Modeling School Reopening Options in Response to COVID-19

Nikhil Shanbhag[1] and Anthony Grabowski[1]

[1]Archbishop Mitty High School, San Jose, California, USA

ABSTRACT

School administrators have the daunting challenge of rolling out school reopening policies and safety standards. While it is important to have children in school to maximize learning, it is critical to ensure that it is done in a manner that prioritizes the safety of staff and students on campus. This research explores models that can be used to examine the spread of COVID-19, and then applies them to evaluate school reopening policies within the United States. Our findings highlight that maintaining a ratio with a higher rate of recovery compared to the rate of infection, also known as the basic reproduction number, is the key to reopening schools. We outline practices to enable this outcome, such as reducing class sizes and quarantining students within 2 or 3 days if they test positive for COVID-19. We show that when the basic reproduction number is greater than 1, the proportion of the population never infected drops rapidly, and we also quantify the impact of quarantining on the basic reproduction number.

Introduction

We analyze the issue of reopening schools utilizing mathematical models. Specifically, we first analyze the Susceptible-Infected-Recovered (SIR) model and then introduce the new scenario of quarantining students to analyze the Susceptible-Infected-Quarantined-Recovered (SIQR) model, both of which already exist. They each help us compare the rate of infection and the rate of recovery in order to derive estimations for the value of the basic reproduction number, but the SIQR model specifically helps us understand why quarantining is important in a population.

A variety of complex factors need to be accounted for, including the student population, the rate at which students are infected, the rate at which students are quarantined, and the proportion of students that are susceptible to COVID-19. Our research seeks to simplify the complexity by expressing multiple quantities using one variable known as the basic reproduction number, R_0.

R_0 has been used to estimate the number of new infected individuals generated from one infected individual in a population where all individuals are susceptible [1]. We adapt R_0 for use as our indicator of how rapidly COVID-19 is spreading throughout a population.

Van den Driessche and Watmough [2] explain the following threshold for the basic reproduction number, R_0:

"If $R_0 < 1$, then on average an infected individual produces less than one new infected individual over the course of its infectious period, and the infection cannot grow."
"If $R_0 > 1$, then each infected individual produce, on average, more than one new infection, and the disease can invade the population."

As such, the former scenario ($R_0 < 1$) is desired for school reopening and its implications are described throughout the course of this research. Our research analyzes the impact of multiple school reopening policies on R_0 and then proposes recommendations in order to limit the spread on campus. Additionally, we are able to quantify the impact of some of these policies such as quarantining on R_0.

Methods to Evaluate the Basic Reproduction Number (R_0)

The SIR and SIQR models can be used to evaluate the spread of COVID-19 and its impact on a given population. The SIR model does not account for quarantining as part of R_0 while the SIQR model accounts for a quarantined group. In this section, comparing the SIR and SIQR models enables us to understand the impact of quarantining individuals in a population versus not quarantining.

SIR Model

The SIR model is a compartmental model shown in Figure 1, and it is the most common type of epidemiology model used by scientists when they scrutinize the behavior of a pandemic [3].

In the SIR Model Diagram (Figure 1) below, let S represent the number of individuals susceptible to COVID-19 per unit area, I represent the number of infected individuals per unit area, and R represent the number of recovered individuals per unit area, where "per unit area" refers to the population density of an area. For example, the population density of New York City is a staggering 27,000 people per square mile [4] while the population density of Montana is a meager 7 people per square mile [5]. a is a constant that represents the rate at which susceptible people are being infected and b is a constant that represents the rate at which infected people are recovering. While it is not necessarily true that every infected person recovers, our model excludes deceased individuals because the death rate of COVID-19 is only 1.4% [6].

Figure 1. SIR Model Diagram. It shows the rate at which people transition from susceptible to infected to recovered.

Using the SIR model, we represent the rate of change (with respect to time) of the number of susceptible, infected, and recovered individuals using differential equations [3] as shown below.

Equation 1: Rate of Change of Susceptible Population for SIR Model
$$\frac{dS}{dt} = -aSI$$

Equation 2: Rate of Change of Infected Population for SIR Model
$$\frac{dI}{dt} = aSI - bI$$

Equation 3: Rate of Change of Recovered Population for SIR Model
$$\frac{dR}{dt} = bI$$

$\frac{dS}{dt}$ in Equation 1 represents the rate of change of the number of susceptible individuals in the population with respect to time, $\frac{dI}{dt}$ in Equation 2 represents the rate of change of the number of infected individuals in the population with respect to time, and $\frac{dR}{dt}$ in Equation 3 represents the rate of change of the number of recovered individuals in the

population with respect to time. The rate aSI accounts for the interactions between susceptible and infected people, because transitioning from a susceptible state to an infected state depends on the susceptible person's contact with other infected individuals in the population and the number of susceptible people. Additionally, the infected group is losing individuals at a rate of bI while the recovered group is gaining these individuals at the same rate. This rate is only proportional to I because infected individuals do not transition to the recovered stage through interactions with susceptible individuals; therefore, there is no need to account for the number of susceptible individuals in the population.

Nondimensionalized SIR Model

"Nondimensionalization" is a technique that makes comparison of data points simpler by converting absolute numbers (with units) into proportions (without units). We will now nondimensionalize Equation 1, 2, and 3, meaning that the values of $S, I,$ and R will be converted to produce $S^*, I^*,$ and R^* , which represent the proportion of susceptible, infected, and recovered people per unit area in the population respectively. The proportions simplify our analysis by allowing us to produce curves that will accurately show the impact of higher values of R_0.

In order to convert these values into proportions, we introduce a new variable N, which represents the total population density ($N = S + I + R$). $S^*, I^*,$ and R^* can be written accordingly: $S^* = \frac{S}{N}, I^* = \frac{I}{N},$ and $R^* = \frac{R}{N}$. Our nondimensional equations (Equation 4, 5, and 6) can be used in order to measure the rate of change of the susceptible, infected, and recovered proportions of the population with respect to time, t^*. The value $\frac{1}{b}$ is a time constant to nondimensionalize the differential equations, and it is equal to the average time to recover [Appendix; Part A].

Equation 4: Rate of Change of Susceptible Proportion for SIR Model
$$\frac{dS^*}{dt^*} = \frac{-aN}{b}S^*I^*$$

Equation 5: Rate of Change of Infected Proportion for SIR Model
$$\frac{dI^*}{dt^*} = \frac{aN}{b}S^*I^* - I^*$$

Equation 6: Rate of Change of Recovered Proportion for SIR Model
$$\frac{dR^*}{dt^*} = I^*$$

These equations are important because they universalize our analysis, meaning that they can be applied to any population of individuals, no matter how large or small the density.

R_0 compares the rate at which people are entering the infection group over the rate at which they are leaving the infection group, and can be estimated by the following ratio [7]:

$$R_0 = \frac{"rate\ of\ infection"}{"rate\ of\ recovery"}.$$

Therefore, in Equation 4 and Equation 5, we see that the value of R_0 is $\frac{aN}{b}$, valid at the beginning of the pandemic when the infected population is vanishingly small.

We ideally want Equation 5, which represents the rate of change of the infected proportion in the population, to have a value less than 0, because that means the curve will put us on a trajectory towards a non-pandemic state. This means we want $\frac{1}{R_0} > S^*$, as follows:

$$\frac{aN}{b} S^* I^* - I^* < 0$$

$$\frac{aN}{b} S^* < 1$$

$$S^* < \frac{b}{aN} = \frac{1}{R_0}$$

In other words, as the proportion of susceptible individuals in the population increases, the basic reproduction number of the virus should decrease.

Simulation of SIR Model Using Euler's Method

Euler's Method can be used to estimate the value of a variable using steps of equal size until there is no significant change in the value of the variable. We use this method to estimate the values of S^*, I^*, and R^* with respect to time in order to produce our curves for the simulation.

In this section, we simulate a scenario in which 95% of the population is initially susceptible while the other 5% is infected with COVID-19. We choose these percentages because at the beginning of a pandemic, almost the entire population is susceptible and only a few members are infected. The transmission of the virus depends on interactions between this largely susceptible pool of people who are not contagious and the few individuals who are contagious. We simulate three different scenarios in order to show how the pandemic is governed by the ratio R_0: (1) $R_0 > 1$; (2) $R_0 = 1$; and (3) $R_0 < 1$.

Figure 2 is the corresponding SIR curve, which compares S^*, I^*, and R^* versus t^* for $R_0 = 2$. It signifies that for every infected person, another two susceptible people are likely to be infected. In this case, the infected curve clearly demonstrates an early spike. The S^* and R^* curves also spike downward and upward respectively due to the initial exponential growth in the proportion of individuals infected. Eventually, the curves level off because there are fewer susceptible individuals that can be infected, and all the individuals who were infected have developed the necessary antibodies in order to transition to the "recovered" stage.
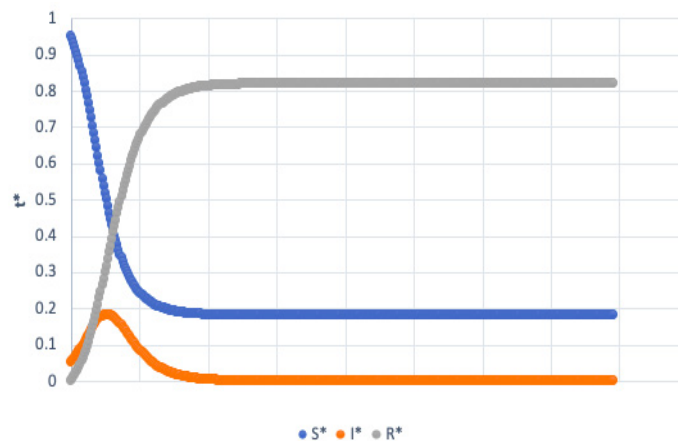

● S* ● I* ● R*

Figure 2. Sample SIR Model for $R_0 = 2$

Figure 3 shows the SIR curve if the value of R_0 is reduced to 1, which signifies that the infection rate is equal to the recovery rate. Notice the flatness of the I^* versus t^* curve. This means that the other two curves are not affected dramatically, because there is no initial spike in infections. Also, it is the borderline value of a virus spreading versus not spreading. Such a situation is the borderline between a non-pandemic and pandemic scenario.
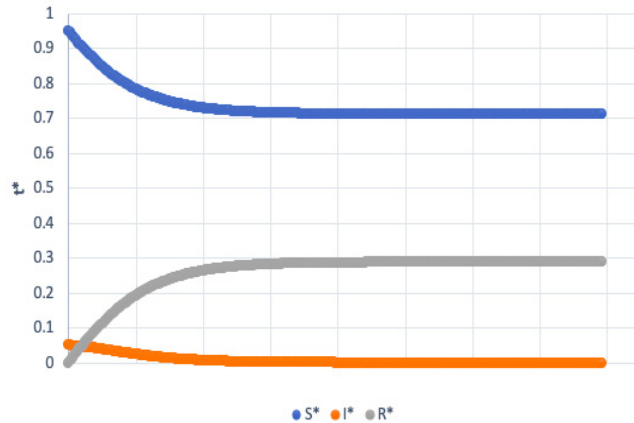

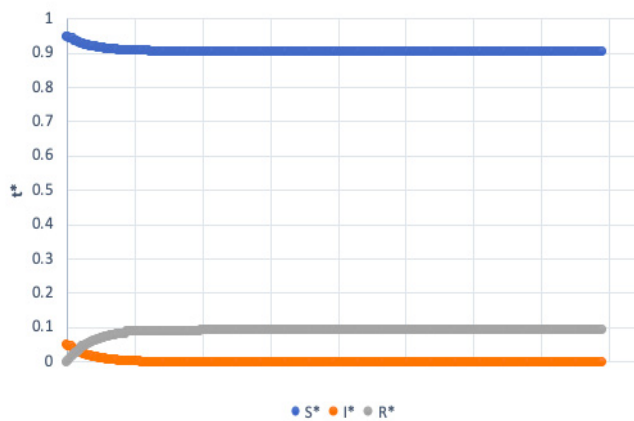**Figure 3.** Sample SIR Model for $R_0 = 1$

Figure 4. Sample SIR Model for $R_0 = 0.5$

Figure 4 highlights what happens when $R_0 = 0.5$. $R_0 < 1$ is desirable because it means that the recovery rate is quicker than the infection rate. As a result, the infected curve approaches 0 immediately and the pandemic dies away.

In-Depth Analysis of SIR Model

Our research now leverages the SIR model to determine the proportion of the population that will never be infected with COVID-19 based on the R_0 by deriving an approximation for S* in terms of R_0, where S* represents the proportion of the population that will never be infected [Appendix; Part B].

Figure 5 demonstrates why it is problematic if $R_0 > 1$. If the basic reproduction number is maintained below 1, essentially the entire population will never be infected, whereas if that number exceeds 1 and continues to increase, more people are guaranteed to be infected in the population at some point. R_0 is graphed on the x-axis and the proportion of the population that is never infected is graphed on the y-axis using Excel. Notice how there is a rapid drop in the proportion never infected after $R_0 = 1$. Once the basic reproduction number increases to 2, 80% of the population

is guaranteed to be infected. If it reaches 4 or 5, which is dangerously high, almost every person in the population is guaranteed to be infected.
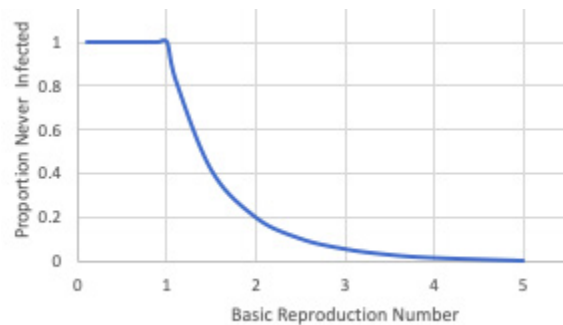

Figure 5. Comparing R_0 to the Proportion Never Infected.

As the number of infected people continues to increase, a virus has a higher chance of mutating. If the mutation rate of the virus increases over time, vaccines become ineffective. This is true of the flu; since it mutates so rapidly, physicians need to administer new vaccinations every year [8]. Even though the novel COVID-19 does not mutate as rapidly as the flu, it is still important to control its spread because developing new forms of vaccines in a strained US healthcare system will be an arduous task. Therefore, as this graph demonstrates, keeping R_0 below 1 is vital because it will prevent the virus from mutating at a rapid rate.

SIQR Model

We want to understand how and why quarantining will slow the infection rate. As such, the infected individuals who require isolation will create the quarantined group, and thus the SIQR Model [9], which is a modified version of the SIR model. Figure 6 shows the diagram for this model. In order to accommodate the quarantined group, let Q represent the number of quarantined individuals per unit area.
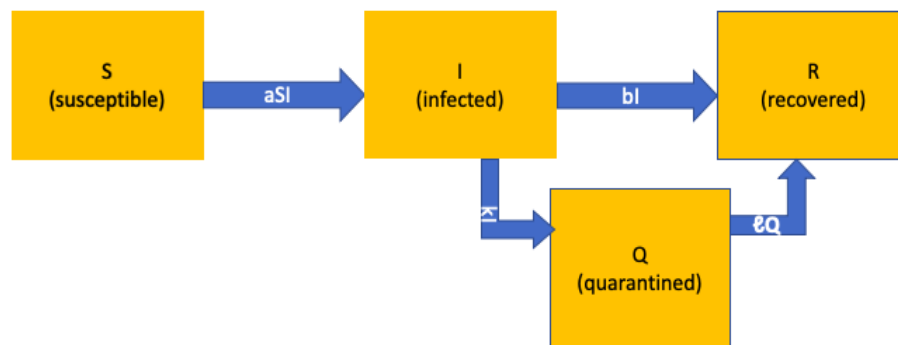

Figure 6. SIQR Model Diagram

There are five parameters to consider when evaluating the SIQR Model:
a: the rate at which individuals are infected (1/days)
b: the rate at which individuals recover (1/days)
N: the population density (individuals per unit area)
k: the rate at which infected individuals are quarantined (1/days)
ℓ: the rate at which quarantined individuals recover (1/days)

As a result, $\frac{1}{a}, \frac{1}{b}, \frac{1}{k}, \frac{1}{\ell}$ would respectively represent the average time for an individual to be infected, for an individual to recover, for an individual to be sent into quarantine or isolated, and for a quarantined individual to recover.

Quarantining is one of the most important actions that a school must undertake in order to avoid a rampant spread of the epidemic. The individuals who test positive for COVID-19 will transition from the susceptible to the infected stage at a rate of aSI. The individuals who recover or quarantine will transition away from the infected stage; hence a loss of bI and kI. The kI individuals are the quarantined population and the bI individuals are the recovered population. The rate at which individuals who recover after quarantining is represented by the amount ℓQ.

Simplified SIQR Model for Quarantining

In order to express the variables described in the previous section in terms of R_0, our research simplifies the SIQR model, which will elucidate the impact of quarantining.

To simplify the SIQR Model, we introduce the parameter T, which combines quarantined individuals with recovered individuals in a population; thus, $T = Q + R$. This is reasonable because quarantined individuals can no longer infect other individuals in the population. While it is not strictly true that all quarantined individuals will recover, we can once again still make this assumption because of the low death rate of COVID-19 [6]. We can call our simplified SIQR Model the SIT Model (Figure 7), and it will be useful when considering how k impacts the R_0.

Figure 7. SIT Model Diagram

Again, the individuals who are susceptible will be infected at a rate of aSI and the quarantined and recovered individuals can be combined into one group as $bI + kI$. We use the nondimensionalized equations for the SIT Model to understand the rate at which the proportion of the susceptible (S^*), infected (I^*), and quarantined/recovered (T^*) populations is changing with respect to time using the time constant $\frac{1}{b+k}$ (Equation 7, 8, and 9). Again, we have that $N = S + I + Q + R$, or in this case, $N = S + I + T$.

Equation 7: Rate of Change of Susceptible Proportion for SIQR Model

$$\frac{dS^*}{dt^*} = -\frac{aN}{b+k}S^*I^*$$

Equation 8: Rate of Change of Infected Proportion for SIQR Model

$$\frac{dI^*}{dt^*} = \frac{aN}{b+k}S^*I^* - I^*$$

Equation 9: Rate of Change of Quarantined/Recovered Proportion for SIQR Model

$$\frac{dT^*}{dt^*} = I^*$$

This is similar to the SIR model, but the new basic reproduction number for quarantining, represented by $R_{0,Q}$, is $\frac{aN}{b+k}$. As a proportion, it can be estimated as the following:

$$R_{0,Q} = \frac{"rate\ of\ infection"}{"rate\ of\ recovery\ \&\ quarantining"}.$$

As made clear by the equation, it is important that schools immediately quarantine an infected student in order to maximize the value of k and consequently reduce $R_{0,Q}$. If the value of k is not maximized, it will be difficult for a population such as a school community to maintain the basic reproduction number below 1.

Results

Now, let us leverage the SIR and SIQR models to evaluate reopening options for schools in the United States by performing both qualitative and quantitative analysis of reopening policies.

Qualitative Analysis of Reopening Policies

The CDC [10] has displayed a variety of recommendations for reopening schools, including the following policies:
1. Hand sanitizer for students and hand washing
2. Face masks
3. Cleaning/disinfecting campus
4. Ventilation
5. Students bring their own meals
6. Staggered scheduling, cohorts
7. Isolate or transport sick students
8. In person, 6 feet distance between desks
9. Closing communal places such as dining halls and playgrounds
10. Worst case: Distance learning only

Table 1 below shows the effects of these 10 virus mitigation approaches on each of the SIRQ model parameters: a, b, N, k, and ℓ (reference section "SIQR Model"). The effect on each parameter is categorized as follows: I = increases; D = decreases; NI = no impact.

Table 1. Measuring the Impact of CDC Policies

	a: infection rate	b: recovery rate	N: student density	k: quarantine rate	ℓ: recovery rate from quarantine
Hand sanitizer/hand washing	D	NI	NI	NI	NI
Face masks	D	NI	NI	NI	NI
Cleaning campus	D	NI	NI	NI	NI
Ventilation in classrooms	D	NI	NI	NI	NI
Students bring their own meals	D	NI	NI	NI	NI
Staggering and cohorts within classrooms	NI	NI	NI	I	NI

Isolate and transport sick students	NI	NI	D	I	NI
Six feet distance between desks in the classrooms	NI	NI	D	NI	NI
Close communal areas	D	NI	NI	NI	NI
Distance learning only	D	NI	NI	NI	NI

While we are unable to quantify the amount by which the SIQR parameters will change as a result of the virus mitigation approaches, it is still possible to determine if each parameter increases, decreases, or stays the same as a result of these approaches. Most of the guidelines recommended above such as hand sanitizer and face masks will only impact the infection rate but none of the other rates. As is noticeable also, none of the policies will effectively improve the recovery rate or the rate at which students recover from quarantine. Isolating and transporting sick students is known as contact tracing, and this helps to both reduce the student density slightly as well as increase the rate at which students are quarantined.

Quantitative Analysis of Reopening Policies

The focus of this section will be on Policy 6 and 7, which are both quantifiable using $b, k,$ and N. For a student, we can assume a two week, or 14 days, recovery on average insofar as they are not in a critical condition [11]. Let R_0 represent the basic reproduction number if a school does not quarantine any students and let $R_{0,Q}$ represent the basic reproduction number if a school effectively quarantines students. We have that $R_0 = \frac{aN}{b}$ from Equation 4 and Equation 5, and $R_{0,Q} = \frac{aN}{b+k}$ from Equation 7 and Equation 8. Therefore, the proportion $\frac{R_{0,Q}}{R_0}$ can be written as a function of b and k: $\frac{R_{0,Q}}{R_0} = \frac{b}{b+k}$. In order for the quarantining to be effective, we would want to minimize the value of $\frac{R_{0,Q}}{R_0}$, which means that k must be maximized. Higher values of k will cause the fraction to approach 0. This means that $R_{0,Q}$ is significantly smaller than R_0, and that the school has quarantined students at an effective rate to slow the spread of the virus.

Suppose schools in the United States identify and quarantine all students who test positive for COVID-19 within two days. This can mean sending them home or isolating them from other non-infected students on campus. In this case, $\frac{R_{0,Q}}{R_0} = \frac{1/14}{1/14 + 1/2} = 0.125$. This means the R_0 would be $\frac{1}{0.125} = 8$ times as much if the school did not quarantine students, which is a significant difference. Similarly, if only three days are required, $\frac{R_{0,Q}}{R_0} = 0.176471$, meaning R_0 would be 5.667 times as much if the school did not quarantine students. Requiring more time to quarantine infected students is problematic, and it causes quarantining to essentially become useless. Suppose schools identify and quarantine students who test positive for COVID-19 within two weeks (14 days), equivalent to recovery time. In this case, $\frac{R_{0,Q}}{R_0} = 0.5$, and thus R_0 would be only twice as much if the school did not quarantine students.

Additionally, in order to reduce R_0 even further, a reduction of class sizes for the non-infected students would be helpful. The same set of students should see each other every time they have classes through staggered scheduling. If the class sizes are reduced, the student population density, N, will be reduced, and the reproduction number would thereby also be reduced. For example, if N is halved then R_0 will be halved. To have a tangible reduction in the R_0, class sizes of non-infected students can be reduced by at least $\frac{1}{2}$, translating to a $\frac{1}{2}$ reduction in R_0 when compared to regular classes.

Conclusion

We are able to estimate the proportion of the population that will never be infected with COVID-19 solely based on the R_0. There is a rapid drop in this proportion after $R_0 = 1$, demonstrating why keeping the R_0 below 1 would be of utmost importance when evaluating the reopening of schools.

Policies such as hand sanitizer for students, face masks, and children bringing their own meals can only be evaluated qualitatively because the models described in this paper cannot be used to quantify how much they reduce the rate of infection by. Nevertheless, these policies are absolutely necessary for a reduction in the R_0, and students must abide by these guidelines when they return to campus. The school should also ensure proper ventilation and disinfect campus to reduce the rate of infection.

Quarantining and transporting sick students can be evaluated quantitatively. It is concluded that US schools should identify and quarantine (send home) all students who test positive for COVID-19 within 2 or 3 days in order to best minimize the spread on campus, assuming that infected students recover within two weeks on average. If schools are able to accomplish this, our findings show that R_0, the ratio of the rate of infection to rate of recovery, would be 5 to 8 times lower as compared to no quarantining. Requiring more than two weeks is not ideal because the reduction in R_0 would be less than 2 times the amount had schools not quarantined, which is insignificant.

A reduction in class size and staggered scheduling impacts student density, which is proportional to the R_0. As such, schools can utilize staggered scheduling to reduce class sizes by at least 50%, or $\frac{1}{2}$, in order to limit R_0.

Acknowledgments

References

[1] Jones, James H. "Notes on R0." *Stanford University*, 1 May 2007, web.stanford.edu/~jhj1/teachingdocs/Jones-on-R0.pdf.

[2] Driessche, P. van den, and James Watmough. "Reproduction Numbers and Sub-Threshold Endemic Equilibria for Compartmental Models of Disease Transmission." *Mathematical Biosciences*, Elsevier, 1 Oct. 2002, www.sci-encedirect.com/science/article/pii/S0025556402001086.

[3] Abou-Ismail, Anas. "Compartmental Models of the COVID-19 Pandemic for Physicians and Physician-Scientists." *SN Comprehensive Clinical Medicine*, Springer International Publishing, 4 June 2020, www.ncbi.nlm.nih.gov/pmc/articles/PMC7270519/.

[4] "Population - New York City Population." *Planning-Population-NYC Population Facts - DCP*, www1.nyc.gov/site/planning/planning-level/nyc-population/population-facts.page.

[5] "Montana Population, Area, and Density." *Montana State Population, Area, and Density Information*, www.states101.com/populations/montana.

[6] "Coronavirus (COVID-19) Mortality Rate." *Worldometer*, 14 May 2020, www.worldometers.info/coronavirus/coronavirus-death-rate/.

[7] Ridenhour, Benjamin, et al. "Unraveling R0: Considerations for Public Health Applications." *American Journal of Public Health*, American Public Health Association, Feb. 2014, www.ncbi.nlm.nih.gov/pmc/articles/PMC3935673/.

[8] Rossman, Jeremy. "COVID-19: Is the Virus Mutating?" *BBC Science Focus Magazine*, 9 Dec. 2020, www.sciencefocus.com/news/covid-19-is-the-virus-mutating/.

[9] Odagaki, Takashi. "Exact Properties of SIQR Model for COVID-19." *Physica A*, Elsevier B.V., 15 Feb. 2021, www.ncbi.nlm.nih.gov/pmc/articles/PMC7680020/.

[10] "Considerations for Schools." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 19 May 2020, www.cdc.gov/coronavirus/2019-ncov/community/schools-childcare/schools.html.

[11] Maragakis, Lisa. "Coronavirus Diagnosis: What Should I Expect?" *Johns Hopkins Medicine*, 22 Oct. 2020, www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/diagnosed-with-covid-19-what-to-expect.