

## Modeling the Effect of Human Behavior on Disease Transmission

Katie Yan

Skidmore College, [kyan@skidmore.edu](mailto:kyan@skidmore.edu)

Follow this and additional works at: <https://scholar.rose-hulman.edu/rhumj>



Part of the [Disease Modeling Commons](#), and the [Health Communication Commons](#)

---

### Recommended Citation

Yan, Katie (2024) "Modeling the Effect of Human Behavior on Disease Transmission," *Rose-Hulman Undergraduate Mathematics Journal*: Vol. 25: Iss. 1, Article 10.

Available at: <https://scholar.rose-hulman.edu/rhumj/vol25/iss1/10>

## Modeling the Effect of Human Behavior on Disease Transmission

---

By *Katie Yan*

**Abstract.** Many infectious disease models build upon the classic Susceptible-Infected-Recovered (SIR) model, a compartmental system that is used to simulate disease transmission in a population. The SIR model focuses on the transmission of disease but rarely includes behavioral or informational components that explore how disease perception influences transmission. In this paper, we propose a six-compartment behavioral SIR model that further segments the classic SIR system based on knowledge of information about the disease, and we explore how sharing information affects disease transmission. We designate two states as *aware* and *unaware* based on whether the relevant information is known by the population. Additionally, we include two types of information: good information that reduces transmission rates and bad information that increases transmission rates. We find that while compliance with good information is useful in decreasing community transmission, compliance with bad information has a greater magnitude of effect in terms of total cases. These results reaffirm that knowledge and human behavior are influential factors in disease transmission and should be included in future human disease models for more accurate transmission representation.

### 1 Introduction

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. They range in complexity from simple algebraic equations to complex systems of ordinary differential equations (ODEs) and are used to predict outcomes under a certain set of assumptions and conditions.

Many mathematical models focus on the spread of infectious diseases. These diseases are caused by infectious agents including viruses, bacteria, and protozoa. 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 spread of disease, we can predict hot-spot locations for the disease, transmission rates, and effective prevention and protection measures.

---

*Mathematics Subject Classification.* 92B05

*Keywords.* Mathematical model, behavioral epidemiology, disease transmission, information sharing

Although many robust mathematical models exist, they often fail to incorporate human behavior. Recently, the National Science Foundation released a “Dear Colleague Letter” prompting proposals that research how human behavior could be incorporated into disease models [1]. As shown throughout the COVID-19 pandemic, the way humans behave is essential to understanding the spread of disease and how best to decrease transmission rates.

In this paper we propose a behavioral model of disease, incorporating the idea of “good” and “bad” information that influences the transmissibility of a disease. We begin by giving an overview of a simple Susceptible-Infected-Recovered (SIR) model and behavioral epidemiology in Section 2. In Section 3 we define our behavioral SIR model which we show in both compartmental and ODE form. We explain simulation results from this model in Section 4 and discuss our results and conclusions in Section 5. Finally, we give a table of variables and compartments in Appendix A, and additional simulation results in Appendix B

## 2 Background

### 2.1 The SIR Model

The SIR compartmental model is one of the most common models used to describe the transmission of a disease within a population. The SIR model is described by the compartmental model shown in Figure 1 [2]. 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 carry the disease and spread it to susceptible individuals. Finally, recovered ( $R$ ) individuals have passed the infectious stage of the disease and no longer transmit the disease. Individuals in the population will move from  $S$  to  $I$  and finally to  $R$  as the disease progresses.

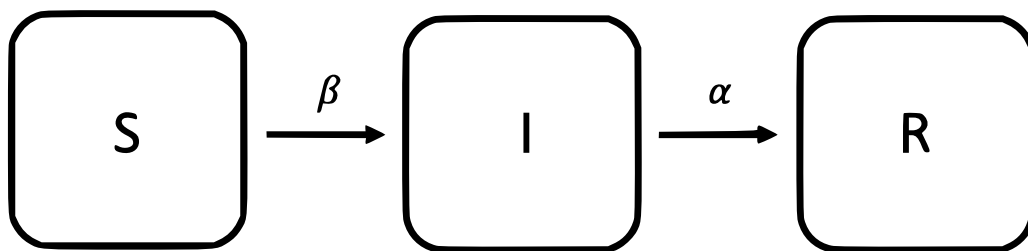


Figure 1: SIR Compartmental Model: In this model, individuals are separated based on infection status. As the disease progresses individuals become infected at rate  $\beta$  and recover at rate  $\alpha$ .

In the SIR compartmental model, people in the susceptible compartment move at

rate  $\beta$  to the infected compartment. Likewise, the recovery rate  $\alpha$  describes the rate at which individuals recover from the disease and are no longer infectious. Several simplifying assumptions are made to the model. These assumptions include:

1. The total population size  $N$ , where  $N = S + I + R$ , is constant. There are no births or deaths.
2. After someone recovers from the disease, they are considered immune to the disease. They cannot reenter the  $S$  compartment.
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 an 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.
6. The recovery rate  $\alpha$  is proportional to the number of infected individuals at time  $t$ .

The SIR compartmental model can be written as a series of ODEs to describe the rate of change of each compartment given by the System of Equations (1) and thus to describe the movement of individuals from one compartment to another. The SIR model as described in (1) has two parameters  $\alpha$  and  $\beta$ . In the model,  $\beta$  has units of one over time per individual, and  $\alpha$  has units of one over time.

$$\begin{aligned}\dot{S} &= -\beta SI \\ \dot{I} &= \beta SI - \alpha I \\ \dot{R} &= \alpha I\end{aligned}\tag{1}$$

We assume, at  $t = 0$ , the population size is a non-negative value, the number of susceptible and infected individuals is non-zero, and the number of recovered individuals is 0. Therefore, the initial conditions for our model are  $S(0) = S_0 > 0$ ,  $I(0) = I_0 > 0$ , and  $R(0) = 0$ .

## 2.2 Behavioral Epidemiology

Behavioral epidemiology is the study of human behavior, its influences, and its impact on the spread of infectious diseases. Developed in the 1970s, behavioral epidemiology

is an interdisciplinary field that draws from mathematics, biology, sociology, psychology, economics, and other disciplines that study human behavior and disease. Modern interpretations of the phrase “behavioral epidemiology” have shifted to a more mathematical and modeling-centered perspective. Specifically, the modern interpretation of behavioral epidemiology aims to understand the effects of determinants on human behavior.

In a classic SIR model and its variations, we operate under the assumption that people within our population of interest behave homogeneously. This means there are no outliers in terms of human behavior, and we expect an even mixing of susceptible, infected, and recovered individuals. However, this is rarely the case in real-life disease scenarios, where individuals will react and change their behavior based on various influences they experience.

In this paper, we use the schema described by behavioral epidemiology to improve the SIR model formulation as we aim to understand the influence of both helpful and harmful information. Specifically, we are motivated to explore the compliance and non-compliance of individuals in following public health guidelines and recommendations throughout the COVID-19 pandemic.

### 3 Behavioral Mathematical Models

In this section, we explore the relationship between disease transmission and information sharing. To model this relationship we propose a six-compartment mathematical model and corresponding system of equations similar to Funk et al. in [3].

#### 3.1 The Behavioral Model of Disease

As seen in Figure 2, this model expands on the basic SIR model depicted in Figure 1 by subdividing each compartment into aware ( $a$ ) and unaware ( $u$ ) groups depending on if the individuals within each compartment are aware or unaware of the information.

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 susceptible individuals moving from the unaware group to the aware group ( $S_u$  to  $S_a$ ). As in the original SIR model, 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 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  $\beta_a$  and  $\beta_u$ . 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 the information being spread is harmful,  $\beta_a > \beta_u$ . Examples of

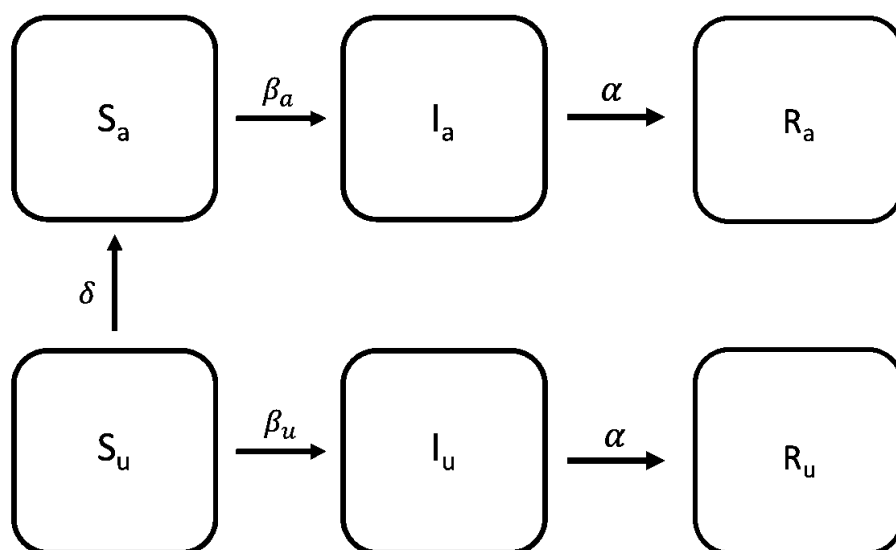


Figure 2: Behavioral SIR Six Compartment Model: Knowledge of information is denoted by subscript  $u$  (unaware) and  $a$  (aware) with different infection rates ( $\beta_u$  or  $\beta_a$ ) for each group. Note that the recovery rate,  $\alpha$ , is the same for both the unaware and aware groups. Information is shared across susceptible groups at rate  $\delta$ .

harmful information include misinformation about the disease such as suggestions that the disease is not real or recommendations to not comply with public health measures.

Similar to the classic SIR model, in the behavioral SIR model after an individual is infected, they recover at rate  $\alpha$ , regardless if they know the information or if the information is helpful or harmful because we model the effects of information on transmission, not recovery. As before we assume the birth and death rates of the whole population are negligible. Finally, the most important difference between our described model and the classic SIR model is the addition of an information-sharing rate  $\delta$ . The parameter  $\delta$  represents the flow of information from the aware population to the unaware group. 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. We assume that information can only be shared between susceptible people since we assume the information only affects disease transmission and not disease recovery. All parameters and their meanings are shown in Table 1.

Using the compartmental model described and shown in Figure 2, we write a system

Table 1: Parameter definitions for Behavioral SIR Model: In this model, we subdivide  $\beta$  into  $\beta_a$  and  $\beta_u$  to allow for the differences in infection rates dependent on knowledge of information. We maintain a population-level  $\alpha$  parameter such that information does not influence one's recovery from the disease. Finally, we include  $\delta$  as our information-sharing rate.

| Parameter | Meaning                               |
|-----------|---------------------------------------|
| $\beta_a$ | Infection rate for aware population   |
| $\beta_u$ | Infection rate for unaware population |
| $\alpha$  | Population recovery rate              |
| $\delta$  | Information sharing rate              |

of six ODEs describing the sharing of information and the spread of disease through the population. Note System (2) essentially combines two SIR models of form (1), one for aware ( $a$ ) and one for unaware ( $u$ ). The only movement between these two models is through the addition of the term  $\delta(S_a S_u)$  to describe the change in aware and unaware susceptible individuals.

$$\begin{aligned}
 \dot{S}_a &= -\beta_a(S_a I_a) + \delta(S_a S_u) \\
 \dot{I}_a &= \beta_a(S_a I_a) - \alpha I_a \\
 \dot{R}_a &= \alpha I_a \\
 \dot{S}_u &= -\beta_u(S_u I_u) - \delta(S_a S_u) \\
 \dot{I}_u &= \beta_u(S_u I_u) - \alpha I_u \\
 \dot{R}_u &= \alpha I_u
 \end{aligned} \tag{2}$$

Like Funk et al.'s six-compartment model, in System of Equations (2) we see that information is shared causing individuals to move from the unaware to the aware susceptible compartment. Akin to the classic SIR model we assume a homogeneously mixed population within the aware and unaware groups.

In our behavioral SIR model, System of Equations 2, we make several simplifying assumptions. These assumptions include:

1. The total population size  $N$ , where  $N = S_a + I_a + R_a + S_u + I_u + R_u$ , is constant. There are no births or deaths.
2. After someone recovers from the disease, they are considered immune to the disease. They cannot reenter the S compartment.
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 for each group, *aware* and *unaware*, are uniformly mixed to ensure an 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.
6. The recovery rate  $\alpha$  is proportional to the number of infected individuals at time  $t$  and is the same across groups.
7. Transmission of the disease from infected to susceptible individuals depends on the group.
8. When information is shared individuals in the susceptible unaware class,  $S_u$ , can move to the susceptible aware class,  $S_a$ , at rate  $\delta$ .

### 3.2 RStudio Model

To simulate epidemic scenarios using System of Equations (2), we use R along with the tidyverse, DescTools, and deSolve packages [4, 5, 6, 7]. We utilize RStudio to investigate the effects of varying parameter values and to determine the final and maximum values population for each compartment. Additionally, the shinySIR package is used to vary the values of  $\beta_a$ ,  $\beta_u$ , and  $\delta$  [8].

### 3.3 Methods

To explore the relationship between human behavior and disease transmission, we solve System (2) numerically with a range of  $\beta_a$  and  $\beta_u$  values. As described previously, we interpret  $\beta_a > \beta_u$  as a situation where the information is harmful and a situation where  $\beta_a < \beta_u$  as one where the information is helpful.

Our initial conditions and parameters replicate G. F. Raggett's 1921 paper, modeling the plague in Eyam, England. The plague in Eyam began in 1665 and within 2 years, 267 of Eyam's 350 residents died from the plague. However, the death toll could have been much higher had the village rector, William Mompesson, not decided to cease all travel in and out of Eyam to protect the surrounding villages. In the summer months, after the initial wave of the plague, Rector Momepesson gathered Eyam's remaining residents to propose a quarantine. In the fall, the plague returned causing a second wave of infections. This wave lasted until November of 1666 when the last death was recorded. Using Rector Mompesson's historical records of plague-related deaths, Raggett created a SIR model of the second wave of infection, which occurred in the late summer to early fall. Using these recorded values, 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 and recovery, which he found to be  $\beta = \frac{2.82}{159} = 0.0177$  (people  $\times$  months) $^{-1}$  and  $\alpha = 2.82$  months $^{-1}$ , respectively.

In our analysis and for all of our simulations, the initial populations remain fixed while we manipulate the rate parameter values. The initial conditions, in number of people, adopted from Raggett's Eyam model [9], are as follows:

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

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

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

We follow the same process for each of our experiments. First, we hold  $\beta_a$ ,  $\beta_u$ , and  $\alpha$  constant and set  $\delta = 0$  to simulate a subdivided population that contains two groups, aware and unaware, that do not share information between them. Next, we hold  $\beta_a$ ,  $\beta_u$ , and  $\alpha$  constant, but we set  $\delta = 0.0001$  to simulate a population with two groups that communicate between themselves, allowing information to spread. We repeat this process for both good and bad information. Then, we complete a sensitivity analysis for  $\beta_a$  and  $\beta_u$  to determine for what ratios an outbreak occurs. The parameter values for each scenario are provided in the description of each model simulation.

For each of our simulations, we also calculate an approximate  $R_0$ ,  $\tilde{R}_0$  for the aware and unaware groups over time.  $R_0$  is classically defined as the number of secondary infections from one initial infection in a fully naive population [10]. However, since we are drawing from Raggett's plague simulations, we begin with 7 infected individuals in an otherwise susceptible population. Thus, we use an approximate value  $\tilde{R}_0$ , which is calculated by

$$\tilde{R}_0(t) \approx \frac{\beta_i(S_i(t) + I_i(t) + R_i(t))}{\alpha}, \text{ where } i = a \text{ or } u \quad (3)$$

As we monitor the  $\tilde{R}_0$  over time, we also consider whether the  $\tilde{R}_0$  for each group exceeds 1. For an infection to spread successfully in a population,  $R_0 > 1$  such that for each infected individual in a fully susceptible population more than one individual is infected by the originally infected individual. When  $R_0 < 1$  an infection is unable to take hold in a population and will not sustain a long chain of transmission. We also monitor if the  $\tilde{R}_0$  value for each group, aware and unaware, has a crossing point such that one  $\tilde{R}_0$  becomes larger than the other.

## 4 Model Simulation Results

In this section, we describe our numerical simulation and results for each of our experiments. Our experimental design is below in Table 2.

Table 2: Experimental design: Our  $2 \times 2$  experimental design varies the type of information (good or bad) and whether it is shared across groups (aware or unaware).

|                        |                               | Type of Information          |                             |
|------------------------|-------------------------------|------------------------------|-----------------------------|
|                        |                               | Good ( $\beta_a < \beta_u$ ) | Bad ( $\beta_a > \beta_u$ ) |
| Sharing of Information | Not Sharing ( $\delta = 0$ )  | Scenario A                   | Scenario B                  |
|                        | Sharing ( $\delta = 0.0001$ ) | Scenario C                   | Scenario D                  |

We begin with exploring the null model results when  $\beta_a = \beta_u$  and we consider when  $\delta = 0$  and  $\delta = 0.0001$ . Then, we introduce good information,  $\beta_a < \beta_u$ , scenarios A and C. We consider the introduction of bad information,  $\beta_a > \beta_u$ , scenarios B and D. When we introduce good and bad information we also conduct a sensitivity analysis, varying the ratio between  $\beta_a$  and  $\beta_u$  from  $1:\frac{2}{3}$  to  $1:\frac{1}{4}$ . In total, 18 unique experiments were conducted. We define the 1:1 scenario as the null scenario, with or without information sharing, where unaware and aware groups transmit at equal rates. Our full results can be found in Appendix B. Here we highlight our main findings.

#### 4.1 The Null Model

When there is no information sharing,  $\delta = 0$ , we observe two decoupled SIR models with no movement of individuals between the aware and unaware groups. In Figure 3 notice that the aware and unaware trajectories are indistinguishable as they are the same SIR system repeated twice. However, when information is shared across groups, even with  $\beta_a = \beta_u$  we notice that the final value of  $R_a$  is much larger than  $R_u$ . The increase in the final  $R_a$  value when  $\delta$  is increased to 0.0001 makes sense as unaware susceptible individuals have three possible paths as time progresses. Either they will become infected or not, or they will move to the susceptible aware class and then potentially become infected. Thus, because individuals cannot move from the aware group to the unaware group, an increase in the final  $R_a$  number is logical.

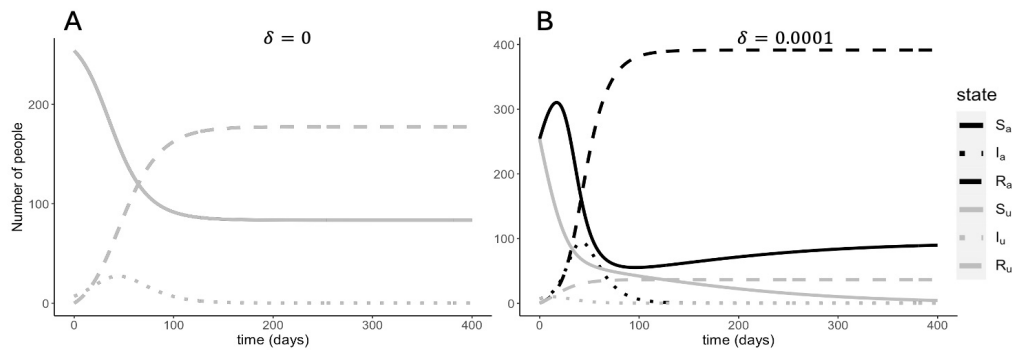


Figure 3: Spread of disease in the null model. In this simulation, we model when  $\beta_a = \beta_u$  and  $\delta$  is equal to 0 (3A) or 0.0001 (3B). Note that when  $\delta = 0$  both the unaware group and aware group produce the same results.

## 4.2 The Effects of Good Information

Recall that good or helpful information is information that leads to behavior that decreases disease transmission amongst aware individuals, such that  $\beta_a < \beta_u$ . Before testing the model, we hypothesize that all levels of 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 recovered than in the unaware population. That is, the final  $R_a < R_u$ .

In our sensitivity analysis, we vary the  $\beta_a$  and  $\beta_u$  ratio from  $1:\frac{2}{3}$  to  $1:\frac{1}{4}$ . See results in Figure 4. We use this relationship between  $\beta$  values to simulate an imagined plague that is as transmissible as the one observed in Eyam, and where the knowledge of good information reduces the infectiousness of the simulated plague.

When  $\beta_a = \frac{2}{3}\beta_u$ , Figure 4A, we see that the final number of infected individuals in the aware group is large compared to when  $\beta_a = \frac{1}{3}\beta_u$ , Figure 4C. We find a boundary for which an epidemic will not occur between  $\beta_a = \frac{1}{2}\beta_u$  and  $\beta_a = \frac{1}{3}\beta_u$ . In general, as the ratio between  $\beta_a$  and  $\beta_u$  changes from  $\frac{2}{3}$  of  $\beta_u$  to  $\frac{1}{3}$  of  $\beta_u$ , we find that  $\tilde{R}_0$  for the aware group does not exceed 1. More simply, there is such a thing as shared information that is "good enough" to stop an epidemic from occurring.

## 4.3 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  $\beta_a > \beta_u$ . We hypothesize that as bad information "worsens" we will observe an increase in the number of overall infected individuals.

For our testing, we assume unaware individuals transmit the disease at a smaller

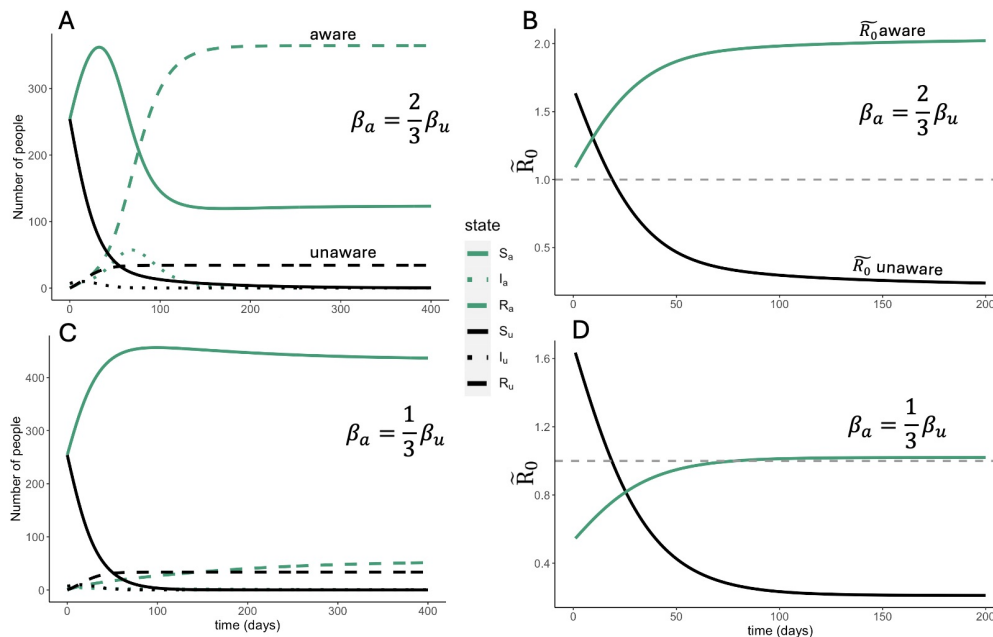


Figure 4: Spread of disease when good information is shared. In this simulation, we model when  $\beta_a < \beta_u$  and  $\delta$  is equal to 0.0001. An epidemic is modeled in 4A and 4B when  $\beta_a = \frac{2}{3}\beta_u$ , but not in 4C and 4D when  $\beta_a = \frac{1}{3}\beta_u$

rate than the aware. These values simulate scenarios where knowing the information is related to infecting others at an increased rate compared to not knowing the information. Here, we hold  $\beta_a$  constant, with the same value as in the Eyam model. In summary, this section aims to understand the influence of information that increases the infection rate on disease transmission.

When bad information is introduced and shared, the number of infected individuals increases slightly as the ratio between  $\beta_a$  and  $\beta_u$  changes from  $\frac{2}{3}$  of  $\beta_a$  to  $\frac{1}{3}$  of  $\beta_a$ . Both outbreaks are of a similar magnitude. However, as we change the ratio between  $\beta_a$  and  $\beta_u$  the  $\tilde{R}_0$  value for the unaware group decreases below 1, the epidemic threshold. In general, when bad information is shared there is little change between the epidemic curves, as shown in Figure 5. We also see this lack of change reflected in the  $\tilde{R}_0$  values over time.

## 5 Discussion and Conclusions

As discussed in the previous subsections, 4.2 and 4.3, the effects of information depend on the type of information being shared — if that information is good and how "good" that information is. Across all simulations, sharing good information will result in a

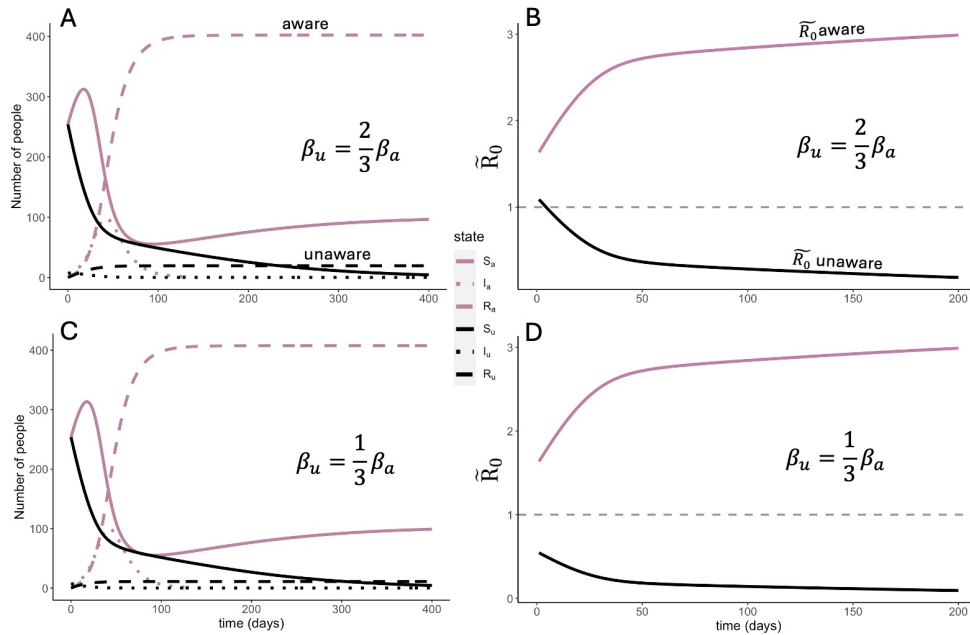


Figure 5: Spread of disease when bad information is shared. In this simulation, we model when  $\beta_a > \beta_u$  and  $\delta$  is equal to 0.0001. An epidemic occurs in all bad information scenarios.

smaller amount of infected individuals compared to sharing bad information. However, until  $\beta_a = \frac{1}{3}\beta_u$  we find that a decoupled SIR system, such that information is not shared, outperforms the sharing of good information when considering the total number of infected individuals.

Figure 6 shows the variation in the total number of infected individuals as the infection rate ratio is increased. When no information or good information is shared, the total number of infected individuals decreases as the ratio decreases. However, when bad information is shared the total number of infected individuals only slightly increases. Comparing the effects of sharing two types of information for a variety of transmission rates, not sharing results in a total of 188 to 354 infected individuals, sharing good information results in a total of 49 to 427 infected individuals, and sharing bad information results in a total of 417 to 427 infected individuals. In general, regardless of information type, as the ratio moves towards 1, which means as the  $\beta_a$  and  $\beta_u$  values approach each other, the total number of infected individuals also increases.

The inclusion of information in the classic SIR model system introduces more complex dynamics, especially when one considers the ability to share and not share information across groups. When information exists, but is not shared, we see a decoupled SIR system that can potentially reduce the total number of infected individuals by 166

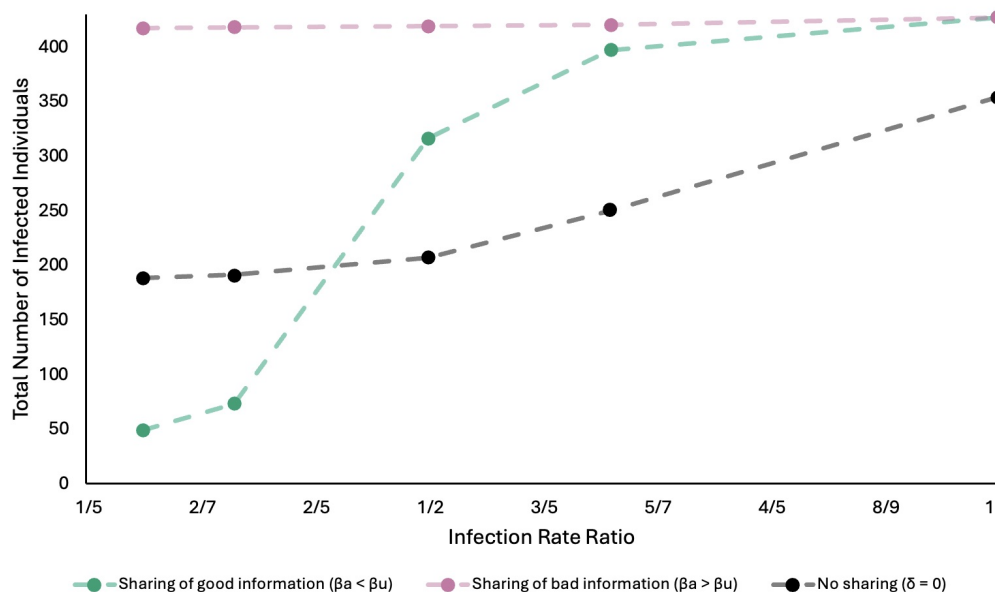


Figure 6: Total Number of Infected Individuals for Differing Infection Rate Ratio.

individuals, a nearly 50% reduction. However, when we qualify information as good or bad we see further nuances in the ways that information can potentially avert infections. Good information is always better than bad information at averting infections. The introduction of shared good information can reduce infections by about 90% while sharing bad information only reduces infections by about 2%.

It should be noted, however, that until  $\beta_a < \frac{1}{2}\beta_u$  the sharing of information, good or bad, produces more infections than when information is not shared across groups. Even in the worst-case scenario,  $\beta_a = \beta_u$ , the total number of infections when no information is shared is 354 compared to 427 infections when good or bad information is shared.

Therefore, from a decision-making framework, if the ability exists to separate those who do and do not know the information, and decouple the SIR system, then when  $\beta_a$  is reduced by  $\frac{1}{2}$  a coupled system is preferred.

A potential example of such a scenario would be when intervention behaviors, such as masking, are tied to information. Imagine that when an individual hears information about masking, they choose to begin masking which reduces their transmission potential from  $\beta_u$  to  $\beta_a$ . If masking is so powerful that it reduces  $\beta_a < \frac{1}{2}\beta_u$ , then a decision maker should feel comfortable in allowing the masked (aware) and unmasked (unaware) groups to mix. However, if masking is not as powerful such that  $\beta_a > \frac{1}{2}\beta_u$ , then susceptible individuals should not be allowed to intermix.

Although the introduction of good information was not always enough to decrease the overall number of infected individuals, Figure 6, our results are still useful when

considering biological and epidemiological interpretations. Particularly, we see that the introduction of good information results in an earlier disease outbreak amongst unaware individuals, emphasizing the importance of information campaigns in slowing the spread of disease. Additionally, the delay in disease outbreaks within the aware group could potentially be further time to share useful information.

When bad information is introduced, the number of total infected individuals aware and unaware is always greater than unshared and shared good information. In general, these simulations suggest that the negative effect of bad information is potentially more pronounced than the benefit of good information and that in an imagined scenario with only bad information, a decision-maker should prioritize stopping the spread of such information.

Our work can also be broadly applied to all disease-mitigation strategies, where we consider the *aware* and *unaware* groups akin to *compliant* and *non-compliant*. Using this framework we can apply our behavioral model to masking, social distancing, and other mitigation methods. Note that parameters for the specific type of mitigation strategy should be calculated and used to fully consider each mitigation method. However, in a purely theoretical system, this model serves as a useful tool to consider the ideal power of mitigation against disease transmission.

In this model, we segregate individuals depending on whether they are *aware* or *unaware*. In doing so we model a population where infected and recovered individuals of different groups do not intermingle. This scenario is representative of when individuals with differing beliefs or knowledge do not mix such that individuals who comply with public health measures only socialize with others who are compliant. Likewise, those who are non-compliant only socialize with those who are also non-compliant. 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 to investigate the spread of information that impacts recovery.

In this paper, we present a six-compartment model of disease transmission and information sharing to analyze the effect of human behavior on disease transmission. Our findings suggest that when separating a population by knowledge of information, good information is not always "good enough" such that there is a boundary between when an outbreak will and will not occur. Conversely, bad information is incredibly powerful such that in our simulations an outbreak will always occur when bad information is present and shared throughout the population. These findings indicate that while both good and bad information are essential to monitor in the context of disease management and decision-making, greater attention should be focused on misinformation and harmful information that increases disease transmission.

## A Variable and Compartment List

Table 3: An index of the variables and compartments used in this manuscript

| Symbol        | Type             | Meaning                         |
|---------------|------------------|---------------------------------|
| $S$           | Compartment      | Susceptible individuals         |
| $I$           | Compartment      | Infected individuals            |
| $R$           | Compartment      | Recovered individuals           |
| $N$           | Population Value | Total population size           |
| $\tilde{R}_0$ | Population Value | Approximate $R_0$               |
| $\beta$       | Parameter        | Infection rate                  |
| $\alpha$      | Parameter        | Recovery rate                   |
| $S_a$         | Compartment      | Susceptible aware individuals   |
| $I_a$         | Compartment      | Infected aware individuals      |
| $R_a$         | Compartment      | Recovered aware individuals     |
| $S_u$         | Compartment      | Susceptible unaware individuals |
| $I_u$         | Compartment      | Infected unaware individuals    |
| $R_u$         | Compartment      | Recovered unaware individuals   |
| $\beta_a$     | Parameter        | Aware infection rate            |
| $\beta_u$     | Parameter        | Unaware infection               |
| $\delta$      | Parameter        | Information sharing rate        |



## B Sensitivity Analysis Results

In our full sensitivity analysis, we used 18 sets of unique parameters to understand how information sharing and type interact and affect disease outbreaks.

In Figure 7 we varied the ratio between  $\beta_a$  and  $\beta_u$  when good information,  $\beta_a < \beta_u$  was shared,  $\delta = 0.001$  across groups.

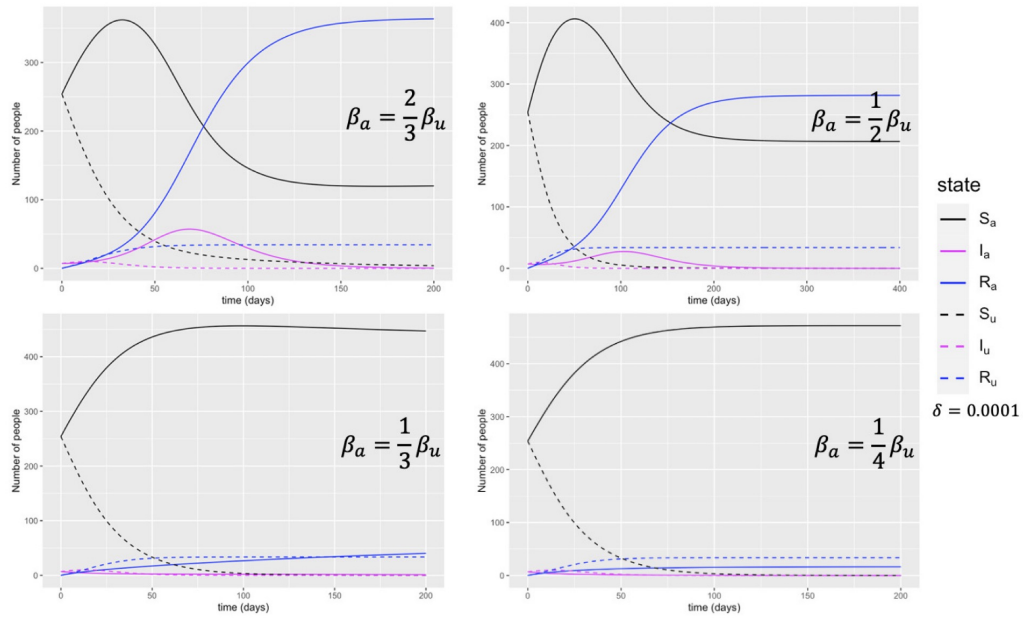


Figure 7: Spread of disease when varying levels of good information are shared.

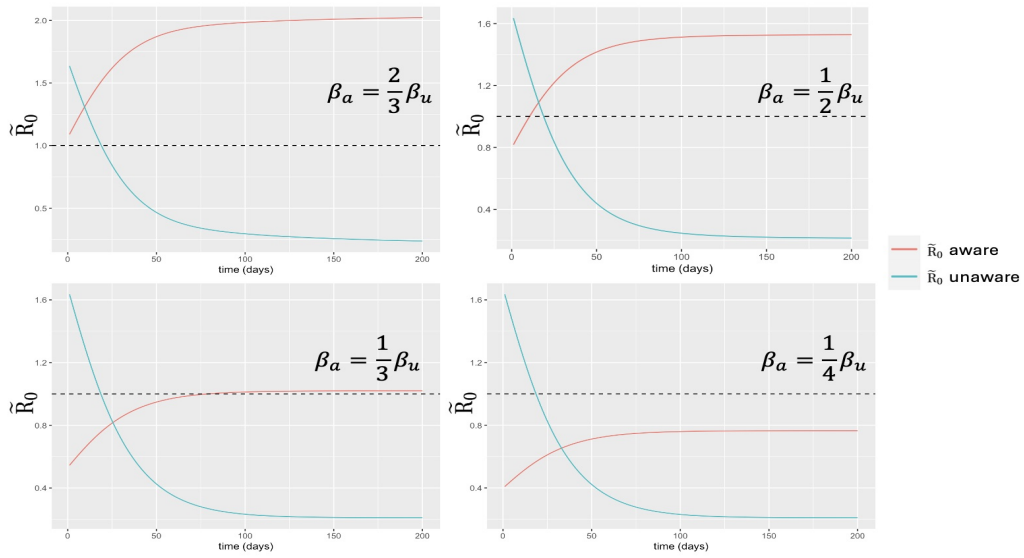


Figure 8:  $\tilde{R}_0$  when varying levels of good information are shared.

Our results from the sharing of good information sensitivity analysis are corroborated by our findings in Figure 8.

In Figure 9 we varied the ratio between  $\beta_a$  and  $\beta_u$  when bad information,  $\beta_a > \beta_u$  was shared,  $\delta = 0.001$  across groups.

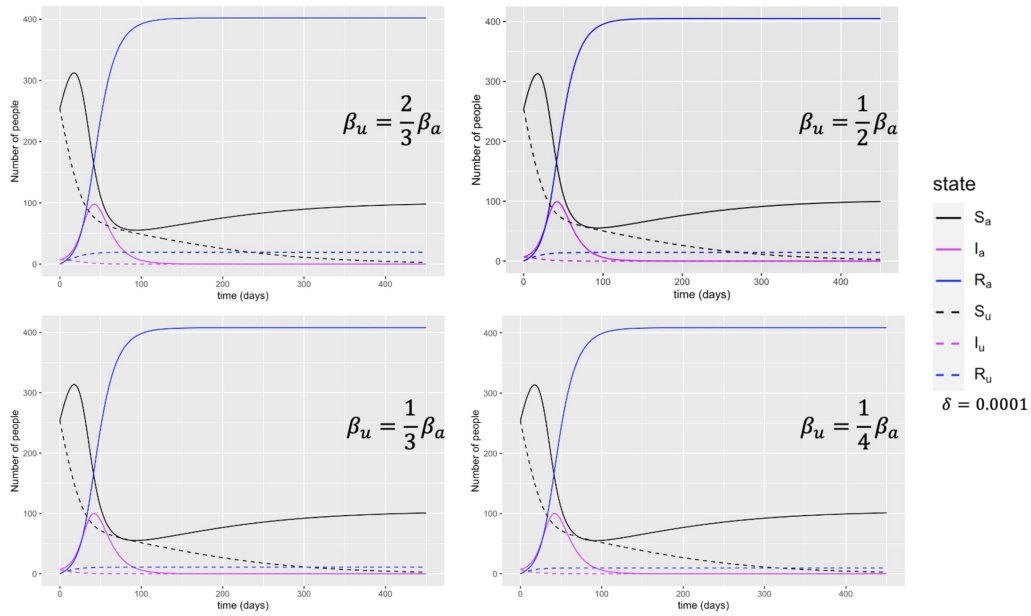


Figure 9: Spread of disease when varying levels of bad information are shared.

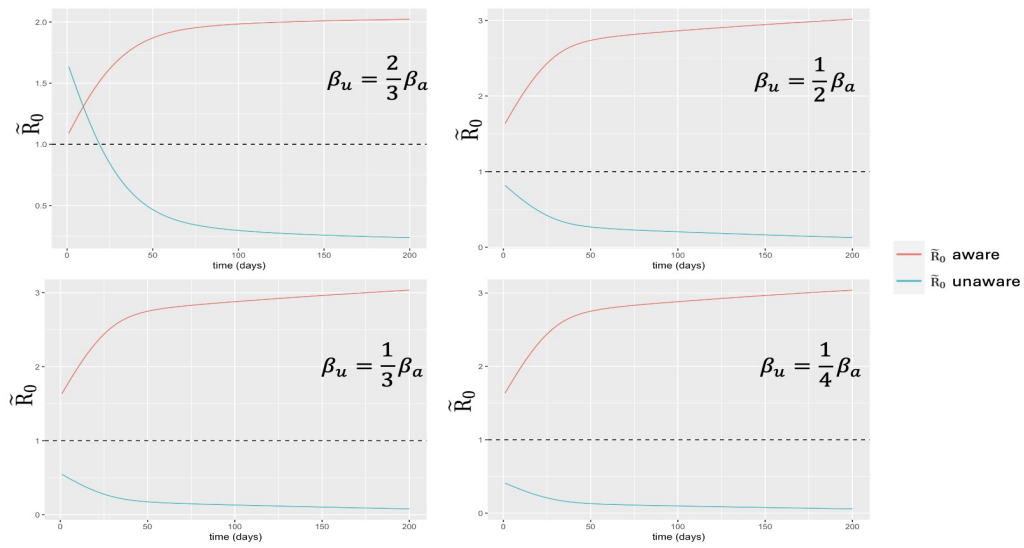


Figure 10:  $\tilde{R}_0$  when varying levels of bad information are shared.

Finally, in Figure 10 we confirmed the lack of qualitative differences between the curves shown in Figure 9.

## References

- [1] Incorporating human behavior in epidemiological models (IHBEM), (Mar 2022). URL: <https://beta.nsf.gov/funding/opportunities/incorporating-human-behavior-epidemiological-models-ihbem>
- [2] W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proceedings of the royal society of London. Series A, Containing papers of a mathematical and physical character* 115 (772) (1927) 700–721.
- [3] G. E. Funk, S., V. Jansen, Endemic disease, awareness, and local behavioral response, *Journal of Theoretical Biology* 264 (2) (2010) 501– 509. doi:<https://doi.org/10.1016/j.jtbi.2010.02.032>.
- [4] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2021). URL <https://www.R-project.org/>
- [5] H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. Francois, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the tidyverse, *J. Open Source Softw.* 4 (43) (2019) 1686. doi:10.21105/joss.01686. URL: <http://dx.doi.org/10.21105/joss.01686>
- [6] S. A. et mult. al., *DescTools: Tools for Descriptive Statistics*, r package version 0.99.44 (2021). URL <https://cran.r-project.org/package=DescTools>
- [7] K. Soetaert, T. Petzoldt, R. W. 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>
- [8] 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>
- [9] G. F. Raggett, Modeling the Eyam plague, *Bulletin of the Institute of Mathematics and its Applications* 18 (1982) 221–226. doi:10.1080/02664768200000021.
- [10] Christophe Fraser et al., Pandemic Potential of a Strain of Influenza A (H1N1): Early Findings. *Science* 324, (2009) 1557–1561. doi:10.1126/science.1176062

**Katie Yan**

Skidmore College, Mathematics & Statistics Department

kyan@skidmore.edu