Review

Theoretical Epidemiology Needs Urgent Attention in China

Yao Wang[1]; Zeyu Zhao[1]; Jia Rui[1]; Tianmu Chen[1,#]

ABSTRACT

The mathematical method to which theoretical epidemiology belongs is one of the three major methodologies in epidemiology. It is of great value in diagnosing infectious disease epidemic trends and evaluating the effectiveness of prevention and control measures. This paper aims to summarize the brief history of the development of theoretical epidemiology, common types of mathematical models, and key steps to develop a mathematical model. It also provides some thoughts and perspectives on the development and application of theoretical epidemiology in China.

Theoretical epidemiology, also referred to as mathematical modeling in epidemiology, is a significant component of epidemiological methodologies. It has played a crucial role in managing infectious disease outbreaks in China, particularly during the coronavirus disease 2019 (COVID-19) pandemic. The field aids in understanding disease transmission mechanisms, assessing epidemics on-site, and refining prevention strategies (*1*–*6*). Despite its contributions, there remain challenges in advancing theoretical epidemiology in China, notably in training public health professionals, applying disease control measures, and sustaining scholarly interest. The future trajectory of interest in this discipline, post the integration of COVID-19 prevention into China's infectious disease management, poses an important consideration.

A BRIEF HISTORY OF THE DEVELOPMENT OF THEORETICAL EPIDEMIOLOGY

Based on our comprehensive review of theoretical epidemiology literature, several seminal contributions have shaped our understanding of disease dynamics. R. A. Ross, in 1902, employed mathematical models to evaluate malaria interventions (*7*), while W. H. Hamer in 1906 highlighted the importance of reducing susceptible populations to control epidemics (*8*–*10*). In 1928, Reed-Frost elucidated infection spread across generations (*11*). McKendrick and Kermack between 1927 and 1935 introduced differential equations to depict disease transmission and the transmission threshold concept (*12*–*14*). George MacDonald further explored transmission dynamics by posing critical questions regarding infected individuals entering a population: What would occur if an infected individual entered a population? How many new infections would result (*15*)? The Threshold Theory in 1970 suggested that 80% vaccine coverage could eradicate smallpox. Mathematical modeling gained prominence following the 2003 SARS outbreak, aiding in transmissibility calculations and evaluating prevention measures (*16*–*20*). The 2009 Influenza A (H1N1) pandemic witnessed extensive modeling to assess vaccination effectiveness and establish influenza surveillance systems (*21*–*24*). The COVID-19 pandemic prompted accelerated development in theoretical epidemiology, with various models predicting trends, simulating transmission, and evaluating control measures to inform government decision-making (*3*–*4*,*25*–*30*). Subsequently, theoretical epidemiology has focused on refining models and addressing previous limitations, such as real-time database optimization and dynamic modeling to monitor disease progress (*31*–*32*).

OVERVIEW OF COMMON TYPES OF MATHEMATICAL MODELS IN THEORETICAL EPIDEMIOLOGY

Upon reviewing the literature, we identified two main types of mathematical models in theoretical epidemiology: "data-driven" models and "mechanism-driven" models.

"Data-driven" models encompass various approaches to investigate the connection between disease occurrence and time. These methods involve temporal regression models, control charts, time series models

like autoregressive integrated moving average model (ARIMA), Monte Carlo algorithmic models, gray theory models, neural network models, and related derivatives (*33*).

This category of methods focuses on analyzing large datasets of case information to comprehend the trends of diseases or health conditions over time. It involves examining the correlation between disease occurrences and time or other variables, developing models, and forecasting disease prevalence trends.

This model type offers the advantage of rapid and simple forecasting based on historical data. However, its primary drawbacks include using time as an independent variable and attributing disease development solely to temporal changes. These models often rely heavily on data, overlooking the infectious and epidemiologic aspects of disease (*34*). Typically, such models focus solely on prediction, utilizing historical data to simulate scenarios for forecasting, and rarely assess the effectiveness of infectious disease prevention and control strategies (*35*–*36*).

"Mechanism-driven" models are grounded in a comprehensive understanding of the disease process. These models simulate population dynamics, progressing from susceptibility to incubation, symptomatic or asymptomatic infection, and eventual recovery or mortality. They incorporate key epidemiological factors such as infectiousness, pathogenicity, and virulence (*1*,*6*,*37*–*39*). Additionally, they integrate practical prevention and control measures, including pharmacological interventions (e.g., antiviral drugs, antibiotics, vaccines) and non-pharmacological strategies (such as contact tracing, testing, school closures, hand hygiene, social distancing, mask-wearing), environmental disinfection, and vector control. These models are parameterized to reflect variations across different populations and adjust for interrelated parameters, ensuring alignment with real-world scenarios of disease transmission and control.

The models can be classified based on the object of study and parameter properties into two categories: 1) models with a group object and deterministic parameters, exemplified by the transmission dynamics model (*1*,*4*,*6*,*40*–*45*); 2) models with an individual object and stochastic parameters, illustrated by agent-based models (*46*–*47*), multiagent system (MAS) (*48*), cellular automata (CA) (*49*–*50*), and others.

Mechanism-driven models offer several key advantages: 1) They excel in uncovering the fundamental mechanisms of diseases and providing detailed explanations of epidemiological occurrences; 2) These models offer high flexibility in modeling. By employing a dynamics model that considers the natural history of the disease, clinical features, and epidemiological characteristics like the "three links" and "two factors," supplemented by biological evidence, and utilizing mathematical tools such as differential equations and computer technology, these models can effectively simulate disease transmission processes. This leads to enhanced predictive reliability, aiding in public health decision-making and the formulation of effective intervention strategies; 3) Mechanism-driven modeling necessitates a multidisciplinary approach, integrating knowledge and techniques from epidemiology, biology, statistics, and computer science. This interdisciplinary collaboration results in a more holistic comprehension of disease complexity, facilitating the proposition of comprehensive solutions.

The model exhibits significant limitations: 1) It is highly sensitive to the initial values and relies heavily on subjective and empirical factors for parameter settings, with unclear significance of individual parameters in practical usage; 2) Its rigorous mathematical and theoretical derivation demands a deep understanding of mathematics and computer science, limiting its universal applicability in frontline systems.

MODEL BUILDING PROCESS

When developing mechanism-driven models, researchers must consider key factors such as the research's objective and the compatibility of the data with the selected model. Subsequently, selecting a suitable model to evaluate transmissibility, morbidity, and other anticipated outcome measures is vital. The precise construction and selection of mechanism-driven models are essential to accurately replicate real-world scenarios and achieve desired outcomes. Hence, the process of model building and assessment calculations should adhere to a standardized and reproducible methodology. As such, our research team has introduced the "MODELS" modeling framework, comprising six primary steps and 19 sub-steps in the modeling process (*51*).

OUTLOOK

The current mathematical model has limitations in practical application. Therefore, it is essential to validate its accuracy using real data and enhance it

through adjustments. The model serves as a theoretical foundation and enhances data support for disease control. Here, personal opinions on its future development are discussed.

The "connotation" of theoretical epidemiology is expected to be further enhanced: The complexity and professional limitations of mathematical modeling have hindered extensive research and application of infectious disease mathematical modeling in Chinese public health. This is attributed to several factors: 1) Studies primarily focus on methodologies rather than practical implementation in public health; 2) Some research lacks clear parameter interpretation and comprehensive system dynamics; 3) Inconsistencies exist among experts regarding the natural history and transmission mechanisms of infectious diseases, for example, the percentage and transmissibility of asymptomatic infections are often overlooked.

The "outreach" of theoretical epidemiology is expected to be further expanded: The emergence of new infectious disease outbreaks present opportunities for the advancement of theoretical epidemiology, encompassing both theory and practical application. The theoretical aspect involves interdisciplinary collaboration, crucial for addressing complex public health issues by integrating various fields like infectious disease epidemiology, ecology, pathogen biology, genetics, applied mathematics, and computer science in the age of real-world and big data (*52*). The practical application aims to extend the reach of theoretical epidemiology from infectious diseases to areas such as chronic non-communicable diseases. This expansion is poised to stimulate the progress of public health strategies and aid in addressing global health challenges.

Theoretical epidemiology will play an important role in the training and pooling of human resources: Our research team surveyed the status of theoretical epidemiology instruction in academic and Center for Disease Control and Prevention (CDC) settings throughout China. Findings reveal a significant demand for theoretical epidemiology research and application within CDCs nationwide. However, there is a notable absence of theoretical epidemiology education in undergraduate programs across Chinese universities, hindering its practical application in epidemiology. Theoretical epidemiology is critical to understanding epidemic trends and transmission patterns of disease, and we therefore call for its inclusion in basic public health and preventive medicine curricula. This integration not only enhances the preparedness of future public health professionals but also reinforces the knowledge base within the discipline. By fostering theoretical epidemiology education, we aim to fortify the public health system to effectively combat infectious disease outbreaks and health emergencies.

Conflicts of interest: No conflicts of interest.

doi: 10.46234/ccdcw2024.096

[#] Corresponding author: Tianmu Chen, chentianmu@xmu.edu.cn.

[1] State Key Laboratory of Vaccines for Infectious Diseases, Xiang An Biomedicine Laboratory, School of Public Health, Xiamen University, Xiamen City, Fujian Province, China.

Submitted: December 28, 2023; Accepted: March 19, 2024

REFERENCES

1. Chen TM, Rui J, Wang QP, Zhao ZY, Cui JA, Yin L. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. Infect Dis Poverty 2020;9(1):24. https://doi.org/10.1186/s40249-020-00640-3.
2. Cheng XQ, Hu JL, Luo L, Zhao ZY, Zhang N, Hannah MN, et al. Impact of interventions on the incidence of natural focal diseases during the outbreak of COVID-19 in Jiangsu Province, China. Parasit Vectors 2021;14(1):483. https://doi.org/10.1186/s13071-021-04986-x.
3. Liu WK, Ye WJ, Zhao ZY, Liu C, Deng B, Luo L, et al. Modelling the emerging COVID-19 epidemic and estimating intervention effectiveness - Taiwan, China, 2021. China CDC Wkly 2021;3(34):716 – 9. https://doi.org/10.46234/ccdcw2021.177.
4. Zhao QL, Yang M, Wang Y, Yao LS, Qiao JG, Cheng ZY, et al. Effectiveness of interventions to control transmission of Reemergent cases of COVID-19 - Jilin Province, China, 2020. China CDC Wkly 2020;2(34):651 – 4. https://doi.org/10.46234/ccdcw2020.181.
5. Zhao ZY, Niu Y, Luo L, Hu QQ, Yang TL, Chu MJ, et al. The optimal vaccination strategy to control COVID-19: a modeling study in Wuhan City, China. Infect Dis Poverty 2021;10(1):140. https://doi.org/10.1186/s40249-021-00922-4.
6. Zhao ZY, Zhu YZ, Xu JW, Hu SX, Hu QQ, Lei Z, et al. A five-compartment model of age-specific transmissibility of SARS-CoV-2. Infect Dis Poverty 2020;9(1):117. https://doi.org/10.1186/s40249-020-00735-x.
7. Ross R. Some a priori pathometric equations. Br Med J 1915;1(2830):546 – 7. https://doi.org/10.1136/bmj.1.2830.546.
8. Hamer WH. The Milroy lectures on epidemic disease in England—the evidence of variability and of persistency of type (LECTURE I). Lancet 1906;167(4305):569-74. https://doi.org/10.1016/S0140-6736(01)80187-2.
9. Foppa IM. A historical introduction to mathematical modeling of infectious diseases: seminal papers in epidemiology. London: Academic Press. 2017. http://dx.doi.org/10.1016/C2014-0-01347-0.
10. Hamer WH. The Milroy lectures on Epidemic disease in England—the evidence of variability and of persistency of type (LECTURE III). Lancet 1906;167(4307):733-9. https://doi.org/10.1016/S0140-6736(01)80340-8.
11. Frost WH. Some conceptions of epidemics in general. Am J Epidemiol 1976;103(2):141 – 51. https://doi.org/10.1093/oxfordjournals.aje.a112212.
12. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Philos Trans Roy Soc A: Math Phys Eng Sci 1927;115(772):700 – 21. https://doi.org/10.1098/rspa.1927.0118.
13. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics. II.—The problem of endemicity. Philos Trans Roy Soc A: Math Phys Eng Sci 1932;138(834):55 – 83. https://doi.org/10.1098/rspa.1932.0171.
14. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics. III.—Further studies of the problem of endemicity. Philos Trans Roy Soc A: Math Phys Eng Sci 1933;141

(843):94 – 122. https://doi.org/10.1098/rspa.1933.0106.
15. MacDonald G. The epidemiology and control of malaria. London: Oxford University Press. 1957. https://searchworks.stanford.edu/view/L110357.
16. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science 2003;300(5627):1966 – 70. https://doi.org/10.1126/science.1086616.
17. Galvani AP, Lei XD, Jewell NP. Severe acute respiratory syndrome: temporal stability and geographic variation in case-fatality rates and doubling times. Emerg Infect Dis 2003;9(8):991 – 4. https://doi.org/10.3201/eid0908.030334.
18. Chowell G, Fenimore PW, Castillo-Garsow MA, Castillo-Chavez C. SARS outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism. J Theor Biol 2003;224(1):1 – 8. https://doi.org/10.1016/s0022-5193(03)00228-5.
19. Donnelly CA, Fisher MC, Fraser C, Ghani AC, Riley S, Ferguson NM, et al. Epidemiological and genetic analysis of severe acute respiratory syndrome. Lancet Infect Dis 2004;4(11):672 – 83. https://doi.org/10.1016/s1473-3099(04)01173-9.
20. Wang WD, Ruan SG. Simulating the SARS outbreak in Beijing with limited data. J Theor Biol 2004;227(3):369 – 79. https://doi.org/10.1016/j.jtbi.2003.11.014.
21. Jin Z, Zhang JP, Song LP, Sun GQ, Kan JL, Zhu HP. Modelling and analysis of influenza A (H1N1) on networks. BMC Public Health 2011;11 Suppl 1(Suppl 1):S9. http://dx.doi.org/10.1186/1471-2458-11-s1-s9.
22. Nishiura H, Iwata K. A simple mathematical approach to deciding the dosage of vaccine against pandemic H1N1 influenza. Euro Surveill 2009;14(45):19396. https://doi.org/10.2807/ese.14.45.19396-en.
23. Sypsa V, Hatzakis A. School closure is currently the main strategy to mitigate influenza A(H1N1)v: a modeling study. Euro Surveill 2009;14(24):19240. https://doi.org/10.2807/ese.14.24.19240-en.
24. Chen TM, Zhang SS, Feng J, Xia ZG, Luo CH, Zeng XC, et al. Mobile population dynamics and malaria vulnerability: a modelling study in the China-Myanmar border region of Yunnan Province, China. Infect Dis Poverty 2018;7(1):36. https://doi.org/10.1186/s40249-018-0423-6.
25. Ren JC, Yan YJ, Zhao HM, Ma P, Zabalza J, Hussain Z, et al. A novel intelligent computational approach to model epidemiological trends and assess the impact of non-pharmacological interventions for COVID-19. IEEE J Biomed Health Inform 2020;24(12):3551 – 63. https://doi.org/10.1109/jbhi.2020.3027987.
26. Odusanya OO, Odugbemi BA, Odugbemi TO, Ajisegiri WS. COVID-19: a review of the effectiveness of non-pharmacological interventions. Niger Postgrad Med J 2020;27(4):261 – 7. https://doi.org/10.4103/npmj.npmj_208_20.
27. AlJohani NI, Mutai K. Effect of non-pharmacological interventions on the COVID-19 epidemic in Saudi Arabia. Epidemiol Infect 2021;149:e252. https://doi.org/10.1017/s0950268821002612.
28. He GH, Zeng FF, Xiao JP, Zhao JG, Liu T, Hu JX, et al. When and how to adjust non-pharmacological interventions concurrent with booster vaccinations against COVID-19 - Guangdong, China, 2022. China CDC Wkly 2022;4(10):199 – 206. https://doi.org/10.46234/ccdcw2022.048.
29. Chowdhury R, Heng K, Shawon MSR, Goh G, Okonofua D, Ochoa-Rosales C, et al. Dynamic interventions to control COVID-19 pandemic: a multivariate prediction modelling study comparing 16 worldwide countries. Eur J Epidemiol 2020;35(5):389 – 99. https://doi.org/10.1007/s10654-020-00649-w.
30. Spiliotis K, Koutsoumaris CC, Reppas AI, Papaxenopoulou LA, Starke J, Hatzikirou H. Optimal vaccine roll-out strategies including social distancing for pandemics. iScience 2022;25(7):104575. https://doi.org/10.1016/j.isci.2022.104575.
31. Bhatia S, Imai N, Watson OJ, Abbood A, Abdelmalik P, Cornelissen T, et al. Lessons from COVID-19 for rescalable data collection. Lancet Infect Dis 2023;23(9):e383 – 8. https://doi.org/10.1016/s1473-3099(23)00121-4.
32. Rakhshan SA, Nejad MS, Zaj M, Ghane FH. Global analysis and prediction scenario of infectious outbreaks by recurrent dynamic model and machine learning models: a case study on COVID-19. Comput Biol Med 2023;158:106817. https://doi.org/10.1016/j.compbiomed.2023.106817.
33. Deng B, Niu Y, Xu JW, Rui J, Lin SN, Zhao ZY, et al. Mathematical models supporting control of COVID-19. China CDC Wkly 2022;4(40):895 – 901. https://doi.org/10.46234/ccdcw2022.186.
34. Baharom M, Ahmad N, Hod R, Arsad FS, Tangang F. The impact of meteorological factors on communicable disease incidence and its projection: a systematic review. Int J Environ Res Public Health 2021;18(21):11117. https://doi.org/10.3390/ijerph182111117.
35. Martens MJ, Logan BR. A unified approach to sample size and power determination for testing parameters in generalized linear and time-to-event regression models. Stat Med 2021;40(5):1121 – 32. https://doi.org/10.1002/sim.8823.
36. Qi BG, Liu NK, Yu SC, Tan F. Comparing COVID-19 case prediction between ARIMA model and compartment model - China, December 2019-April 2020. China CDC Wkly 2022;4(52):1185 – 8. https://doi.org/10.46234/ccdcw2022.239.
37. Zhao ZY, Chen Q, Wang Y, Chu MJ, Hu QQ, Hannah MN, et al. Relative transmissibility of shigellosis among different age groups: a modeling study in Hubei Province, China. PLoS Negl Trop Dis 2021;15(6):e0009501. https://doi.org/10.1371/journal.pntd.0009501.
38. Yang M, Cheng XQ, Zhao ZY, Li PH, Rui J, Lin SN, et al. Feasibility of controlling hepatitis E in Jiangsu Province, China: a modelling study. Infect Dis Poverty 2021;10(1):91. https://doi.org/10.1186/s40249-021-00873-w.
39. Wang Y, Zhao ZY, Wang MZ, Hannah MN, Hu QQ, Rui J, et al. The transmissibility of hepatitis C virus: a modelling study in Xiamen City, China. Epidemiol Infect 2020;148:e291. https://doi.org/10.1017/s0950268820002885.
40. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics—I. Bull Math Biol 1991;53(1-2):33 – 55. https://doi.org/10.1007/bf02464423.
41. Wang Y, Zhao ZY, Zhang H, Lin Q, Wang N, Ngwanguong Hannah M, et al. Estimating the transmissibility of hepatitis C: a modelling study in Yichang City, China. J Viral Hepat 2021;28(10):1464 – 73. https://doi.org/10.1111/jvh.13582.
42. Nakamura GM, Cardoso GC, Martinez AS. Improved susceptible-infectious-susceptible epidemic equations based on uncertainties and autocorrelation functions. Roy Soc Open Sci 2020;7(2):191504. https://doi.org/10.1098/rsos.191504.
43. Yang F, Yang Q, Liu XX, Wang P. SIS evolutionary game model and multi-agent simulation of an infectious disease emergency. Technol Health Care 2015;23 Suppl 2:S603-13. http://dx.doi.org/10.3233/thc-150999.
44. Brown GD, Porter AT, Oleson JJ, Hinman JA. Approximate Bayesian computation for spatial SEIR(S) epidemic models. Spat Spatiotemporal Epidemiol 2018;24:27 – 37. https://doi.org/10.1016/j.sste.2017.11.001.
45. Zhao ZY, Chen Q, Zhao B, Hannah MN, Wang N, Wang YX, et al. Relative transmissibility of shigellosis among male and female individuals: a modeling study in Hubei Province, China. Infect Dis Poverty 2020;9(1):39. https://doi.org/10.1186/s40249-020-00654-x.
46. Tracy M, Cerdá M, Keyes KM. Agent-based modeling in public health: current applications and future directions. Annu Rev Public Health 2018;39:77 – 94. https://doi.org/10.1146/annurev-publhealth-040617-014317.
47. Yin L, Zhang H, Li Y, Liu K, Chen TM, Luo W, et al. A data driven agent-based model that recommends non-pharmaceutical interventions to suppress Coronavirus disease 2019 resurgence in megacities. J Roy Soc Interface 2021;18(181):20210112. https://doi.org/10.1098/rsif.2021.0112.
48. Venkatramanan S, Sadilek A, Fadikar A, Barrett CL, Biggerstaff M, Chen JZ, et al. Forecasting influenza activity using machine-learned mobility map. Nat Commun 2021;12(1):726. https://doi.org/10.1038/s41467-021-21018-5.
49. Gwizdałła T. Viral disease spreading in grouped population. Comput Methods Programs Biomed 2020;197:105715. https://doi.org/10.1016/j.cmpb.2020.105715.
50. Escobar Ospina ME, Perdomo JG. A growth model of human papillomavirus type 16 designed from cellular automata and agent-based models. Artif Intell Med 2013;57(1):31 – 47. https://doi.org/10.1016/j.artmed.2012.11.001.
51. Rui J, Li KG, Wei HJ, Guo XH, Zhao ZY, Wang Y, et al. MODELS: a six-step framework for developing an infectious disease model. Infect Dis Poverty 2024;13(1):30. https://doi.org/10.1186/s40249-024-01195-3.
52. Arshad S, Khalid S, Javed S, Amin N, Nawaz F. Modeling the impact of the vaccine on the COVID-19 epidemic transmission via fractional derivative. Eur Phys J Plus 2022;137(7):802. https://doi.org/10.1140/epjp/s13360-022-02988-x.