Understanding Epidemics with the SIR Model:
A Simulation of Disease Spread

Rossa Hyeonji Lee[1] [*]

[1]Shady Side Academy, Pittsburgh, PA, USA
*Corresponding Author: rossalee0@gmail.com

Advisor: Heegyeong Kim, hkbreana@gmail.com

Abstract

The COVID-19 pandemic has significantly impacted society, highlighting the need to understand disease dynamics. This paper explored the Susceptible-Infected-Recovered (SIR) model, which categorizes populations to simulate disease spread. By analyzing the transmission (β) and recovery rates (γ), the model provided insights into epidemic behavior. Using COVID-19 data from South Korea (February-April 2020), the study examined the impact of social distancing on the effective reproduction number (R_t). The findings showed that interventions effectively reduced R_t, demonstrating that the SIR model's predictions align with observed outcomes in managing outbreaks. The paper concluded that even though the SIR model is valuable for showing basic and simple trends for the disease, incorporating R_t and more complex alternative models provides a more comprehensive understanding and enhances response strategies. This approach is crucial for preparing for future pandemics and developing effective public health policies.

Keywords: SIR model, COVID-19, Reproduction Number, Epidemic, Effective Reproduction Number, Public Health

1. Introduction

The outbreak of COVID-19, caused by the novel Coronavirus SARS-CoV-2, has posed unprecedented societal problems, economic crises, and global health issues. For teenagers, school closures and the shift to remote learning have posed significant challenges. It has impacted their academic progress, mental and physical health, social skills, and overall well-being (Lee et al., 2021; Copeland et al., 2021). As for the economy, major depressions in various industries such as retail, hospitality, and tourism have led to a global economic crisis.

The understanding and management of infectious disease outbreaks have long been supported by the development of epidemiological models. One of the most fundamental of these models is the Susceptible-Infected-Recovered (SIR) model, introduced by Kermack and McKendrick in 1927 (Kermack & McKendrick, 1927). The SIR model, among others, has been instrumental in predicting the spread of infectious diseases and guiding public health interventions (Anderson & May, 1991). For instance, during the Spanish Flu, early versions of epidemic models helped in understanding transmission dynamics (Mills, Robins, & Lipsitch, 2004), while in more recent times, models similar to the SIR have been used to manage the H1N1 influenza pandemic (Chowell & Nishiura, 2008). By providing insights into potential future scenarios, these models have played a crucial role in shaping responses to outbreaks. The SIR model's continued relevance, especially during the COVID-19 pandemic, underscores the importance of such tools in predicting and mitigating the spread of diseases (Fraser et al., 2004; Heesterbeek, 2002).

Many scholars warn that trending global health issues such as climate change and drug resistance could induce more hazardous and frequent pandemics (Prillama, 2022, Georgetown University Medical Center, 2022, Wong 2024). As a result, there has been an increasing focus on models that can predict how an infectious disease might spread and

how much of the population could get infected. Such models have become a topic of discussion in the news and other media. Fortunately, a lot of this work has already been done in the fields of mathematical biology and epidemiology. The SIR model helps us comprehend how diseases spread through a population over time. It divides the population into three groups: those susceptible to the disease , those infected and infectious, and those who have recovered and gained immunity. By understanding the dynamics of these groups, scientists can predict how an epidemic might unfold, identify potential control measures, and estimate the impact on healthcare systems.

While the SIR model has been widely used, this paper uniquely analyzes COVID-19 transmission dynamics using data from South Korea. Previous studies have often focused on Western countries, leaving a gap in understanding the behavior of the disease in different cultural and social contexts. This study aims to address this gap and provide insights that can be generalized to other regions and future pandemics. By analyzing the results of both theoretical simulation and empirical data, an examination of the accuracy and applicability of these models has been conducted, especially in understanding and predicting the spread of COVID-19. The goal of this paper is to identify the implications and limitations of preparing for future pandemics and formulate policies and interventions that will protect adolescents from outbreaks.

2. SIR Model

The SIR model was introduced by Kermack and McKendrick in 1927 (Wade, 2023). This model is based on a hypothetical disease that has no incubation period, and once cured, the patient develops herd immunity. The SIR model mathematically examines this phenomenon through a set of ordinary differential equations (ODEs) that establish connections among the variables mentioned above. Numerous studies have expanded upon this theory, aiming to explore diverse facets of disease propagation using the straightforward framework of ODEs. In this model, individuals fall into three categories: Susceptible (S), Infected (I), or Recovered (R) with acquired immunity. Those who are infected transmit the

Figure 1. Diagram explaining the relation between parameters (Weiss, 2013).

disease to susceptible individuals they encounter at a rate represented by β, which means the number of contacts per person per time by the probability of disease transmission in a contact between a susceptible and an infective. The patient recovery rate is indicated by γ, which is the reciprocal of the infectious period. The relation between the parameters and the mathematical equations is indicated in Figure 1.

$$ds/dt = -\beta SI/N \qquad (1)$$

$$dI/dt = \beta SI/N - \gamma I \qquad (2)$$
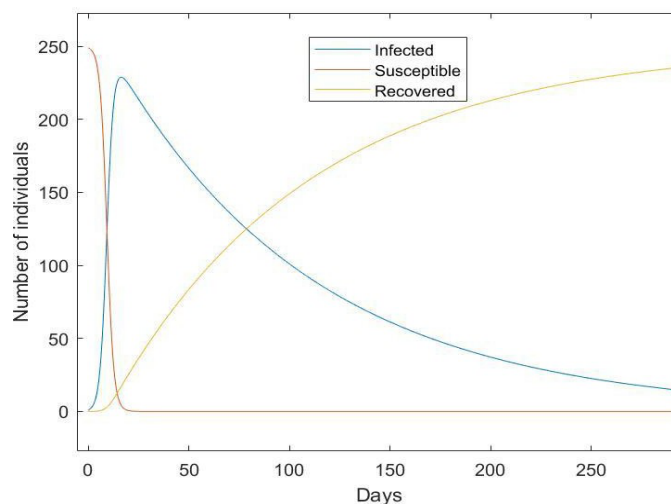
$$dR/dt = \gamma I \qquad (3)$$

Figure 2. The number of individuals susceptible, infected, and recovered over time of 250 populations where the transmission rate β is 0.6, and the recovery rate γ is 0.01.

The ODEs of the SIR model and their corresponding graph provide a comprehensive understanding of disease dynamics.

1. Susceptible Equation (dS/dt = -βSI/N): This equation, depicted by the red line in Figure 2, represents the rate of change in the susceptible population. The decrease in the number of susceptible individuals is proportional to the increase in newly infected individuals. This inverse relationship is encapsulated by the term βSI/N in the equation.

2. Infected Equation (dI/dt = βSI/N - γI): This equation, represented by the blue line in the graph, describes the rate of change in the infected population. The infected population increases from the number of new infections (βSI/N). However, this increase is counterbalanced by the number of individuals who recover or die (γI), thereby reducing the infected population. The net change in the infected population is the difference between these two values.

3. Recovered Equation (dR/dt = γI): This equation signifies the rate of change in the removed (recovered or deceased) population. As the number of infected individuals decreases due to recovery or death, the number of removed individuals correspondingly increases, which is represented by γI in the equation.

These equations and their graphical representations offer a dynamic view of how an infectious disease propagates and gets controlled within a population. The susceptible population decreases as the disease spreads, the infected population initially increases but then decreases as individuals recover or die, and the removed population increases over time as the disease is brought under control.

The two variables that had the greatest impact on the spread and control of pandemics were the transmission rate β and recovery rate γ. This figure shows the number of infected patients according to the infection equation and recovery equation. As the infection ratio increases, the number of infected patients increases, which makes the disease more contagious. As the recovery ratio increases, the speed of spreading gets slow and the number of infected patients decreases.

This correlation shows two ways to reduce the number of infected patients and slow down the speed of spreading. The first way is to reduce the number of new infections. This is equivalent to reducing β in the SIR model introduced above, meaning that the fewer contacts an infected patient has with uninfected people, the lower the maximum number of infections. An example of practicing this is social distancing. Another way is to increase the number of infected patients cured in a day, for example, by making γ larger. As shown in figure 2, the number of peak infections decreases as γ increases (left side of the graph), and the number of peak infections increases as γ decreases (right side of the graph).

2.1 Reproduction Numbers

Another parameter that uses this correlation is called Basic Reproduction Number(R_0), which shows how contagious the disease is. R_0 is β divided by γ and is an indicator of the average number of additional infections that occur after an individual is infected. For example, measles has a R_0 of 12 to 18, indicating that an infected individual can cause as many as 18 new infections (D'Arienzo and Coniglio, 2020).
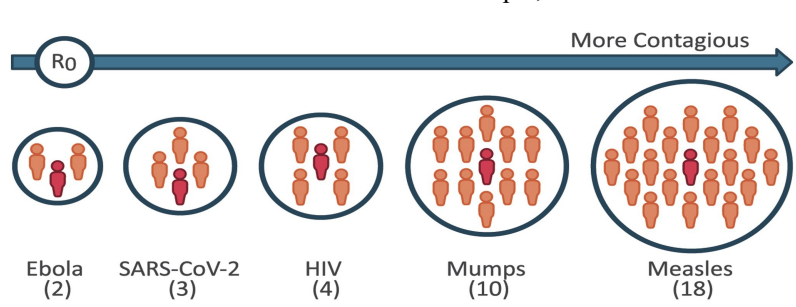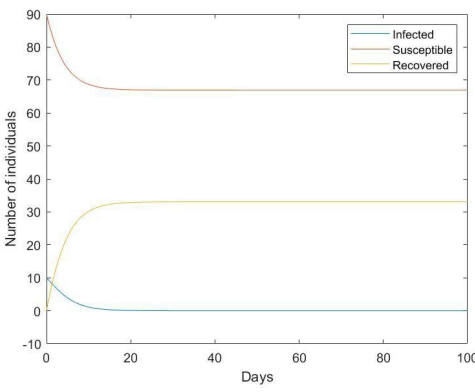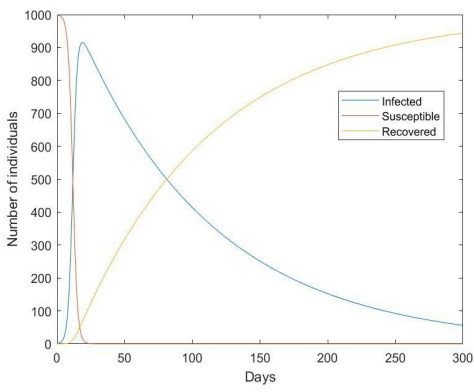

Figure 3. Basic reproduction number (R_0) of Ebola, SARS, HIV, Mumps, and Measles. Red represents the initial infection and orange represents the secondary infection (Patel et al.,2020).

If β becomes low and γ becomes high so that R_0 is less than 1, the number of infected patients will not increase and will die out naturally. In a spreadable graph, the decrease in the number of infected patients is due to a decrease in the number of people who can potentially be infected. In contrast, the graph of a self-dying epidemic,

depicted in the left graph in Table 1, shows that the disease cannot spread due to the fact that the number of recovering patients is rising faster than the number of people getting infected. The table below shows how R_0 affects the spread of a disease. In most cases, this can be divided into three types. If $R_0 > 1$, an infected individual can spread the disease to more than one person, which indicates the possibility of an epidemic. If $R_0 = 1$, such as seasonal influenza, the disease stays stable and has a low chance of inducing an epidemic. If $R_0 < 1$, the disease will eventually be eradicated.

Table 1. The effect of the basic reproduction number on spreading disease

	Non-spreadable	Spreadable
$R_0 = \beta/\gamma$	$\beta = 0.6; \gamma = 0.67; R_0 = 0.9 \leq 1$	$\beta = 0.6; \gamma = 0.01; R_0 = 60 \geq 1$

For values of R_0 less than 1, as depicted in the left column of Table 1, a person can infect less than one person. Therefore, the number of infected people slowly decreases and disappears. Building herd immunity is one of the most well-known ways to have $R_0 < 1$. Herd immunity is the protection mechanism that can arise when a sufficient number of susceptible populations are immune to certain diseases (Randolph and Barreiro, 2020). On the other hand, for R_0 values greater than 1, as depicted in the left column of Table 1, a person can infect more than one person. In other words, the number of infections increases at a rate that varies depending on the proportion, peaks, and then disappears.

2.2 Effective Reproduction Number

The effective reproduction number, often abbreviated as R_{eff} or R_t, is a pivotal epidemiological measurement that quantifies the average number of secondary infections generated by a single infectious individual within a specific population. Unlike the basic reproduction number (R_0), which estimates the average number of secondary infections in a wholly susceptible population, R_t considers factors such as immunity within the population, interventions like vaccination or public health measures, or changes in social behavior and environment.

This measurement is crucial for understanding the transmission dynamics of infectious diseases, particularly during epidemics or pandemics. If R_t exceeds 1, it indicates that each infectious individual, on average, is transmitting the disease to more than one susceptible individual, leading to sustained transmission. This state may potentially lead to an epidemic or outbreak. Conversely, if R_t is less than 1, the number of new infections decreases over time, suggesting that the disease is declining within the population.

Mathematically, R_t can be expressed as the product of R_0 and the proportion of the population that is susceptible. This formulation accounts for the reduction in transmission potential as individuals become immune to the disease through vaccination or prior infection.

Understanding R_t is crucial for developing public health interventions and strategies aimed at controlling and

mitigating the spread of infectious diseases. By monitoring changes in R_t over time and implementing appropriate measures to keep it below 1, health authorities can effectively manage outbreaks, protect vulnerable populations, and ultimately reduce the burden of disease within communities.

From the definition of R_t shown in equation (4), R_t represents the ratio between the transmission rate (β(t)) and the recovery rate (γ(t)) at a specific time (t) (Chen et al., 2020; Heesterbeek, 2002). This value reflects the dynamic relationship between the rates of infection and recovery over time. Additionally, it influences the probability of encountering an individual who is susceptible to the infection.

$$R_t(t) = \frac{\beta(t)}{\gamma(t)} \cdot \frac{S}{N} \tag{4}$$

By combining equation (1) and equation (2) and applying chain rules and derivative inverse function theorem, we can get equation (5).

$$\frac{dI}{dS} = \frac{dI}{dt} \cdot \frac{dt}{dS} = \left(\frac{\beta SI}{N} - \gamma I\right) \cdot \left(-\frac{N}{\beta SI}\right) = \frac{\gamma N}{\beta S} - 1 \tag{5}$$

Substituting $\gamma N / \beta S$ with R_t, equation (5) can be written with respect to R_t over a sufficiently small time interval, i, can be written:

$$R_t(t) = \frac{1}{\frac{\Delta_i I}{\Delta_i S} + 1} \tag{6}$$

Taking into account death (D) in the classical SIR model, the population balance will follow a discrete difference $\Delta_i S + \Delta_i I + \Delta_i R + \Delta_i D = 0$. Equation (6) can be simplified as

$$R_t(t) = \frac{-\Delta_i S}{\Delta_i R + \Delta_i D} \tag{7}$$

which means

$$R_t(t) = \frac{New\ Infections}{New\ Recoveries + New\ Death} \qquad \text{(Contreras et al., 2020).} \tag{8}$$

3. Analysis

This section used empirical data from South Korea to analyze COVID-19 transmission dynamics, especially on how well it follows the SIR model. This study also calculated the R_t using equation (8) and compared it with other research results which have also been conducted within the same period. The analysis focuses on the period of two months: from February 18, 2020, when the outbreak spread, to April 19, 2020, when the outbreak subsided slightly due to increased social distancing (Kissler et al., 2020).

The 18th of February was the day when the 31st patient, who was eventually called the "super spreader", was reported. Since then, the number of newly infected cases has increased exponentially as well as the R_t. From the 29th

Figure 4. Effective reproduction numbers (R_t) from the 18th of February to the 19th of April in 2020.

of February to the 21st of March, the South Korean government implemented social distancing measures. While there was no full lockdown unlike other countries, rules were introduced to limit indoor gatherings in public facilities and regulate operating hours for restaurants, aiming to minimize situations where people gather. As depicted in the
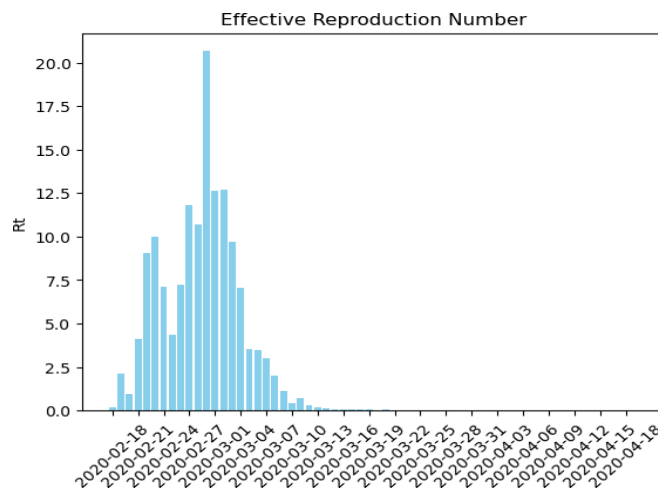
graph in Figure 4, these measures had a significant impact in reducing the spread of COVID-19. From March 22nd, the South Korean government further strengthened social distancing rules. Although new infections continued to occur, the number of recovered individuals, indicating those who have developed immunity, increased significantly. As a result, as illustrated in Figure 4, the reproduction number converged towards zero.
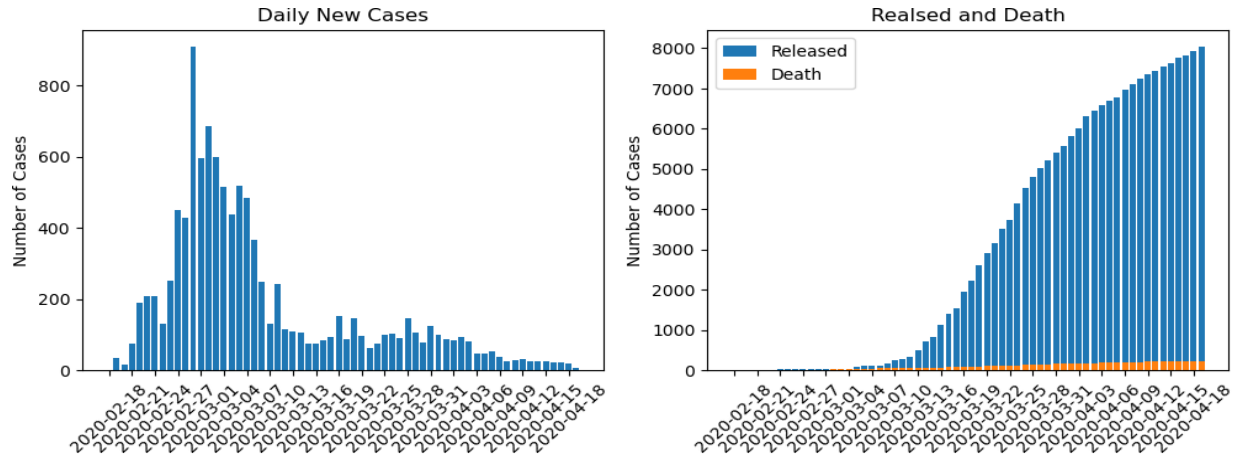

Figure 5. The left graph shows the number of new infections during the period and the right graph shows the number of people who recovered or died during the same period.
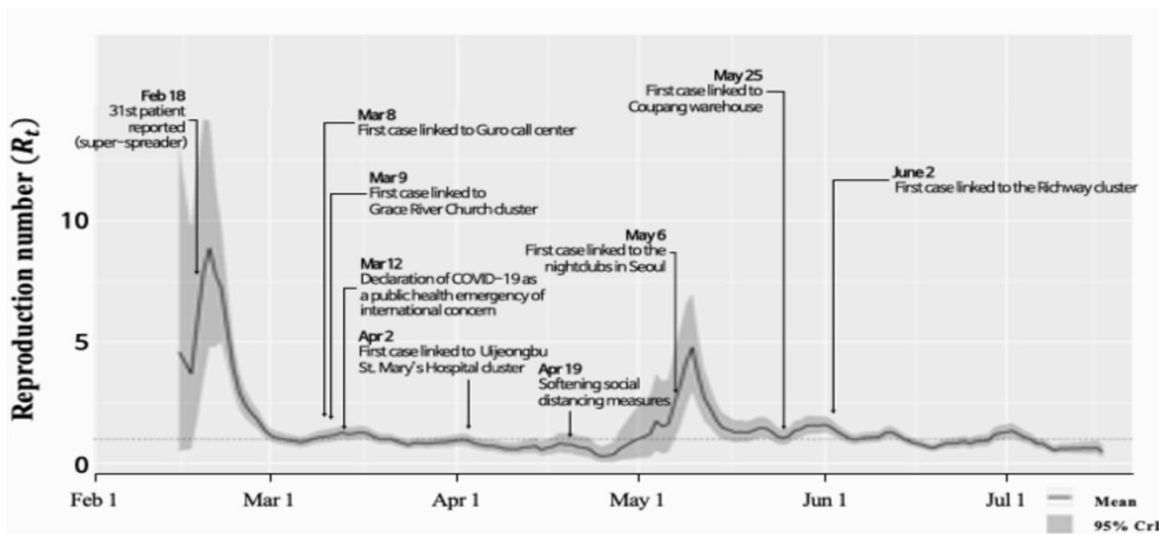

Figure 6. The effective reproduction numbers (R_t) with a 95 % confidence interval in Gyeonggi Province (Shim, E. et al., 2021).

Equation (8) is favored over other methods due to its simplicity and practicality, as it does not require specialized mathematical or scientific expertise to estimate R_t a given population during an epidemic or pandemic. However, because it relies on real-time data, uncertainties in input values can greatly impact the results. Various factors, such as inconsistencies in case-reporting protocols, delays in diagnosis, and biases in reporting or measurement, can contribute to uncertainty in the variables that determine R_t. Considering all the factors, Shim et al. (2021) showed different values of R_t over the same period, as depicted in Figure 6. This study analyzes four different regions in South Korea: Seoul, Gyeonggi Province, Gyeongbuk Province, and Daegu. While the values may differ, the overall trends observed in the graph of Gyeonggi Province, which has the largest population, remained similar. This consistency is significant, as it indicates a general pattern of the impact of interventions and the progression of the epidemic despite variations in specific values.

4. Conclusion

The SIR model, a fundamental tool in epidemiology, offers a mathematical lens through which one can explore the intricate dynamics of disease transmission. The COVID-19 pandemic, which struck the globe with unprecedented force, served as a stark reminder of the necessity of understanding and modeling the spread of disease.

The study focused on South Korea, a region that effectively managed the initial outbreak through interventions such as social distancing. However, the findings have broader applicability, as the principles of the SIR model, such as the impact of transmission and recovery rates on disease spread, are universally relevant. By understanding these dynamics, other regions can adapt the findings to their specific contexts and improve their pandemic response strategies. Future research can build on this work by examining how cultural, social, and economic factors influence the model's parameters in different settings.

The analysis conducted on COVID-19 transmission dynamics in South Korea, utilizing the SIR model, offered valuable insights into the effectiveness of distancing measures and the behavior of the virus within the population. This analysis highlighted the significant impact of social distancing measures enforced during a pandemic, leading to a reduction of new infections and the gradual convergence of R_t towards zero.

During the peak of the virus outbreak, stringent measures were imposed, including school shutdowns, limitations on large gatherings, and advisories against unnecessary public gatherings. As the virus spread rapidly, these restrictions were crucial in curbing transmission. However, as time progressed and a significant portion of the population gained immunity through natural infection, the need for such strict measures diminished.

Upon analyzing the simulation data, a decrease in the transmission rate parameter (β) was observed, coinciding with the increase in population immunity. This decline in β led to a rapid decrease in the effective reproduction number (R_t), as depicted in Figure 5 above. From the graph, it is evident that the natural acquisition of immunity within society conferred a form of herd immunity, resulting in a significant reduction in β.

Before widespread immunity was achieved, restrictions and social distancing measures served as essential tools in mitigating transmission, adding to the protective effect of vaccinations. However, as immunity became more prevalent within the population, these measures became less stringent and more voluntary. This shift underscores the importance of population immunity in controlling the spread of infectious diseases and highlights the potential role of natural immunity in complementing vaccination efforts.

The SIR model, while useful, has several limitations. It assumes no incubation period, meaning individuals become infectious immediately upon contracting the disease, which is not always accurate. Additionally, the model assumes perfect immunity post-recovery, which may not hold for all diseases. These assumptions can impact the accuracy of the model's predictions. For example, diseases with long incubation periods or partial immunity may spread differently than the model suggests. Understanding these limitations is crucial for interpreting the model's results and applying them to real-world scenarios.

A comparative analysis revealed that while the SIR model provides a basic understanding, alternative models offer more detailed insights, especially for diseases with complex transmission patterns. Beyond the SIR model, these alternative modeling frameworks include the SEIR model, modified SEIR model, MSEIR model, and MSEIRS model, which have been proposed to capture nuanced aspects of disease dynamics (He et al., 2020). These models enrich our understanding of epidemiology by incorporating factors such as incubation periods, temporary immunity, and newborn immunity (Almeida et al., 2019).

Moreover, the effective reproduction number (R_t) emerged as a critical metric in understanding epidemic dynamics, offering insights into the trajectory of disease spread and the effectiveness of interventions. Integrating R_t analysis into epidemic modeling enhances our ability to evaluate control measures and anticipate future transmission trends.

In conclusion, the application of mathematical models, such as the SIR model, alongside considerations of R_t and exploration of alternative frameworks, empower us to comprehensively understand disease dynamics and allow for evidence-based public health strategies. While navigating the challenges posed by infectious diseases like COVID-19, these tools remain invaluable in our pursuit of global health security.

References

Almeida, R., et al. (2019). "An epidemiological MSEIR model described by the Caputo fractional derivative." *International Journal of Dynamics and Control, 7*, 776-784.

Anderson, R.M. (1991) "Discussion: the Kermack-McKendrick epidemic threshold theorem." *Bulletin of Mathematical Biology, 53*(1), 1–32.

Anderson, R. M., & May, R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press.

Chen, Y., et al. (2020). "A time-dependent SIR model for COVID-19 with undetectable infected persons." *IEEE Transactions on Network Science and Engineering, 7*(4)*, 3279-3294.

Chowell, G., & Nishiura, H. (2008)**. "Quantifying the transmission potential of pandemic influenza." *Physics of Life Reviews*, 5(1), 50-77.

Contreras, S., et al. (2020). "Real-time estimation of Rt for supporting public-health policies against COVID-19." *Frontiers in Public Health, 8,* 556689.

Copeland, W. E., et al. (2021). "Impact of COVID-19 pandemic on college student mental health and wellness." *Journal of the American Academy of Child & Adolescent Psychiatry, 60*(1), 134-141.

D'Arienzo, M., & Coniglio, A. (2020). "Assessment of the SARS-CoV-2 basic reproduction number, R0, based on the early phase of COVID-19 outbreak in Italy." *Biosafety and Health, 2*(2), 57-59.

Epstein, J.M. (2009). "Modeling to contain pandemics." *Nature, 460*(7256), 687.

Fraser, C., Riley, S., Anderson, R. M., & Ferguson, N. M. (2004). "Factors that make an infectious disease outbreak controllable." *Proceedings of the National Academy of Sciences*, 101(16), 6146-6151.

Heesterbeek, J.A.P. "A brief history of r 0 and a recipe for its calculation." (2002). *Acta Biotheor, 50,* 189–204.
He, S., Peng, Y., & Sun, K. (200). "SEIR modeling of the COVID-19 and its dynamics." *Nonlinear dynamics, 10*, 1667-1680.

Kermack, W. O., & McKendrick, A. G. (1927). A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *115*(772), 700–721. http://www.jstor.org/stable/94815

Kissler, S., et al. (2020). "Social distancing strategies for curbing the COVID-19 epidemic." *MedRxiv*, 2020-03.

Lee, J., et al. (2021). "Effects of learning attitudes and COVID-19 risk perception on poor academic performance among middle school students." *Sustainability, 13*(10), 5541.

Mills, C. E., Robins, J. M., & Lipsitch, M. (2004). "Transmissibility of 1918 pandemic influenza." *Nature*, 432(7019), 904-906.

Patel, S. S., Kalma, J., & Bluman, E. M. (2020). Understanding COVID-19 vaccines and their development. *JBJS*, *102*(20), 1759-1769.

Randolph, H. E., & Barreiro, L. (2020). "Herd immunity: understanding COVID-19." *Immunity, 52*(5), 737-741.

Shim, E., Tariq, A., & Chowell, G. (2021). "Spatial variability in reproduction number and doubling time across two waves of the COVID-19 pandemic in South Korea, February to July, 2020." *International Journal of Infectious Diseases*, 102, 1-9.

Prillaman, M. (2022). Climate change is making hundreds of diseases much worse. Nature. https://www.nature.com/articles/d41586-022-02167-z#:~:text=Climate%20change%20has%20exacerbated%20more,to%20a%20rise%20in%20cases.

Georgetown University Medical Center. (2022, April 28). Climate change could spark the next pandemic, new study finds. ScienceDaily.

Wong, C. (2024). Antibiotic resistance is a growing threat-is climate change making it worse? Nature. https://www.nature.com/articles/d41586-023-04077-0

Wade, Robert Hunter (2023) The world development report 2022: finance for an equitable recovery in the context of the international debt crisis. *Development and Change*, 54 (5). pp. 1354-1373.

Weiss, H. H. (2013). "The SIR model and the foundations of public health." *Materials Mathematics*, 1–17.