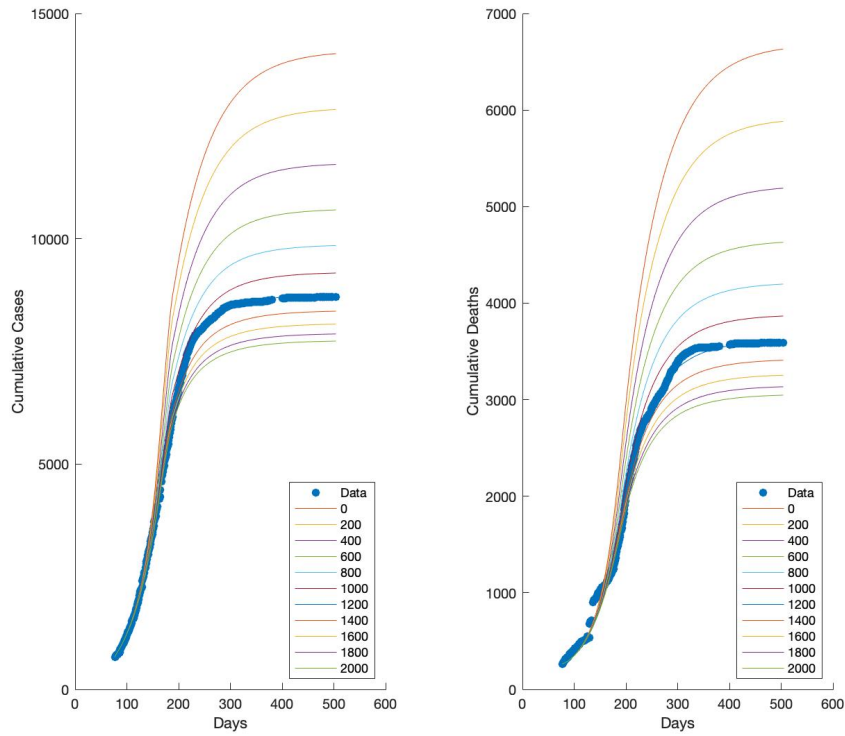


Figure 4.5: Effect of varying the number of contact tracers available from 0 to 2000, with 1200 as the corresponding number in our model.

4.4.2 Importance of contact tracing

Figure 4.5 shows potential trajectories for epidemics with different numbers of contact tracer workers available, either more or fewer than were actually available during the epidemic. We varied the number of these workers from 0 to 2000, and note that 1200 is the corresponding number in our model. Without contact tracing at all, the highest blue curve, there would have been thousands more cases and deaths. Even a much smaller work force than existed would have made a dramatic improvement on the trajectory of the epidemic from what would have happened without contact tracing. Once the number of contact tracers reaches about 1000, each increase in the number of workers has much less dramatic effect. More tracers still would have been better, but the difference in trajectories is much less dramatic than the difference between 0 tracers and 200 tracers.

The number of persons traced from each hospitalization (κ_1) and the num-

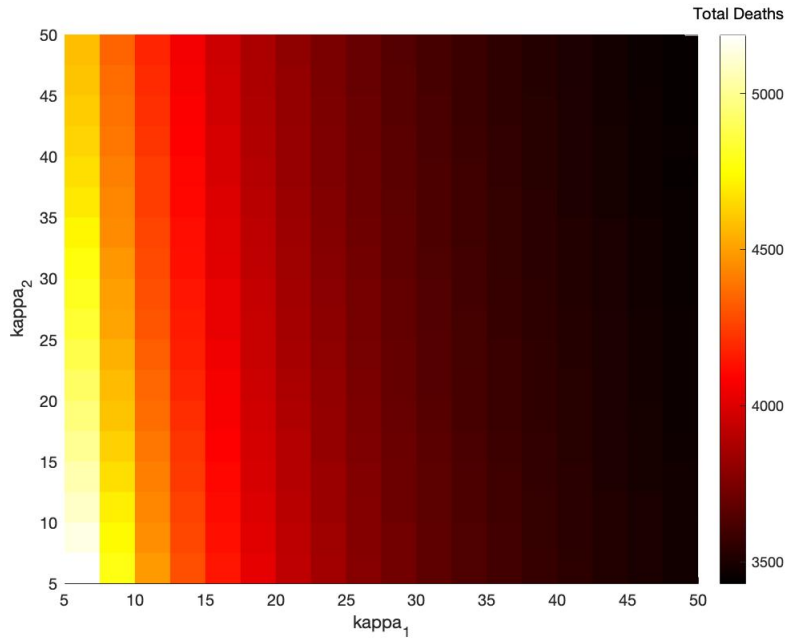


Figure 4.6: Effect of varying contact tracing parameters κ_1 and κ_2 on the total number of deaths by day 504 of the epidemic.

ber from each funeral (κ_2), were estimated as $\kappa_1 = 29.7$ early, $\kappa_1 = 44.9$ late, $\kappa_2 = 44.6$ early, and $\kappa_2 = 16.6$ late in the model. We varied those numbers from 5 to 50 to see the effect on the epidemic. If we hold each of the contact tracing parameters κ_1 and κ_2 constant at the values in Figure 4.6, the heat map shows the total number of deaths by day 504 of the outbreak. Increasing each of the two parameters reduces the total number of deaths, but κ_1 has a much more dramatic effect than κ_2 . This seems to indicate that more deaths resulted from people having contact with infected individuals than resulted from people having contact with dead bodies.

Chapter 5

COVID-19 changing the face of the world. Can sub-Saharan Africa cope?

5.1 Introduction

The coronavirus pandemic has disrupted global economies and health systems in unprecedented ways. As of 19 November 2021 there were 256 507 133 recorded cases with 5 150 613 deaths and 231 700 776 recoveries, [116].

Developed countries including the United States of America (USA), Italy, Spain, United Kingdom (UK), France, Germany, Russia, Brazil and China have experienced high numbers of COVID-19 cases. Most of these developed countries have excellent health facilities. However, the COVID-19 epidemic in Italy, Spain, France and the USA has demonstrated that the current medical facilities were not designed to serve the populations during a pandemic. The current Western health systems performed well during the initial stages of the coronavirus disease progression but have since been challenged due to the acute rise in infection rates. Various countries have had to make decisions over who is offered or not offered a bed in an Intensive Care Unit (ICU). The decisions were based on hospital bed capacity to avoid hospital overload and were dependent on early testing and isolating those who test positive as was the case in South Korea, Germany and China [117].

The epidemic in these developed countries has had very serious effects on the overall infrastructure and livelihoods of the people living there, and we believe that the effects on the sub-Saharan Africa region could very likely overwhelm this region where the economies are very weak and their ICU capacity compared to population sizes averages only 9 beds per 10 000 inhabitants [118]. The current economic landscape will likely impact the ability of the various health systems in the long term to supply personal protective equipment (PPE) required by the frontline workers as they take care of infected patients.

Sub-Saharan African governments do not have the capacity to construct overflow hospital bed capacity. This study will address the following objectives:

- (i) to determine the total number of individuals likely to be infected with COVID-19.
- (ii) to find possible hospitalization strategies that would not overload hospital bed capacity and the number of infected individuals who would need to be safely isolated.
- (iii) to find a strategy which yields the lowest number of deaths.

5.2 Do people learn from previous epidemics: Ebola

To date Africa has registered very few cases of Covid-19 infections and deaths. One wonders whether frequent exposure of the African population to epidemics has prepared it to prevention protocols. We give examples from Ebola in the Democratic Republic of Congo (DRC) and the Sudan, where prior exposure to the epidemic reduced the disease caseload.

The data for various epidemics in the DRC and the Sudan presented in Figures 5.1(a) and 5.1(b) shows a very high peak of total infections for the first epidemic and declining number of total infected during subsequent epidemics. The data suggests that the populations in the two countries have learnt how to manage an epidemic (Ebola) which requires high standards of hygiene and social distancing. Levy et al, [115] have shown that the severity of an outbreak

is linked to the level of prior knowledge and education of the general population as well as preparedness of health care facilities.

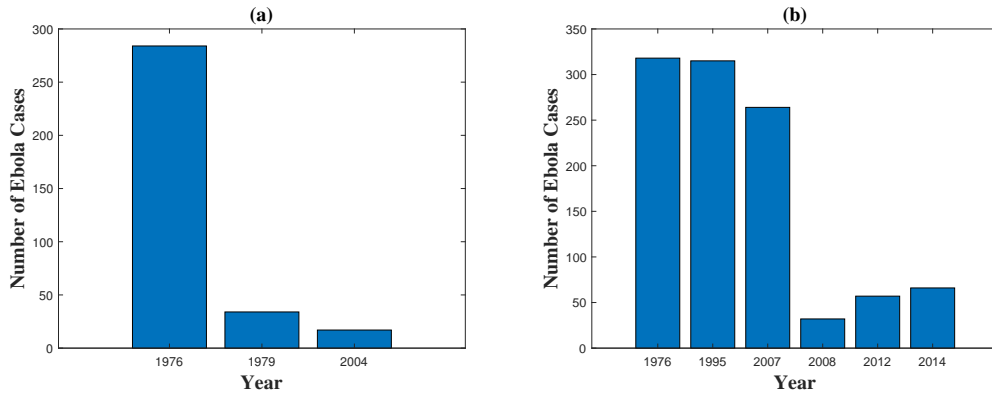


Figure 5.1: Ebola outbreaks plots for (a) Sudan and (b) the Democratic Republic of Congo.

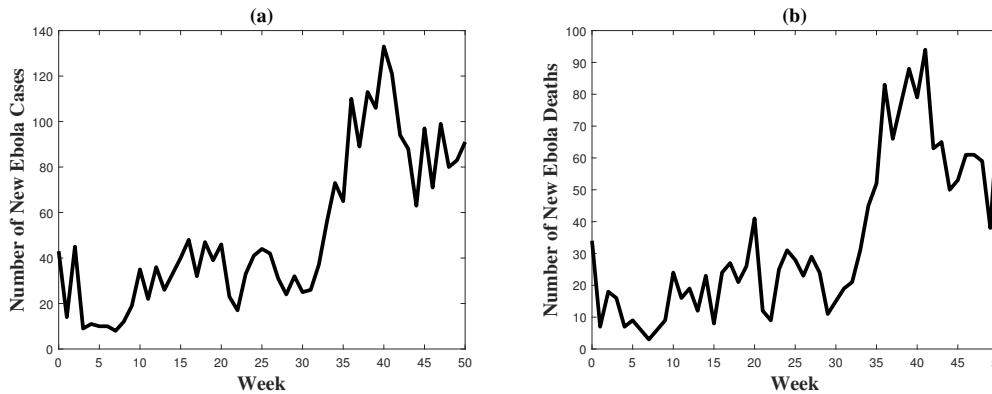


Figure 5.2: Democratic Republic of Congo plots for (a) new Ebola cases (b) new Ebola deaths.

Using data from the Democratic Republic of Congo (DRC) for new Ebola cases and new Ebola deaths (5 August 2018 to 21 July 2019), we notice a trend of rising to a peak and then declining, indicating two phases

- (i) naivety to the virus early in the infection and
- (ii) experience towards adherence to prevention protocols (Figures 5.2(a) and 5.2(b)).

Ebola and COVID-19 management are similar in many ways. They both rely on very high standards of hygiene, avoiding hand shaking and low density occupancy in residential homes. The difficulty one encounters when modeling Ebola or COVID-19 is modeling the infections arising from poor hygiene, social

distancing etc. From Figures 5.1 and 5.2, we suggest modeling infections from poor hygiene and lack of social distancing by a function fitted to the data and depicted in Figure 5.3. This function represents high infections due to poor understanding of the prevention measures early into the epidemic, but as the infection progresses individuals who apply the knowledge from previous epidemics adapt to avoid infections as was the case for Ebola. The function in Figure 5.3 mimics the process observed for Ebola and is adopted in this study.

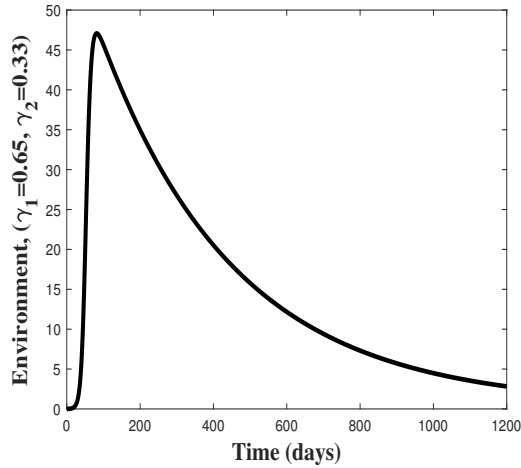


Figure 5.3: Time course for the environment

5.3 Quality of the Lockdown

Since SARS-2 is a respiratory disease, we want to incorporate the effects of household transmission due to an infected individual i , in household h at time t during the lockdown. The rate of exposure from a member of the family j to i is given by:

$$\nu_h = \eta_g \times \psi_H \times \psi_{i,\text{inf}} \quad (5.1)$$

where ψ_H depends on household size and $\psi_{i,\text{inf}}$ represents asymptomatic or symptomatic status of the infection. An individual in household h is also exposed to community transmission because a member of the household has to leave the house to transact with the community, for example to buy food or to go

to work (if they are front-line workers) during the lockdown. This movement is inevitable in poverty stricken Africa where food insecurity is common and food must be sourced daily. The rate of community exposure is given by:

$$C_h = \varepsilon_g \times f_g(t) \times \psi_{i,\text{age age group}}, \quad (5.2)$$

where ε_g is the baseline exposure from the community, f_g is a time dependent curve that modifies the community rate of exposure over time and $\psi_{i,\text{age}}$ accounts for disease susceptibility depending on age.

The rate of exposure of individual i in household h in which a member goes out to transact with the community is given by:

$$\lambda = S_{i,g(t)} \left[M_{i,h(t)} \sum_{j \neq i} \nu_h + C_h \right] \quad (5.3)$$

(3) has been used to moderate transmission of respiratory syncytial virus which is similar to SARS-2, [119]. The approach described in (5.1), (5.2) and (5.3) is described in detail in [119]. We have used the data from [119] to incorporate the quality of the lockdown and to explain why the number of COVID-19 cases exploded after the lockdown.

5.4 Model Description

We develop a simple model in the context of the sub-Saharan Africa environment which consists of a class of individuals, S , who are susceptible to the disease, a class of individuals, E , who have been exposed to the disease, a class of asymptomatic individuals, I_{as} . These are individuals who are not showing symptoms but are transmitting the virus. A class of symptomatic infectives, I_s , a class of individuals who require hospitalization or to be isolated, H , and a class of recovered individuals, R . We note that I_{as} and I_s can infect S directly and indirectly by contaminating the environment ϕ . For simplicity of notation, let $x(t) = (x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t)) = (S(t), E(t), I_{as}(t), I_s(t), H(t), R(t))$, $x_7(t) = \phi(t)$ and $N(t) = x_1(t) + x_2(t) + x_3(t) + x_4(t) + x_5(t) + x_6(t)$.

We consider the following model:

$$\frac{dx_1}{dt} = \pi - \frac{(\beta_1 x_3 + \beta_2 x_4)x_1}{N} - \beta_3 x_7 x_1 - \mu x_1, \quad (5.4)$$

$$\frac{dx_2}{dt} = \frac{(\beta_1 x_3 + \beta_2 x_4)x_1}{N} + \beta_3 x_7 x_1 - (\alpha_1 + \mu)x_2, \quad (5.5)$$

$$\frac{dx_3}{dt} = \alpha_1 x_2 - (\gamma_1 + \alpha_2 + \mu + \kappa_1)x_3, \quad (5.6)$$

$$\frac{dx_4}{dt} = \alpha_2 x_3 - (\gamma_2 + \delta_1 + \kappa_2)x_4, \quad (5.7)$$

$$\frac{dx_5}{dt} = \gamma_1 x_3 + \gamma_2 x_4 - (\delta_2 + \kappa_3)x_5, \quad (5.8)$$

$$\frac{dx_6}{dt} = \kappa_1 x_3 + \kappa_2 x_4 + \kappa_3 x_5 - \mu x_6, \quad (5.9)$$

where

$$(\pi, \mu) = \begin{cases} (0, 0) & \text{during lockdown} \\ \neq (0, 0) & \text{no lockdown, inter-zonal} \\ & \text{movement allowed.} \end{cases} \quad (5.10)$$

For short lockdown periods one can assume that $\mu = 0$. This assumption is not valid for long lockdowns. The quality of the environment is described by the equation:

$$\frac{dx_7}{dt} = \frac{\alpha_3(x_3 + x_4)}{N} - \mu_1 x_7. \quad (5.11)$$

The parameter α_3 measures the rate of contaminating the environment, that is, the rate at which the amount of pathogens are released into the environment by both asymptomatic and symptomatic infectives. Equation 5.11 is also supported by Berge et al. [120] for their model on Ebola. The model flow diagram is given below: where $\theta = \frac{(\beta_1 x_3 + \beta_2 x_4)x_1}{N} + \beta_3 x_7 x_1$.

Equation (5.4) describes the rate of change in the susceptible population, x_1 . The first term in (5.4) represents recruitment of people into the susceptible class through movement of people from different zones when this is allowed. This term is zero during a lockdown but may be nonzero if inter-zone movements are allowed. The second term represents loss due to infection of susceptible individuals by asymptomatic infectives, x_3 , at the rate β_1 and symptomatic infectives, x_4 , at the rate β_2 . The third term represents indirect infections due to an unclean environment at the rate β_3 . As time has progressed, it has become

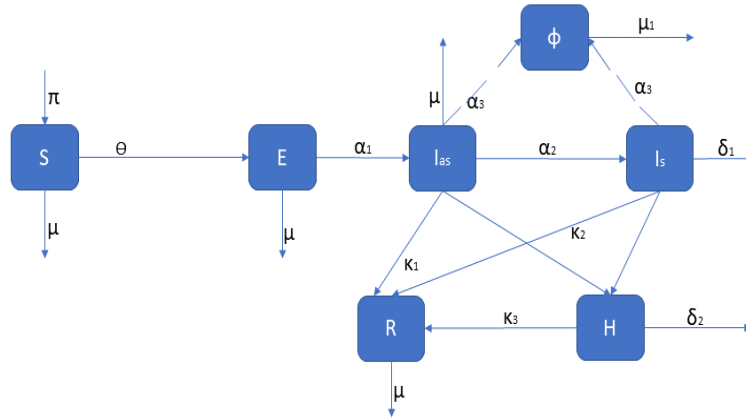


Figure 5.4: Model flow diagram

necessary to account for loss due to other causes such as natural death at the rate μ . There has been confusion over the number of covid deaths and it has become pathologically necessary to ascertain that a covid infected individual actually died of covid complications, [121, 122]. However, μ was ignored during early stages of the pandemic and every covid infected individual was assumed to have died of covid.

Equation (5.5) describes the rate of change of the exposed class, x_2 . The first two terms represent gain from infection of susceptible individuals and the third term represents loss due to sero-conversion to the asymptomatic state at the rate α_1 and natural death at a constant rate μ , respectively. The same comment regarding μ in equation (5.4) applies in this case.

Equation (5.6) describes the rate of change of the asymptomatic infected class, x_3 . The first term represents gain from exposed individuals who are converting to sero-positive status without exhibiting symptoms. The second term represents loss due to conversion to symptomatic state and hospitalization at the rates α_2 and γ_1 , respectively. κ_1 represents losses from this class due to recovery

and μ represents natural death as explained in (5.10).

The first term on the right hand side of equation (5.7) represents gain from the asymptomatic state. The second term represents loss due to hospitalization or isolation and a blanket term representing loss due to both natural death and disease induced death at the rates γ_2 and $\delta_1 = (\mu + \delta_{x_4})$, respectively. Loss from this class due to recovery is assumed to occur at a constant rate κ_2 .

Equation (5.8) describes the rate of change of the hospitalized and isolated class. The first two terms represent gain from testing and contact tracing of both asymptomatic and symptomatic individuals. The third term represents loss due to blanket death at the rate $\delta_2 = (\mu + \delta_{x_5})$ and recovery at the rate κ_3 .

Equation (5.9) describes the rate of change of the recovered individuals. The first three terms represent gain from recovery of asymptomatic, symptomatic and hospitalized cases. The recovered class loses people through natural death at a rate μ for a long pandemic.

Equation (5.11) describes the rate of contaminating the environment due to the release of pathogens into the environment by symptomatic and asymptomatic infectives at the rate α_3 . The second term represents cleaning of the environment (naturally or due to interventions) at the rate μ_1 .

5.5 Model Analysis

5.5.1 Positivity of solutions

Denote by \mathfrak{R}_+^6 the points $x(t) = (x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t))$ in \mathfrak{R}^6 with positive coordinates and consider the system (5.4) -(5.9) with initial values

$$x^0 = (x_1^0, x_2^0, x_3^0, x_4^0, x_5^0, x_6^0) \in \mathfrak{R}_+^6.$$

We can state the following Lemma:

Lemma 5.1. *If $x_i^0 \geq 0$, $i = 1, \dots, 6$ then $x_i(t) \geq 0$ for $t > 0$, $i = 1, \dots, 6$.*

Proof:

First, we want to show that $0 \leq x_7(t) \leq \frac{\alpha_3}{\mu_1}$ where the lower bound signifies

a clean environment and the upper bound signifies a Covid-19 contaminated environment. From (5.11), we have

$$\begin{aligned}\frac{dx_7}{dt} &= \frac{\alpha_3(x_3 + x_4)}{N} - \mu_1 x_7 \\ &\leq \alpha_3 - \mu_1 x_7.\end{aligned}$$

The solution for a totally contaminated COVID-19 environment is given by

$$x_7(t) \leq x_7(0) e^{-\mu_1 t} + \frac{\alpha_3}{\mu_1} (1 - e^{-\mu_1 t}).$$

As

$$t \rightarrow \infty, x_7(t) \leq \frac{\alpha_3}{\mu_1}.$$

For a clean environment, (5.11) becomes

$$\begin{aligned}\frac{dx_7}{dt} &= \frac{\alpha_3(x_3 + x_4)}{N} - \mu_1 x_7 \\ &\geq -\mu_1 x_7 \\ x_7(t) &= x_0 e^{-\mu_1 t} \geq 0.\end{aligned}$$

Hence, we have

$$0 \leq x_7(t) \leq \frac{\alpha_3}{\mu_1}.$$

From equation (5.4), we have

$$\begin{aligned}\frac{dx_1}{dt} &= \pi - \frac{(\beta_1 x_3 + \beta_2 x_4)x_1}{N} - \beta_3 x_7 x_1 \\ &\quad - \mu x_1 \\ &\geq \pi - \left(\beta_1 + \beta_2 + \beta_3 \frac{\alpha_3}{\mu_1} \right) x_1 \\ &\geq -w x_1, \quad w = \left(\beta_1 + \beta_2 + \beta_3 \frac{\alpha_3}{\mu_1} \right).\end{aligned}$$

The solution is

$$x_1(t) \geq x_1^0 e^{-wt} \geq 0, \forall t \geq 0.$$

Similarly, we can show that for $i = 2, \dots, 6$

$$x_i(t) \geq 0, i = 2, \dots, 6.$$

This completes the proof.

5.5.2 Invariance

The total population, $N(t)$, at time t is given by

$$\begin{aligned} \frac{dN}{dt} &= \pi - \mu N - \delta_{x_4} x_4 - \delta_{x_5} x_5 \\ &\leq \pi - \mu N. \end{aligned}$$

By Gronwall inequality, it is easy to show that

$$0 \leq N(t) \leq \frac{\pi}{\mu}.$$

For the existence of a unique bounded solution, we infer that any solution of the system (5.4)-(5.11) is non negative and bounded in $\Omega = \left\{ (x_1, x_2, x_3, x_4, x_5, x_6) \in \mathfrak{R}_+^6 : N \leq \frac{\pi}{\mu}, x_7 \leq \frac{\alpha_3}{\mu_1} \right\}$.

5.5.3 Disease Free Equilibrium (DFE)

5.5.3.1 Disease Free Equilibrium for $\pi = 0$

During a strict lockdown, we can take $\pi = 0$. The system (5.4) - (5.11) has two equilibrium points, the disease eradication point $\zeta_0 = (0, 0, 0, 0, 0, 0, 0)$ and the disease endemic point $\zeta_1 = (x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, x_6^*, x_7^*)$. To establish the stability of ζ_0 , we use the Hartmann-Grobmann theorem which roughly states that near an equilibrium point, the dynamics of the original (nonlinear) system are the same as those for the linearized system. The linearized system for (5.4)-(5.11) is given by:

$$\frac{dy(t)}{dt} = Ay(t),$$

where

$$y(t) = (x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t))^T,$$

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\alpha_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha_1 & -b_{11} & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_2 & -c_{11} & 0 & 0 & 0 \\ 0 & 0 & \gamma_1 & \gamma_2 & -d_{11} & 0 & 0 \\ 0 & 0 & \kappa_1 & \kappa_2 & \kappa_3 & 0 & 0 \\ 0 & 0 & \alpha_3 & \alpha_3 & 0 & 0 & -\mu_1 \end{bmatrix}$$

and

$$b_{11} = (\gamma_1 + \alpha_2 + \kappa_1)$$

$$c_{11} = (\gamma_2 + \kappa_2 + \delta_1)$$

$$d_{11} = (\delta_2 + \kappa_3)$$

The eigenvalues of $|A - \lambda I|$ are given by $(0, 0, -\alpha_1, -b_{11}, -c_{11}, -d_{11}, -\mu_1)$. The reproduction number R_0 , is given by the largest spectral radius of the matrix $|A - \lambda I|$. In this instance the possible values of R_0 are $\max\{\alpha_1, b_{11}, c_{11}, d_{11}, \mu_1\}$. Since $\lambda_i < 0, i = 1, 2, \dots, 7$, the system (5.4) - (5.11) is stable and tends to ζ_0 as $t \rightarrow \infty$. Provided the population maintains the covid control protocols and keeps R_0 below 1, the disease will fail to establish in the population.

5.5.3.2 Disease Free Equilibrium for $\pi \neq 0$

When inter-zone movements are allowed, that is, $\pi \neq 0$, the system (5.4) - (5.11) has a disease free equilibrium point $\zeta_0 = (\frac{\pi}{\mu}, 0, 0, 0, 0, 0, 0)$.

We use the technique by Van den Driessche and Watmough [80] to find the model reproduction number. The matrix for new infections is given by F whilst the matrix for other transitions is given by V , where

$$\mathbf{F} = \begin{bmatrix} \frac{(\beta_1 x_3 + \beta_2 x_4) x_1}{N} + \beta_3 x_7 x_1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} a x_2 \\ b x_3 - \alpha_1 x_2 \\ c x_4 - \alpha_2 x_3 \\ \mu_1 x_7 - \frac{\alpha_3 (x_3 + x_4)}{N} \end{bmatrix}$$

and

$$a = (\alpha + \mu)$$

$$b = (\gamma_1 + \alpha_2 + \kappa_1 + \mu)$$

$$c = (\gamma_2 + \kappa_2 + \delta_1).$$

The Jacobian matrices \mathbf{F} of F and \mathbf{V} of V are computed at the point ζ_0 with respect to the infected classes (x_2, x_3, x_4, x_7) . Thus, the basic reproduction number R_0 , given by the spectral radius of the matrix \mathbf{FV}^{-1} is the maximum of the moduli of the eigenvalues of that matrix \mathbf{FV}^{-1} given by,

$$R_0 = R_{x_3} + R_{x_4} + R_{x_7},$$

where,

$$\begin{aligned} R_{x_3} &= \frac{\alpha_1 \beta_1}{a b} \\ R_{x_4} &= \frac{\alpha_1 \alpha_2 \beta_2}{a b c} \\ R_{x_7} &= \frac{\alpha_1 \alpha_3 \beta_3 (c + \alpha_2)}{a b c \mu_1}. \end{aligned}$$

Based on R_0 we can state the following theorem:

Theorem 5.2. *The DFE point, ζ_0 , is locally asymptotically stable if $R_0 < 1$ and unstable if $R_0 > 1$.*

Remark 1: R_0 can be greater than 1 even if $R_{x_3} < 1$, $R_{x_4} < 1$, $R_{x_7} < 1$. It is necessary for each sub-reproduction number R_{x_3} , R_{x_4} and R_{x_7} to be less than 1 and their sum to be less than 1 for the disease to fail to establish in the population.

Remark 2: Unlike the case $\pi = 0$, now the adherence to the covid protocols

by each infected sub-population, x_i , $i = 3, 4$, is more strict (adherence to social distancing, mask wearing etc) and the hygiene measures must be more strictly observed for the stability of ζ_0 to be achieved.

5.5.3.3 Global stability of the Disease Free Equilibrium

Theorem 5.3. *The DFE point, ζ_0 , of system (5.4) - (5.11) is globally asymptotically stable for $R_0 < 1$.*

Proof. To prove theorem 6.3, we use the Kamgang-Sallet stability theorem, [87]. Let $Z = (Z_1, X_2)$ with $Z_1 = (x_1, x_6) \in \mathfrak{R}^2$ and $Z_2 = (x_2, x_3, x_4, x_5, x_7) \in \mathfrak{R}^5$. In terms of Z_1 and Z_2 system (5.4) - (5.11) can be written as:

$$\begin{aligned}\dot{Z}_1 &= A_1(Z)(Z_1 - Z_1^*) + A_{12}(Z)Z_2 \\ \dot{Z}_2 &= A_2(Z)(Z_2)\end{aligned}$$

where $Z_1^* = (\frac{\pi}{\mu}, 0)$, with

$$A_1(Z) = \begin{bmatrix} -\mu & 0 \\ 0 & -\mu \end{bmatrix}, \quad A_{12}(Z) = \begin{bmatrix} 0 & -\frac{\beta_1 x_1}{N} & -\frac{\beta_2 x_1}{N} & 0 & -\beta_3 x_1 \\ 0 & \kappa_1 & \kappa_2 & \kappa_3 & 0 \end{bmatrix}$$

and

$$A_2(Z) = \begin{bmatrix} -a & \frac{\beta_1 x_1}{N} & \frac{\beta_2 x_1}{N} & 0 & \beta_3 x_1 \\ \alpha_1 & -b & 0 & 0 & 0 \\ 0 & \alpha_2 & -c & 0 & 0 \\ 0 & \gamma_1 & \gamma_2 & -d & 0 \\ 0 & \frac{\alpha_3}{N} & \frac{\alpha_3}{N} & 0 & -\mu_1 \end{bmatrix},$$

where $d = (\delta_2 + \kappa_3)$. We want to show that the five sufficient conditions of Kamgang-Sallet Theorem in [87] are satisfied as follows:

- (i) The system (5.4) - (5.11) is a dynamical system on Ω , as defined and shown in section V.
- (ii) The eigenvalues of $A_1(Z)$ are real and negative, thus the system $\dot{Z}_1 =$

$A_1(Z)(Z_1 - Z_1^*) + A_{12}(Z)Z_2$ is globally asymptotically stable at the equilibrium Z_1^* .

(iii) The matrix $A_2(Z)$ is a Metzler matrix, i.e. a matrix such that off diagonal terms are non negative and is irreducible for any given $Z \in \Omega$.

(iv) There exists a matrix \bar{A}_2 , which is an upper bound for the set $M = A_2(Z) : Z \in \Omega$. Indeed,

$$\bar{A}_2 = \begin{bmatrix} -a & \beta_1 & \beta_2 & 0 & \beta_3 x_1^* \\ \alpha_1 & -b & 0 & 0 & 0 \\ 0 & \alpha_2 & -c & 0 & 0 \\ 0 & \gamma_1 & \gamma_2 & -d & 0 \\ 0 & \frac{\alpha_3}{N^*} & \frac{\alpha_3}{N^*} & 0 & -\mu_1 \end{bmatrix}$$

is an upper bound for M .

(v) For $R_0 < 1$, λ is the eigenvalue of \bar{A}_2 ,

$$\alpha(\bar{A}_2) = \max \{ \text{Re}(\lambda) : \lambda \in \sigma(\bar{A}_2) \}$$

To check condition (v) we will use the following lemma which is a characterization of Metzler stable matrices:

Lemma 5.4. *Let M be a square matrix written in block form*

$$D = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

with A and D being square matrices. M is Metzler stable if and only if matrices A and $D - CA^{-1}B$ are Metzler stable.

Using Lemma 6.4, matrix \bar{A}_2 can be expressed in the form of matrix M with:

$$A = \begin{bmatrix} -a & \beta_1 & \beta_2 \\ \alpha_1 & -b & 0 \\ 0 & \alpha_2 & -c \end{bmatrix}, \quad B = \begin{bmatrix} 0 & \beta_3 x_1^* \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & \gamma_1 & \gamma_2 \\ 0 & \frac{\alpha_3}{N^*} & \frac{\alpha_3}{N^*} \end{bmatrix}, \quad D = \begin{bmatrix} -d & 0 \\ 0 & -\mu_1 \end{bmatrix}.$$

Clearly A , is a stable Metzler matrix and after computations we obtain $D - CA^{-1}B$ is a stable Metzler matrix if and only if

$$\begin{aligned} R_0^{max} &= \frac{\alpha_1 \beta_1}{ab} + \frac{\alpha_1 \alpha_2 \beta_2}{abc} \\ &\quad + \frac{\alpha_1 \alpha_3 \beta_3 (c + \alpha_2)}{abc \mu_1} \\ &\leq 1. \end{aligned}$$

5.5.4 Endemic Equilibrium Point (EEP)

Solving for the system of equations (5.4) - (5.11) by equating the RHS to zero, we find the coordinates of the EEP given by $\zeta_1 = (x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, x_6^*, x_7^*)$, where

$$x_1^* = \frac{abcN^*\mu_1}{\alpha_1 a_{33}} \quad (5.12)$$

$$x_2^* = \frac{a_{22}(R_0 - 1)}{a\alpha_1 a_{33}} \quad (5.13)$$

$$x_3^* = \frac{a_{22}(R_0 - 1)}{ab a_{33}} \quad (5.14)$$

$$x_4^* = \frac{\alpha_2 a_{22}(R_0 - 1)}{abc a_{33}} \quad (5.15)$$

$$x_5^* = \frac{a_{22} a_{44}(R_0 - 1)}{abcd a_{33}} \quad (5.16)$$

$$x_6^* = \frac{\kappa_3 a_{22} a_{44} a_{55}(R_0 - 1)}{abcd \mu a_{33}} \quad (5.17)$$

$$x_7^* = \frac{\alpha_3(\alpha_2 + c)a_{22}(R_0 - 1)}{abcN^*\mu_1 a_{33}} \quad (5.18)$$

and

$$a_{22} = abcN^*\mu\mu_1$$

$$a_{33} = (c + \alpha_2)\alpha_3\beta_3 + (c\beta_1 + \alpha_2\beta_2)\mu_1$$

$$a_{44} = (\alpha_2\gamma_2 + c\gamma_1)$$

$$a_{55} = (cd\kappa_1 + d\kappa_2\alpha_2).$$

The coordinates (5.13) - (5.18) exist if and only if $R_0 > 1$.

5.5.4.1 Global stability of the Endemic Equilibrium Point

The global stability of the EEP is explored by proving the following theorem:

Theorem 5.5. *If $R_0 > 1$ then the EEP given by ζ_1 is globally asymptotically stable in the region Ω .*

Proof. Following the work of [47], we construct a Lyapunov function L of the type:

$$L(x_i) = \sum_{i=1}^7 \left(x_i - x_i^* - x_i^* \ln \frac{x_i}{x_i^*} \right).$$

Differentiating L with respect to x_i gives:

$$\frac{dL}{dt} = \sum_{i=1}^7 \left(\frac{x_i - x_i^*}{x_i} \right) \frac{dx_i}{dt}.$$

Substituting for $\frac{dx_i}{dt}$, $i = 1, \dots, 7$, we get

$$\begin{aligned} \frac{dL}{dt} &= \left(\frac{x_1 - x_1^*}{x_1} \right) \left[\pi - \frac{(\beta_1 x_3 + \beta_2 x_4)x_1}{N} - \beta_3 x_7 x_1 - \mu x_1 \right] \\ &+ \left(\frac{x_2 - x_2^*}{x_2} \right) \left[\frac{(\beta_1 x_3 + \beta_2 x_4)x_1}{N} + \beta_3 x_7 x_1 - a x_2 \right] \\ &+ \left(\frac{x_3 - x_3^*}{x_3} \right) [\alpha_1 x_2 - b x_3] + \left(\frac{x_4 - x_4^*}{x_4} \right) [\alpha_2 x_3 - c x_4] \\ &+ \left(\frac{x_5 - x_5^*}{x_5} \right) [\gamma_1 x_3 + \gamma_2 x_4 - d x_5] + \left(\frac{x_6 - x_6^*}{x_6} \right) \\ &[\kappa_1 x_3 + \kappa_2 x_4 + \kappa_3 x_5 - \mu x_6] \\ &+ \left(\frac{x_7 - x_7^*}{x_7} \right) \left[\frac{\alpha_3(x_3 + x_4)}{N} - \mu_1 x_7 \right] \\ &= A_{33} - A_{22}, \end{aligned}$$

where

$$\begin{aligned}
A_{22} = & \left(\frac{x_1 - x_1^*}{x_1} \right)^2 \left[\frac{(\beta_1 x_3 + \beta_2 x_4)}{N} + \beta_3 x_7 + \mu \right] \\
& + \left[\frac{x_1^*}{x_1} \pi + \frac{(\beta_1 x_3 + \beta_2 x_4) x_1^*}{N} + \beta_3 x_7 x_1^* + \mu x_1^* \right] \\
& + \left(\frac{x_2 - x_2^*}{x_2} \right)^2 a + a x_2^* \\
& + \frac{x_2^*}{x_2} \left[\frac{\beta_1 x_3 + \beta_2 x_4}{N} x_1 + \beta_3 x_7 x_1 \right] \\
& + \left(\frac{x_3 - x_3^*}{x_3} \right)^2 b + b x_3^* + \frac{x_3^*}{x_3} \alpha_1 x_2 \\
& + \left(\frac{x_4 - x_4^*}{x_4} \right)^2 c + c x_4^* + \frac{x_4^*}{x_4} \alpha_2 x_3 \\
& + \left(\frac{x_5 - x_5^*}{x_5} \right)^2 d + d x_5^* + \frac{x_5^*}{x_5} (\gamma_1 x_3 + \gamma_2 x_4) \\
& + \left(\frac{x_6 - x_6^*}{x_6} \right)^2 \mu + \mu x_6^* \\
& + \frac{x_6^*}{x_6} (\kappa_1 x_3 + \kappa_2 x_4 + \kappa_3 x_5) \\
& + \left(\frac{x_7 - x_7^*}{x_7} \right)^2 \mu_1 + \mu_1 x_7^* + \frac{x_7^*}{x_7} \frac{\alpha_3 (x_3 + x_4)}{N}
\end{aligned}$$

and

$$\begin{aligned}
A_{33} = & \pi + \frac{x_1^{*2}}{x_1} \left[\frac{(\beta_1 x_3 + \beta_2 x_4)}{N} + \beta_3 x_7 + \mu \right] \\
& + \left[\frac{(\beta_1 x_3 + \beta_2 x_4)}{N} + \beta_3 x_7 + \mu \right] \\
& + \frac{x_2^{*2}}{x_2} a + \alpha_1 x_2 + \frac{x_3^{*2}}{x_3} b + \alpha_2 x_3 + \frac{x_4^{*2}}{x_4} c \\
& + \gamma_1 x_3 + \gamma_2 x_4 + \frac{x_5^{*2}}{x_5} d + \kappa_1 x_3 + \kappa_2 x_4 \\
& + \kappa_3 x_5 + \frac{x_6^{*2}}{x_6} \mu + \frac{\alpha_3 (x_3 + x_4)}{N} + \frac{x_7^{*2}}{x_7} \mu_1.
\end{aligned}$$

Since all the parameters used in the system (5.4)-(5.11) are non negative we have $\frac{dL}{dt} \leq 0$ for $A_{33} \leq A_{22}$ and $\frac{dL}{dt} = 0$ if and only if $x_1 = x_1^*, x_2 = x_2^*, x_3 = x_3^*, x_4 = x_4^*, x_5 = x_5^*, x_6 = x_6^*, x_7 = x_7^*$. Thus by La-Salle's invariance principle, the EEP is globally asymptotically stable.

5.6 Numerical Simulations

We want to find optimal isolation rates γ_1 and γ_2 , using Matlab programs, under which the number of infected individuals will not overwhelm the hospital capacity, H . We performed numerical simulations using data from the Italian coronavirus epidemic from 31st January to the 15th of May 2020 to demonstrate that our model can accurately reproduce the recorded data on numbers of infected and dead individuals. The parameter values used for numerical simulations are given in Table 6.1.

Table 5.1: Numerical values for the parameters of the Italian case

Parameter	Value/range	Source
β_1	0.492	[48]
β_2	1.30	[50]
β_3	0	estimate
δ_1	0.015	[51]
δ_2	$\frac{4827}{53578}$	[123]
α_1	$\frac{1}{3.21}$	[50]
α_2	$\frac{1}{2.27}$	[50]
α_3	0	estimate
γ_1	[0, 0.6]	estimate
γ_2	[0, 0.6]	estimate
κ_1	$\frac{1}{6}$	[48]
κ_2	$\frac{1}{10}$	[48]
κ_3	$\frac{1}{14}$	[48]
μ	0.00003032	[124]
μ_1	0	estimate
π	$\mu \times 400\,000$	[49]

5.6.0.1 Sensitivity Analysis

Five parameters, β_3 , α_3 , γ_1 , γ_2 and μ_1 in Table 6.1 have been estimated based on the number of hospitalized cases in the Lombardy region of Northern Italy. We have analyzed how sensitive the reproduction number is to the changes in

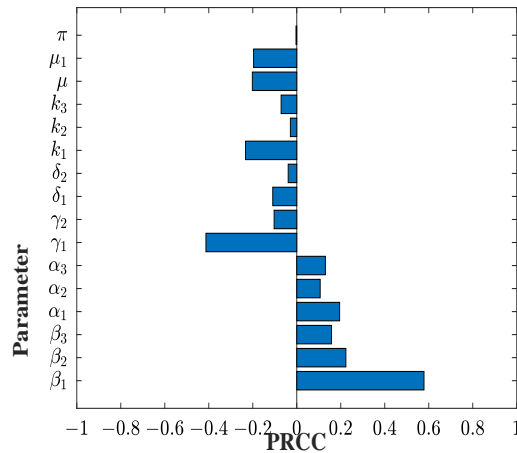


Figure 5.5: A diagram showing the sensitivity of R_0 to various model parameters.

these five parameters (Figure 5.5). The parameters α_3 and β_3 are important and provide insight into understanding how the sub-Saharan Africa situation will differ from the Italian or in general the situation in Western countries, where we assumed $\alpha_3 = 0$.

5.6.1 Italy

Lombardy is a region with a population of approximately 10 million. By an iterative procedure, since we know the number of people infected with the coronavirus from 31st January to 22nd March 2020, we have estimated that the number of people susceptible to infection by the virus, through their failure to observe prevention measures such as self quarantine, social distancing etc at the start of the coronavirus epidemic, was about 400 000. As at 31st January 2020, the number who were potentially exposed to the infection is estimated to be around 200 individuals. On the 31st January 2020, Italy recorded its first 3 cases of individuals infected with the coronavirus. Based on this initial data, we present examples of strategies by fixing the rate of isolating symptomatic infectives and then finding the corresponding rate of isolating asymptomatic infectives which ensures that the combined number of infectives does not exceed the hospital bed capacity and *vice versa*. The examples discussed here are not unique but are typical of other scenarios and provide insight into the following:

- (i) how the infection curve can be flattened to ensure that the hospital bed capacity is not exceeded.
- (ii) For non optimal cases, where hospital bed capacity is exceeded, to quantify the number of infectives for each non optimal case which must be safely isolated outside hospital facilities.

According to [125], Italy has 12.5 beds per 100 000 individuals of Intensive Care Unit (ICU) or critical care beds (CCB) beds and 3.18 beds per 1 000 individuals ordinary hospital beds. For Lombardy, this data (combining both ordinary and ICU hospital beds) gives 33 182 beds.

Figure 5.6 presents an example for a fixed rate $\gamma_2 = 0.33$ of isolating symptomatic infectives. We find that the optimal rate of securing asymptomatic infectives when there is a lockdown, $\pi = 0$, should be $\gamma_1 = 0.46$. Converting these rates to time, we see that asymptomatic infectives should be identified, hospitalized in a time almost 1.5 times faster than the time of hospitalizing symptomatic infectives. Currently, every nation is experiencing a lack of materials and equipment to conduct tests. This rate of hospitalizing asymptomatic infectives would be difficult to achieve as there has been a shortage of testing materials, and most governments have decided they would only test individuals who present with COVID-19 symptoms. Moreover, the time it is taking to obtain test results 2 – 3 days is slowing down the testing significantly.

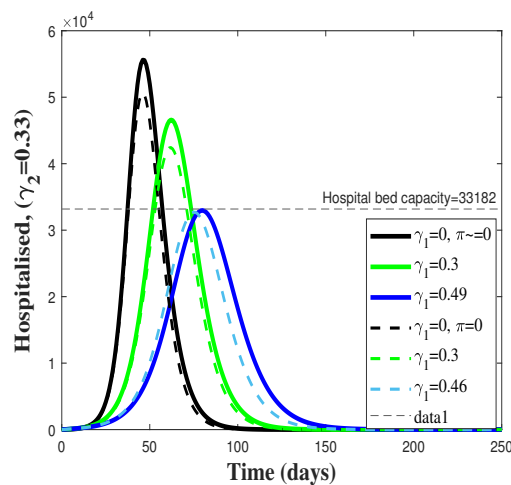


Figure 5.6: Population level plots for hospitalized individuals for varying values of γ_1

For this strategy, we have determined the number of infected people that should be safely secured for varying values of γ_1 at peak infection. For $\gamma_1 = 0$, the number in excess of hospital bed capacity that must be secured is 22 408, for $\gamma_1 = 0.3$, the number in excess of hospital bed capacity that must be secured is 13 407. For the non optimal cases, if the hospital overflow bed capacity is increased by 50% (as was done in most Western countries) the only case that would have accommodated the overflow is $(\gamma_1, \gamma_2) = (0.3, 0.33)$. Figure 5.6 shows that when inter zonal movement is allowed, $\pi \neq 0$, the number of people needing hospitalization or to be safely isolated increases. The optimum removal rates with inter-zonal movement are given by $(\gamma_1^*, \gamma_2^*) = (0.49, 0.33)$, implying that asymptomatic infectives must be isolated in a time 1.6 times faster than when there is no inter zonal movement. The optimal solutions in Figure 5.6 show that the peak hospitalization capacity occurs earlier when there is a lockdown ($\pi = 0$) than when there is inter-zonal movement ($\pi \neq 0$).

Figure 5.7 presents a strategy where the rate of hospitalizing symptomatic infectives is fixed at $\gamma_2 = 0.6$. The optimal rate for γ_1 which flattens the hospital admission curve is found to be $\gamma_1 = 0.16$. In other words, almost one third of the effort must be devoted to testing, contact tracing and hospitalizing asymptomatic infectives.

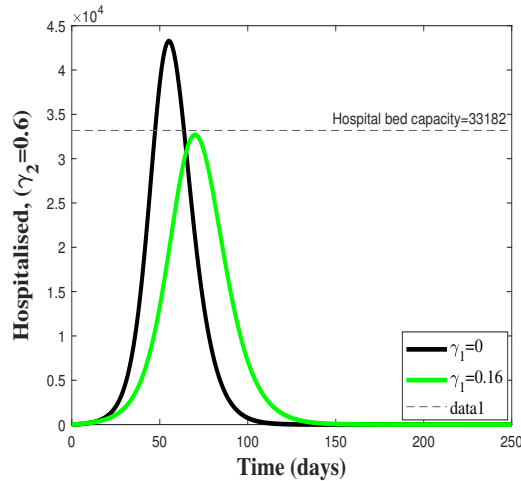


Figure 5.7: Population level plots for hospitalized individuals for varying values of γ_1

For this strategy the situation regarding the non optimal cases is as follows: for $\gamma_1 = 0$ the number needed to be secured is 10 097, which is too high to accom-

modate.

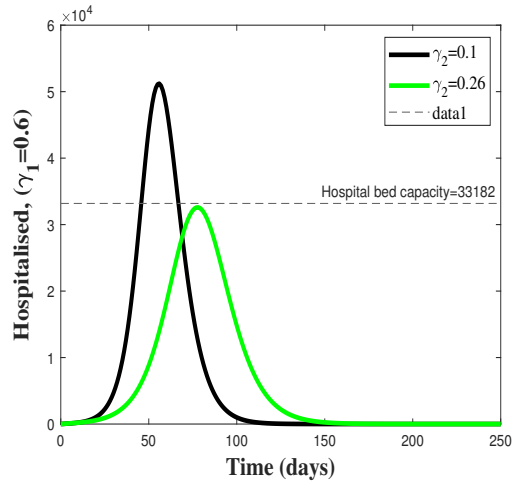


Figure 5.8: Population level plots for hospitalized individuals for varying values of γ_2

Using optimization techniques to find γ_1^* and γ_2^* we have found and demonstrated in Figure 5.8 that the optimum rates are $(\gamma_1^*, \gamma_2^*) = (0.6, 0.26)$. This optimum pair implies that the effort devoted to testing, contact tracing and isolating symptomatic infectives must be 2.3 times higher than that of isolating asymptomatic infectives.

Figure 5.9 shows that for $\gamma_1 = 0.3$ the optimum pair is $(\gamma_1^*, \gamma_2^*) = (0.3, 0.45)$. This

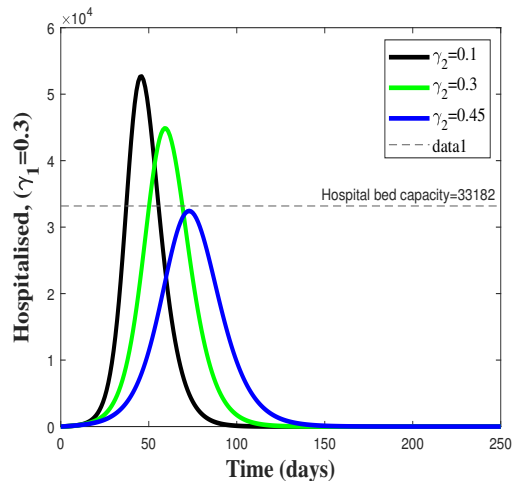


Figure 5.9: Population level plots for hospitalized individuals for varying values of γ_2

strategy targets to remove symptomatic infectives faster than the asymptomatic

infectives. This strategy removes symptomatic infectives in a time at least one and half times faster than the time of isolating asymptomatic infectives.

To illustrate the impact of non adherence to prevention measures, such as social distancing, not wearing masks etc, on the optimal case given in Figure 5.9, we considered how the case $\beta_3 \neq 0$ for Italy would have altered the conclusions presented in Figures 5.6 to 5.9. Figure 5.10 shows that if we vary the parameter α_3 the hospital bed capacity for Lombardy in Italy would have been exceeded for any $\alpha_3 \geq 0.01$. This suggests that high standards of hygiene are key to controlling COVID-19 infections.

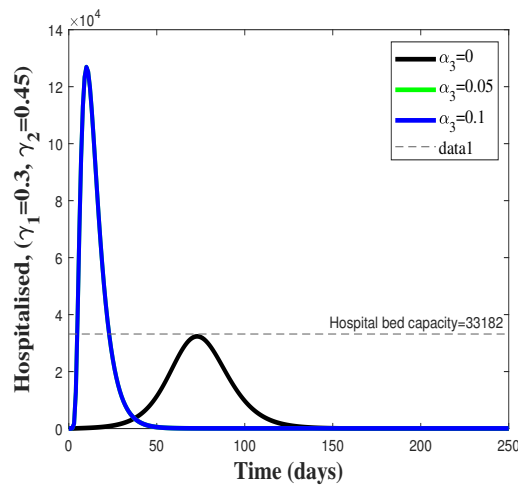


Figure 5.10: Population level plots for hospitalized individuals for varying values of α_3

We can see from the examples of the four strategies illustrated in Figures 5.6, 5.7, 5.8, 5.9 that there is no unique way of flattening the curve in order to protect the hospital bed capacity. The question we address is how does one choose the best strategy among the optimal strategies? Table 7.1 gives the number of deaths resulting from the four strategies above and compares each strategy with the actual number of recorded deaths. We conclude that the best strategy is one which reduces the number of deaths. In this case (Table II) any strategy that isolates the symptomatic infectives faster is preferable as it is more economical.

Table 5.2: Comparison of deaths at different optimal cases

Strategy	γ_1	γ_2	30 days $\delta = 0.01$	60 days $\delta = 0.03$	90 days $\delta = 0.02$	105 days $\delta = 0.015$ Lockdown	105 days $\delta = 0.09$ No lockdown
1	0.49	0.33	210 (29)	10 126 (11 591)	32 108 (27 682)	31 433 (31 368)	120 200
2	0.16	0.6	226 (29)	12 861 (11 591)	33 994 (27 682)	31 202 (31 368)	115 690
3	0.6	0.26	200 (29)	8 989 (11 591)	30 734 (27 682)	31 162 (31 368)	120 640
4	0.3	0.45	217 (29)	11 179 (11 591)	32 771 (27 682)	31 135 (31 368)	117 600

5.6.2 Optimal Control

5.6.2.1 Optimal control without incorporating household and community exposure

To study measures that reduce disease transmission, such as lockdown or isolation of infectives, we introduced two controls, u_1 and u_2 , on the infection coefficients β_1 and β_2 as shown below:

$$\frac{dx_1}{dt} = \pi - (1 - u_1) \frac{\beta_1 x_3 x_1}{N} - (1 - u_2) \frac{\beta_2 x_4 x_1}{N} - \beta_3 x_7 x_1 - \mu x_1 \quad (5.19)$$

$$\frac{dx_2}{dt} = (1 - u_1) \frac{\beta_1 x_3 x_1}{N} + (1 - u_2) \frac{\beta_2 x_4 x_1}{N} + \beta_3 x_7 x_1 - (\alpha_1 + \mu) x_2 \quad (5.20)$$

$$\frac{dx_3}{dt} = \alpha_1 x_2 - (\gamma_1 + \alpha_2 + \mu + \kappa_1) x_3 \quad (5.21)$$

$$\frac{dx_4}{dt} = \alpha_2 x_3 - (\gamma_2 + \delta_1 + \kappa_2) x_4 \quad (5.22)$$

$$\frac{dx_5}{dt} = \gamma_1 x_3 + \gamma_2 x_4 - (\delta_2 + \kappa_3) x_5 \quad (5.23)$$

$$\frac{dx_6}{dt} = \kappa_1 x_3 + \kappa_2 x_4 + \kappa_3 x_5 - \mu x_6 \quad (5.24)$$

and

$$\frac{dx_7}{dt} = \frac{\alpha_3 (x_3 + x_4)}{N} - \mu_1 x_7 \quad (5.25)$$

We want to minimize the objective functional given by

$$J(u_1, u_2) = \int_0^{t_T} A_1 x_2 + A_2 x_3 + A_3 x_4 + \frac{1}{2} A_4 u_1^2 + \frac{1}{2} A_5 u_2^2. \quad (5.26)$$

The goal is to find a set of controls that minimize the number of susceptible individuals who come into contact with infected individuals. Let u_1^* , u_2^* be the optimal controls. The problem is to find

$$J(u_1^*, u_2^*) = \min J(u_1, u_2), (u_1, u_2) \in U, \quad (5.27)$$

subject to system (5.19) - (5.25), where U is the set of measure functions defined from $[0, t_T]$ to $[0, 1]$.

The optimality conditions are given by,

$$u_1^* = \min \left\{ \max \left[0, (\lambda_{x_2} - \lambda_{x_1}) \left(\frac{\beta_1 x_3 x_1}{A_4 N} \right) \right], u_1 \max \right\},$$

$$u_2^* = \min \left\{ \max \left[0, (\lambda_{x_2} - \lambda_{x_1}) \left(\frac{\beta_2 x_4 x_1}{A_5 N} \right) \right], v_2 \max \right\},$$

Calculation for u_1^* and u_2^* is based on Pontryagin's Maximum Principle (see [94]) for a detailed description.

5.6.2.2 Optimal control incorporating household and community exposure.

To study the impact that household and community exposure has on disease transmission control measures, we modify equations (5.19) - (5.25) by adding an additional term λ given in (3) which captures household and community exposure of an individual. The modified equations are given by:

$$\begin{aligned}\frac{dx_1}{dt} &= \pi - (1 - u_1 + \lambda) \frac{\beta_1 x_3 x_1}{N} - (1 - u_2 + \lambda) \frac{\beta_2 x_4 x_1}{N} - \beta_3 x_7 x_1 \\ &\quad - \mu x_1\end{aligned}\tag{5.28}$$

$$\begin{aligned}\frac{dx_2}{dt} &= (1 - u_1 + \lambda) \frac{\beta_1 x_3 x_1}{N} + (1 - u_2 + \lambda) \frac{\beta_2 x_4 x_1}{N} + \beta_3 x_7 x_1 \\ &\quad - (\alpha_1 + \mu) x_2\end{aligned}\tag{5.29}$$

$$\frac{dx_3}{dt} = \alpha_1 x_2 - (\gamma_1 + \alpha_2 + \mu + \kappa_1) x_3\tag{5.30}$$

$$\frac{dx_4}{dt} = \alpha_2 x_3 - (\gamma_2 + \delta_1 + \kappa_2) x_4\tag{5.31}$$

$$\frac{dx_5}{dt} = \gamma_1 x_3 + \gamma_2 x_4 - (\delta_2 + \kappa_3) x_5\tag{5.32}$$

$$\frac{dx_6}{dt} = \kappa_1 x_3 + \kappa_2 x_4 + \kappa_3 x_5 - \mu x_6\tag{5.33}$$

and

$$\frac{dx_7}{dt} = \frac{\alpha_3 (x_3 + x_4)}{N} - \mu_1 x_7\tag{5.34}$$

where λ is given by (3).

The optimality conditions that we want to satisfy are the same as those for equations (5.19) - (5.25). We consider the non optimal strategy in Figure 5.6 using equations (5.19) - (5.25) and equations (5.28) - (5.34) for $\gamma_1 = 0.3$ and $\gamma_2 = 0.33$ where $u_1 = 0$ and $u_2 = 0$. If we use the controls $u_1^* \geq 0.2$, $u_2^* \geq 0.2$ and $(\gamma_1, \gamma_2) = (0.3, 0.33)$ we obtain Figures 5.11(a) - 5.11(e). It is clear from Figure 5.11(e) that household and community disease transmission would increase the number hospitalized though the numbers would still be below the hospital bed capacity for optimal values of γ_1 and γ_2 . The number of susceptible individuals (Figure 5.11(a)) would decline due to increased infection rates.

5.6.3 Sub-Saharan Africa: South Africa as an example.

To provide insight into the sub-Saharan Africa outlook, simulations were done using data from South Africa, one of the countries in this region with the highest number of coronavirus cases and with the best medical facilities. It is our view that if South Africa cannot cope then most, if not all, countries in sub-Saharan

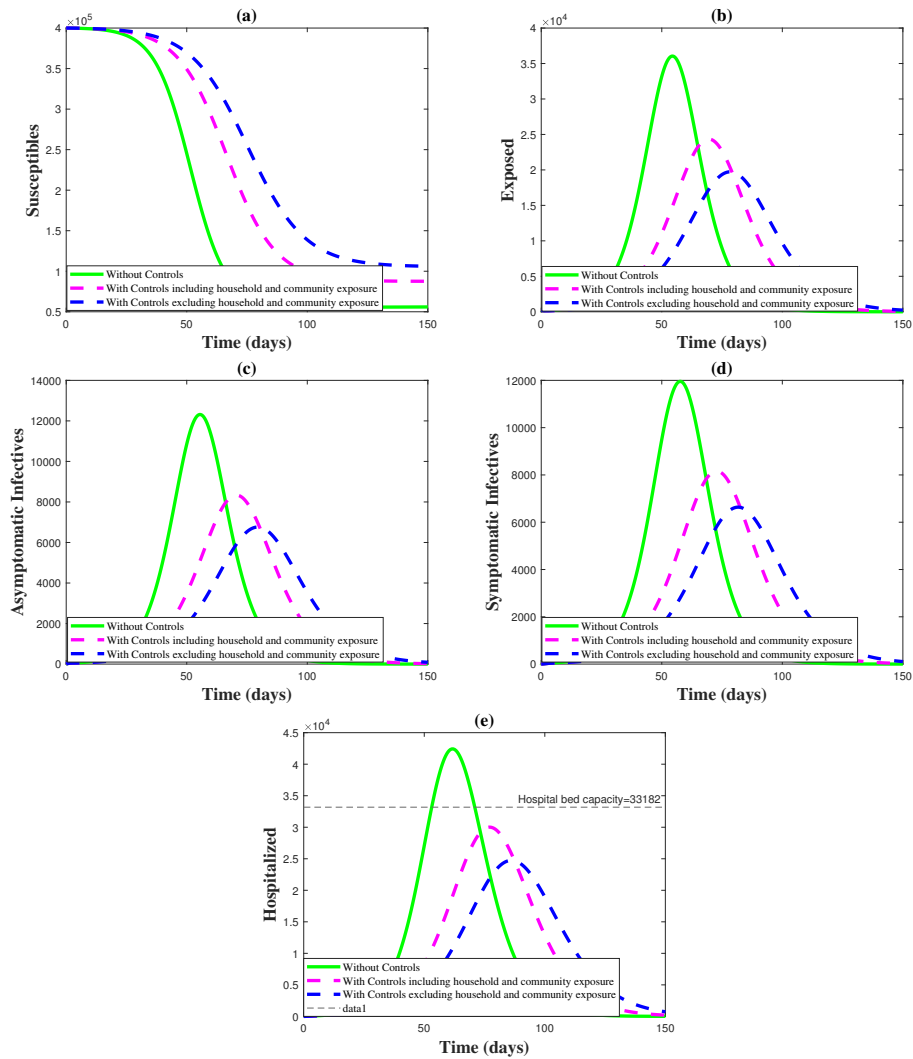


Figure 5.11: A comparison of state variables with and without controls

Africa would not cope. On the 4th of March 2020, South Africa recorded its first case of the coronavirus. As of the 1st of April 2020 the highest case counts of coronavirus had been reported in 3 provinces namely Gauteng, KwaZulu Natal and the Western Cape. The 3 provinces had 1 380 reported coronavirus cases distributed as follows: 645 in Gauteng, 326 in KwaZulu Natal and 186 in the Western Cape. These 3 provinces have a total population of 33 309 473 people. The number of people infected with the coronavirus from the 4th of March to the 1st April 2020 is known, we have estimated that the number of people susceptible to infection by the virus through failure to self quarantine, self isolate, observe social distancing etc was about 1 200 000. At that time the number who were potentially exposed to the infection is estimated to be about 1 000 individuals.

Except for β_3 , π , α_3 , μ , μ_1 and δ_2 , we use the parameters in Table 6.1 for Italy.

Table 5.3: Numerical values for the parameters of the South African case

Parameter	Value/range	Source
β_1	0.492	[48]
β_2	1.30	[50]
β_3	0.001006082	[49]
δ_1	0.015	[51]
δ_2	$\frac{5}{1380}$	[123]
α_1	$\frac{1}{3.21}$	[50]
α_2	$\frac{1}{2.27}$	[50]
α_3	5.3346×10^{-6}	estimate
γ_1	[0, 0.77]	estimate
γ_2	[0, 0.33]	estimate
κ_1	$\frac{1}{6}$	[48]
κ_2	$\frac{1}{10}$	[48]
κ_3	$\frac{1}{14}$	[48]
μ	0.00004290	[126]
μ_1	0.00274	[49]
π	$\mu \times 1\,200\,000$	[49]

This is justified on the basis that Italian family bonds are similar to sub-Saharan Africa. The values used for this simulation are given in Table 5.3. According to [127], Gauteng province, KwaZulu Natal Province and the Western Cape province combined have a hospital bed capacity of 62 787. This number includes both ICU and ordinary hospital beds.

Figure 5.12 compares real time epidemic curves for Italy, Spain, the United Kingdom and South Africa for the first 40 days of the epidemic. Each country implemented the lockdown strategy at different stages of infection. Italy introduced the lockdown on the 9th of March 2020 when the total number of individuals infected with COVID-19 was 9 172. Spain introduced the lockdown on the 15th of March 2020 when the total number of those infected was 7 798. The United Kingdom introduced the lockdown on the 23rd of March 2020 when the number of individuals infected was 6 650. South Africa introduced the lockdown

on the 27th of March 2020 when the number infected was only 1 170. Figure 5.12 shows different infection trends for the four countries.

It is obvious that South Africa which introduced the lockdown early enough displays an epidemic which rises at a gentle rate. We consider the problem of finding the hospitalization rates γ_1 and γ_2 for which the hospital bed capacity of 62 787 would suffice.

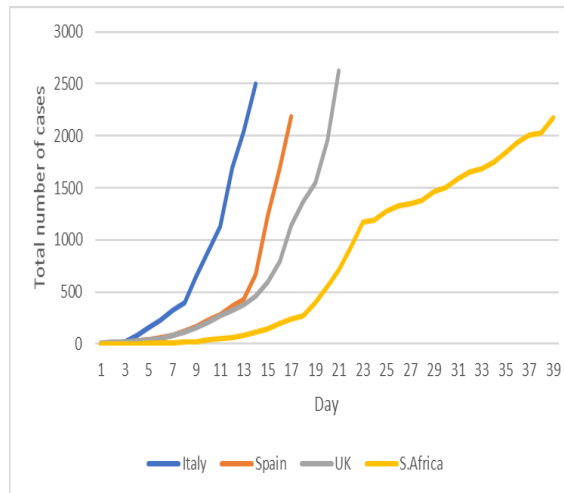


Figure 5.12: Total covid cases, UK, Spain, Italy, South Africa

5.6.3.1 Scenario with no lockdown measures

Figure 5.13 shows that for a fixed rate $\gamma_2 = 0.33$ of symptomatic infectives, the least rate of isolating asymptomatic infectives should be $\gamma_1 = 0.77$ for $\phi = 0$ (implying that the population must observe all the prevention measures ,social distancing, mask wearing etc) but that for $\phi \neq 0$ the hospital bed capacity is never sufficient.

Table 5.4: Number of people who must be hospitalized.

γ_1	γ_2	H	$H \times 1.5$	H peak (H_P)	Overflow
0	0.33	62 787	94 180	246 174	151 994
0.3	0.33	62 787	94 180	207 654	113 474
0.6	0.33	62 787	94 180	112 952	18 772

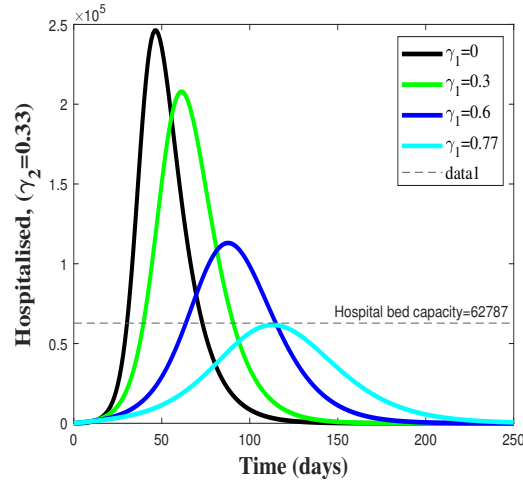


Figure 5.13: Population level plots for hospitalized individuals for varying values of γ_1

For the strategies in Figure 5.13, the numbers of people that must be isolated are given in Table 5.4.

Table 5.4 presents cases when the hospital bed capacity, denoted by H , is increased by 50% as was done in Italy, Spain and the State of New York, USA. The overflow is given by $(H_P - H \times 1.5)$. We can see that no pair of isolating infectives γ_1 and γ_2 would accommodate the number of infected individuals in the 3 provinces of South Africa within hospital facilities. The overflow capacity would be too large for any non-optimal pair γ_1 and γ_2 and $\phi \neq 0$, to safely isolate the overflow infected population. The only feasible solution is hospitalizing at the optimal rates $(\gamma_1^*, \gamma_2^*) = (0.77, 0.33)$ with $\phi = 0$. This option requires isolating asymptomatic infectives faster than symptomatic infectives, a strategy which requires perfect hygiene and sanitary measures. This requirement cannot be met in sub-Saharan Africa given the state of the economies.

To illustrate the impact of non adherence to disease prevention measures on the optimal case in Figure 5.13, we varied the parameter α_3 . Figure 5.14 shows how the optimal solution in Figure 5.13 is altered by varying α_3 . The hospital

bed capacity is exceeded for any case $\alpha_3 \neq 0$. For the case $\phi \neq 0$, no matter how small ϕ is, no optimal strategy exists for the three provinces of South Africa. It is unrealistic to expect perfect adherence to social distancing, wearing of masks and lockdown measures, in a region where the population survives on less than \$1 per day and there are no food banks.

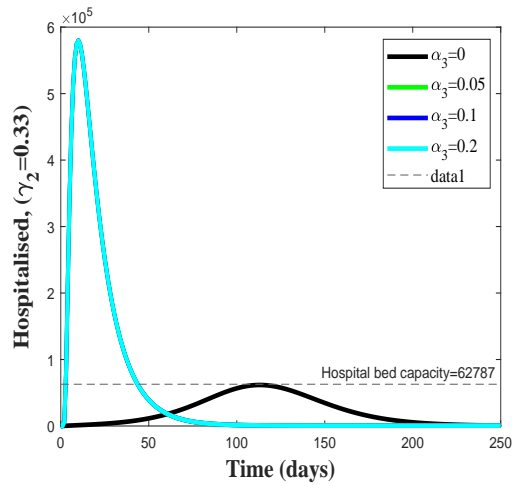


Figure 5.14: Population level plots for hospitalized individuals for varying values of α_3

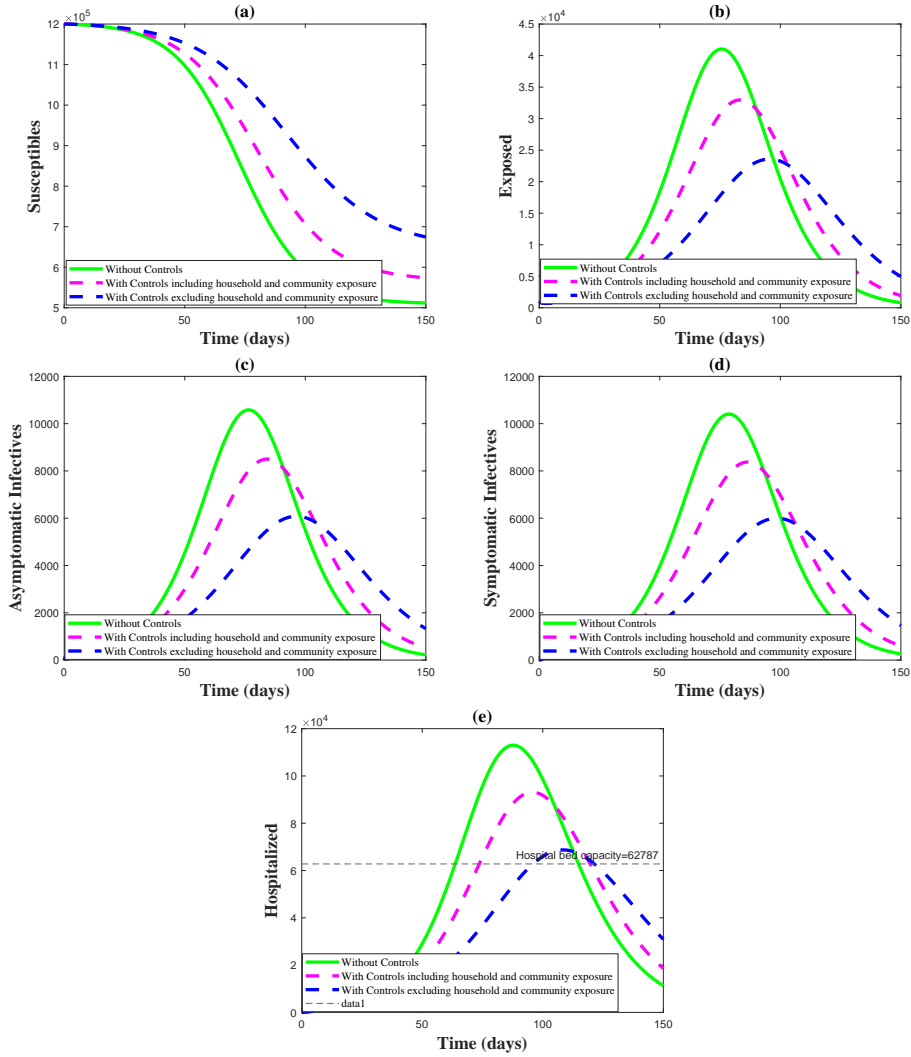


Figure 5.15: A comparison of state variables with and without controls

5.6.3.2 Effect of early lockdown

South Africa introduced the lockdown very early. The strategy discussed in section 5.6.3.1 is therefore not relevant. Hence, we consider the modified system with control given in (5.19) - (5.25) and (5.28) - (5.34). We consider hypothetically how the lockdown could have altered a non optimal strategy in Figure 5.12. We chose $\gamma_1 = 0.6$ and $\gamma_2 = 0.33$ ($u_1^* \geq 0.1$, $u_2^* \geq 0.1$) to illustrate this example (see Figures 5.15(a) - 5.15(e)).

The number of deaths for a model with controls for South Africa for different optimal cases of γ_1 and γ_2 is given in Table 5.5. Table 5.5 compares the actual number of recorded deaths, the controlled number of deaths and the uncontrolled number of deaths (no lockdown). It is clear that the lockdown was

Table 5.5: Comparison of deaths at different optimal cases

γ_1	γ_2		30 days $\delta = 0.003$	60 days $\delta = 0.01$	72 days $\delta = 0.01$	90 days $\delta = 0.01$
0.77	0.33	no lockdown	178	3 358	5 930	12 394
		lockdown	7	151	211	327
		actual	(5)	(123)	(238)	(705)
0.4	0.6	no lockdown	187	3 805	6 834	14 287
		lockdown	7	153	216	339
		actual	(5)	(123)	(238)	(705)

very effective.

Chapter 6

The effect of smoking habits on lung cancer incidence.

6.1 Introduction

Tobacco smoking is the world's leading cause of preventable disease and is the biggest risk factor for lung cancer [128]. This epidemic, including exposure to secondhand smoke, accounts for 8 million non-communicable diseases deaths globally every year and if preventive measures are not enforced, is projected to increase markedly in the coming years (World Health Organisation (WHO) [6]). Tobacco smoke was suspected as a cause for lung cancer as early as the late 1920s following an increasing trend in the numbers of patients with this type of rare disease [129]. Proctor's [130] chronological study of cigarette-lung cancer trends has revealed that the epidemic started between the 1940s and 1950s, a conclusion reached from a variety of studies. Globally, there has been a noticeable increase in the number of lung cancer cases and deaths since 2011. The increase is attributed mainly to an increasing number of smokers in developing countries [131].

We used data provided by the Office for National Statistics (ONS) on smoking habits in Wales to understand better the effect of smoking on lung cancer development. Lung cancer is the third most common cancer in men and second most common cancer in women in Wales (Welsh Cancer Intelligence and Surveillance Unit (WCISU) [132]). At present around 50% of the people diag-

nosed with lung cancer in Wales die from the disease within six months of diagnosis and almost 75% within a year [132] and the survival rates are amongst the lowest in Europe [133]. The ill-effects of smoking on health, concern not only the smokers but the entire population sharing the same environment and economy. Its effect on healthcare, productivity (smokers take more sick leave), social care (smokers require extra care later in life due to smoke-related illness) and house fires, is a huge economic burden. For example, smoking costs the National Health Service of Wales (NHSW) an estimated £302 million per year [134], 7% of its total healthcare expenditure. To stop lung cancer being a major public health problem in the future, tobacco control is still needed to bring the rates down significantly [132]. This study seeks to understand the factors that fuel or maintain this epidemic.

Not much has been done in terms of modeling the dynamics of smoking and lung cancer at a population level. In this study we developed two models motivated by the modeling framework in [73, 76]. Although the intervention measures on reducing smoking and consequently lung cancer cases are known, the efficacies of these measures have not been addressed quantitatively. This study addresses the efficacy of these measures and asks the question “what number of new potential smokers should not be exceeded for the intervention measures to succeed,” that is, for the epidemic to clear.

For the stochastic model, we developed a system of five equations consisting of moderate and heavy smokers, lung cancer sufferers in limited and extensive stages and lung cancer associated deaths, in line with the available data. In this model, moderate and heavy smokers were defined as follows: moderate smokers are those who smoke on average less than or equal to 10 cigarettes per day and generally comprising individuals in the age group 18 – 34 years [135] and heavy smokers are those who smoke on average more than 10 cigarettes per day and generally comprising individuals over 35 years. Stage I - III clinical lung cancer cases were defined as the limited stage lung cancer whilst stage IV was defined as the extensive stage lung cancer. For the deterministic model, we developed a system of seven non-linear ordinary differential equations, represented by the susceptible population, moderate and heavy smokers, limited stage lung cancer

sufferers and extensive stage lung cancer sufferers, ex-smokers and lung cancer survivors.

6.2 The Stochastic Model

Consider a stochastic model consisting of five classes namely, moderate smokers, $x_1(t)$, heavy smokers, $x_2(t)$, limited stage lung cancer patients, $x_3(t)$, extensive stage lung cancer patients, $x_4(t)$, and smoking lung cancer deaths, $x_5(t)$. To illustrate this, we fitted the model to ONS data in Wales. The plots are shown in Figure 6.1. The actual data representing moderate and heavy smokers are explained by linear functions with negative slopes (Figures 6.1(a) and 6.1(b)) for the period 2001 to 2019. The populations of limited stage lung cancer, extensive stage lung cancer and lung cancer deaths are explained by random processes. To account for stochasticity in these populations, we propose to formulate a stochastic model of the type:

$$dx_t = \mu(x) dt + \sigma(x) dB, \quad (6.1)$$

where

$$\begin{aligned} x_t &= (x_1, x_2, x_3, x_4, x_5)^T, & \mu(x) &= (f_1(x_1), f_2(x_2), f_3(x_3), f_4(x_4), f_5(x_5))^T, \\ \sigma(x) &= \text{diag}_{i=1..5}(\sigma_i), & dB &= (dB_1, dB_2, dB_3, dB_4, dB_5)^T, \end{aligned}$$

where $B = (B_1, B_2, \dots, B_5)$ is a 5-D Brownian motion.

The functions $f_i(x_i)$, $i = 1, \dots, 5$ are at most of linear growth. As an example, we will illustrate this by finding functions $f_i(x_i)$, $i = 1, \dots, 5$, that explain the actual Wales data in Figure 6.1. This fitting can be done to any data set.

First, we want to prove that the stochastic model (6.1) does not give negative values of the state variables x_i , $i = 1..5$, despite the inclusion of the noise terms. We denote by \mathfrak{R}_+^5 the points

$$x_t = (x_1(t), x_2(t), x_3(t), x_4(t), x_5(t)) \in \mathfrak{R}_+^5$$

with positive coordinates and consider the stochastic model with initial values

$$x^0 = (x_1^0, x_2^0, x_3^0, x_4^0, x_5^0) \in \mathfrak{R}_+^5.$$

We want to prove the following Lemma:

Lemma 6.1. (*Positivity of Solutions*)

For any finite $T > 0$, the solution x_t of the system (6.1) with initial conditions $x^0 \in \mathfrak{R}_+^5$ remains in \mathfrak{R}_+^5 for all $t < T \wedge \tau$ (where $T \wedge \tau$ denotes the smaller between T and τ) so that the components of x_t satisfy $x_i > 0$ if $t < T \wedge \tau$ for $i = 1, \dots, 5$.

Proof. We define a Lyapunov-type function as follows:

$$\begin{aligned} V(x_1, \dots, x_5) &= x_1 - k_1 \log x_1 + x_2 - k_2 \log x_2 + x_3 - k_3 \log x_3 + x_4 \\ &\quad - k_4 \log x_4 + x_5 - k_5 \log x_5 \\ &= \sum_{i=1}^5 (x_i - k_i \log x_i). \end{aligned} \tag{6.2}$$

Applying Ito's formula, we have

$$\begin{aligned} dV &= \sum_{i=1}^5 \left(\left(1 - \frac{k_i}{x_i}\right) dx_i + \frac{1}{2} k_i \alpha_i^2 dt \right) \\ &= \left(1 - \frac{k_1}{x_1}\right) (\mu_1 x_1 dt + \sigma_1 x_1 dB_1) + \frac{1}{2} k_1 \alpha_1^2 dt + \left(1 - \frac{k_2}{x_2}\right) \\ &\quad (\mu_2 x_2 dt + \sigma_2 x_2 dB_2) + \frac{1}{2} k_2 \alpha_2^2 dt + \left(1 - \frac{k_3}{x_3}\right) (\sigma_3 x_3 dB_3) \\ &\quad + \frac{1}{2} k_3 \alpha_3^2 dt + \left(1 - \frac{k_4}{x_4}\right) (\sigma_4 x_4 dB_4) + \frac{1}{2} k_4 \alpha_4^2 dt + \left(1 - \frac{k_5}{x_5}\right) \\ &\quad (\sigma_5 x_5 dB_5) + \frac{1}{2} k_5 \alpha_5^2 dt. \\ &= \mu_1 x_1 dt - k_1 \mu_1 dt + \sigma_1 x_1 dB_1 - k_1 \sigma_1 dB_1 + \frac{1}{2} k_1 \alpha_1^2 dt \\ &\quad + \mu_2 x_2 dt - k_2 \mu_2 dt + \sigma_2 x_2 dB_2 - k_2 \sigma_2 dB_2 + \frac{1}{2} k_2 \alpha_2^2 dt \\ &\quad + \sigma_3 x_3 dB_3 - k_3 \sigma_3 dB_3 + \frac{1}{2} k_3 \alpha_3^2 dt + \sigma_4 x_4 dB_4 - k_4 \sigma_4 dB_4 \\ &\quad + \frac{1}{2} k_4 \alpha_4^2 dt + \sigma_5 x_5 dB_5 - k_5 \sigma_5 dB_5 + \frac{1}{2} k_5 \alpha_5^2 dt. \end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{2} k_1 \alpha_1^2 dt + \sigma_1(x_1 - k_1) dB_1 + \frac{1}{2} k_2 \alpha_2^2 dt + \sigma_2(x_2 - k_2) dB_2 \\
&\quad + \frac{1}{2} k_3 \alpha_3^2 dt + \sigma_3(x_3 - k_3) dB_3 + \frac{1}{2} k_4 \alpha_4^2 dt + \sigma_4(x_4 - k_4) dB_4 \\
&\quad + \frac{1}{2} k_5 \alpha_5^2 dt + \sigma_5(x_5 - k_5) dB_5. \\
&= M dt + \sum_{i=1}^5 \sigma_i(x_i - k_i) dB_i,
\end{aligned}$$

where

$$M = \frac{1}{2} \sum_{i=1}^5 k_i \alpha_i^2, \quad x_j < k_j, j = 1, 2.$$

Integrating from 0 to $T \wedge \tau$, we get

$$\int_0^{T \wedge \tau} dV \leq \int_0^{T \wedge \tau} M dt + \int_0^{T \wedge \tau} \sum_{i=1}^5 \sigma_i(x_i - k_i) dB_i. \quad (6.3)$$

Taking expectations, we get

$$E[V(x_{(T \wedge \tau)})] \leq E[V(x^0)] + M(T \wedge \tau). \quad (6.4)$$

Note that if a path x_t is such that it exits \mathfrak{R}_+^5 at $T \wedge \tau$ then by definition (6.2) the function V becomes ∞ at the exit point. In view of (6.3) this probability is zero. Hence x_t does not exit \mathfrak{R}_+^5 . \square

We fitted the stochastic model to the actual Wales data and estimated the drift coefficients, $\mu(x)$, that fit the data to be:

$$\left. \begin{aligned}
\mu(x) &= (0.5754 - 0.0183 x_1, 1.0592 - 0.0336 x_2, 0, 0, 0) \\
\sigma(x) &= \text{diag}_{i=1, \dots, 5}(\sigma_i), \quad dB = (dB_1, dB_2, dB_3, dB_4, dB_5)^T.
\end{aligned} \right\} \quad (6.5)$$

The solution set for the stochastic model (6.1) with drift coefficients (6.5) is given in the Appendix C. The plots for the actual data and the simulated data from the model (6.1) with drift coefficients (6.5) are compared in Figure 6.1 for the period 2001 – 2019.

A linear fit was first used for the actual data, the stochastic model was then developed with the linear fit being used as the deterministic part to assess how the introduction of noise would compare with the actual data. The σ_i 's, $i =$

1, ..., 5 were varied to improve data fit. The estimation of the parameters σ_i is significantly difficult but the best fit was for $\sigma_i = 0.001, i = 1, \dots, 5$.

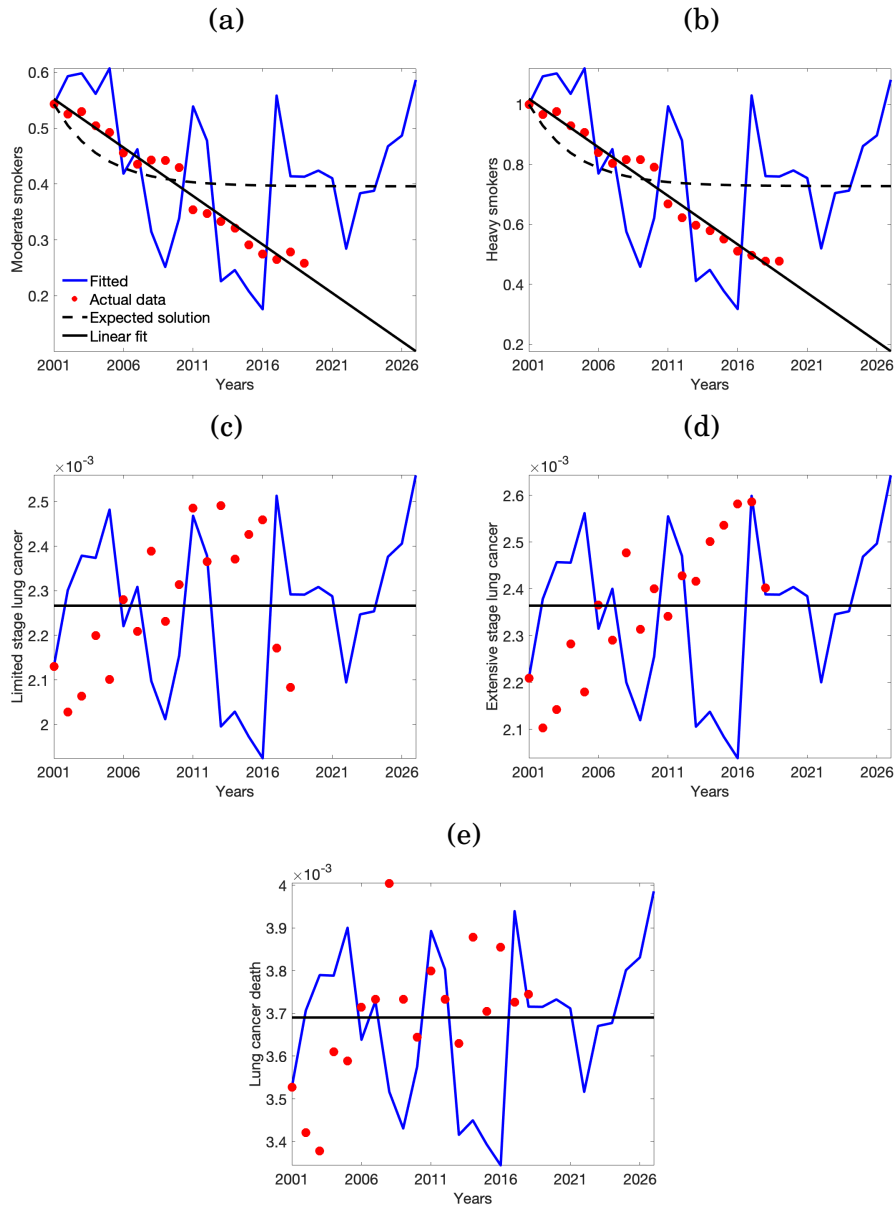


Figure 6.1: Fitting the stochastic model to proportions of (a) moderate smokers, (b) heavy smokers, (c) new cases for limited stage lung cancer, (d) new cases for extensive stage lung cancer and (e) lung cancer deaths

We can see that while the actual data yields a linear fit which predicts an end to the epidemic as the populations of moderate and addicted smokers become extinct, the expected solution suggests an initial decline but eventually tending to a non-zero steady state for moderate and addicted smokers, implying that there remains a proportion of smokers (both moderate and addicted) who never

quit the habit of smoking. This point is well illustrated by the expected solution represented by dotted lines in Figures 6.1 (a) and 6.1 (b).

In the absence of any intervention, the expected solution shows that the population of moderate smokers declined by 28% and that of addicted smokers by 22%. A proportion between 72 – 78% of the initial populations of moderate and addicted smokers remain to fuel the epidemic. The number of lung cancer cases, that is, the limited and extensive stages do not exhibit any trend (rising or falling). The expected solutions for lung cancer cases and lung cancer deaths are given by constant horizontal lines for both cases (Figures 6.1(c), 6.1(d)) and lung cancer deaths (Figure 6.1(e)). This result suggests that the levels of lung cancer cases and lung cancer deaths would remain the same for some time despite a decrease in the population of both moderate and heavy smokers.

First, this analysis shows the danger associated with interpreting results obtained from short data sets. However, if statistical approaches are used simultaneously with mathematical methods, long term patterns and trends can be inferred to limit the danger of over-interpretation. The use of a stochastic model fitted to the actual data has led to identification of long term trends for the five classes of data.

These results do not take into account the role played by the susceptible population towards sustaining the smoking population and ultimately lung cancer cases. Therefore, it is not correct to draw definitive conclusions without including the population of susceptible individuals who feed into the classes of smokers. In the next section, we intend to develop a deterministic model that accounts for susceptible individuals. The justification for developing a deterministic model is provided by the fact that the populations of moderate and addicted smokers progressed to stable states as suggested by the expected solutions in Figures 1(a) and 1(b).

The deterministic model will investigate how the conclusions reached for the system 6.1 for smokers are altered when intervention measures are introduced.

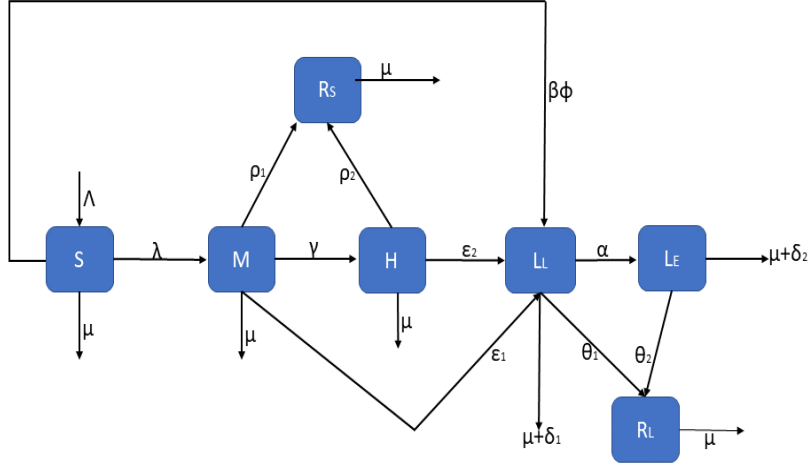


Figure 6.2: Flowchart of population interactions for the deterministic model.

6.3 The Deterministic Model

We consider a population consisting of seven classes with the flow of interactions described in Figure 6.2. The class, $S(t)$, represents susceptibles, that is, individuals who do not smoke but have the potential to become smokers due to influence of those who smoke. The class, $M(t)$, represents moderate smokers, that is, individuals who smoke occasionally but are not yet addicted to smoking. The class, $H(t)$, represents individuals who smoke heavily and are addicted to the habit. The class, $L_L(t)$, represents individuals who have developed limited stage lung cancer. The class, $L_E(t)$, represents individuals who have developed extensive stage lung cancer. The class, $R_S(t)$, represents those who have stopped smoking due to intervention campaigns or treatment. Finally, the class $R_L(t)$ represents those who have recovered from lung cancer due to treatment. The population size, excluding deaths, at time t denoted by $N(t)$ is given by,

$$N(t) = S(t) + M(t) + H(t) + L_L(t) + L_E(t) + R_S(t) + R_L(t).$$

In the deterministic model the total lung cancer cases and deaths are handled differently from the stochastic model. Therefore, a change in variable names has been implemented to avoid confusion with the stochastic model. However, the change still allows for comparison of some stochastic and deterministic classes. In what follows, we simplify the variable notation as follows: $y_0(t) := S(t)$, $y_1(t) := M(t)$, $y_2(t) := H(t)$, $y_3(t) := L_L(t)$, $y_4(t) := L_E(t)$, $y_5(t) := R_S(t)$ and $y_6(t) := R_L(t)$. The system of equations describing the dynamics in Figure 6.2 is given by,

$$\frac{dy_0}{dt} = \Lambda - (\mu + \lambda + \beta \phi(y_1, y_2)) y_0 \quad (6.6)$$

$$\frac{dy_1}{dt} = \lambda y_0 - a y_1 \quad (6.7)$$

$$\frac{dy_2}{dt} = \gamma y_1 - b y_2 \quad (6.8)$$

$$\frac{dy_3}{dt} = \beta \phi(y_1, y_2) y_0 + \varepsilon_1 y_1 + \varepsilon_2 y_2 - c y_3 \quad (6.9)$$

$$\frac{dy_4}{dt} = \alpha y_3 - d y_4 \quad (6.10)$$

$$\frac{dy_5}{dt} = \rho_1 y_1 + \rho_2 y_2 - \mu y_5 \quad (6.11)$$

$$\frac{dy_6}{dt} = \theta_1 y_3 + \theta_2 y_4 - \mu y_6, \quad (6.12)$$

where

$$\begin{aligned} a &= (\mu + \varepsilon_1 + \gamma + \rho_1), & b &= (\mu + \varepsilon_2 + \rho_2), & c &= (\mu + \delta_1 + \theta_1 + \alpha) \\ d &= (\mu + \delta_2 + \theta_2), & \lambda &= \frac{\beta_1 y_1}{1 + k y_1} + \frac{\beta_2 y_2}{1 + k y_2}. \end{aligned}$$

All the parameters, Λ , μ , ε_1 , ε_2 , γ , ρ_1 , ρ_2 , α , δ_1 , δ_2 , θ_1 , θ_2 , β , β_1 , β_2 and k are assumed to be positive. The parameters β_1 and β_2 define rates at which a susceptible becomes a smoker due to influence from the moderate and heavy smokers, respectively. A detailed description of the model parameters is given in Table 6.1. The terms $\frac{1}{1+k y_1}$ and $\frac{1}{1+k y_2}$ describe the inhibition effect from the behavioral change of the susceptible individuals when their number increases or from the crowding effect of the infected individuals.

Equation (6.6) describes the rate of change in the susceptible class. The first

Table 6.1: Model variables and parameters

Variable or Parameter	Description	Value	Source
S or y_0	Susceptible individuals		
M or y_1	Moderate smokers		
H or y_2	Heavy smokers		
L_L or y_3	Limited stage lung cancer sufferers		
L_E or y_4	Extensive stage ung cancer sufferers		
R_S or y_5	Ex-smokers		
R_L or y_6	Lung cancer survivors		
Λ	Recruitment rate into S	0.0125	[136, 137]
β_1	Progression rate from S to M due to M	0.372858	estimate
β_2	Progression rate from S to M due to H	0.299826	estimate
ϕ	Level of ETS	$0 \leq \phi \leq 1$	variable
β	Progression rate from S to M due to ETS	2.72766×10^{-7}	estimate
μ	Natural death rate of the general population	0.0125 yr^{-1}	[136, 137]
ρ_1	Recovery rate for moderate smokers	$0.07653913 \text{ yr}^{-1}$	estimate
ρ_2	Recovery rate for heavy smokers	0.11 yr^{-1}	estimate
γ	Progression rate from M to H	0.15611077	estimate
ε_1	Progression rate from M to L_L	0.0011	estimate
ε_2	Progression rate from H to L_L	0.0011	estimate
k	Inhibition factor	0.25	estimate
θ_1	Recovery rate for limited stage lung cancer	0.015 yr^{-1}	[138]
θ_2	Recovery rate for extensive stage lung cancer	0.01 yr^{-1}	[139]
α	Progression rate from L_L to L_E	3.7174×10^{-4}	estimate
δ_1	Disease induced death rate for L_L	$2.89982 \times 10^{-5} \text{ yr}^{-1}$	estimate
δ_2	Disease induced death rate L_E	0.0001 yr^{-1}	estimate
N_0	Initial size of total population	2 910 200	[132, 136, 138]

term represents recruitment into this class through births or immigration at a constant rate Λ . The second term represents losses due to natural death at a constant rate μ and addiction to smoking at a rate λ . $\beta \phi(y_1, y_2) y_0$ represents loss due to the effects of secondhand smoke, and we have assumed that these individuals develop limited stage lung cancer. The function $\phi(y_1, y_2)$ is an indication of the level of environmental tobacco smoke (ETS), defined as:

$$\phi(y_1, y_2) = \begin{cases} 0 & \text{if } y_1 = y_2 = 0, \\ 1 & \text{if } y_1 \neq y_2 \neq 0. \end{cases} \quad (6.13)$$

Equation (6.7) describes the rate of change in the moderate smokers' class. The first term represents gain of individuals from the susceptible class who acquire smoking habits. The second term represents total losses from this class consisting of those who die naturally at a constant rate μ , those who develop limited stage lung cancer at a constant rate ε_1 , those whose addiction worsens and become heavy smokers at a constant rate γ and those who quit smoking at a constant rate ρ_1 .

Equation (6.8) describes the rate of change in the heavy smokers class. The first term represents gain from conversion of moderate smokers to heavy smok-

ers as their addiction worsens at a constant rate γ . The second term represents losses due to natural death at a constant rate μ , some heavy smokers developing limited stage lung cancer at a constant rate ε_2 and some heavy smokers quitting smoking at a constant rate ρ_2 .

Equation (6.9) describes the rate of change in the limited stage lung cancer class. The first three terms represent gains from susceptibles who develop limited stage lung cancer from secondhand smoke and moderate and heavy smokers who develop limited stage lung cancer at constant rates ε_1 and ε_2 , respectively. The fourth term represents total losses resulting from natural death at a constant rate μ , disease induced death at a constant rate δ_1 , some limited stage lung cancer sufferers recovering at a constant rate θ_1 and some limited stage lung cancer sufferers progressing to the extensive stage lung cancer class at a constant rate α .

Equation (6.10) describes the rate of change in the extensive stage lung cancer class. The first term represents gain from the limited stage lung cancer as the disease worsens at a constant rate α . The second term represents total losses from natural death at a constant rate μ , disease induced death at a constant rate δ_2 and recovery at a constant rate θ_2 .

Equation (6.11) describes the rate of change in the recovered smokers class. The first two terms represent gain from individuals who quit smoking from the moderate and heavy smokers classes at constant rates ρ_1 and ρ_2 , respectively. The third term represents loss from natural death at a constant rate μ .

Equation (6.12) describes the rate of change in the recovered lung cancer sufferers class. The first two terms represent gain from individuals who recover from limited stage lung cancer and extensive stage lung cancer at constant rates θ_1 and θ_2 , respectively. The third term represents loss from natural death at a constant rate μ .

6.3.1 Model Analysis

6.3.1.1 Positivity and Boundedness of Solutions

We will show that all state variables of the system (6.6)-(6.12) are non-negative for all $t \geq 0$. Denote by \mathfrak{R}_+^7 the points,

$y(t) = (y_0(t), y_1(t), y_2(t), y_3(t), y_4(t), y_5(t), y_6(t)) \in \mathfrak{R}^7$ with positive coordinates and consider the system (6.6)-(6.12) with initial values,

$y^0 = (y_0^0, y_1^0, y_2^0, y_3^0, y_4^0, y_5^0, y_6^0) \in \mathfrak{R}_+^7$. We prove the following theorem:

Theorem 6.2. *If $y_i^0 \geq 0$, $i = 0, \dots, 6$ then $y_i(t) \geq 0$ for $t > 0$, $i = 0, \dots, 6$.*

Proof. We can write the system of equations (6.6)-(6.12) in the form:

$$\frac{dy}{dt} = A(y)y + B,$$

where $y = (y_0, y_1, y_2, y_3, y_4, y_5, y_6)^T$,

$$A(y) = \begin{bmatrix} -(\mu + \lambda + \beta\phi) & 0 & 0 & 0 & 0 & 0 & 0 \\ \lambda & -a & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma & -b & 0 & 0 & 0 & 0 \\ \beta\phi & \varepsilon_1 & \varepsilon_2 & -c & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & -d & 0 & 0 \\ 0 & \rho_1 & \rho_2 & 0 & 0 & -\mu & 0 \\ 0 & 0 & 0 & \theta_1 & \theta_2 & 0 & -\mu \end{bmatrix} \text{ and } B = \begin{bmatrix} \Lambda \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

We can see that all the off-diagonal elements of $A(y)$ are non-negative, hence the matrix $A(y)$ is a Metzler matrix implying that the given system can not admit negative solutions y for $t \geq 0$, (see Ngeleja et al. [140]) for details. \square

6.3.2 The Reproduction Number

At the disease free equilibrium (DFE) there are no secondary smokers ($\phi(y_1, y_2) = 0$), hence, the DFE is obtained by equating the right hand side of

(6.6)-(6.12) to zero. This gives

$$\zeta_0 = (y_{0_0}, y_{1_0}, y_{2_0}, y_{3_0}, y_{4_0}, y_{5_0}, y_{6_0}) = \left(\frac{\Lambda}{\mu}, 0, 0, 0, 0, 0, 0 \right). \quad (6.14)$$

Using the technique by Van den Driessche and Watmough [80], we have computed the basic reproduction number, R_0 , as shown below. From the system (6.6)-(6.12), we define the matrices for new infections \mathbf{F} and other transitions \mathbf{V} as follows:

$$\mathbf{F} = \begin{bmatrix} \left(\frac{\beta_1 y_1}{1+k y_1} + \frac{\beta_2 y_2}{1+k y_2} \right) y_0 \\ 0 \\ \beta \phi y_0 \\ 0 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} a y_1 \\ b y_2 - \gamma y_1 \\ c y_3 - (\varepsilon_1 y_1 + \varepsilon_2 y_2) \\ d y_4 - \alpha y_3 \end{bmatrix}.$$

The Jacobians for \mathbf{F} , \mathbf{V} and the inverse of \mathbf{V} are given by:

$$\mathbf{F} = \begin{bmatrix} \frac{\beta_1 y_{0_0}}{1+k y_{1_0}} & \frac{\beta_2 y_{0_0}}{1+k y_{2_0}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} a & 0 & 0 & 0 \\ -\gamma & b & 0 & 0 \\ -\varepsilon_1 & -\varepsilon_2 & c & 0 \\ 0 & 0 & -\alpha & d \end{bmatrix}, \quad \mathbf{V}^{-1} = \begin{bmatrix} \frac{1}{a} & 0 & 0 & 0 \\ \frac{\gamma}{ab} & \frac{1}{b} & 0 & 0 \\ \frac{\gamma \varepsilon_2 + \varepsilon_1 b}{abc} & \frac{\varepsilon_2}{bc} & \frac{1}{c} & 0 \\ \frac{\alpha(\gamma \varepsilon_2 + \varepsilon_1 b)}{abcd} & \frac{\alpha \varepsilon_2}{bcd} & \frac{\alpha}{cd} & \frac{1}{d} \end{bmatrix}.$$

The spectral radius of the matrix $\mathbf{F}\mathbf{V}^{-1}$ denoted by ρ is the maximum of the moduli of the eigenvalues of matrix $\mathbf{F}\mathbf{V}^{-1}$. Thus, $R_0 = \rho(\mathbf{F}\mathbf{V}^{-1})$ given

$$\mathbf{F}\mathbf{V}^{-1} = \begin{bmatrix} \frac{\beta_1 y_{0_0}}{a(1+k y_{1_0})} + \frac{\gamma \beta_2 y_{0_0}}{ab(1+k y_{2_0})} & \frac{\beta_2 y_{0_0}}{b(1+k y_{2_0})} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

is given by $R_0 = \frac{\beta_1 \Lambda}{\mu a} + \frac{\gamma \beta_2 \Lambda}{\mu ab}$.

Remark: $R_{0_M} = \frac{\beta_1 \Lambda}{\mu a}$ is a sub-reproduction number for secondary infection generated by influence of moderate smokers and $R_{0_H} = \frac{\gamma \beta_2 \Lambda}{\mu ab}$ is a sub-reproduction number for secondary infection generated by influence of heavy smokers.

Both $\frac{\beta_1 \Lambda}{\mu a}$ and $\frac{\gamma \beta_2 \Lambda}{\mu ab}$ can be less than one but the sum can be greater than 1. It is important to stress that the disease clears if the sum of the two fractions does

not exceed 1. The epidemic can only clear if both groups of smokers adhere to the intervention measures.

Theorem 6.3. *If $R_0 < 1$ then the DFE, ζ_0 , is stable, otherwise it is unstable.*

6.3.2.1 Global Stability of the DFE

We present the conditions that guarantee the global asymptotic stability of the DFE. We can write the system (6.6)-(6.12) in the following form:

$$\begin{aligned} Y'(t) &= F(Y, Z) \\ Z'(t) &= G(Y, Z), \quad G(Y, 0) = 0, \end{aligned}$$

where $Y = (y_0, y_5, y_6)$ and $Z = (y_1, y_2, y_3, y_4)$. The components of $Y \in \mathfrak{R}_+^3$ represent the classes of uninfected states and the components of $Z \in \mathfrak{R}_+^4$ represent the classes of infected states. Note that, the DFE can be written from the components of ζ_0 in (6.14) as

$$\zeta_0 = (Y^*, 0, 0), \quad \text{where } Y^* = \frac{\Lambda}{\mu}.$$

Following Castillo-Chavez et al. [141], two conditions $H1$ and $H2$ below must be met to guarantee global asymptotic stability:

$$\left. \begin{aligned} (H1) \quad &Y'(t) = F(Y, 0), \quad Y^* \text{ is locally asymptotically stable.} \\ (H2) \quad &G(Y, Z) = AZ - G(\hat{Y}, Z), \quad G(\hat{Y}, Z) \geq 0 \quad \text{for } (Y, Z) \in \Omega. \end{aligned} \right\} \quad (6.15)$$

where $A = D_Z G(Y^*, 0, 0)$ is an M -matrix and Ω is the region where the model is well-posed.

Theorem 6.4. *The DFE, $\zeta_0 = (Y^*, 0, 0)$ of the system (6.6)-(6.12) is not globally asymptotically stable (GAS).*

Proof. We have from (H1) and the definition of $A = D_Z G(Y^*, 0, 0)$,

$$F(Y, 0) = \begin{bmatrix} \Lambda - \mu y_0 \\ -\mu y_5 \\ -\mu y_6 \end{bmatrix}, \quad A = \begin{bmatrix} -a & 0 & 0 & 0 \\ \gamma & -b & 0 & 0 \\ \varepsilon_1 & \varepsilon_2 & -c & 0 \\ 0 & 0 & \alpha & -d \end{bmatrix}.$$

Condition (H2) simplifies to,

$$G(\hat{Y}, Z) = \begin{bmatrix} G_1(\hat{Y}, Z) \\ G_2(\hat{Y}, Z) \\ G_3(\hat{Y}, Z) \\ G_4(\hat{Y}, Z) \end{bmatrix} = \begin{bmatrix} -\lambda y_0 \\ 0 \\ -\beta \phi \\ 0 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Clearly the condition H_2 in (6.15) is not met, and we conclude that ζ_0 is not GAS. \square

Remark: Theorem 6.4 has implications regarding the control of the smoking lung cancer epidemic. It is clear from this theorem that control of this epidemic depends on the initial conditions. Hence, campaigns to educate susceptible individuals so that they do not start smoking and government regulation on the price of tobacco products should be intensified in schools and hospitality places such as bars and restaurants. The question is “what level of intervention is needed to clear the epidemic?” This question is addressed in section 6.4.3 on optimal control.

6.3.3 Existence of the Endemic Equilibrium Point (EEP)

At the endemic equilibrium point, we can take without loss of generality the maximum possible value for $\phi = 1$. Solving the system (6.6)-(6.12) by equating the right hand side to zero, we find the EEP given by $\zeta_1 =$

$(y_0^*, y_1^*, y_2^*, y_3^*, y_4^*, y_5^*, y_6^*)$, where in terms of λ^* is given by:

$$\begin{aligned} y_0^* &= \frac{\Lambda}{\beta + \lambda^* + \mu}, & y_1^* &= \frac{\Lambda \lambda^*}{a(\beta + \lambda^* + \mu)}, & y_2^* &= \frac{\Lambda \lambda^* \gamma}{ab(\beta + \lambda^* + \mu)} \\ y_3^* &= \frac{\Lambda a_0}{ab c(\beta + \lambda^* + \mu)}, & y_4^* &= \frac{\Lambda a_0 \alpha}{abcd(\beta + \lambda^* + \mu)}, \\ y_5^* &= \frac{\Lambda \lambda^* (b \rho_1 + \gamma \rho_2)}{ab \mu(\beta + \lambda^* + \mu)}, & y_6^* &= \frac{\Lambda a_0 (d \theta_1 + \alpha \theta_2)}{abcd \mu(\beta + \lambda^* + \mu)}, \end{aligned}$$

where

$$\lambda^* = \frac{\beta_1 y_1^*}{1 + k y_1^*} + \frac{\beta_2 y_2^*}{1 + k y_2^*}.$$

and

$$a_0 = ab\beta + b\lambda^* \varepsilon_1 + \gamma \lambda^* \varepsilon_2.$$

Using (6.8) we have found a polynomial for λ^* given by:

$$A \lambda^{*2} + B \lambda^* + C = 0, \quad (6.16)$$

where the coefficients A, B and C are given by:

$$\left. \begin{aligned} A &= -[k^2 \Lambda^2 \gamma + ak \Lambda \gamma + abk \Lambda + a^2 b]. \\ B &= a^2 b \mu (R_0 - 1) + k \gamma \beta_1 \Lambda^2 + k \Lambda^2 \beta_2 \gamma - (a^2 b \beta + a \beta k \Lambda \gamma + a^2 b \beta). \\ C &= a^2 b \mu^2 (R_0 - 1) + ab \mu \beta_1 \Lambda + a \beta \beta_2 \Lambda \gamma - (a^2 b \beta^2 + a^2 b \beta \mu + a^2 \beta \mu). \end{aligned} \right\} \quad (6.17)$$

Proposition 6.5. *If $C < 0$ and $B > 0$ then the system (6.6)-(6.12) possesses two real positive endemic equilibrium points.*

Remark: We have not calculated the two endemic equilibrium points explicitly because it is tedious to do so even with numerical software but instead, we want to determine the conditions under which backward bifurcation can occur if one of the endemic equilibrium points is unstable and exists for $R_0 < 1$.

6.3.3.1 Bifurcation Analysis

We shall investigate the nature of bifurcation involving the DFE, ζ_0 , using the centre manifold theory described in Buonomo and Lacitignola [142]. Using the transcritical point $R_0 = 1$, we can restate Theorem 6.3 as follows:

Lemma 6.6. *The DFE, ζ_0 , is locally stable if $\beta_2 < \frac{b(\mu a - \beta_1 \Lambda)}{\gamma \Lambda}$ and unstable if $\beta_2 > \frac{b(\mu a - \beta_1 \Lambda)}{\gamma \Lambda}$.*

The proof of Lemma 6.6 is given in Appendix D. Lemma 6.6 gives the important relationship between the infection coefficient for heavy smoker, β_2 , and the infection coefficient for moderate smokers, β_1 . In terms of the sub-reproduction number, R_{0_M} , for moderate smokers, Lemma 6.6 can be restated as follows:

Lemma 6.7. *The DFE, ζ_0 , is locally stable if $\beta_2 < \frac{b\beta_1}{\gamma R_{0_M}}(1 - R_{0_M})$ and unstable if $\beta_2 > \frac{b\beta_1}{\gamma R_{0_M}}(1 - R_{0_M})$.*

Remark: It is important to note that the positivity of β_2 means that $R_{0_M} < 1$.

Theorem 6.8. *The epidemic persists for $\Lambda > \Lambda^* = \sqrt{\frac{\mu a}{\beta_1 k}}$, even for $R_0 < 1$. In other words one of the endemic equilibrium states of proposition 6.5 undergoes backward bifurcation.*

The proof of Theorem 6.8 is in two parts:

- (i) to determine the right eigenvector $w = (w_0, w_1, w_2, w_3, w_4, w_5, w_6)$ and the left eigenvector $v = (v_0, v_1, v_2, v_3, v_4, v_5, v_6)$ associated with the Jacobian (7.1) at β_2^* .
- (ii) to use the components w_i and v_i , $i = 0, \dots, 6$ to define quadratic forms Z_1 and Z_2 as defined in [143] and to use the positivity of Z_1 and Z_2 to show the existence of backward bifurcation.

Proof. (i) Let $w = (w_0, w_1, w_2, w_3, w_4, w_5, w_6)$ and $v = (v_0, v_1, v_2, v_3, v_4, v_5, v_6)$ be the right eigenvector and the left eigenvector, respectively associated with the Jacobian (7.1) in Appendix E at β_2^* . Following [142] and [143], we can set up a system of linear equations for w and v given below:

$$R w^T = 0 \tag{6.18}$$

and

$$L v^T = 0 \quad (6.19)$$

where R and L are given in Appendix E.

Solving the system (6.18) in terms of $w_2 > 0$ gives

$$\left. \begin{aligned} w_0 &= \frac{a a_2}{\mu (a_1 - a)} w_2, & w_1 &= -\frac{a_2}{(a_1 - a)} w_2, & w_3 &= \frac{\Gamma_2}{c (a_1 - a)} w_2 \\ w_4 &= \frac{\alpha \Gamma_2}{c d (a_1 - a)} w_2, & w_5 &= \frac{\rho_2 (a_1 - a) - a_2 \rho_1}{\mu (a_1 - a)} w_2 \\ w_6 &= \frac{b_2 \Gamma_2}{c d \mu (a_1 - a)} w_2, \end{aligned} \right\} \quad (6.20)$$

where

$$\begin{aligned} b_2 &= d \theta_1 + \theta_2 \alpha, \\ \Gamma_2 &= \varepsilon_2 (a_1 - a) - \varepsilon_1 a_2. \end{aligned}$$

and solving the system (6.19) in terms of $v_2 > 0$ gives

$$\left. \begin{aligned} v_1 &= \frac{-\gamma}{(a_1 - a)} v_2, & v_0 &= v_3 = v_4 = v_5 = v_6 = 0 \end{aligned} \right\} \quad (6.21)$$

(ii) Following (Castillo-Chavez and Song [143]), we can define the quadratic forms Z_1 and Z_2 as follows:

$$Z_1 = \sum_{k,i,j=0}^n v_k w_i w_j \frac{\partial^2 f_k}{\partial y_i \partial y_j} (\zeta_0, \beta_2^*), \quad Z_2 = \sum_{k,i=0}^n v_k w_i \frac{\partial^2 f_k}{\partial y_i \partial \beta_2} (\zeta_0, \beta_2^*). \quad (6.22)$$

In terms of (6.20) and (6.21) we obtain:

$$Z_1 = \frac{2 v_2 w_2^2 \gamma}{\mu (a_1 - a)^2} \left(\frac{D}{\mu (a_1 - a)} \right), \quad Z_2 = -\frac{\gamma \Lambda}{\mu (a_1 - a)} v_2 w_2. \quad (6.23)$$

where

$$D = a a_2 \beta_1 \mu + a_2^2 \beta_1 \Lambda k \mu (a_1 - a) + \beta_2 \Lambda k (a_1 - a)^2 - \mu a a_2 \beta_2 (a_1 - a), \quad (6.24)$$

$$a_1 = \frac{\beta_1 \Lambda}{\mu}, \quad (6.25)$$

$$a_2 = \frac{\beta_2 \Lambda}{\mu}, \quad (6.26)$$

$$a = (\mu + \varepsilon_1 + \gamma + \rho_1). \quad (6.27)$$

The coefficients of the right and left eigenvectors must always be positive [143], so that:

$$Z_2 > 0, \quad \text{if } (a_1 - a) < 0 \quad (6.28)$$

and

$$Z_1 > 0, \quad \text{if } \frac{D}{\mu(a_1 - a)} > 0, \quad (6.29)$$

The inequalities (6.23) and (6.28) mean that heavy smokers, for some reasons, which could include heavy smokers reducing the number of cigarettes, peer influence etc are becoming moderate smokers, that is $\gamma \rightarrow -\gamma$. The signs of the terms in (6.24) are as follows: the first term is positive, the second term is negative, the third term is positive and the fourth term together with the sign is positive. We can solve the inequality $\frac{D}{\mu(a_1 - a)} > 0$ in (6.29) from the second and fourth terms in (6.24) as follows:

$$\begin{aligned} W &= a_2^2 \beta_1 \Lambda k \mu (a_1 - a) - \mu a a_2 \beta_2 (a_1 - a) \\ &= (a_1 - a) (a_2^2 \beta_1 \Lambda k \mu - \mu a a_2 \beta_2) \\ &= \left(\frac{\beta_1 \Lambda}{\mu} - a \right) \left(\frac{\Lambda^3 \beta_1 \beta_2^2 a k - \Lambda \beta_2^2 a^2 \mu}{\mu} \right) \\ &\leq \left(\frac{\Lambda \beta_1}{\mu} + \frac{\Lambda \beta_2 \gamma}{\mu b} - a \right) \left(\frac{\Lambda^3 \beta_1 \beta_2^2 a k - \Lambda \beta_2^2 a^2 \mu}{\mu} \right) \\ &= a (R_0 - 1) \left(\frac{\Lambda^3 \beta_1 \beta_2^2 a k - \Lambda \beta_2^2 a^2 \mu}{\mu} \right) \\ &< 0. \end{aligned} \quad (6.30)$$

The inequality (6.30) is true if either:

$$(i) \quad R_0 - 1 > 0 \quad \text{and} \quad \left(\frac{\Lambda^3 \beta_1 \beta_2^2 a k - \Lambda \beta_2^2 a^2 \mu}{\mu} \right) < 0 \quad (6.31)$$

or

$$(ii) \quad R_0 - 1 < 0 \quad \text{and} \quad \left(\frac{\Lambda^3 \beta_1 \beta_2^2 a k - \Lambda \beta_2^2 a^2 \mu}{\mu} \right) > 0 \quad (6.32)$$

From (6.31) we have $R_0 > 1$ and $\Lambda < \Lambda^* = \sqrt{\frac{\mu a}{\beta_1 k}}$. This is the usual endemic equilibrium requirement. It implies that the endemic equilibrium point will persist even for recruitment levels into the susceptible smokers class below $\Lambda < \Lambda^* = \sqrt{\frac{\mu a}{\beta_1 k}}$. From (6.32), we obtain $R_0 < 1$ and $\Lambda > \sqrt{\frac{\mu a}{\beta_1 k}}$. The endemic equilibrium point persists for $R_0 < 1$ if the recruitment levels into the susceptible smokers class exceeds $\sqrt{\frac{\mu a}{\beta_1 k}}$. This completes the proof. \square

Remark: The possibility of backward bifurcation has implications regarding clearance of the disease from the population as this can lead to wrong recommendations arising from complacency by public health officials who may be led to believe that the epidemic is under control. This study is suggesting a routine survey to estimate the recruitment into the susceptible smokers class.

Figures 6.3(a) and 6.3(b) illustrate the fact that the reproduction number, R_0 , increases with increasing numbers of potential new susceptible smokers. It is clear that if the number of new susceptible smokers exceeds $\Lambda^* = 11\,919$ the reproduction number exceeds one. Figure 6.3 shows that the endemic state can be maintained even at low recruitment levels of new susceptible smokers for $R_0 < 1$.

6.4 Numerical Simulations

We performed numerical simulations for the system (6.6)-(6.12), with the initial population values based on ONS data in Wales for the period 2001 – 2019 [132,

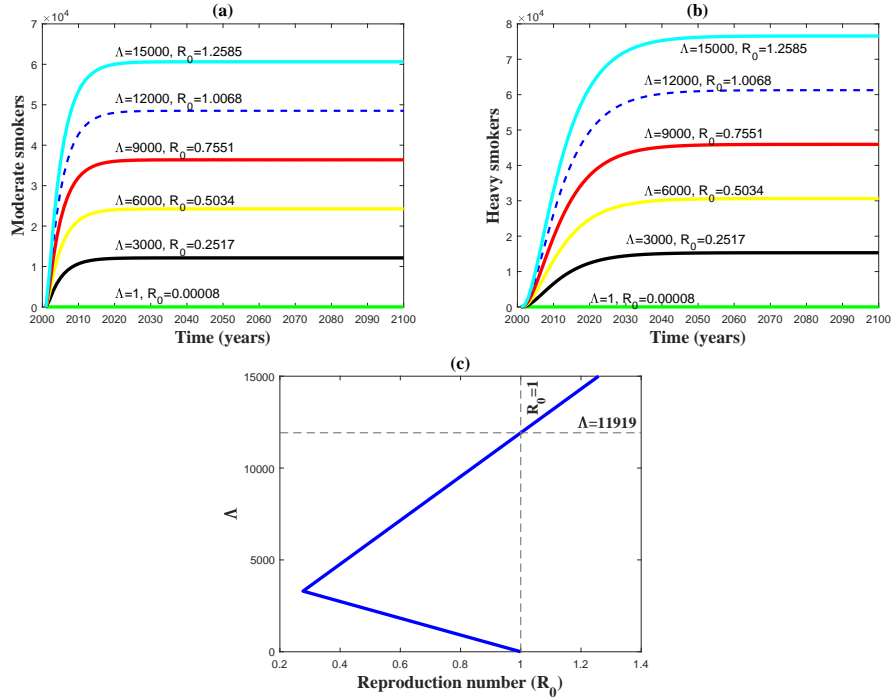


Figure 6.3: (a) moderate smokers, (b) heavy smokers and (c) R_0 for varying values of Λ .

138, 144]. The initial values for the seven populations, y_0, \dots, y_6 , are given by:

$$\begin{aligned}
 y^0 &= (y_0^0, y_1^0, y_2^0, y_3^0, y_4^0, y_5^0, y_6^0) \\
 &= (765\,942, 275\,419, 507\,425, 1\,081, 1\,121, 1\,359\,063, 149).
 \end{aligned}$$

6.4.1 Sensitivity Analysis

We conducted sensitivity analysis to assess how the reproduction number is correlated to the model parameters since many of the parameters in Table 6.1 have been estimated. We have analyzed how sensitive R_0 is to the changes in these parameters.

We used the method of Partial Correlation Coefficients (PRCC) to achieve the objective mentioned above (See Table 7.1 in Appendix F). The sensitivity analysis revealed that the parameters β_1 and Λ have a moderately strong positive correlation, while the parameter μ has a moderately strong negative correlation with R_0 . We stress that caution must be exercised when interpreting the results as changes in β_1 and Λ can overestimate the severity of the epidemic. The parameter μ is known clinically and is not an estimated parameter. Neverthe-

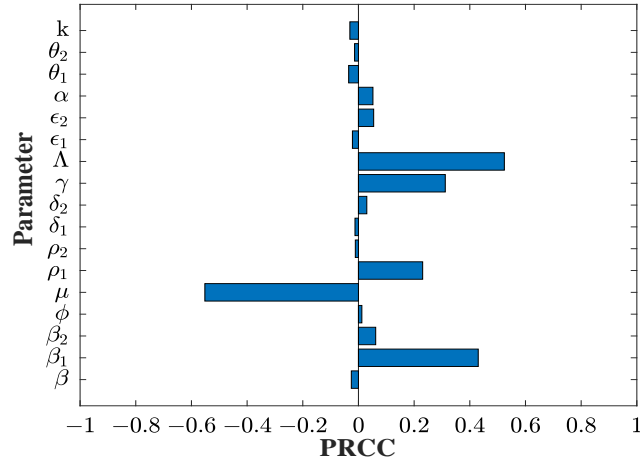


Figure 6.4: PRCC showing the linear relationship between R_0 and certain model parameters, when the effect of the other parameters is discounted.

less, long life for smokers has the effect of prolonging the smoking lung cancer pandemic. It is surprising that R_0 is weakly correlated to δ_1 and δ_2 , the disease induced death rates.

6.4.2 Determination of lung cancer cases and deaths

To fit the deterministic model to data, we defined two additional state variables representing the total number of lung cancer cases (y_7) and lung cancer deaths (y_8) that are derived from the solution of the system (6.6)-(6.12):

$$\begin{aligned} \frac{dy_7}{dt} &= \beta \phi y_0 + \epsilon_1 y_1 + \epsilon_2 y_2 - (\mu + \delta_1 + \theta_1) y_3 \\ &\quad - (\mu + \delta_2 + \theta_2) y_4 \end{aligned} \quad (6.33)$$

$$\frac{dy_8}{dt} = \delta_1 y_3 + \delta_2 y_4. \quad (6.34)$$

This step (equations (6.33) and (6.34)) is necessary because the data for lung cancer cases is given as a sum for each year, and we want to account for the two sources of lung cancer deaths. Figure 6.5 shows how our simulated results compare to the actual data for moderate and heavy smokers, lung cancer cases and lung cancer deaths. Figures 6.5(a) and 6.5(b) show a linear decline in both moderate and heavy smokers. The simulated results (using equations (6.33)

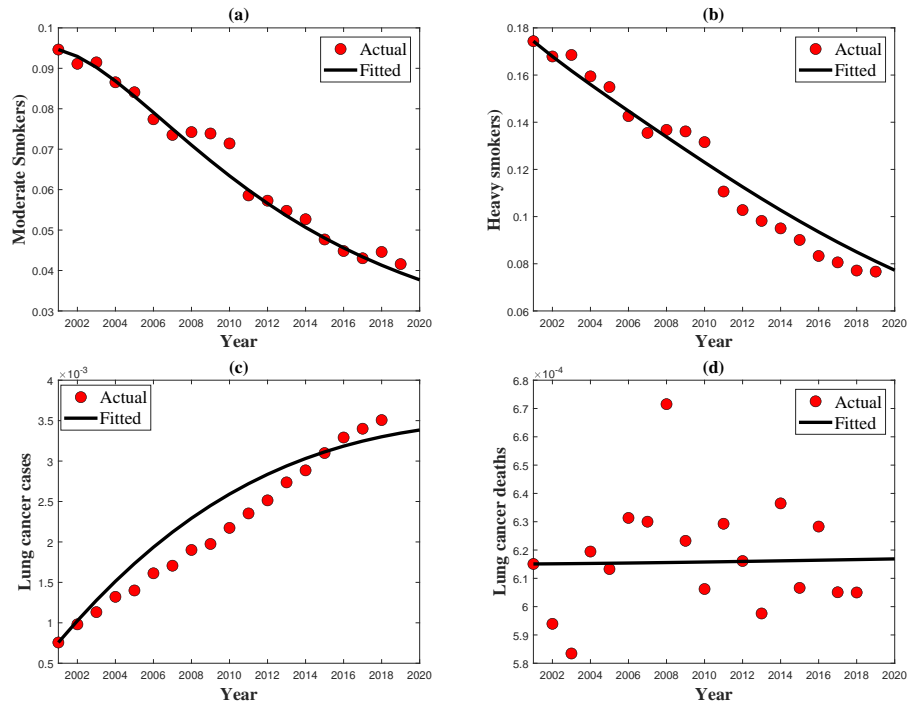


Figure 6.5: Simulations showing the comparison of actual and simulated cases of (a) moderate smokers, (b) heavy smokers (c) lung cancer sufferers and (d) lung cancer deaths using the deterministic model (6.6)-(6.12).

and (6.34))show a very good fit. The correlation coefficient between the actual and the simulated data for moderate smokers is $r = 0.9870$. Thus $r^2 = 0.974$ or 97.4% of the variability of the recorded moderate smokers data can be explained by the simulated moderate smokers data or vice versa. The correlation coefficient between the actual and the simulated data for heavy smokers is $r = 0.9875$. Thus $r^2 = 0.975$ or 97.5% of the variability of the recorded moderate smokers data can be explained by the simulated moderate smokers data or vice versa. For lung cancer cases $r = 0.972$, implying that 94.4% of the variability of the recorded lung cancer cases can be explained by the simulated lung cancer cases or vice versa. Figure 6.5(d) shows a stochastic behaviour for lung cancer deaths which agrees with the stochastic result (Figure 6.1(e)). The correlation analysis shows that a good fit for the actual data to the proposed deterministic model (6.6)-(6.12).

If we simulate to longer times (Figure 6.6), we see that a steady state is achieved as opposed to extinction suggested by the actual data in Figure 6.1 and Figures 6.5(a) and 6.5(b). Figures 6.6(a)-(c) show that for moderate smok-

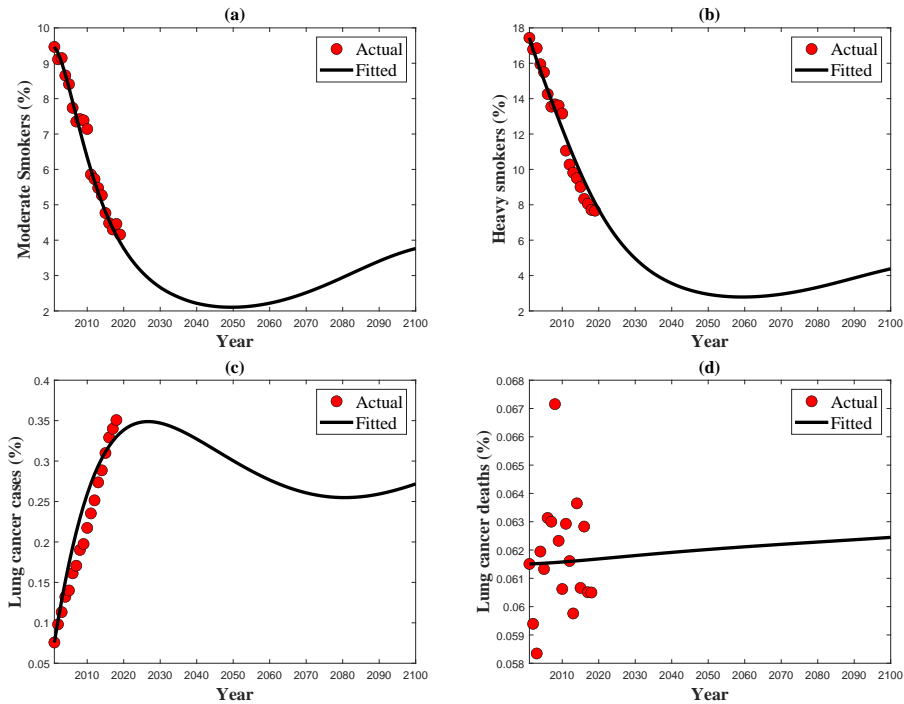


Figure 6.6: Longterm model predictions of actual and simulated cases of (a) moderate smokers, (b) heavy smokers (c) lung cancer sufferers and (d) lung cancer deaths.

ers, heavy smokers and lung cancer cases, there is no permanent stability. In particular, in Figures 6.6(a)-(b) the population for moderate and heavy smokers decrease towards extinction but eventually these populations move away from the axis of extinction. These results imply that interventions to curb the smoking epidemic can work but extinction of the populations of smokers is not achievable as there remains a group of die-hard smokers who will not quit smoking. This is also confirmed by the stochastic model results in Figure 6.1.

6.4.3 Optimal Control

Tobacco control has become an important priority at both national and international levels although the level of action varies from country to country [145]. The most efficient method to prevent lung cancer is to eliminate cigarette smoking and exposure to secondhand smoke. This can result in reduced morbidity and mortality not only of lung cancer, but also chronic obstructive pulmonary disease, coronary heart disease, cerebrovascular disease and other conditions [16]. We will use the system (6.6)-(6.12) to determine optimal control strategies.

Two possible controls are introduced to decrease the tendency towards smoking, these are u_1 and u_2 , representing social and medical treatment support, respectively. The governing equations for the control variables associated with the system (6.6)-(6.12) are given by:

$$\frac{dy_0}{dt} = \Lambda - \left[\mu + (1 - u_1) \left(\frac{\beta_1 y_1}{1 + k y_1} + \frac{\beta_2 y_2}{1 + k y_2} \right) + \beta \phi \right] y_0 \quad (6.35)$$

$$\begin{aligned} \frac{dy_1}{dt} &= (1 - u_1) \left(\frac{\beta_1 y_1}{1 + k y_1} + \frac{\beta_2 y_2}{1 + k y_2} \right) y_0 - (\mu + \varepsilon_1 + \gamma) y_1 \\ &\quad - (1 + u_2) \rho_1 y_1 \end{aligned} \quad (6.36)$$

$$\frac{dy_2}{dt} = \gamma y_1 - (\mu + \varepsilon_2) y_2 - (1 + u_2) \rho_2 y_1 \quad (6.37)$$

$$\frac{dy_3}{dt} = \beta \phi y_0 + \varepsilon_1 y_1 + \varepsilon_2 y_2 - c y_3 \quad (6.38)$$

$$\frac{dy_4}{dt} = \alpha y_3 - d y_4 \quad (6.39)$$

$$\frac{dy_5}{dt} = (1 + u_2) \rho_1 y_1 + (1 + u_2) \rho_2 y_2 - \mu y_5 \quad (6.40)$$

$$\frac{dy_6}{dt} = \theta_1 y_3 + \theta_2 y_4 - \mu y_6 \quad (6.41)$$

Our goal is to minimize the objective functional defined as:

$$J(u_1, u_2) = \int_0^{t_T} A_1 y_1 + A_2 y_2 + \frac{1}{2} A_3 u_1^2 + \frac{1}{2} A_4 u_2^2,$$

where A_1 and A_2 are the weight constants associated with the number of moderate and heavy smokers, respectively. The constants A_3 and A_4 are the weight constants of the control variables u_1 and u_2 , respectively. Let u_1^* , u_2^* be the optimal controls. We want to find a set of control functions such that

$$J(u_1^*, u_2^*) = \min_{(u_1, u_2) \in U} J(u_1, u_2)$$

subject to (6.35)-(6.41), where U is the set of measure functions defined from $[0, t_T]$ to $[0, 1]$. We use Pontryagin's Maximum Principle to solve the optimal control problem that satisfies the necessary conditions. The Hamiltonian H

from the objective function subject to the model system of equations is given by:

$$\begin{aligned}
H &= A_1 y_1 + A_2 y_2 + \frac{1}{2} A_3 u_1^2 + \frac{1}{2} A_4 u_2^2 + \sum_{i=0}^6 \lambda_{y_i} \frac{dy_i}{dt} \\
&= A_1 y_1 + A_2 y_2 + \frac{1}{2} A_3 u_1^2 + \frac{1}{2} A_4 u_2^2 \\
&\quad + \lambda_{y_0} \left[\Lambda - (1 - u_1) \left(\frac{\beta_1 y_1}{1 + k y_1} + \frac{\beta_2 y_2}{1 + k y_2} \right) y_0 - \beta \phi y_0 - \mu y_0 \right] \\
&\quad + \lambda_{y_1} \left[(1 - u_1) \left(\frac{\beta_1 y_1}{1 + k y_1} + \frac{\beta_2 y_2}{1 + k y_2} \right) y_0 - (\mu + \varepsilon_1 + \gamma) y_1 - (1 + u_2) \rho_1 y_1 \right] \\
&\quad + \lambda_{y_2} \left[\gamma y_1 - (\mu + \varepsilon_2) y_2 - (1 + u_2) \rho_2 y_2 \right] \\
&\quad + \lambda_{y_3} (\beta \phi y_0 + \varepsilon_1 y_1 + \varepsilon_2 y_2 - c y_3) \\
&\quad + \lambda_{y_4} (\alpha y_3 - d y_4) \\
&\quad + \lambda_{y_5} ((1 + u_2) \rho_1 y_1 + (1 + u_2) \rho_2 y_2 - \mu y_5) \\
&\quad + \lambda_{y_6} (\theta_1 y_3 + \theta_2 y_4 - \mu y_6)
\end{aligned}$$

where $\lambda_{y_0}, \lambda_{y_1}, \lambda_{y_2}, \lambda_{y_3}, \lambda_{y_4}, \lambda_{y_5}, \lambda_{y_6}$ are adjoint variables satisfying the following adjoint system:

$$\frac{d\lambda_{y_0}}{dt} = -\frac{\partial H}{\partial y_0}, \dots, \frac{d\lambda_{y_6}}{dt} = -\frac{\partial H}{\partial y_6},$$

where

$$\begin{aligned}
\frac{d\lambda_{y_0}}{dt} &= (\lambda_{y_0} - \lambda_{y_1}) (1 - u_1) \left(\frac{\beta_1 y_1}{1 + k y_1} + \frac{\beta_2 y_2}{1 + k y_2} \right) + (\lambda_{y_0} - \lambda_{y_3}) \beta \phi + \lambda_{y_0} \mu \\
\frac{d\lambda_{y_1}}{dt} &= -A_1 + (\lambda_{y_0} - \lambda_{y_1}) (1 - u_1) \frac{\beta_1 y_0}{1 + k y_1} + (\lambda_{y_1} - \lambda_{y_5}) (1 + u_2) \rho_1 + \lambda_{y_1} (\mu + \varepsilon_1 + \gamma) \\
&\quad - \lambda_{y_2} \gamma - \lambda_{y_3} \varepsilon_1 \\
\frac{d\lambda_{y_2}}{dt} &= -A_2 + (\lambda_{y_0} - \lambda_{y_1}) (1 - u_1) \frac{\beta_2 y_0}{1 + k y_2} + (\lambda_{y_1} - \lambda_{y_5}) (1 + u_2) \rho_2 + \lambda_{y_2} \mu - \lambda_{y_3} \varepsilon_2 \\
\frac{d\lambda_{y_3}}{dt} &= \lambda_{y_3} c - \lambda_{y_4} \alpha \\
\frac{d\lambda_{y_4}}{dt} &= \lambda_{y_4} d \\
\frac{d\lambda_{y_5}}{dt} &= \lambda_{y_5} \mu \\
\frac{d\lambda_{y_6}}{dt} &= \lambda_{y_6} \mu
\end{aligned}$$

The transversality conditions are $\lambda_{y_i}(t_T) = 0$ for $i = 0 \dots 6$. The optimality conditions require that $\frac{\partial H}{\partial u_1} = \frac{\partial H}{\partial u_2} = 0$ at $u_1 = u_1^*$, $u_2 = u_2^*$, where

$$u_1^* = \min \left\{ \max \left[0, (\lambda_{y_1} - \lambda_{y_0}) \left(\frac{\beta_1 y_0 y_1}{A_3(1 + k y_1)} + \frac{\beta_2 y_0 y_2}{A_3(1 + k y_2)} \right) \right], \max u_1 \right\}$$

$$u_2^* = \min \left\{ \max \left[0, (\lambda_{y_1} - \lambda_{y_5}) \frac{\rho_1 y_1}{A_4} + (\lambda_{y_2} - \lambda_{y_5}) \frac{\rho_2 y_2}{A_4} \right], \max u_2 \right\}$$

Figure 6.7 shows the positive impact that control measures ($u_1, u_2 \geq 0.4$) have on the population of moderate smokers, heavy smokers, lung cancer sufferers, lung cancer deaths and susceptibles. Figures 6.7(a)-6.7(b) show that the combined social and health policy measures to enhance smoking cessation can reduce the number of moderate and heavy smokers by up to 95%. This is 17% lower for moderate smokers and 15% lower for the heavy smokers compared to pre-control levels. Figures 6.7(c)-(d) show that these measures can reduce the lung cancer cases and deaths by about 50%.

Figure 6.8 shows the profiles of the treatments u_1 and u_2 .

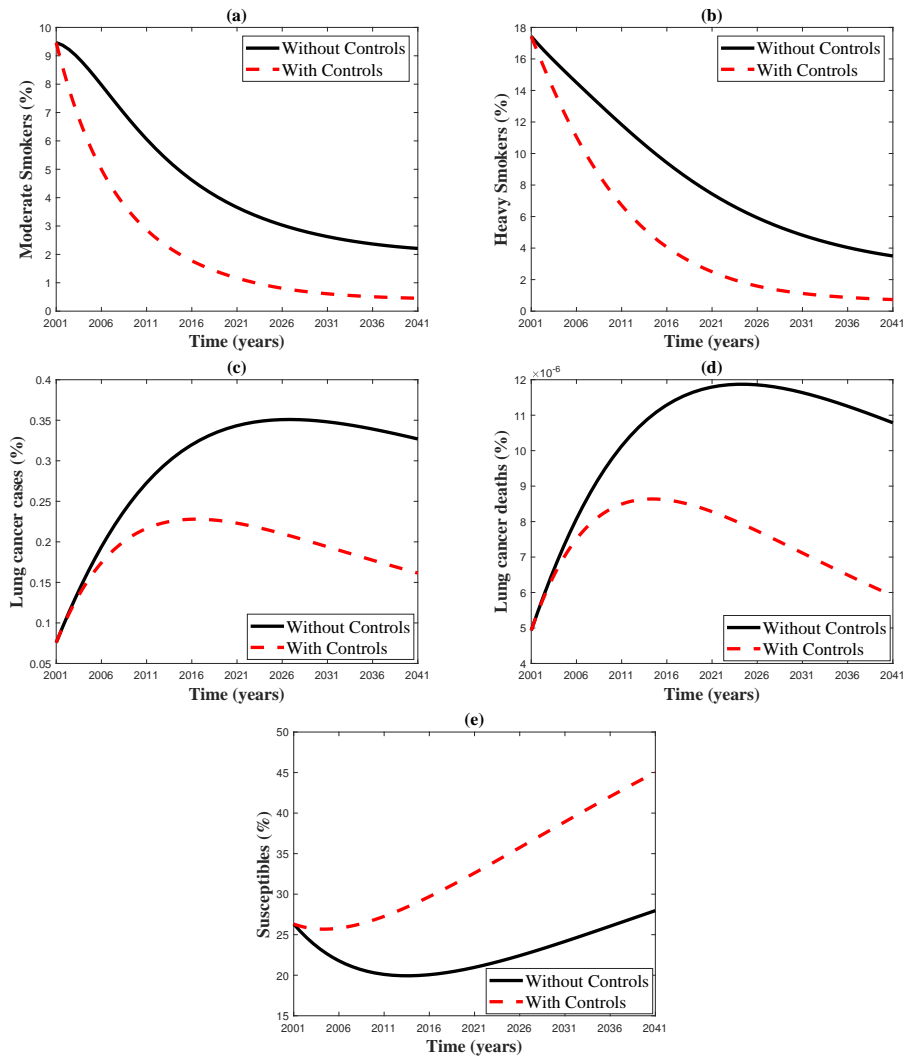


Figure 6.7: Optimal control diagrams showing the effect of control measures on (a) moderate smokers, (b) heavy smokers population (c) lung cancer cases, (d) lung cancer deaths and (e) susceptible population.

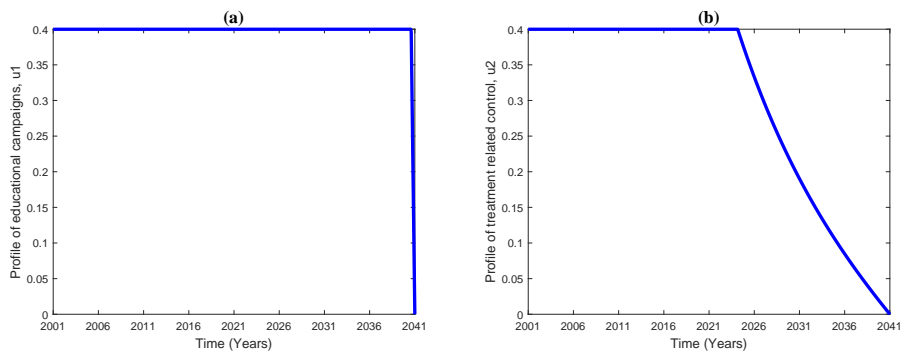


Figure 6.8: Profiles for controls representing (a) social and (b) health policy measures.

Chapter 7

Discussion

The good health of a population depends on a number of important factors such as disease outbreaks, their severity and transmissibility as well as to mount a plan to execute a response. As the COVID-19 pandemic has demonstrated, a key test for health systems worldwide is their resilience in simultaneously addressing emergencies and ongoing health needs [146]. Within a few months of the COVID-19 pandemic the documented rates demonstrated that hospitalization, added need for special care and deaths for people with pre-existing conditions like hypertension, diabetes, chronic respiratory diseases or cancer were quite high to the extent that facilities to attend to these needs were overwhelmed. This in itself highlights the convergence that occurs between non communicable diseases and infectious (communicable) diseases. The COVID-19 pandemic has shown that we can not deal with one and ignore the other.

During the early stages of pandemics, pharmaceutical interventions such as vaccines are not available, and in resource-limited settings, such interventions are rarely readily available or affordable. Moreover, due to the emergence of new viral strains /variants present challenges that render the pharmaceutical interventions obsolete in some cases. In such instances non pharmaceutical interventions become the most important public health interventions that governments, communities and/or individuals can adopt.

In this study we explored one of the key non pharmaceutical intervention used in disease management. The practical utility of the model we developed for contact tracing during an Ebola epidemic is its use in infectious disease manage-

ment such as COVID-19. One conclusion of our model is that behavior change over the course of an outbreak significantly impacts dynamics and should be considered when formulating models and management responses. It would be interesting to retrospectively analyze other past outbreaks, where data is available, to investigate the effect of allowing for time-dependent parameters. One could try to connect behavior change with specific information campaigns. Figure 4.5 shows clearly how a linear decrease in the amount of adequate contact tracing during an outbreak can result in a nonlinear increase in the number of cases and deaths. As a result, our time-dependent modeling approach can be used in future outbreaks to assess the amount of contact tracing that should be conducted in order to limit the total number of cases and deaths.

There was a period when the contact tracing infrastructure was overwhelmed by cases, as seen in the plateaus in Figures 4.4. More contact tracers available to work would have prevented this plateau, but the number of contact tracers available was sufficient to prevent many more cases and deaths from occurring. Increasing either κ_1 or κ_2 would have decreased the number of deaths that occurred, but κ_1 had a stronger effect than κ_2 . Overall this work makes a strong contribution to understanding the effects of contact tracing and changes in behavior on disease management.

In the future further exploration of the role of contact tracing in epidemics could be undertaken, by adding international spread features, through considering mobility data [147]. A model with a more realistic form of the function f which represents how contact tracing capacity grows in response to an epidemic could be developed. The mechanisms of contact tracing procedures for other diseases might be quite different and require the development of disease-specific models. Optimization techniques (such as optimal control) could be used to design management strategies for contact tracing.

Human respiratory virus infections lead to a spectrum of respiratory symptoms and disease severity, contributing to substantial morbidity, mortality and economic losses worldwide as seen in the COVID-19 pandemic [148]. The greatest challenge presented by these infectious diseases is their reliance on controlling infections through good hygiene and sanitary practices which most resource

limited countries are unable to meet as well as the fact that they can be transmitted through the air.

In order to understand how sub-Saharan Africa, South Africa as an example, could cope with the COVID-19 pandemic in view of limited resources and compromised health systems we explored different strategies that could be used to protect hospital bed capacity and help flatten the infection curve. The analysis revealed that there are several strategies that can flatten the infection curve and protect hospital bed capacity. However, we found that of all the strategies, the preferred strategy is either that which removes asymptomatic infectives much faster than the symptomatic infectives or the strategy that removes symptomatic infectives faster than asymptomatic infectives as either of these strategies gives the least number of deaths.

South Africa introduced the lockdown early. This is evident from the flatness of the infection curve (Figure 5.12), depicting that early lockdown slowed down the number of infections. The question is why despite this measure the number of infected individuals has kept on rising. The number of infections in fact rose rapidly after the lockdown. It is clear from Figure 5.14 that the lockdown was ineffective because the individuals in various households came out to mingle with individuals from other households at community places such as shops, markets etc. The lockdown did not have the desired effect. From Figure 5.14 it is evident that if household and community transmission had been avoided due to higher levels of adherence to control measures the situation would have resulted in 33% fewer asymptomatic infectives, 33% fewer symptomatic infectives and 62% less hospitalizations.

The difficulty for South Africa, and indeed any sub-Saharan African country, will be how to ensure the populations living in high density areas observe social distancing and good hygiene practices. It is easier in developed countries with social security arrangements to maintain disease prevention measures. In sub-Saharan Africa where the majority of the people live a subsistence life and are self employed, it will be hard to enforce the measures employed in developed countries.

Worldwide between 82000 and 99000 young people begin smoking everyday,

80% of them from low-middle income countries (LMICs). If current trends continue more than 200 million young people under the age of 20 will die prematurely from tobacco related diseases [149]. Adolescents's exposure to norms coming from their own social environment is the major factor responsible for their taking up early smoking. They are especially vulnerable to social influence as they try to fit in with their peers [150, 151]. In addition to other effects of cigarette smoking, early age at initiation of smoking increases risk of lung cancer [16]. Any effort towards reducing lung cancer incidence or mortality should start with promoting public health interventions that reduce smoking initiation for young people and encourages smoking cessation efforts.

Sensitivity analysis has revealed that parameters representing the rate at which susceptible individuals become smokers due to interaction with moderate smokers as well as the rate at which moderate smokers become addicted have the largest impact on the development of the smoking epidemic. Therefore, to reduce the number of smokers, in addition to stricter measures on tobacco advertising, tobacco pricing and programs offering help to quit tobacco use, this study recommends targeting susceptibles to avoid taking up smoking as a habit in the first place. Our model has shown that these measures can result in the decline of lung cancer cases and deaths in the long run. Because we used the data for Wales to fit our model, we want to caution that more publicly available data on smoking habits for other countries is needed to extend the findings of our study to a wider area.

We formulated a stochastic model and used it to predict smoking trends and its consequences using ONS data in Wales for the period 2001 – 2019. The stochastic solution revealed that the long-term outcome could be a quasi steady state, where all classes will persist but at low levels. The deterministic model suggests the existence of backward bifurcation which has the potential to complicate the decision making regarding the effectiveness of control measures of smoking and smoking lung cancer. This study recommends routine surveys to find out the recruitment levels of new susceptible smokers. These surveys can be paid for through value added tax on tobacco products. This levy of tobacco products can also be used to strengthen the existing counseling clinics. The pos-

sibility of backward bifurcation is a new result which relates the reproduction number to the recruitment rate of new potential smokers. This result suggests that the epidemic is sustained in two ways: first by the high influence rate of the smokers and secondly by the high number of individuals who succumb to peer influence. The health officials have the responsibility to monitor and protect the non smokers from being converted to smoking and from the effect of second hand smoke.

Our recommendation is for health officials to set targets towards the elimination/reduction of smoking cancer as is being done for climate change. This way the success of the campaigns will be quantifiable and the elimination tools can be refined depending on the level of success. Some of the countries growing tobacco as a cash crop will be reluctant to stop growing it but if incentives to change to other cash crops are offered, the production of tobacco can be reduced significantly. This will raise the price of tobacco and reduce consumption.

The optimal control analysis indicated that the effects of tobacco control interventions in the form of u_1 (educational campaigns) and u_2 (treatment options) can reduce the number of smokers by approximately 75% and the number of lung cancer cases and related deaths by about 50% in the long run. If this is coupled with the recommendation regarding the growing of tobacco the success suggested by the optimal control would be achievable.

Bibliography

- [1] Bigna JJ, Noubiap JJ. The rising burden of non communicable diseases in sub-Saharan Africa. *Lancet Global Health*. 2019 Oct;7(10):e1295-e1296. doi:10.1016/S2214-109X(19)30370-5. PMID:31537347
- [2] Yosef T. Prevalence and associated factors of chronic non-communicable diseases among cross-country truck drivers in Ethiopia. *BMC Public Health*. 2020 OCT;20, 1564. doi:10.1186/s12889-020-09646-w.
- [3] Centre for Disease Control and Prevention. [Internet] How COVID-19 spreads. [cited 2021 Sept 9] Available from: tools.cdc.gov/api/v2/resources/media/407478/content.
- [4] Excler JL, Saville M, Berkely S, Kim JH. Vaccine development for emerging infectious diseases. *Nat Med*. 2021 Apr;27(4): 591-600. doi:10.1038/s41591-021-01301-0. Epub 2021 Apr 12. PMID:33846611
- [5] Nikoloski Z, Alquinabet AM, Abdulrahman Alfawaz R, Almudarra SS, Herbst CH, El-Saharty S, Alsukait R, Algwizani A. COVID-19 and non-communicable diseases: evidence from a systematic literature review. *BMC Public Health*. 2021 Jun;21,1068. doi:10.1186/s12889-021-11116w
- [6] World Health Organisation. [Internet] Non Communicable Diseases Fact Sheet. [cited 2020 Aug 19] Available from: www.who.int/mediacentre/factsheets/fs355/en/
- [7] World Odometer. [Internet] COVID-19 coronavirus pandemic.[cited 2021 July 27] Available from: <https://www.worldometersinfo/coronavirus/>

- [8] United Nations [Internet] Good health and well being. [cited 2021 Aug 31] Available from: unric.org/en/sdg-3/#:text=Goal%203%3A%20Ensure%20healthy%20lives,is%20essential%20to%20sustainable%20development
- [9] Usman MS, Siddiqi TJ, Khan MS, Patel UK, Shahid I, Ahmed J, Kalra A, Michos ED (2020) Is there a smoker's paradox in COVID-19?. *BMJ Evidence-Based Medicine*. doi: 10.1136/bmjebm-2020-111492
- [10] Torre LA, Siegel RL, Ward EM, Jemal A. Global cancer incidence and mortality rates and trends-an update. *Cancer Epidemiol Biomarkers Prev* 2016 Jan;25(1):16-27. doi:10.1158/1055-9965.EPI-15-0578.
- [11] Anderson P, deBruijn A, Angus K, Ross G, Hastings G. Impact of alcohol advertising and media exposure on adolescent alcohol use a systematic review of longitudinal studies. *Alcohol Alcohol*. 2009 May-Jun 44(3):229-243. doi:10.1093/alcalc/agn115.
- [12] Gordon R, Hastings G, Moodie C. Alcohol marketing and young people's drinking: what the evidence base suggests for policy. *J Public Affairs*. 2009 Feb;10: 88-101. doi:10.1002/pa.338.
- [13] Ozturk O, Fidanci I. The Relationship between Smoking and Cancer: Mini Review. *Cancer Surgery*. 2016 Aug. doi:10.4172/2573-542X.1000108.
- [14] Tobacco Atlas [Internet] Prevalence. [cited 2021 June 20] Available from: tobaccoatlas.org/topic/prevalence/
- [15] Sheikh ZD, Branston JR, Gilmore AB. Tobacco industry pricing strategies in response to excise tax policies: a systematic review. *Tob Control*. 2021 Aug 9:tobaccocontrol-2021-056630. doi:10.1136/tobaccocontrol-2021-056630.
- [16] Leon ME, Peruga A, McNeill A, Kralikova E, Guha N, Minozzi S, Espina C, Schüz, J. European Code against Cancer (4th ed). *Cancer Epidemiol*. 2015 Dec;39Suppl1:S20-S33. doi:10.1016/j.canep.2015.06.001.

- [17] Global Cancer Observatory [Internet] World. [cited 2021 June 20] Available from: gco.iarc.fr/today/data/factsheets/populations/900-world-factsheets.pdf
- [18] The Lancet. COVID-19: A new lense for non communicable diseases. The Lancet Editorial Volume 396(10252). 2020 Sept; doi:10.1016/S0140-6736(20)31856-0.
- [19] World Health Organisation. [Internet] Coronavirus disease (COVID-19): How it is transmitted? [cited 2021 August 6] Available from: who.int/news-room/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted
- [20] Talisuna AO, Okiro EA, Yahaya AA, Stephen M, Bonkougou B, Musa EO, Minkoulou EM, Okeibunor J, Impouma B, Djingarey HM, Yao NKM, Oka S, Yoti Z, Fall IS. Spatial and temporal distribution of infectious disease epidemics, disasters and other potential public health emergencies in the World Health Organisation Africa region, 2016-2018. *Global Health*. 2020 Jan 15;16(1):9. doi: 10.1186/s12992-019-0540-4.
- [21] Fenollar F, Mediannikov O. Emerging infectious diseases in Africa in the 21st century. *New Microbes New Infect*. 2018 Sep 21;26:S10-S18. doi: 10.1016/j.nmni.2018.09.004.
- [22] Tangwa GB, Munung NS. COVID-19: Africa's relation with epidemics and some imperative ethics considerations of the moment. *Research Ethics*. 2020 Jun;16(3-4): 1-11. doi:10.1177/1747016120937391.
- [23] Emediegwu L, Oni E M. [Internet] Are there differences in Africa's responses to Ebola and COVID-19. [cited 2021 Aug 31] Available from: economicsobservatory.com/are-there-differences-in-africas-responses-to-ebola-and-covid-19
- [24] Simbine SL, Muridzo N, Chikadzi V, Mafa I. Responding to the COVID-19 pandemic in Zimbabwe. Lessons learnt from 2018 Cholera Outbreak. *Journal of social development in Africa*. 2021 Jun; Special issue.

- [25] World Health Organisation. [Internet] . WHO report on the global tobacco epidemic 2019: offer help to quit tobacco use. [cited 2021 August 6] Available from: <https://www.who.int/teams/health-promotion/tobacco-control/who-report-on-the-global-tobacco-epidemic-2019>
- [26] Cummings KM, Ballin S, Sweanor D. The past is not the future in tobacco control. *Prev Med.* 2020 Nov;140:106183. doi: 10.1016/j.ypmed.2020.106183.
- [27] Piñeros, M, Sierra MS, Forman D. Descriptive epidemiology of lung cancer and current status of tobacco control measures in Central and South America. *Cancer Epidemiol.* 2016 Sep;44 Suppl 1:S90-S99. doi: 10.1016/j.canep.2016.03.002.
- [28] Eames KT, Keeling MJ. Contact tracing and disease control. *Proceedings. Biological Sciences.* 2003 Dec;270(1533):2565-2571. DOI: 10.1098/rspb.2003.2554.
- [29] Wang V (2020) Modeling the testing and contact-tracing needed to suppress COVID-19. DOI: 10.26904/RF-137-1595947741.
- [30] Porco TC, Holbrook KA, Fernyak SE, Portnoy DL, Reiter R, Aragón TJ. Logistics of community smallpox control through contact tracing and ring vaccination: a stochastic network model. *BMC Public Health.* 2004 Aug 6;4:34. doi: 10.1186/1471-2458-4-34. PMID: 15298713; PMCID: PMC520756.
- [31] Armbruster B, Brandeau ML. Contact tracing to control infectious disease: when enough is enough. *Health Care Manag Sci.* 2007 Dec;10(4):341-55. doi: 10.1007/s10729-007-9027-6. PMID: 18074967; PMCID: PMC3428220.
- [32] Centre for Disease Control and Prevention. [Internet] Ebola (Ebola virus disease): 2014-2016 ebola outbreak in West Africa. [cited 2021 September 9] Available from: <https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/index.html>

- [33] Saurabh S, Prateek S. Role of contact tracing in containing the 2014 Ebola outbreak: a review. *Afr Health Sci.* 2017 Mar;17(1):225-236. doi: 10.4314/ahs.v17i1.28. PMID: 29026397; PMCID: PMC5636234.
- [34] Centre for Disease Control and Prevention. [Internet] Ebola (Ebola virus disease): Signs and Symptoms. [cited 2021 September 9] Available from: <https://www.cdc.gov/vhf/ebola/symptoms/index.html>
- [35] Müller J, Kretzschmar M. Contact tracing - Old models and new challenges. *Infect Dis Model.* 2020 Dec 30;6:222-231. doi: 10.1016/j.idm.2020.12.005. PMID: 33506153; PMCID: PMC7806945.
- [36] Chowell G, Nishiura H. Transmission dynamics and control of Ebola virus disease (EVD): a review. *BMC Med.* 2014 Oct 10;12:196. doi: 10.1186/s12916-014-0196-0.
- [37] Rivers CM, Lofgren ET, Marathe M, Eubank S, Lewis BL. Modeling the impact of interventions on an epidemic of ebola in sierra leone and liberia. *PLoS Curr.* 2014 Nov 6;6:ecurrents.outbreaks.4d41fe5d6c05e9df30ddce33c66d084c. doi: 10.1371/currents.outbreaks.4d41fe5d6c05e9df30ddce33c66d084c.
- [38] Browne C, Gulbudak H, Webb G. Modeling contact tracing in outbreaks with application to Ebola. *J Theor Biol.* 2015 Nov 7;384:33-49. doi: 10.1016/j.jtbi.2015.08.004. Epub 2015 Aug 18. PMID: 26297316.
- [39] Webb G, Browne C, Huo X, Seydi O, Seydi M, Magal P. A model of the 2014 ebola epidemic in west Africa with contact tracing. *PLoS Curr.* 2015 Jan 30;7:ecurrents.outbreaks.846b2a31ef37018b7d1126a9c8adf22a. doi: 10.1371/currents.outbreaks.846b2a31ef37018b7d1126a9c8adf22a. PMID: 25685636; PMCID: PMC4323422.
- [40] Hoffman SJ, Silverberg SL. Delays in Global Disease Outbreak Responses: Lessons from H1N1, Ebola, and Zika. *Am J Public Health.* 2018 Mar;108(3):329-333. doi: 10.2105/AJPH.2017.304245. Epub 2018 Jan 18. PMID: 29345996; PMCID: PMC5803810.

- [41] Largent EA. EBOLA and FDA: reviewing the response to the 2014 outbreak, to find lessons for the future. *J Law Biosci.* 2016 Sep 16;3(3):489-537. doi: 10.1093/jlb/lsw046. PMID: 28852537; PMCID: PMC5570698.
- [42] Stehling-Ariza T, Rosewell A, Moiba SA, Yorpie BB, Ndomaina KD, Jimissa KS, Leidman E, Rijken DJ, Basler C, Wood J, Manso D. The impact of active surveillance and health education on an Ebola virus disease cluster - Kono District, Sierra Leone, 2014-2015. *BMC Infect Dis.* 2016 Oct 27;16(1):611. doi: 10.1186/s12879-016-1941-0. PMID: 27784275; PMCID: PMC5082353.
- [43] Medecins Sans Frontier (MSF). [Internet] Six lessons learned as Ebola outbreak in northeastern DRC ends. [cited 2021 October 6] Available from: <https://www.msf.org/Six-lessons-learned-drc-ebola-outbreak-ends>
- [44] Olu OO, Lamunu M, Nanyunja M, et al. Contact Tracing during an Outbreak of Ebola Virus Disease in the Western Area Districts of Sierra Leone: Lessons for Future Ebola Outbreak Response. *Frontiers in Public Health.* 2016 ;4:130. DOI: 10.3389/fpubh.2016.00130. PMID: 27446896; PMCID: PMC4916168.
- [45] Swanson KC, Altare C, Wesseh CS, Nyenswah T, Ahmed T, Eyal N, Hamblion EL, Lessler J, Peters DH, Altmann M. Contact tracing performance during the Ebola epidemic in Liberia, 2014-2015. *PLoS Negl Trop Dis.* 2018 Sep 12;12(9):e0006762. doi: 10.1371/journal.pntd.0006762. PMID: 30208032; PMCID: PMC6152989.
- [46] Green A. WHO and partners launch Ebola response plan. *Lancet.* 2014 Aug;384:481. doi:10.1016/S0140-6736(14)61322-2
- [47] Deressa CT, Duressa GF. Modeling and optimal control of transmission dynamics of COVID-19: The case of Ethiopia. *Alexandria Eng. J.* 2021 Feb;60(1):719-732. doi:10.1016/j.aej.2020.10.004
- [48] Garba SM, Lubuma JM, Tsanou B. Modeling the transmission dynamics of the COVID-19 Pandemic in South Africa. *Math Biosci.* 2020 Oct;328:108441. doi: 10.1016/j.mbs.2020.108441. Epub 2020 Aug 4. PMID: 32763338; PMCID: PMC7402282.

- [49] Kassa SM, Njagarah JBH, Terefe YA. Analysis of the mitigation strategies for COVID-19: From mathematical modelling perspective. *Chaos Solitons Fractals*. 2020 Sep;138:109968. doi: 10.1016/j.chaos.2020.109968. Epub 2020 Jun 5. PMID: 32536760; PMCID: PMC7274644.
- [50] Mukandavire Z, Nyabadza F, Malunguza NJ, Cuadros DF, Shiri T, Musuka G. Quantifying early COVID-19 outbreak transmission in South Africa and exploring vaccine efficacy scenarios. *PLoS One*. 2020 Jul 24;15(7):e0236003. doi: 10.1371/journal.pone.0236003. PMID: 32706790; PMCID: PMC7380646.
- [51] Ngonghala CN, Iboi E, Eikenberry S, Scotch M, MacIntyre CR, Bonds MH, Gumel AB. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. *Math Biosci*. 2020 Jul;325:108364. doi: 10.1016/j.mbs.2020.108364. Epub 2020 May 1. PMID: 32360770; PMCID: PMC7252217.
- [52] Ahmed N, Maqsood A, Abduljabbar T, Vohra F. Tobacco Smoking a Potential Risk Factor in Transmission of COVID-19 Infection. *Pak J Med Sci*. 2020 May;36(COVID19-S4):S104-S107. doi: 10.12669/pjms.36.COVID19-S4.2739. PMID: 32582324; PMCID: PMC7306971.
- [53] Zeb A, Alzahrani A. Non-standard finite difference scheme and analysis of smoking model with reversion class. *Results Phys*. 2021 Feb;21:103785. doi: 10.1016/j.rinp.2020.103785. Epub 2021 Jan 6. PMID: 33816094; PMCID: PMC8009654.
- [54] Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, Curtis HJ, Mehrkar A, Evans D, Inglesby P, Cockburn J, McDonald HI, MacKenna B, Tomlinson L, Douglas IJ, Rentsch CT, Mathur R, Wong AYS, Grieve R, Harrison D, Forbes H, Schultze A, Croker R, Parry J, Hester F, Harper S, Perera R, Evans SJW, Smeeth L, Goldacre B. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020 Aug;584(7821):430-436. doi: 10.1038/s41586-020-2521-4. Epub 2020 Jul 8. PMID: 32640463; PMCID: PMC7611074.

- [55] Miyara M, Tubach F, Pourcher V, Morelot-Panzini C, Pernet J, Haroche J. Low incidence of daily active tobacco smoking in patients with symptomatic COVID-19. *Qeios*. 2020 Jun;1-13. doi:10.1101/2020.06.10.20127514
- [56] Tonnesen P, Marott JL, Nordestgaard B, Bojesen SE, Lange P. Secular trends in smoking in relation to prevalent and incident smoking-related disease: A prospective population-based study. *Tob Induc Dis*. 2019 Oct 7;17:72. doi: 10.18332/tid/112459. PMID: 31768164; PMCID: PMC6830353.
- [57] Castillo-Garsow C, Jordàn-Salivia G, Rodriguez-Herrera A. Mathematical models for the dynamics of tobacco use, recovery, and relapse. [Internet] 1997 [cited 2020 August 19] Available from: ecommons.cornell.edu/bitstream/handle/1813/32095/BU-1505-M.pdf;sequence=1
- [58] Sharomi O, Gumel AB. Curtailing smoking dynamics: A mathematical modeling approach. *Applied Mathematics and Computation*. 2008 Feb;195:475-499. doi:10:1016/j.amc.2007:05:012
- [59] Zeb A, Zaman G, Momani S. Square-root dynamics of a giving up smoking model. *Applied Mathematical Modelling*. 2013 Apr;37:5326-5384. doi:10.1016/j.apm.2012.10.005
- [60] Alkhudari Z, Al-Sheikh S, Al-Tuwairqi S. The effect of heavy smokers on the dynamics of a smoking model. *International Journal of Differential Equations and Applications*. 2015 Jan;14(4):343-356. doi:10.127321ijdea.v14:4.2577
- [61] Sikander W, Khan U, Ahmed N, Mohyud-Din S T. Optimal solutions for a bio mathematical model for the evolution of smoking habit. *Results in Physics*. 2017 Jan;7(2017):510-517. doi:10.1016j.rinp.2017.01.001
- [62] Ullah, R, Khan M, Zaman G, Islam S, Khan MA, Jan S, Gul T. Dynamical features of a mathematical model on smoking. *J. Appl. Environ. Biol. Sci*. 2016 Jan;6(1):92-96. doi:10.1007/978-3-319-24223-1.

- [63] Matintu SD. Smoking as Epidemic: Modelling and simulation study. *American Journal of Applied Mathematics*. 2017 Feb;5(1):31 – 38. doi:10.11648/j.ajam20170501.14
- [64] Ihsanjaya MMM, Susyanto N. Mathematical model of changes in smoking behavior which involves smokers who temporarily and permanently quit smoking. *AIP Conference Proceedings* 2192,060011. doi:10.1063/1.5139157.
- [65] Verma V, Agarwal M. Global dynamics of a mathematical model on smoking with media campaigns. *Research Desk*. 2015;4(1):500-512.
- [66] Sun C, Jia J. Optimal control of a delayed smoking model with immigration. *Journal of Biological Dynamics*. 2019 Jun;13(1):447-460. doi:10.1080/17513758.2019.1629031.
- [67] Verma V, Bhadauria AS. Global dynamics of a mathematical model on smoking: impact of anti-smoking campaign. *Journal of Mathematical Modeling*. 2019 Mar;7(1):49-62. doi:10.22124/jmm.2018.10117.1153.
- [68] Alzahrani E, Zeb A. Stability analysis and prevention strategies of tobacco smoking model. *Boundary Value Problems*. 2020 Jan;2020:3. doi:10.1186/s13661-019-01315-1.
- [69] Fatimah M, Aldila D, Handari BD. Backward bifurcation arises from the smoking transmission model considering media campaigns. *Journal of Physics: Conference Series* 1722(2021)012004. doi:10.1088/1742-6596/1722/1/012004.
- [70] Centre for Disease Control and Prevention. [Internet] Smoking and Cancer:Overviews of Diseases/Conditions:Tips from former smokers: CDC. [cited 2020 August 19] Available from: <https://www.cdc.gov/tobacco/campaign/tips/diseases/cancer.html>
- [71] Ma Y, Li MD. Establishment of a Strong Link Between Smoking and Cancer Pathogenesis through DNA Methylation Analysis. *Sci Rep*. 2017 May

- 12;7(1):1811. doi: 10.1038/s41598-017-01856-4. PMID: 28500316; PMCID: PMC5431893.
- [72] Darby SC, Pike MC. Lung cancer and passive smoking: predicted effects from a mathematical model for cigarette smoking and lung cancer. *Br J Cancer*. 1988 Dec;58(6):825-31. doi: 10.1038/bjc.1988.319. PMID: 3224084; PMCID: PMC2246877.
- [73] Acevedo-Estephanian CA, Gonzalez K, Rios-Soto KR, Summerville ED. A mathematical model for lung cancer: The effects of second hand smoke and education. [Internet] 1999 [cited 2020 June 20] Available from: <https://mtbi.asu.edu/sites/default/files/>
- [74] Ott W H, Steinemann AC, Wallace LA. Exposure Analysis. [Internet] 2006 [cited 2020 June 20] Available from: repace.com/pdf/EXPOSURE_TO_SECONDHAND_SMOKE.pdf
- [75] Andest JN. A mathematical model on cigarette smoking and nicotine in the lung. *International Journal of Engineering and Science*. 2013;2(6):01-03.
- [76] Sebastian E, Victor P. A discrete time mathematical model on Lung cancer incorporating smokers and non-smokers. *International Journal on Recent and Innovation trends in Computing and Communication*. 2017;5(5):1260-1265.
- [77] Trisilowati. Stability analysis and optimal control of lung cancer growth model with education. *IOP Conf.Series: Materials Science and Engineering*. 2019 Jun;546(2019) 052081. doi:10.1088/1757-899X/546/052081.
- [78] Ahmed J, Biswas HA. Mathematical modeling and analysis of the effect of smoking for the dynamics of lung cancer. *Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management Singapore, March 7-11 2021*. [Internet] 2021 [cited 2021 July 28] Available from: ieomsociety.org/singapore2021/papers/251.pdf
- [79] Martcheva M. An introduction to mathematical epidemiology. *Texts in Applied Mathematics* 2015;61. doi:10.1007/978-1-4899-7612-3

- [80] van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math Biosci.* 2002 Nov-Dec;180:29-48. doi: 10.1016/s0025-5564(02)00108-6. PMID: 12387915.
- [81] Diekmann o, Heesterbeek JAP. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analyis and Interpretation.* 1999. Wiley, New York.
- [82] Nemytskii VV, Stepanov VV. *Qualitative theory of differential equations.* 1960. Princeton University Press, New Jersey.
- [83] Chicone C. *Ordinary differential equations with applications.* 2006 Vol 34 of *Texts in Applied Mathematics*, springer, New York, second ed.
- [84] Birkoff G and Rota GC. *Ordinary differential equations.* 1978 Ginn Boston.
- [85] Boyce W E and Di Prima R C (2009) *Elementary differential equations and boundary value problems.* Ninth ed., John Wiley & Sons, Inc.
- [86] Perko L. *Differential equations and dynamical systems.* 2001 Third ed., Springer-Verlag, New York.
- [87] Kamgang JC, Sallet G. Computation of threshold conditions for epidemiological models and global stability of the disease-free equilibrium (DFE). *Math Biosci.* 2008 May;213(1):1-12. doi: 10.1016/j.mbs.2008.02.005. Epub 2008 Feb 23. PMID: 18405926.
- [88] Gomero, B.: Latin hypercube sampling and partial rank correlation coefficient analysis applied to an optimal control problem (2012)
- [89] Pereira A, Broed R. *Methods for uncertainty and sensitivity analysis: review and recomendations for implementation in ecology.* Fysikum, Stockholm. [Internet] 2006 [cited 2021 Oct 20]. Available from: <http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-1079>
- [90] Islam MR, Peace AL. *Parameter Sensitivity Analysis and Control Strategies of a Multi-Stage Epidemic Model.* Poster, Conference: Multiscale Dy-

namics of Infections, Ohio State University, Columbus, OH, 23-27 April 2018.

- [91] Blower SM, Dowlatabadi H. Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: An HIV Model, as an Example. *International Statistical Review*. 1994 Aug;62:229-243. doi:10.2307/1403510
- [92] Wu J, Dhingra R, Gambhir M, Remais JV. Sensitivity analysis of infectious disease models: methods, advances and their application. *Journal of the Royal Society, Interface*. 2013 Sep;10(86):20121018. DOI: 10.1098/rsif.2012.1018. PMID: 23864497; PMCID: PMC3730677.
- [93] Clarke F H. *Optimal control and nonsmooth analysis*. 1990 SIAM, Philadelphia.
- [94] Lenhart S, Workman JT. *Optimal Control Applied to Biological Problems*. Chapman & Hall/CRC, Boca Raton, Fla, USA, 2007.
- [95] Oksendal B. *Stochastic differential equations: an introduction with applications*. 2013 Springer Science & Business media.
- [96] Burton D, Lenhart S, Edholm C, Levy B, Washington ML, Greening BR, White KAJ, Lungu E, Chimbola O, Kgosimore M, Chirove F, Ronoh M, Machingauta MH. A mathematical model of contact tracing during the 2014-2026 West African Ebola out break. *Mathematics*. 2021 Mar;9:608. doi:10.3390/math9060608.
- [97] Frieden TR, Damon I, Bell BP, Kenyon T, Nichol S. Ebola 2014—new challenges, new global response and responsibility. *N Engl J Med*. 2014 Sep 25;371(13):1177-80. doi: 10.1056/NEJMp1409903. Epub 2014 Aug 20. PMID: 25140858. doi:10.1056/NEJMp1002530.
- [98] WHO Ebola Response Team, Aylward B, Barboza P, Bawo L, Bertherat E, Bilivogui P, Blake I, Brennan R, Briand S, Chakauya JM, Chitala K, Conteh RM, Cori A, Croisier A, Dangou JM, Diallo B, Donnelly CA, Dye C, Eckmanns T, Ferguson NM, Formenty P, Fuhrer C, Fukuda K, Garske T, Gasasira A, Gbanyan S, Graaff P, Heleze E, Jambai A, Jombart T, Kasolo

- F, Kadiobo AM, Keita S, Kertesz D, Koné M, Lane C, Markoff J, Massaquoi M, Mills H, Mulba JM, Musa E, Myhre J, Nasidi A, Nilles E, Nouvellet P, Nshimirimana D, Nuttall I, Nyenswah T, Olu O, Pendergast S, Perea W, Polonsky J, Riley S, Ronveaux O, Sakoba K, Santhana Gopala Krishnan R, Senga M, Shuaib F, Van Kerkhove MD, Vaz R, Wijekoon Kannangarage N, Yoti Z. Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections. *N Engl J Med*. 2014 Oct 16;371(16):1481-95. doi: 10.1056/NEJMoa1411100. Epub 2014 Sep 22. PMID: 25244186; PMCID: PMC4235004.
- [99] Bogoch II, Creatore MI, Cetron MS, Brownstein JS, Pesik N, Miniota J, Tam T, Hu W, Nicolucci A, Ahmed S, Yoon JW, Berry I, Hay SI, Anema A, Tatem AJ, MacFadden D, German M, Khan K. Assessment of the potential for international dissemination of Ebola virus via commercial air travel during the 2014 west African outbreak. *Lancet*. 2015 Jan 3;385(9962):29-35. doi: 10.1016/S0140-6736(14)61828-6. Epub 2014 Oct 21. Erratum in: *Lancet*. 2015 Jan 3;385(9962):28. PMID: 25458732; PMCID: PMC4286618.
- [100] Brainard J, Hooper L, Pond K, Edmunds K, Hunter PR. Risk factors for transmission of Ebola or Marburg virus disease: a systematic review and meta-analysis. *Int J Epidemiol*. 2016 Feb;45(1):102-16. doi: 10.1093/ije/dyv307. Epub 2015 Nov 20. PMID: 26589246; PMCID: PMC4795563.
- [101] Dietz PM, Jambai A, Paweska JT, Yoti Z, Ksiazek TG. Epidemiology and risk factors for Ebola virus disease in Sierra Leone-23 May 2014 to 31 January 2015. *Clin Infect Dis*. 2015 Dec 1;61(11):1648-54. doi: 10.1093/cid/civ568. Epub 2015 Jul 15. Erratum in: *Clin Infect Dis*. 2016 Mar 1;62(5):673. Ksiazek, Thomas G [corrected to Ksiazek, Thomas G]. PMID: 26179011.
- [102] Drake JM, Bakach I, Just MR, O'Regan SM, Gambhir M, Fung IC. Transmission Models of Historical Ebola Outbreaks. *Emerg Infect Dis*. 2015 Aug;21(8):1447-50. doi: 10.3201/eid2108.141613. PMID: 26196358; PMCID: PMC4517740.

- [103] Gomes MF, Pastore Y Piontti A, Rossi L, Chao D, Longini I, Halloran ME, Vespignani A. Assessing the international spreading risk associated with the 2014 west african ebola outbreak. *PLoS Curr.* 2014 Sep 2;6:ecurrents.outbreaks.cd818f63d40e24aef769dda7df9e0da5. doi: 10.1371/currents.outbreaks.cd818f63d40e24aef769dda7df9e0da5. PMID: 25642360; PMCID: PMC4169359.
- [104] Skrip LA, Fallah MP, Gaffney SG, et al. Characterizing risk of Ebola transmission based on frequency and type of case-contact exposures. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences.* 2017 May;372(1721). DOI: 10.1098/rstb.2016.0301. PMID: 28396472; PMCID: PMC5394639.
- [105] Li SL, Bjørnstad ON, Ferrari MJ, Mummah R, Runge MC, Fonnesebeck CJ, Tildesley MJ, Probert WJM, Shea K. Essential information: Uncertainty and optimal control of Ebola outbreaks. *Proc Natl Acad Sci U S A.* 2017 May 30;114(22):5659-5664. doi: 10.1073/pnas.1617482114. Epub 2017 May 15. PMID: 28507121; PMCID: PMC5465899.
- [106] Ebola Situation Report, Ministry of Health and Sanitation, Sierra Leone. web.archive.org/web/20150314233800, 2015.
- [107] Ebola Situation Report, Ministry of Health and Sanitation, Sierra Leone. web.archive.org/web/2016050901463
- [108] Senga M, Pringle K, Ramsay A, Brett-Major DM, Fowler RA, French I, Vandi M, Sellu J, Pratt C, Saidu J, Shindo N, Bausch DG; Sierra Leone Kenema District Task Force and Kenema Government Hospital. Factors Underlying Ebola Virus Infection Among Health Workers, Kenema, Sierra Leone, 2014-2015. *Clin Infect Dis.* 2016 Aug 15;63(4):454-9. doi: 10.1093/cid/ciw327. Epub 2016 May 18. PMID: 27193749; PMCID: PMC4967603.
- [109] Lokuge K, Caleo G, Greig J, Duncombe J, McWilliam N, Squire J, Lamin M, Veltus E, Wolz A, Kobinger G, de la Vega MA, Gbabai O, Nabieu S, Lamin M, Kremer R, Danis K, Banks E, Glass K. Successful Control of

Ebola Virus Disease: Analysis of Service Based Data from Rural Sierra Leone. *PLoS Negl Trop Dis*. 2016 Mar 9;10(3):e0004498. doi: 10.1371/journal.pntd.0004498. PMID: 26959413; PMCID: PMC4784943.

[110] Washington, M. personal communication, 2017.

[111] Diekmann, O. *Mathematical epidemiology of infectious diseases : model building, analysis, and interpretation*; Wiley series in mathematical and computational biology, John Wiley: Chichester, 2000.

[112] Diekmann, O.; Heesterbeek, H.; Britton, T. *Mathematical tools for understanding infectious disease dynamics*; Vol. 7, Princeton University Press, 2012.

[113] van den Driessche, P.; Watmough, J. Further notes on the basic reproduction number. *Lecture Notes in Mathematics* 2008, 1945, 159–178. doi:10.1007/978-3-540-78911-6_6.

[114] CDC. *Increases in Heroin Overdose Deaths — 28 States, 2010 to 2012*. 2014. 63, 849–854.

[115] Levy B, Edholm C, Gaoue O, Kaondera-Shava R, Kgosimore M, Lenhart S, Lephodisa B, Lungu E, Marijani T, Nyabadza F. Modeling the role of public health education in Ebola virus disease outbreaks in Sudan. *Infect Dis Model*. 2017 Jun 29;2(3):323-340. doi: 10.1016/j.idm.2017.06.004. PMID: 29928745; PMCID: PMC6001965.

[116] World Odometer. [Internet] COVID-19 coronaviruss pandemic. [cited 2021 Nov 19] Available from <https://www.worldometersinfo/coronavirus/2-utmcampaign=homeAdUOa?Si>

[117] McKinsey & Company. [Internet] Covid-19: Briefing materials. Global health and crisis response, March 25, 2020. [cited 2020 Apr 22] Available from <https://www.mckinsey.com/~/media/McKinsey/BusinessFunctions/Risk/OurInsights/Covid19Implicationsforbusiness/Covid19March30/COVID-19-Facts-and-Insights-March-25-v5.ashx>

- [118] World Health Organisation. [Internet] Health situation analysis in the African region. atlas of health statistics, 2011. [cited 2020 Apr 7] Available from: <https://apps.who.int/medicinedocs/documents/s18863en/s11863en.pdf>
- [119] Kombe IK, Munywoki PK, Baguelin M, Nokes DJ, Medley GF. Model-based estimates of transmission of respiratory syncytial virus within households. *Epidemics*. 2019 Jun;27:1-11. DOI: 10.1016/j.epidem.2018.12.001. PMID: 30591267; PMCID: PMC6543068.
- [120] Berge T, Lubuma JM, Moremedi GM, Morris N, Kondera-Shava R. A simple mathematical model for Ebola in Africa. *J Biol Dyn*. 2017 Dec;11(1):42-74. doi: 10.1080/17513758.2016.1229817. PMID: 29067875.
- [121] Pulla P. What counts as a covid-19 death? *BMJ*. 2020 Jul 17;370:m2859. doi: 10.1136/bmj.m2859. PMID: 32680851.
- [122] Veeranna CH, Rani S. Cause of Death Certification in COVID-19 Deaths. *Indian J Crit Care Med*. 2020 Sep;24(9):863-867. doi: 10.5005/jp-journals-10071-23561. PMID: 33132574; PMCID: PMC7584845.
- [123] World Health Organisation. [Internet] Rolling updates on coronavirus disease (COVID-19). [cited 2020 March 23] Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>
- [124] Worldbank. [Internet] The World Bank life expectancy at birth, total (years) - Italy. [cited 2020 June 29] Available from: <https://data.worldbank.org/indicator/SP.DYN.LE00IN?locations=IT>
- [125] Rhodes A, Ferdinande P, Flaatten H, Guidet B, Metnitz PG, Moreno RP. The variability of critical care bed numbers in Europe. *Intensive Care Med*. 2012 Oct;38(10):1647-53. doi: 10.1007/s00134-012-2627-8. Epub 2012 Jul 10. PMID: 22777516.

- [126] Worldbank. [Internet] The World Bank life expectancy at birth, total (years) - South Africa. [cited 2020 June 29] Available from: <https://data.worldbank.org/indicator/SP.DYN.LE00IN?locations=ZA>
- [127] Dell AJ, Kahn D. Geographical maldistribution of surgical resources in South Africa: A review of the number of hospitals, hospital beds and surgical beds. *S Afr Med J*. 2017 Nov 27;107(12):1099-1105. doi: 10.7196/SAMJ.2017.v107i12.12539. PMID: 29262964.
- [128] Thun MJ, Henley SJ, Calle EE. Tobacco use and cancer: an epidemiologic perspective for geneticists. *Oncogene*. 2002 Oct 21;21(48):7307-25. doi: 10.1038/sj.onc.1205807. PMID: 12379875.
- [129] Miller YE. Pathogenesis of lung cancer: 100 year report. *Am J Respir Cell Mol Biol*. 2005 Sep;33(3):216-23. doi: 10.1165/rcmb.2005-0158OE. PMID: 16107574; PMCID: PMC2715312.
- [130] Proctor RN. The history of the discovery of the cigarette-lung cancer link: evidentiary traditions, corporate denial, global toll. *Tob Control*. 2012 Mar;21(2):87-91. doi: 10.1136/tobaccocontrol-2011-050338. Erratum in: *Tob Control*. 2013 Jan;22(1):62. PMID: 22345227.
- [131] Bade BC, Dela Cruz CS. Lung Cancer 2020: Epidemiology, Etiology, and Prevention. *Clin Chest Med*. 2020 Mar;41(1):1-24. doi: 10.1016/j.ccm.2019.10.001. PMID: 32008623.
- [132] Welsh Cancer Intelligence and Surveillance Unit, Lung cancer in Wales:A detailed analysis of population trends of incidence and stage at diagnosis up to and including 2012. (2015b) [Online article]: cited on 6 November 2020, ISBN:978-0-9928835-6-0.
- [133] Welsh Cancer Intelligence and Surveillance Unit. [Internet] Lung cancer. [cited 2020 Nov 6] Available from <https://phw.nhs.wales/services-and-teams/welsh-cancer-intelligence-and-surveillance-unit-wcisu/awareness-posters-and-information/lung-cancer/>

- [134] Welsh Government. [Internet] Tobacco control delivery plan for Wales 2017 - 2020. [cited 2020 Nov 6] Available from: <https://gov.wales/sites/default/files/publications/2017-11/tobacco-control-delivery-plan-for-wales-2017-to-2020>.
- [135] Scottish Government [Internet] Scottish Health Survey 2018: Main report-revised 2020. [cited 2020 Nov 6] Available from: <https://www.gov.scot/publications/scottish-health-survey-2018-volume-1-main-report/pages/33/>
- [136] Office for National Statistics. [Internet] Life expectancy at birth and at age 65 by local areas in England and Wales: 2011 to 2013. (2012) [cited 2020 Nov 6] Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/lifeexpectancyatbirthandatage65bylocalareasinenglandandwales/previousReleases>.
- [137] Office for National Statistics. [Internet] Census: population and household estimates for Wales, March 2011. [cited 2020 Nov 6] Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/2011censuspopulationandhouseholdestimatesforwalesmarch2011>.
- [138] Welsh Cancer Intelligence and Surveillance Unit. [Internet] Lung cancer in Wales-Lung cancer survival and survival by stage. [cited 2020 Nov 6] ISBN:978-0-9928835-8-4.
- [139] Cancer Research UK. [Internet] Lung cancer survival statistics. [cited 2020 Aug 19] Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lungcancer/survival#heading-Zero>.
- [140] R. C. Ngeleja, L. Luboobi and Y. Nkansah-Gyekye, Stability Analysis of Bubonic Plague Model with the Causing Pathogen *Yersinia pestis* in the Environment. *Advances in Infectious Diseases*. 2016 Sept;6(2016):120-137. doi:10.4326/aid.2016.63016

- [141] C. Castillo-Chavez, Z. Feng and W. Huang. On the computation of R_0 and its role on global stability. [Internet] 2002 [cited 2020 June 20] Available from: math.la.asu.edu/chavez2002/JB276.pdf
- [142] Buonomo B, Lacitignola D. On the backward bifurcation of a vaccination model with nonlinear incidence. *Nonlinear Anal. Model. Control.* 2011 Jan;16(1):30-46. doi:10.15388/NA.16.1.14113
- [143] Castillo-Chavez C, Song B. Dynamical models of tuberculosis and their applications. *Math Biosci Eng.* 2004 Sep;1(2):361-404. doi: 10.3934/mbe.2004.1.361. PMID: 20369977.
- [144] Office for National Statistics. [Internet] Smoking prevalence in each country of the UK by age and sex, 2011 to 2019. [cited 2020 Nov 6] Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectanciesand-methodology>
- [145] G. Fowler. Proven strategies for smoking cessation: Adopting a global approach. *European Journal of Public Health.* 2000 Sept;10(3) :3-4. doi:10.1093/eurpub/10.suppl_3.3.
- [146] Hick JL, Biddinger PD. Novel Coronavirus and Old Lessons - Preparing the Health System for the Pandemic. *The New England Journal of Medicine.* 2020 May;382(20):e55. DOI: 10.1056/nejmp2005118. PMID: 32212515.
- [147] Halloran ME, Vespignani A, Bharti N, Feldstein LR, Alexander K, Ferrari M, Shaman J, Drake JM, Porco T, Eisenberg JN, et al. Ebola Mobility Data. *Science.* 2014;346:433.
- [148] Leung NHL. Transmissibility and transmission of respiratory viruses. *Nat Rev Microbiol.* 2021 Aug;19(8):528-545. doi: 10.1038/s41579-021-00535-6. Epub 2021 Mar 22. PMID: 33753932; PMCID: PMC7982882.

-
- [149] Hymowitz N. Cigarette Smoking and Lung Cancer: Pediatric Roots. *Lung Cancer Int.* 2012;2012:790841. doi: 10.1155/2012/790841. Epub 2012 Aug 30. PMID: 26316938; PMCID: PMC4437397.
- [150] Aral S, Walker D. Identifying influential and susceptible members of social networks. *Science.* 2012 JUL;337(6092):337-341. doi:10.1126/science.1215842
- [151] Gardner M, Steinberg L. Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Dev Psychol.* 2005 Jul;41(4):625-35. doi: 10.1037/0012-1649.41.4.625. Erratum in: *Dev Psychol.* 2012 Mar;48(2):589. PMID: 16060809.

Appendix

Appendix A

Initial Fitting Results

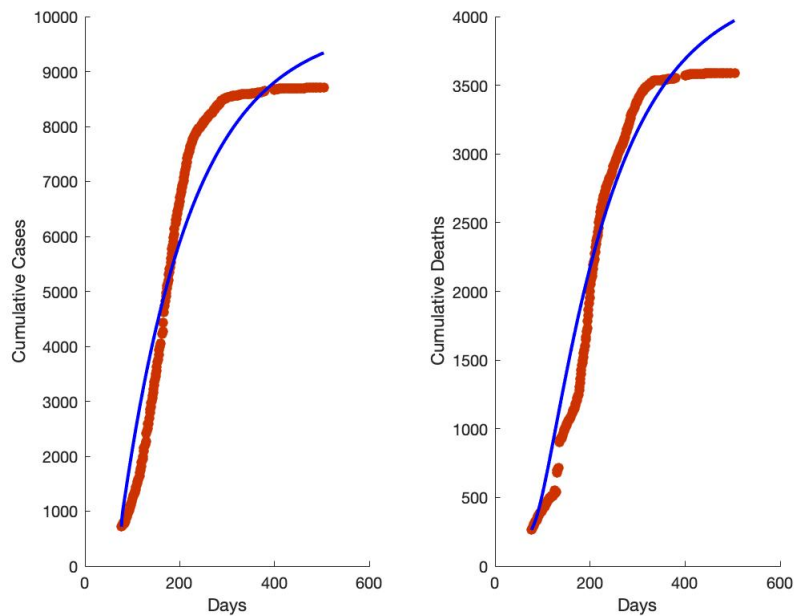


Figure 7.1: First attempt match to the data of cumulative cases and cumulative deaths with all parameters constant. The value of J is 0.1963.

Note that in Figure 7.1 the curves are still increasing at day 500, indicating that the epidemic would have continued.

Appendix B

Date	Day	Cumulative Cases	Cumulative Deaths	Date	Day	Cumulative Cases	Cumulative Deaths
12-Aug-14	77	717	264	11-Feb-15	260	8183	3009
13-Aug-14	78	733	273	12-Feb-15	261	8193	3018
14-Aug-14	79	747	280	13-Feb-15	262	8208	3030
15-Aug-14	80	757	287	14-Feb-15	263	8213	3036
16-Aug-14	81	775	297	15-Feb-15	264	8226	3043
17-Aug-14	82	778	305	16-Feb-15	265	8230	3050
18-Aug-14	83	783	312	17-Feb-15	266	8237	3058
19-Aug-14	84	804	320	18-Feb-15	267	8239	3063
20-Aug-14	85	813	322	19-Feb-15	268	8244	3066
21-Aug-14	86	823	329	20-Feb-15	269	8260	3079
22-Aug-14	87	881	333	21-Feb-15	270	8275	3088
23-Aug-14	88	904	336	22-Feb-15	271	8289	3095
24-Aug-14	89	935	341	23-Feb-15	272	8301	3103
25-Aug-14	90	955	355	24-Feb-15	273	8308	3113
26-Aug-14	91	961	363	25-Feb-15	274	8320	3124
27-Aug-14	92	988	372	27-Feb-15	276	8349	3151
28-Aug-14	93	1018	377	28-Feb-15	277	8353	3164
29-Aug-14	94	1033	383	1-Mar-15	278	8370	3180
30-Aug-14	95	1077	387	2-Mar-15	279	8374	3188
31-Aug-14	96	1106	388	3-Mar-15	280	8383	3199
1-Sep-14	97	1115	396	4-Mar-15	281	8389	3210
2-Sep-14	98	1146	399	5-Mar-15	282	8398	3222
3-Sep-14	99	1174	404	7-Mar-15	284	8416	3245
5-Sep-14	101	1234	413	8-Mar-15	285	8428	3263
6-Sep-14	102	1276	426	9-Mar-15	286	8444	3279
7-Sep-14	103	1287	428	10-Mar-15	287	8463	3289
8-Sep-14	104	1305	433	11-Mar-15	288	8469	3297
9-Sep-14	105	1341	436	12-Mar-15	289	8472	3303
10-Sep-14	106	1367	445	13-Mar-15	290	8476	3312
11-Sep-14	107	1401	450	15-Mar-15	292	8487	3325
12-Sep-14	108	1432	459	16-Mar-15	293	8501	3327
13-Sep-14	109	1464	463	17-Mar-15	294	8502	3336
14-Sep-14	110	1513	468	19-Mar-15	296	8508	3360
15-Sep-14	111	1542	474	20-Mar-15	297	8515	3370
16-Sep-14	112	1571	483	21-Mar-15	298	8518	3376
17-Sep-14	113	1585	489	22-Mar-15	299	8520	3381
18-Sep-14	114	1618	495	23-Mar-15	300	8528	3393
19-Sep-14	115	1640	497	24-Mar-15	301	8529	3398
20-Sep-14	116	1696	501	25-Mar-15	302	8532	3407
21-Sep-14	117	1745	502	26-Mar-15	303	8535	3413

Date	Day	Cumulative Cases	Cumulative Deaths	Date	Day	Cumulative Cases	Cumulative Deaths
12-Aug-14	77	717	264	11-Feb-15	260	8183	3009
13-Aug-14	78	733	273	12-Feb-15	261	8193	3018
14-Aug-14	79	747	280	13-Feb-15	262	8208	3030
15-Aug-14	80	757	287	14-Feb-15	263	8213	3036
16-Aug-14	81	775	297	15-Feb-15	264	8226	3043
17-Aug-14	82	778	305	16-Feb-15	265	8230	3050
18-Aug-14	83	783	312	17-Feb-15	266	8237	3058
19-Aug-14	84	804	320	18-Feb-15	267	8239	3063
20-Aug-14	85	813	322	19-Feb-15	268	8244	3066
21-Aug-14	86	823	329	20-Feb-15	269	8260	3079
22-Aug-14	87	881	333	21-Feb-15	270	8275	3088
23-Aug-14	88	904	336	22-Feb-15	271	8289	3095
24-Aug-14	89	935	341	23-Feb-15	272	8301	3103
25-Aug-14	90	955	355	24-Feb-15	273	8308	3113
26-Aug-14	91	961	363	25-Feb-15	274	8320	3124
27-Aug-14	92	988	372	27-Feb-15	276	8349	3151
28-Aug-14	93	1018	377	28-Feb-15	277	8353	3164
29-Aug-14	94	1033	383	1-Mar-15	278	8370	3180
30-Aug-14	95	1077	387	2-Mar-15	279	8374	3188
31-Aug-14	96	1106	388	3-Mar-15	280	8383	3199
1-Sep-14	97	1115	396	4-Mar-15	281	8389	3210
2-Sep-14	98	1146	399	5-Mar-15	282	8398	3222
3-Sep-14	99	1174	404	7-Mar-15	284	8416	3245
5-Sep-14	101	1234	413	8-Mar-15	285	8428	3263
6-Sep-14	102	1276	426	9-Mar-15	286	8444	3279
7-Sep-14	103	1287	428	10-Mar-15	287	8463	3289
8-Sep-14	104	1305	433	11-Mar-15	288	8469	3297
9-Sep-14	105	1341	436	12-Mar-15	289	8472	3303
10-Sep-14	106	1367	445	13-Mar-15	290	8476	3312
11-Sep-14	107	1401	450	15-Mar-15	292	8487	3325
12-Sep-14	108	1432	459	16-Mar-15	293	8501	3327
13-Sep-14	109	1464	463	17-Mar-15	294	8502	3336
14-Sep-14	110	1513	468	19-Mar-15	296	8508	3360
15-Sep-14	111	1542	474	20-Mar-15	297	8515	3370
16-Sep-14	112	1571	483	21-Mar-15	298	8518	3376
17-Sep-14	113	1585	489	22-Mar-15	299	8520	3381
18-Sep-14	114	1618	495	23-Mar-15	300	8528	3393
19-Sep-14	115	1640	497	24-Mar-15	301	8529	3398
20-Sep-14	116	1696	501	25-Mar-15	302	8532	3407
21-Sep-14	117	1745	502	26-Mar-15	303	8535	3413

Date	Day	Cumulative Cases	Cumulative Deaths	Date	Day	Cumulative Cases	Cumulative Deaths
12-Aug-14	77	717	264	11-Feb-15	260	8183	3009
13-Aug-14	78	733	273	12-Feb-15	261	8193	3018
14-Aug-14	79	747	280	13-Feb-15	262	8208	3030
15-Aug-14	80	757	287	14-Feb-15	263	8213	3036
16-Aug-14	81	775	297	15-Feb-15	264	8226	3043
17-Aug-14	82	778	305	16-Feb-15	265	8230	3050
18-Aug-14	83	783	312	17-Feb-15	266	8237	3058
19-Aug-14	84	804	320	18-Feb-15	267	8239	3063
20-Aug-14	85	813	322	19-Feb-15	268	8244	3066
21-Aug-14	86	823	329	20-Feb-15	269	8260	3079
22-Aug-14	87	881	333	21-Feb-15	270	8275	3088
23-Aug-14	88	904	336	22-Feb-15	271	8289	3095
24-Aug-14	89	935	341	23-Feb-15	272	8301	3103
25-Aug-14	90	955	355	24-Feb-15	273	8308	3113
26-Aug-14	91	961	363	25-Feb-15	274	8320	3124
27-Aug-14	92	988	372	27-Feb-15	276	8349	3151
28-Aug-14	93	1018	377	28-Feb-15	277	8353	3164
29-Aug-14	94	1033	383	1-Mar-15	278	8370	3180
30-Aug-14	95	1077	387	2-Mar-15	279	8374	3188
31-Aug-14	96	1106	388	3-Mar-15	280	8383	3199
1-Sep-14	97	1115	396	4-Mar-15	281	8389	3210
2-Sep-14	98	1146	399	5-Mar-15	282	8398	3222
3-Sep-14	99	1174	404	7-Mar-15	284	8416	3245
5-Sep-14	101	1234	413	8-Mar-15	285	8428	3263
6-Sep-14	102	1276	426	9-Mar-15	286	8444	3279
7-Sep-14	103	1287	428	10-Mar-15	287	8463	3289
8-Sep-14	104	1305	433	11-Mar-15	288	8469	3297
9-Sep-14	105	1341	436	12-Mar-15	289	8472	3303
10-Sep-14	106	1367	445	13-Mar-15	290	8476	3312
11-Sep-14	107	1401	450	15-Mar-15	292	8487	3325
12-Sep-14	108	1432	459	16-Mar-15	293	8501	3327
13-Sep-14	109	1464	463	17-Mar-15	294	8502	3336
14-Sep-14	110	1513	468	19-Mar-15	296	8508	3360
15-Sep-14	111	1542	474	20-Mar-15	297	8515	3370
16-Sep-14	112	1571	483	21-Mar-15	298	8518	3376
17-Sep-14	113	1585	489	22-Mar-15	299	8520	3381
18-Sep-14	114	1618	495	23-Mar-15	300	8528	3393
19-Sep-14	115	1640	497	24-Mar-15	301	8529	3398
20-Sep-14	116	1696	501	25-Mar-15	302	8532	3407
21-Sep-14	117	1745	502	26-Mar-15	303	8535	3413

Date	Day	Cumulative Cases	Cumulative Deaths	Date	Day	Cumulative Cases	Cumulative Deaths
12-Aug-14	77	717	264	11-Feb-15	260	8183	3009
13-Aug-14	78	733	273	12-Feb-15	261	8193	3018
14-Aug-14	79	747	280	13-Feb-15	262	8208	3030
15-Aug-14	80	757	287	14-Feb-15	263	8213	3036
16-Aug-14	81	775	297	15-Feb-15	264	8226	3043
17-Aug-14	82	778	305	16-Feb-15	265	8230	3050
18-Aug-14	83	783	312	17-Feb-15	266	8237	3058
19-Aug-14	84	804	320	18-Feb-15	267	8239	3063
20-Aug-14	85	813	322	19-Feb-15	268	8244	3066
21-Aug-14	86	823	329	20-Feb-15	269	8260	3079
22-Aug-14	87	881	333	21-Feb-15	270	8275	3088
23-Aug-14	88	904	336	22-Feb-15	271	8289	3095
24-Aug-14	89	935	341	23-Feb-15	272	8301	3103
25-Aug-14	90	955	355	24-Feb-15	273	8308	3113
26-Aug-14	91	961	363	25-Feb-15	274	8320	3124
27-Aug-14	92	988	372	27-Feb-15	276	8349	3151
28-Aug-14	93	1018	377	28-Feb-15	277	8353	3164
29-Aug-14	94	1033	383	1-Mar-15	278	8370	3180
30-Aug-14	95	1077	387	2-Mar-15	279	8374	3188
31-Aug-14	96	1106	388	3-Mar-15	280	8383	3199
1-Sep-14	97	1115	396	4-Mar-15	281	8389	3210
2-Sep-14	98	1146	399	5-Mar-15	282	8398	3222
3-Sep-14	99	1174	404	7-Mar-15	284	8416	3245
5-Sep-14	101	1234	413	8-Mar-15	285	8428	3263
6-Sep-14	102	1276	426	9-Mar-15	286	8444	3279
7-Sep-14	103	1287	428	10-Mar-15	287	8463	3289
8-Sep-14	104	1305	433	11-Mar-15	288	8469	3297
9-Sep-14	105	1341	436	12-Mar-15	289	8472	3303
10-Sep-14	106	1367	445	13-Mar-15	290	8476	3312
11-Sep-14	107	1401	450	15-Mar-15	292	8487	3325
12-Sep-14	108	1432	459	16-Mar-15	293	8501	3327
13-Sep-14	109	1464	463	17-Mar-15	294	8502	3336
14-Sep-14	110	1513	468	19-Mar-15	296	8508	3360
15-Sep-14	111	1542	474	20-Mar-15	297	8515	3370
16-Sep-14	112	1571	483	21-Mar-15	298	8518	3376
17-Sep-14	113	1585	489	22-Mar-15	299	8520	3381
18-Sep-14	114	1618	495	23-Mar-15	300	8528	3393
19-Sep-14	115	1640	497	24-Mar-15	301	8529	3398
20-Sep-14	116	1696	501	25-Mar-15	302	8532	3407
21-Sep-14	117	1745	502	26-Mar-15	303	8535	3413

Date	Day	Cumulative Cases	Cumulative Deaths	Date	Day	Cumulative Cases	Cumulative Deaths
12-Aug-14	77	717	264	11-Feb-15	260	8183	3009
13-Aug-14	78	733	273	12-Feb-15	261	8193	3018
14-Aug-14	79	747	280	13-Feb-15	262	8208	3030
15-Aug-14	80	757	287	14-Feb-15	263	8213	3036
16-Aug-14	81	775	297	15-Feb-15	264	8226	3043
17-Aug-14	82	778	305	16-Feb-15	265	8230	3050
18-Aug-14	83	783	312	17-Feb-15	266	8237	3058
19-Aug-14	84	804	320	18-Feb-15	267	8239	3063
20-Aug-14	85	813	322	19-Feb-15	268	8244	3066
21-Aug-14	86	823	329	20-Feb-15	269	8260	3079
22-Aug-14	87	881	333	21-Feb-15	270	8275	3088
23-Aug-14	88	904	336	22-Feb-15	271	8289	3095
24-Aug-14	89	935	341	23-Feb-15	272	8301	3103
25-Aug-14	90	955	355	24-Feb-15	273	8308	3113
26-Aug-14	91	961	363	25-Feb-15	274	8320	3124
27-Aug-14	92	988	372	27-Feb-15	276	8349	3151
28-Aug-14	93	1018	377	28-Feb-15	277	8353	3164
29-Aug-14	94	1033	383	1-Mar-15	278	8370	3180
30-Aug-14	95	1077	387	2-Mar-15	279	8374	3188
31-Aug-14	96	1106	388	3-Mar-15	280	8383	3199
1-Sep-14	97	1115	396	4-Mar-15	281	8389	3210
2-Sep-14	98	1146	399	5-Mar-15	282	8398	3222
3-Sep-14	99	1174	404	7-Mar-15	284	8416	3245
5-Sep-14	101	1234	413	8-Mar-15	285	8428	3263
6-Sep-14	102	1276	426	9-Mar-15	286	8444	3279
7-Sep-14	103	1287	428	10-Mar-15	287	8463	3289
8-Sep-14	104	1305	433	11-Mar-15	288	8469	3297
9-Sep-14	105	1341	436	12-Mar-15	289	8472	3303
10-Sep-14	106	1367	445	13-Mar-15	290	8476	3312
11-Sep-14	107	1401	450	15-Mar-15	292	8487	3325
12-Sep-14	108	1432	459	16-Mar-15	293	8501	3327
13-Sep-14	109	1464	463	17-Mar-15	294	8502	3336
14-Sep-14	110	1513	468	19-Mar-15	296	8508	3360
15-Sep-14	111	1542	474	20-Mar-15	297	8515	3370
16-Sep-14	112	1571	483	21-Mar-15	298	8518	3376
17-Sep-14	113	1585	489	22-Mar-15	299	8520	3381
18-Sep-14	114	1618	495	23-Mar-15	300	8528	3393
19-Sep-14	115	1640	497	24-Mar-15	301	8529	3398
20-Sep-14	116	1696	501	25-Mar-15	302	8532	3407
21-Sep-14	117	1745	502	26-Mar-15	303	8535	3413

Appendix C

The solution set for the stochastic model (6.5):

$$\begin{aligned}\ln\left(E[x_1(t)] - \frac{0.5754}{0.0183}\right) &= \ln\left(x_{1_0} - \frac{0.5754}{0.0183}\right) - 0.0183t \\ \ln\left(E[x_2(t)] - \frac{1.0592}{0.0336}\right) &= \ln\left(x_{2_0} - \frac{1.0592}{0.0336}\right) - 0.0336t \\ \ln(E[x_i(t)]) &= \ln(x_{i_0}) - \frac{\sigma_i^2}{2}t, \quad i = 3, 4, 5\end{aligned}$$

Appendix D

Proof. The system (6.6)-(6.12) can be rewritten in the form

$$\frac{d}{dt}(y_0, y_1, y_2, y_3, y_4, y_5, y_6)^T = (f_0, f_1, f_2, f_3, f_4, f_5, f_6)^T,$$

where the functions f_i , $i = 0, \dots, 6$ are given by

$$\begin{aligned}f_0 &= \Lambda - \mu y_0 - \beta \phi(y_1, y_2) y_0 - \frac{\beta_1 y_1 y_0}{1 + k y_1} - \frac{\beta_2 y_2 y_0}{1 + k y_2}, \\ f_1 &= \frac{\beta_1 y_1 y_0}{1 + k y_1} + \frac{\beta_2 y_2 y_0}{1 + k y_2} - a y_1, \\ f_2 &= \gamma y_1 - b y_2, \quad f_3 = \beta \phi(y_1, y_2) y_0 + \varepsilon_1 y_1 + \varepsilon_2 y_2 - c y_3, \\ f_4 &= \alpha y_3 - d y_4, \quad f_5 = \rho_1 y_1 + \rho_2 y_2 - \mu y_5, \quad f_6 = \theta_1 y_3 + \theta_2 y_4 - \mu y_6.\end{aligned}$$

We want to investigate the occurrence of the transcritical bifurcation at $R_0 = 1$. The Jacobian of the system of equations (6.6)-(6.12) evaluated at the DFE is given by

$$J(\zeta_0) = \begin{bmatrix} -\mu & -a_1 & -a_2 & 0 & 0 & 0 & 0 \\ 0 & (a_1 - a) & a_2 & 0 & 0 & 0 & 0 \\ 0 & \gamma & -b & 0 & 0 & 0 & 0 \\ 0 & \varepsilon_1 & \varepsilon_2 & -c & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & -d & 0 & 0 \\ 0 & \rho_1 & \rho_2 & 0 & 0 & -\mu & 0 \\ 0 & 0 & 0 & \theta_1 & \theta_2 & 0 & -\mu \end{bmatrix}, \quad (7.1)$$

where

$$a_1 = \beta_1 y_0^*, \quad a_2 = \beta_2 y_0^*.$$

Now, let's consider $R_0 = 1$ and take $\beta_2 = \beta_2^*$ as the bifurcation parameter. Then, we have

$$\beta_2^* = \frac{b(\mu a - \beta_1 \Lambda)}{\gamma \Lambda}.$$

It follows from Lemma 6.6 that the DFE, ζ_0 , is locally stable when $\beta_2 < \frac{b(\mu a - \beta_1 \Lambda)}{\gamma \Lambda}$ and unstable when $\beta_2 > \frac{b(\mu a - \beta_1 \Lambda)}{\gamma \Lambda}$. \square

Appendix E

$$R = \begin{bmatrix} -\mu & -a_1 & -a_2 & 0 & 0 & 0 & 0 \\ 0 & (a_1 - a) & a_2 & 0 & 0 & 0 & 0 \\ 0 & \gamma & -b & 0 & 0 & 0 & 0 \\ 0 & \varepsilon_1 & \varepsilon_2 & -c & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & -d & 0 & 0 \\ 0 & \rho_1 & \rho_2 & 0 & 0 & -\mu & 0 \\ 0 & 0 & 0 & \theta_1 & \theta_2 & 0 & -\mu \end{bmatrix},$$

$$L = \begin{bmatrix} -\mu & 0 & 0 & 0 & 0 & 0 & 0 \\ -a_1 & (a_1 - a) & \gamma & \varepsilon_1 & 0 & \rho_1 & 0 \\ -a_2 & a_2 & -b & \varepsilon_2 & 0 & \rho_2 & 0 \\ 0 & 0 & 0 & -c & \alpha & 0 & \theta_1 \\ 0 & 0 & 0 & 0 & -d & 0 & \theta_2 \\ 0 & 0 & 0 & 0 & 0 & -\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\mu \end{bmatrix},$$

Appendix F

Table 7.1: Parameters and their ranges used for the sensitivity analysis

Parameter	Range	PRCC	p-value
β	[0, 0.0000003]	-0.02560785	0.5984
β_1	[0, 1]	0.43027179	0.0000
β_2	[0, 0.5]	0.06185022	0.01485
ϕ	[0, 0.001]	0.01214754	0.07253
μ	[0, 0.05]	-0.55180	0.0000
ρ_1	[0, 1]	0.23058413	0.00000000000006523
ρ_2	[0, 1]	-0.01121879	0.07523
δ_1	[0, 0.0001]	-0.01248264	0.07523
δ_2	[0, 0.001]	0.02990188	0.5391
γ	[0, 1]	0.31186786	0.0000
Λ	[0, 0.05]	0.52410196	0.0000
ε_1	[0, 0.01]	-0.02162308	0.6515
ε_2	[0, 0.01]	0.05458266	0.2114
ε_3	[0, 0.001]	0.040342646	0.4632
α	[0, 0.01]	0.05194083	0.2199
θ_1	[0.0146, 0.0421]	-0.03522090	0.5096
θ_2	[0, 0.0146]	-0.03086953	0.7253
k	[0, 1]	-0.03086953	0.5391