Epidemic Modelling and Digital Intervention Strategies for Infectious Disease: *A COVID-19 Case Study of Three English Cities*

Tahsinur Rahman Khan

Nottingham Geospatial Institute, University of Nottingham, UK

Dissertation submitted to The University of Nottingham in partial fulfilment of the degree of Master of Research in Geospatial Systems

August 2020

ABSTRACT

The emergence of the novel coronavirus and the resulting global pandemic has shown the importance of epidemic modelling and how sound, scientifically driven policies can aid in response efforts. In purely quantitative terms, an epidemic event can be considered as one of the most complex geospatial events to be modelled and make prediction about. However, mathematical and computational models have been applied in the past with success in understanding and tackling such events. In this report we will cover the two main approaches for epidemic modelling – mathematical and agent-based computational models and compare and contrast the results in the context of three English cities – Leicester, Bradford and Blackburn, that have been heavily impacted by the Covid-19 pandemic. We will also discuss some plans on future research directions on constructing a robust pandemic resiliency framework that can be deployed across local geographical regions to aid public health authorities deal with pandemic outbreaks.

ACKNOWLEDGMENTS

I would like to sincerely thank my primary supervisor – Professor Stuart Marsh who has been a great mentor and guide throughout my first year at the Geospatial Systems CDT. I would also like to extend my gratitude to my secondary supervisor – Professor Theodore Kypraios for taking me on as a supervisee on top of his already extremely hectic schedule and for providing valuable advise for my MRes. I would also like to thank my colleagues at the Nottingham Geospatial Institute for always patiently answering all my questions and advising me throughout my research.

I would like to dedicate this research to my parents, who have given me so much throughout my life; and my uncle – who gave so much to his family, his people and his country. Rest in peace General.

Glossary of Terms

Symbols	Definition
S	Susceptible
I_S	Infected with Symptoms
I_{NS}	Infected with No Symptoms
T	Treatment
R	Recovered
λ_S	Force of infection for infected cases with symptoms
λ_{NS}	Force of infection for infected cases without symptoms
p	Proportion of infected population with symptoms
$1-p$	Proportion of infected population with no symptoms
h	Inverse of incubation period
γ	Recovery rate. Inverse of number of days needed for recovery
μ_{cov}	Mortality rate due to Covid-19
$\mu_{non-cov}$	Mortality rate due to non-Covid causes
ρ_{NS}	Transmission rate for non-symptomatic cases
ρ_S	Transmission rate for symptomatic cases
C_{NS}	Contact Matrix for Non-Symptomatic Cases
C_S	Contact Matrix for Symptomatic Cases

Table 1 Glossary of all epidemic modelling terms referenced in the report

CONTENTS

List of Figures

List of Tables

1. INTRODUCTION

The two main domains that have seen wide application in epidemic modelling include the mathematical and the computational domains (Epstein *et al.*, 2008). Methodologies belonging to these domains have proven to aid researchers and public health officials for understanding the transmission of an infectious diseases across communities. Such information can play a critical role during a pandemic outbreak by allowing healthcare agencies to enact effective policies to control the spread of the pathogen.

Although both the mathematical and computational approaches have the same goal of aiding policy makers tackle epidemic scenarios, they take diverging paths to achieve such a goal. While mathematical models are able to provide broad overview of the probable scenarios that might play out in pandemic scenarios, computational approaches can provide further granular level details to refine the broader pictures provided by the mathematical models.

The main advantage of mathematical models includes their ability to leverage only few fundamental factors and simplistic assumptions to start building out a functional framework for projecting a wide range of pandemic scenarios. These key factors and assumptions are embedded into the models through parametrized variables, that are then used to predict macroscopic details such as the infection rate threshold of a potential outbreak and the probable final size of the epidemic event in terms of the number of infected, specific disease related mortality rate and the like. While these numbers are crucial to draw up large scale management plans for deploying efficient intervention strategies and for reducing the burden on healthcare facilities, they are dependent on initial assumptions. Hence, hidden biases in those assumptions can cause the results derived from mathematical models to give wrong impressions about the ground realities. Moreover, most mathematical models only consider homogeneous community structures where the interaction pattern in population is considered to be uniform for modelling purposes. However, in reality that is hardly the case since subpopulations in communities (le Polain de Waroux *et al.*, 2018) mix in heterogeneous ways.

To enable the mathematical models to give more accurate picture of a pandemic scenario, researchers have made use of computational methodologies that can simulate minute level details of populations under study. Through such simulations, the initial assumptions of the mathematical models can be refined and any hidden biases can be gotten rid of. Two types of computational models are generally applied for the study of epidemics – network based model and agent-based model. In such approaches, epidemic models are built by considering each individual in a population as an artificial agent. These agents are then *authored* in specific ways to mimic the expected decision making of individuals in a pandemic situation (Conte and Paolucci, 2014). The whole population can be divided into smaller subpopulations representing individuals from different age bands and socioeconomic groups. Then stochastic processes are employed and the agents are allowed to mix in artificially created societies to observe their behaviour and potentially uncover the transmission dynamics. The authoring of agents and the stochastic mixing patterns serve an important purpose of replicating everyday movement and behaviour patterns in humans. Thus, allowing for a virtual, privacy preserving monitoring system. For all the benefits that computational approaches provide, their main drawbacks include their dependence on high volumes of data and computational resources. But as progresses are made in both of those fronts – availability of big data and advances in hardware technologies, computational models are being adopted for the study of epidemics.

The emergence of the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), that have been declared a global pandemic by the World Health Organization (WHO) in March, 2020 (World Health Organization, 2020a), have engaged the wider research communities in the biosciences and technology spheres to try and find an answer to control the spread of the virus. Research institutions have collaborated to conduct wide scale study of various aspects of the disease ranging from finding vaccines and antivirals to predicting the spread of the disease across different geographical locations (New York Times, 2020).

In this report we will review both the mathematical and computational approaches for predicting the spread of an infectious virus. Our models will take into consideration the available data for the novel coronavirus that have been vetted by reliable health authorities. Through such diverse modelling approach we aim to quantitatively highlight the advantages and limitations of the two models. We hope that the groundwork we lay through such research work will be leveraged to build more complex epidemic models that can aid public health authorities streamlined pandemic response systems.

The entire report is broken down into 6. Section 2 covers the literature review where we cover a wide breadth of research in mathematical and computational approaches. We detail our methodology, for both our mathematical model and our computational, agent-based model, in section 3. In section 4 we present the results from our models and discuss the accuracies and limitations of our findings. Section 5 includes the future plan on how we plan on building upon the groundwork laid through this research. Section 6 includes the concluding remarks. All code and additional tables that are not inside the main report are included in the appendices.

2. LITERATURE REVIEW

In this section we have broken down the discussion of Mathematical Epidemic Modelling and Computational Epidemic Modelling into three sections. Section 2.1 will cover purely mathematical models – both deterministic and stochastic. Sections 2.2 and 2.3 will cover two types of computational approaches – the Network based Models and the Agent-based Models. The literature review section will conclude with a discussion of the advantages and limitations of all three approaches in section 2.4.

2.1. Mathematical Modelling for Epidemics

Mathematical modelling has been employed for formulating the spread of epidemics for quite some time. Some of the earliest known cases of mathematical modelling for spread of disease include Daniel Bernoulli's model to gauge the efficacy of vaccination for smallpox virus in 1766, the 1906 discrete time model that was proposed by Hamer for discerning the recurrent pattern of measles epidemics, Ross's 1911 model that was based on differential equations to explore the effects of various intervention strategies for malaria; and the Reed-Frost model from the 1920s that was devised by Lowell Reed and Wade Hampton Frost to outline the ways how contagious diseases spread across different demographics within a population (Löscher and Prüfer-Krämer, 2010). The Reed-Frost model was further developed by Kermack and McKendrick to form the basis of the widely used compartmental model today, which describes the system level dynamics of the transmission of infectious diseases (Barbour and Utev, 2004). What McKendrick and Ross found was that, if the basic reproduction number (R0) which is used to represent the average number of secondary infections caused by a single infected case in an otherwise susceptible population, was greater than a certain threshold then only a novel pathogen can cause an epidemic event in a population. These early epidemic models have been extended to include many more details about the different aspects of a society – such as household sizes, age groups, contact mixing patterns among other, to further refine the models in order to make better projections of the number of people who might fall under different health categories – such as susceptible to, infected of and recovered from a disease. In general, all mathematical epidemic models are divided into either deterministic or stochastic models.

2.1.1. Deterministic Models

By the sheer number of publications and amount of research, compartment models are the most extensively researched topic in mathematical epidemic modelling (Dimitrov and Meyers, no date; Keeling and Danon, 2009; Garnett *et al.*, 2011; Siettos and Russo, 2013). In a simple, deterministic epidemic model, the population under study is assumed to be homogeneous in terms of social and biological characteristics and mixture pattern. The entire population is then divided into different compartments according to their individual heath states, namely susceptible, infected and removed. The relationship between different compartments and the transition of the population between the compartments are established through differential equations. In such models there are three main differential equations that keep track of the change in each of the compartments with respect to time. Each of those equations combines different quantities such as the infection and recovery rates, force of infection, size of the population, the basic reproduction number and the proportion of contacts made by individuals in each of those compartment with sub-populations that are classed under other compartments.

Among the most popular compartment models are the SIR (susceptible, infectious and removed) model, SEIR (susceptible, exposed, infectious, removed) model, the SIRS (susceptible, infected, susceptible) model among others. Figure 1 below shows these models below. Here beta denotes the infection rate, gamma denotes the recovery rate and alpha represents the incubation period which denotes the time between exposure and the onset of symptoms.

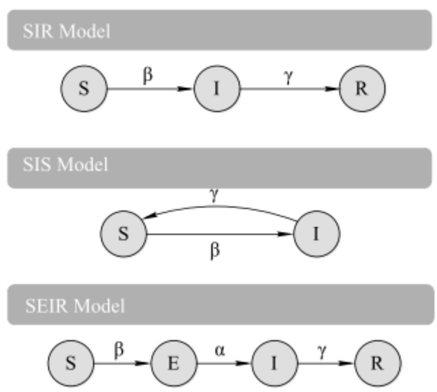


Figure 1 Basic compartmental models. Reproduced from (Duan *et al.*, 2015)

In a SIR model, individuals are placed in one of the three health states at any arbitrary point in time during the time frame of simulation. Simple models start off by placing only one individual in the infected compartment and the rest of the population are placed in the susceptible compartment. The susceptible individuals are then transferred to the infectious compartment at a rate equivalent to the infection rate of the disease (β). After a period of time, the infected individuals are moved to the recovered state at a rate known as the recovery rate (γ). After reaching the recovered state, it is considered that individuals have immunity from the disease and remain in that state indefinitely. The graph in Figure 2 shows the temporal variation in the number of people in each of the three compartments. During the initial stages, the proportion of population who are in the susceptible category is very high (close to 1, representing almost 100 percentage of the population). However, over time the proportion of infected cases rises until it reaches a peak and then gradually declines. The population proportion in the recovered state varies from very low numbers at the start of an epidemic event and gradually plateaus at levels that are close to the levels of the initial susceptible proportions.

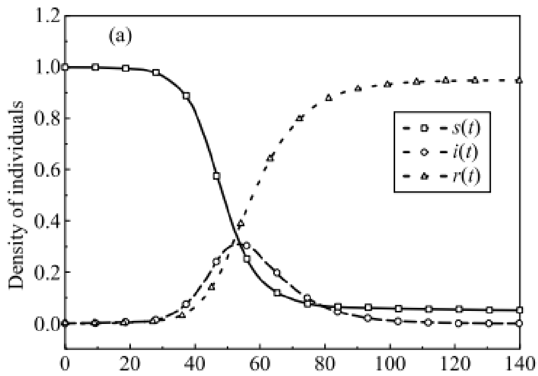


Figure 2 Results from an SIR model simulation showing the variation in each of the S, I and R compartments. Reproduced from (Duan *et al.*, 2015)

The SEIR model is similar to the SIR model, however there is one additional "exposed" stage between the susceptible and infected stages. The exposed state simply refers to the time between a susceptible individual getting exposed to a disease and the onset of symptoms. Many diseases like chicken pox, various strains of influenza and Covid-19 (Mitchell and Ross, 2016) have an exposed state that corresponds to the time in between the virus entering the body and its manifestation through symptoms. The other main type of mathematical epidemic model is the SIS model. In this model, the population goes through a cyclical process from susceptible to infected states. The framework of the model suggests that it is suitable for diseases that are recurrent in the population. In such models, infected individuals return back to the susceptible pool after recovery. This means, the population does not develop immunity for the disease and continuously oscillate between the states. Seasonal flu might be a good example for such a disease, where there is no long lasting vaccine that can aid in developing lifelong immunity (Barberis *et al.*, 2016).

2.1.2. Stochastic Models

In real life, complex scenarios such as an epidemic event is rarely deterministic. Simply relying on differential equations will provide either vastly over or underestimated results. Therefore, to factor in random events into the prediction models it is useful to introduce stochasticity. Stochastic processes such as Markov Chains, Monte Carlo methods and hybrid mixture of both – Markov Chain Monte Carlo methods have been widely used for such purposes (O'Neill, 2002; Jewell *et al.*, 2009; Britton, 2010; Rorres, Pelletier and Smith, 2011; Korostil *et al.*, 2013). Simple stochastic epidemic models make the assumption that contact events between susceptible and infected individuals can be modelling probabilistically such that every contact between individuals belonging to those groups does not definitively result in transmission of disease. Although rudimentary in approach, this is in line with natural biological processes of diseases transmission where each and every contact does not necessarily result in diseases being passed on (Allen, 2017).

A preliminary sketch for stochastic modelling can be made by thinking of a population consisting of N individuals. In such a population, every individual has the same uniform probability to come in contact with one another. Starting with one infected individual in that population, the disease spreads when mixing happens in the population. However, not all contacts result in an infected individual transferring the disease to a susceptible one. Therefore, a probabilistic trend emerges in transmission dynamics – which are of interest to modelers since that resembles real life events. Similar to SIR models, the infected individuals then recover and attain immunity.

As mentioned earlier, if we consider there is only one infected individual in the population initially then we can say $I(0) = 1$. The initial number of susceptible will be $S(0) = N - 1$. If we consider the per contact, between a susceptible and infected individual, probability of contracting the disease be p, then the probability of a susceptible not contracting the disease will be $1 - p$. Now if we extend our assumption to any arbitrary time t, then the the probability that any susceptible individual is infected will be $P_{t+1} = p^{I(t)}$ and the probability that they won't be infected will be $Q_{t+1} = 1 - p^{I(t)}$. Here, $I(t)$ represents the incidence of number of individuals at time t. We can then extend our model to project the expected number of infected cases at time $t + 1$ to according to equation 1 below,

$$E\big(I(t+1)\big|I(t), S(t)\big) = S(t)\big(1 - q^{I(t)}\big) \quad 1$$

2.2. Complex Network Models for Epidemics

Owing to the advent of ubiquitous computing that can track movement patterns (Spreitzer and Theimer, 2007), complex network models for epidemic events have seen a rise in interest in recent times (Eubank *et al.*, 2004; Leskovec *et al.*, 2007). In these graph-like network models, nodes are used to represent individuals in a population and the links between the nodes represent all possible interactions between a node and all of its neighbors. All complex network models for epidemics can be broadly classed under two categories – namely the dynamic complex networks and the numerically simulated complex networks. The dynamic networks mainly rely on mean-field theory to demonstrate the spread of epidemics in a network like contact structure within local populations (Pastor-

Satorras and Vespignani, 2001). Just like in basic SIR compartmental models, individuals are divided into different health stages. In a homogeneously mixing structure, the infectiousness of nodes that are classified as infected are dependent on the node degree of the network – which is basically a count of the number of edges connected to the node. However, in real life situations, mixing in the population are rarely homogeneous. Rather, the entire population is broken down into demographic subgroups (Colizza *et al.*, 2007). To model such non-uniform, heterogenous interactions, in more complex dynamic models the notes are split into subgroups with different node degrees. The other main category of complex network based models include the numerically simulated complex networks that leverage off of *individual-based models* to express the contact patterns in a population. In such models, the links between susceptible and infected nodes represent the infection probability between susceptible and infected individuals in the population (Deijfen, 2011).

2.2.1. Role of Network Topologies on Epidemic Modelling

In complex network models, the underlying networks play significant roles in determining the temporal spread of the epidemic throughout the network (Pastor-Satorras and Vespignani, 2001; Chung-Yuan Huang, 2004; Zhou *et al.*, 2006). The common network topologies that have been studied widely for modelling epidemics include random networks, scale – free networks (Barabási and Albert, 1999) and small-world networks (Watts and Strogatz, 1998). Research into these variety of network structures have revealed that during early stages of an epidemic event, scale free networks, where characteristics of network remain constant regardless of the expanding size have a higher propensity to facilitate the spread of infectious diseases than in small-world networks, where connection between nodes are much more sparse.

Among the well-known complex networks that have been covered in epidemic modelling literature include Pastor-Satorras et al.'s SIS dynamic complex network model that employed mean-field theory to predict the spread of disease over a homogeneously mixing population (Pastor-Satorras and Vespignani, 2001). In this model Pastor-Satorras et al. studied the behavior of high-dimensional stochasticity of an epidemic model first by building simpler, networked model and adding more complexity to it by averaging over degrees of freedom that represent various social factors that lead to more social contacts in real life population dynamics.

Effects of different vaccination strategies in complex networked models have also been covered widely (Madar *et al.*, 2004). Additionally, many recent research works have dived into incorporating heterogeneity into complex network models by applying hierarchical and modular networks that resemble real world demographic structure (Palla *et al.*, 2005; Y. Wang *et al.*, 2011; Griffin and Nunn, 2012).

2.2.2. Weighted Network Models

An alternative approach to complex network modelling for epidemics include the weighted networks, where relative weights of the node edges serve as indicators to the interaction strength. Examples of such weighted networks can be seen in other areas such as in airport networks where weights represent the flight traffic between airports (Barrat *et al.*, 2004); collaboration network in the scientific community where node edges represent the density of co-authored papers (Li *et al.*, 2005); trade networks where higher the bilateral trade is between any two country, greater will be the edge weights (Latora and Marchiori, 2003).

Introducing weighted edges can have the advantage of capturing the detailed interaction patterns within a population, as the corresponding weighted edges between nodes represent the amount of interactions between any two adjoining nodes. Among the well-established weighted networks for epidemic projection include the Barrat, Barthelemy, and Vespignani (BBV) networks (Barrat, Barthélemy and Vespignani, 2004) that had been repurposed by Yan et al. (Kamp, Moslonka-Lefebvre and Alizon, 2013) to study epidemic spread. The BBV networks are a class of scale-free networks that allow for evolution in edge weights depending on the population interaction pattern. In such BBV network derived models, Yan et al. assumed a constant population that had been exposed to types of diseases that exhibit SIS cyclical disease pattern, where infected individuals return to the pool of susceptible after recovering. On performing random numerical simulations, Yan et al. concluded that larger dispersion of weights within the network slows the spread of the disease, which might indicate that disease clusters – that have the potential to cause epidemic events – are avoidable when regions of highly weighted, joined nodes are averted; meaning in real life scenarios, if frequency of contacts between individuals are reduced across demographic subpopulations, then that might help in preventing an epidemic. Figure 3 represents the distributions in node degrees, node connectivity and weights of the edges in a BBV network.

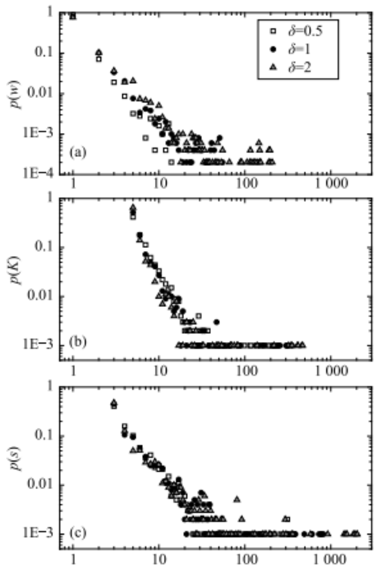


Figure 3 The distribution of edge weights, node connectivity and node strengths in a BBV network. Reproduced from (Duan *et al.*, 2015)

Chu et al. (Chu *et al.*, 2009) took the model from Yan et al. further by incorporating community structure into the model. In their model, Chu et al., classified the edges as either internal or external edges in a weighted community network. The internal edges are meant to represent the interactions between individuals in the same demographic subgroup, while external edges represent the interaction between different subgroups of the entire population. In such a network, the transmission rates are defined by the corresponding internal and external edges as well. Results from the study indicated that external node weights played the major role in determining the temporal projection of the spread of the epidemic.

Weighted networks have also been used for evaluating vaccination strategies. Eames et al. (Eames, Read and Edmunds, 2009) leveraged data from a survey of population contact behaviors to study the effects of targeted vaccination strategies in contact networks with varying configurations. Other significant works related to the spread of epidemics in weighted networks include Fournie et al.'s analysis on the spread of H5N1 influenza in live bird markets in northern Vietnam (Fournie *et al.*, 2013); and Duan et al.'s study on epidemic events from the Severe Acute Respiratory Syndrome (SARS) (Duan, Qiu, *et al.*, 2013). Several other works have focused on different aspects of an epidemic event, such as the epidemic threshold – which denotes the minimum number of infected individuals that are needed to cause an epidemic in a geographical location, incidence and prevalence levels in weighted networks and the impact of weight distributions and network topology on the final size of an epidemic (Kamp, Moslonka-Lefebvre and Alizon, 2013; Zhu *et al.*, 2013).

2.2.3. Adaptive Network Models

Pandemic events are complex events that affects changes in human behavior as well. However, the models mentioned so far are static and does not consider the changing behavior of human beings in societies, in response to outbreak events. To incorporate this factor of changing behavior, epidemic modelers have made use of adaptive networks. Adaptive networks are capable of capturing the epidemic dynamics by integrating a feedback loop that takes into account any natural phenomenon that might influence the contact network between individuals in a population (Gross and Blasius, 2008; Schwartz and Shaw, 2010; Shaw and Schwartz, 2010).

Adaptive networks for modelling epidemic spread works by rewiring the connections between nodes that mirrors the change in human behaviors. Gross et al. applied such adaptive networks to model an SIS cyclical epidemic event (Gross and Blasius, 2008). Their findings revealed that different rates of infection and rewiring can lead to a range of different situations that include the proportion of healthy individuals left in a population after an epidemic, a permanent endemic state and oscillatory states where the public health state at a geographic location varies according to seasons. In their research, Gross et al. also investigated the relationship between infection and rewiring rates to uncover conditions where both endemic and disease-free states could co-exists.

Further work combining adaptive networks for projecting disease spread has been done by Shaw and Schwartz, where they considered an SIR model instead of SIS (Shaw and Schwartz, 2008). In their SIR adaptive network model, Shaw and Schwartz also studied the effects of various vaccination strategies on populations. Marceau et al. and Yang et al. have focused their efforts on the temporal evolution of diseases under various network populations (Marceau *et al.*, 2010; Yang, Tang and Zhang, 2012). In these models, compartmentalization of the entire population has not only been based on the individual node's health state but also the health states of the neighboring nodes.

Although most research has focused on leveraging adaptive random networks, several researchers have explored various other network topologies. Song et al. (Song, Jiang and Xu, 2011) has studied epidemic spread on range of networks that include adaptive random networks, nearest neighbor networks and small world networks. They extended their study to include evolutionary rewiring strategies in the underlying network by employing techniques such as cellular automata. Various other adaptive networks that have produced significant insights into transmission dynamics include the preferred degree adaptive networks and multi-type networks that contain variety of nodes with differing characteristics (B. Wang *et al.*, 2011).

In recent times, advances have been made to incorporate information feedback onto adaptive networks to control rewiring strategies (van Segbroeck, Santos and Pacheco, 2010; Zhang *et al.*, 2012). Segbroeck et al. (van Segbroeck, Santos and Pacheco, 2010) studied SIS and SIR models in adaptive networks where state of neighboring nodes were used as information that influenced the evolution the disease among localized regions of the network. Zhang et al and Zanette and Risau-Gusman explored several rewiring rules and their effects in adaptive networks where contact networks evolved not through rewiring but by breaking and recovering existing links (Zanette and Risau-Gusmán, 2008; Zhang *et al.*, 2012). They concluded that during outbreak events individuals should reduce contacts with others and should consider isolation activities rather than coordinative activities with others in the population.

2.3. Agent based Models for Epidemics

In recent times, with the increase in computational power and availability of big data pertaining to public health and demographics, agent based models for simulating epidemics have gain popularity (Dunham, 2005). Agent based simulation allows for simulating and predicting the complex evolutionary processes of transmission of diseases. As such, agent-based models can be useful for illustrating the spatio-temporal ground situation that might arise in a real-life pandemic outbreak. In agent-based models, each agent is considered to represent an individual within a population. Hence such models are capable of capturing the heterogeneity in the population mixture, such as evolving characteristics of individuals and their behaviors towards one another, in a stochastic manner (Duan *et al.*, 2015). This ability of agent-based models to capture microscopic trends in societies serve an important role to uncover underlying causes that are responsible for macroscopic events during an epidemic.

2.3.1. Large Scale Agent-based Simulation

Quite a few number of agent based simulation systems have been developed to study the transmission of diseases across geographical locations. Among the popular ones are MASON, EpiSims, BioWar, Geograph and Flute (Hunter, Namee and Kelleher, 2017).

Tools have been developed in Geograph, to investigate geospatial behavior among agents that are exposed to diseases that follow the SEIR cycle. The EpiSims simulation system makes use of traffic networks to model local human urban mobility patterns. This mobility patterns are then used as a proxy to simulate the transmission dynamics of an infectious disease. Flute simulation system are particularly suited for modelling the diffusion of influenza type virus across major cities. To do this, Flute makes expansive use of census data to create detailed population and community structures. Mason is a Java based, parallel programmable platform that contains a large sent of simulation libraries capable of modelling complex, large scale interaction among agents in a simulated epidemic event. BioWar is mainly used to simulate biological attacks on populated metropolitan areas. Due to the large number of complex processes that take place during each time step in a human population, simulating such events can be quite computationally and data expensive. To handle such complexities, the BioWar platform integrate social network models, information media, disease transmission parameters, population data, weather data, urban infrastructure models and map data.

The agent-based simulation systems described above rely heavily on census, spatial and infrastructure data. GIS techniques have been used in the past to illustrate outbreaks across geolocations. The high-resolution details that are displayed through agent-based models require efficient algorithms and considerable computational powers. As such, work has been done to develop high performance, scalable and parallelly computable epidemic simulation systems. For instance, EpiSimdemics and EpiFast algorithms, that make use of parallel computation, have been specifically constructed to increase the performance of EpiSims (Barrett C *et al.*, 2008; Bisset *et al.*, 2009). Zeng et al., have developed a spatial agent based system specifically covering the entire location of the city of Beijing, in order to investigate the transmission of infectious disease transmission (Zeng *et al.*, 2009). Figure 4 shows the map of Beijing and the spread of influenza in the city. In order to build the system, Zeng et al. integrated census data, transportation networks and boundaries of different regions within the city. Performing simulation runs on such a detailed system has allowed Zeng et al., to not only predict the spread pattern of influenza across Beijing but also come up with efficient intervention strategies that can help mitigate the spread under various scenarios.

Figure 4 Agent based simulation of influenza epidemic in Beijing. Reproduced from (Zeng *et al.*, 2009)

2.3.2. Contact Patterns in Agent-based Simulations

Contact patterns in populations play significant roles in prompting the spread of epidemics. Hence, it is of great importance to precisely replicate real life human contact patterns in simulated environments in order to create range of possible scenarios with reasonable accuracy. To capture such minute details, researchers gather data on population behaviors through participatory surveys, wearable sensors and ubiquitous computing devices such as smartphones. Edmunds et al. have conducted a large scale survey of various segments of populations across several English cities to approximate contact patterns (Edmunds, O'Callaghan and Nokes, 1997). Social networks are also widely used as a proxy to codify contact patterns in agents (Mei *et al.*, 2010; Duan, Qiu, *et al.*, 2013). In such approaches, the research community is split into two groups where one group assumes that each agent comes in contact with all of its neighbors at each time steps; while the other group assumes contact patterns are stochastic and heterogeneous, where contacts between certain subpopulations of agents depends on various real-life social contexts that they try to replicate (Beutels *et al.*, 2006).

Weighted networks have been used widely in recent times to capture the heterogeneous contact patterns in agent populations. In these networks, the contact probabilities between agents are usually a function of the edge weights as shown in equation 2 below (Duan, Cao, *et al.*, 2013)

$$p_{ij} = \frac{w_{ij}}{\sum_{k \epsilon N_i'(t)} w_{ik}} \qquad 2$$

Here p_{ij} represents the probability that agent i and j comes into contacts with each other; w_{ij} is equivalent to the weight of the edge linking the two agent nodes i and j; the $N_i'(t)$ denotes the set neighbors that agent i came in contact with at any arbitrary time t. As can be imagined from real life scenarios, individuals come in contact with different sets of people throughout the day and hence $N_i'(t)$ is a temporally changing quantity. Temporal networks (Masuda, Klemm and Eguíluz, 2013) and localized spatial contact patterns (Wang *et al.*, 2013) have also been explored to represent detailed agent contact patterns.

2.3.3. Mobility Patterns in Agent-based Simulation

Human mobility trends are one of the most significant factors that determine the transmission of a virus in the population (Codling, Plank and Benhamou, 2008; Balcan *et al.*, 2009; Merler and Ajelli, 2010). One of the more popular models used for studying mobility patterns is Codling et al.'s, random walk model. In this model, movement patterns of individuals are mathematically formulated as all possible paths, within a finite space, that consists of sequences of random steps. Another promising random walk model is the Levy flight model that breaks down the human movement displacements as a set of independent identically distributed random variables (Brockmann, Hufnagel and Geisel, 2006). In this model, mobility distances are well approximated by power law distribution; meaning that probability densities at any of the random variables that denote the human movement distances are proportional; that is, the relative likelihood between any of the movements events, characterized by the random variables, is the same.

Another encouraging direction for predicting agent mobility patterns is the weighted two mode networks (Opsahl, 2013). Such networks combine both spatial and contact networks in order to illustrate spatio-temporal contact patterns. Weighted two mode networks separate out the nodes into two types – the agent nodes and the location nodes. There are interconnecting edges between agent nodes and location nodes and between agent nodes themselves. The corresponding weights of the edges between agent nodes and location nodes represent the relative amounts of agent movements at a particular location. The weights on the interconnecting edges between the agent nodes themselves denote the contact probabilities between agents at the same location. In such networks, the movements of the agents can be modelled as discrete time events where agent behaviors can be defined by activity based models (Min, Goh and Vazquez, 2011).

2.4. Advantages and Limitations of the Different Approaches

In the previous sections we covered details about three most prevalent ways of modelling an epidemic. Each approach had their own advantages and disadvantages. Although the purely mathematical modelling approach has the advantage of providing holistic, macroscopic events of an outbreak, they lack the flexibility to model heterogeneous mixing among populations. On the other hand, contact network-based models have the advantage of incorporating the heterogeneous mixing and community structure but these types of models have to make lots of initial assumptions which might introduce bias further downstream in the calculation. Finally, agent-based models are able to capture granular level details of epidemic models by modelling individuals in populations as agents; however, agent based models are computationally very expensive. In the section below we have covered the advantages of each of the three modelling techniques in greater details.

2.4.1. Literature Evaluation of Mathematical Models

Mathematical models are particularly suitable to provide quick, generalized view of a pandemic event by defining macroscopic quantities such as epidemic threshold, basic reproduction number and estimated final epidemic sizes. However, although these quantities are able to provide public health officials and policy makers to implement policies initially, they quickly become deprecated if not continuously updated with new ground data. Furthermore, these predicted quantities are derived through various averaging and aggregation techniques; hence, factors of vital importance such as average infection and recovery rates cannot properly describe the complex heterogenous, micro-events.

Expanding on the limitations of mathematical models, they ignore microscopic aspects of pandemic events such as change in behavior and movement patterns of individuals in response to a pandemic. Ignoring such details make most mathematical models inadequate to capture microscopic factors that might contribute to transmission dynamics, which is one of the most important elements of a disease outbreak event to consider.

Researchers have focused on some of the limitations of mathematical models by focusing on extending the widely used compartmental model. (Zhang *et al.*, 2005; Chowell *et al.*, 2009) have taken into consideration the individual characteristics and behaviors and constructing community structures according to population health states, ages, occupation statuses and potential of certain segments of the population to be *super spreaders*. Researchers such as Fenichel et al. went further to incorporate human behavior trends to deduce the effective infection rates (Fenichela *et al.*, 2011). Their approach has been effective in illustrating the temporal variation in progression of epidemics.

2.4.2. Literature Evaluation of Network-based Models

Unlike purely mathematical epidemic models, complex network models can capture details on heterogeneity in population structures and the interaction patterns by formulating evolving node degrees and edge weights. Numerical simulations performed on individual based models (IBMs), that are built on top of complex networks, can be used to explore detailed temporal evolution of spread pattern of a disease during an outbreak event (Rahmandad and Sterman, 2008).

However, epidemic events in human societies are highly complex. While complex network models can be used to capture some of the complexity, they still lack the abilities to fully explain all the intricacies of complex human activities that plays a part in epidemic spread. With the advent of more computational powers, researchers have tried to incorporate more features into complex network models to better capture the highly dynamic human contact patterns over large temporal scales (Lee *et al.*, 2012).

2.4.3. Literature Evaluation of Agent-based Models

Out of the three types of epidemic models discussed so far, agent-based models provide the greatest level of details with regards to projecting wide range of future scenarios that might play out. Researchers have been able to combine the strengths of both mathematical and complex network models in agent based models (Duan and Qiu, 2012). Such hybrid agent based models, have the advantage of not only monitoring the effects of human mobility patterns on epidemics but also map out detailed spatio-temporal spread of disease over geographical locations (Perez and Dragicevic, 2009).

The one major disadvantage for agent-based models is that they are very computationally expensive and overly dependent on data. However, with the rise of ubiquitous computing and high power computational resources these limitations are slowly being addressed (Campuzano *et al.*, 2011). And although algorithmic complexity for agent based models have made the application of such models, for real life applications, prohibitive; recent advances in smarter algorithms are lowering the barrier of employing agent based models for epidemic event simulations (Bonabeau, 2002).

3. Methodology

As mentioned in the introduction, this report will focus on two approaches for modelling epidemic events – mathematical models based on partial differential equations and computational approach leveraging agent-based modelling using Netlogo, a multi-agent modelling integrated development environment (IDE). The aim is to compare the two approaches and highlight potential benefits and pitfalls.

Since mathematical models are more suited to represent macroscopic quantities of an epidemic outbreak, so we have used various published demographic and Covid-19 related data for three cities in England – Leicester, Bradford and Blackburn, in order to model the cumulative total number of infected in those three cities till August 13th, 2020. Table 2 below shows the total number of cumulative cases of per 100,000 population in each of the regions in England. For comparison, the cases per 100,000 for the three cities under study are added in the corresponding positions to the regions they are located in. Figure 5 below illustrates the spatial variation of the Covid-19 cases across the country, highlighting the regions of the three cities. As can be seen from Table 2, we chose these three cities in particular because these places have the highest recorded cumulative cases per 100,000 of the population till August 13th (Cities, 2020).

One thing to note, the mathematical model does not take into account the granular details such as the change in behavior of people in response to pandemic event and the government-imposed lockdown in England. Hence, the numbers represented by the model could be comparable to potentially catastrophic events that might have taken place if no change in social norms were observed during the Covid-19 outbreak.

Region	Population size	Total Number of Covid-19	Per, 100,000	City Under Study	Cases per 100,000
North East	2,570,100	15895	618.46		
North West	7,073,900	51346	725.85	Blackburn	1129
Yorkshire and the Humber	5,317,300	35092	659.96	Bradford	1063
East Midlands	4,671,800	25709	550.30	Leicester	1263
West Midlands	5,733,100	29450	513.68		
East of England	6,087,300	25709	422.34		
London	8,810,800	38049	431.85		
South East	8,913,000	36978	414.88		
South West	5,451,500	14241	261.23		

Table 2 Variation of Covid-19 across England

For the agent-based modelling, we have tried to include a few more details into our model. We have used alternative approach to incorporate the change in behavior in the population with the emergence of a pandemic scenario. We have performed data mining operations on an online news portal articles to measure the level of Covid-19 news coverage across the community. Depending on the level of coverage, we have designed our model in a way that agents reacted to such news by wearing masks and maintaining social distancing and self-isolation. The idea was to leverage proxy data to build concept of social awareness among the artificial agents and observe whether or not any changes in social behavior occurs and whether that leads to any observed changes in the number of infected cases of an infectious disease amongst a population. Use of proxy dataset is being increasingly used by researchers to predict emergence and spread of infectious diseases in recent times (Nsoesie *et al.*, 2020).

Therefore, the methodology part will be broken down into four separate sections, each dealing with different aspects of epidemic modelling. Section 3.1 will cover data used to compute the terms that have been used for both the mathematical model and the agent-based model. The epidemic modelling terms that we have used are primarily related to the ongoing Covid-19 pandemic and have been gathered from publicly available datasets from various sources such as World Health Organization (WHO), British Medical Journal (BMJ) and Office of National Statistics (ONS).

Section 3.2 will cover details on how we have developed our mathematical model. This includes details into the compartmental SIR model that we have constructed, the partial differential equations used in the model and the epidemic modelling quantities that we have used. This section will also contain the base case scenarios that our model produced for the three cities.

Section 3.3 will detail how we have gathered news articles covering Covid-19. We have detailed on the data mining methodologies that we have used to gather all the news data pertaining to the peak period of the pandemic from March to June, 2020 (London School of Hygiene & Tropical Medicine, 2020). The idea was to model the effects of news coverage in changing the social norms of the population during times of pandemic event. The results from this section was integrated with the agent based modelling, in section four, to see if changes in social behavior can lead to better outcomes in terms of number of infected cases by reducing the transmission rate of the disease.

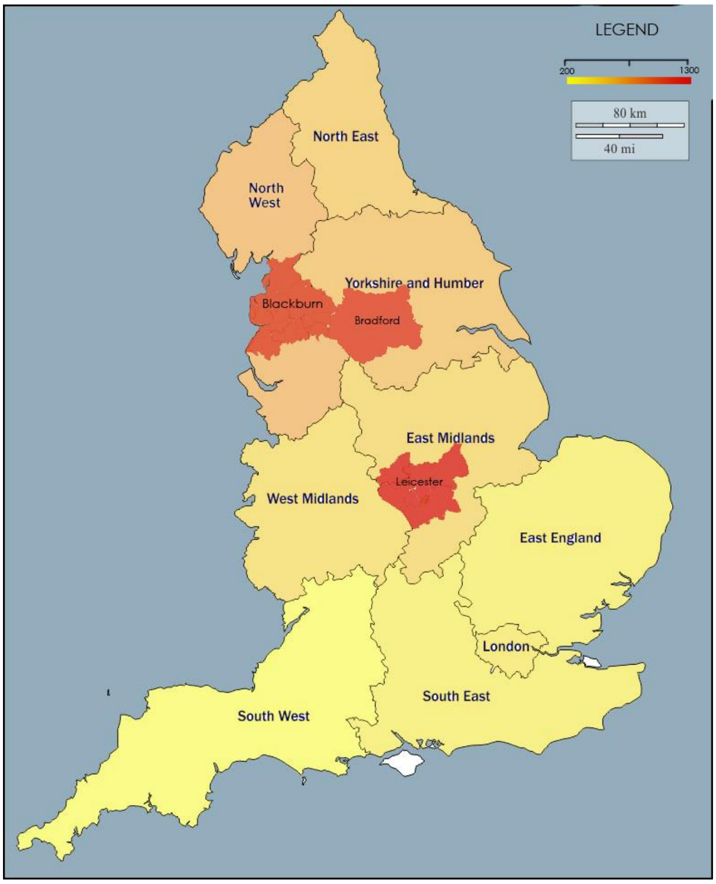


Figure 5 Spatial Variation in the number of Covid-19 cases per 100,000 of population across different regions of England. For illustration only.

Finally in section 3.4 we have detailed how we have designed our agent-based model. This includes details on splitting up the entire population into three subgroups of – children, adult and elderly. Authoring the agents, belonging to the different subgroups, with different characteristics and contact patterns. Then we integrate the results from section three to our agent-based model and see how individual agents from each subgroup react to the news coverage of Covid-19. Similar to mathematical modelling in section two, we produce graphs of cumulative number of infected individuals over the study period.

3.1. Epidemic Modelling Data and Information

This section contains all the data and information that we have sourced from publicly available resources related to Covid-19.

3.1.1. City Population and Demographic Breakdown

We began by gathering the basic demographic data for the areas under consideration – the cities of Leicester, Bradford and Blackburn. The data for each of the cities was gathered from (Urbistat, 2020). Table 3 below shows the demographic breakdown of the three cities according to three age groups – Children (0 – 17 years), Adults (18 – 64 years) and Elderly (65 and above).

City	Total Population	Age Group	Age Group Name	Percentage of the Population	Count
		0 - 17	Children	23.71	83,778
Leicester	353,540	18 - 64	Adult	64.56	228,226
		65 and above	Elderly	11.75	41,536
		0 - 17	Children	26.48	141,606
Bradford	534,800	18 - 64	Adult	58.88	314,875
		65 and above	Elderly	14.64	78,319
		0 - 17	Children	25.98	38,657
Blackburn	148,772	18 - 64	Adult	59.71	88,824
		65 and above	Elderly	14.31	21,291

Table 3 Demographic Breakdown of the three cities according to age

3.1.2. Contact Rates among Different Age Groups

According to the widescale survey study conducted by Mossong et al. (Mossong *et al.*, 2008), different age groups interact with each other in different ways. Table 4below shows the number of daily contacts between individuals in each age band in Great Britain.

Age Bands	Children	Adult	Elderly
Children	7	5	1
Adult	2	9	1
Elderly	1	3	2

Table 4 Daily Average Number of Contacts between Individuals in Different Age Bands

The diagonal in the table indicate the daily number of contacts made by each individual with others in the same category. The non-diagonal numbers indicate the number of contacts made by individuals from separate categories.

3.1.3. Probability of Transmission per Contact (ρ)

The probability of transmission per contact indicates the per contact of a susceptible person becoming infected when they come in contact with an infected person (Kucharski *et al.*, 2020). This quantity depends on two factors, mainly the closeness of contact and the level of infectiousness of the disease in general. Since it is not yet possible to say for certain whether the probability of transmission is different across genders and age groups in England (*Coronavirus (COVID-19) infections in the community in England - Office for*

National Statistics, no date), so we have relied on the numbers presented by He et al. (He *et al.*, 2020) for both symptomatic and asymptomatic cases which has been averaged for the entire world. The equations below define the two transmission rates obtained for our calculations.

$$\rho_{symptomatic} = \frac{126}{2001} = 0.063 \quad 3$$

$$\rho_{asymptomatic} = \frac{6}{146} = 0.041 \quad 4$$

3.1.4. Proportion of Symptomatic to Asymptomatic Cases:

It is more or less established that the number of asymptomatic cases far outweigh the number of symptomatic cases. Through rigorous field studies of Covid-19 test results, a study published in the British Medical Journal has suggested that the ratio of asymptomatic to symptomatic cases is roughly at 78:22 according to their findings (Day, 2020). For our purposes of epidemic modelling, we will stick with this ratio.

3.1.5. Recovery Rate from Covid – 19 (γ)

Since it is still to early to properly measure the geographic variation in recovery rates for both asymptomatic and symptomatic cases, so we have stuck with the globally established rate of recovery of 14 days (*Coronavirus Recovery: Rate, Time, and Outlook*, 2020)]. Therefore, recovery rate, γ is defined as,

$$\gamma = \frac{1}{Number\ of\ days\ to\ recover\ from\ Covid - 19} = \frac{1}{14}\ days^{-1} \quad 5$$

3.1.6. Incubation Period

The incubation period is defined as the number of days between exposure to an infection to the onset of infections. The widely established number for England is set at 5 days by the NHS (*Coronavirus (COVID-19) - NHS*, 2020). Therefore, we will use this number for our calculations.

The incubation period is important since it has been proven by WHO that exposed individuals are also liable for transmission of Covid-19 (World Health Organization, 2020b). With England's contact-tracing system being implemented only recently, exposed individuals in the past have unknowingly spread the virus before symptoms set in. For our modelling purposes, we use the term *h,* which is simply defined as,

$$h = \frac{1}{incubation\ period} = \frac{1}{5}\ days^{-1} \qquad 6$$

3.1.7. Age Standardized Mortality Rate for Covid – 19:

The age standardized mortality rate for Covid – 19 is estimated to be at 781.9 per 100,000 for males ($\mu_{covid-19,male,England}$) and 439 per 100,000 for females ($\mu_{covid-19,female,England}$) in England (Office for National Statistics, 2020b). Since it is not yet clear whether these rates vary for different age bands, so we will average it out for both males and females across all age groups living in Leicester, Bradford and Blackburn. Table 5 below shows the Covid – 19 related mortality rates for the three cities.

City	Male Population	Female Population	Calculated Covid - 19 Mortality Rate
Leicester	176,820	176,720	0.006105
Bradford	263707	271,093	0.006086
Blackburn	74,495	74,277	0.006107

Table 5 Covid-19 Related Mortality Rates in Leicester, Bradford and Blackburn

The equation for calculating the Covid-19 related mortality rate is as shown below.

$$\mu_{covid-19} = \frac{male\ population}{total\ population} * Age\ standardized\ \mu_{covid-19,male,England} + \frac{female\ population}{total\ population} * Age\ standardized\ \mu_{covid-19,female,England} \qquad 7$$

3.1.8. Non-covid related mortality:

The non-covid related number of deaths in the three cities are obtained from ONS data from March to June (Office for National Statistics, 2020c) and extrapolated out to August. Table 6 below shows the total number of non-covid related deaths. The non-covid related mortality rate ($\mu_{non-cov}$) is then calculated using equation 8 shown below the table.

City	Total non-Covid deaths till June	Total non-Covid deaths extrapolted to August	Calculated non-covid Mortality Rate
Leicester	806	1343	0.0038
Bradford	1603	2672	0.005
Blackburn	1346	2243	0.004772

Table 6 Non-Covid-19 related Mortality Rates in the three Cities

$$\mu_{non-cov} = \frac{Total\ non-Cvoid\ deaths\ till\ August}{Total\ Population\ of\ the\ city} \quad (8)$$

3.1.9. Cumulative Number of Cases of Covid – 19 in the Three Cities

The total number of cumulative cases till August 13th is listed in Table 7 below. The data was obtained from (Cities, 2020).

City	Cumulative Cases per 100,000	Cumulative Number of Cases till Aug 13
Leicester	1263	4465
Bradford	1063	5685
Blackburn	1129	1680

Table 7 Cumulative Number of Covid-19 cases in the three cities till Aug 13

3.1.10. Timeframe of Analysis:

The timeframe for the study was set at 150 days, roughly representing from March 13th to Aug 13th – that is the period covering the start of the rise of Covid-19 in England (Office for National Statistics, 2020a) to the latest date.

3.2. Mathematical Model

For our mathematical model, we constructed an extended SIR model with a few additional terms and states. Table 8 below shows the schematic diagram of our model with all the necessary information. As mentioned earlier, for each of the cities the entire population is broken down into three age bands – children (0 – 17 years of age), adults (17 – 64 years of age) and elderly (65 and above). After that, nearly the whole population are placed in the susceptible category (S box) and two infected cases are introduced, one in each of the I_S and I_{NS} boxes. For our simulations, we traversed through all combinations of infected cases from each of the age groups. For example, the initial simulation run for each city contained a combination of one symptomatic index case from the adult age band and one non-symptomatic index case from the children age band. Table 8 below shows all the combinations of index cases that we have simulated for in each city. Therefore, for each city, nine simulation runs were performed to capture all possible outbreak scenarios that might play out.

Symtomatic Infected	Non-symptomatic Index
Adult	children Adult Elderly
Children	children Adult Elderly
Elderly	children Adult Elderly

Table 8 All combinations of initially infected cases that were used for simulation runs for each city

As infected people move about in their communities, they come in contact with the susceptible population and transmit the disease to them. We setup our model to account for the average daily contacts between individuals from each of the age bands according to the data from Table 4. Since Covid-19 is a complex disease that give rise to both asymptomatic and symptomatic cases, so we predict that asymptomatic cases will unknowingly interact with other in the normal rate. As for symptomatic individuals, we have modelled them to reduce their contact rates with others significantly. The C_{NS} and C_S matrices in equations 9 and 10 below define the contact rates for the non-symptomatic and symptomatic cases respectively. Notice, C_{NS} is equivalent to the normal daily contact rates.

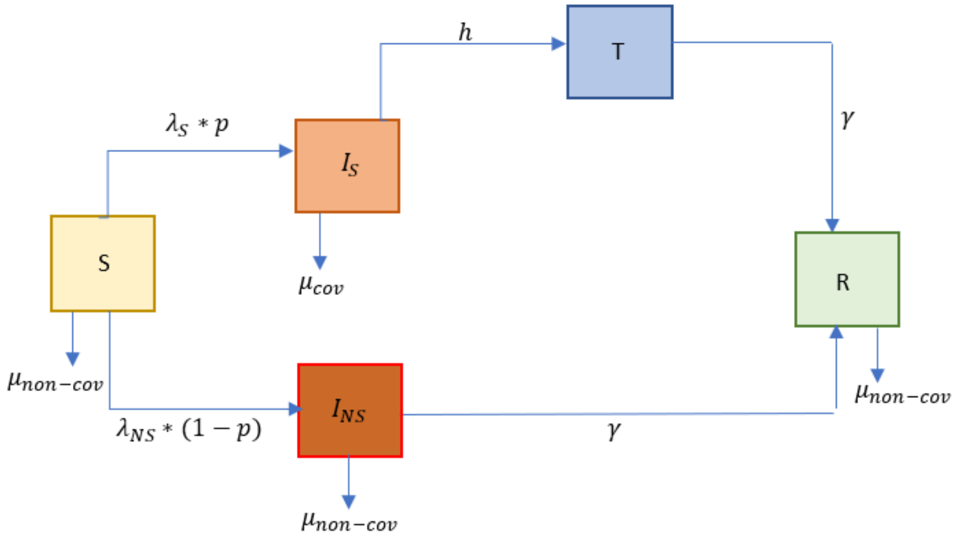


Figure 6 SIR Model developed for Modelling Cumulative number of infective cases in the three cities

The rate at which susceptible individuals are infected is defined by the λ term, which is the force of infection. Since we have two type of infected cases, therefore we have defined two λ terms - λ_S and λ_{NS} for symptomatic and non-symptomatic cases. The *p* term represents the proportion of the cases who are symptomatic and the *(1 - p)* term refers to the proportion who are asymptomatic. Therefore, the $\lambda_S * p$ and $\lambda_{NS} * (1- p)$ represents rates at which individuals are shifted from the susceptible category to each of the infected categories.

Infected individuals with symptoms usually have more complications and hence take more time to recover than asymptomatic cases. The entire recovery period for both symptomatic and asymptomatic cases can last anywhere between 14 – 21 days (Bi *et al.*, 2020). Therefore, in our model, infected individuals go through an additional treatment stage (T box) before recovering. The rate at which individuals move from I_S to T is defined by *h*, which slows down the process by further five days. From the treatment stage, infected individuals are transferred to the recovery state (R box) at a rate of γ. On the other hand, non-symptomatic cases move to the recovered state directly at a rate equal to γ.

Few significant things to note for our model are – we assumed that the while deaths, from both Covid-19 and non-Covid, are included in the model there are no additions to the population from new births or migration. We also assumed that Covid-19 related deaths, denoted by μ_{cov}, are only associated with individuals who are symptomatic. While the non-Covid related deaths, represented by $\mu_{non-cov}$, occur from the susceptible, asymptomatically infected and recovered populations. Equations 9 – 17 defined below were used to construct our mathematical epidemic model.

We used the programming language R for coding up our model. Specific packages that we have made use of include the *deSolve* package for solving the differential equations, the *ggplot2* for plotting graphs and illustrations and the *reshape* package for handling dataframe.

$$C_{NS} = \begin{pmatrix} 7 & 5 & 1 \\ 2 & 9 & 1 \\ 1 & 3 & 2 \end{pmatrix} \quad (9)$$

$$C_S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (10)$$

$$\lambda_{NS} = \rho_{NS} * C_{NS} * \frac{I_{NS}}{N} \quad (11)$$

$$\lambda_S = \rho_S * C_S * \frac{I_S}{N} \quad (12)$$

$$\frac{dS}{dt} = -\lambda_S * p * S - \lambda_{NS} * (1-p) * S - \mu_{non-cov} * S \quad (13)$$

$$\frac{dI_S}{dt} = \lambda_S * p * S - h * I_S - \gamma * I_S - \mu_{cov} * I_S \quad (14)$$

$$\frac{dI_{NS}}{dt} = \lambda_{NS} * (1-p) * S - \gamma * I_{NS} - \mu_{non-cov} * I_{NS} \quad (15)$$

$$\frac{dT}{dt} = h * I_S - \gamma * T \quad (16)$$

$$\frac{dR}{dt} = \gamma * T + \gamma * I_{NS} - \mu_{non-cov} * R \quad (17)$$

3.3. Data Mining:

For the purposes of data mining for Covid-19 related terms we had to decide one specific website as we did not have the computational resources to perform extraction from multiple sources. We had to make sure that the source is reliable and widely read by the population across England. We also had to make sure that the website we were going to choose also allowed for web scrapping to be performed. We found that the Times online news portal[1] matched all of our criterion. The portal was ranked as the third most popular online daily by YouGov in 2018 (YouGov, 2018). Also, the online portal for the Times has an html-sitemap section where it archives all its published articles, making it ideal for web scraping and mining for Covid-19 related information. The data extraction was performed from the period covering the peak times between March through June, 2020.

For our purposes we have used python-based libraries such as *Selenium* and *Gekodriver* for mining the data and *Regular Expressions* to find patterns in extracted text data.

In the first step, we analyzed the latest results from Google Trends (Google Trends, 2020) to find the most searched terms related to Covid-19. Although we received many such terms, for brevity we only selected the top few. Table 9 below shows the most searched terms. After getting the terms, we pre-processed the text data to include all variations of caps lock as well to create a bag of words that we were going to leverage for our data mining operation.

Count	Terms
1	Coronavirus
2	Covid-19
3	social distance
4	lockdown
5	mask
6	quarantine
7	pandemic

Table 9 Bags of Words related to Covid-19

We have broken down the entire data mining operation in an ordered series of five steps that are expanded below. For clarity, we have also developed a Unified Modelling Language (UML) diagram showing the systemic overview of the entire data mining operation. Figure X below contains the UML diagram. Each step in the section below corresponds to the steps numbered in the UML diagram.

[1] https://www.thetimes.co.uk/html-sitemap

Step 1

We used the Selenium webdriver to open a browser in the background to access the archived articles from the Time website through the Uniform Resource Locator (URL) addresses of those articles. The start date was set for the third week of March. After accessing the URLs of each article, the scrapper moves on the next steps below.

Step 2

Since UK is still under the GDPR regulations, accessing the archived storage leads to a pop-up page asking for permission to accept cookies. Handle the pop-up pages, we located the button that accepts the terms and conditions for the archive page, as well as all webpages stored in the archived section. A fail-safe redundant method was also developed to avert any potential disruption such unresponsive webpages.

Step 3

Once the consent form is out of the way, the program goes through the list of news articles per week. Each article title is stored is held in an array and each element of the array is checked to see if they contain any of the words that are in the created bag of words. If an article contains any of the words then it is separated out and placed in a different array. The program then performs the same operation for all the articles in that particular week.

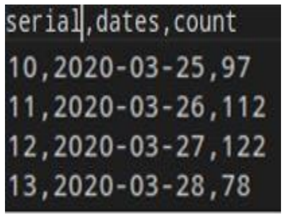
Step 4

After extracting all the Covid-19 related articles, the program moves on to extracting the dates of the extracted articles. During this step, the program also aggregates all the articles from each date and keeps a total count of those per day. This step was done to see the daily variation in the number of Covid-19 related published articles in the Times.

Step 5

When we examined the HTML page we found that the developer of the website used two types of classes to handle the webpage's datetime information. They are called *responsiveweb__DatePublicationContainer-tglil3-1* and *css-901oao* respectively. These two classes (a characteristic of HTML elements) differ in their names and the way they are interpreted by the webdriver program. When the driver fetches the elements with class name *responsiveweb__DatePublicationContainer-tglil3-1*, then only the datetime information can be seen. But when the *class css-901oao* is fetched, lots of information other than the datetime are also fetched. Therefore, we had to develop two different methods to handle the two separate classes. We designed method one to handle just the date and method two to find the datetime information from an array of other information.

Method one was used to handle the *responsiveweb__DatePublicationContainer-tglil3-1* class and method two was used to handle the *css-901oao* class. The raw results are then converted to csv format for further analysis. Figure 7 shows the raw output from a section of the text mining operation. The serial represents the timestep along with the date and the count of the number of Covid-19 related articles published in each particular date.



```
serial,dates,count
10,2020-03-25,97
11,2020-03-26,112
12,2020-03-27,122
13,2020-03-28,78
```

Figure 7 Raw output from the program

Steps one through five above are repeated until all the Covid-19 related articles are traversed through and accounted for. The scraper program is run from the third week of March to the second week of June. After the mining operation is performed for the last week (i.e. The second week of June), the program exits.

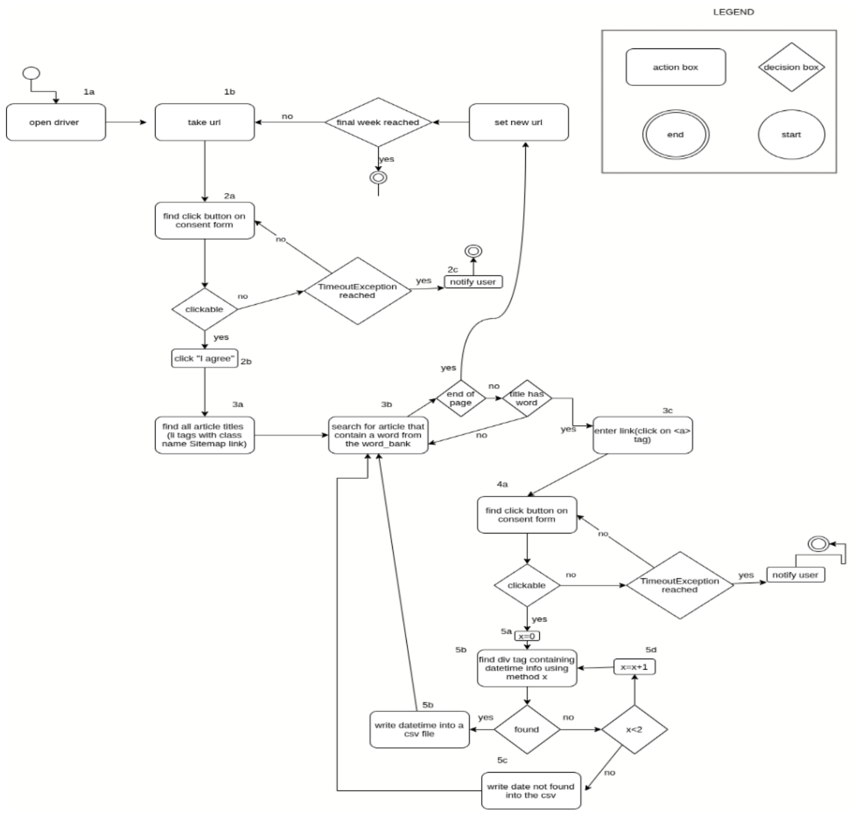


Figure 8 UML diagram of the entire text mining operation from the Times online news portal

After obtaining the CSV file, containing all the dates and the corresponding number of articles, a graph was plotted and the best fit line to the datapoints were drawn. The equation of the line is then used in the next section – Agent Based Modelling, to define the level of social awareness among an artificially generated population representing all the three cities. Figure X below shows the plot of the datapoints and the equation of the best fitting line.

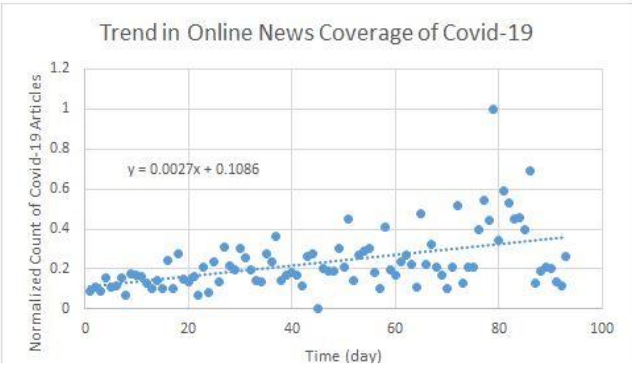


Figure 9 Plot of number of Covid-19 related articles in the Times versus time, from March to June

3.4. Agent Based Modelling

We have created an Agent Based Modelling using NetLogo to simulate the spread of Covid-19 in the city of Bradford, Leicester and Blackburn separately based on SIR model of disease spread. SIR based model assumes that the population contains three types of agents – Susceptible to the disease, Infected and Recovered who become immune to the disease and cannot get infected again or infect others. In our model, the population is further subdivided into three sub groups – children, adults and elderly. The agents were designed in a way such that members of the population can only interact with each other following the contact matrices described in equations 9 and 10.

The simulation world has 32 by 32 patches so the area is 1024 square unit. The world wraps around both horizontally and vertically. The simulation is setup by creating the agents scattered randomly in the world. The population and number of members from each group is fixed for each city. One adult agent is randomly chosen to be infected initially.

The spread of Covid-19 from the initially infected agent is then observed for 90 timesteps. Here, each timestep is considered as one day. In each timestep, the agents can move a maximum number of 12 steps (that's the maximum number of contacts anyone from any of the group can make with others – inline with the contact matrix) in random direction. Infected agents are then modelled to infect other agents within a radius of 1.5 unit; with an infection probability of 4.584% as base case, 0.6876% if either one of them has any means of mitigation and 0.1031% if both have any means of mitigation. Here mitigation refers to a combination of social distancing and wearing masks.

The 4.584% rate was found by averaging over the infectibility rates of both asymptomatic and symptomatic cases. The 0.6876% probability is found by reducing the base case infectivity by 85% and the 0.1031% number represents a further reduction of the probability of infection by 85%. These numbers reflect the findings from a WHO backed study, that suggested that wearing of masks and maintaining of social distancing by either in a pair of interacting individuals reduce the chances of transmission by 85% (Chu *et al.*, 2020). We have also incorporated the incubation period by making each infected agent move into quarantined state after 5 days of getting infected. While quarantined, the infected agents do not infect others. After 14 days of quarantine, they recover and gains long lasting immunity, meaning they do not infect others and neither do they get infected by others.

Figure 10 below shows screenshot of the world at a random timestamp during the simulation. For a good visual observation, different shapes and color are used for different types of agents. Three different shape is used to indicate Child, Adult or Elderly agent. Susceptible agents are in gray, infected agents are in red, quarantined agents are in yellow and recovered agents are in green color. If an agent is using any means of mitigation then a blue mask is used on the agent's head.

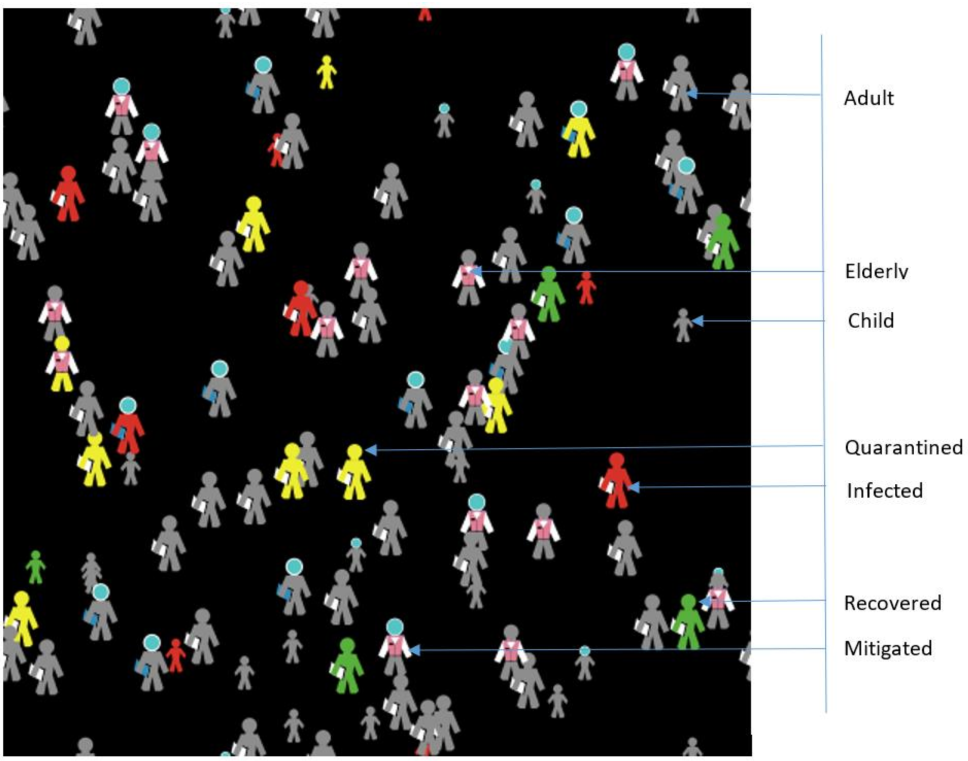


Figure 10 Screenshot of the subsection of agent population during a simulation run

For each city, the first simulation is run without any mitigation strategy in effect and another one with an increasing rate of mitigation. The rate of mitigation is calculated using following equation,

$$\textit{Mitigation Rate} = 0.0027 * \textit{days} + 0.1086$$ [18]

The equation is determined from the trend line found from the previous section where we performed data mining. We consider that as the pandemic progresses and the number of published articles related to Covid-19 increases, the social awareness increases in proportion to that. Therefore, mitigating strategies, such as wearing of masks and maintaining of social distancing, are adopted to reflect that change in social behavior among individuals in a population during an epidemic event.

One thing to note is that, due to lack in computation resources we had to scale the population and subpopulations of the three cities according to their real population numbers. Table X below shows the relevant data that were employed to model each of the three cities.

City	Real Populaton	Scaled Population	Age Band	Age Band Percentage	Age Band Population	Probability of Infection	Incubation Period	Recovery Days
Leicester	353,540	66	Children	24%	16	0.04584	5 days	14 days
			Adult	65%	43			
			Elderly	11%	7			
Bradford	534,800	100	Children	26%	26	0.04584	5 days	14 days
			Adult	59%	59			
			Elderly	15%	15			
Blackburn	148,772	28	Children	26%	7	0.04584	5 days	14 days
			Adult	60%	17			
			Elderly	14%	4			

Table 10 Demographic breakdown of the three cities and their corresponding scaled down population numbers

4. Results and Discussion

The results from the mathematical modelling and the agent modelling are presented separately in sections 4.1 and 4.2. Section 0 presents a combined discussion on results from our two models. Finally, section 4.4 presents a discussion on the assumptions and limitations of the report.

4.1. Mathematical Modelling

The mathematical model that was presented in the methodology section was run with each of the nine configurations of initially infected cases, as mentioned in Table 8 from section 3.2. For all of the cities under consideration we noticed that the final number of infected population were very similar. Therefore, for brevity and because the majority of the population in all the cities were in the Adult age band of 18 – 64 years of age, we will only present the simulation runs where both asymptomatically and symptomatically initially infected cases were from that age group. Graphs and tables related to the rest of the simulation runs are presented in Appendix 1. Figure 11shows the simulation run for our model when applied to Leicester, Bradford and Blackburn. Table 11 below provides the simulation results from our model. The table breaks down the number of predicted infected cases in each of the age bands, as well as the total number of people predicted to be infected from Covid-19 in each of the three cities.

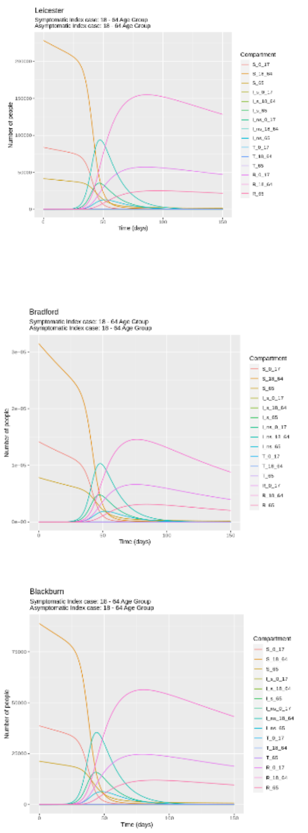


Figure 11 Simulation Graphs from SIR Model

City	Symptomatic Age Group	Asymptomatic Age Group	Children Infected	Adult Infected	Elderly Infected	Total Number of People Infected
Leicester	18 - 64	18 - 64	83675.36	227609.6	39813.35	351098.3
Bradford	18 - 64	18 - 64	141490	314454.6	76677.56	532622.1
Blackburn	18 - 64	18 - 64	38591.64	88618.02	20524.37	147734

Table 11 Simulation Outcomes predicting the number of infected cases in each of the three cities

4.2. Agent Based Modelling

The agent-based model that we had developed were applied to scaled population of each city from Table 10. Unlike the mathematical model, the initially infected cases were randomly selected. Graphs in Figure 12 to Figure 23 show the cumulative cases for each of each cities as well as the corresponding SIR curves of each of those cities for both when mitigation strategies were put into action and without any mitigation strategies. Although we ran many simulations for each of the cities, since most of the simulation produced similar results we have only provided one simulation run per city in this section.

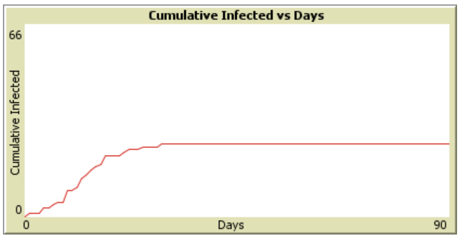


Figure 12 Cumulative number of infected cases in scaled down population in Leicester, without any mitigation strategies

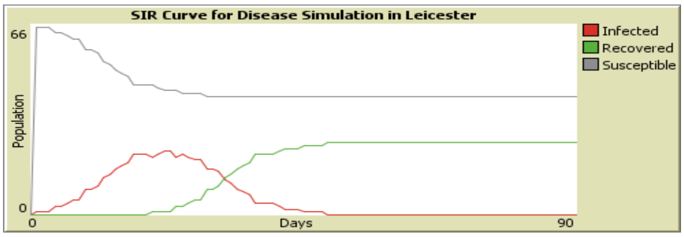


Figure 13 Corresponding SIR curve for Leicester without any mitigation strategy

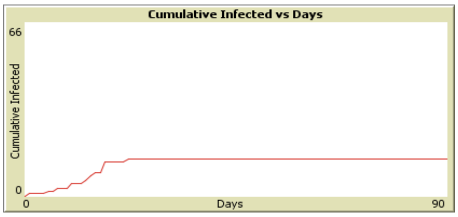


Figure 14 Cumulative number of infected cases in scaled down population in Leicester with mitigation

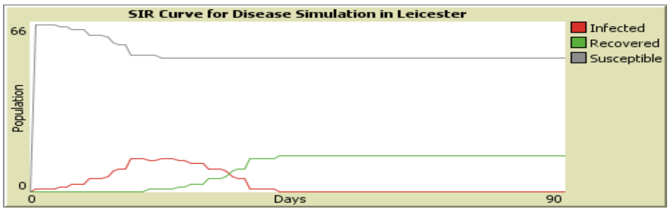


Figure 15 Corresponding SIR curve for Leiceter with mitigation strategy

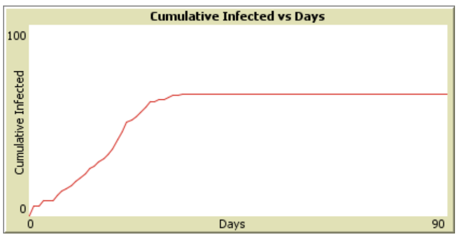


Figure 16 Cumulative number of infected cases in scaled down population in Bradford, without any mitigation strategies

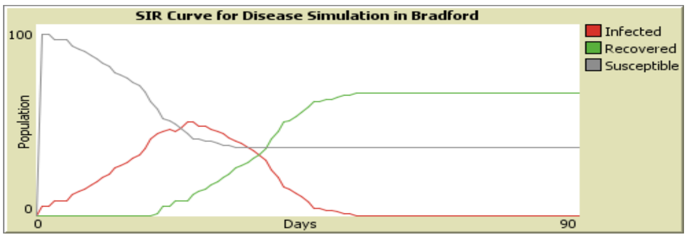


Figure 17 Corresponding SIR curve for Bradford without any mitigation strategy

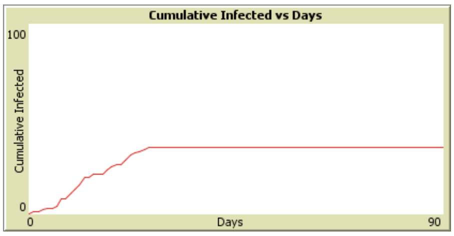


Figure 18 Cumulative number of infected cases in scaled down population in Bradford with mitigation

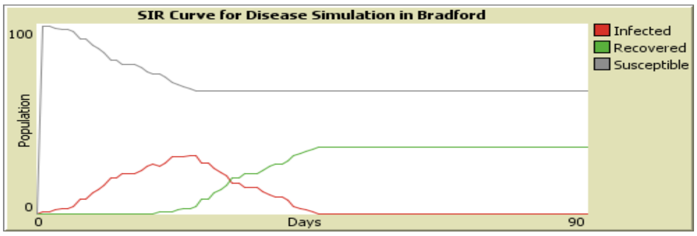


Figure 19 Corresponding SIR curve for Bradford with mitigation strategy

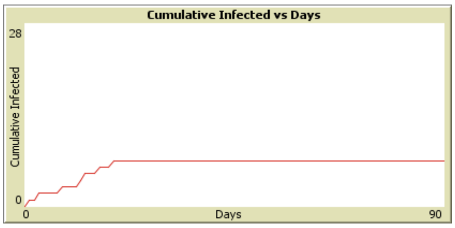


Figure 20 Cumulative number of infected cases in scaled down population in Blackburn, without any mitigation strategies

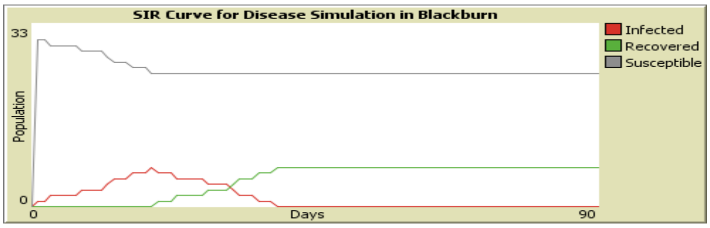


Figure 21 Corresponding SIR curve for Blackburn without any mitigation strategy

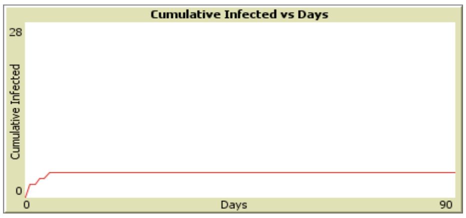


Figure 22 Cumulative number of infected cases in scaled down population in Blackburn with mitigation

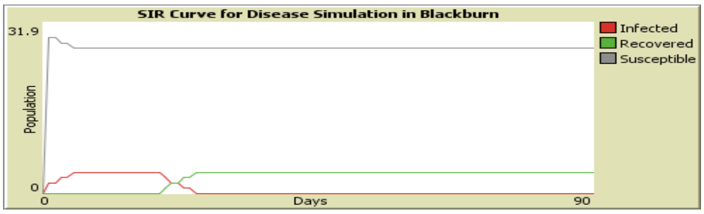


Figure 23 Corresponding SIR curve for Blackburn with mitigation strategy

4.3. Combined Results and Discussion

To conduct a thorough evaluation of all of our models, we measured the error between the predicted number of infected cases from our mathematical model and agent based model separately. The predictions from the agent based model was further broken down into two separate scenarios – the first scenario where no mitigation strategy has been applied to the population and the second scenarios where mitigation strategies such as wearing of masks and maintaining of social distancing rules have been taken into account. Table 12 and Table 13 below shows the error rates for our mathematical model and the agent based models respectively.

City	Total Population	Total Number of Reported Infected Cases	Total Number of People Infected from Mathematical Model as Percentage	Error Rate for Mathematical Model
Leicester	353540	4465	351098.3	98.72827638
Bradford	534800	5685	532622.1	98.93263911
Blackburn	148772	1680	147734	98.86282102

Table 12 Performance of Mathematical Model

City	Scaled Down Population	Scaled down Total Number of People Infected from Agent Based Model without Mitigation	Approximate Number of People Infected from ABM without Mitigation	Error Rate for Agent Based Model without Mitigation	Scaled down Total Number of People Infected from ABM with Mitigation	Approximate Number of People Infected from ABM with Mitigation	Error Rate for Agent Based Model with Mitigation
Leicester	66	25	133914	96.66577057	14	74988	94.04571398
Bradford	100	64	342270	98.33903059	35	187176	96.96275163
Blackburn	28	7	37188	95.48241368	4	21252	92.09486166

Table 13 Performance of Agent based Models under two different scenarios

As can be seen from the tables above, the agent-based model under the mitigation strategy scenario performed the best. This might be due to the fact that agent-based models are able to capture granular level details about a population through simulation. The results prove that the mitigation strategies that were factored into our agent-based model were successful in reducing the transmission of the virus in the simulated environment. The restrictions put in place such as wearing masks reducing the number of contacts through social distancing goes to show that mitigation strategies do work to reduce the number of infected cases during a pandemic event. The next best estimate was from the agent-based model under the scenario that no mitigation strategy was placed in the population. Finally, our mathematical model was the worst performing one. Although we have taken into consideration that contact rates between symptomatic, infected individuals decrease, it was not enough to stop the spread of the epidemic in our model. As mentioned in earlier sections, mathematical models are a good way to generalize a situation and present worst case scenarios. Therefore, results from our mathematical model can be used as a baseline for further model development during the PhD.

4.4. Limitation of our Models

As mentioned in the introduction, the primary aim of this report was to focus on the different approaches of epidemic modelling. The mathematical model that we had developed took into consideration only three different age bands and the contact structure between each. However, more sophisticated models can be developed that can further break down the population into more number of age groups. Moreover, some of the assumptions that we have made are based on global metrics such as the recovery rate, ratio of symptomatic to asymptomatic cases, the incubation period and transmission rate among others.

Moreover, in our mathematical model, we initially consider everyone in the population to be susceptible, which is not necessarily the case as aged population are more vulnerable to Covid-19 than young children and adults (Esteve *et al.*, 2020). Also, the contact rates that we have used for our calculation represents an average figure and might be off for the cities we have studied. One omission that we have felt that could have improved the model was inclusion of an exposed state in between susceptible and infected states. The model can also be further tested by introducing stochasticity where individuals in the each of the subpopulation would be chosen randomly to be infected. Since data related to Covid-19 is still not very reliable, in the future Bayesian based modelling together with Markov Chain Monte Carlo (MCMC) based fitting methods can be explored to develop more reliable models.

In our agent based model, we had to scale down the populations of the three cities in order to run the simulations. Since Bradford has the highest population among all the cities, at 534800, so it's population was set at the maximum limit of 100 agents that is allowable in the Netlogo programming environment. The rest of the cities' populations were scaled down accordingly.

In our agent-based model we tried to incorporate the different age bands, contact rates as well as the introduction of mitigation strategies to capture granular patterns of disease spread in the population. However, our results from our ABM suffered due to the same reasons as our mathematical model – mainly depending on second-hand data and approximations. Furthermore, the equation we derived for measuring the social awareness, which in turn was used to dictate the mitigation rate in the population, was based on only one source of information. We also assumed that as more Covid-19 related news was generated, the population became more "aware" and started taking safety measure proportionally. But in reality that is hardly the case. Another important extension that we can make is to include spatial layout of each of the city into our model. In the future, to incorporate more complexity into the agent based models, we think it might be worthwhile to explore additional framework related to ABM such as MASON, Repast, EpiSims and other similar ones. Incorporating such features might produce better results through more sophisticated spatial agent-based models. One last thing to note is that, the timeframes for mathematical and agent based simulation differed by a bit. While the mathematical simulations were run for 150 time steps, the agent based models were only run for 90 time steps. This was because agent based model simulations are much more computationally expensive and our (underpowered) laptop was only capable of handling limited simulation runs. However, the numbers related to total number of infected cases that were obtained from the agent based simulation runs were linearly extrapolated to match those obtained from the mathematical models.

Although our models were not the most accurate, the methodologies that we had followed were robust and can be built upon in our future research. The confirmation that we received from the results of our agent-based model under the mitigation strategy scenario is inline with the findings from major world health agencies such as WHO. This partially validates that our models are able to capture broad level trends of an epidemic event, albeit with a high error rate. Furthermore, in the future models we can incorporate more complex evaluation methodologies such as sensitivity analyses to further evaluate our models.

5. Additional Section

This section contains an approximate framework that can be developed during the PhD to comprehensively tackle any future events of contagious disease outbreaks. The section has been detailed by considering the current case of Covid-19.

5.1. Future Plan and Further Development of Framework during PhD

An epidemic event is a highly complex and dynamic event. As such, making a geographical location pandemic resilient is a multidisciplinary undertaking. According to our research a comprehensive pandemic resilient framework might consist of a blend of three different disciplines: Mathematical and agent-based modelling and simulation; Bioinformatics; and GIS Analysis.

5.1.1. Mathematical and Agent Based Modelling and Simulation

Through modelling and simulation, generative SEIR epidemiological models can be developed that can be used a policy decision making and risk prediction tool during potential pandemic outbreaks. High resolution census, infrastructure, urban mobility data and public health data can be incorporated among others. The developed model will be run in

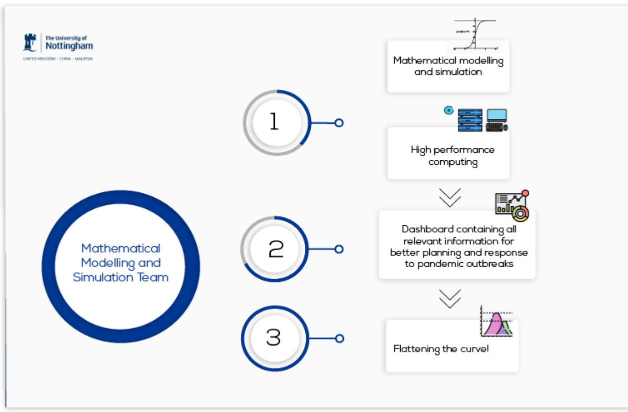


Figure 24 Mathematical and ABM based Research for Developing Pandemic Resiliency Framework

simulated environments under various conditions, to come up with robust ways to –

I. predict spread patterns across demographics and local sub-regions over time
II. predict vital metrics such as virus reproductive number (RO) in the local communities.

The work done through such modelling and simulation-based research will aim to provide a reliable macroscopic view to local policy makers to visualize and predict any potential clusters and enact policies to prevent new outbreaks. Figure 1 shows the workflow of such modelling and simulation research.

5.1.2. Bioinformatics Research

Through the collection of patient health data, testing results and other omics data, techniques in bioinformatics can be leveraged to uncover any hidden correlation between different strains of viruses that are prevalent at different geographical locations (Kupferschmidt, 2020). Such bioinformatics based research will used to – (i) perform genome sequencing of Sars-COV2

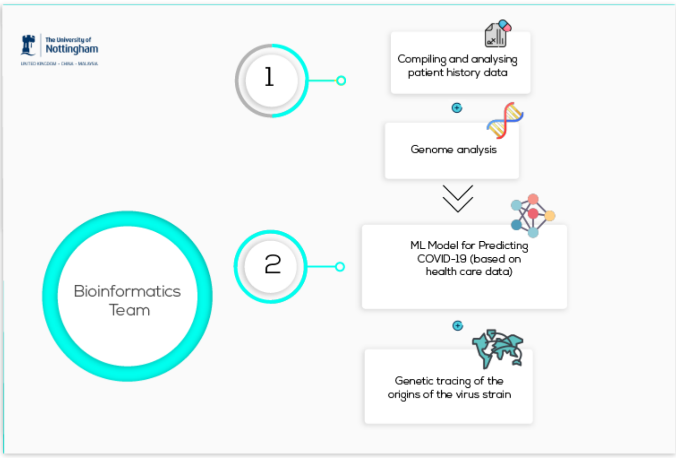


Figure 25 Bioinformatics Research for Developing Pandemic Resiliency Framework

samples to understand the particular strain of virus prevalent in the local population. (ii) Perform genetic tracing of the virus to understand the originating source and location of the virus in the region. (iii) develop a ML model for predicting new cases of nCov by solely assessing reported symptomatic and medical history data. This type of research will be essential to understand the transmission dynamics of the infection, aid in digital contact tracing efforts and help public healthcare officials in deciding the most optimized ways of enforcing control mechanisms such as local lockdowns and travel restrictions from certain high-risk geographies. Figure 2 illustrates the workflow of the bioinformatics research for such purposes.

5.1.3. GIS Analysis and Research

The GIS research can be leveraged to build out a comprehensive digital track and trace system. Such GIS research will involve –

I. Developing a smartphone app that will allow users to enter their symptomatic data and medical history. Then using the ML model developed by the bioinformatics team (step (iii) in the preceding paragraph), the app will try and predict the possibility of the user being infected. In the cases of possible Covid-19 infected users, the app will collect location data, with the permission from the users and by following all local and national data privacy laws. Through the collection of real-time user data, virtual social networks of the local population can be developed to understand granular level patterns of interaction, which can significantly aid in the overall contact tracing efforts.

II. Digitally map out – location optimized mobile testing centers for launching faster testing of potential new clusters.

III. Help the mathematical modelling and simulation team to incorporate the spatial interaction network in order to continuously update the overall model for more reliable prediction. Figure 26 shows the workflow of such a GIS based research.

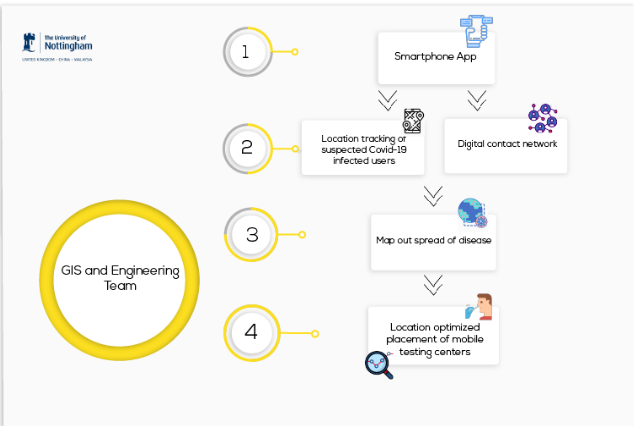


Figure 26 GIS Analysis and Research for Developing Pandemic Resiliency Framework

Figure 27 below shows the entire schematic flow of the three different strand of research and the collaborative links between each. In our final PhD research proposal, to be submitted in January 2020, we will focus on different elements of each of the three strands mentioned here to define the exact research questions that we will be answering from each of the disciplines.

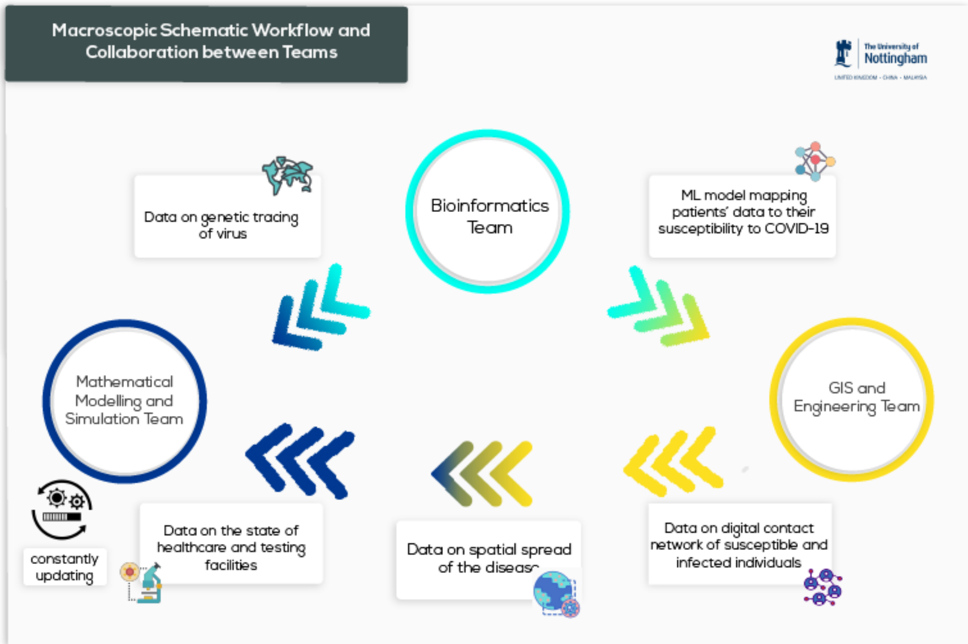


Figure 27 Schematic Workflow of a Pandemic Resiliency Framework

6. CONCLUSIONS

In this paper we conducted two of the main ways of epidemic modelling, namely the mathematical modelling and agent based modelling to predict the number of Covid-19 infected cases at three English cities – Leicester, Bradford and Blackburn. The three cities were chosen since they have reported the highest number of cumulative cases per 100,000 of the population in the country. Four our modelling we used publicly available datasets from major health agencies such as WHO and relied on statistical data related to Covid-19 in local communities through the Office for National Statistics. We also made use of global, macroscopic quantities related to Covid-19 to establish our age-stratified SIR compartmental model where individuals within each group had different rates of contact with others. We performed a total of 27 simulations for each of the cities with different types of initial infected cases – asymptomatic and symptomatic, coming from different age groups. However, we found that since the model was deterministic and depended on partial differential equations so the initial conditions player little part in determining the final number of infected cases in each of the cities. We investigated the results from our model and discussed their main limitations and covered some of the ways that the model can be improved in future iterations.

Although the SIR model was suitable for getting a generalized view of the pandemic situation such as obtaining the final number of cumulative infected cases over a certain period, there were only limited opportunities to include heterogeneity into the model. We made use of the contact matrices for modelling the transmission rates between the two different types of infected individuals, asymptomatic and symptomatic, and the susceptible population. However, this was not adequate to capture the complex heterogeneity that exists in everyday interacting pattern in human population. Therefore, to incorporate such heterogeneity we developed an agent based modelled that was age-stratified. We ran multiple simulations on each of the three cities both with and without mitigation strategies such as wearing of masks and maintaining social distancing. The mitigation rate was defined by an equation that we obtained by plotting the number of online news articles related to Covid-19 during a three month period from March to June. The idea was that as Covid-19 gets more coverage, more people will adhere to the mitigation strategies put forth by health care officials and the government. Although it was bit simplistic metric, the resulting agent based simulation runs with that mitigation rate produced results that were inline with recommendations from WHO – that mitigation strategies do reduce the number of infected cases compared to scenarios when no mitigation strategies are taken at all. We then discussed some future directions for agent based modelling such as incorporating spatial elements such as maps and infrastructure information and leveraging more sophisticated tools that will allow for code parallelization and faster and more detailed simulation runs.

Finally we talked about a comprehensive framework that can be developed for making geographical location pandemic resilient. Such framework will need to blend elements from - (i) Mathematical and agent based modelling and simulation; (ii) Bioinformatic and (iii) GIS Analysis and research. Each of the three components are vital for a fully functional framework and we provided a generalized overview of how those components are collected. For future we will delve a bit more into those different strands of research and try to come up with a holistic research proposal for the entirety of our PhD research.

7. REFERENCES

- Allen, L. J. S. (2017) 'A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis', *Infectious Disease Modelling*. KeAi Communications Co., 2(2), pp. 128–142. doi: 10.1016/j.idm.2017.03.001.
- Balcan, D. *et al.* (2009) 'Multiscale mobility networks and the spatial spreading of infectious diseases', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 106(51), pp. 21484–21489. doi: 10.1073/pnas.0906910106.
- Barabási, A. L. and Albert, R. (1999) 'Emergence of scaling in random networks', *Science*. Science, 286(5439), pp. 509–512. doi: 10.1126/science.286.5439.509.
- Barberis, I. *et al.* (2016) 'History and evolution of influenza control through vaccination: From the first monovalent vaccine to universal vaccines', *Journal of Preventive Medicine and Hygiene*. Pacini Editore S.p.A., pp. E115–E120. doi: 10.15167/2421-4248/jpmh2016.57.3.642.
- Barbour, A. D. and Utev, S. (2004) 'Approximating the Reed-Frost epidemic process', *Stochastic Processes and their Applications*. North-Holland, 113(2), pp. 173–197. doi: 10.1016/j.spa.2004.03.013.
- Barrat, A. *et al.* (2004) 'The architecture of complex weighted networks', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 101(11), pp. 3747–3752. doi: 10.1073/pnas.0400087101.
- Barrat, A., Barthélemy, M. and Vespignani, A. (2004) 'Weighted evolving networks: Coupling topology and weight dynamics', *Physical Review Letters*. American Physical Society, 92(22), p. 228701. doi: 10.1103/PhysRevLett.92.228701.
- Barrett C *et al.* (2008) 'EpiSimdemics: An efficient and scalable framework for simulating the spread of infectious disease on large social networks', *International Conference for High Performance Computing, Networking, Storage and Analysis (SC08)*, (November). doi: 10.1145/1413370.1413408.
- Beutels, P. *et al.* (2006) 'Social mixing patterns for transmission models of close contact infections: Exploring self-evaluation and diary-based data collection through a web-based interface', *Epidemiology and Infection*. Cambridge University Press, 134(6), pp. 1158–1166. doi: 10.1017/S0950268806006418.
- Bi, Q. *et al.* (2020) 'Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study', *The Lancet Infectious Diseases*. Lancet Publishing Group, 20(8), pp. 911–919. doi: 10.1016/S1473-3099(20)30287-5.
- Bisset, K. R. *et al.* (2009) 'EpiFast: A fast algorithm for large scale realistic epidemic simulations on distributed memory systems', in *Proceedings of the International Conference on Supercomputing*, pp. 430–439. doi: 10.1145/1542275.1542336.
- Bonabeau, E. (2002) 'Agent-based modeling: Methods and techniques for simulating human systems', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 99(SUPPL. 3), pp. 7280–7287. doi: 10.1073/pnas.082080899.
- Britton, T. (2010) 'Stochastic epidemic models: A survey', *Mathematical Biosciences*. Math Biosci, 225(1), pp. 24–35. doi: 10.1016/j.mbs.2010.01.006.
- Brockmann, D., Hufnagel, L. and Geisel, T. (2006) 'The scaling laws of human travel', *Nature*. Nature Publishing Group, 439(7075), pp. 462–465. doi: 10.1038/nature04292.
- Campuzano, F. *et al.* (2011) 'Flexible simulation of ubiquitous computing environments', in *Advances in Intelligent and Soft Computing*. Springer, Berlin, Heidelberg, pp. 189–196. doi: 10.1007/978-3-642-19937-0_24.
- Chowell, G. *et al.* (2009) 'Adaptive Vaccination Strategies to Mitigate Pandemic Influenza: Mexico as a Case Study', *PLoS ONE*. Edited by H. E. Gendelman. Public Library of Science, 4(12), p. e8164. doi: 10.1371/journal.pone.0008164.
- Chu, D. K. *et al.* (2020) 'Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis', *The Lancet*. Lancet Publishing Group, 395(10242), pp.

1973–1987. doi: 10.1016/S0140-6736(20)31142-9.

- Chu, X. *et al.* (2009) 'Epidemic spreading in weighted scale-free networks with community structure', *Journal of Statistical Mechanics: Theory and Experiment*. IOP Publishing, 2009(7), pp. 7–43. doi: 10.1088/1742-5468/2009/07/P07043.
- Chung-Yuan Huang, C.-T. S. J.-L. H. and H. L. (2004) 'Simulating SARS: Small-World Epidemiological Modeling and Public Health Policy Assessments', *Journal of Artificial Societies and Social Simulation*. JASSS, 7.
- Cities, C. for (2020) *What does the COVID-19 crisis mean for the economies of British cities and large towns?* Available at: https://www.centreforcities.org/coronavirus/ (Accessed: 23 August 2020).
- Codling, E. A., Plank, M. J. and Benhamou, S. (2008) 'Random walk models in biology', *Journal of the Royal Society Interface*. Royal Society, pp. 813–834. doi: 10.1098/rsif.2008.0014.
- Colizza, V. *et al.* (2007) *EPIDEMIC PREDICTABILITY IN META-POPULATION MODELS WITH HETEROGENEOUS COUPLINGS: THE IMPACT OF DISEASE PARAMETER VALUES*, *International Journal of Bifurcation and Chaos*. Available at: http://www.iata.org. (Accessed: 23 August 2020).
- Conte, R. and Paolucci, M. (2014) 'On agent-based modeling and computational social science.', *Frontiers in psychology*. Frontiers Research Foundation, 5(JUL), p. 668. doi: 10.3389/fpsyg.2014.00668.
- *Coronavirus (COVID-19) - NHS* (2020). Available at: https://www.nhs.uk/conditions/coronavirus-covid-19/ (Accessed: 23 August 2020).
- *Coronavirus (COVID-19) infections in the community in England - Office for National Statistics* (no date). Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19infectionsinthecommunityinengland/july2020 (Accessed: 23 August 2020).
- *Coronavirus Recovery: Rate, Time, and Outlook* (2020). Available at: https://www.webmd.com/lung/covid-recovery-overview#1 (Accessed: 23 August 2020).
- Day, M. (2020) 'Covid-19: four fifths of cases are asymptomatic, China figures indicate', *BMJ (Clinical research ed.)*. NLM (Medline), 369, p. m1375. doi: 10.1136/bmj.m1375.
- Deijfen, M. (2011) 'Epidemics and vaccination on weighted graphs', *Mathematical Biosciences*, 232(1), pp. 57–65. doi: 10.1016/j.mbs.2011.04.003.
- Dimitrov, N. B. and Meyers, L. A. (no date) 'Mathematical Approaches to Infectious Disease Prediction and Control', *INFORMS*. doi: 10.1287/educ.1100.0075.
- Duan, W., Cao, Z., *et al.* (2013) 'An ACP approach to public health emergency management: Using a campus outbreak of h1n1 influenza as a case study', *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, 43(5), pp. 1028–1041. doi: 10.1109/TSMC.2013.2256855.
- Duan, W., Qiu, X., *et al.* (2013) 'Heterogeneous and stochastic agent-based models for analyzing infectious diseases' super spreaders', *IEEE Intelligent Systems*, 28(4), pp. 18–25. doi: 10.1109/MIS.2013.29.
- Duan, W. *et al.* (2015) 'Mathematical and computational approaches to epidemic modeling: a comprehensive review', *Front. Comput. Sci*, 9(5), pp. 806–826. doi: 10.1007/s11704-014-3369-2.
- Duan, W. and Qiu, X. (2012) 'Fostering artificial societies using social learning and social control in parallel emergency management systems', *Frontiers of Computer Science*, 6(5), pp. 604–610. doi: 10.1007/s11704-012-1166-3.
- Dunham, J. B. (2005) 'An Agent-Based Spatially Explicit Epidemiological Model in MASON', *Journal of Artificial Societies and Social Simulation*. JASSS, 9(1).
- Eames, K. T. D., Read, J. M. and Edmunds, W. J. (2009) 'Epidemic prediction and control in weighted networks', *Epidemics*. Epidemics, 1(1), pp. 70–76. doi: 10.1016/j.epidem.2008.12.001.
- Edmunds, W. J., O'Callaghan, C. J. and Nokes, D. J. (1997) 'Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections', *Proceedings of the Royal Society B: Biological Sciences*. Royal

Society, 264(1384), pp. 949–957. doi: 10.1098/rspb.1997.0131.

- Epstein, J. M. *et al.* (2008) 'Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations', *PLoS ONE*. Edited by A. P. Galvani. Public Library of Science, 3(12), p. e3955. doi: 10.1371/journal.pone.0003955.
- Esteve, A. *et al.* (2020) 'National age and coresidence patterns shape COVID-19 vulnerability', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 117(28), pp. 16118–16120. doi: 10.1073/pnas.2008764117.
- Eubank, S. *et al.* (2004) 'Modelling disease outbreaks in realistic urban social networks', *Nature*. Nature Publishing Group, 429(6988), pp. 180–184. doi: 10.1038/nature02541.
- Fenichela, E. P. *et al.* (2011) 'Adaptive human behavior in epidemiological models', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 108(15), pp. 6306–6311. doi: 10.1073/pnas.1011250108.
- Fournie, G. *et al.* (2013) 'Interventions for avian influenza A (H5N1) risk management in live bird market networks', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 110(22), pp. 9177–9182. doi: 10.1073/pnas.1220815110.
- Garnett, G. P. *et al.* (2011) 'Mathematical models in the evaluation of health programmes', *The Lancet*. Lancet, pp. 515–525. doi: 10.1016/S0140-6736(10)61505-X.
- Google Trends (2020) *Coronavirus Search Trends - Google Trends*. Available at: https://trends.google.com/trends/story/US_cu_4Rjdh3ABAABMHM_en (Accessed: 24 August 2020).
- Griffin, R. H. and Nunn, C. L. (2012) 'Community structure and the spread of infectious disease in primate social networks', *Evolutionary Ecology*. Springer, 26(4), pp. 779–800. doi: 10.1007/s10682-011-9526-2.
- Gross, T. and Blasius, B. (2008) 'Adaptive coevolutionary networks: a review', *Journal of The Royal Society Interface*. Royal Society, 5(20), pp. 259–271. doi: 10.1098/rsif.2007.1229.
- He, D. *et al.* (2020) 'The relative transmissibility of asymptomatic COVID-19 infections among close contacts', *International Journal of Infectious Diseases*. doi: 10.1016/j.ijid.2020.04.034.
- Hunter, E., Namee, B. Mac and Kelleher, J. (2017) 'A taxonomy for agent-based models in human infectious disease epidemiology', *Journal of Artificial Societies and Social Simulation*. University of Surrey, 20(3). doi: 10.18564/jasss.3414.
- Jewell, C. P. *et al.* (2009) 'Bayesian Analysis for Emerging Infectious Diseases', *Bayesian Analysis*, 4(4), pp. 465–496. doi: 10.1214/09-BA417.
- Kamp, C., Moslonka-Lefebvre, M. and Alizon, S. (2013) 'Epidemic Spread on Weighted Networks', *PLoS Computational Biology*. Public Library of Science, 9(12), p. 1003352. doi: 10.1371/journal.pcbi.1003352.
- Keeling, M. J. and Danon, L. (2009) 'Mathematical modelling of infectious diseases', *British Medical Bulletin*. Oxford Academic, 92(1), pp. 33–42. doi: 10.1093/bmb/ldp038.
- Korostil, I. A. *et al.* (2013) 'Adaptive Markov chain Monte Carlo forward projection for statistical analysis in epidemic modelling of human papillomavirus', *Statistics in Medicine*, 32(11), pp. 1917–1953. doi: 10.1002/sim.5590.
- Kucharski, A. J. *et al.* (2020) 'Early dynamics of transmission and control of COVID-19: a mathematical modelling study', *The Lancet Infectious Diseases*. Lancet Publishing Group, 20(5), pp. 553–558. doi: 10.1016/S1473-3099(20)30144-4.
- Kupferschmidt, K. (2020) 'Mutations can reveal how the coronavirus moves—but they're easy to overinterpret', *Science*. American Association for the Advancement of Science (AAAS). doi: 10.1126/science.abb6526.
- Latora, V. and Marchiori, M. (2003) 'Economic small-world behavior in weighted networks', *European Physical Journal B*. Springer, 32(2), pp. 249–263. doi: 10.1140/epjb/e2003-00095-5.
- Lee, S. *et al.* (2012) 'Exploiting Temporal Network Structures of Human Interaction

to Effectively Immunize Populations', *PLoS ONE*. Edited by Y. Moreno. Public Library of Science, 7(5), p. e36439. doi: 10.1371/journal.pone.0036439.

- Leskovec, J. *et al.* (2007) 'Cost-effective outbreak detection in networks', in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, New York, USA: ACM Press, pp. 420–429. doi: 10.1145/1281192.1281239.
- Li, M. *et al.* (2005) 'Weighted networks of scientific communication: The measurement and topological role of weight', *Physica A: Statistical Mechanics and its Applications*. North-Holland, 350(2–4), pp. 643–656. doi: 10.1016/j.physa.2004.11.039.
- London School of Hygiene & Tropical Medicine (2020) *Peak in COVID-19 deaths occurring in English hospitals passed on 8 April | LSHTM*. Available at: https://www.lshtm.ac.uk/newsevents/news/2020/peak-covid-19-deaths-occurring-english-hospitals-passed-8-april (Accessed: 23 August 2020).
- Löscher, T. and Prüfer-Krämer, L. (2010) 'Emerging Emerging infectious diseases and Re-emerging Infectious Diseases', in Krämer, A., Kretzschmar, M., and Krickeberg, K. (eds) *Modern Infectious Disease Epidemiology SE - 3*. Springer New York (Statistics for Biology and Health), pp. 39–67. doi: 10.1007/978-0-387-93835-6_3.
- Madar, N. *et al.* (2004) 'Immunization and epidemic dynamics in complex networks', in *European Physical Journal B*, pp. 269–276. doi: 10.1140/epjb/e2004-00119-8.
- Marceau, V. *et al.* (2010) 'Adaptive networks: Coevolution of disease and topology', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. American Physical Society, 82(3), p. 036116. doi: 10.1103/PhysRevE.82.036116.
- Masuda, N., Klemm, K. and Eguíluz, V. M. (2013) 'Temporal networks: Slowing down diffusion by long lasting interactions', *Physical Review Letters*. American Physical Society, 111(18), p. 188701. doi: 10.1103/PhysRevLett.111.188701.
- Mei, S. *et al.* (2010) 'Complex agent networks explaining the HIV epidemic among homosexual men in Amsterdam', *Mathematics and Computers in Simulation*. North-Holland, 80(5), pp. 1018–1030. doi: 10.1016/j.matcom.2009.12.008.
- Merler, S. and Ajelli, M. (2010) 'The role of population heterogeneity and human mobility in the spread of pandemic influenza', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 277(1681), pp. 557–565. doi: 10.1098/rspb.2009.1605.
- Min, B., Goh, K. I. and Vazquez, A. (2011) 'Spreading dynamics following bursty human activity patterns', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. American Physical Society, 83(3), p. 036102. doi: 10.1103/PhysRevE.83.036102.
- Mitchell, L. and Ross, J. V. (2016) 'A data-driven model for influenza transmission incorporating media effects', *Royal Society Open Science*. Royal Society Publishing, 3(10). doi: 10.1098/rsos.160481.
- Mossong, J. *et al.* (2008) 'Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases', *PLoS Medicine*. Edited by S. Riley. Public Library of Science, 5(3), p. e74. doi: 10.1371/journal.pmed.0050074.
- New York Times (2020) *Mapping the Social Network of Coronavirus - The New York Times*. Available at: https://www.nytimes.com/2020/03/13/science/coronavirus-social-networks-data.html (Accessed: 25 August 2020).
- Nsoesie, E. O. *et al.* (2020) *Analysis of hospital traffic and search engine data in Wuhan China indicates early disease activity in the Fall of 2019*. Available at: http://nrs.harvard.edu/urn-3:HUL.InstRepos:42669767 (Accessed: 23 August 2020).
- O'Neill, P. D. (2002) 'A tutorial introduction to Bayesian inference for stochastic epidemic models using Markov chain Monte Carlo methods', *Mathematical Biosciences*. Elsevier Inc., 180(1–2), pp. 103–114. doi: 10.1016/S0025-5564(02)00109-8.
- Office for National Statistics (2020a) *Coronavirus (COVID-19) roundup: Deaths and health - Office for National Statistics*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19roundupdeathsandhealth/2020-06-26

(Accessed: 23 August 2020).

- Office for National Statistics (2020b) *Deaths involving COVID-19, England and Wales - Office for National Statistics*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19englandandwales/deathsoccurringinapril2020 (Accessed: 23 August 2020).
- Office for National Statistics (2020c) *Deaths involving COVID-19 by local area and socioeconomic deprivation - Office for National Statistics*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19bylocalareasanddeprivation/latest (Accessed: 23 August 2020).
- Opsahl, T. (2013) 'Triadic closure in two-mode networks: Redefining the global and local clustering coefficients', *Social Networks*, 35(2), pp. 159–167. doi: 10.1016/j.socnet.2011.07.001.
- Palla, G. *et al.* (2005) 'Uncovering the overlapping community structure of complex networks in nature and society', *Nature*. Nature Publishing Group, 435(7043), pp. 814–818. doi: 10.1038/nature03607.
- Pastor-Satorras, R. and Vespignani, A. (2001) 'Epidemic spreading in scale-free networks', *Physical Review Letters*. American Physical Society, 86(14), pp. 3200–3203. doi: 10.1103/PhysRevLett.86.3200.
- Perez, L. and Dragicevic, S. (2009) 'An agent-based approach for modeling dynamics of contagious disease spread', *International Journal of Health Geographics*. BioMed Central, 8(1), p. 50. doi: 10.1186/1476-072X-8-50.
- le Polain de Waroux, O. *et al.* (2018) 'Characteristics of human encounters and social mixing patterns relevant to infectious diseases spread by close contact: A survey in Southwest Uganda', *BMC Infectious Diseases*. BioMed Central Ltd., 18(1), p. 172. doi: 10.1186/s12879-018-3073-1.
- Rahmandad, H. and Sterman, J. (2008) 'Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models', *Management Science*. INFORMS , 54(5), pp. 998–1014. doi: 10.1287/mnsc.1070.0787.
- Rorres, C., Pelletier, S. T. K. and Smith, G. (2011) 'Stochastic modeling of animal epidemics using data collected over three different spatial scales', *Epidemics*. NIH Public Access, 3(2), pp. 61–70. doi: 10.1016/j.epidem.2011.02.003.
- Schwartz, I. and Shaw, L. (2010) 'Rewiring for adaptation', *Physics*. American Physical Society (APS), 3(17). doi: 10.1103/physics.3.17.
- van Segbroeck, S., Santos, F. C. and Pacheco, J. M. (2010) 'Adaptive contact networks change effective disease infectiousness and dynamics', *PLoS Computational Biology*. Public Library of Science, 6(8), p. 1000895. doi: 10.1371/journal.pcbi.1000895.
- Shaw, L. B. and Schwartz, I. B. (2008) 'Fluctuating epidemics on adaptive networks', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. American Physical Society, 77(6), p. 066101. doi: 10.1103/PhysRevE.77.066101.
- Shaw, L. B. and Schwartz, I. B. (2010) 'Enhanced vaccine control of epidemics in adaptive networks', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. NIH Public Access, 81(4), p. 046120. doi: 10.1103/PhysRevE.81.046120.
- Siettos, C. I. and Russo, L. (2013) 'Mathematical modeling of infectious disease dynamics', *Virulence*. Taylor and Francis Inc., 4(4), pp. 295–306. doi: 10.4161/viru.24041.
- Song, Y. R., Jiang, G. P. and Xu, J. G. (2011) 'An epidemic spreading model in adaptive networks based on cellular automata', *Acta Physica Sinica*. 物理学报, 60(12), pp. 120509–120509. doi: 10.7498/aps.60.120509.
- Spreitzer, M. and Theimer, M. (2007) 'Providing Location Information in a Ubiquitous Computing Environment', in *Mobile Computing*. Springer US, pp. 397–423. doi: 10.1007/978-0-585-29603-6_15.
- Urbistat (2020) *Maps, analysis and statistics about the resident population in UK*. Available at:

https://ugeo.urbistat.com/AdminStat/en/uk/demografia/popolazione/blackburn-with-darwen/46/4 (Accessed: 23 August 2020).

- Wang, B. *et al.* (2011) 'Epidemic spread in adaptive networks with multitype agents', *Journal of Physics A: Mathematical and Theoretical*. IOP Publishing, 44(3), p. 35101. doi: 10.1088/1751-8113/44/3/035101.
- Wang, L. *et al.* (2013) 'How human location-specific contact patterns impact spatial transmission between populations?', *Scientific Reports*. Nature Publishing Group, 3(1), pp. 1–10. doi: 10.1038/srep01468.
- Wang, Y. *et al.* (2011) 'The impact of community structure of social contact network on epidemic outbreak and effectiveness of non-pharmaceutical interventions', in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Nature Publishing Group, pp. 108–120. doi: 10.1007/978-3-642-22039-5_12.
- Watts, D. J. and Strogatz, S. H. (1998) 'Collective dynamics of 'small-world9 networks', *Nature*. Macmillan Magazines Ltd, 393(6684), pp. 440–442. doi: 10.1038/30918.
- World Health Organization (2020a) *Archived: WHO Timeline - COVID-19*. Available at: https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19 (Accessed: 25 August 2020).
- World Health Organization (2020b) *Situation Report-73 HIGHLIGHTS*. doi: 10.3201/eid2606.200239.
- Yang, H., Tang, M. and Zhang, H. F. (2012) 'Efficient community-based control strategies in adaptive networks', *New Journal of Physics*, 14. doi: 10.1088/1367-2630/14/12/123017.
- YouGov (2018) *The most popular newspapers in the UK | Media | YouGov Ratings*. Available at: https://yougov.co.uk/ratings/media/popularity/newspaper/all (Accessed: 24 August 2020).
- Zanette, D. H. and Risau-Gusmán, S. (2008) 'Infection spreading in a population with evolving contacts', *Journal of Biological Physics*. Springer, 34(1-2 SPEC. ISS.), pp. 135–148. doi: 10.1007/s10867-008-9060-9.
- Zeng, D. *et al.* (2009) 'Disease surveillance based on spatial contact networks: A case study of Beijing 2003 SARS epidemic', *IEEE Intelligent Systems*. Institute of Electrical and Electronics Engineers Inc., 24(6), pp. 77–82. Available at: https://arizona.pure.elsevier.com/en/publications/disease-surveillance-based-on-spatial-contact-networks-a-case-stu (Accessed: 23 August 2020).
- Zhang, H. *et al.* (2012) 'Modeling the influence of information on the coevolution of contact networks and the dynamics of infectious diseases', *Physica D: Nonlinear Phenomena*. Elsevier B.V., 241(18), pp. 1512–1517. doi: 10.1016/j.physd.2012.05.011.
- Zhang, J. *et al.* (2005) 'A compartmental model for the analysis of SARS transmission patterns and outbreak control measures in China', *Applied Mathematics and Computation*. Elsevier, 162(2), pp. 909–924. doi: 10.1016/j.amc.2003.12.131.
- Zhou, T. *et al.* (2006) 'Behaviors of susceptible-infected epidemics on scale-free networks with identical infectivity', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. American Physical Society, 74(5). doi: 10.1103/PhysRevE.74.056109.
- Zhu, G. *et al.* (2013) 'Epidemic spreading on contact networks with adaptive weights', *Journal of Theoretical Biology*. Academic Press, 317, pp. 133–139. doi: 10.1016/j.jtbi.2012.09.036.

Appendix 1

City	Symptomatic Age Group	Asymptomatic Age Group	Children Infected	Adult Infected	Elderly Infected	Total Number of People Infected
Leicester	0 - 17	0 - 17	3675.37	227609.8	39813.56	351098.7
		18 - 64	83675.34	227609.6	39813.35	351098.3
		65+	83675.32	227609.5	39813.15	351098
	18 - 64	0 - 17	83675.39	227609.7	39813.56	351098.7
		18 - 64	83675.36	227609.6	39813.35	351098.3
		65+	83675.33	227609.5	39813.15	351098
	65+	0 - 17	83675.39	227609.8	39813.46	351098.6
		18 - 64	83675.36	227609.6	39813.24	351098.3
		65+	83675.33	227609.5	39813.05	351097.9
Bradford	0 - 17	0 - 17	141490	314454.6	6677.57	532622.2
		18 - 64	141490	314454.6	76677.56	532622.2
		65+	141490	314454.4	76677.02	532621.3
	18 - 64	0 - 17	141490	314454.6	76677.57	532622.2
		18 - 64	141490	314454.6	76677.56	532622.1
		65+	141490	314454.3	76677.01	532621.3
	65+	0 - 17	141490	314454.6	76677.51	532622.1
		18 - 64	141490	314454.6	76677.5	532622.1
		65+	141490	314454.4	76676.96	532621.3
Blackburn	0 - 17	0 - 17	38591.62	88618.04	20524.39	147734.1
		18 - 64	38591.62	88618.04	20524.37	147734
		65+	38591.6	88617.96	20524.2	147733.8
	18 - 64	0 - 17	38591.64	88618.03	20524.38	147734
		18 - 64	38591.64	88618.02	20524.37	147734
		65+	38591.61	88617.95	20524.19	147733.7
	65+	0 - 17	38591.64	88618.05	20524.3	147734
		18 - 64	38591.64	88618.04	20524.29	147734
		65+	38591.61	88617.97	20524.11	147733.7

Table 14 Complete Results of Mathematical Model Simulation Runs for the 3 cities

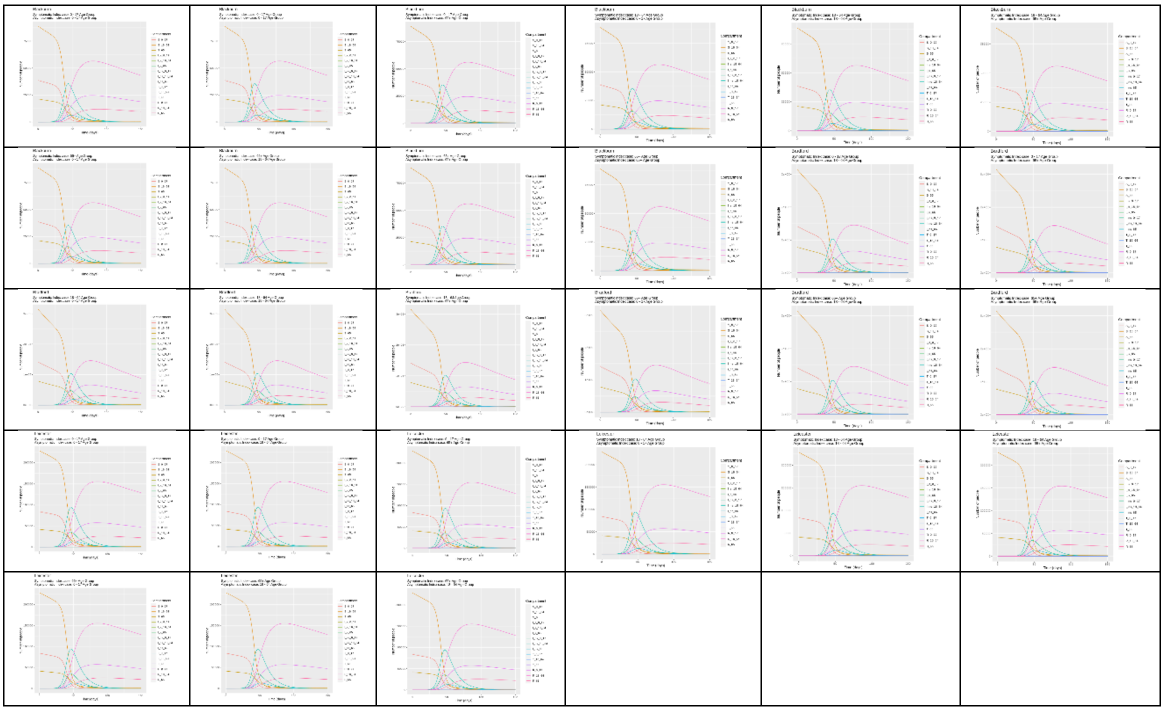


Figure 28 Graphs from Mathematical Simulation runs for all 3 cities with every combination of initial infected cases

	date	count	Normalize
0	15/03/2020	32	0.088889
1	16/03/2020	36	0.111111
2	17/03/2020	32	0.088889
3	18/03/2020	44	0.155556
4	19/03/2020	36	0.111111
5	20/03/2020	37	0.116667
6	21/03/2020	44	0.155556
7	22/03/2020	28	0.066667
8	23/03/2020	48	0.177778
9	24/03/2020	47	0.172222
10	25/03/2020	45	0.161111
11	26/03/2020	39	0.127778
12	27/03/2020	35	0.105556
13	28/03/2020	42	0.144444
14	29/03/2020	35	0.105556
15	30/03/2020	60	0.244444
16	31/03/2020	35	0.105556
17	01/04/2020	66	0.277778
18	02/04/2020	43	0.15
19	03/04/2020	41	0.138889
20	04/04/2020	45	0.161111
21	05/04/2020	28	0.066667
22	06/04/2020	54	0.211111
23	07/04/2020	31	0.083333
24	08/04/2020	58	0.233333
25	09/04/2020	40	0.133333
26	10/04/2020	72	0.311111
27	11/04/2020	55	0.216667
28	12/04/2020	51	0.194444
29	13/04/2020	70	0.3
30	14/04/2020	62	0.255556
31	15/04/2020	51	0.194444
32	16/04/2020	42	0.144444
33	17/04/2020	41	0.138889
34	18/04/2020	66	0.277778
35	19/04/2020	58	0.233333
36	20/04/2020	81	0.361111
37	21/04/2020	42	0.144444
38	22/04/2020	46	0.166667
39	23/04/2020	49	0.183333
40	24/04/2020	47	0.172222
41	25/04/2020	37	0.116667
42	26/04/2020	63	0.261111
43	27/04/2020	66	0.277778
44	28/04/2020	16	0
45	29/04/2020	53	0.205556
46	30/04/2020	50	0.188889
47	01/05/2020	50	0.188889
48	02/05/2020	71	0.305556
49	03/05/2020	54	0.211111
50	04/05/2020	97	0.45
51	05/05/2020	42	0.144444
52	06/05/2020	64	0.266667
53	07/05/2020	68	0.288889
54	08/05/2020	71	0.305556
55	09/05/2020	49	0.183333
56	10/05/2020	35	0.105556
57	11/05/2020	90	0.411111
58	12/05/2020	51	0.194444
59	13/05/2020	46	0.166667
60	14/05/2020	58	0.233333
61	15/05/2020	64	0.266667
62	16/05/2020	56	0.222222
63	17/05/2020	36	0.111111
64	18/05/2020	102	0.477778
65	19/05/2020	56	0.222222
66	20/05/2020	74	0.322222
67	21/05/2020	54	0.211111
68	22/05/2020	47	0.172222
69	23/05/2020	34	0.1
70	24/05/2020	54	0.211111
71	25/05/2020	109	0.516667
72	26/05/2020	39	0.127778
73	27/05/2020	54	0.211111
74	28/05/2020	54	0.211111
75	29/05/2020	88	0.4
76	30/05/2020	114	0.544444
77	31/05/2020	96	0.444444
78	01/06/2020	196	1
79	02/06/2020	78	0.344444
80	03/06/2020	122	0.588889
81	04/06/2020	112	0.533333
82	05/06/2020	97	0.45
83	06/06/2020	98	0.455556
84	07/06/2020	88	0.4
85	08/06/2020	140	0.688889
86	09/06/2020	39	0.127778
87	10/06/2020	50	0.188889
88	11/06/2020	54	0.211111
89	12/06/2020	52	0.2
90	13/06/2020	40	0.133333
91	14/06/2020	37	0.116667
92	15/06/2020	63	0.261111

Table 15 Data mining results containing the number of Covid-19 related articles from March to June from Times online news portal. The last column contains the normalized count for finding the equation of mitigation rate

City	Population	Age Band	Count	Covid Mortality Rate	Non-Covid Mortality Rate	Total Cumulative Cases
Leicester	353,540	Children	83,778	0.006105	0.0038	4465
		Adult	228,226			
		Elderly	41,536			
Bradford	534,800	Children	141,606	0.006086	0.005	5685
		Adult	314,875			
		Elderly	78,319			
Blackburn	148,772	Children	38,657	0.006107	0.004772	1680
		Adult	88,824			
		Elderly	21,291			

Table 16 Snapshot of the different mortality rate and cumulative cases in each of the three cities

Appendix 2

7.1. Code for Mathematical Model

```
######################################################################
###############
library(deSolve)
library(reshape2)
library(ggplot2)

require(deSolve)
require(reshape2)
require(ggplot2)

# INPUT

# contact matrix
contact_matrix_asym <- matrix(0,nrow=3,ncol=3)
# Fill in the contract matrix
contact_matrix_asym[1,1] = 7
contact_matrix_asym[1,2] = 5
contact_matrix_asym[1,3] = 1
contact_matrix_asym[2,1] = 2
contact_matrix_asym[2,2] = 9
contact_matrix_asym[2,3] = 1
contact_matrix_asym[3,1] = 1
contact_matrix_asym[3,2] = 3
contact_matrix_asym[3,3] = 2



contact_matrix_sym <- matrix(0,nrow=3,ncol=3)
# Fill in the contract matrix
```

```
contact_matrix_sym[1,1] = 1
contact_matrix_sym[1,2] = 1
contact_matrix_sym[1,3] = 0
contact_matrix_sym[2,1] = 1
contact_matrix_sym[2,2] = 1
contact_matrix_sym[2,3] = 0
contact_matrix_sym[3,1] = 0
contact_matrix_sym[3,2] = 0
contact_matrix_sym[3,3] = 1


######################################################################
############################
# Parameters for Leicester
parameters_leicester <- c(rho_symtomatic = 0.063,
                rho_asymptomatic = 0.041,
                contact_matrix = contact_matrix,
                gamma = 1/14,
                mu_cov = 0.006105,
                mu_normal = 0.003800,
                p_symptomatic = 0.22,
                p_asymptomatic = 0.78,
                h = 1/5)
######################################################################
#############################

# Run simulation for 5 months
times <- seq(from = 0, to = 150, by = 1)

# MODEL FUNCTION
sir_model_age_separated <- function(time, state, parameters) {

  with(as.list(parameters), {

    n_agegroups <- 3                  # number of age groups
    S <- state[1:n_agegroups]
    I_s <- state[(n_agegroups+1):(2*n_agegroups)]
    I_ns <- state[(2*n_agegroups+1):(3*n_agegroups)]
    T <- state[(3*n_agegroups+1):(4*n_agegroups)]
    R <- state[(4*n_agegroups+1):(5*n_agegroups)]

    N <- S + I_s + I_ns + T + R

lambda_symptomatic <- rho_symtomatic * contact_matrix_sym %*%
as.matrix(I_s/N)
lambda_asymptomatic <- rho_asymptomatic * contact_matrix_asym %*%
as.matrix(I_ns/N)

    # The differential equations

    dS <- -lambda_symptomatic * p_symptomatic * S -
lambda_asymptomatic * p_asymptomatic * S - mu_normal * S
    dI_s <- lambda_symptomatic * p_symptomatic * S - h * I_s - gamma *
I_s - mu_cov * I_s
    dI_ns <- lambda_asymptomatic * p_asymptomatic * S - gamma * I_ns -
mu_normal * I_ns
```

```
    dT <- h * I_s - gamma * gamma * T
    dR <- gamma * T + gamma * I_ns + gamma * I_s - mu_normal * R

    # Output
    return(list(c(dS, dI_s, dI_ns, dT, dR)))
  })
}

######################################################################
#############################
## Leicester
######################################################################
################
######################################################################
#############################

N <- 353540
N_0_17 <- 0.2371*N
N_18_64 <- 0.6456*N
N_65 <- 0.1175*N


initial_state_values <- c(S_0_17 = N_0_17,
                          S_18_64 = N_18_64,
                          S_65 = N_65,

                          I_s_0_17 = 0,
                          I_s_18_64 = 0,
                          I_s_65 = 1,

                          I_ns_0_17 = 1,
                          I_ns_18_64 = 0,
                          I_ns_65 = 0,

                          T_0_17 = 0,
                          T_18_64 = 0,
                          T_65 = 0,

                          R_0_17 = 0,
                          R_18_64 = 0,
                          R_65 = 0)

# Run model output
output_leicester <- as.data.frame(ode(y = initial_state_values,
                            times = times,
                            func = sir_model_age_separated,
                            parms = parameters_leicester))

# Graph of incidence

output_leicester_long <- melt(as.data.frame(output_leicester), id =
"time")

# Plot number of people in all compartments over time
ggplot(data = output_leicester_long,
```

```
       aes(x = time, y = value, colour = variable, group = variable))
+
  geom_line() +
  xlab("Time (days)")+
  ylab("Number of people") +
  ggtitle("Leicester", subtitle = "Symptomatic Index case: 65+ Age
Group \nAsymptomatic Index case: 0 - 17 Age Group") +
  labs(colour = "Compartment")


# Calculate cumulative incidence in each age group:
results_leicester <- data.frame(child_cum_inc =
output_leicester$S_0_17[1]-
output_leicester$S_0_17[nrow(output_leicester)],
                               adult_cum_inc =
output_leicester$S_18_64[1]-
output_leicester$S_18_64[nrow(output_leicester)],
                               elderly_cum_inc =
output_leicester$S_65[1]-
output_leicester$S_65[nrow(output_leicester)],
                               total_cum_inc =
sum(output_leicester[1,c("S_0_17", "S_18_64", "S_65")])-
sum(output_leicester[nrow(output_leicester),c("S_0_17", "S_18_64",
"S_65")]))

print(results_leicester)

######################################################################
#############################
## Bradford
######################################################################
#################
######################################################################
#############################

######################################################################
#############################
# Parameters for Bradford
parameters_bradford <- c(rho_symtomatic = 0.063,
                          rho_asymptomatic = 0.041,
                          contact_matrix = contact_matrix,
                          gamma = 1/14,
                          mu_cov = 0.006086,
                          mu_normal = 0.00500,
                          p_symptomatic = 0.22,
                          p_asymptomatic = 0.78,
                          h = 1/5)
######################################################################
#############################

N <- 534800
N_0_17 <- 0.2648*N
N_18_64 <- 0.5888*N
N_65 <- 0.1464*N
```

```
initial_state_values <- c(S_0_17 = N_0_17,
                          S_18_64 = N_18_64,
                          S_65 = N_65,

                          I_s_0_17 = 0,
                          I_s_18_64 = 0,
                          I_s_65 = 1,

                          I_ns_0_17 = 0,
                          I_ns_18_64 = 1,
                          I_ns_65 = 0,

                          T_0_17 = 0,
                          T_18_64 = 0,
                          T_65 = 0,

                          R_0_17 = 0,
                          R_18_64 = 0,
                          R_65 = 0)

# Run model output
output_bradford <- as.data.frame(ode(y = initial_state_values,
                                      times = times,
                                      func = sir_model_age_separated,
                                      parms = parameters_bradford))

# Graph of incidence

output_bradford_long <- melt(as.data.frame(output_bradford), id =
"time")

# Plot number of people in all compartments over time
ggplot(data = output_bradford_long,
       aes(x = time, y = value, colour = variable, group = variable))
+
  geom_line() +
  xlab("Time (days)")+
  ylab("Number of people") +
  ggtitle("Bradford", subtitle = "Symptomatic Index case: 65+ Age
Group \nAsymptomatic Index case: 18 - 64 Age Group") +
  labs(colour = "Compartment")


# Calculate cumulative incidence in each age group:
results_bradford <- data.frame(child_cum_inc =
output_bradford$S_0_17[1]-
output_bradford$S_0_17[nrow(output_bradford)],
                               adult_cum_inc =
output_bradford$S_18_64[1]-
output_bradford$S_18_64[nrow(output_bradford)],
                               elderly_cum_inc =
output_bradford$S_65[1]-output_bradford$S_65[nrow(output_bradford)],
                               total_cum_inc =
sum(output_bradford[1,c("S_0_17", "S_18_64", "S_65")])-
sum(output_bradford[nrow(output_bradford),c("S_0_17", "S_18_64",
"S_65")]))
```

```
print(results_bradford)

#######################################################################
#############################
## Blackburn
#######################################################################
################
#######################################################################
#############################

#######################################################################
#############################
# Parameters for Blackburn
parameters_blackburn <- c(rho_symtomatic = 0.063,
                         rho_asymptomatic = 0.041,
                         contact_matrix = contact_matrix,
                         gamma = 1/14,
                         mu_cov = 0.006107,
                         mu_normal = 0.004772,
                         p_symptomatic = 0.22,
                         p_asymptomatic = 0.78,
                         h = 1/5)

#######################################################################
#############################

N <- 148772
N_0_17 <- 0.2598*N
N_18_64 <- 0.5971*N
N_65 <- 0.1431*N


initial_state_values <- c(S_0_17 = N_0_17,
                          S_18_64 = N_18_64,
                          S_65 = N_65,

                          I_s_0_17 = 0,
                          I_s_18_64 = 0,
                          I_s_65 = 1,

                          I_ns_0_17 = 0,
                          I_ns_18_64 = 0,
                          I_ns_65 = 1,

                          T_0_17 = 0,
                          T_18_64 = 0,
                          T_65 = 0,

                          R_0_17 = 0,
                          R_18_64 = 0,
                          R_65 = 0)

# Run model output
output_blackburn <- as.data.frame(ode(y = initial_state_values,
                                     times = times,
                                     func = sir_model_age_separated,
```

```
                                           parms = parameters_blackburn))

# Graph of incidence

output_blackburn_long <- melt(as.data.frame(output_blackburn), id =
"time")

# Plot number of people in all compartments over time
ggplot(data = output_blackburn_long,
       aes(x = time, y = value, colour = variable, group = variable))
+
  geom_line() +
  xlab("Time (days)")+
  ylab("Number of people") +
  ggtitle("Blackburn", subtitle = "Symptomatic Index case: 65+ Age
Group \nAsymptomatic Index case: 65+ Age Group") +
  labs(colour = "Compartment")



# Calculate cumulative incidence in each age group:
results_blackburn <- data.frame(child_cum_inc =
output_blackburn$S_0_17[1]-
output_blackburn$S_0_17[nrow(output_blackburn)],
                               adult_cum_inc =
output_blackburn$S_18_64[1]-
output_blackburn$S_18_64[nrow(output_blackburn)],
                               elderly_cum_inc =
output_blackburn$S_65[1]-
output_blackburn$S_65[nrow(output_blackburn)],
                               total_cum_inc =
sum(output_blackburn[1,c("S_0_17", "S_18_64", "S_65")])-
sum(output_blackburn[nrow(output_blackburn),c("S_0_17", "S_18_64",
"S_65")]))
print(results_blackburn)
```

7.2. Code for Data Mining

```
from selenium import webdriver

import bs4 as bs

import time

import re

import requests

from selenium.webdriver.support.ui import WebDriverWait, Select

from selenium.webdriver.support import ui

from selenium.webdriver.support import expected_conditions as EC
```

```
from selenium.common.exceptions import TimeoutException,
NoSuchElementException, WebDriverException

from selenium.webdriver.common.keys import Keys

from selenium.webdriver.common.action_chains import ActionChains

import csv

from selenium.webdriver.common.by import By

import os

duration = 2  # seconds

freq = 440


#The urls holding the articles are defined here

urls = ["https://www.thetimes.co.uk/html-sitemap/2020-03-
3","https://www.thetimes.co.uk/html-sitemap/2020-03-
4","https://www.thetimes.co.uk/html-sitemap/2020-04-
1","https://www.thetimes.co.uk/html-sitemap/2020-04-
2","https://www.thetimes.co.uk/html-sitemap/2020-04-
3","https://www.thetimes.co.uk/html-sitemap/2020-04-
4","https://www.thetimes.co.uk/html-sitemap/2020-05-
1","https://www.thetimes.co.uk/html-sitemap/2020-05-
2","https://www.thetimes.co.uk/html-sitemap/2020-05-
3","https://www.thetimes.co.uk/html-sitemap/2020-05-
4","https://www.thetimes.co.uk/html-sitemap/2020-06-
1","https://www.thetimes.co.uk/html-sitemap/2020-06-2"]


#for each url the scraper is run

for url in urls:

    #words to be searched in each article title is defined here

    word_bank =
["coronavirus","Coronavirus","\"coronavirus","\"Coronavirus","Covid-
19","covid-19","social distance","Social
Distance","lockdown","Lockdown","wear mask","Wear Mask","stay
indoors","Stay
Indoors","quarantine","Quarantine","pandemic","Pandemic","social
distancing","Social Distancing"]

    #browser is opened in the background using a webdriver

    options = webdriver.FirefoxOptions()
```

```
    options.add_argument('-headless')

    driver = webdriver.Firefox(options=options)

    #the driver fetches the url

    driver.get(url)

    time.sleep(3)


    #the driver presses the consent button

    try:

driver.switch_to.frame(driver.find_element_by_id("sp_message_iframe_21
6133"))

        driver.find_element_by_xpath('//*[@title="I Agree"]').click()

        driver.switch_to.default_content()

    except (NoSuchElementException,TimeoutException):


        print("cant click in i agree")

        pass


    #fetches the elements containing the article titile

    tags = driver.find_elements_by_css_selector('li.Sitemap-link')


    with open('times4_1.csv', 'w', newline='') as file:

        writer = csv.writer(file)

        writer.writerow(["Date", "Occurence"])


    #the tags variable is iterated to find the article titles that
have words matching the words from the word_bank

    i=0

    for item in tags:

        cant_click=0
```

```
        for word in word_bank:
            if re.search(word,item.text):
                    #when the word is found the driver opens another
windows and opens the article
                if i>0:
                    print("found",word,"in",item.text)
                    driver2 = webdriver.Firefox(options=options)
                    try:

driver2.get(item.find_element_by_tag_name('a').get_attribute('href'))
                    except WebDriverException:
                        pass
                    time.sleep(3)


                    #click on the consent page
                    try:
                        iframe = WebDriverWait(driver2,
10).until(EC.presence_of_element_located((By.ID,
"sp_message_iframe_216133")))
                        driver2.switch_to.frame(iframe)
                        driver2.switch_to.default_content()
                    except (NoSuchElementException,TimeoutException)
                        cant_click=1


                    #method 1 to extract the datetime info
                    try:
                        date =
driver2.find_element_by_class_name('responsiveweb__DatePublicationCont
ainer-tglil3-1')
                        date_cleaned =
date.get_attribute('innerHTML').split('>')[1].split(',')[0]
```

```
                    # method 2 to extract the info

                    except (NoSuchElementException):

                        date =
driver2.find_elements_by_class_name('css-901oao')

                        for j in date:



                            if
re.search("<time",j.get_attribute('innerHTML')):

                                date=j

                                date_cleaned =
date.get_attribute('innerHTML').split('>')[1].split(',')[0]



                    #if the page is unresponsive and exception is
caught a "cant click" entry is made into the csv

                    if cant_click==1 and date_cleaned==None:

                        with open('times4_1.csv', 'a') as file:

                            writer = csv.writer(file)

                            writer.writerow(["cant click","1"])

                    else:

                        with open('times4_1.csv', 'a') as file:

                            writer = csv.writer(file)

                            writer.writerow([date_cleaned,"1"])

                    #when the date is extracted from the article the
browser window is closed

                    driver2.close()



                    break

                i+=1

    #when all the articles are parsed on that page the window is
closed and the next url is fetched

    driver.close()
```

```
os.system('play -nq -t alsa synth {} sine {}'.format(duration, freq))
```

7.3. Code for Agent-based Model

```
globals [
  population
  masked_agent
  new_mask
  mask_rate
]

turtles-own [
  infected?
  days_infected
  quarantined?
  cured?
  mask_on?
  contactable_child
  contactable_adult
  contactable_old
  contacted_child
  contacted_adult
  contacted_old
  disable_contact?
]

breed [children child]
breed [adults adult]
breed [olds old]

; function for setting up new simulation
to setup
  ; resetting previous states
  clear-all
  reset-ticks

  ; creating agents from each sub group
  create-child-agents
  create-adult-agents
  create-old-agents

  ; counting total population for later use
  set population count turtles

  ask turtles [setxy random-xcor random-ycor]   ; setting random
position

  ; setting attributes
  ask turtles [set color gray]
  ask turtles [set infected? false]
  ask turtles [set quarantined? false]
  ask turtles [set mask_on? false]
  ask turtles [set cured? false]
  ask turtles [set disable_contact? false]
```

```
  ; infect one random agent
  (ifelse
    first_infected = "child"
    [
      ask one-of children [
        set infected? true
        set color red
      ]
    ]

    first_infected = "adult"
    [
      ask one-of adults [
        set infected? true
        set color red
      ]
    ]
    first_infected = "old"
    [
      ask one-of olds [
        set infected? true
        set color red
      ]
    ]
   )
end

; function for running one day cycle
to one-day
  ; stop simulation after 90 days
  if ticks = 90 [stop]

  ; increase agents with mitigation
  if mitigation
  [
    set masked_agent count turtles with [mask_on?]                ;
currently masked agents
    set mask_rate 0.0026 * ticks + 0.1697                         ;
mask_rate = 0.0026 * time + 0.1697
    set new_mask ((population * mask_rate) - masked_agent)        ;
new_mask = (currently expected masked agents) - (already masked
agents)

    ask n-of new_mask turtles with [not mask_on?]
    [
      set mask_on? true
      set shape (ifelse-value breed = children ["child masked"] breed
= adults ["adult masked"] ["old masked"])
    ]
  ]

  ; reset dailty contact counter
  ask turtles [
    set contacted_child 0
    set contacted_adult 0
```

```
    set contacted_old 0
  ]
  ask turtles with[not quarantined?] [set disable_contact? false]

  ; count infected days for infected agents
  ask turtles with [infected?] [set days_infected days_infected + 1]

  ; send to quarantine after 5 days of infection
  ask turtles with [days_infected = 6] [
    set quarantined? true
    set disable_contact? true
    set color yellow
  ]

  ; recover after 14 days of quarantine
  ask turtles with [days_infected = 6 + 14] [
    set infected? false
    set quarantined? false
    set cured? true
    set disable_contact? false
    set color green
    set days_infected 0
  ]

  ; simulate 12 steps
  repeat 12 [
    ; take a walk
    ask turtles with [not quarantined?] [walk]

    ; Infect others
    ask turtles with [infected? and not quarantined?] [
      if contacted_child < contactable_child [infect-children]
      if contacted_adult < contactable_adult [infect-adults]
      if contacted_old < contactable_old [infect-olds]
    ]

    ; Count number of people that came in contact in this step
    ask turtles with [not quarantined?] [
      set contacted_child contacted_child + count other children in-
radius 1.5 with [not quarantined? and contacted_child <
contactable_child]
      set contacted_adult contacted_adult + count other adults in-
radius 1.5 with [not quarantined? and contacted_adult <
contactable_adult]
      set contacted_old contacted_old + count other olds in-radius 1.5
with [not quarantined? and contacted_old < contactable_old]
    ]
  ]

  ; day counter
  tick
end

; helper functions

to walk
```

```
  ifelse (random 2 = 0) [right random 45] [left random 45]
  forward 1
end

to create-child-agents
  create-children 16
  ask children [set shape "child"]
  ask children [set size 1.5]
  ask children [
    set contactable_child 7
    set contactable_adult 5
    set contactable_old 1
  ]
end

to create-adult-agents
  create-adults 43
  ask adults [set shape "adult"]
  ask adults [set size 2.5]
  ask adults [
    set contactable_child 2
    set contactable_adult 9
    set contactable_old 1
  ]
end

to create-old-agents
  create-olds 7
  ask olds [set shape "old"]
  ask olds [set size 2.5]
  ask olds [
    set contactable_child 1
    set contactable_adult 3
    set contactable_old 2
  ]
end

to infect-children
  (ifelse
    ; Infected agent has mask
    mask_on?
    [
      ask other children in-radius 1.5 with [not infected? and not
cured?]
      [
        (ifelse
          ; Uninfected agent has mask
          mask_on?
          [
            if random-float 100 < 0.10314     ; 0.10314% Probability
            [
              set infected? true
              set color red
            ]
          ]
          ; Uninfected agent doesn't have mask
```

```
          (not mask_on?)
          [
            if random-float 100 < 0.6876    ; 0.6876% Probability
            [
              set infected? true
              set color red
            ]
          ]
        )
      ]
    ]

    ; Infected agent doesn't have mask
    (not mask_on?)
    [
      ask other children in-radius 1.5 with [not infected? and not
cured?]
      [
        (ifelse
          ; Uninfected agent has mask
          mask_on?
          [
            if random-float 100 < 0.6876    ; 0.6876% Probability
            [
              set infected? true
              set color red
            ]
          ]
          ; Uninfected agent doesn't have mask
          (not mask_on?)
          [
            if random-float 100 < 4.584    ; 4.584% Probability
            [
              set infected? true
              set color red
            ]
          ]
        )
      ]
    ]
  )
end

to infect-adults
  (ifelse
    ; Infected agent has mask
    mask_on?
    [
      ask other adults in-radius 1.5 with [not infected? and not
cured?]
      [
        (ifelse
          ; Uninfected agent has mask
          mask_on?
          [
            if random-float 100 < 0.10314     ; 0.10314% Probability
```

```
          [
            set infected? true
            set color red
          ]
        ]
        ; Uninfected agent doesn't have mask
        (not mask_on?)
        [
          if random-float 100 < 0.6876    ; 0.6876% Probability
          [
            set infected? true
            set color red
          ]
        ]
      )
    ]
  ]

  ; Infected agent doesn't have mask
  (not mask_on?)
  [
    ask other adults in-radius 1.5 with [not infected? and not
cured?]
    [
      (ifelse
        ; Uninfected agent has mask
        mask_on?
        [
          if random-float 100 < 0.6876    ; 0.6876% Probability
          [
            set infected? true
            set color red
          ]
        ]
        ; Uninfected agent doesn't have mask
        (not mask_on?)
        [
          if random-float 100 < 4.584    ; 4.584% Probability
          [
            set infected? true
            set color red
          ]
        ]
      )
    ]
  ]
 )
end

to infect-olds
 (ifelse
  ; Infected agent has mask
  mask_on?
  [
    ask other olds in-radius 1.5 with [not infected? and not cured?]
    [
```

```
        (ifelse
          ; Uninfected agent has mask
          mask_on?
          [
            if random-float 100 < 0.10314     ; 0.10314% Probability
            [
              set infected? true
              set color red
            ]
          ]
          ; Uninfected agent doesn't have mask
          (not mask_on?)
          [
            if random-float 100 < 0.6876    ; 0.6876% Probability
            [
              set infected? true
              set color red
            ]
          ]
        )
      ]
    ]

    ; Infected agent doesn't have mask
    (not mask_on?)
    [
      ask other olds in-radius 1.5 with [not infected? and not cured?]
      [
        (ifelse
          ; Uninfected agent has mask
          mask_on?
          [
            if random-float 100 < 0.6876    ; 0.6876% Probability
            [
              set infected? true
              set color red
            ]
          ]
          ; Uninfected agent doesn't have mask
          (not mask_on?)
          [
            if random-float 100 < 4.584    ; 4.584% Probability
            [
              set infected? true
              set color red
            ]
          ]
        )
      ]
    ]
  )
end
```