CODEE Journal

Volume 17 *Engaging the World: Differential Equations Can Influence Public Policies* Article 1

1-4-2024

To Open or Not to Open: Developing a COVID-19 Model Specific to Small Residential Campuses

Christina Joy Edholm
Scripps College

Maryann Hohn

Nicole Lee Falicov

Emily Lee

Lily Natasha Wartman

See next page for additional authors

To Open or Not to Open: Developing a COVID-19 Model Specific to Small Residential Campuses

Authors

Christina Joy Edholm, Maryann Hohn, Nicole Lee Falicov, Emily Lee, Lily Natasha Wartman, and Ami Radunskaya

To Open or Not to Open: Developing a COVID-19 Model Specific to Small Residential Campuses

Christina J. Edholm
Nicole L. Falicov
Lily N. Wartman
Scripps College, Claremont, CA, USA

Maryann E. Hohn
Institute for Defense Analyses, Center for Computing Sciences, MD, USA

Emily Y. Lee
Claremont McKenna College, Claremont, CA, USA

Ami E. Radunskaya
Pomona College, Claremont, CA, USA

Keywords: ODE Model, COVID-19 Pandemic, Parameter Estimation
Manuscript received on July 19, 2023; published on January 4, 2024.

Abstract: In May 2020, administrators of residential colleges struggled with the decision of whether or not to open their campuses in the Fall semester of 2020. To help guide this decision, we formulated an ODE model capturing the dynamics of the spread of COVID-19 on a residential campus. In order to provide as much information as possible for administrators, the model accounts for the different behaviors, susceptibility, and risks in the various sub-populations that make up the campus community. In particular, we start with a traditional SEIR model and add compartments representing relevant variables, such as quarantine compartments and a hospitalized compartment. We then duplicated the model for ten interacting sub-populations, result-ing in a large system of differential equations. The model predicts possible outcomes based on hypothetical administrative policies such as masking, social distancing, and quarantining. As the pandemic developed, we updated the model to account for new policies, such as testing and vaccination and calibrated the model to data gathered from local sources. To complete the modeling process, we describe the parameter-fitting procedure, in which we used publicly available data from the county, as well as specific descriptions of our student body, faculty, and staff. The final stage of the work involved performing numerical simulations and designing an interactive application that allows non-mathematicians to experiment with a range of scenarios. We then extrapolate the findings of our model to a general audience, which along with our plots and app makes model conclusions accessible to all, democratizing the policy-making process.

Contents

1 Introduction

First appearing in 2019 in Wuhan, China COVID-19 lead to an international pandemic, spanning the globe and years [22, 44]. There have been a number of studies tracking the pandemic and exploring the virus on a cellular level, along with reflecting on similar historic outbreaks [22, 44]. For the United States, much of the general public information could be found in credible news sources and from government agencies, for instance the Centers for Disease Control and prevention (CDC) [17, 2]. During the pandemic, a number of mathematical models were proposed to study the disease dynamics, using a variety of techniques, data, and analysis options. The ordinary differential equation (ODE) model with three compartments, Susceptible-Infected-Recovered (SIR), has been used since the early 20^{th} century to study epidemics [8]. These models are explored in a variety of mathematical modeling texbooks and resources for reference [11, 31, 30, 42]. For the COVID-19 pandemic, researchers used the SIR model as a basis to study features of the spread of the disease, considering additional variables to accurately capture the displayed characteristics.

One such expanded model from the SIR was to include an Exposed class, resulting in an SEIR model which can then include additional dynamics. A variety of studies were done using the SEIR model format, which led to some new insights. Wang et al. used an SEIR model for modeling cases in Wuhan, China at the start of the pandemic [41]. Ivorra et al. expanded on SEIR by adding Hospitalized and Dead compartments, and then considered control/management options for a case study of China [21]. Roda et al. discussed model choice and pros/cons of SIR versus SEIR, specifically using AIC for model selection with a case study of parameters from China [38]. Later, Yang et al. performed a review of models focusing on SIR and SEIR, along with models including additional extensions to the SEIR [43]. Additionally, other models expanded on the SEIR to include asymptomatic classes and consider case studies in other locations, such as South Africa, Italy, and British Columbia [15, 19].

Later in the pandemic, vaccines were introduced for COVID-19 updating the dynamics and management options. The CDC is a resource for data on vaccinations: from their website, we can find the specific dates on which vaccines were implemented according to the phases given by the CDC [16]. Dooling et al. outlined the phases and the tiers for COVID-19 vaccinations [13]. Pritchard et al. described COVID in the UK. They used a larger community/sample size to find a more accurate efficacy of each vaccine [37]. Milman et al. showed that with the vaccine, there is a population-level effect, cross-protection for those who are not vaccinated, specifically the effects of the Pfizer vaccine [29]. Pilishvili et al. examined the likelihood of transmission of COVID-19 after a person received the first and second dose, data was taken from the healthcare professionals/ frontline workers for January 2021 through March 2021 [34]. There are many more ODE models examining vaccinations for COVID-19 and the role vaccinations play in management, for instance, Edholm et al. [15].

Our focus in this paper is on the Claremont Colleges, a consortium of five residential undergraduate colleges in Los Angeles (LA) County. In May 2020 we started working on the question of whether or not the colleges should be remote or in-person for the Fall semester of 2020: our goal was to develop and implement a model that administrators could use to help in their policy decisions. During the summer of 2020, we presented our work to presidents of the colleges and eventually learned the colleges would be remote for the 2020-2021 school year. During the summers of 2021 and 2022, we had undergraduate research assistants who extended the models past Fall 2020, utilizing new data, updating dynamics, and constructing coding interfaces. Since the start of our collaboration, a number of other researchers have considered what will happen with COVID-19 on college campuses [18, 6, 7, 24]. Some of the models use ODE dynamics and others use agent-based models or other mathematical frameworks. For instance, in 2021-2022 a model was formulated to examine dining dynamics with an agent-based model for one of the Claremont Colleges [25]. One key model for our research examined the spread of COVID-19 in a closed environment with distinct classes of individuals, a prison system [26, 27]. Lofgren et al. capture the dynamics of an isolated location where different individuals with different roles interact in specific ways [26, 27]. The manuscript by Lofgren et al. used the jail system in Allegheny County, PA, and showed how to develop such models across the US. We implement similar dynamics by dividing the college population into students, administrators/teachers, and staff, and considering the closed interactions that

happen on a residential campus. We are fortunate to have access to a variety of data through the California Open Data portal [2, 3, 4, 5].

In the rest of the paper, we will tell the story of our model development, and how we calibrated it to local data and then used it to make predictions. Hopefully, this story will help future mathematicians (and college administrators?) engage with public health emergencies in an effective way. In Section 2 we outline the ODE models we formulated first for parameter estimation with LA County data, with no vaccination, Sections 2.1 and 2.3, and then for the Claremont Colleges 2.2 and 2.4. Sections 2.1 and 2.2 outline models without vaccination, while models in Sections 2.3 and 2.4 include new variables and parameters related to the COVID-19 vaccines. In Section 3 we present the parameter estimation from LA County data without vaccines in Section 3.1 and with vaccine data in Section 3.2. Next, in Section 3.3 we extend the parameters from LA County to the Claremont Colleges model and detail the process. Section 3.4 includes the numerical simulations we completed for the Claremont Colleges models, and then Section 3.5 provides an overview of the accompanying app we formulated in MATLAB® [20]. Section 4 includes our results and conclusions in Section 4.1, then future directions to consider in Section 4.2. Lastly, Section 4.3 contains suggestions for instructors and students on how to use the research in this paper and our insights from the process.

2 Model Formulation

In this section, we outline the mathematical models we use to capture COVID-19 dynamics in LA County and the Claremont Colleges. The first model in Section 2.1 is for LA County at the start of the COVID-19 pandemic, with no vaccination, which we expanded upon in Section 2.2 to the Claremont Colleges. Similarly, Section 2.3 focuses on LA County with the advent of the vaccine, and Section 2.4 expands the vaccine model to the Claremont Colleges.

2.1 LA County Model with No Vaccination

During the advent of the COVID-19 pandemic, to capture the dynamics of the outbreak, we formulated an ordinary differential equation model describing four state variables: Susceptible (can be infected), Exposed (infected by the virus, but not yet infectious and not showing symptoms), Infected (show symptoms and infectious), and Recovered (no longer infectious) individuals. Since Infected individuals both show symptoms and can transmit, to capture infected individuals who never show symptoms we included a transition rate from Exposed to Recovered individuals. To match available data on hospitalizations and COVID deaths, we wanted to keep track of individuals in medical facilities and those who passed away due to the disease. Therefore, we added two more state variables: Medical (hospitalized) and Dead (those who have died). A flow diagram of our model is shown in Figure 1. A table of our state variables, denoted by capital letters, can be seen in Table 1, and the parameters, denoted by lower case Greek letters (with one exception: $M_{\max}$) are listed in Table 2.

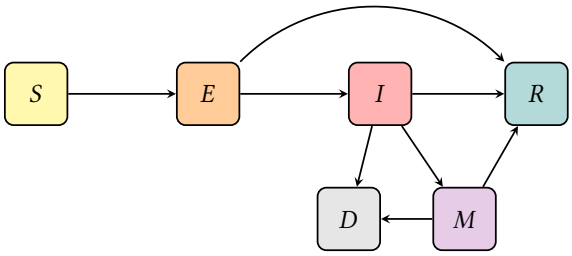

Figure 1: Flow diagram for COVID-19 model for LA County with no vaccination.

The equations for our model are

$$\begin{aligned}
\frac{dS}{dt} &= -\beta S(g + I + \alpha E) \\
\frac{dE}{dt} &= \beta S(g + I + \alpha E) - (\gamma_I + \gamma_R)E \\
\frac{dI}{dt} &= \gamma_I E - (\delta_I + \mu_I)I - \omega_I e^{-(M/M_{max})^2} I \\
\frac{dR}{dt} &= \delta_I I + \delta_M M + \gamma_R E \\
\frac{dM}{dt} &= \omega_I e^{-(M/M_{max})^2} I - (\delta_M + \mu_M)M \\
\frac{dD}{dt} &= \mu_I I + \mu_M M
\end{aligned} \tag{2.1}$$

Model	Variable	Description
Section 2.1	S	Number of Susceptible individuals
	E	Number of Exposed individuals
	I	Number of Infected individuals
	R	Number of Recovered individuals
	M	Number of Medical center individuals
	D	Number of Dead individuals
Section 2.2	H_S	Number of Susceptible individuals being held
	H_E	Number of Exposed individuals being held
	H_I	Number of Infected individuals being held from contact tracing - *only for non-compliant students*
Section 2.3	V_1	Number of partially vaccinated individuals
	V_2	Number of fully vaccinated individuals
	E_1	Number of partially vaccinated and exposed individuals
	E_2	Number of fully vaccinated and exposed individuals
	I_v	Number of Infected vaccinated individuals

Table 1: Description of the variables for both no vaccination and vaccination models depicted in Figures 1, 4, 5, 6.

Model	Symbol	Description
Section 2.1	α	fraction change in transmission during asymptomatic period
	β	transmission rate
	g	greater community interaction probability.
	γ_I	transfer rate of Exposed to Infected individuals
	γ_R	transfer rate of Exposed to Recovered individuals
	δ_I	recovery rate for Infected individuals
	δ_M	recovery rate for Medical center individuals
	M_{max}	hospital capacity parameter
	μ_I	death rate of Infected individuals
	μ_M	death rate of Medical center individuals
	ω_I	transfer rate of Infected individuals to Medical center
Section 2.2	κ	effectiveness of contact tracing
	ρ_S	(length of quarantine $)^{-1}$
	σ	average number of individuals in contact with an infected person per day.
Section 2.3	$\check{\xi}_1$	vaccination rate of susceptible individuals
	$\check{\xi}_2$	vaccination rate of partially vaccinated individuals
	b_1	scale of transmission rate of partially vaccinated individuals
	b_2	scale of transmission rate of fully vaccinated individuals
	c_1	scale of infection rate of exposed partially vaccinated individuals
	c_2	scale of infection rate of exposed fully vaccinated individuals
	d_1	scale of recovery rate of exposed partially vaccinated individuals
	d_2	scale of recovery rate of exposed fully vaccinated individuals
	δ_v	recovery rate for vaccinated Infected individuals

Table 2: Parameter descriptions for both no vaccination and vaccination models depicted in Figures 1, 4, 5, 6.

The right-hand side of each equation shows the rate at which individuals enter or leave each compartment. These equations match up with the flow diagram: an arrow pointing towards a compartment will appear as a positive term in the rate equation, and an arrow pointing away from a compartment will appear as a negative term in the rate equation. With this in mind, the equations should make intuitive sense, after referring to Table 2. Here are a few comments explaining some of the terms:

- The sum of the right-hand sides of the equations is zero. This says that we are considering a "closed system", i.e. we assume the total population under consideration is constant. In other words, we are ignoring new births, immigration and emigration, and deaths that are not due to COVID. This assumption is reasonable if we are trying to predict the effect of the disease over a relatively short period of time in a closed community such as a residential college.

- We described the E (Exposed) compartment as "infected but asymptomatic and not infective". However, as the pandemic evolved, it was apparent that asymptomatic individuals did, in fact, transmit the disease. Therefore, we consider the E class to

be infective, and include the transmission term αE in the first equation.

- The parameter g, accounts for outside influence on disease transmission. For this first model, we assume that LA county is self-contained, so we set $g = 0$. This parameter will play more of a role in the Claremont Colleges models in Sections 2.2 and 2.4, where transmission rates could increase when individuals from the Colleges interact with people outside of the community.

- The first equation describes the rate at which susceptible individuals become infected. We can interpret this equation as follows: transmission of the disease is proportional to the number of encounters between infective and susceptible individuals. We approximate the number of possible encounters by the number of possible susceptible-infected pairs, i.e. the product of the number of susceptibles, S, with the number of infective individuals, I. As mentioned above, we add to the infectives the infected but asymptomatic individuals, E, and infectives outside the community, g. To take into account differences in the probability of encountering individuals in different classes, as well as differences in transmission probabilities, we multiply the number in the exposed class by a factor, α. At different stages in the pandemic, α might be smaller than 1 (asymptomatic individuals are less likely to transmit the disease since they carry less viral load), or α might be bigger than 1 (asymptomatic individuals might be more likely to transmit the disease, since they are more likely to be out and about, they don't know they are infected, and yet the virus is easily transmissible). The parameter β can be interpreted as the probability that an encounter between a susceptible and an infected takes place *and* that this encounter results in an infection. These parameters are discussed more fully in Section 3.

- We include the term $e^{-(M/M_{max})^2}$ in the $\frac{dM}{dt}$ and $\frac{dI}{dt}$ equations to account for the number of available hospital beds, which was a concern throughout the pandemic, but particularly at the start. As the number of hospitalized individuals approaches $M_{\max}$, the rate at which people can enter the Medical class decreases exponentially. Note that M_{max} is not exactly the maximum number of hospital beds: even if $M = M_{max}$, there still is a positive probability that an infected person can enter the hospital.

2.2 Model 1 College COVID 5C with No Vaccination

When expanding the model in System 2.1 to the Claremont Colleges, we considered the different types of individuals present on the campuses as well as interactions between these groups. We have three communities which consist of administrators and teaching staff; housing, dining and groundskeeping staff; and students. These groups are depicted in Figure 2. Each community is then divided into categories based on risk and student interactions, see Figure 3. For administrators and teaching, housekeeping, dining staff, we further divide the populations into high/low student interaction (HS/LS) and then further into high/low risk categories (HR/LR), see Figure 3. We assume that groundskeepers will have minimal student interaction, so we create only high/low risk categories for them.

For students, we assume only low risk students will attend the college and for those students we divide them into two categories - those who will be compliant with social distancing and other safety measures, Section 2.2.1, and those who will not comply, Section 2.2.2. Furthermore, we assume all individuals who are not students will be compliant, Section 2.2.1. Lastly, we are assuming the model only considers individuals on campus at the Claremont Colleges, so the g parameter allows us to include interactions with the surrounding community by any individuals.

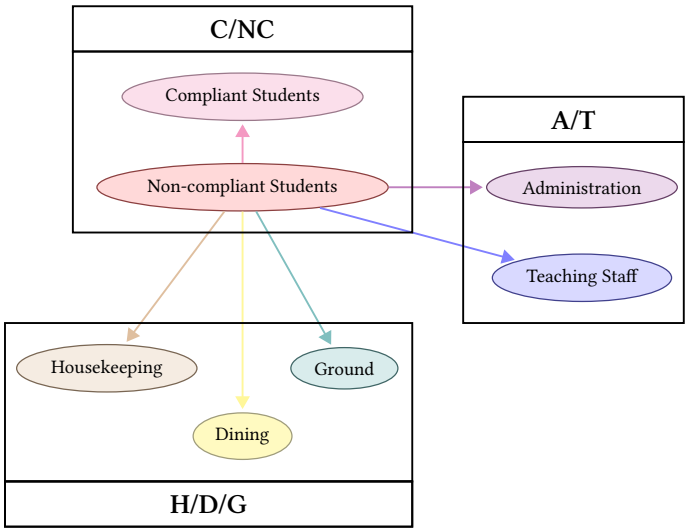

Figure 2: Three communities at the Claremont Colleges.

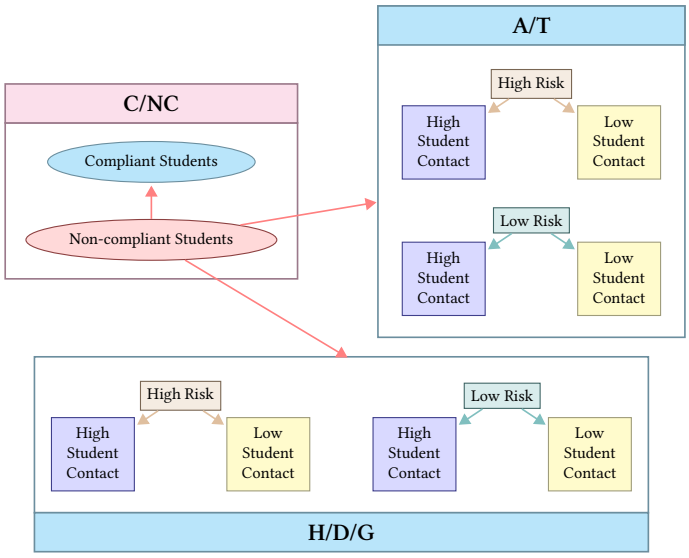

Figure 3: Three communities at the Claremont Colleges with High and Low Risk along with High and Low Student Contact included. Blue denotes those who are compliant with health care protocols, and pink denotes those who are low risk individuals.

2.2.1 Compliant

We assumed non-students and a subgroup of students would be compliant with public health protocols, affecting their behavior in terms of quarantining. As stated previously, this model was constructed in Summer 2020, when there were no tests or vaccinations for COVID-19, so quarantining or holding individuals through contact tracing was the main form of management we could implement. We use the model from System 2.1 as the basis and then add additional Held classes for removing individuals from the general population who were in contact with an infected individual. The Held Class, or Held State, encompasses the uncertain nature of the disease, allowing for quarantine, which is for the sick, and also those being isolated after being in contact with a sick person; this latter group will return to the Susceptible class. Since compliant individuals follow healthcare procedures, we assume that if they are infected they will quarantine on their own and do not require an additional "held" class. Susceptible individuals who are isolated in the Held Class will eventually return to the Susceptible Class since they do not carry the pathogen, while exposed individuals who are in the Held Class from contact tracing will remain in that class (i.e. remain isolated) until they recover. The updated flow diagram can be seen in Figure 4.

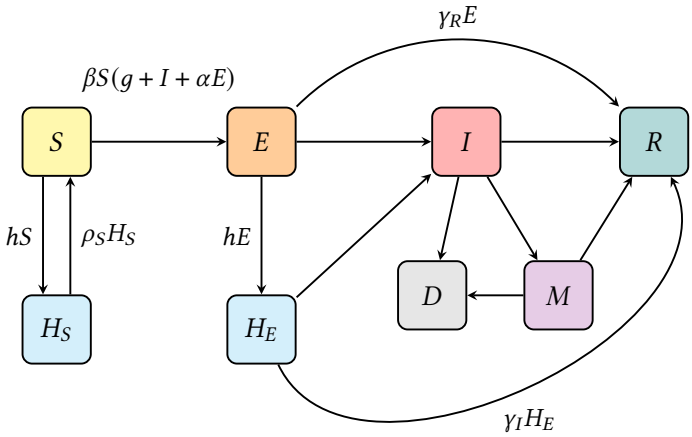

Figure 4: Flow diagram for the Claremont Colleges Compliant individuals with no vaccination.

The updated equations for compliant individuals are as follows:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta S(g + I + \alpha E) - hS + \rho_S H_S \\
\frac{dE}{dt} &= \beta S(g + I + \alpha E) - (\gamma_I + \gamma_R)E - hE \\
\frac{dI}{dt} &= \gamma_I(E + H_E) - (\delta_I + \mu_I)I - \omega_I e^{-(M/M_{max})^2} I \\
\frac{dR}{dt} &= \delta_I I + \delta_M M + \gamma_R(E + H_E) \\
\frac{dM}{dt} &= \omega_I e^{-(M/M_{max})^2} I - (\delta_M + \mu_M)M \\
\frac{dD}{dt} &= \mu_I I + \mu_M M \\
\frac{dH_S}{dt} &= hS - \rho_S H_S \\
\frac{dH_E}{dt} &= hE - (\gamma_R + \gamma_I)H_E
\end{aligned} \tag{2.2}$$

where $h(S, E, I, R) = \left(\frac{\kappa\sigma}{1+S+E+I+R}\right)\left(\frac{I}{1+I}\right)$. The function h describes the rate at which the students go into the Held class and is influenced by the effectiveness of contract tracing κ, and the average number of individuals in contact with infected person. Note that we added one in the denominators to avoid dividing by zero.

2.2.2 Non-Compliant

While we assume a subgroup of the students will be compliant with public health protocols, we consider the remaining students will be non-compliant through any number of decisions. There have been studies focused on the proportion of college/university students who were non-compliant [10, 36, 32]. To reflect the difference in behavior, we included an additional class in the model to hold Infected individuals since non-compliant individuals might not voluntarily quarantine and require a held class, see Figure 5. Additionally, in Section 3 we will explore different parameter values for these different subgroups of individuals.

The updated equations for non-compliant individuals are as follows:

$$
\begin{aligned}
\frac{dS}{dt} &= -\beta S(g + I + \alpha E) - hS + \rho_S H_S \\
\frac{dE}{dt} &= \beta S(g + I + \alpha E) - (\gamma_I + \gamma_R)E - hE \\
\frac{dI}{dt} &= \gamma_I E - (\delta_I + \mu_I)I - hI - \omega_I e^{-(M/M_{max})^2} I \\
\frac{dR}{dt} &= \delta_I(I + H_I) + \delta_M M + \gamma_R(E + H_E) \\
\frac{dM}{dt} &= \omega_I e^{-(M/M_{max})^2}(I + H_I) - (\delta_M + \mu_M)M \\
\frac{dD}{dt} &= \mu_I(I + H_I) + \mu_M M \\
\frac{dH_S}{dt} &= hS - \rho_S H_S \\
\frac{dH_E}{dt} &= hE - (\gamma_I + \gamma_R)H_E \\
\frac{dH_I}{dt} &= hI + \gamma_I H_E - \left(\omega_I e^{-(M/M_{max})^2} + \delta_I + \mu_I\right) H_I
\end{aligned}
\tag{2.3}
$$

where $h = \left(\frac{\kappa\sigma}{1+S+E+I+R}\right)\left(\frac{I}{1+I}\right)$.

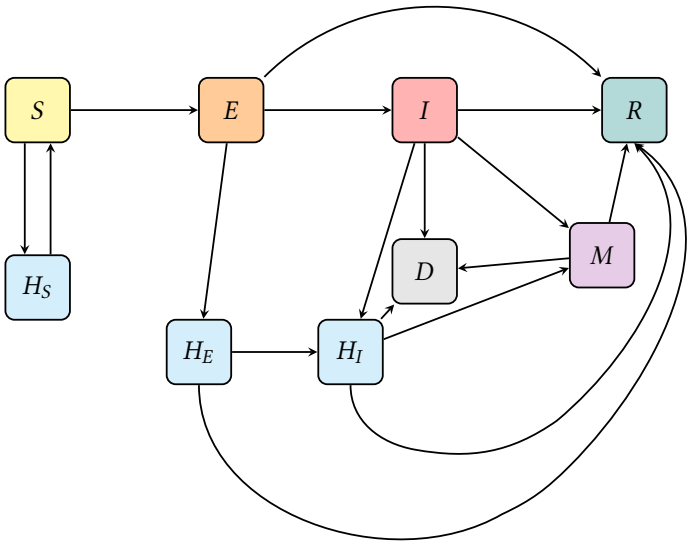

Figure 5: Flow diagram for the Claremont Colleges Non-Compliant individuals with no vaccination.

Note that the first model of LA County, with six state variables, has now been expanded to one with 81 variables to capture all the communities depicted in Figure 2. The 81 variables for Figure 3 comes from there are ten communities for the colleges where the

nine compliant communities will have a copy of the model in Figure 4 with eight classes, while the non-compliant students utilize the model in Figure 5 with nine classes.

2.3 LA County Model with Vaccination

As the pandemic evolved, we needed to update our model to reflect changes to the dynamics, specifically, we focused on the novel vaccines introduced in Spring of 2021. Our new model has eleven population classes, which represent the number of individuals in a state at any time, see Figure 6. The population class S represents individuals susceptible to COVID-19. The classes V_1 and V_2, respectively, are the population of partially vaccinated and vaccinated individuals against COVID-19. When someone is exposed to COVID-19 they move into an exposed class, either E, E_1, or E_2 based on vaccination status. The classes I and I_v are the infected population. It was critical to separate the infected population into two classes based on vaccination status (partially and fully vaccinated and infected are I_V and unvaccinated/susceptible and infected are I) because medical research shows duration of illness, severity, and death rate from COVID-19 is significantly different depending on one's vaccination status. As before, the population classes R, M, and D represent recovered individuals, individuals in the medical center, and deaths, respectfully. See Table 1 for descriptions of the additional variables and the parameters are shown in Table 2.

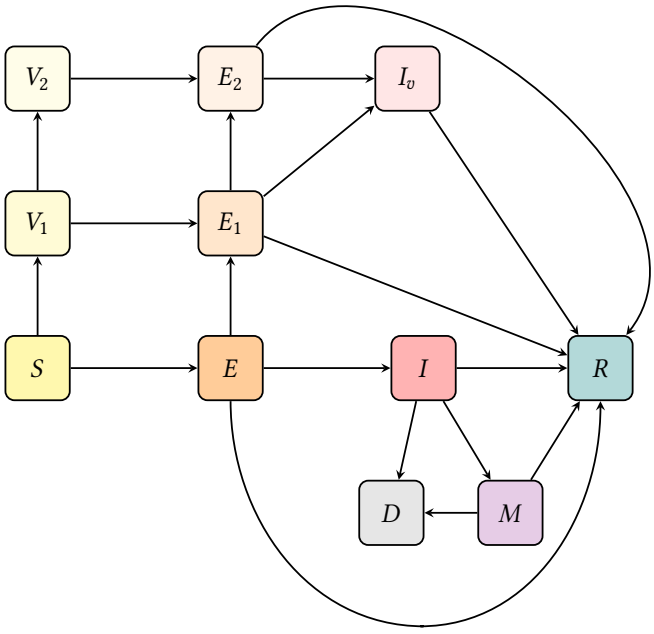

Figure 6: LA County Model with vaccination flow diagram for COVID-19.

The model with vaccines for LA County is described by the following equations, which are similar to Equations 2.1.

$$
\begin{aligned}
\frac{dS}{dt} &= -\lambda S - \xi_1 S \\
\frac{dE}{dt} &= \lambda S - (\gamma_I + \gamma_R)E \\
\frac{dI}{dt} &= \gamma_I E - (\delta_I + \mu_I)I - \omega_I e^{-(M/M_{max})^2} I \\
\frac{dR}{dt} &= \delta_I I + \delta_M M + \gamma_R E + d_1\gamma_R E_1 + d_2\gamma_R E_2 + \delta_v I_v \\
\frac{dM}{dt} &= \omega_I e^{-(M/M_{max})^2} I - (\delta_M + \mu_M)M \\
\frac{dD}{dt} &= \mu_I I + \mu_M M \\
\frac{dV_1}{dt} &= \xi_1 S - \xi_2 V_1 - b_1\lambda V_1 \\
\frac{dE_1}{dt} &= b_1\lambda V_1 - (c_1\gamma_I + d_1\gamma_R)E_1 \\
\frac{dV_2}{dt} &= \xi_2 V_1 - b_2\lambda V_2 \\
\frac{dE_2}{dt} &= b_2\lambda V_2 - (c_2\gamma_I + d_2\gamma_R)E_2 \\
\frac{dI_v}{dt} &= c_1\gamma_I E_1 + c_2\gamma_I E_2 - \delta_v I_v
\end{aligned}
\tag{2.4}
$$

where $\lambda = \beta(g + I + \alpha(E + E_1 + E_2))$.

2.4 Model 2 College COVID 5C with Vaccination

As with the models outlined in Section 2.2, we subdivided the population at the Claremont Colleges displayed in Figure 3. Since we are simulating during the 2021-2022 school year, we use the model from System 2.4 for each of the classes. Additionally, we do not include any held classes, rather assuming that testing and school policies incorporated being held into the exposed and infectious classes. With the addition of vaccination, we expect to observe smaller outbreaks and a safer environment in terms of disease mortality. At the Claremont Colleges, vaccines played a key role in opening the campus and updating healthcare protocols. In the discussion, we will include future configurations and suggested changes for this base model, Section 4.

3 Parameters and Numerical Simulation

Due to the changing nature of COVID-19 policies, human behavioral responses, and strains of the virus throughout the pandemic, using deterministic equations to fit long periods of time does not work. Rather, we split time into smaller segments where behavior was constant and then consider the parameters during that time. Thus, we computed numerous

parameter fittings for System 2.1 and additional parameter fittings when we use System 2.4 which includes vaccinations. We estimated some parameters from the literature and others we fit using data from LA County for System 2.1 and System 2.4. We then used these parameters to simulate the Claremont College Models. For our simulations, we were fortunate to have access to a variety of data through the California Open Data portal [3, 5, 4].

3.1 LA County Model with No Vaccination Parameters Literature and Estimation

Our first model captures dynamics at the onset of the COVID-19 pandemic, specifically in LA County. For some parameters in the model, we established values from existing literature and public health guidelines. Since we were initially working in response to campus closures in Spring 2020 with an eye to opening or closing the campus in Fall 2020, we started with time ranges early in the pandemic April 20 - May 16, 2020 and June 13 - June 30, 2020. These ranges also capture changes in policy in terms of closures prior to any testing or vaccine management options. We also included August 1 - September 6, 2020 and September 7 - November 1, 2020 to examine what happened just prior to and during the semester, after our campuses had elected to stay remote for Fall 2020.

As mentioned, we use data from the California Open Data portal to estimate initial conditions and parameters [3]. For initial conditions, to find the initial number of infected, we take the difference between the start date we are fitting and 10 days before for the cumulative case data, and use a scalar of this for the initial number of exposed individuals (presymptomatic and asymptomatic). For the initial number of recovered individuals, we scale the number of individuals from cumulative infected data to express that these have individuals who have recovered, and due to our short fitting periods, we do not have waning immunity. The initial conditions for the Medical facilities, Dead individuals, and Cumulative infected (used in fitting) all come from the data, so the initial Susceptible comes from taking the total population and subtracting the other class initial conditions.

Parameters that were unobtainable through literature or simple calculations on data were estimated by fitting the model to Los Angeles County Public Health data on COVID-19 cumulative cases, cumulative deaths, and hospitalizations [39]. The method we used has been implemented on a variety of models for COVID-19 and other diseases [1, 9, 14, 15, 19, 23]. To compare the data with the model, we estimated cumulative cases by integrating the arrivals into the infected compartment (I). In other words, we add a differential equation to the system representing the cumulative infected cases, CI:

$$\frac{dCI}{dt} = g\beta(I + \alpha E)S$$

with $CI(0)$ set equal to the cumulative cases on the date we start fitting data. Since the compartment representing deaths is cumulative (there is no outflow from this compartment), we compare the cumulative deaths in LA to $D(t)$. Finally, the number of hospitalizations is reported as a daily rate, so this is directly compared to the model variable, $M(t)$. To find the parameter values, we minimize the difference between the data and model outputs, J_1,

$$J_1 = \left\| \frac{CI_{\text{data}} - CI_{\text{model}}}{\overline{CI_{\text{model}}}} \right\| + \left\| \frac{D_{\text{data}} - D_{\text{model}}}{\overline{D_{\text{model}}}} \right\| + \left\| \frac{M_{\text{data}} - M_{\text{model}}}{\overline{M_{\text{model}}}} \right\|. \tag{3.1}$$

where $||\cdot||$ is the L^2 norm: $||x|| = \sqrt{\sum_i x_i^2}$, and $\bar{\cdot}$ is the average/mean: $\overline{x} = \frac{x_1+\ldots+x_N}{N}$. We use the built-in function `fmincon` in MATLAB® which, given a range and initial estimate for each parameter, finds a local minimum of J_1, the least squares difference between the data and model output measures, with respect to the parameter values. Furthermore, we implement the routine `Multistart` which runs `fmincon` for multiple iterations, using a grid of initial estimates, to search the parameter space for multiple minima, returning the parameter combination that minimizes the distance between the data and the model outputs, J_1. The parameter values of the resulting fits are in Table 3, and simulations using the estimated parameters are displayed in Figure 7. For a discussion of these comparison graphs (Figures 7, 8, 15), as well as other subsequent figures, see Section 4.

Symbol	4/20-5/16/20	6/13-6/30/20	8/1-9/06/20	9/7-11/1/20
α	0.0389	2.2096	0.2231	0.6358
β	4.9486×10^{-8}	2.787×10^{-9}	5.0163×10^{-9}	6.4247×10^{-9}
δ_I	0.2669	0.046	0.0887	0.0562
δ_M	0.0129	0.0129	0.038	0.038
γ_I	0.0092	0.01332	0.0048	0.0049
γ_R	0.0378	0.0013	0.0375	0.0212
M_{max}	2548	5694	10,000	10,000
μ_I	3.4126×10^{-4}	1.0016×10^{-8}	6.2976×10^{-4}	6.7116×10^{-5}
μ_M	0.0186	0.0144	0.0128	0.0177
ω_I	0.0433	0.006	0.0031	0.0039

Table 3: Parameter values fit to LA County Data, dates listed, with no vaccination. The values of the scaled residuals, J_1, given in Equation (3.1) for each time period were $J_1 = 0.1518$ (04/20 - 05/16/20), $J_1 = .0002$ (06/13-06/30/20), $J_1 = 0.247$ (08/01-09/06/20) and $J_1 = 0.3453$ (09/07-11/01/20).

3.2 LA County with Vaccination Parameters Literature and Estimation

When using the model from System 2.4, we included additional state variables and parameters given in Tables 1 and 3, respectively. Some of these values were established from preexisting literature or calculations from data. For the rest of the parameters, we use the method described in Section 3.1. For this updated model with vaccination, and the campuses resuming in-person instruction, we considered time ranges immediately before and during the semester of Fall 2021: August 1 - September 1, 2021; September 1 - October 30, 2021; and October 31- November 30, 2021. Additionally, we considered a high travel

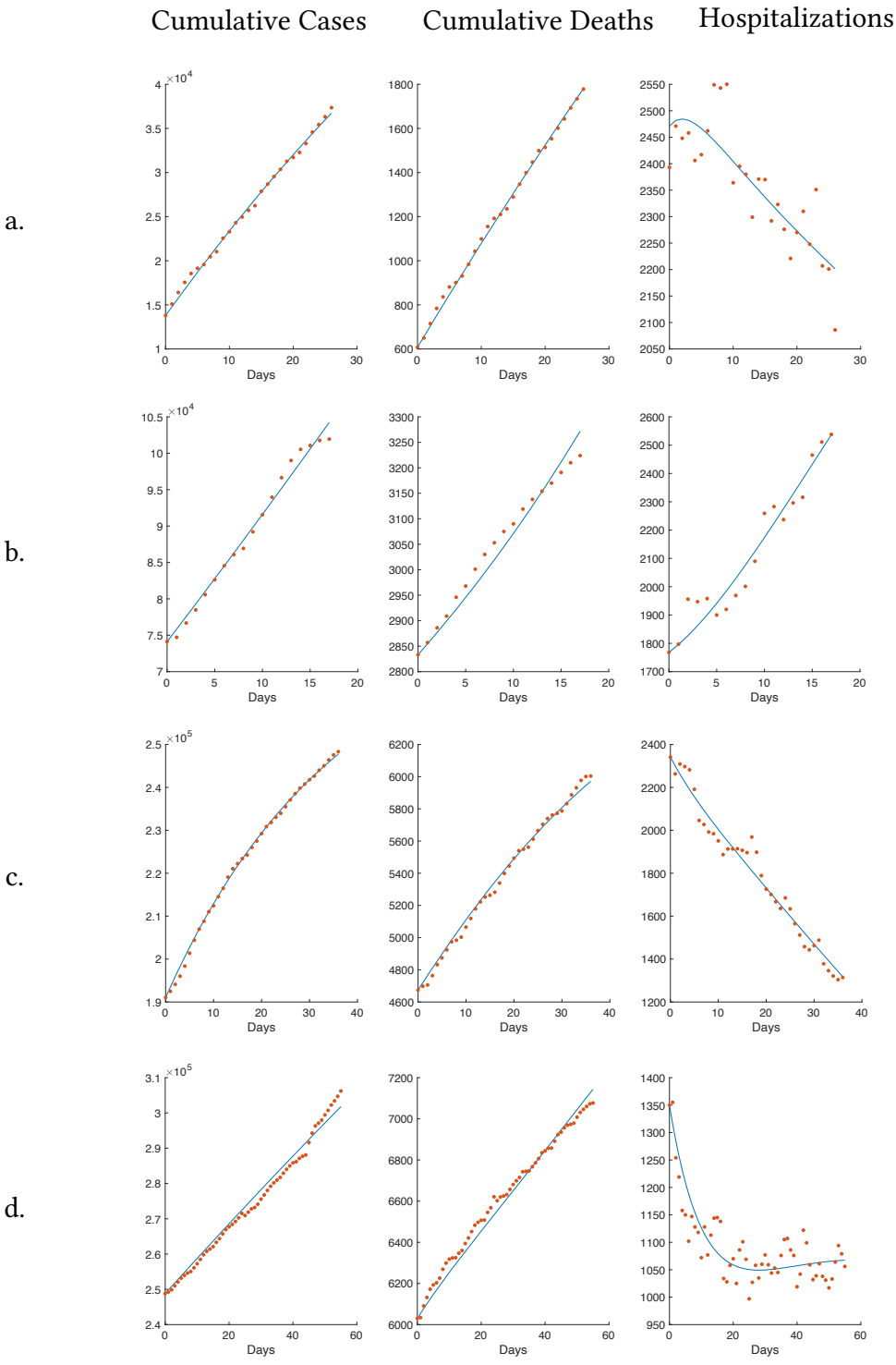

Figure 7: Model fits to four data sets from Los Angeles County in 2020, with no vaccination. (a.) April 20 - May 16, 2020; (b.) June 13 - June 30, 2020; (c.) August 1 - September 6, 2020; (d.) September 7 - November 1, 2020. The model was fit to publicly available data: cumulative confirmed cases, cumulative deaths due to COVID and daily hospitalizations.

period and the start of the Spring 2021 semester, using data from December 26, 2021 - January 23, 2022, during which our campuses elected to do remote instruction for the first two weeks of the semester.

For those parameters established from literature, we used a variety of sources, based on news and scientific literature [13, 16, 34, 37, 39, 40]. Additionally, we considered the parameter fits we had previously estimated for the model from System 2.1 from Table 3 and which parameters had stabilized. For instance, with this model, we have $\delta_I = 1/14$, to account for the fact that during these time periods, once you test positive or are diagnosed, you had to quarantine for 14 days. Similarly, the other recovery rates were assumed to be $d_1 = d_2 = 14/10$ (10 days of recovery) and $\delta_V = 1/10$ (10 days of recovery). Also, at this later stage in the pandemic we assume that hospital capacity remained constant, with $M_{max} = 10^4$. The parameter c_1 was found by multiplying the percentage of people who received each company's vaccine (Pfizer, Moderna, Johnson and Johnson) by the effectiveness of the first dose of the vaccine against hospitalization (weighted average).

$$c_1 = 1 - ((0.53)(0.82) + (0.4)(0.82) + (0.08)(0.72)) = 0.1798$$

Meanwhile, parameter c_2 was found by multiplying the percentage of people who got each company's vaccine (Pfizer, Moderna, Johnson and Johnson) by the effectiveness of the vaccine against hospitalization (weighted average).

$$c_2 = 1 - ((0.53)(1) + (0.4)(1) + (0.08)(0.86)) = 0.0012$$

We found baseline values for additional parameters, we would then fit. The baselines for parameters b_1 and b_2 were estimated by multiplying the percentage of individuals who received each company's vaccine (Pfizer, Moderna, Johnson and Johnson) by the effectiveness of the vaccine (weighted average):

$$b_1 = 1 - ((0.53)(0.82) + (0.4)(0.82) + (0.08)(0.72)) = 0.1798$$
$$b_2 = 1 - ((0.53)(0.95) + (0.4)(0.941) + (0.08)(0.72)) = 0.0625.$$

When we include the vaccine in our model, we update the initial conditions. The initial infected are now scaled by the proportion vaccinated in the population which is infected, so we can also include the initial vaccinated infected term. We scale these values to get the exposed individuals with and without the vaccine. We then can find the vaccinated class and susceptible classes through subtraction of established initial conditions.

For the remaining parameters, we fit System 2.4, using the method described in Section 3.1, requiring an updated equation for the difference between the data and model outputs, J_2,

$$J_2 = \left\| \frac{CI_{\text{data}} - CI_{\text{model}}}{\overline{CI_{\text{model}}}} \right\| + \left\| \frac{D_{\text{data}} - D_{\text{model}}}{\overline{D_{\text{model}}}} \right\| + \left\| \frac{M_{\text{data}} - M_{\text{model}}}{\overline{M_{\text{model}}}} \right\| + \left\| \frac{V_{\text{data}} - V_{\text{model}}}{\overline{V_{\text{model}}}} \right\| \quad (3.2)$$

Note that we added in data on vaccinations, so we have an additional term for J_2 from J_1 in Equation 3.1 [2, 3, 4, 5]. The parameter values of the resulting fits are in Table 4, and the fits are displayed in Figure 8.

Symbol	8/1-9/1/21	9/1-10/30/21	10/31-11/30/21	12/26/21-1/23/22
α	1.0230×10^{-5}	0.0963	0.0983	1.0001×10^{-5}
β	1.0000×10^{-8}	1.0025×10^{-9}	1.0053×10^{-9}	4.9667×10^{-6}
δ_M	0.0018	3.8970×10^{-5}	3.9663×10^{-5}	0.0576
γ_I	0.0031	0.0018	0.0017	0.0231
γ_R	0.0041	0.0127	0.0403	0.0010
μ_I	1.0107×10^{-6}	1.7285×10^{-4}	8.5829×10^{-4}	1.0003×10^{-6}
μ_M	0.0111	0.0196	0.0083	0.0087
ω_I	0.0051	1.0049×10^{-5}	2.9366×10^{-4}	0.0233
ξ_1	9.9996×10^{-4}	9.9586×10^{-4}	3.8410×10^{-4}	1.0131×10^{-7}
ξ_2	0.0173	0.0155	0.0120	0.0109
b_1	0.0029	0.0940	0.8886	0.0091
b_2	0.0014	0.0416	0.3515	0.0061

Table 4: Parameter values fit to LA County Data with vaccination [2, 3, 4, 5] used in System 2.4 based on the specific data period. Values for J_2 from equation 3.2 based on time period were $J_2 = 0.1500$ for (8/1-9/1/21), $J_2 = 0.5836$ for (9/1-10/30/21), $J_2 = 0.1765$ for (10/31-11/30/21), and $J_2 = 0.3973$for (12/26/21-1/23/22). The fixed parameters established from the literature were $\delta_I = 1/14$, $M_{max} = 10^4$, $d_1 = 14/10$, $d_2 = 14/10$, $\delta_V = 1/10$, $c_1 = 0.1798$, and $c_2 = 0.0012$.

3.3 Parameters for the Claremont Colleges Extensions

When considering the models for the Claremont Colleges, we used the parameters from the LA County models and then modified them based on the different communities in the Claremont Colleges. Note that the parameter selection will affect the outcomes of the model, which we explore in Sections 3.4, and these outcomes informed the advice we gave to administrators.

There are five of the original model parameters that we expect to differ between the community groups. These parameters had to do with the expected differences between high and low-risk groups, namely the death rates due to COVID (μ_I and μ_M), recovery rates (δ_I and δ_M), and hospitalization rates (ω_I). Since we did not have access to medical records of any individuals at the Colleges, we used age as a proxy for high and low risk. Anyone over 65 was categorized as “high risk", while all the rest were considered "low risk". Of course, this is a generalization that is not at all quantitatively accurate. Nevertheless, given the data at hand, the available knowledge, and the need to make a decision about whether or not to reopen the campus, this generalization (and others!) were unavoidable. LA County posted hospitalizations, death rates, and length of hospital stays broken down by age group ([33]), and we used these, together with age demographics of Claremont staff and faculty, to estimate parameters for the high and low-risk communities. Thus, these five parameters became vectors of length 13, with different values for each sub-population. The parameter g, which describes the expected number of daily interactions with the greater (outside of the Claremont Colleges) community, also varies between community groups. In fact, g factors into how we characterized “following the rules" or “misbehaving";

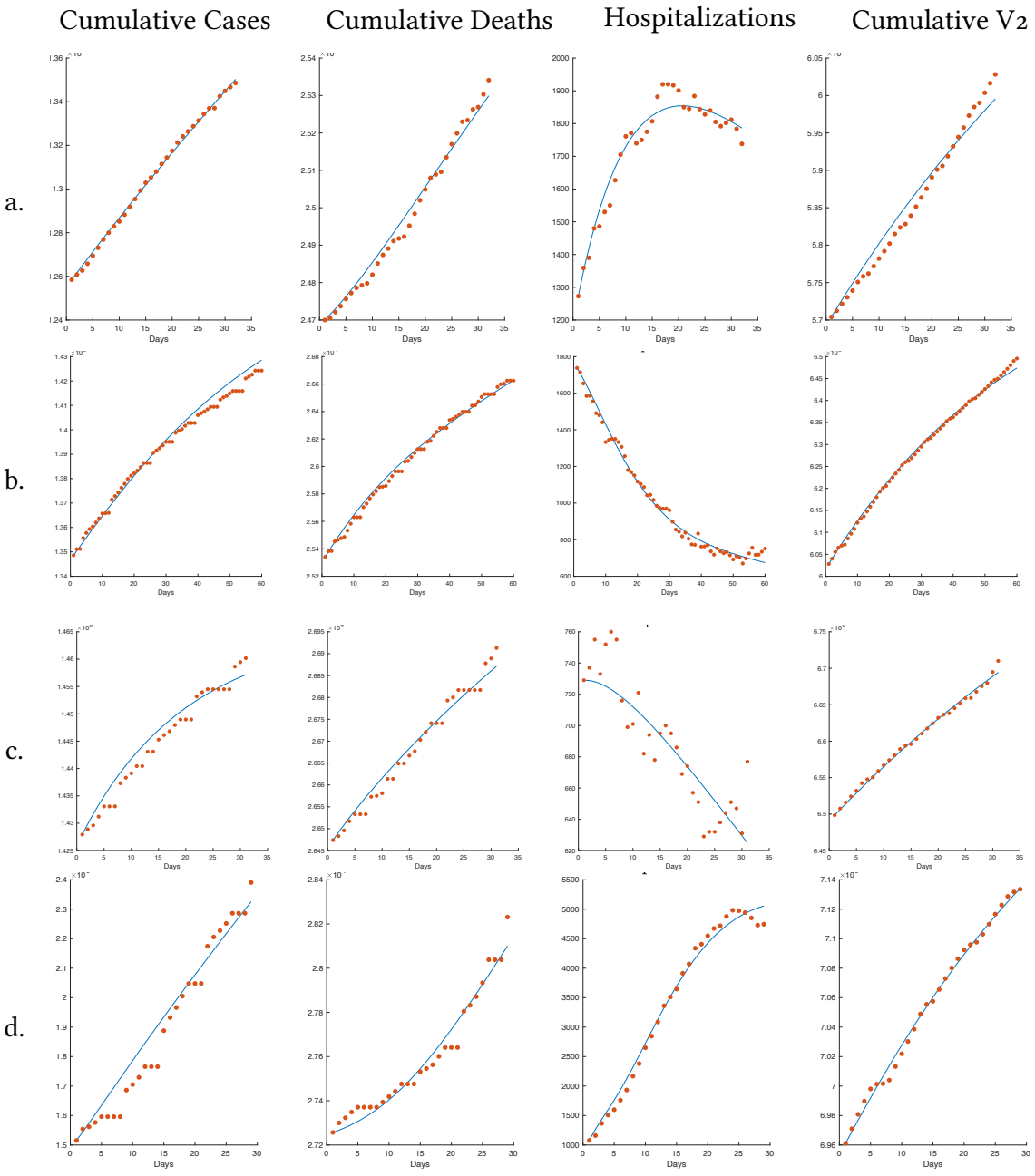

Figure 8: Model fits for four data sets from LA County in 2021 with vaccination, using parameters in Table 4. (a.) August 1 - September 1, 2021; (b.) September 1 - October 30, 2021; (c.) October 31- November 30, 2021; (d.) December 26, 2021 - January 23, 2022. The model was fit to publicly available data: cumulative confirmed cases, cumulative deaths due to COVID, hospitalizations, and cumulative number of fully vaccinated individuals.

we assumed that individuals who did not comply with social distancing guidelines would have a greater number of these interactions, on average. For the model presented in this

paper, only students were assumed to be non-compliant, with the assumption that faculty and staff would adhere to prescribed COVID guidelines.

The transmission rate, β, is slightly more complicated.

3.3.1 Transmission Rate β

The parameter β represents the rate of transmission from infected to susceptible individuals. In the full model, where the population is divided into different groups whose vulnerability and behavior is different, the values of β depend on which groups are interacting. These values are organized into a matrix, Betas, where Betas_{ij} is the transmission rate from infected individuals in group i to susceptible individuals in group j. To determine the relative values of these transmission rates, we follow the approach taken in [26].

A base level for the transmission rate is first estimated by fitting the extended SEIR model given in System 2.1 to LA County data; this base rate is denoted by β_{LA}. In the model equations, this rate is multiplied by the number of susceptible individuals and the number of infected individuals, so we interpret it as a "per-susceptible" infection rate. This base rate, β_{LA}, is then multiplied by the population of LA County to get a "per-population" infection rate: $\beta^* = \beta_{LA} \cdot n_{LA}$, where n_{LA} is the population of LA County.

To build the group-dependent transmission rates, we consider the *relative proclivity* of individuals in a specific group, q, to be the fraction of people in group q who are infected, relative to the fraction of people in the entire population who are infected:

$$\frac{\#\text{ cases in } q}{\#\text{ people in } q} \times \frac{\text{total }\#\text{ people}}{\text{total }\#\text{ cases}} = \text{relative proclivity of } q \text{ to be infected.}$$

Note that in our case "total number of people" means the total number in LA county, since that is the group we used for our base transmission estimate.

To estimate the transmission rate from group i to group j, we use the estimated β^* as a relative per-population infection rate, multiply by the relative proclivity of group i, and then multiply by the average number of times (per day) a person in group i is in contact with a person in group j. We then divide by n_j, the total population in category j, to get the per-susceptible infection rate:

$$\text{Betas}_{ij} = \beta^* \cdot (\text{relative proclivity of } i) \cdot (\#\text{ contacts between } i \text{ and } j \text{ per time unit}) \cdot \frac{1}{n_j}.$$

In full transparency, the relative proclivities and number of contacts between groups had to be roughly estimated for our scenario. Estimating these values typically results in a healthy discussion on data-gathering techniques. For example, we estimated that the transmission rate between non-compliant students was more than twice the transmission rate between compliant students, based on their relative proclivity and more frequent close contacts. Note also that, when estimating the number of contacts between individuals, these values should be *relative to the same quantities in the total population* (LA County, in our case).

3.4 Numerical Simulations

With the established models, we considered three main scenarios for how individuals might behave when on campus. We labeled these: Realistic Goal, Students Misbehave, and Everyone Misbehaves as seen in Table 5. For each scenario, we are scaling the initial conditions for all variables, the amount of outside interaction (beyond the Claremont Colleges), and specific model parameters. The initial conditions reflect the state of incoming students and faculty, which could be managed with testing before coming to campus. The amount of outside interaction could be varied with healthcare campaigns along with shutdown restrictions. The other parameters we varied corresponded to student compliance, which could be affected by administrative action and policies.

		Realistic Goal	**Students Misbehave**	**Everyone Misbehaves**
Initial Conditions	$E_{AT/HR}$	0	0	0.015
	$R_{AT/HR}$	0.01	0.01	0.05
	$E_{AT/LR}$	0.002	0.002	0.015
	$R_{AT/LR}$	0.05	0.05	0.05
	$E_{HDG/HR}$	0	0	0.015
	$R_{HDG/HR}$	0.01	0.01	0.05
	$E_{HDG/LR}$	0.002	0.002	0.015
	$R_{HDG/HR}$	0.05	0.05	0.05
	$E_{S/C}$	0.015	0.015	0.015
	$R_{S/C}$	0.08	0.08	0.05
	$E_{S/NC}$	0.015	0.015	0.015
	$R_{S/NC}$	0.08	0.08	0.05
Outside Interactions	$g_{AT/HR}$	0	0	0.12
	$g_{AT/LR}$	1/7	1/7	1
	$g_{HDG/HR}$	0	0	0.12
	$g_{HDG/LR}$	1/7	1/7	1
	$g_{S/C}$	1/7	3/7	1
	$g_{S/NC}$	3/7	10/7	2
Model Parameters	$prop_{S/NC}$	0.25	0.75	0.5
	κ	0.8	0.8	0.5

Table 5: Here, we list starting values used to simulate three different hypothetical scenarios. In these numerical simulations, we vary initial conditions and parameters in the model. Note, the notation here is to reflect the different subgroups, see Figure 3. For instance, $E_{AT/HR}$ is the proportion of the population in the exposed class for the Administration and Teaching subgroup who are also high-risk. Meanwhile, $g_{S/C}$ changes the outside contact rate for students who are compliant with healthcare protocols. The term $prop_{NC}$ denotes changing the proportion of students who are non-compliant in the population of students. Using our app, results in these three scenarios are depicted in Figures 12, 13 and 14.

In Table 5, for the Realistic Goal, we provide values describing a scenario where the population is following healthcare protocols and the number of exposed individuals on the campus at the start of the semester is very low. Meanwhile, in Students Misbehave, we highlight what might happen if more of the student population is non-compliant with healthcare protocols through initial conditions, outside exposures, and the amount of the student population who are non-compliant. Lastly, Everyone Misbehaves allows for the scenario if the total population is less compliant, resulting in higher initial conditions of exposed, greater outside interactions, along with students being more non-compliant. Naturally, these scenarios do not capture all possibilities, but rather serve as a starting point into investigation of what could happen during a semester.

3.4.1 Model 1 College COVID 5C wih No Vaccination

For the model described by System 2.1, we explore the three scenarios outlined in Table 5: Realistic Goal, Students Misbehave, Everyone Misbehaves. The parameters for our simulation come from Table 3, specifically the dates from 4/20-5/16/20 when we first modeled these scenarios, the results plotted in Figure 9.

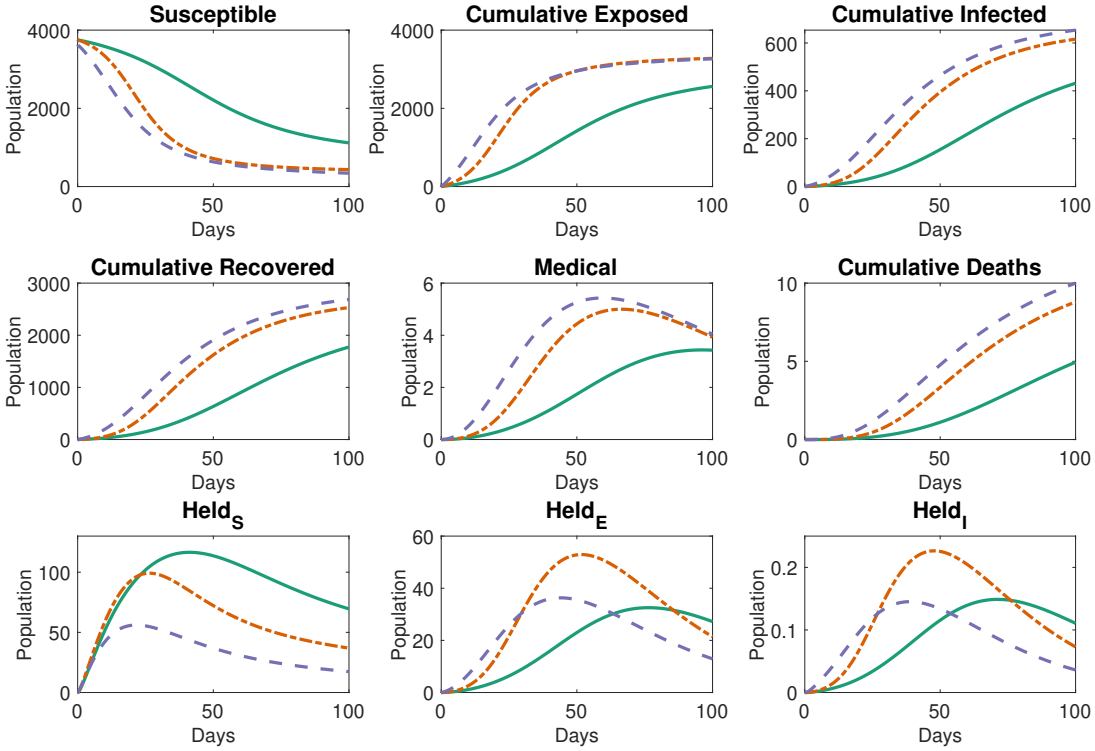

Figure 9: Simulations showing nine state variables over time under the three scenarios described in Table 5, with no vaccination. Green represents our Realistic Goal. Blue represents the situation where only students misbehave, and red shows the scenario Everyone Misbehaves. These simulations are run over 100 days, which is approximately the length of one semester, intended to illustrate what might happen under various scenarios if the colleges were to open during the Fall semester.

In Figure 9, by day 50, we see that the difference between the number of Cumulative Recovered individuals in the Realistic Goal scenario and the Everyone Misbehaves is over 1000 students. Also, if we aim for the Realistic Goal scenario, over 2000 individuals will not get COVID. Additionally, the held classes vary with the three scenarios reflecting different management options. Specifically, more susceptible individuals are held with the Realistic Goal scenario, while for the two misbehaving scenarios, we hold more exposed and infected individuals. We encourage future simulations and suggestions for different numerical experiments in Section 4.

To further investigate options such as the effectiveness of contact tracing, we started with the Realistic Goal scenario and then scaled the parameter for contact tracing. In Figure 10, we scaled the parameter κ, the contact tracing effectiveness parameter, from 0 (no contact tracing) to 1 (maximum contact tracing) by increments of 0.1.

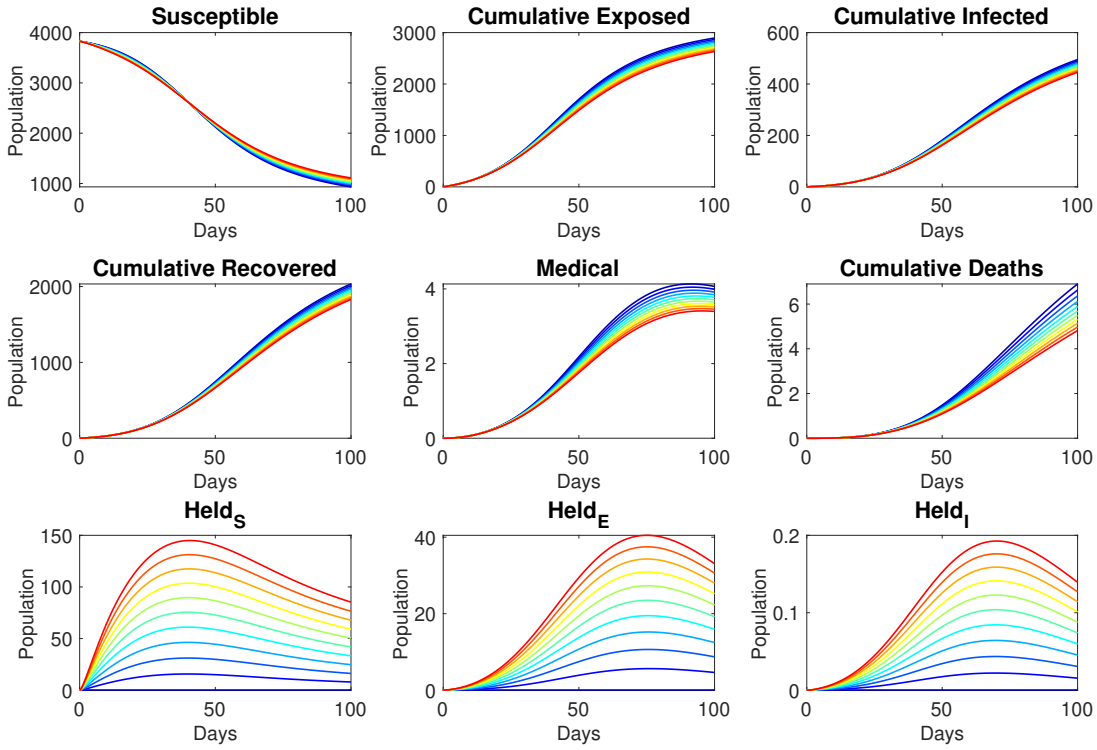

Figure 10: Simulations showing nine state variables over time, where the contact tracing parameter, κ, is multiplied by a factor ranging from 0 to 1 in increments of 0.1, with no vaccination. Dark purple represents multiplying the contact tracing parameter by a factor of 0 (no contact tracing), and red represents multiplying the contact tracing parameter by 1 (very effective contact tracing). As expected, the populations in the Held classes increases as κ increases, while Infected, Deaths and Recovered classes decrease. Other parameters are as listed in the "Realistic Goal" column of Table 5. These simulations are run over 100 days, which is approximately the length of one semester, intended to illustrate what might happen under various scenarios if the colleges were to open during the Fall semester.

3.4.2 Model 2 College COVID 5C with Vaccination

For the model described by System 2.4, we explore the scenarios outlined in Table 5. The parameters for our simulation come from Table 4, specifically the (8/1/2021-9/1/2021), we chose this time range as it is right when our semester started. We encourage future simulations and give suggestions for different numerical experiments in Section 4. We simulate the Realistic Goal scenario and then explore the effect of vaccination, seen in Figure 11.

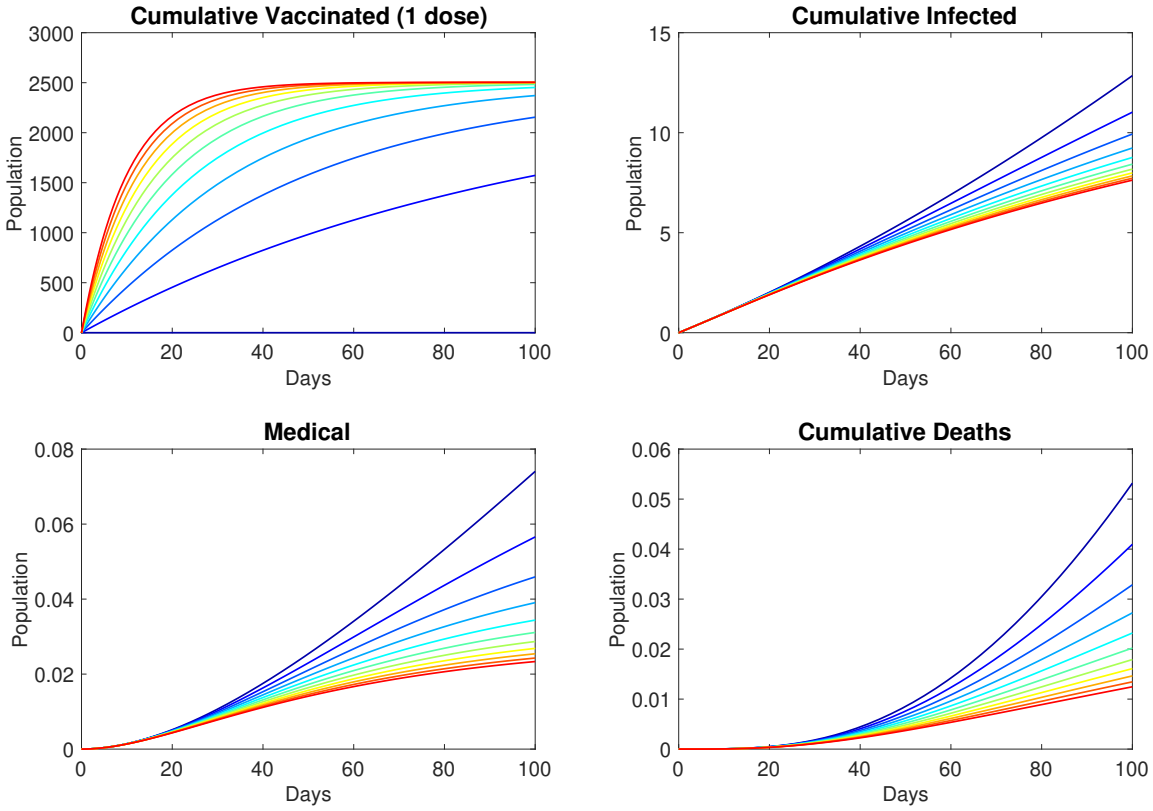

Figure 11: Simulations of Realistic Goal after vaccination where the vaccination rate for the first dose, ξ_1, is multiplied by a factor sweeping from 0 to 100 in increments of 10. Dark purple represents multiplying the vaccination rate by a factor of 0, and red represents multiplying the vaccination rate by 100. Note that in the upper left panel showing Cumulative Vaccinated, the dark purple graph is a horizontal line at zero, since nobody is vaccinated. While the number of vaccinated individuals increases as ξ_1 increases, the Infected, Medical and Death classes decrease.

3.5 Model App Analysis

A key part of mathematical modeling research is communicating findings to the appropriate audience. For our project, the initial goal was to give insight into opening the

Claremont Colleges for Fall 2020 to in-person instruction, meaning we wanted to communicate with the administration at our colleges. To aid with displaying possible scenarios, we decided to build an interactive app, which would allow the user to change initial conditions, parameters, and scenarios easily with an interactive interface. By varying parameter values, we could display a variety of scenarios for administrators so they could consider different management options. During the meetings, we had an open dialogue about the assumptions and limitations of the model, along with the potential insights into the pandemic on the college campuses. An important part of the limitations we discussed were realistic ranges for parameter values and what outcomes were feasible.

We used MATLAB® to build the app, and screenshots of the interface are shown in Figure 12. In the Settings panel on the left, the user can modify some key model parameters and initial conditions for all subgroups to see their effects on outcomes. In particular, a user can scale the number of infected contacts, along with adjusting for the subgroup the percent non-compliant, tracing effectiveness, asymptomatic infectiousness, and the transmission rate (β). In this figure, we see the settings that reflect the Realistic Scenario proposed in Section 3.4.1.

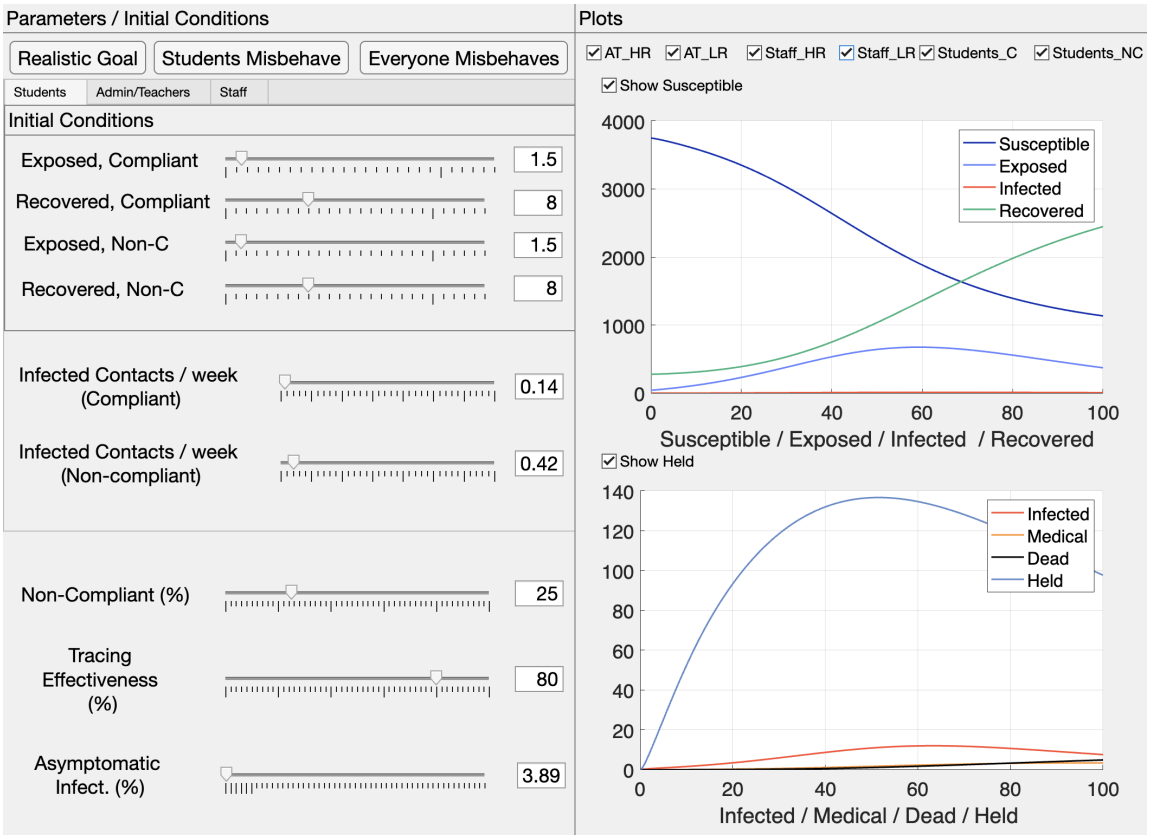

Figure 12: Screenshot of our app for the Realistic Scenario values from Table 5, with no vaccination.

As the values are changed, there are two plots on the right that react to the changes

showing the values of the state variables. By checking appropriate boxes, the user can select which variables to view. The upper plot allows the user to select variables from the SEIR compartments in each of the subgroups. We added a toggle for the susceptible class since that group is often so much larger than the others. Meanwhile, the lower plot displays the IMDH classes, where a user can toggle on or off a display of the held classes.

When talking to College Administrators, we were able to use the app to quickly show the consequences of different scenarios. For example, Figure 13 compares the consequences of the Realistic scenario to the Students Misbehave scenario (the relevant initial values and parameter values are given in Table 5). By using the check-boxes, we can show what happens to those in the high risk classes only. On the left, the model predicts that, in the Realistic scenario, four people in the high risk class, on average, would die in 100 days (approximately one semester). By contrast, in the Students Misbehave scenario, where the proportion of non-compliant students is increased from .25 to .75, the number of predicted deaths increases to slightly over 7 in a semester (note that the app automatically scales the vertical axes).

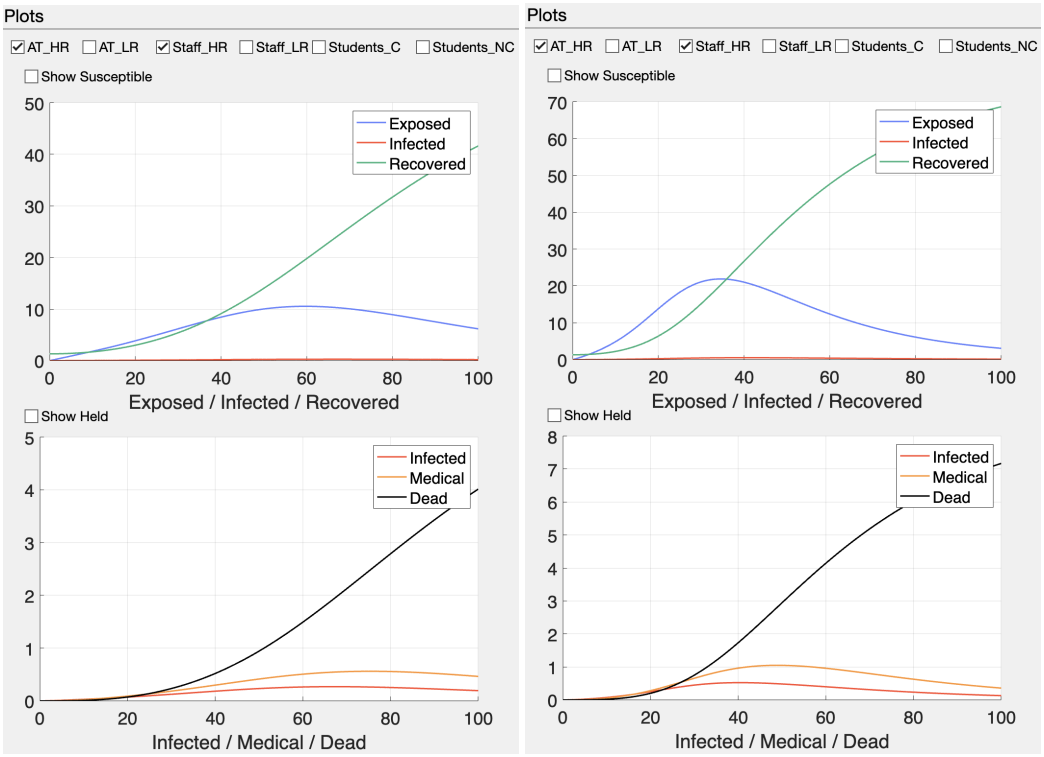

Figure 13: Screenshots of our app for the Realistic (left) and Students Misbehave (right) scenarios, with no vaccination. We focus on the High Risk individuals only and plot the number recovered (top graph in upper panels) and those who pass away (top graph in lower panels) in each scenario. We see that deaths double from the Realistic to Students Misbehave scenarios.

To see the impact of contact tracing and quarantines on dormitory capacities, we can

check the display choices to show students in the Held classes. Figure 14 compares the use of the “Show Susceptible" and “Show Held" categories to aid in visualisation. In these plots we show the results of the Everyone Misbehaves scenario, described in the last column of Table 5. In contrast to Figure 13, here we have selected all the low risk classes at the top of the panels. On the left two panels, we have checked “Show Susceptibles" in the top graph, and “Show Held" in the bottom graph. In the lower left panel, we see that in this scenario the model predicts that there will be nearly 70 low risk individuals in the Held class when it reaches its maximum at around Day 30 (one month into the semester). In the lower right panel, we have unchecked the “Show Held" box to better observe the predictions for the number of infected individuals in the low risk class (over 20 at the semester’s peak), and the number of predicted deaths among this class.

In our discussions with College administrators, these predictions were critical in the decision-making process. During the summer of 2022, we worked on implementing the app for the vaccination model from Section 2.4.

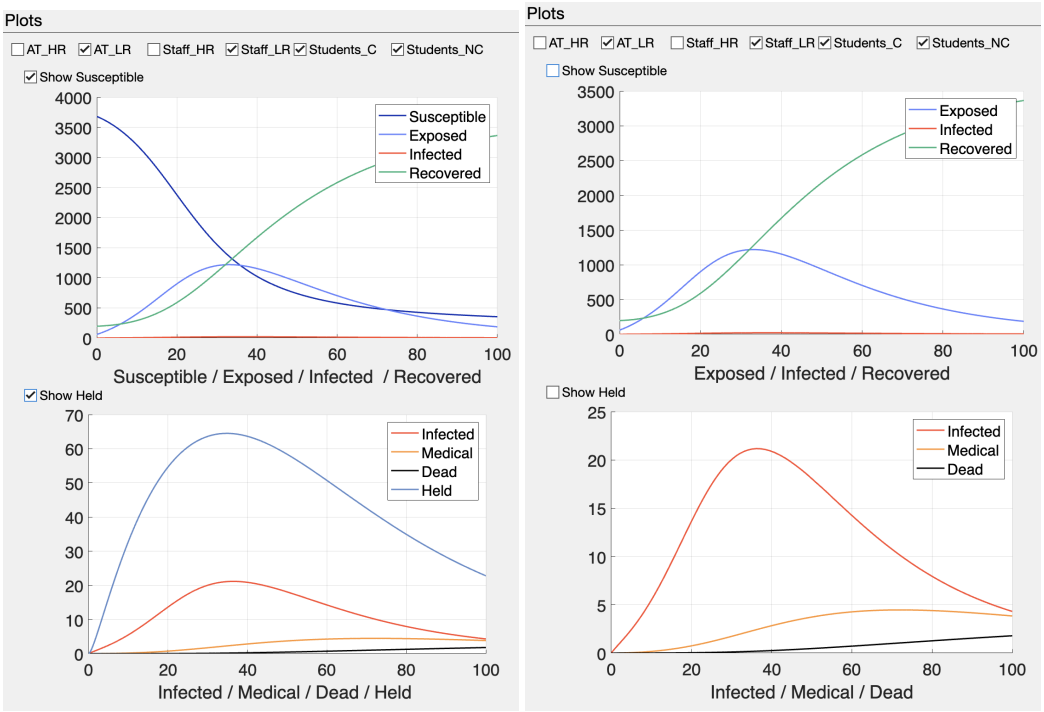

Figure 14: Screenshots of our app for the Everyone Misbehave scenario, with no vaccination. The left has the boxes “Show Susceptible" and “Show Held" checked, while on the right these categories are unchecked. For this display, we focus on the Low Risk individuals by checking only the “LR" boxes at the top.

4 Discussion

4.1 Results and Conclusions

Through the models presented in Sections 2.1, 2.2, 2.3, and 2.4 we examined the dynamics of the COVID-19 pandemic, with and without vaccines and at the LA County level and residential college level of the Claremont Colleges. The first step we took in analysis was to parameterize the models in Sections 2.1 and 2.3 using data from LA County [3, 5, 4]. For the model in Section 2.1 our fits for four time periods resulted in Figure 7. We can see the model fits the data trends for the time periods capturing for instance the changing behavior in the number of hospitalizations data, also reflected in the parameter values in Table 3. In Figure 7, we can see that in subplots (a.)-(c.) all capture the hospitalization data clearly with the model, but subplot (d.) captures the trend with noisy data and fits for a longer time period. We mentioned that for our model we need to fit to shorter time periods due to the deterministic nature. If we use the ODE model format for longer times, we lose accuracy in the model capturing data dynamics as seen in Figure 15 (a.). In Figure 7 we break the time period from August to November into two fits, but in Figure 15 we display the fit if we try and use all the data for one model parameter estimation.

Similarly, we fit the model in Section 2.3 to four data sets with resulting Figure 8. As with the previous fits, we can see the model is capturing the overall dynamics, especially the changes in the hospitalizations for the different time periods, reflected in Table 4 as well. An interesting plot in Figure 8 is hospitalizations for subplot (a.), where the data and model capture the changing number of hospitalizations with an inflection point. Note in subplot (d.) the end of the time frame has some changing dynamics which the model smooths with it's deterministic nature. When considering longer time frames with the vaccination data, we ran into a number of issues, we can see the plot for fitting September to November in Figure 15 (b.). We notice the model includes more cumulative cases than the data, along with the data discrepancy for the cumulative deaths. The data for cumulative deaths was retroactively changed during the pandemic, leading to the non-monotonic data for a cumulative value. Figure 15 highlights the importance of a modeler examining the data critically along with the model, to ensure the disease dynamics are captured. When partitioning data, we considered regulation changes throughout the pandemic, along with social changes. Overall, our parameter estimation results for models from Sections 2.1 and 2.3 resulting in Figures 7 and 8 demonstrated our model capturing the disease dynamics for the COVID-19 pandemic time periods and allowed us to move on to numerical simulations in Section 3.4.

When we scaled the contact tracing parameter κ in Section 3.4.1, Figure 10 shows that increasing contact tracing decreases the number of Cumulative Exposed by about 500 individuals and increases the number of Held individuals. That is, the college administration would need to weigh decreasing the number of infected individuals with the housing needs of Held individuals on campus. The figure also shows that the number of cumulative deaths does decrease with contact tracing, albeit by only one or two individuals.

When we scaled the vaccination rate ξ_1 in Section 3.4.2, Figure 11 depicts a very different time in the COVID epidemic. Here, the parameters that determine the spread and severity of the disease differ from the previous time periods, before vaccines were

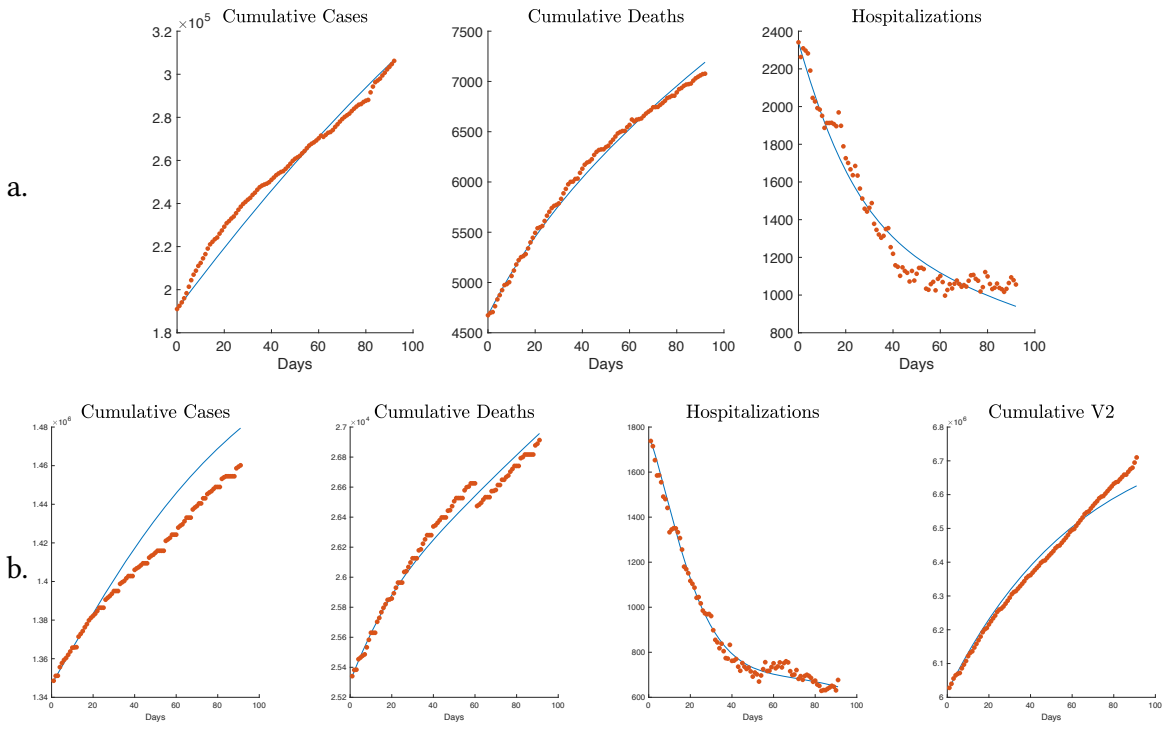

Figure 15: The parameters of the COVID model are not constant in time, as discussed above, with no vaccination. We illustrate this in the two graphs in this figure. The (a.) panel shows fits to three months of data from early in the pandemic, August 1 - November 1, 2020, with $J_2 = 0.8949$. The month of September 2020 shows a clear difference in hospitalization rates. The (b.) panel shows fits to data from September 1 to November 30, 2021, $J_2 = 1.5239$. In this case, we see a discontinuity in the data on cumulative deaths around November 1 and changes in the rate of new cases and hospitalizations. Note that the scaled residuals are significantly larger for these two fits than for the ones presented in Tables 3 and 4, an indication that parameter values might be changing over time.

available. For example, compare the first column of Table 4 that contain the parameters used in this simulation (just before the Fall 2021 semester) with the first column of Table 3, which were fit to data from late spring, 2020. The base transmission rate, β, was five times higher in Spring 2020, and both the mortality rate, μ_I, and the hospitalization rate, ω_I, were an order of magnitude larger. This could be due to many factors, including increased adherence to safety measures, the development of anti-viral treatments, and a better knowledge of how to prevent and treat the disease in general.

Our plots show that vaccination does decrease the number of COVID cases (Cumulative Infected individuals) as well as hospitalizations (Medical) and deaths. However, the number of individuals in each category is small, even without vaccination, i.e. when $\xi_1 = 0$. Since the numbers shown are predicted averages, we can interpret the predictions as probabilities: without vaccination we predict an average of approximately thirteen symptomatic COVID cases (assuming the other parameters are as listed in Table 4 and in

the “Realistic Goal" column of Table 5). However, if vaccinations are acquired quickly - as was the case on our campuses where they were required with the first few weeks of school, the average number of symptomatic COVID cases goes down to 6 or 7.

4.2 Future Expansions

Many extensions and modifications to our models are possible. One modification is to implement the models at other residential colleges or other residential communities. We outline how one can develop models with various classes and new information such as the evolution of the model in Section 2.2 to the model in Section 2.4. When including vaccines in the model, we did make key assumptions, similar to when we made the base model, which allow for alterations to the model to address different questions.

For instance, in the base model in Section 2.1, we include the infection of asymptomatic individuals through a parameter value with transmission, but this could also be accomplished with an asymptomatic individual variable. Furthermore, when considering the outside community transmission in the Section 2.2 model, we used a parameter value, but this interaction could be it’s own SEIR model as well coupled to our residential community. These changes would explore different dynamics and allow the modeler to examine new questions and analyses.

For the models including vaccination in Sections 2.3 and 2.4, we made key assumptions about how to include the COVID-19 vaccines into our model. For instance, when expanding to the Claremont Colleges, we removed the variable held classes and included them intrinsically in the model through varying parameter values. For instance, allowing E to be a held variable and I to be quarantined by changing parameter values. An extension would be to include 6 held variables coupled with the S, V_1, V_2, E_1, E_2 variables. Alternatively, a modeler could include a variable for the infected individual who are vaccinated, I_V, So E_1 and E_2 would then transition to I_V rather than directly to I. There structural changes reflect different options for incorporating the vaccines and focus on different problems during the pandemic analysis.

A potential numerical extension to Sections 3.4.1 and 3.4.2 would be to complete more local sensitivity analysis, along with in depth sensitivity and uncertainty analysis. In our presented numerical simulations in Section 3.4, we explore varying contact tracing for Section 3.4.1 and varying vaccination rate for Section 3.4.2. There are various other parameters of interest in these initial models, and an in depth sensitivity analysis would allow a modeler to ascertain both potential outcomes and which parameters are key to outcomes and need precision in their measurements. Furthermore, our numerical simulations utilize one set of parameters from our parameter estimation in Tables 3 and 4. A potential numerical extension would be to implement other parameter combinations to explore different stages of the pandemic, along with potential management strategies. Additionally, for Section 3.5 with the app development, an extension would be streamlining the app process and implementing additional parameter toggles with different plotting outcomes or measurements.

4.3 For Instructors and Students

Mathematical modeling took a front line during the COVID epidemic, and this provided fertile ground for teaching the impact of mathematical models on decision-making. Mathematical modeling can be useful to inform other decisions on a college campus, and we hope that the ideas here can be adapted to the spread of other diseases and infections. For this COVID study, we gathered data from our institutions as well as from the County. This sharing of data is an important step in the model development and calibration process.

Here are some reflections on how we used modeling efforts to engage students. While we focus on COVID-19, our insights can be utilized to model other pandemics and epidemics, along with other applications.

Parameter fitting Mathematical models are often used to make predictions. In order to make these predictions meaningful, the relevant situation. In our case, we needed to use available data from the county and estimates of inter-group behaviors to predict outcomes on our campus. In general, this is a good exercise for students of modeling, who need to consider the following questions:

1. What data is available?
2. How do we use these data to estimate model parameters?
3. How can we infer missing parameter values?
4. How can we quantify the uncertainty in our model predictions?

We typically provide students with some fitting routines and a sample data set, and encourage them to find current and relevant data to use for their own model-fitting. Model sensitivity and uncertainty can be quantified using a variety of analysis techniques. Initial quantification can be done by testing the local sensitivity of varying values with respect to specific outcomes. For more in depth analysis, there are multiple tutorials and resources we share with students [28, 12, 23, 14, 1].

Model refinement Any model we present to our students will be a simplification of reality. As we gain understanding from theoretical analysis and observations, we are often led to add components to the model. For example, during the COVID pandemic, we saw interventions introduced (masking, vaccinations, anti-viral medications) that could be important to include in the model. Providing students with a base model that they can refine to include emerging treatments and new insights encourages them to think about the process of model development as well as how to critically compare several models.

Communication of results This particular modeling effort was initiated in order to help college administrators make an informed decision about reopening our campus. A critical part of this effort was to communicate what we learned from the simulations to college leadership. Students can appreciate this, and can learn to use a mathematical model to effectively tell a story by giving talks to (fictitious or real) stake-holders, or to make interactive apps that non-mathematicians can use to explore a variety of virtual scenarios.

For our particular purpose, we used the MATLAB® App Designer tool. This allowed us to create something that an administrator who is not familiar with modeling or pro-

gramming can use to test different scenarios by manipulating parameters. The user can adjust parameters using sliders, and the app outputs graphs that can be interpreted by non-mathematicians. A note of caution: designing an app is certainly an extra level of programming, and if this is part of a project, students will need to allow time for the implementation. There are also other tools for designing interactive simulations that might have a shallower learning curve and that use open source software. For example, web apps can be written in Shiny [35], which interfaces with either R or Python.

Potential for expansion to research projects Some model refinements and explorations can go beyond what can be done in a course. Several research projects have evolved from the current model, including statistical analysis of the robustness of the model under perturbations and the introduction of vaccines, the emergence of new variants of the virus, and waning immunity. In fact, a search using the keywords "mathematical model COVID pandemic" yields over 250,000 results. After learning the basic COVID model, students can explore other research questions through this extensive literature and/or by developing their own model extensions.

Acknowledgements

Thanks to Ezra Buchla for help implementing the app. EYL was supported by CMC SIE program for summer 2021. CJE was supported by the AMS-Simons Travel Grants, which are administered by the American Mathematical Society with support from the Simons Foundation.

References

[1] Alison Adams, Quiyana M Murphy, Owen P Dougherty, Aubrey M Sawyer, Fan Bai, Christina J Edholm, Evan P Williams, Linda JS Allen, and Colleen B Jonsson. Data-driven models for replication kinetics of Orthohantavirus infections. *Mathematical Biosciences*, 349:108834, 2022.

[2] Government Operations Agency. Statewide COVID-19 Vaccines Administered By County. https://data.ca.gov/dataset/covid-19-vaccine-progress-dashboard-data/resource/c020ef6b-2116-4775-b11d-9df2875096ab, 2020-2023. Accessed: 2022-01-03.

[3] Government Operations Agency. COVID-19 time-series metrics by county and state. https://data.ca.gov/dataset/covid-19-time-series-metrics-by-county-and-state, 2020-2023. Accessed: 2023-05-31.

[4] Government Operations Agency. COVID-19 post-vaccination infection data (archived). https://data.ca.gov/dataset/covid-19-post-vaccination-infection-data-archived, 2020-2023. Accessed: 2023-05-31.

[5] Government Operations Agency. COVID-19 Vaccine Progress Dashboard Data. https://data.ca.gov/dataset/covid-19-vaccine-progress-dashboard-data, 2020-2023. Accessed: 2023-05-31.

[6] Riti Bahl, Nicole Eikmeier, Alexandra Fraser, Matthew Junge, Felicia Keesing, Kukai Nakahata, and Lily Reeves. Modeling COVID-19 spread in small colleges. *Plos one*, 16(8):e0255654, 2021.

[7] Molly Borowiak, Fayfay Ning, Justin Pei, Sarah Zhao, Hwai-Ray Tung, and Rick Durrett. Controlling the spread of COVID-19 on college campuses. *Mathematical Biosciences and Engineering*, 18(1):551–563, 2021. ISSN 1551-0018. doi: 10.3934/mbe.2021030. URL https://www.aimspress.com/article/doi/10.3934/mbe.2021030.

[8] Fred Brauer. Mathematical epidemiology: Past, present, and future. *Infectious Disease Modelling*, 2(2):113–127, 2017.

[9] Danielle Burton, Suzanne Lenhart, Christina J Edholm, Benjamin Levy, Michael L Washington, Bradford R Greening Jr, KA Jane White, Edward Lungu, Obias Chimbola, Moatlhodi Kgosimore, et al. A mathematical model of contact tracing during the 2014–2016 West African Ebola outbreak. *Mathematics*, 9(6):608, 2021.

[10] Robert E Davis, Manoj Sharma, Kayla E Simon, and Amanda H Wilkerson. Conceptualization of college students' COVID-19 related mask-wearing behaviors using the multi-theory model of health behavior change. *Health promotion perspectives*, 11(2):194, 2021.

[11] Gerda De Vries, Thomas Hillen, Mark Lewis, Johannes Müller, and Birgitt Schönfisch. *A course in mathematical biology: quantitative modeling with mathematical and computational methods*. SIAM, 2006.

[12] An Dela, Blerta Shtylla, and Lisette de Pillis. Multi-method global sensitivity analysis of mathematical models. *Journal of Theoretical Biology*, 546:111159, 2022.

[13] Kathleen Dooling, Mona Marin, Megan Wallace, and et al. The advisory committee on immunization practices' updated interim recommendation for allocation of COVID-19 vaccine — United States, December 2020. *MMWR Morb Mortal Wkly Rep*, 69, 2021.

[14] Christina Edholm, Benjamin Levy, Ash Abebe, Theresia Marijani, Scott Le Fevre, Suzanne Lenhart, Abdul-Aziz Yakubu, and Farai Nyabadza. A risk-structured mathematical model of Buruli ulcer disease in Ghana. *Mathematics of Planet Earth: Protecting Our Planet, Learning from the Past, Safeguarding for the Future*, pages 109–128, 2019.

[15] Christina J Edholm, Benjamin Levy, Lee Spence, Folashade B Agusto, Faraimunashe Chirove, C Williams Chukwu, David Goldsman, Moatlhodi Kgosimore, Innocent

Maposa, KA Jane White, et al. A vaccination model for COVID-19 in Gauteng, South Africa. *Infectious Disease Modelling*, 7(3):333–345, 2022.

[16] Centers for Disease Control and prevention. How CDC is making COVID-19 vaccine recommendations. https://www.cdc.gov/coronavirus/2019-ncov/vaccines/recommendations-process.html, 2020-2021. Accessed: 2021-06-15.

[17] Centers for Disease Control and prevention. COVID-19. https://www.cdc.gov/coronavirus/2019-ncov/index.html, 2020-2023. Accessed: 2023-05-31.

[18] Peter I Frazier, J Massey Cashore, Ning Duan, Shane G Henderson, Alyf Janmohamed, Brian Liu, David B Shmoys, Jiayue Wan, and Yujia Zhang. Modeling for COVID-19 college reopening decisions: Cornell, a case study. *Proceedings of the National Academy of Sciences*, 119(2):e2112532119, 2022.

[19] Karen KL Hwang, Christina J Edholm, Omar Saucedo, Linda JS Allen, and Nika Shakiba. A hybrid epidemic model to explore stochasticity in COVID-19 dynamics. *Bulletin of Mathematical Biology*, 84(9):91, 2022.

[20] The MathWorks Inc. MATLAB version: 9.9.0 (r2020b), 2020. URL https://www.mathworks.com.

[21] Benjamin Ivorra, Miriam R Ferrández, María Vela-Pérez, and Angel Manuel Ramos. Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. the case of China. *Communications in nonlinear science and numerical simulation*, 88:105303, 2020.

[22] Mujeeb Khan, Syed F Adil, Hamad Z Alkhathlan, Muhammad N Tahir, Sadia Saif, Merajuddin Khan, and Shams T Khan. COVID-19: a global challenge with old history, epidemiology and progress so far. *Molecules*, 26(1):39, 2020.

[23] Benjamin Levy, Christina Edholm, Orou Gaoue, Roselyn Kaondera-Shava, Moatlhodi Kgosimore, Suzanne Lenhart, Benjamin Lephodisa, Edward Lungu, Theresia Marijani, and Farai Nyabadza. Modeling the role of public health education in Ebola virus disease outbreaks in Sudan. *Infectious Disease Modelling*, 2(3):323–340, 2017.

[24] Chuan-Yao Li and Jie Yin. A pedestrian-based model for simulating COVID-19 transmission on college campus. *Transportmetrica A: Transport Science*, 19(1): 2005182, 2023.

[25] Reia Li, Ruth Efe, and Zintan Mwinila-Yuori. COVID-19, crowdedness, and CMC dining: An agent-based model approach to reducing the spread of COVID-19. *SIURO*, 15, 2022.

[26] Eric Lofgren, Kristian Lum, Aaron Horowitz, Brooke Madubuonwu, Kellen Myers, and Nina H Fefferman. The epidemiological implications of jails for community, corrections officer, and incarcerated population risks from COVID-19. *medRxiv*, pages 2020–04, 2021.

[27] Eric T Lofgren, Kristian Lum, Aaron Horowitz, Brooke Mabubuonwu, Kellen Meyers, and Nina H Fefferman. Carceral amplification of COVID-19: Impacts for community, corrections officer, and incarcerated population risks. *Epidemiology (Cambridge, Mass.)*, 33(4):480, 2022.

[28] Simeone Marino, Ian B Hogue, Christian J Ray, and Denise E Kirschner. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of theoretical biology*, 254(1):178–196, 2008.

[29] Oren Milman, Idan Yelin, Noga Aharony, Rachel Katz, Esma Herzel, Amir Ben-Tov, Jacob Kuint, Sivan Gazit, Gabriel Chodick, Tal Patalon, et al. Community-level evidence for SARS-CoV-2 vaccine protection of unvaccinated individuals. *Nature Medicine*, pages 1–3, 2021.

[30] James D. Murray. *Mathematical Biology I. An Introduction*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer, New York, 3 edition, 2002. doi: 10.1007/b98868.

[31] NIMBioS Webinar Series - Nina Fefferman. The role of applied math in real-time pandemic response: How basic disease models work. https://www.youtube.com/watch?v=Ewuo_2pzNNw&feature=youtu.be, 2020. Accessed: 2023-10-06.

[32] Amy Nivette, Denis Ribeaud, Aja Murray, Annekatrin Steinhoff, Laura Bechtiger, Urs Hepp, Lilly Shanahan, and Manuel Eisner. Non-compliance with COVID-19-related public health measures among young adults in Switzerland: Insights from a longitudinal cohort study. *Social science & medicine*, 268:113370, 2021.

[33] County of Los Angeles Public Health. COVID19 dashboard. http://publichealth.lacounty.gov/media/Coronavirus/, 2020. Accessed: 2020-07-04.

[34] Tamara Pilishvili, Katherine E Fleming-Dutra, Jennifer L Farrar, Ryan Gierke, Nicholas M Mohr, David A Talan, Anusha Krishnadasan, Karisa K Harland, Howard A Smithline, Peter C Hou, et al. Interim estimates of vaccine effectiveness of Pfizer-BioNTech and Moderna COVID-19 vaccines among health care personnel—33 us sites, January–March 2021. *Morbidity and Mortality Weekly Report*, 70(20):753, 2021.

[35] Posit.co. Easy web apps for data science without the compromises. https://shiny.posit.co, Accessed 2023-07-10.

[36] Jagajeet Prasad Singh, Anshuman Sewda, and Dutt Gupta Shiv. Assessing the knowledge, attitude and practices of students regarding the COVID-19 pandemic. *Journal of Health Management*, 22(2):281–290, 2020.

[37] Emma Pritchard, Philippa C Matthews, Nicole Stoesser, David W Eyre, Owen Gethings, Karina-Doris Vihta, Joel Jones, Thomas House, Harper VanSteenHouse,

Iain Bell, et al. Impact of vaccination on new SARS-CoV-2 infections in the United Kingdom. *Nature Medicine*, pages 1–9, 2021.

[38] Weston C Roda, Marie B Varughese, Donglin Han, and Michael Y Li. Why is it difficult to accurately predict the COVID-19 epidemic? *Infectious disease modelling*, 5:271–281, 2020.

[39] LA Times. Tracking coronavirus vaccinations in California. https://www.latimes.com/projects/california-coronavirus-cases-tracking-outbreak/covid-19-vaccines-distribution/#company-comparison, 2020-2023. Accessed: 2021-07-30.

[40] LA Times. The Los Angeles Times' open-source archive of California coronavirus data. https://github.com/datadesk/california-coronavirus-data, 2020-2023. Accessed: 2023-03-27.

[41] Huwen Wang, Zezhou Wang, Yinqiao Dong, Ruijie Chang, Chen Xu, Xiaoyue Yu, Shuxian Zhang, Lhakpa Tsamlag, Meili Shang, Jinyan Huang, et al. Phase-adjusted estimation of the number of coronavirus disease 2019 cases in Wuhan, China. *Cell discovery*, 6(1):10, 2020.

[42] Howard Weiss. The SIR model and the Foundations of Public Health. 2013. URL https://api.semanticscholar.org/CorpusID:69125770.

[43] Wuyue Yang, Dongyan Zhang, Liangrong Peng, Changjing Zhuge, and Liu Hong. Rational evaluation of various epidemic models based on the COVID-19 data of China. *Epidemics*, 37:100501, 2021.

[44] Dhanusha Yesudhas, Ambuj Srivastava, and M Michael Gromiha. COVID-19 outbreak: history, mechanism, transmission, structural studies and therapeutics. *Infection*, 49:199–213, 2021.