Spring 4-22-2022

Modeling the Effect of Human Behavior on Disease Transmission

Katie Yan
kyan@skidmore.edu

Modeling the Effect of Human Behavior on Disease Transmission

by

Katie Yan

April 22, 2022

A senior thesis submitted to the

Faculty of the Mathematics and Statistics Department of

Skidmore College

to fulfill the honors requirement for the

mathematics major

Bachelors of Arts

Department of Mathematics and Statistics

Acknowledgments

I would like to thank Professor Rachel Roe-Dale for her support and guidance in helping me create this thesis. I am particularly grateful to Dr. Roe-Dale for the opportunity to participate in the 2019 Summer Research Program where we began exploring disease models after our return from London, England. I would also like to thank the members of the Department of Mathematics and Statistics for their generosity and kindness in supporting my research throughout my time at Skidmore College.

Land Acknowledgement

I would like to acknowledge that this research at Skidmore College was performed on land stolen from the Haudenosaunee, Mohawk, and Mohican peoples. In offering this land acknowledgement, I affirm Indigenous sovereignty, history and experiences and I invite all to recognize our shared responsibility to acknowledge, honor, and make visible Skidmore College's relationship to Native peoples.

Table of Contents

List of Figures

Abstract

Many infectious disease models build upon the classical Susceptible-Infected-Recovered (SIR) model. The SIR model is a compartmental model that is used to model disease transmission throughout a population. The SIR model and its variations often focus on the transmission of disease but rarely include behavioral or informational components that explore how the perception of a disease influences transmission. In this thesis we propose a six compartment SIR model that segments the classical SIR model based on knowledge of information to explore the sharing of information and its ability to increase and decrease transmission. We designate these two model states as aware and unaware based on whether the information is known by the population. We find that while good behavior is useful in decreasing community transmission, bad behavior is significantly more damaging for the community in terms of disease transmission. These preliminary results suggest that more research is needed on the effect of information and behavior on disease transmission.

Chapter 1

Mathematical models of disease

"All models are wrong, but some are useful"

George E. P. Box, British statistician (1919-2013)

Mathematical models provide quantitative solutions to real-world problems. Examples range from determining the growth rate of a cell culture to predicting changes in stock price. Models range in complexity from simple algebraic equations to complex systems of ordinary differential equations (ODEs), and models are used to predict or forecast outcomes under a certain set of assumptions and conditions.

In this thesis, we focus on the mathematical modeling of infectious disease in humans. These diseases are caused by infectious agents including viruses, bacteria, and protozoa. Often we model the spread or transmission of diseases from human to human hosts, and sometimes models will incorporate secondary animal hosts. In this thesis, we model human to human transmission.

Recently the relevance of mathematical modeling research has increased due to the ongoing COVID-19 pandemic. This resurgence has brought to light the importance and limitations of such models. In forecasting the possible spread of disease, we are able to predict

hot-spot locations for the disease, transmission rates, and effective prevention and protection measures. However, for each of these predictions, the model assumptions must be analyzed. All mathematical models of disease use assumptions to decrease computational load and reduce scenario complexity. For example, a model with few assumptions will quickly grow in complexity as more scenario factors are taken into account. Conversely, a model that makes too many assumptions might fail to describe the scenario, leaving out key ideas that influence the model's behavior. It is for these reasons that it is important that we thoroughly analyze and question our assumptions when creating our model. By adjusting and simplifying our models using appropriate assumptions, we can reduce computation time and make our models more accessible for a variety of disciplines.

1.1 The SIR Model

The Susceptible-Infected-Recovered (SIR) compartmental model is one of the most common models used to describe the transmission of a disease within a population. Originally created by Kermack and McKendrick of the Royal College of Physicians in Edinburgh, UK in 1927, the SIR model can be most simply described by the compartmental model shown in Figure 1.1 [1]. The susceptible (S) compartment represents individuals in a population who are susceptible to disease infection but are not currently infectious or infected. The infected (I) compartment represents individuals in the population who both carry the disease and can spread it to susceptible individuals. Finally, recovered (R) individuals have surpassed the infectious stage of the disease and no longer can transmit the disease. Individuals in the population will move from the S class to the I and finally the R as the disease progresses.

In this SIR compartmental model, people in the susceptible compartment move at rate β to the infected compartment. Thus we can think of β as the infection rate of our disease. Likewise, the recovery rate α describes the rate at which individuals recover from the disease and are no longer infectious. As with all models, several simplifying assumptions are made.

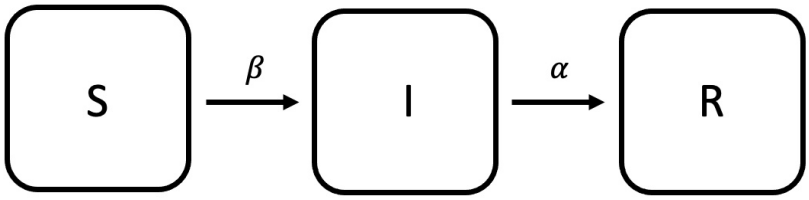

Figure 1.1: SIR Compartmental Model
Individuals become infected at rate β and recover at rate α.

These assumptions include:

1. Population size N is constant. There are no births or deaths.

2. After someone recovers from the disease, they are considered immune to the disease.

3. The incubation period of the disease is short enough to be ignored. The incubation period is defined as the period in which an individual is infected but not infectious; during this time the individual might or might not be aware they are infected.

4. The SIR compartments are uniformly mixed to ensure equal risk of catching the disease.

5. Healthy individuals in the susceptible compartment get sick at a rate proportional to the product of S and I. This follows from the assumption that the population is uniformly mixed to ensure an equal risk of getting the disease.

The simple SIR compartmental model can be transformed into a series of ODEs to describe the rate of change of each compartment. The ODE is given by System of Equations 1.1 and describes the movement of individuals from one compartment to another. The SIR model as described in System of Equations 1.1 has two parameters α and β. In the model, β has units of one over time per individual and α has units of one over time.

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - \alpha I \\ \frac{dR}{dt} &= \alpha I \end{aligned} \tag{1.1}$$

$$N(t) = S(t) + I(t) + R(t) \tag{1.2}$$

In the SIR model our total population size is denoted by N. Note from assumption 1, N is a constant. To show that N is constant, we will show that $\frac{dN}{dt} = 0$. First, recall that the equation for N is described in Equation 1.2. Taking the derivative of Equation 1.2 and applying the derivatives of S, I, and R from System 1.1 we find

$$\begin{aligned} \frac{dN}{dt} &= \frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} \\ \frac{dN}{dt} &= -\beta SI + \beta SI - \alpha I + \alpha I \\ \frac{dN}{dt} &= 0 \end{aligned}$$

Since population size is a non-negative variable, the initial conditions for our model are $S(0) = S_0 > 0$, $I(0) = I_0 > 0$, and $R(0) = 0$. That is the number of susceptible and infected individuals are greater than 0, while the initial number of recovered individuals is 0. A key value we determine from our system of equations is R_0, "R-Naught", the basic reproductive rate of infection. R_0 is defined as the number of secondary infections resulting from a single initial infection in a population where $S(0) = N-1$ and $I(0) = 1$. Typically, R_0 describes the infectiousness of a disease, it is defined as $R_0 = \frac{\beta S_0}{\alpha}$. If $R_0 > 1$, each infectious individual on average will infect more than one other individual. Previous studies on infectious diseases have shown that R_0 often varies from disease to disease. For example, the calculated R_0 for the 1918 influenza was between 2 and 3 [2]. Utilizing systems of equations to describe

the transmission of a disease throughout a population, we can determine if there will be an epidemic, how many people an infected individual will infect, and various parameter tests that alter the values of α and β.

1.2 Example SIR Model, Eyam Second Wave

As describe previously, we can apply System of Equations 1.1 to real-world scenarios. A noted example of an SIR model is G. F. Raggett's 1921 Eyam second wave model [3]. Eyam, England located in the Sheffield region is also where Raggett hails from. Raggett's work was later expanded on by D. Sulsky who graphed his proposed system of equations and parameters, exploring the similarities between the real world data and the SIR model's predictions. Sulsky's work will be discussed later in this thesis.

Eyam's plague epidemic began in 1665 with a package containing a damp cloth from London that was delivered to Eyam's tailor. The cloths were opened and laid out to dry by the tailor's assistant. It is believed that the cloth contained fleas or eggs which were infected with the bubonic plague. A few days after the cloth had been delivered, the tailor's assistant was dead. What followed was one of the most deadly plague outbreaks in European history. In the small village of approximately 350 residents, by the end of the epidemic only 83 villagers were still alive. However, the death toll could have been much higher had the village rector, William Mompesson, not made the decision to cease all travel to and from Eyam to protect the surrounding villages. In the summer months, after the initial wave of the plague, Rector Momepesson gathered all of Eyam's remaining residents to propose a self-imposed village-wide quarantine. Encouraged by the summer weather and declining number of plague infections, the villagers agreed. Nonetheless, in the fall the plague returned causing a second wave of infections. This wave lasted until November of 1666 when the last plague related death was recorded.

From Rector Mompesson's historical records of plague related deaths, Raggett was able to

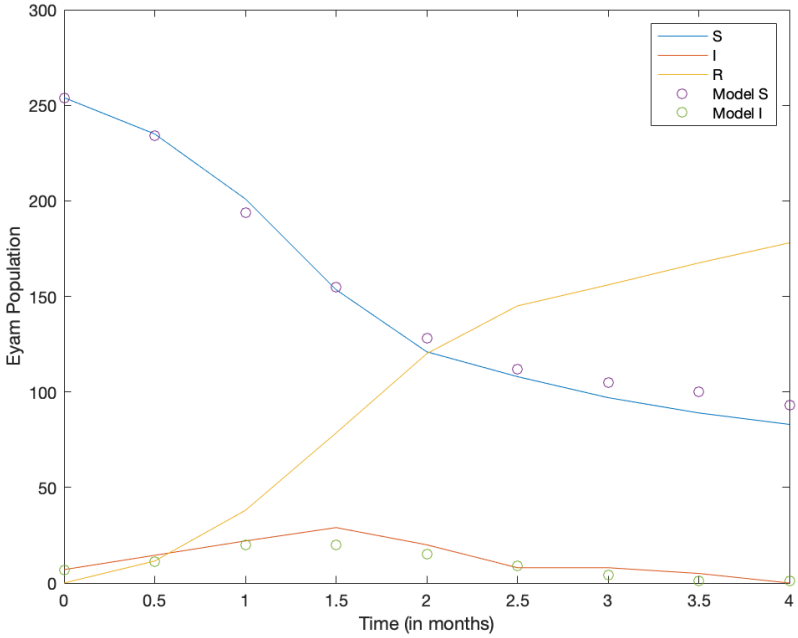

Figure 1.2: Eyam Predictions and Historical Records
Comparison of Rector Mompesson's historical Eyam records (solid lines) and Raggett's SIR model predictions (circles).

create his SIR model of the second wave of infection. These results are reproduced in Figure 1.2 using Matlab. Using these values Raggett found that at the start of the second wave of infection there were 254 remaining uninfected villagers and 7 infected villagers. Raggett also used these death records to estimate the average rate of infection, $\alpha = 2.82$ months^{-1} and death rate, $\beta = \frac{2.82}{159}$ (people $\times$ months)$^{-1}$.

As shown in this section, SIR models of disease can be fit to real world data providing insight and prediction power to similar scenarios. For example, using Raggett's β and α parameters we could simulate the spread of the plague in a population of 1000 individuals, 999 susceptible and 1 infected. We could also manipulate β and α to fit a different disease such as COVID-19.

1.3 Epidemic and Pandemic Definitions

Although the term *epidemic* was used previously to describe the plague outbreak in Eyam, there is no universally agreed upon quantified definition for when a disease outbreak is classified as an epidemic. Nor is there an agreed definition for *pandemic*. However, pandemics are typically classified as epidemics that span a wide area.

In this thesis, we consider an outbreak an *epidemic* when the infection curve has a sustained positive slope. When $\frac{dI}{dt} > 0$, this indicates that the number of individuals being infected at each time step is greater than the number of individuals who recover from the disease. Thus, we see an increase in the number of infected individuals. This increase must be sustained for a significant duration of time, which is unique to the disease. Eventually, the peak will plateau and decline as the number of susceptible individuals decreases. However, until the curve begins its downward trajectory we classify the rapid increase of infected individuals as an epidemic.

Beyond defining epidemics we can also classify them into to broad categories. In cases where $\frac{dI}{dt}$ is large we can expect a sharp increase in the number of infected individuals. However, large spikes in the number of infected individuals do not often lead to long-term increases in disease transmission. Rather, spikes tend to quickly increase then decrease. Conversely, a small $\frac{dI}{dt}$ is usually associated with a long term period of infection where only a small proportion of the population is infected at any given point in time, but the disease persists in the population for a significant period oftentimes. In future studies it would be useful to explore for what values of $\frac{dI}{dt}$ a significant proportion of the is population infected and the differences between the two behaviors described above.

1.4 Variations on the SIR Model

The classic SIR model can be modified to add or remove assumptions. In this section we give an overview of a few popular variations on this classic model. In particular, we examine SIR models that remove assumptions 1-3. Note that each model we present does not build on the previous model. Instead, we consider three separate models and the scenarios they describe.

1.4.1 SIR with demography

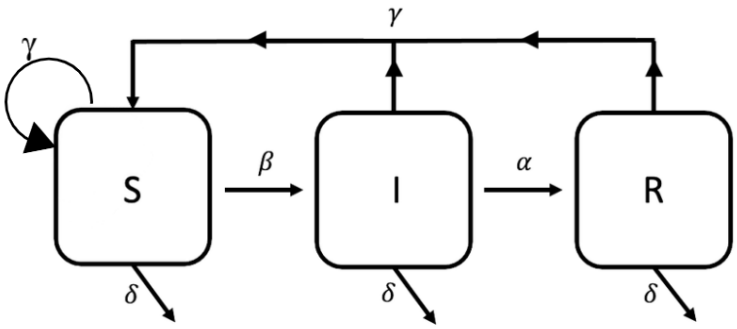

Figure 1.3: Open SIR Compartmental Model
Adding births (at rate γ) and deaths (at rate δ).

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI + \gamma(S+I+R) - \delta S \\ \frac{dI}{dt} &= \beta SI - \alpha I - \delta I \\ \frac{dR}{dt} &= \alpha I - \delta R \end{aligned} \tag{1.3}$$

The classic SIR model is a closed model. This qualification means N is a constant, such that there are no births or deaths within the population. Open versions of the SIR models will account for births and deaths in the population. Adding these properties yields the compartmental model displayed in Figure 1.2. From this model we write the equations

described in System of Equations 1.3, where γ is the population birth rate of new susceptible individuals and δ is the death rate from non-disease related causes. Note that it is assumed in this model that newborns are susceptible to the disease and are therefore not born with inherited immunity.

1.4.2 SIRS Model

Removing assumption 2, we can include waning immunity to the disease over time, as displayed in Figure 1.3. This model is typically referred to as the SIRS model whereby individuals transition from the recovered class back to the susceptible class as their immunity to the disease decreases.

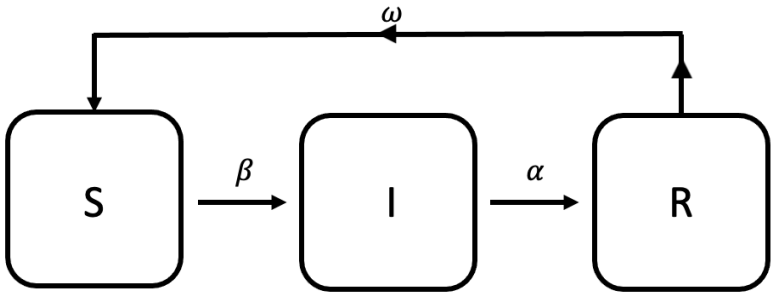

Figure 1.4: SIRS Compartmental Model
Including waning immunity (at rate ω) where individuals can become reinfected with the disease.

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI + \omega R \\ \frac{dI}{dt} &= \beta SI - \alpha I \\ \frac{dR}{dt} &= \alpha I - \omega R \end{aligned} \tag{1.4}$$

ODEs derived from this compartmental model are described in System of Equations 1.4 where ω represents the rate of waning immunity of individuals in the recovered compartment. This model is commonly used for diseases where individuals can be reinfected. Examples of

such diseases include the seasonal influenza, STDs, and the common cold.

1.4.3 SEIR Model

Finally, we remove assumption 3, and instead consider a non-negligible incubation period where individuals are infected but not infectious, meaning they cannot spread the disease. This model is typically called the SEIR model where E represents the exposed class [4]. The SEIR model gives the compartmental diagram as seen in Figure 1.4.

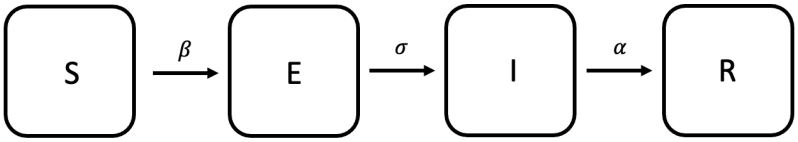

Figure 1.5: SEIR Compartmental Model
Adding an exposed compartment E for individuals who are not yet infectious and latency rate σ.

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dE}{dt} &= \beta SI - \sigma E \\ \frac{dI}{dt} &= \sigma E - \alpha I \\ \frac{dR}{dt} &= \alpha I \end{aligned} \tag{1.5}$$

In this system of equations, we include $\frac{dE}{dt}$ to represent the rate of change of the exposed class. Notice that individuals in the exposed compartment do not infect individuals in the susceptible compartment. To become infected a susceptible individual must still contact an infectious individual. Thus, our rate of infection is still β and the rate of individuals becoming infectious is σ.

The classic SIR model and its variations are a frequently used template for modeling infectious disease transmission. These models can be further modified to represent different

conditions and assumptions. In this section, we described three common modifications that can be applied to the classic SIR model. Moving forward we will consider the classic SIR model in a nondimensionalized form and use behavioral epidemiology techniques to explore the effect of human behavior and information flow on the spread of disease.

1.5 Behavioral Epidemiology

Behavioral epidemiology (BE) is the study of human behavior, its influences, and its impact on the spread of infectious diseases. Developed in the 1970s, BE has grown as an interdisciplinary field that draws from mathematics, biology, sociology, psychology, economics, and other disciplines that study human behavior and disease [5]. In 2000, Sallis et al. published a systematic framework describing the five major phases of BE research. The first phase aims to find connections between human behavior and health. The next phase quantifies or measures the behavior of interest. The third phase further explores the behavior by studying the factors that influence it. In the fourth phase interventions that change or prevent the behavior are studied. Finally, in the fifth phase, the four previous phases are used to inform future practices that mitigate or prevent the behavior.

A recent example of research following the five phases proposed by Sallis et al. is the ongoing COVID-19 pandemic and problematic behaviors that increase disease transmission. In the first phase of COVID-19 research, close contact was linked to higher rates of disease transmission. Then, further research quantified the behavior, revealing that a distance of 6 feet or fewer between two people was considered close contact. As information and misinformation spread about COVID-19, a popular piece of misinformation was that social distancing was ineffective in stopping the spread of COVID-19. Thus, the misinformation influenced antisocial distancing behavior. As the pandemic continued, public health messaging was utilized as an intervention technique, combining Sallis' intervention and implementation phases.

While Sallis et al.'s framework continues to serve as a useful tool for describing research

phases, the modern interpretation of the phrase "behavioral epidemiology" has shifted to a more mathematical and modeling-centered perspective. Specifically, the modern interpretation of BE still aims to understand the effects of determinants on human behavior, but instead of an empirical approach, modern studies tend to use a computational approach. Common determinants of human behavior include: the acquisition of information, risk perception, benefits perception, and trustworthiness of information. Modern BE research is an interdisciplinary study of epidemiology from a human-centered perspective grounded upon the belief that human behavior drives the transmission of disease.

Looking further at modern BE applications, in classic SIR models and variations we operate under the assumption that people within our population of interest behave as we would expect, homogeneously. This means there are no outliers in terms of human behavior and we expect even mixing of susceptible, infected, and recovered individuals. However, this is rarely the case in real-life disease scenarios. More often than not in real-world epidemic and pandemic scenarios, individuals will react and change their behavior based on various influences they experience. Within the current COVID-19 pandemic we see common examples of influences of human behavior causing individuals to act in such a way that causes both positive and negative effects on the spread of COVID-19. A positive effect increases the spread of disease. An example of an influence leading to a positive effect includes anti-masking and anti-vaccine information. On the other hand, a negative effect like the practice of social distancing decreases the spread.

Funk, Gilad and Jansen use BE in disease modeling in their 2010 paper that describes a six compartmental model whereby a population is divided into two main sub-classes, aware and unaware [6]. These sub-classes denote whether individuals in each class know the information. Each sub class is then further divided based on the standard SIR model. This model is shown in Figure 1.6, which was reproduced from [6].

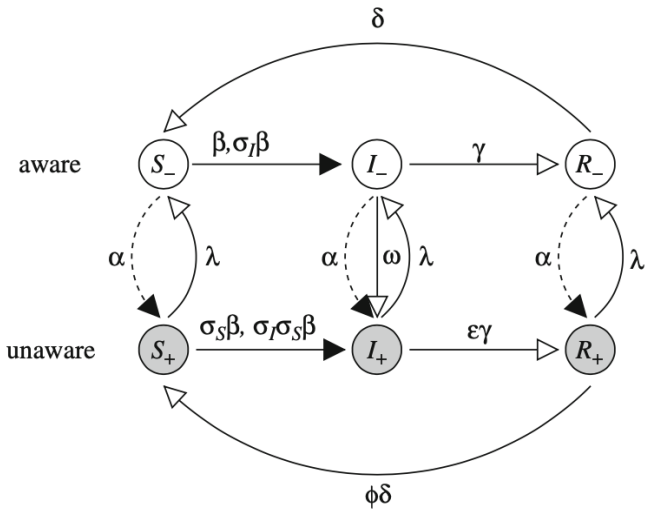

Figure 1.6: Six Compartment SIR Model including Behavior
Knowledge is denoted by *aware* and *unaware* groups and information is shared across groups (at rate λ).

The described six compartment model features one SIRS model for those who are aware and one for those who are unaware. Awareness spreads like a disease and the transmission of information is akin to a susceptible-infected-susceptible (SIS) model where after one becomes "infected," the information fades from memory after a period of time. Because the information that is being transmitted is undefined in its quality and truthfulness, the authors perform multiple tests using beneficial and detrimental information. If information is useful, disease transmission decreases within the aware population. If information is harmful, disease transmission increases amongst the aware population. Within the model, the authors define 11 parameters shown in Figure 1.6.

Funk et al. modeled an endemic disease, a persistent disease in a population, to understand the effect of information on disease transmission. Using mean-field analysis, Funk et al. found that information changes the conditions for disease-free and endemic equilibrium. Generally, this suggests that information can prevent disease establishment in the population. Additionally, Funk et al. found that when information reduced the intensity of infection, caused a shorter duration of infection, or reduced the susceptibility of non-infected individ-

uals, it was difficult for the disease to establish itself. A combination or triad of these effects of information made it even more difficult for the establishment of the disease. Finally, Funk et al. found that as the disease approached equilibrium, its rate of spread decreased. This meant that if information was able to spread thoroughly throughout the population, the size of the disease outbreak is prominently due to the rate of disease transmission. These findings suggest that mathematical models incorporating behavioral epidemiology are a useful tool in understanding the complexities of disease transmission and suggest the importance of public health messaging to curb the spread of disease.

When mathematical models of diseases fail to account for human behavior, these models fail to fully describe the population of interest. BE has grown from a theoretical field to an applied set of metrics that can be used to build robust mathematical models of diseases to predict the effect of human behavior on disease transmission such as in Funk et al [6]. In this thesis we will use the schema described by BE to inform our model creation as we aim to understand the influence of both helpful and harmful information. Specifically, we are inspired by the recent rise in misinformation throughout the COVID-19 pandemic and its potential effect on disease transmission.

Chapter 2

Nondimensionalized Model

The study of dynamics analyzes systems that evolve over time. In this thesis we will use techniques from dynamical systems to examine the evolution of a disease within a population over time. Using the classic SIR model first proposed by Kermack and McKendrick, we determine for what values of α and β an epidemic will occur. We first solve the SIR model, then perform a fixed point analysis, and finally determine the threshold of an epidemic on the nondimensionalized equation. We used a nondimensionalized form to remove units from our equations and to simplify our analyses.

Moving forward we adopt the dot notation such that, for example $\dot{S} = \frac{dS}{dt}$. Recall, the SIR model given by System of Equations 1.1 is:

$$\dot{S} = -\beta SI$$

$$\dot{I} = \beta SI - \alpha I$$

$$\dot{R} = \alpha I$$

To start we look for solutions of the form:

$$S(t) = S_0 \cdot \exp\frac{-\beta R(t)}{\alpha} \tag{2.1}$$

To find solutions of this form we take the derivative of equation 2.1, and then substitute in for $\frac{dR}{dt}$ from System of Equations 1.1.

$$\begin{aligned}
\dot{S} &= S_0 \cdot \exp\frac{-\beta R(t)}{\alpha} \cdot \frac{-\beta}{\alpha} \cdot \dot{R} \\
&= S(t) \cdot \frac{-\beta}{\alpha} \cdot \dot{R} \\
&= S(t) \cdot \frac{-\beta}{\alpha} \cdot \alpha I \\
&= S(t) \cdot -\beta I \\
&= -\beta S I
\end{aligned}$$

Using our new equation for $\dot{S}$ and the fact that $N(t) = S(t) + I(t) + R(t)$, we can rewrite $\dot{R}$ as

$$\begin{aligned}
\dot{R} &= \alpha I \\
&= \alpha[N - R - S] \\
&= \alpha[N - R - S_0 \cdot \exp\frac{-\beta R(t)}{\alpha}]
\end{aligned}$$

We next distribute α throughout.

$$\frac{dR}{dt} = \alpha N - \alpha R - S_0 \cdot \alpha \cdot \exp\frac{-\beta R(t)}{\alpha}$$

As previously mentioned nondimensionalization involves the removal of units from our parameters and variables. Our next step is to make a change of variables, introducing a dimensionless state variable u. Here we let

$$u = \frac{\beta R}{\alpha} \text{ so that } R = \frac{\alpha u}{\beta} \text{ and differentiating gives: } \frac{du}{dt} = \frac{\beta}{\alpha}\frac{dR}{dt} \tag{2.2}$$

Next we nondimensionalize time by introducing τ:

$$\tau = S_0\beta t \text{ and differentiating gives: } d\tau = S_0\beta dt \text{ and } dt = \frac{d\tau}{S_0\beta} \tag{2.3}$$

Using these new variables u and τ and their differentials we substitute them into our previous equation to find

$$\begin{aligned}
\frac{\alpha}{\beta}\frac{du}{dt} &= \alpha N - \alpha(\frac{\alpha u}{\beta}) - S_0\alpha e^{-u} \\
\frac{1}{S_0\alpha}[\frac{\alpha}{\beta}\frac{du}{dt} &= \alpha N - \alpha(\frac{\alpha u}{\beta}) - S_0 \cdot \alpha e^{-u}] \\
\frac{1}{S_0\beta}\frac{du}{dt} &= \frac{N}{S_0} - \frac{1}{S_0}(\frac{\alpha u}{\beta}) - e^{-u} \\
\frac{du}{d\tau} &= \frac{N}{S_0} - \frac{\alpha u}{S_0\beta} - e^{-u}
\end{aligned}$$

Finally, we set a and b to be

$$\begin{aligned}
a &= \frac{N}{S_0} \\
b &= \frac{\alpha}{S_0\beta}
\end{aligned}$$

Substituting in these values results in our dimensionless equation

$$\frac{du}{d\tau} = a - bu - e^{-u} = 0 \tag{2.4}$$

We note that $a \geq 1$ as there will always be more people in the population than in the susceptible compartment since $N \geq S_0 > 0$. Additionally, $b > 0$ as we know that $\alpha, \beta, S_0 > 0$ from our model assumptions. Finally, recall that $R_0 = \frac{S_0\beta}{\alpha}$ and that $b = \frac{1}{R_0}$.

From the dimensionless model Equation 2.4, we can find fixed points $u^\star$. Fixed points

represent equilibrium solutions for Equation 2.4. Equilibrium solutions are points where $\frac{du}{d\tau} = 0$. These points can either be unstable or stable depending on if a response to a small disturbance away from $u^\star$ causes the solution to return to equilibrium or move away over time. Here $\frac{du}{d\tau} = 0$ means that $a - bu - e^{-u} = 0$ and that at a fixed point,

$$a - bu = e^{-u} \tag{2.5}$$

To find fix points, we use roots of Equation 2.4. Using this form, equilibrium solutions can be found by setting $a - bu$ and e^{-u} equal to each other. This form is easy to analyze graphically and so we plot the two sides of Equation 2.5 and look for intersections. Intersections represent points where Equation 2.4 is equal to 0.

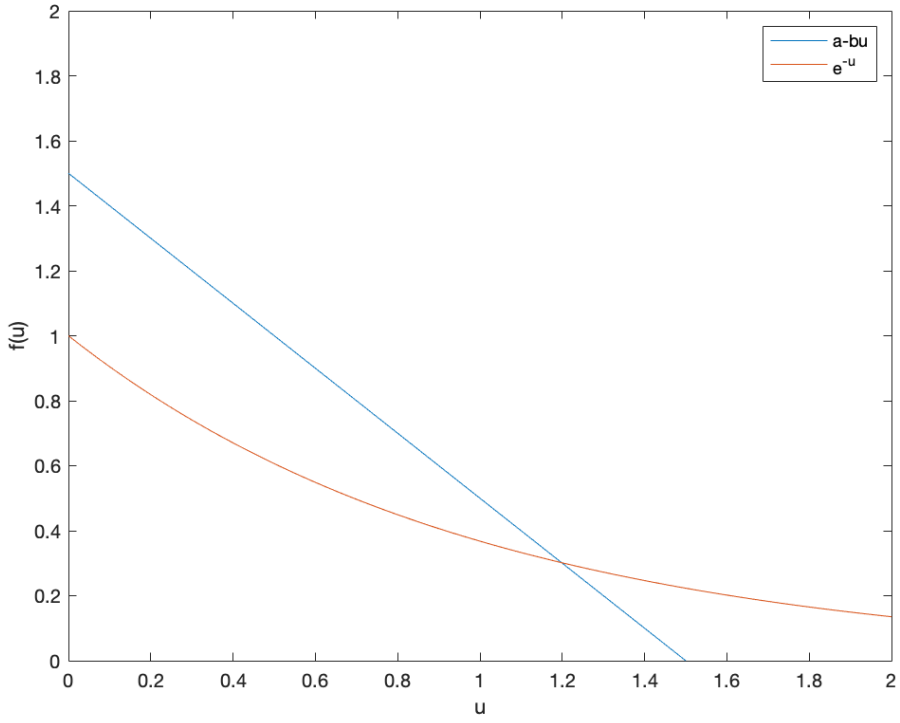

Figure 2.1: Fixed Point Analysis of u
Where $a = 1.5$ and $b = 1$ are typical parameter values. In this scenario the fixed point corresponds to $u = 1.2$.

As mentioned previously, we can analyze equation 2.5 using graphical methods. From

Figure 2.1, we see there is one fixed point, $u^\star$, where $a - bu$ and e^{-u} intersect. Note that $a - bu$ and e^{-u} intersect at another point, not shown, but because at that point $u < 0$, it is not relevant to our analysis.

Thus far we have converted part of the classic SIR system of equations to a dimensionless ODE modeling rate of change in R. Now we will use the dimensionless form to determine for what values of b an epidemic occurs. Remember that $b = \frac{1}{R_0}$. Given that Equation 2.4 is a rescaled version of $\dot{R}(t)$, we use $\frac{du}{d\tau}$ as our new recovery rate. First, we denote t_{peak} as the time of the epidemic peak, when $\frac{du}{d\tau}$ is at its maximum. Let us now consider three cases: $b < 1$, $b > 1$, and $b = 1$. In our first case $b < 1$, which means $R_0 > 1$.

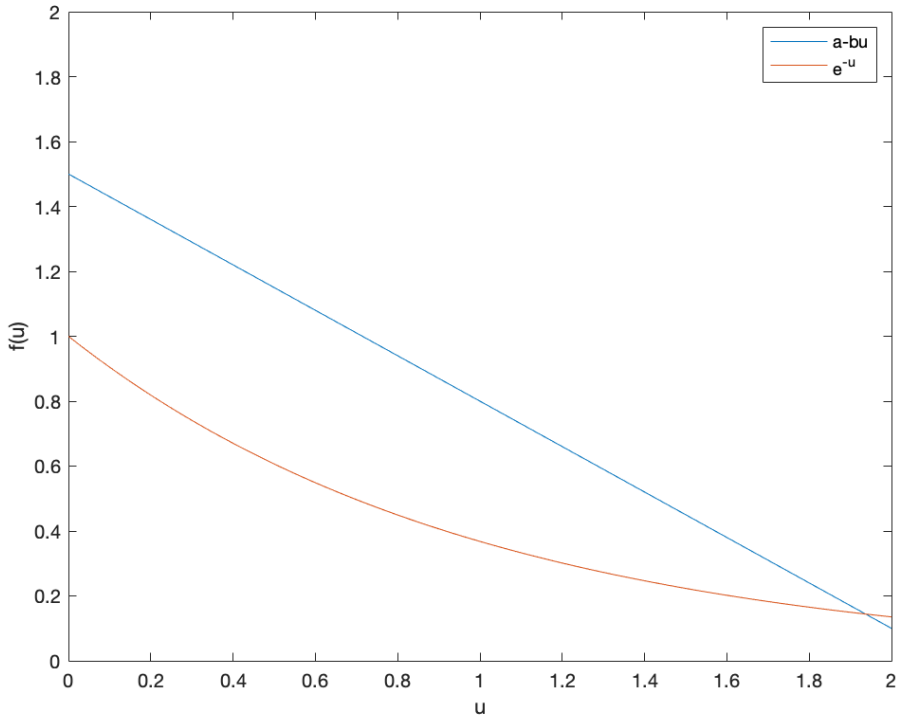

Figure 2.2: Outbreak Analysis when $b < 1$
$\frac{du}{d\tau}$ when $b = 0.7$ and $a = 1.5$. Note that the fixed point occurs at $u = 1.9$.

Equation 2.4 is maximized when the distance between $a - bu$ and e^{-u} is the largest. As shown in Figure 2.2 this occurs between $u = 0.2$ and 0.6, and then the difference decreases to 0 as u approaches $u^\star$. Note that $u^\star$ is approximately 1.9. Thus, $t_{peak} > 0$ because the distance

between $a - bu$ and e^{-u} first increases then decreases. In this scenario t_{peak} is approximately 0.35 for the specific set of parameters used in Figure 2.2. For all scenarios where $b < 1$, we observe the same pattern of the death rate increasing then decreasing to $u^{\star}$. This indicates that the spread of disease first increases, causing an epidemic, then decreases as the number of infected individuals decrease such that there are not enough infected individuals for the disease to continue spreading. In essence, in this scenario things must get worse before they can get better.

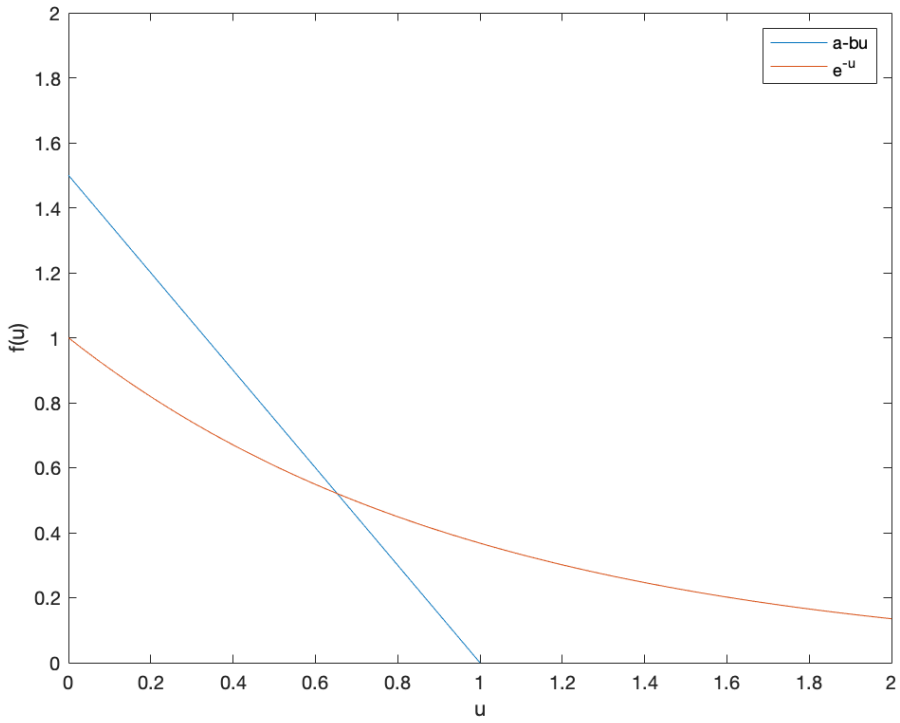

Figure 2.3: Outbreak Analysis when $b > 1$
$\frac{du}{d\tau}$ when $b = 1.5$ and $a = 1.5$. Note that the fixed point occurs at $u = 0.7$.

In our second case $b > 1$ and $R_0 < 1$. We find $t_{peak} = 0$ because the distance between $a - bu$ and e^{-u} decreases such that the greatest distance between the two functions is when $u = 0$. This can also be shown as $f'(u) < 0$ so the difference will always be decreasing. Although there is disease transmission, we would describe this scenario as a small outbreak, not an epidemic as the disease spreads steadily but does not have an increase as in Figure

2.2.

Finally, let us consider case three when $b = 1$. We return to Figure 2.1 where $b = 1$ so $R_0 = 1$. We know from previous analysis that when $b > 1$ an epidemic will not occur whereas for $b < 1$ an epidemic will occur. Because of the behavior exhibited when b is less than or greater to 1 we consider $b = 1$ the threshold for an epidemic. Remember earlier we defined $b = \frac{\alpha}{S_0\beta} = \frac{1}{R_0}$. When $b > 1$ we define this scenario as when the number of people who are becoming infected is less than the number of individuals recovering from the disease. It follows that an epidemic would not occur. Conversely, when $b < 1$ more individuals are becoming infected than recovering, creating an epidemic. In this chapter we have nondimensionalized our System of Equations 1.1 to perform an epidemic analysis. In the following chapters we will present our Behavioral SIR Model and explore the effects of human behavior on disease transmission.

Chapter 3

Behavioral SIR Model

As previously shown, mathematical models of disease rarely consider the effect of behavior or the sharing of information. In this chapter we investigate disease and information sharing. To facilitate this investigation we propose a 6 compartment mathematical model and corresponding system of equations similar to Funk et al. in [6]. Finally, we confirm that our proposed model allows for the use of different time scales and their corresponding parameter values.

One's behavior is often influenced by the nature of or context of the information they receive, the source of the information, and their openness to accepting new information. For example, an individual who does not believe in science will be a skeptic of any information received from a scientific source. In the midst of a pandemic, the decision to not trust factual public health information could lead to individuals engaging in risky behaviors. Conversely, if someone believes in science and hears the same public health information, they might be more likely to engage in less risky behavior. Ultimately, one's openness to receiving new information, regardless of its trustworthiness or legitimacy, is more indicative of their behavior than the specifics of the message. In both scenarios the information is received but only in the case of the individual who believes in science is any action taken. Thus, individuals who engage in risky behaviors are more likely to have a higher disease transmission rate

than those who protect themselves. When considering a mathematical model of disease, information about behavior is an essential component to include when one considers the volatility of human to human transmission.

3.1 The Behavioral Model of Disease

Our proposed behavioral model of disease features 6 compartments, similar to Funk et al.'s model in [6]. As seen in Figure 3.1, the Susceptible, Infected, and Recovered compartments are subdivided into aware (a) and unaware (u) groups depending on if the individuals within each compartment are aware of the information.

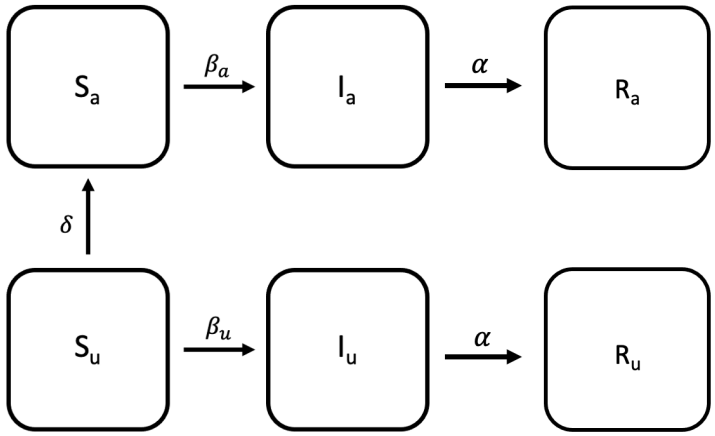

Figure 3.1: Behavioral SIR Six Compartment Model
Knowledge is denoted by subscript u (unaware) and a (aware) with different infection rates (β_u or β_a) for each group. Information is shared across groups at rate δ.

Individuals begin in either the aware or unaware group. As time progresses, individuals spread the information and the disease. The spread of information is represented by individuals moving from the unaware group to the aware group. Similarly, the spread of disease is represented by individuals moving from the Susceptible compartment to the Infected compartment corresponding to their aware or unaware state. Once an individual is infected, regardless of if they know the information or not, they progress through the I compartment

to the R compartment as they would progress in the classic SIR model.

To differentiate the types of information, we use two infection rate parameters β_a and β_u. In scenarios where knowing the information is helpful $\beta_a < \beta_u$. For example, useful information could be recommending protective measures like social distancing, isolation, or quarantine when an individual is infectious or a close contact of an infectious individual. When the information being spread is harmful $\beta_a > \beta_u$. Examples of harmful information include misinformation about the disease such as suggestions that the disease is not real or recommendations to not follow public health measures.

Similar to the classic SIR model, in the behavioral SIR model after an individual is infected, they recover at rate α, regardless if they know the information and if the information is helpful or harmful. Additionally, we assume the birth and death rates are negligible. Finally, the most important difference between our described model and the classic SIR is the addition of an information sharing term δ. The parameter δ represents the flow of information from the aware population to the unaware. In this model, we assume that information can spread only between susceptible people, individuals in the Infected and Recovered compartments are unable to share information, and individuals cannot forget the information over time. All parameters and their meanings are shown in Table 3.1

Parameter	Meaning
β_u	Infection rate for unaware population
β_a	Infection rate for aware population
α	Population recovery rate
δ	Information sharing rate

Table 3.1: Parameter definitions

Using the compartmental model described and shown in Figure 3.1, we write a system of six equations describing the sharing of information and spread of disease through the population. Notice, that S_a, I_a and R_a form a classic SIR model as do S_u, I_u and R_u. The only movement between these two models is through the addition of the term $\delta(S_a S_u)$.

$$\begin{aligned}
\dot{S}_a &= -\beta_a(S_a I_a) + \delta(S_a S_u) \\
\dot{I}_a &= \beta_a(S_a I_a) - \alpha I_a \\
\dot{R}_a &= \alpha I_a \\
\dot{S}_u &= -\beta_u(S_u I_u) - \delta(S_a S_u) \\
\dot{I}_u &= \beta_u(S_u I_u) - \alpha I_u \\
\dot{R}_u &= \alpha I_u
\end{aligned} \tag{3.1}$$

Like Funk et al.'s six compartment model, in System of Equations 3.1 we see that information is shared from the aware susceptible compartment to the unaware susceptible compartment. Akin to the classic SIR model we assume a homogeneously mixed population within the aware and unaware groups. However, this model assumes that the disease flows within the aware and unaware groups with no disease transmission between them.

3.2 RStudio Model

To simulate epidemic scenarios using System of Equations 3.1, we use RStudio along with the tidyverse, DescTools, and deSolve packages [7, 8, 9]. We utilize RStudio to generate plots of the simulations over a specified length of time to investigate the effects of varying parameter values and to determine final and maximum values for each compartment. Additionally, the shinySIR package is used to vary the values of β_a, β_u, δ, and α using sliders [10].

3.2.1 Model Validation

To create these simulations we use shinySIR to plot a figure that updates as parameter slider values are adjusted. A screenshot of an example shinySIR output is shown in Figure 3.2. Beyond changing parameter values, model conditions can also be adjusted. For example the

number of time points for integration and the parameter maximum and minimum values can be set to specific values.

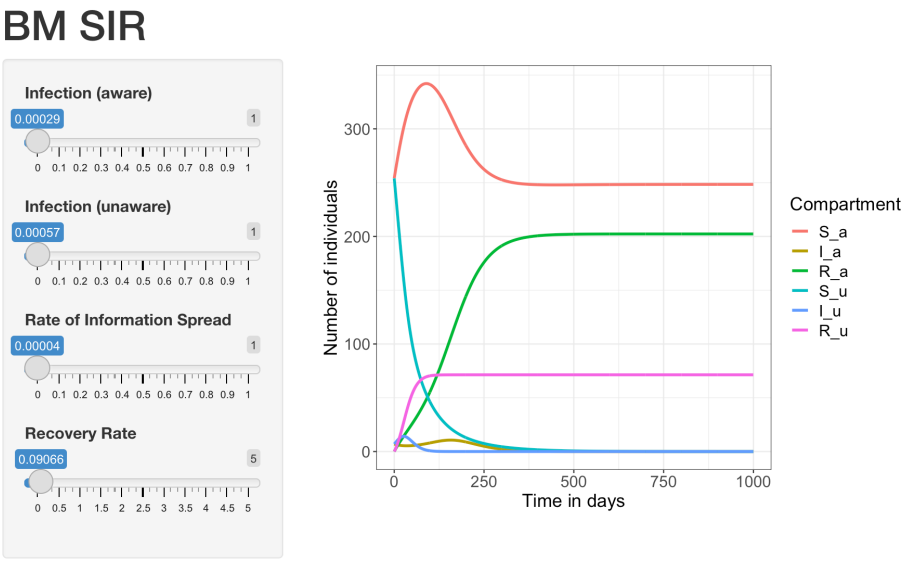

Figure 3.2: Example shinySIR Output
Parameter sliders on the left and graph on the right. Compartment labels are shown on the far right with differing colors. As parameter sliders are changed the graph will update in real-time.

To demonstrate our methods on known data, we recreated Figure 3.3**A** with our algorithms to replicate Sulsky's analysis of the second wave of the plague in Eyam, England [11]. Originally proposed by Raggett, Sulsky's work graphs the SIR model. In his original paper, Raggett determines $\alpha = 2.82$ (months^{-1}) and $\beta = \frac{\alpha}{159}$ (people $\times$ months)$^{-1}$ as the recovery and infection rates respectively. The initial conditions for the second wave of the Eyam epidemic were $S_0 = 254$ people, $I_0 = 7$ people, and $R_0 = 0$ people. For our simulations using the Eyam data, as shown in Figure 3.3 we use the unaware group as the overall population, setting the shinySIR initial conditions so that there are no susceptible or infected individuals in the aware group.

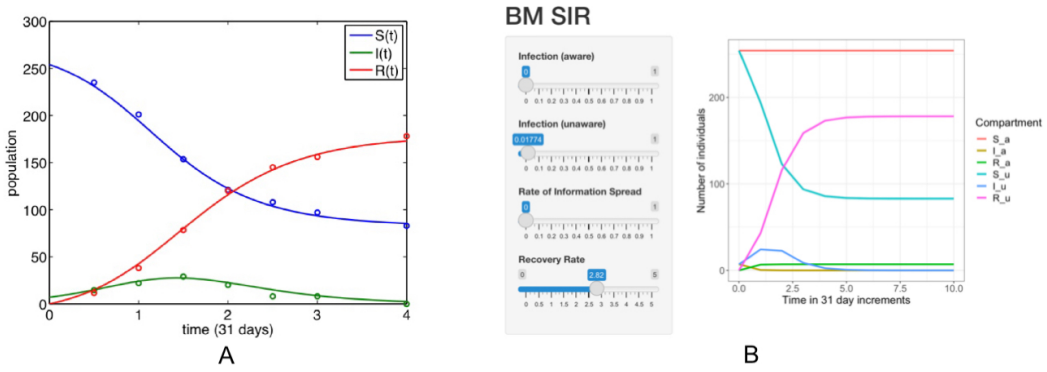

Figure 3.3: Eyam Second Wave Predictions
Figure 3.3A shows Sulksy's plot of the SIR model, solid lines, compared to data from Rector Mompesson's historical records, open circles as given in [11]. Figure 3.3B shows shinySIR recreation of Sulsky's plot using the unaware group as a proxy, not including the data points from Mompesson's historic records.

3.2.2 Time Scaling in the Model Simulations

Before testing the model via parameter analysis, we converted the time scale to daily increments. By converting our time scale we showed that the model can be used to simulate a variety of simulations on differing time scales including days, weeks, months, and years. In Raggett and Sulsky's Eyam models, parameters were determined using 31 day increments as shown in Figure 3.3**A** and **B**. Because of the original scaling where one increment equals 31 days, to change the increment to 1 day, we divided all of our rate parameters by 31. This gave us new parameter values of $\beta = 0.00057$ (people $\times$ day)$^{-1}$ per individual and $\alpha = 0.09097$ day^{-1}. We note that the resulting graph in Figure 3.4, where the model is simulated for 500 days looks almost identical to Figure 3.3**A** which uses months as the increment. Because of the shorter increments between integration steps, Figure 3.4 appears smoother than 3.3**B**. However, the final values for S, I, and R compartments match across both time scales.

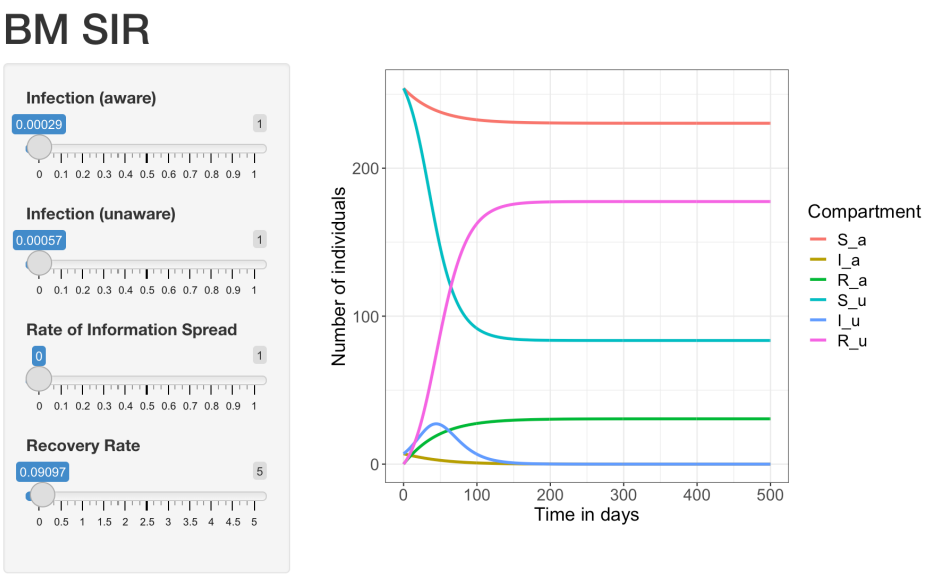

Figure 3.4: Behavioral SIR Model with Time in Days
Using modified Eyam parameters ($\beta = 0.00057$ (people $\times$ day)$^{-1}$ per individual and $\alpha = 0.09097$ day^{-1}).

Chapter 4

Results and Discussion

After developing the computer algorithms to solve our system, we began simulating cases with different β_a and β_u values. As described previously, we interpret $\beta_a > \beta_u$ as a situation where the information is harmful. Conversely, when $\beta_a < \beta_u$, we interpret the scenario as one where knowing the information is helpful.

In our analysis and for all of our simulations, the initial conditions remain fixed while we manipulate the specific parameter values. The initial conditions, in number of people, adopted from Ragget's Eyam model are as follows:

$$S_a(0) = S_u(0) = 254$$

$$I_a(0) = I_u(0) = 7$$

$$R_a(0) = R_u(0) = 0.$$

In our initial experiments we hold β_a, β_u, and α constant while δ is varied to better understand the impact of sharing information on disease transmission. The values for β_a, β_u, and α are provided in the description of each testing scenario.

4.1 The Effects of Good Information

Recall that good or helpful information is information that leads to behavior that decreases disease transmission amongst aware individuals. Prior to testing the model, we hypothesize that good information will help reduce the intensity of disease outbreaks. We define a reduction of disease intensity as when fewer individuals in the aware group are infected than the unaware. In our testing we use parameter values $\beta_u = 0.00057$ (people $\times$ day)$^{-1}$ and $\beta_a = \frac{1}{2}\beta_u$ (people $\times$ day)$^{-1}$ where β_u is the Eyam plague infection rate in days, and β_a is one half of that value. We use this relationship between β values to simulate an imagined plague that is half as transmissible as the one observed in Eyam. Choosing $\beta_a = \frac{1}{2}\beta_u$ (people $\times$ day)$^{-1}$ was an arbitrary choice and in later sections we explore choosing different relationships between β_a and β_u.

4.1.1 Disease Transmission Without Sharing Good Information

When there is no sharing of information, we observe what occurs when good information is present in one group but not shared between groups, and our model is reduced to two decoupled SIR models with no movement of individuals between them. For the unaware group the spread of disease, as shown in Figure 4.1, gives the same results as Sulsky's data in Figure 3.3**A**. The graphs are identical because there is no information being shared between the aware and unaware groups. In our unaware group we have a small initial outbreak, but no long-term disease transmission lasting greater than 365 days. Overall, in Figure 4.1 we note a large increase in the number of unaware infected individuals near the start of our simulation. Therefore, most people who became infected are from the unaware group. The outbreak peak for the unaware population was at day 44 when 27 individuals were infected. On the other hand, the aware population had its peak at day 0 when 7 individuals were infected. In total, 178 unaware individuals and 29 aware individuals were infected over the course of 500 days. Note that the disease was no longer transmissible after approximately

150 days when the total number of infected individuals across both aware and unaware groups reached 0. As hypothesized, the final number of susceptible, infected, and recovered individuals from the unaware population matched the final values from Sulsky and Raggett's Eyam model as there was no information being shared.

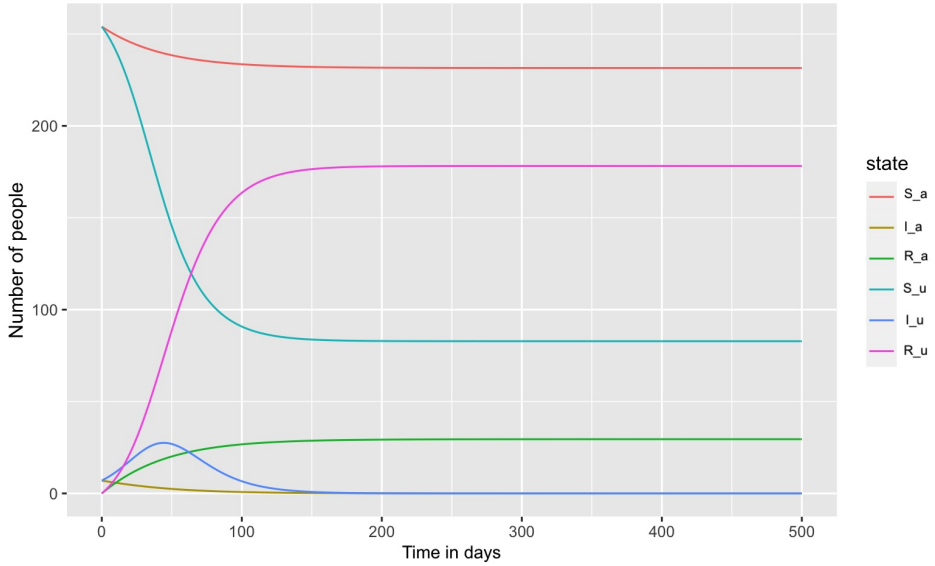

Figure 4.1: Spread of Disease with no Good Information being Shared
In this simulation we model the presence of good information where $\beta_a < \beta_u$ but information is not shared, hence $\delta = 0$.

4.1.2 Disease Transmission With Sharing Good Information

To test the effect of sharing information we introduce $\delta \neq 0$. The parameter δ represents the rate that individuals move from S_u to S_a as they share information and move from unaware to aware. Once again we set $\beta_u = 0.00057$ (people $\times$ day)$^{-1}$ and $\beta_a = \frac{1}{2}\beta_u$ (people $\times$ day)$^{-1}$ the results are shown in Figure 4.1. Overall, when good information is shared, most people who become infected are unaware as seen in the large increase of I_u individuals near the start of the simulation. However, people quickly moved from the unaware group to the aware group as shown in the rapid decrease of S_u. Note that some of the decrease in S_u can be attributed to individuals becoming infected. The outbreak peak for the unaware group shifted from day 44 in the previous simulation to day 14, representing the quick sharing of

information in the population. However, the outbreak peak for the aware group occurred at day 103. On day 14 the maximum number of infected individuals was 10 for the unaware group and on day 103 there were 27 infected individuals in the aware group. The delay in an aware group outbreak is likely due to a lag in disease infection as individuals first must move from the unaware to aware group then become infected with the disease. In total, 316 individuals were infected with the disease, 282 of which were from the aware group and 34 from the unaware group.

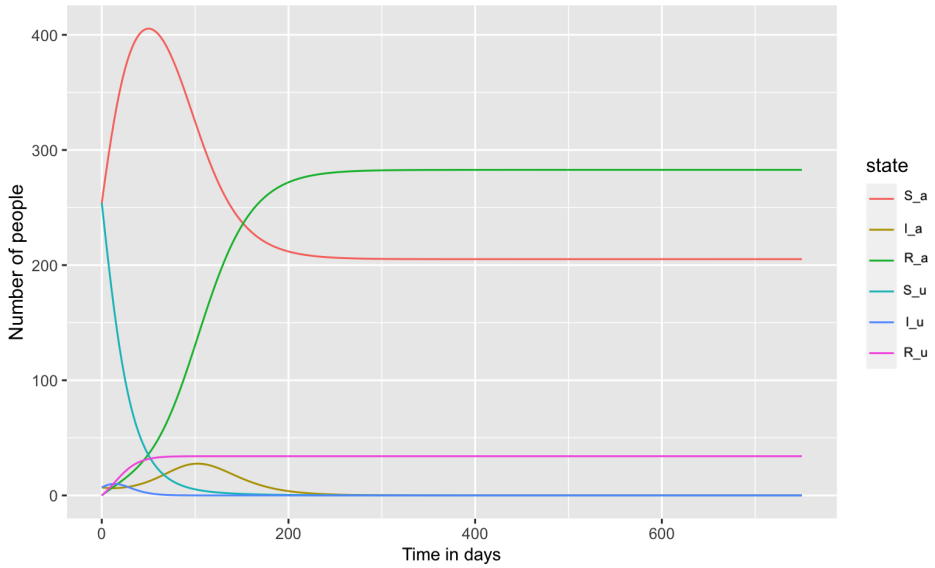

Figure 4.2: Spread of Disease with Good Information being Shared
In these simulations we model the presence of good information, $\beta_a < \beta_u$ and the sharing of information, $\delta = 0.0001$ (people $\times$ day)$^{-1}$.

4.1.3 Conclusions

The infection pattern with the sharing of good information versus the pattern without sharing information shows a dampened infection curve. While ultimately more people were infected when information transmission was introduced, the lag period (day 0 - 103) in infection might serve as a crucial time to increase public health measures and develop vaccine strategies. Although the introduction of good information was not enough to decrease the overall number of infected individuals, our results are still useful when considering the bi-

ological and epidemiological interpretation. Particularly, we see that the introduction of information results in an earlier disease outbreak amongst unaware individuals, emphasizing the importance of information campaigns in slowing the spread of a disease. Additionally, the delay in disease outbreak within the aware group could potentially be further time to spread useful information. In a later section we will further explore the idea of dampening the infection curve via β or infection rate analysis.

4.2 The Effects of Bad Information

Bad or harmful information, as defined previously is information that causes a specific behavior that increases the disease transmission rate amongst aware individuals such that $\beta_u < \beta_a$. We hypothesize that bad information will increase the number of overall infected individuals and for our testing we use $\beta_a = 0.00057$ (people $\times$ day)$^{-1}$ and $\beta_u = \frac{1}{2}\beta_a$ (people $\times$ day)$^{-1}$. In summary, this section aims to understand the influence of bad information on disease transmission.

4.2.1 Disease Transmission Without Sharing Bad Information

Again, when there is no sharing of information, our model is reduced to two decoupled SIR models as discussed in Section 4.1.1. As expected our results match the scenario with good information and no sharing of information, but with the aware and unaware groups swapped as shown in Figure 4.4. Thus, the unaware group has a smaller number of total infected individuals than the aware, which matches Raggett and Sulsky's Eyam values.

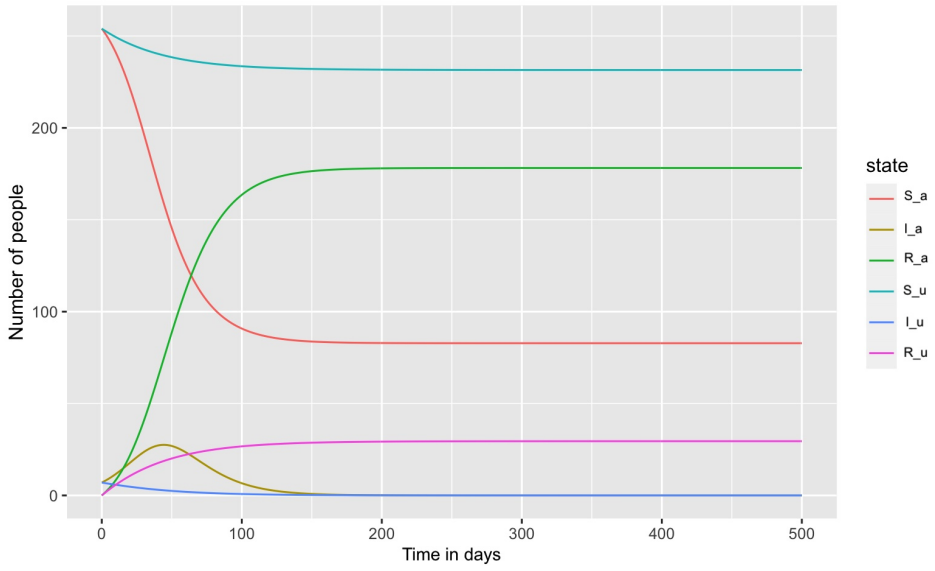

Figure 4.3: Spread of Disease with no Bad Information being Shared
In these simulations we model the presence of bad information, $\beta_a > \beta_u$ and the sharing of information, $\delta = 0$.

4.2.2 Disease Transmission With Sharing Bad Information

To determine the effect of sharing bad information on disease transmission, we again introduce $\delta \neq 0$. This allows individuals to flow from S_u to S_a as they share the bad information which increases disease transmission. To show this increase in transmission, we again set $\beta_a = 0.00057$ (people $\times$ day)$^{-1}$ and $\beta_u = \frac{1}{2}\beta_a$ (people $\times$ day)$^{-1}$.

When bad information is introduced, the unaware group infection peaks at day 0 while the aware group's infection peaks at day 42. In total, 419 individuals were infected, 405 from the aware group and 14 from the unaware group. Unlike the effect of good information, bad information not only spreads through the population quickly but also causes a more immediate outbreak compared to good information, which peaked at day 103.

4.2.3 Conclusions

Again translating our results from numerical values to biological and epidemiological concepts, when bad information exists and is shared, we observe a sharp outbreak of the disease

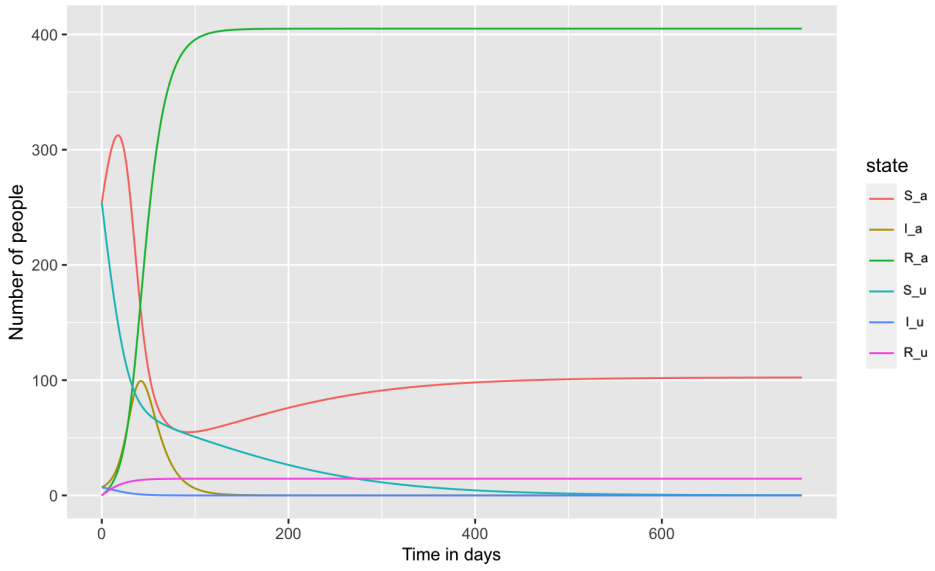

Figure 4.4: Spread of Disease with Bad Information being Shared
In these simulations we model the presence of bad information, $\beta_a > \beta_u$ and the sharing of information, $\delta = 0.0001$ (people $\times$ day)$^{-1}$.

within a short period of time. Comparisons of outbreak peaks from the four modeled scenarios are shown below in Table 4.1. Because of the lack of a lag period in the I_a curve at day 42 as seen in Figure 4.2, we deduce that bad information is far more harmful than good information is benefical. When bad information is introduced, the number of total infected individuals aware and unaware is over 100 more than the sum of infections in our good information simulation. Beyond the greater number of overall infected individuals, we also observe that the maximum number of individuals infected at the peak of the outbreak increased from 27 when good information is shared to 99 in the bad information simulation. In general, these simulations suggest that the effect of bad information is more pronounced than the benefit of good information. As we see in Table 4.1, bad information not only causes more people to become infected but also a faster outbreak of infection. These results suggest that mitigating bad information should be prioritized to decrease the outbreak magnitude caused by bad information. We explore the effect of good information more thoroughly below.

Group	Information Type	δ (people $\times$ day)$^{-1}$	Infection peak	Total infected
Aware	Good	0	day 0	29 people
Unaware	Good	0	day 44	178 people
Aware	Good	0.0001	day 103	282 people
Unaware	Good	0.0001	day 14	34 people
Aware	Bad	0	day 44	178 people
Unaware	Bad	0	day 0	29 people
Aware	Bad	0.0001	day 42	405 people
Unaware	Bad	0.0001	day 0	14 people

Table 4.1: Infection peak summary

4.3 Effects of Good Information on Infectiousness

As previously discussed even when good information is introduced in the population and information sharing can occur, we still observe an outbreak in the aware group. This outbreak leads to a greater total number of infected individuals ($I_a + I_u$) than if no information sharing occurred. In this section we aim to explore when good information is "good enough" such that it prevents an increase in the total number of aware individuals who become infected. To perform this experiment we look for the values of β_a that prevent an outbreak. That is, we find the values of β_a that give $R_a < R_u$.

Using the ShinySIR model we varied β_a, holding all other parameter values constant, until we observed visible a peak in infection. A visible peak in infection means that I_a has a visible increase followed by a decrease. We then used the β_a value that generated this plot in our model to determine final values for R_a and R_u. In doing so we found $R_a < R_u$ when $\beta_a \leq 0.00017$ (people $\times$ day)$^{-1}$. In comparison our base β_u value was 0.00057 (people $\times$ day)$^{-1}$. So β_a needs to be approximately $\frac{1}{3}\beta_u$ to prevent an outbreak in the aware group. Using $\beta_a = 0.00017$ (people $\times$ day)$^{-1}$ and $\beta_u = 0.00057$ (people $\times$ day)$^{-1}$, the final number of infected individuals for the aware and unaware groups were $R_a = 27$ and $R_u = 34$.

In our studies, good information, is a slight misnomer as the introduction of any information regardless of its type increases the total number of infected individuals as seen in Table 4.1 where when $\delta = 0$ the total number of infected individuals was 207 and when $\delta = 0.0001$

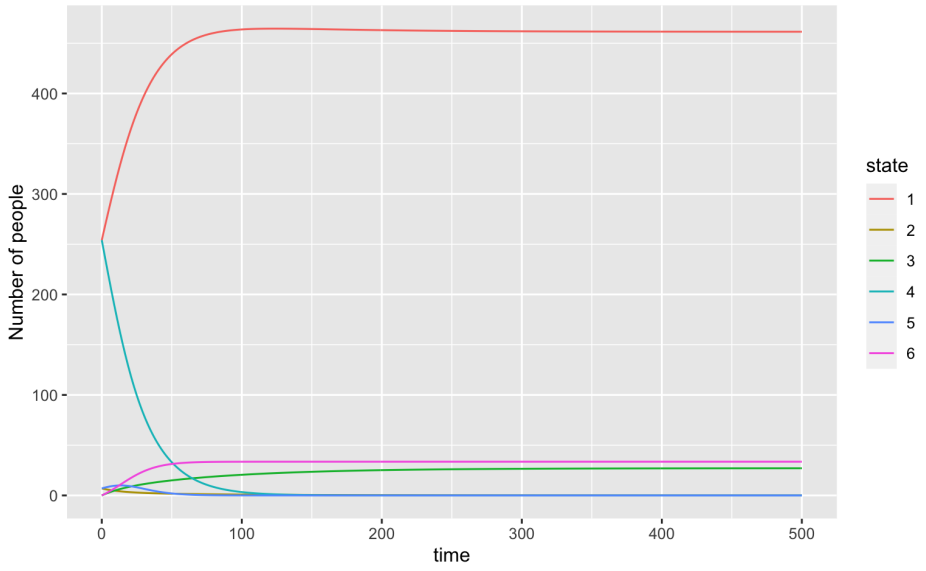

Figure 4.5: "Good enough" Information
In this simulation we model when good information is "good enough". We find that $\beta_a = 0.00017$ (people $\times$ day)$^{-1}$ and $\beta_u = 0.00057$ (people $\times$ day)$^{-1}$ for $R_a < R_u$.

the total number of infected individuals increased to 316 with the introduction of good information and 419 with bad information. However, when we further examine the impact of good information, we see that in fact, information simply needs to be "good enough" to stop an outbreak. Simply put, information that reduces the disease infection rate by $\frac{1}{2}$ does not do enough good to fully reduce the total number of infected individuals. Instead, we have shown that for our specific set of parameters, good information needs to reduce the infection rate by approximately $\frac{1}{3}$ to dampen the infection curve.

From a public health and epidemiological perspective, decreasing the infection rate of a disease is not impossible and can be done through the implementation of messaging, personal protective equipment, social distancing, among other strategies. Depending on the disease that is present within the population, different recommendations can be made. For example, in the recent COVID-19 pandemic, the Centers for Disease Control (CDC) recommended that individuals wear face masks to decrease aerosol spray and spread of COVID-19 [12]. The use of personal protective equipment can be incredibly useful in decreasing disease transmission, and this model provides a useful tool to understand the threshold of disease

transmission to stop an outbreak.

Chapter 5

Conclusion

In this thesis we have developed a six compartment model of disesae transmission and information sharing to analyze the effect of human behavior on disease transmission. Current mathematical models of disease often assume a homogeneous population with population parameters that apply to every individual regardless of their beliefs. Our findings suggest that separating the population by knowledge, shows that when good information exists, the information has to be "good enough" at lowering the disease transmission rate to prevent a severe outbreak. Conversely, bad information is incredibly powerful such that an outbreak is likely to occur when bad information is present and shared throughout the community. These findings indicate that while both good and bad information are essential to monitor in the context of disease management, greater attention should be focused on misinformation and harmful information that increases disease transmission.

5.1 Model Limitations

A key limitation in our analyses is that the parameters used within our experiments come from the 1666 Eyam second wave plague outbreak. Although the plague is still present in communities today, most notably Madagascar, public health measures and medical advancements have reduced the transmission of the disease. Instead, these parameters could be

altered specifically for another highly contagious disease that spreads quickly throughout a population.

An underlying assumption of the model is that when individuals are presented with information, they act on that information. However, this is not always the case as individuals will often be presented with both good and bad information and fail to act on the information they receive. Thus, a more appropriate interpretation of the scenarios modeled might be the spread of action or behavior rather than information.

5.2 Future Work

In the future to continue expanding this work, a model featuring separate recovery rate parameters for aware and unaware groups could be developed. This model could be used for scenarios where the information not only effects disease transmission but also recovery. As shown in Section 1.4.1 births and deaths can also be incorporated or other variations of the classic SIR model could also be considered.

Additionally, information sharing could be expanded beyond only susceptible individuals. Allowing infected and recovered individuals to share information would describe a scenario in which individuals are able to contact one another regardless of disease status. Combined with differing recovery parameters this would also allow for individuals who are in the unaware infected class to move to the aware recovered class if they became aware of the information while infected.

Appendix A

Behavior Model Code

Analyses from the behavior six compartment model are described in Chapter 4. For all tests the parameter initial conditions were changed as described and results were recorded. In some instances, interesting changes in the final resulting figure were further explored using the shinySIR model for faster changes in parameter values.

First, we define our model's variables and parameters.

```
DensityDepSIR <- function(t,x,params){
  # Args:
  # t: vector of time points in the integration
  # x: vector of variables in the differential equation
  # params: vector of parameter values
  #
  # Returns:
  # The rate of change of the variables S_a, I_a, R_a, S_u, I_u,
  # and R_u.
  # Where a = aware and u = unaware

  # Local Variables
```

```
S_a <- x[1]
I_a <- x[2]
R_a <- x[3]
S_u <- x[4]
I_u <- x[5]
R_u <- x[6]

# Local parameters
beta_a <- params[1]
beta_u <- params[2]
delta <- params[3]
alpha <- params[4]

with(
as.list(params),
{
  # Equations
  dS_a <- -beta_a * (S_a * I_a) + delta * (S_u * S_a)
  dI_a <- beta_a * (S_a * I_a) + (-alpha * I_a)
  dR_a <- alpha * I_a
  dS_u <- -beta_u * (S_u * I_u) - delta * (S_u * S_a)
  dI_u <- beta_u * (S_u * I_u) + (-alpha * I_u)
  dR_u <- alpha * I_u

  dx <- c(dS_a, dI_a, dR_a, dS_u, dI_u, dR_u)
  # combine results into a single vector dx
  list(dx) #return result as a list,
```

```
    }
    )
}
```

Then we initialize our variables including integration points, initial conditions, and parameter values.

```
# time points for the integration in days
tmin <- 0
tmax <- 500
t.SIR <-seq(tmin,tmax,by=1)

# Initial conditions for S,I,R aware and unaware
xstart<- as.numeric(c(S_a=254, I_a=7, R_a=0, S_u=254,
        I_u=7, R_u=0))

# parameter values
# transformed from Raggett paper, 1 time unit = 31 days
SIRparameters <- c(beta_a=(0.0002), beta_u=(((2.82/159))/31),
        delta=0.0001, alpha=(2.82/31))
```

Finally, we run our model. This code outputs a figure displaying the change in population size per compartment. Additionally, a table comprising the final values and maximum value for each compartment is generated.

```
DensityDepSIR(t.sir, xstart, SIRparameters)

# simulate an outbreak
SIRsim <- as.data.frame(lsoda(y = xstart, times = t.SIR,
        func = DensityDepSIR, parms = SIRparameters))
```

```
# Plot SIR
 SIRsim %>%
    gather(key = state, value = value, "1":"6") %>%
     ggplot(data = ., aes(time, value, color = state)) +
     geom_line() + labs(y = "Number of people")


#create a table with final values
labels <- c("S_a", "I_a", "R_a", "S_u", "I_u", "R_u")
vals <- c(SIRsim[tmax,2], SIRsim[tmax,3], SIRsim[tmax,4],
          SIRsim[tmax,5], SIRsim[tmax,6], SIRsim[tmax,7])
final_vals <- round(vals, 2)
max_vals <- c(max(SIRsim[,2]), max(SIRsim[,3]), max(SIRsim[,4]),
          max(SIRsim[,5]), max(SIRsim[,6]), max(SIRsim[,7]))
final_vals <- data.frame(labels, final_vals, max_vals)
print(final_vals)
```

Appendix B

shinySIR Model Code

To generate a shinySIR output with sliders for each variable, we first define a new function, mySIRS containing our necessary equations.

```
mySIRS <- function(t, y, parms) {

  with(as.list(c(y, parms)),{

  # Equations
   dS_a <- -beta_a * (S_a * I_a) + delta * (S_u * S_a)
   dI_a <- beta_a * (S_a * I_a) + (-alpha * I_a)
   dR_a <- alpha * I_a
   dS_u <- -beta_u * (S_u * I_u) - delta * (S_u * S_a)
   dI_u <- beta_u * (S_u * I_u) + (-alpha * I_u)
   dR_u <- alpha * I_u


  return(list(c(dS_a, dI_a, dR_a, dS_u, dI_u, dR_u)))
  })
```

```
}
```

We then initialize all of our variables and parameters. Additionally, we name all of our variables.

```
# t=0 starting values
xstart<- c(S_a=254, I_a=7, R_a=0, S_u=254, I_u=7, R_u=0)


SIRparameters <- c(beta_a=0, beta_u=2.82/159, delta=0.000,
alpha=(2.82))
parameter_names <- c("Infection (aware)", "Infection (unaware)",
    "Rate of Information Spread", "Recovery Rate"
```

Finally, we create the shinySIR model that is outputted from the following code chunk.

```
run_shiny(model = "BM SIR",
            neweqns = mySIRS,
            ics = xstart,
            timestep = 1,
            tmax = 10,
            parm0 = SIRparameters,
            parm_names = parameter_names,
            xlabel = "Time in 31 day increments",
            parm_min = c(beta_a=0, beta_u=0, delta=0, alpha=0),
            parm_max = c(beta_a=1, beta_u=1, delta=1, alpha=5),
            slider_steps = c(beta_a=0.00001, beta_u=0.00001,
                    delta=0.00001, alpha=0.00001))
```

Bibliography

[1] W. O. Kermack and A. G. McKendrick. "A contribution to the mathematical theory of epidemics". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 115.772 (1927), 700–721. DOI: 10.1098/rspa.1927.0118.

[2] Wagner B. G. Coburn B. J. and S. Blower. "Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1)". In: *BMC Med* 7.30 (2009), pp. 1–8. DOI: https://doi.org/10.1186/1741-7015-7-30.

[3] G. F. Raggett. "Modeling the Eyam Plague". In: *Bulletin of the Institute of. Mathematics and its Applications* 18 (1982), pp. 221–226. DOI: 10.1080/02664768200000021.

[4] O. N. Bjørnstad et al. "The SEIRS model for infectious disease dynamics". In: *Nature Methods* 17.6 (2020), 557–558. DOI: 10.1038/s41592-020-0856-2.

[5] J. F. Sallis, N. Owen, and M. J. Fotheringham. "Behavioral epidemiology: A systematic framework to classify phases of research on Health Promotion and Disease Prevention". In: *Annals of Behavioral Medicine* 22.4 (2000), 294–298. DOI: 10.1007/bf02895665.

[6] Gilad E. Funk S. and V.A.A. Jansen. "Endemic disease, awareness, and local behavioural response". In: *Journal of Theoretical Biology* 264.2 (2010), pp. 501–509. ISSN: 0022-5193. DOI: https://doi.org/10.1016/j.jtbi.2010.02.032.

[7] H. Wickham et al. "Welcome to the tidyverse". In: *J. Open Source Softw.* 4.43 (Nov. 2019), p. 1686. ISSN: 2475-9066. DOI: 10.21105/joss.01686. URL: http://dx.doi.org/10.21105/joss.01686.

[8] Signorell A. et mult. al. *DescTools: Tools for Descriptive Statistics*. R package version 0.99.44. 2021. URL: https://cran.r-project.org/package=DescTools.

[9] K. Soetaert, T. Petzoldt, and R. Woodrow Setzer. *deSolve: General solvers for initial value problems of ordinary differential equations (ODE), partial differential equations (PDE), differential algebraic equations (DAE) and delay differential equations (DDE)*. R package version 1.13. 2016. URL: http://cran.r-project.org/web/packages/deSolve/deSolve.pdf.

[10] S. E. Morris. *shinySIR: Interactive Plotting for Mathematical Models of Infectious Disease Spread*. R package version 0.1.2. 2020. URL: https://github.com/SineadMorris/shinySIR.

[11] D. Sulsky. "Using Real Data in an SIR Model". In: *Unpublished Report* (2012), pp. 1–13.

[12] K. L. Andrejko et al. "Effectiveness of face mask or respirator use in indoor public settings for prevention of SARS-COV-2 infection — California, February–December 2021". In: *MMWR. Morbidity and Mortality Weekly Report* 71.6 (2022), 212–216. DOI: 10.15585/mmwr.mm7106e1.