

MATHEMATICAL MODELS FOR PREDICTING THE EPIDEMIC TRENDS OF BRUCELLOSIS

Sijai Liu^{1*}, Sarisak Soontornchai², Somchai Bovornkitti³, Wang Xuemei⁴

^{1,2,3}Public Health Program, Suan Sunandha Rajabhat University, Bangkok, Thailand,

⁴School of Public Health, Inner Mongolia Medical University, China, Inner Mongolia

E-Mail: s64584949023@ssru.ac.th^{1*}, sarisak.so@ssru.ac.th², somchai.bo@ssru.ac.th³,
wangxm_zsu@163.com⁴

ABSTRACT

Objectives: This review aims to provide a comprehensive analysis of mathematical models in epidemiology, specifically focusing on Brucellosis. The objectives include examining the classification, characteristics, and practical applications of various mathematical models, distinguishing between data-driven and mechanism models, and exploring the implications for enhancing public health efforts in addressing Brucellosis outbreaks. **Methods:** A thorough literature review was conducted by searching PubMed and the China National Knowledge Infrastructure for relevant scientific articles. The selected studies were analyzed to identify key aspects of mathematical models. The focus was on understanding how these models contribute to predicting and managing the spread of infectious diseases, with a special emphasis on Brucellosis. **Findings:** The review identifies and categorizes mathematical models used in epidemiology, providing insights into their unique characteristics and practical applications. Distinctions between data-driven and mechanism models are explored based on various criteria, including their basis, interpretability, generalization, data requirements, and applicability. **Novelty:** This review contributes novelty by synthesizing information on mathematical models in epidemiology, offering a comprehensive overview with a specific focus on Brucellosis. The exploration of distinctions between data-driven and mechanism models, along with their implications for public health, adds a unique perspective to the existing literature.

Keywords: Brucellosis, Model, Epidemic Trends

INTRODUCTION

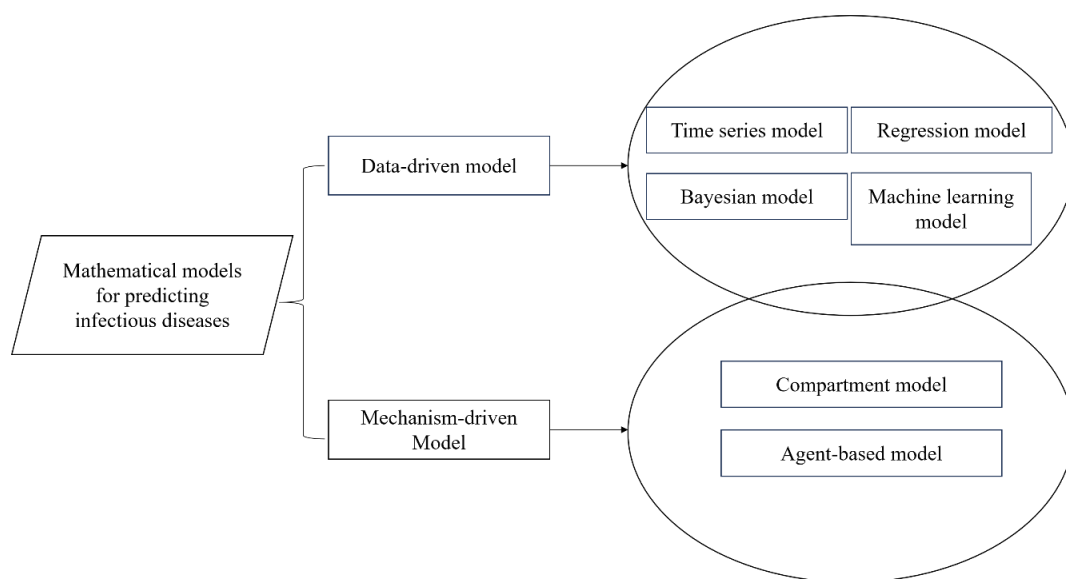
Brucellosis, also known as Malta fever, is the most widespread and mainly neglected zoonotic infectious disease in the world (Peng et al., 2021). The primary source of infection is domestic animals. Brucellosis typically spreads among livestock or wildlife, and then involves humans (Galińska & Zagórski, 2013). Humans acquire infection chiefly through direct/indirect contact with infected animals and their secretions, consumption of raw meat and dairy products, or inhalation of aerosols containing *Brucella*. At the early stage of brucellosis infection in humans, primary clinical symptoms include fever, excessive sweating, fatigue, and joint/muscle pain (Chen et al., 2016). Without prompt and effective treatment, complications such as gastrointestinal bleeding, endocarditis, reproductive system damage, and joint damage may develop (Jin et al., 2023). Brucellosis not only impairs labor ability and quality of life in mild cases, reducing economic income and increasing medical expenses, but in severe cases, it becomes a threat to human life and safety, resulting in significant disease burden and mortality. This poses significant challenges to public health worldwide. Despite the global impact of brucellosis on public health and socio-economic well-being, there exists a research gap in understanding and predicting its epidemic trends. The complex dynamics of brucellosis transmission, involving both animal and

human populations, necessitate a nuanced approach for effective control. Existing literature highlights the severity of the disease and its consequences but lacks a comprehensive exploration of mathematical models specifically tailored for predicting the epidemic trends of brucellosis. This study aims to provide a comprehensive analysis of mathematical models in brucellosis by examining their classification, characteristics, and practical applications for enhancing public health efforts in addressing Brucellosis outbreaks.

Mathematical Models for Predicting The Transmission Of Brucellosis

We have categorized mathematical models into two main types: data-driven models and mechanism-driven models. Data-driven models encompass time series models, regression models, Bayesian models and Machine learning. On the other hand, mechanism-driven models consist of compartment models and agent-based models. Figure 1 illustrates our systematic classification of forecasting models used for infection.

Figure 2 Classification of forecasting mathematical models used for brucellosis.



Data-Driven Models

“Data-driven model” is a broad concept referring to a variety of models that investigate the correlation between disease occurrences and time. Data-driven models rely heavily on real-world data to make predictions and inform decision-making(Phoon et al., 2022). These models leverage historical and real-time epidemiological data, demographic information, spatial information and other relevant datasets. The mean absolute error (MAE), Akaike information criterion (AIC), and r-squared (R²) can be used to assess accuracy of model fitting and prediction(Deng, 2022). Here are some typical data-driven models.

Time series models analyze past observations in time series data to make short-term forecasts(Chimmula & Zhang, 2020), particularly useful for predicting disease incidence. Further, many infectious diseases exhibit seasonal patterns. Time series models, such as seasonal ARIMA (SARIMA), can capture and quantify these seasonal variations, help to be implemented for real-time monitoring of disease trends. Common models for brucellosis include ARIMA(Wang, 2021), SARIMA(Zheng et al., 2021), Exponential Smoothing(Guan et al., 2018) and ARIMAX(Lee et al., 2013). However, their robustness is constrained by assumptions about the statistical attributes of the data and the statistical variability associated with estimating unidentified parameters(Siettos & Russo, 2013).

Regression models, such as Linear Regression, Poisson Regression, and Logistic Regression, utilize various input factors to estimate and predict disease outcomes(Nunez et al., 2011), providing valuable insights into the relationships between predictor variables and disease trends. These models contribute to forecasting, risk assessment, and decision-making in brucellosis control interventions(Kansiime et al., 2015; Megahed et al., 2022; Peng et al., 2020). However, it is crucial to consider factors like the transience assumption in the relationship between disease outcomes and independent variables, careful selection of model types (linear or nonlinear), and adherence to underlying

assumptions regarding the statistical properties of random error terms, including independence, normal distribution, and constant variance (Siettos & Russo, 2013).

Bayesian models, grounded in Bayesian probability theory, use probability distributions to reason and make inferences. These mathematical representations of real-world systems integrate prior knowledge, observational data, and probabilistic reasoning through a likelihood function to determine the posterior distribution (van de Schoot & Märtens, 2021). The posterior aids in parameter estimation, predictions, and uncertainty quantification. In brucellosis research and epidemiology, Bayesian models find diverse applications, including disease surveillance, epidemic forecasting, parameter estimation for epidemiological models, spatial analysis, health policy assessment, and optimal resource allocation (Cross et al., 2010; Sanogo & Abatih, 2013).

Machine learning for predicting infectious disease trends involves using computational algorithms and statistical models to analyze data and make predictions. Applications in brucellosis include supervised learning (e.g., decision trees, random forests, support vector machines) and deep learning (e.g., artificial neural networks, long short-term memory networks) (Bagheri et al., 2019; ZHANG et al., 2022). While machine learning shows promise, challenges include the need for large, high-quality datasets, potential overfitting, and the "black box" nature of some models, particularly deep learning. Interpretable models are crucial in epidemiology for informed public health decision-making.

Mechanism-Driven Models

Mechanism-driven models in epidemiology are designed to capture the underlying biological, behavioral and environmental processes that drive the spread of infectious diseases. These models aim to provide a deeper understanding of how disease transmission occurs and how interventions can be effectively implemented (Shi et al., 2023). Because some models of mechanism-driven incorporate detailed knowledge of the disease-causing agent, host factors, vectors (if applicable), and the interactions between them, they consider human behavior, healthcare practices, and interventions that can influence disease dynamics by changing model parameters or assumptions (Prybutok et al., 2022). At the same time, mechanism-driven models also require the estimation of model parameters based on biological or experimental data.

Compartment models are also known as epidemic dynamic models (Ma et al., 2009) and differential equation models (Deng, 2022). They are mathematical frameworks used to describe and understand the dynamics of disease transmission within the population. Basic compartmental models commonly include SIR Model (Susceptible-Infectious-Recovered) and SEIR Model (Susceptible-Exposed-Infectious-Removed). Meanwhile, there are many extended compartmental models such as SIRS and SIS, etc. While the mechanisms underlying various infectious diseases may differ, compartmental models are well-suited for analyzing transmission dynamics of brucellosis, forecasting epidemic trends (Li & Zheng, 2014), and optimizing preventative and control strategies for brucellosis (Ma et al., 2022). Predicting infectious diseases using compartmental models is a valuable tool, but the accuracy and relevance of predictions depend on careful model development, parameterization, validation, and a deep understanding of the specific disease dynamics under investigation.

Agent-Based Models (ABMs) are computational epidemiological modeling frameworks that simulate the behavior and interactions of individual agents within a population to study the spread and dynamics of infectious diseases (Hunter et al., 2018). The agents move through a simulated environment, interact with each other, and potentially become infected through the transportation network, the mobility of the population, demographics, and transmission between hosts, allowing for a detailed examination of disease transmission dynamics. ABMs reflect the real-world diversity of populations, manifested in agents varying in terms of age, sex, behavior, susceptibility, and contact patterns, and capture the intricate social, behavioral, and spatial factors that influence disease transmission. Therefore, they provide a powerful tool for exploring and understanding epidemic dynamics, evaluating control measures, and supporting informed public health responses (Havas et al., 2014).

Contrasting Characteristics Of Data-Driven And Mechanism-Driven Models

In the realm of infectious disease modeling, the choice between data-driven and mechanism-driven models involves navigating a spectrum of considerations. Each approach offers distinct advantages and limitations, shaping their suitability for specific aspects of epidemic prediction. Next, we delve into the contrasting characteristics of these models across key dimensions.

1) Basis of Modeling. Data-driven models primarily rely on historical data to identify relationships within the data and make predictions and cannot reflect the underlying processes driving disease dynamics. Mechanism-driven models are Rooted in understanding underlying mechanisms of disease transmission, using mathematical equations representing these mechanisms.

2) Interpretability. Data-driven models are lower interpretability in contrast to mechanism-driven models(Shi et al., 2023), especially deep learning models.

3) Generalization. Data-driven models excel at generalizing from historical data, and may struggle when faced with novel or unforeseen situations(Deng, 2022). Mechanism-driven models are designed to capture fundamental principles, allowing them to generalize to different scenarios and diseases. So, the generalization of data-driven models is lower compared to mechanism driven models.

4) Data Requirements. Data-driven models require large, high-quality datasets for training and validation. While data is still important for parameter estimation and validation, mechanism-driven models can often make predictions even with limited data, relying on prior knowledge and assumptions, sometimes the simulation outcomes of mechanisms can be incorrect(Shi et al., 2023).

5) Applicability. Data-driven models are Suitable for short-term predictions(Deng, 2022), so they are often utilized for early warning systems, short-term forecasting, and anomaly detection. Mechanism-driven models are Valuable for gaining insights into long-term dynamics and the impacts of interventions, leading to their use in in-depth epidemiological studies and assessments of policy impacts.

CONCLUSION

This study provides a comprehensive analysis of mathematical models for predicting the epidemic trends of Brucellosis. The research objectives focused on examining the classification, characteristics, and practical applications of various mathematical models, with a particular emphasis on distinguishing between data-driven and mechanism-driven models. The findings contribute valuable insights to enhance public health efforts in addressing Brucellosis outbreaks.

The review identified and categorized mathematical models into two main types: data-driven models and mechanism-driven models. Data-driven models, including time series models, regression models, Bayesian models, and machine learning, rely on real-world data for predictions and decision-making. On the other hand, mechanism-driven models, such as compartment models and agent-based models, aim to capture the underlying biological, behavioral, and environmental processes driving disease spread.

The contrasting characteristics of data-driven and mechanism-driven models, such as interpretability, generalization, data requirements, and applicability, were discussed. Data-driven models excel in short-term predictions, while mechanism-driven models provide valuable insights into long-term dynamics and the impacts of interventions.

However, it is important to acknowledge certain limitations. Data-driven models heavily depend on the availability of large, high-quality datasets, which may pose challenges in resource-limited settings. Mechanism-driven models, while rooted in underlying processes, require accurate parameter estimation based on biological or experimental data.

Future research directions should focus on bridging the gap between these models, integrating the strengths of both data-driven and mechanism-driven approaches. Additionally, efforts should be made to develop models that account for the complex interplay between animal and human populations, considering the unique dynamics of Brucellosis transmission.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to Suan Sunandha Rajabhat University for invaluable help throughout this research.

This study was supported by Natural Science Foundation of Inner Mongolia (2023QN07004).

REFERENCES

- Bagheri, H., Tapak, L., Karami, M., Amiri, B., & Cherghi, Z. (2019). Epidemiological features of human brucellosis in Iran (2011-2018) and prediction of brucellosis with data-mining models. *Journal of Research in Health Sciences*, 19(4), e00462.
- Chen, Q. L., S.Yin, W.Zhou, H.Li, , Y.Mu, & D.Li, Z. Y., H.Yang, W. (2016). Epidemic characteristics, high-risk townships and space-time clusters of human brucellosis in Shanxi Province of China, 2005-2014. *BMC Infect Dis*, 16(1), 760. <https://doi.org/10.1186/s12879-016-2086-x>
- Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons Fractals*, 135, 109864.
- Cross, P. C., Heisey, D. M., Scurlock, B. M., Edwards, W. H., Ebinger, M. R., & Brennan, A. (2010). Mapping brucellosis increases relative to elk density using hierarchical Bayesian models. *PLoS One*, 5(4), e10322.
- Deng, B. N., Yan Xu, Jingwen Rui, Jia Lin, Shengnan Zhao, Zeyu Yu, Shanshan Guo, Yichao Luo, Li Chen, Tianmu (2022). Mathematical Models Supporting Control of COVID-19. *China CDC Weekly*, 4(40), 895.
- Galińska, E. M., & Zagórski, J. (2013). Brucellosis in humans - Etiology, diagnostics, clinical forms [Review]. *Annals of Agricultural and Environmental Medicine*, 20(2), 233-238. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84878931964&partnerID=40&md5=760a32a600c082bceec56af1b8f62b3e>
- Guan, P., Wu, W., & Huang, D. (2018). Trends of reported human brucellosis cases in mainland China from 2007 to 2017: an exponential smoothing time series analysis. *Environmental health preventive medicine*, 23(1), 1-7.
- Havas, K., Boone, R., Hill, A., Salman, M., & health, p. (2014). A Brucellosis Disease Control Strategy for the Kakheti Region of the Country of Georgia: An Agent-Based Model. *Zoonoses and public health*, 61(4), 260-270.
- Hunter, E., Mac Namee, B., & Kelleher, J. D. (2018). A comparison of agent-based models and equation-based models for infectious disease epidemiology.
- Jin, M., Fan, Z., Gao, R., Li, X., Gao, Z., & Wang, Z. (2023). Research progress on complications of Brucellosis. *Frontiers in Cellular Infection Microbiology*, 13, 1136674.
- Kansiime, C., Rutebemberwa, E., Asiimwe, B. B., Makumbi, F., Bazira, J., & Mugisha, A. (2015). Annual trends of human brucellosis in pastoralist communities of south-western Uganda: a retrospective ten-year study. *Infectious diseases of poverty*, 4, 1-8.
- Lee, H. S., Her, M., Levine, M., & Moore, G. E. (2013). Time series analysis of human and bovine brucellosis in South Korea from 2005 to 2010. *Preventive Veterinary Medicine*, 110(2), 190-197.
- Li, M. S., G. Zhang, J. Jin, Z. Sun, X. Wang, Y.Huang, B., & Zheng, Y. (2014). Transmission dynamics and control for a brucellosis model in Hinggan League of Inner Mongolia, China. *Math Biosci Eng*, 11(5), 1115-1137. <https://doi.org/10.3934/mbe.2014.11.1115>
- Ma, X., Sun, G.-Q., Wang, Z.-H., Chu, Y.-M., Jin, Z., & Li, B.-L. (2022). Transmission dynamics of brucellosis in Jilin province, China: Effects of different control measures. *Communications in Nonlinear Science and Numerical Simulation*, 114. <https://doi.org/10.1016/j.cnsns.2022.106702>

- Ma, Z., Zhou, Y., & Wu, J. (2009). Modeling and dynamics of infectious diseases (Vol. 11). World Scientific.
- Megahed, A., Kandeel, S., Alshaya, D. S., Attia, K. A., AlKahtani, M. D., Albohairy, F. M., & Selim, A. (2022). A comparison of logistic regression and classification tree to assess brucellosis associated risk factors in dairy cattle. *Preventive Veterinary Medicine*, 203, 105664.
- Nunez, E., Steyerberg, E. W., & Nunez, J. (2011). Regression modeling strategies. *Revista Española de Cardiología*, 64(6), 501-507.
- Peng, C., Li, Y.-J., Huang, D.-S., & Guan, P. (2020). Spatial-temporal distribution of human brucellosis in mainland China from 2004 to 2017 and an analysis of social and environmental factors. *Environmental health preventive medicine*, 25(1), 1-14.
- Peng, L., Liang, X., Zhu, L., Liang, C., Liu, C., & Chi, X. (2021). Risk Factor Analysis of Brucellosis in Hulunbuir, China, Using the Propensity Score Matching Method. *Jundishapur Journal of Microbiology*, 14(6). <https://doi.org/10.5812/jjm.113528>
- Phoon, K.-K., Ching, J., & Shuku, T. (2022). Challenges in data-driven site characterization. *Georisk: Assessment Management of Risk for Engineered Systems Geohazards*, 16(1), 114-126.
- Prybutok, A. N., Cain, J. Y., Leonard, J. N., & Bagheri, N. (2022). Fighting fire with fire: deploying complexity in computational modeling to effectively characterize complex biological systems. *Current Opinion in Biotechnology*, 75, 102704.
- Sanogo, M. T., Eric Achi, Yaba L Fretin, David Michel, Patrick, & Abatih, E. B., Dirk Saegerman, Claude (2013). Bayesian estimation of the true prevalence, sensitivity and specificity of the Rose Bengal and indirect ELISA tests in the diagnosis of bovine brucellosis. *The Veterinary Journal*, 195(1), 114-120.
- Shi, X.-L., Wei, F.-F., & Chen, W.-N. (2023). A swarm-optimizer-assisted simulation and prediction model for emerging infectious diseases based on SEIR. *Complex Intelligent Systems*, 9(2), 2189-2204.
- Siettos, C. I., & Russo, L. (2013). Mathematical modeling of infectious disease dynamics. *Virulence*, 4(4), 295-306. <https://doi.org/10.4161/viru.24041>
- van de Schoot, R. D., Sarah King, Ruth Kramer, Bianca, & Märtens, K. T., Mahlet G Vannucci, Marina Gelman, Andrew Veen, Duco Willemsen, Joukje. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, 1(1), 1.
- Wang, Y. W., Y. Zhang, L. Wang, A. Yan, Y. Chen, Y. Li, X. Guo, A. Robertson, I. D. (2021). An epidemiological study of brucellosis on mainland China during 2004-2018. *Transbound Emerg Dis*, 68(4), 2353-2363. <https://doi.org/10.1111/tbed.13896>
- ZHANG, R., WANG, X., ZHANG, Y., & LI, Y. (2022). Comparison of random forests and SARIMA in Predicting Brucellosis Incidence. *J Journal of Public Health Preventive Medicine*, 1-5.
- Zheng, Y., Zhang, L., Wang, C., Wang, K., Guo, G., Zhang, X., & Wang, J. (2021). Predictive analysis of the number of human brucellosis cases in Xinjiang, China. *Scientific reports*, 11(1), 11513.