MATHEMATICAL DISEASE MODELING FOR PUBLIC HEALTH EPIDEMIOLOGISTS: DIFFERENTIAL EQUATIONS SUPPORTING PUBLIC HEALTH INTERVENTIONS

GV Fant

US Department of Veterans Affairs, Washington, DC

ABSTRACT

Differential equations are an increasingly essential tool used in public health epidemiology, enabling the modeling of communicable disease spread in a population and informing the development, implementation, and assessment of public health interventions. As public health professionals increasingly rely on GenAI for data management and data analysis, understanding the mathematical equations underlying a communicable disease model becomes crucial. This essay provides an accessible overview of ordinary first-order differential equations, including the linear first-order differential equation (i.e., exponential growth model) and the nonlinear first-order differential equation (i.e., SIR model), for the public health epidemiologist. We explain each type of differential equation and show a step-by-step solution to facilitate comprehension. A real-world application of the use of first-order differential equations in public health is, briefly, presented. Then, we discuss how real-world data, first-order differential equations, data analysis and predictive data modeling can contribute to the development of insights that guide the development, deployment, and assessment of public health interventions by public health decision-makers for a population. In the appendix, sample prompts are provided to help leverage the power of GenAI in the use of mathematical models in public health epidemiology. This technical review aims to provide the epidemiologist with an understanding of first-order differential equations function in the context of communicable disease modeling at a community- or population-level for public health decision-making and action.

KEYWORDS

Public health epidemiology; First-order differential equations; SIR model; Mathematical models; GenAI

1. INTRODUCTION

Differential equations have been fundamental tools in public health epidemiology, providing a mathematical framework to model the spread of infectious diseases, predict outcomes, and guide interventions for a population. These equations describe how the disease experience changes in a population over time, accounting for factors such as transmission rates, recovery rates, and population dynamics. In fact, many epidemiological models (such as exponential growth models, SIR models) are built on first-order differential equations [1,2]. While mathematical, the results and interpretation of these equations make them ideal for communications between public health epidemiologists and public health professionals and the public [3,4].

The ordinary, first-order differential equation is especially useful with public health epidemiological surveillance data because these equations model the dynamics of communicable disease spread over time by capturing the first derivative of disease states in a population over

time. The first-order differential equation can be solved easily with modern analytic software and GenAI especially when the datasets are large and this makes the use of differential equation ideal for use with modern computational tools and algorithms that process public health surveillance data. Surveillance systems based on real-time incorporation of electronic laboratory reporting (ELR) data of new cases and recoveries can be easily ingested by first-order differential equations to update, for example, SIR model parameters (such as transmission rate or recovery rate) in real time [5-10].

Historically, disease modeling used both deterministic and stochastic models in the context of infectious disease outbreaks such as pandemic influenza, HIV/AIDS, and other communicable diseases. The limitations faced by these early uses dealt with issues of data availability, computational power, model accuracy, and the challenges of predicting complex systems [11-14]. Mathematical models, including those based on differential equations, have become important tools for public health decision-makers, offering quantitative insights to inform both policy decisions and operational public health actions [15]. Various mathematical models used in public health epidemiology have been reviewed in the context of epidemics such as H1N1, Ebola and Zika to illustrate how these models were used to predict disease spread and evaluate intervention strategies [16]. Interestingly, these mathematical models can be used to address health disparities during an infectious disease outbreak by considering factors such as race, ethnicity, and socio-economic status in the modeling efforts [17]. The equations model both the progression of disease in a population and enable public health officials to make informed decisions about interventions, such as implementing quarantine measures or allocating resources.

By understanding these mathematical models, public health professionals can better anticipate the course of an epidemic, evaluate the potential impact of interventions, and make data-driven decisions to protect public health. This technical review is designed to discuss key concepts offirst-order differential equation for public health epidemiologists. This review contains four parts: 1) linear first-order differential equation with step-by-step worked problem; 2) nonlinear first-order differential equation with step-by-step worked problem; 3) a case study from the published literature; and 4) framework connecting mathematical models in epidemiology with real-time data for public health interventions. In the appendix, sample prompts are provided to help leverage the power of GenAI in the use mathematical models in public health epidemiology. This technical review aims to provide the epidemiologist with an understanding of first-order differential equations function in the context of communicable disease modeling at a community- or population-level for public health decision-making and action.

2. LINEAR FIRST-ORDER DIFFERENTIAL EQUATIONS

The published, scientific literature has many examples of the description and study out communicable disease outbreak data that demonstrate exponential growth of the disease in a population of a specific geographic location [18]. Exponential growth is a common application of the exponential function that is common in the study and practice of applied calculus in public health epidemiology [1-2,19-21]. The exponential growth function is a common linear first-order differential equation.

2.1. Introduction to Linear First-Order Differential Equations

Linear first-order differential equations describe situations where the first derivative of a function represents the outcome variable of interest (say, the number of individuals with a communicable disease in a population) with respect to time is proportional to the quantity itself. The first derivative ($\frac{dy}{dt}$) is also known as the rate of change. The first-order differential equation is useful

for modeling the early stages of a communicable disease outbreak, where the spread of the disease shows the pattern of exponential growth or decay. Consider the basic first-order differential equation:

$$\frac{dy}{dt} = \beta \, y$$

Where $\frac{dy}{dt}$ represents the rate of change in the number of infected individuals over time t, and β is a constant representing the transmission rate of the disease. This equation models the situation where each infected individual spreads the disease to others at a constant rate, leading to exponential growth in the number of infections within the population. In other words, the exponential growth functions described in the above equation allows the epidemiologist to predict how fast a communicable disease might spread in a population and guiding the deployment of community disease control measures.

2.2. Solving Linear First-Order Differential Equations

Problem: As found in a public health disease surveillance system, the number of cases of a communicable disease is growing at a rate proportional to the current number of cases. Suppose the growth rate of the communicable disease is 0.3 per day. If 100 cases are reported today, model the number of cases N(t) after t days.

Below is a table (Data Table 1) that shows the number of communicable disease counts over time, given a daily growth rate of 0.3 per day.

Data Table 1: Daily counts of new cases of communicable disease by day

Day	Number of Cases N(t)	$\frac{dN}{dt}$ (Rate of Change)
0	100	0.3 x 100 = 30
1	100 x e $^{0.3}$ ≈ 134.99	0.3 x 134.99 ≈ 40.50
2	160 x e $^{0.6}$ ≈ 182.21	0.3 x 182.21 ≈ 54.66
3	190 x e $^{0.9}$ ≈ 245.96	0.3 x 245.96 ≈ 73.79
4	220 x e $^{1.2}$ ≈ 332.87	0.3 x 332.87 ≈ 99.86

Step-by-Step Solution

1. Set-up the differential equation: The above problem describes a communicable disease with exponential growth, which can be modeled by the differential equation:

$$\frac{dN}{dt} = 0.3N$$

2. Solve the equation: This is a separable differential equation. Integrating both sides yields:

$$\int \frac{1}{N} \, dN = \int 0.3 \, dt$$

Solving the integrals gives:

$$\ln(N) = 0.3t + C$$

Exponentiating both sides gives:

$$N(t) = e^{0.3t + C} = e^C e^{0.3t}$$

Letting $e^C = N_o$ (the initial number of cases), we have:

$$N(t) = N_0 e^{0.3t}$$

3. Apply initial conditions: Given $N(0) = 100$, we have $N_0 = 100$. See Data Table for details.

4. Interpret the solution: From the Data Table 1, we see that the number of cases grows exponentially, reflecting the early stages of a communicable disease outbreak where the disease is spreading rapidly. Public health officials can use the model to forecast the number of cases in the coming days, allowing them to allocate resources for the community.

Summary: For monitoring disease spread, using the linear first-order differential equation allows the public health epidemiologist to predict the number of future infections and help the public health decision-maker manage public health resources effectively.

3. NONLINEAR FIRST-ORDER DIFFERENTIAL EQUATIONS

A nonlinear first-order differential equation used to model epidemics is the SIR (Susceptible-Infected-Removed) model and related to first-order differential equations from applied calculus [21]. The SIR model can be applied to the outbreak of influenza in a population [22]. All models of infectious disease use simplified mathematical models to represent key features of complexity of the infection in the population, including the pattern of transmission [4-6; 23-24].

3.1. Introduction to Nonlinear First-Order Differential Equations

The SIR model (Susceptible-Infected-Recovered) is a classic example of a nonlinear differential equation system used to describe how a disease spreads through a population. It divides the population into three groups:

1. **Susceptible (S):** These are people who can catch the disease.
2. **Infected (I):** These are people who currently have the disease and can spread it to others.
3. **Recovered (R):** These are people who have recovered from the disease and are now immune (or have died, depending on the context).

The mathematical model consists of three, coupled first-order differential equations:

$$\frac{dS}{dt} = -\beta\,SI$$

$$\frac{dI}{dt} = \beta\,SI - \gamma\,I$$

$$\frac{dR}{dt} = \gamma\,I$$

These equations represent the rates of change for susceptible, infected, and recovered populations, respectively.

$\frac{dS}{dt}$: This equation represents the rate of change of the susceptible population. It decreases as susceptible individuals come into contact with infected individuals and become infected.

$\frac{dI}{dt}$: This equation represents the rate of change of the infected population. It increases as susceptible individuals become infected and decreases as infected individuals recover.

$\frac{dR}{dt}$: This equation represents the rate of change of the recovered population. It increases as infected individuals recover and gain immunity.

In each case, the first derivative of one group (S, I, or R) depends on the current size of that group and possibly others. The terms of the right-side of these equations, known as the rate of change expressions, are crucial in dictating the dynamics of how the population in each compartment in the mathematical model (S, I, R) evolves over time. The right-side terms determine the rate at which individuals transition between groups, directly influencing the trajectory of the disease spread. In real-world public health applications, the transmission rate (β) and recovery rate (γ) might vary over time due to factors such as changes in population behavior, intervention strategies, or the emergence of new variants.

This basic model serves as an introduction, but more complex models might be needed for specific scenarios. In practice, these coupled, first-order equations are part of the broader system dynamics that govern the spread of infectious diseases. When applying the SIR model to real-world scenarios, it is essential to consider data variability, potential inaccuracies, and the need for continuous updates, especially when using real-time data from Electronic Lab Reporting (ELR) systems. This analysis helps epidemiologists understand the dynamics of the epidemic and the impact of public health interventions

3.2. Solving Nonlinear First-Order Differential Equations

Problem: Suppose we have a small town with 1,000 people, and using the public health communicable disease surveillance system disease tracking in ongoing using ELR data. On Day 1, we have the following:

- 990 people are susceptible (S = 990)
- 10 people are infected (I = 10)
- 0 people have recovered (R = 0)

The transmission rate (β) is 0.001, meaning that the chance of disease transmission during contact between a susceptible and an infected individual is relatively low.

The recovery rate (γ) is 0.1, meaning that each infected person has a 10% chance of recovering on any given day.

Now, using the table (Data Table 2), let's see what happens over the next few days.

Data Table 2: SIR Model-Predicting community disease spread in a population by tracking the number of susceptible, infected, and recovered individuals in a population over time

Day	Susceptible (S)	Infected (I)	Recovered (R)	Transmission Rate (β)	Recovery Rate (γ)
1	990	10	0	0.001	0.1
2	980	19	1	0.001	0.1
3	961	36	3	0.001	0.1
4	926	67	7	0.001	0.1

Step-by-Step Calculations for Day 2

1. Calculate the change in susceptible people (dS):

$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dS}{dt} = -0.001 \times 990 \times 10$$

$$\frac{dS}{dt} = -9.9 \text{ (per day)}$$

So, about 10 susceptible people would get infected over this period.

2. Calculate the change in infected people (dI):

$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dI}{dt} = 0.001 \times 990 \times 10 - 0.1 \times 10$$

$$\frac{dI}{dt} = 9.9 - 1$$

$$\frac{dI}{dt} = 8.9$$

So, the number of infected people increases by 9 over this period.

3. Calculate the change in recovered people (dR):

$$\frac{dR}{dt} = \gamma I$$

$$\frac{dR}{dt} = 0.1 \times 10$$

$$\frac{dR}{dt} = 1$$

So, 1 person recovers in this period.

Updated Counts for Day 2:

- New S = 990 - 10 = 980
- New I = 10 + 9 = 19
- New R = 0 + 1 = 1

These counts for Day 2 along with those for the remaining days allow public health officials to forecast when the number of infected persons for the disease might peak. This knowledge could help officials better prepare for vaccine distribution and consider the possible needs of additional numbers of healthcare beds in a specific area where the communicable disease is present.

(Note: following the above pattern of calculations will yield the data for Day 3 and Day 4.)

Summary: The SIR model allows the public health epidemiologist to predict how a disease might spread in a population by tracking the number of susceptible, infected, and recovered individuals in a population over time (Data Table 2). By using real-time data, such as from ELR systems, this model helps public health officials anticipate the course of a communicable disease outbreak and adjust population-level resources and interventions. For instance, knowing how the transmission rate (β) and the recovery rate (γ) might impact the spread of the disease enables public health decision-makers to allocate resources, implement control measures like quarantine, or adjust public health messaging. The SIR model may provide critical insights for both forecasting communicable disease progression and evaluating the impact of public health interventions. This model plays a vital role in public health surveillance by providing a framework for monitoring the progression of disease outbreaks and evaluating the effectiveness of interventions.

4. CASE STUDY-USE OF DIFFERENTIAL EQUATIONS IN MODELING THE SPREAD OF INFECTIOUS DISEASES

Researchers examined the spread of Hepatitis C virus (HCV) among a group of at-risk individuals, aged 15-29, in Michigan, between 2000 and 2016. Their aim was to study how HCV spreads within this demographic cohort by analyzing disease transmission data and predicting disease prevalence using health surveillance data and ordinary differential equations (ODE) [25]. A HCV ODE transmission model was developed, and the researchers described how surveillance data were incorporated into the ODE transmission model. They also explained how continuous data updates and dynamic model adjustment could refine and enhance the predictive accuracy of their ODE model. The case study aligns with the principles and techniques of differential equations used in public health epidemiology as discussed in this paper. The public health implications resulting from the analytic efforts of the researchers can be found in their published paper [25].

5. LEVERAGING DIFFERENTIAL EQUATION MODELS AND ELR SYSTEMS FOR PUBLIC HEALTH INTERVENTIONS

Concept Map: Framework Connecting Mathematical Models in Epidemiology with Real-Time Data for Public Health Interventions
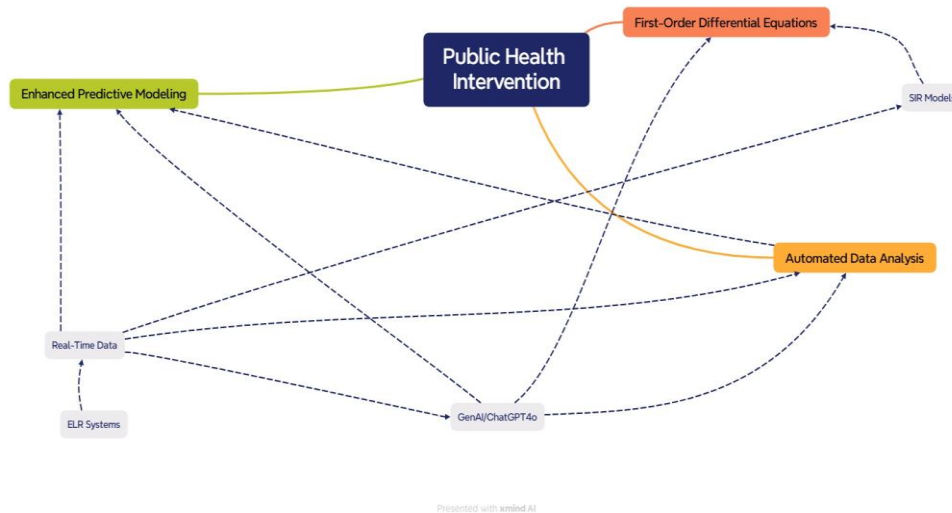

Figure 1.Concept map to visually connecting mathematical models in epidemiology with real-time data for public health interventions

The concept map (Figure 1) illustrates the possible interplay between real-time data, first-order differential equations, automated data analysis and predictive modeling for a public health intervention. For example, adapting population-level disease control strategies to an evolving communicable disease outbreak is a typical application of this approach. The themes or topics discussed in this essay can be visualized using a concept map, where the concepts, enclosed in boxes with lines, show possible relationships between key elements that contribute to public health decision-making for interventions.

Public health, epidemiology, and the operation of a public health disease surveillance system has been described [26-27]. In a modern public health surveillance system, where Electronic Laboratory Reporting (ELR) data is used as a primary source at an enterprise-level and GenAI is an available tool, the public health data structure and response could be different from a traditional approach. ELR enabled public health surveillance systems could provide real-time data, which can be fed into mathematical models, such as first-order differential equations (e.g., exponential growth models, SIR models), to track and predict the of communicable disease in a population. The main elements from the concept map could be connected is a possible way that could leverage real-time data, GenAI, and first-order differential equations to inform the development of public health interventions for the population based, in part, on the interpretation of these surveillance system data. The elements of the concept map are described:

5.1. Element: Core Epidemiological Modeling Techniques

Concept Map Element: First-Order Differential Equations and SIR Models
This element is the foundational importance of first-order differential equations and SIR models for public health epidemiology. The mathematical tools help model the spread of communicable

diseases and inform public health interventions. This is the core of mathematical models in epidemiology for public health practice for communicable disease prevention activities.

Differential equations, including SIR models, have a foundational role in the use of mathematical modeling for public health epidemiology and subsequent public health action. The application of mathematical modelling in understanding the disease dynamics of the COVID-19 pandemic has been documented [28-29].

5.2. Element: Real-Time Disease Data Integration

Concept Map Element: ELR Systems

ELR (Electronic Laboratory Reporting) systems are crucial for collecting and transmitting near, real-time data on communicable diseases. After ensuring that these data are of high quality and internally consistent prior to use, ELR data sources have the potential to update the parameters used in mathematical epidemiologic modeling in order to improve the precision and relevance of public health interpretation for public health action. The ELR data quality require, at least, consistent temporal, test results for the individual records corresponding to the test results and included with HL7 data transmission protocols.

Investigators found that integrating near, real-time data from ELR systems along with dynamic global data improve the accuracy and reliability of epidemiological models [30-31]. This serves to underscore the implied role of ELR data in public health surveillance systems that use mathematical epidemiologic models as the basis of public health epidemiologic analysis and insights.

5.3. Element: Real-Time Predictive Model Calibration

Concept Map Element: Real-Time Data

The ingestion of near, real-time data from ELR systems allow for predictive model calibration (e.g., SIR model calibration). GenAI algorithms could automatically adjust key model parameters, such as the transmission rate (beta) and recovery rate (gamma), based on incoming data. These dynamic model adjustments improve the accuracy of model predictions and help public health epidemiologists analyze and interpret data for public health action.

While the need for continuous model calibration has been acknowledged, it is important to understand the inherent challenges of real-time model calibration during a communicable disease outbreak [30,32]. GenAI-driven calibration, along with the simple act of comparing model results with past data and assessing how small changes might affect subsequent predictions, can help make the model predictions accurate and reliable. Calibrating predictive epidemiological models with continuous real-time data is important.

5.4. Element: Epidemiological Methods and Applications with GenAI

Concept Map Element: Automated Data Analysis and GenAI/ChatGPT4o
AI-tools like GenAI/ChatGPT4o can automate the analysis of complex, vast datasets to uncover trends and patterns that may not be apparent using traditional methods (see appendix). GenAI can assist the public health epidemiologist in the analysis and consideration of data findings and leads to the development of insights that can be contextualized for public health decision-makers as they consider public health interventions for the population.

Investigators examining the COVID-19 outbreak in a geographic location mentioned building trustworthy data-driven epidemiological models discussing concepts such as model development, calibration, real-time data integration, sensitivity analysis that AI supports or automates [30].Other investigators focused on using SIR models to examine the changes in the basic reproduction number (Ro) across different countries during the COVID-19 pandemic, mentioning "automatic figures" and "predicted trends" that suggest the use of advanced computational techniques, including AI-driven algorithms to handle large datasets and complex calculations [33]. By automating data management, data analysis, and model updates, GenAI helps public health epidemiologists respond faster in assessing and interpreting public health surveillance data in order to develop insights for public health decision-makers.

5.5. Element: Communicable Disease Forecasting with GenAI

Concept Map Element: Enhanced Predictive Modeling

AI techniques enhance traditional forecasting efforts and predictive modeling by adjusting the parameters dynamically. This allows the public health epidemiologist to spend more time confirming the appropriate use of forecasting and predictive modeling techniques, considering the output, interpreting the findings for public health action, and framing the overall findings and insights in a way that is most useful for public health decision-makers.

While traditional statistical methods can improve forecasting accuracy and help refine predictive models for understanding the prevalence and incidence of communicable diseases in a population, GenAI may also play an increasingly important role in the same analytic activities [31-32]. GenAI can be leveraged to provide dynamic adjustment of epidemiological model parameters that can lead to improvements in forecasting.

5.6. Element: Public Health Interventions

Concept Map Element: Public Health Interventions

Ultimately, the analytic insights generated using this process must be framed by the public health epidemiologist for the public health decision-maker to guide the development of public health actions and interventions for the population. Interventions are designed to control communicable disease spread in a population, evaluate the effectiveness of public health measures, such as community vaccination campaigns, improve the health status of the population, and protect the health of the population.

SIR models and other epidemiological models were used during the COVID-19 pandemic and similar outbreaks to inform the development of analytic models and guide the implementation of public health interventions to control these diseases [29,32].

ELR data provides real-time information that is critical for parameterizing and validating differential equation models. By incorporating this data into these models for improved predictions, decision-makers can adjust public health strategies.

6. CONCLUSION

Differential equations, both linear and nonlinear, are essential tools in public health epidemiology that help inform public health action. These mathematical models enable public health epidemiologists to model disease dynamics and predict outcomes. The models also contribute to

the development of insights supporting effective public health decisions. By incorporating real-time data from sources like ELR, these mathematical epidemiologic models can be continually refined to improve public health data analysis and insights. As GenAI becomes incorporated into the work of public health, it will be critical for professionals, like public health epidemiologist, to be knowledgeable of mathematical modeling in public health epidemiology in order to evaluate the first-order differential equation outcomes produced by GenAI. This technical review provided a brief introduction to mathematical models in public health epidemiology and described a framework connecting mathematical models in epidemiology with the data found in a public health surveillance that contribute to insights for public health decision-makers. In the appendix, sample prompts are provided to help leverage the power of GenAI in the use mathematical models in public health epidemiology.

ACKNOWLEDGEMENT

REFERENCES

[1]   Neuhauser C. Calculus for Biology and Medicine, 2nd ed. Upper Saddle River, NJ, USA: Pearson Education; 2004.

[2]   Bittinger ML, Brand L, Quintanilla J. Calculus for the Life Sciences, 2nd ed. Upper Saddle River, NJ, USA: Pearson Education; 2006.

[3]   Diekmann O, Heesterbeek JA, Britton T. Mathematical tools for understanding infectious disease dynamics. Princeton University Press; 2012.

[4]   Vynnycky E, White R. An Introduction to Infectious Disease Modelling. Oxford University Press; 2011.

[5]   Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton University Press; 2011.

[6]   Brauer F, Castillo-Chavez C, Feng Z. Mathematical Models in Epidemiology. Springer; 2019.

[7]   Hethcote HW. The mathematics of infectious diseases. SIAM Rev. 2000;42(4):599-653. doi:10.1137/S0036144500371907.

[8]   Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science. 2020;368(6493):860-868. doi:10.1126/science.abb5793.

[9]   Gostic KM, McGough L, Baskerville EB, et al. Practical considerations for measuring the effective reproductive number, Rt. PLOS Computational Biology. 2020;16(12) doi:10.1371/journal.pcbi.1008409.

[10]   Bertozzi AL, Franco E, Mohler G, Short MB, Sledge D. The challenges of modeling and forecasting the spread of COVID-19. Proc Natl Acad Sci U S A. 2020;117(29):16732-16738. doi:10.1073/pnas.2006520117.

[11]   Grassly N, Fraser C. Mathematical models of infectious disease transmission. Nature Reviews: Micorbiology. 2008. 6:477-487.

[12]   Star L, Moghadas S. The role of mathematical modelling in public health planning and decision making. The Purple Paper. 2010. 22:1-6

[13]   Moghadas S, Laskowski M. Review of terms used in modelling influenza infection. 2014. National Collaborating Centre for Infectious Diseases, Winnipeg, Manitoba.

[14]   Fischer L, et al. CDC Grand Rounds: Modeling and Public Health Decision-Making. MMWR. 2016; 65:1374-1377

[15]   Kretzschmar M. Disease modeling for public health: added value, challenges, and institutional constraints. J Public Health Policy. 2020;41(1):39-51. doi:10.1057/s41271-019-00206-0.

[16]   Mata D, Dourado A. Mathematical modeling applied to epidemics: an overview. Epidemiol Infect. 2021;149. doi:10.1017/S0950268821000157.

[17] Richard DM, Lipsitch M. Using infectious disease mathematical modeling to address health disparities. Int J Epidemiol. 2024;53(1). doi:10.1093/ije/dyad180.

[18] Utsunomiya YT, Utsunomiya ATH, Torrecilha RBP, Paulan SC, Milanesi M, Garcia JF. Growth Rate and Acceleration Analysis of the COVID-19 Pandemic Reveals the Effect of Public Health Measures in Real Time. Front Med (Lausanne). 2020 May 22;7:247. doi: 10.3389/fmed.2020.00247.

[19] Goldstein LJ, Lay DC, Schneider DI. *Calculus and its Applications*. 3rd ed. NJ, USA: Prentice-Hall; 1984.

[20] Fant GV. Perspective: Applied calculus for public health epidemiology-empowering public health professionals with mathematical insights. Public Health Open J. 2024;9(1):6-13.

[21] Prodanov D. Analytical solutions and parameter estimation of the SIR epidemic model. In: Mathematical Analysis of Infectious Diseases. 1st ed. Amsterdam: Elsevier; 2022. p. 163-189.

[22] White PJ. Mathematical models in infectious disease epidemiology. Infectious Diseases. 2017; 49-53.e1. doi:10.1016/B978-0-7020-6285-8.0005-8.

[23] Tung KK. Topics in mathematical modeling. Princeton: Princeton University Press, 2007.

[24] Matthew PF, Gower EW. Infectious disease epidemiology. In: Modern Epidemiology, fourth edition. Philadelphia: Wolters Kluwer 2021.

[25] Gicquelais RE, et al. Hepatitis C Transmission in Young People who Inject Drugs: Insights Using a Dynamic Model Informed by State Public Health Surveillance." Epidemics. 2019 June; 27:86-95.

[26] Senthilingam M. Outbreaks and pandemics. Malaysia: UniPress Books, 2023.

[27] Fant G. Concepts of public health surveillance to support infectious disease cluster control efforts in rural Rajasthan, India. Intl J Sci and Healthcare Res. 2020; 5(2): 119-127.

[28] Anastassopoulou C, Russo L, Tsakris A, Siettos C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. PLoS One. 2020;15(3). doi:10.1371/journal.pone.0230405.

[29] Demongeot J, Magal P. Data-driven mathematical modeling approaches for COVID-19: A survey. Phys Life Rev. 2024;50:166-208. doi:10.1016/j.plrev.2024.08.004.

[30] Zhang S, Ponce J, Zhang Z, Lin G, Karniadakis G. An integrated framework for building trustworthy data-driven epidemiological models: Application to the COVID-19 outbreak in New York City. PLoSComput Biol. 2021;17(9). doi:10.1371/journal.pcbi.1009334.

[31] AlQadi H, Bani-Yaghoub M. Incorporating global dynamics to improve the accuracy of disease models: Example of a COVID-19 SIR model. PLoS One. 2022 Apr 8;17(4). doi:10.1371/journal.pone.0265815.

[32] Overton CE, Stage HB, Ahmad S, Curran-Sebastian J, Dark P, Das R, et al. Using statistics and mathematical modelling to understand infectious disease outbreaks: COVID-19 as an example. Infect Dis Model. 2020;5:409-441. doi:10.1016/j.idm.2020.06.008.

[33] Leung TC. Comparing the change in R0 for the COVID-19 pandemic in eight countries using an SIR model for specific periods. COVID. 2024;4(7):930-951. doi:10.3390/covid4070065.

AUTHOR

GV Fant, DSc, PhD, MACE: Epidemiologist, Population Health Informatics, Digital Health Office, Veterans Health Administration, US Department of Veterans Affairs, Washington DC.

Disclaimer: The views expressed in this paper are those of the author and do not represent the official position of the US Government nor the US Department of Veterans Affairs.

APPENDIX: GENAI/CHATGPT4O PROMPTS TO INITIATE ASPECTS OF FRAMEWORK

These simple prompts are included to assist public health epidemiologists in leveraging the power of GenAI in professional work.

Sample prompt for data screening: This prompt will help identify inconsistencies and outliers in the data related to infections, transmission rate, and recovery rate all crucial for accurate modeling.

"Using a de-identified, public use dataset from a public health surveillance system (already uploaded), identify missing or inconsistent data related to reported cases of communicable disease. Focus on identifying outliers in variables number of infections over time, transmission

rate, and recovery rate. Flag any data discrepancies that could affect the accuracy of subsequent differential equations modeling (e.g., for linear models or SIR models)."

<u>Sample prompt for first-order differential equations:</u> This prompt uses the uploaded data to model exponential communicable disease spread with instructions on using a growth rate of 0.3 to solve the differential equation.

"Given the current, de-identified, public use data (uploaded earlier) showing the number of infected individuals over time, solve a linear first-order differential equation that models exponential communicable disease spread. Use an initial infection count of 100 cases and a daily growth rate of 0.3. Predict the number of cases over the next five days and generate a plot to visualize the trend."

<u>Sample prompt for data analysis:</u> This prompt guides the user through analyzing SIR model output, focusing on key patterns, such as infection peaks and recovery rates.

"Using the output from a SIR model simulation, analyze the changes in the number of susceptible (S), infected (I), and recovered (R) individuals over a four-day period. Input values include S=900, I=10, R=0, transmission rate (β)=0.001, and recovery rate (γ)=0.1. Identify the patterns such as wen the infection rate peaks or how the recovery rate influences the decrease in the number communicable disease counts. Generate a chart that visualizes these changes to better understand communicable disease dynamics."

<u>Sample prompt for predictive modeling:</u> This prompt advises on updating the SIR model parameters based on new data (these data would be uploaded).

"Using a current, de-identified, public use data (uploaded earlier) from an ELR system, update the parameters of a SIR model as new cases and recoveries are reported. Adjust the transmission rate (β) and recovery rate (γ) based on changes in the population, such as vaccination campaign efforts. Generate forecasts for the next 14 days and visualize how these changes might alter the communicable disease trajectory."