

## Editorial

# The Epidemiology of Infectious Diseases Meets AI: A Match Made in Heaven

Ankur Bothra <sup>1,\*</sup> , Yongguo Cao <sup>2</sup> , Jiří Černý <sup>3</sup> and Gunjan Arora <sup>4,\*</sup><sup>1</sup> National Institute of Allergy and Infectious Diseases (NIAID), Bethesda, MD 20892, USA<sup>2</sup> College of Veterinary Medicine, Jilin University, Changchun 130062, China<sup>3</sup> Centre for Infectious Animal Diseases, Faculty of Tropical AgriSciences, Czech University of Life Sciences Prague, 165 00 Prague, Czech Republic<sup>4</sup> Section of Infectious Diseases, Yale School of Medicine, Yale University, New Haven, CT 06519, USA

\* Correspondence: ankur.bothra@nih.gov (A.B.); arorag1983@gmail.com (G.A.)

## An Opinion

Infectious diseases remain a major threat to public health [1–3]. This Special Issue on the Epidemiology of Infectious Disease will cover studies related to the emergence, transmission, and containment of infectious diseases, including new research showing potential therapeutic interventions. This Issue will encompass viral, bacterial, and parasitic diseases with an emphasis on emerging research areas such as modeling, clinical studies, longitudinal cohort, and case–control studies, systems biology approaches, artificial intelligence (AI), machine learning, and other molecular and immunological studies [4–6].

AI and machine learning can be employed to study complex interactions between different biological systems, such as signaling pathways and metabolic networks, to advance our understanding of various biological phenomena and improve the diagnosis and treatment of diseases [5,7,8]. These technologies have the potential to significantly impact biological research in a variety of areas, including infectious diseases and epidemiology, as highlighted in the Special Issue of the MDPI journal *Pathogens* entitled “Papers on the Epidemiology of Infectious Diseases”.

AI and machine learning can be used to analyze large datasets, such as genomic data, to identify patterns and trends relevant to the understanding and treatment of infectious diseases [9–12]. For example, machine learning algorithms have been utilized to identify potential drug targets for SARS-CoV-2, which causes COVID-19 [13,14]. In addition, AI and machine learning can be employed to predict the likelihood of certain outcomes, such as the spread of a disease, based on historical data and by analyzing datasets generated by epidemiological studies. This can aid epidemiologists in preventing or mitigating outbreaks of infectious diseases, such as influenza and HIV.

AI can also be utilized to build predictive models that help researchers understand the relationships between different variables, such as gene expression and disease risk, interactions between pathogens and host organisms at the molecular level, and complex molecular interactions within biomolecules. Examples of the use of AI in biological research include AlphaFold [15,16], which can predict the secondary and tertiary structure of proteins with a high level of confidence [17,18], and DeepMind, which analyzes images of cells or tissues to identify specific features or patterns relevant to research.

An application that recently received media attention is AI’s capability in processing natural languages. In this regard, Open AI’s chatbot, named ChatGPT, can process natural language text and can be used to perform complex analysis and help non-English-speaking epidemiologists to draft articles. ChatGPT can provide the definitions of scientific terms, generate prevalence and risk factor maps of any disease, and so on. These efforts can revolutionize biological science research, but the output from such AI platforms needs to be verified, especially in many social, economic, behavioral, and epidemiological studies



**Citation:** Bothra, A.; Cao, Y.; Černý, J.; Arora, G. The Epidemiology of Infectious Diseases Meets AI: A Match Made in Heaven. *Pathogens* **2023**, *12*, 317. <https://doi.org/10.3390/pathogens12020317>

Received: 27 January 2023

Revised: 8 February 2023

Accepted: 10 February 2023

Published: 15 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

that deal with large datasets. Therefore, it demands the assessment of security issues and other challenges that arise in the use of ChatGPT. The validation of AI-based research using standard conventional methods is critical to check the flow of misinformation. The goal for the future development of such AIs is to train them to use relevant data sources; otherwise, we will jump from the post-factual age directly into the non-factual age. Anecdotally, these tools have been found to provide non-existing references in digital epidemiology.

Overall, the use of AI and machine learning in biological research, including epidemiology, has the potential to speed up the research process, improve accuracy and precision, and allow researchers to tackle more complex questions. We encourage our readers to use AI and publish their research in the MDPI journal *Pathogens*.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Dutta, D.; Naiyer, S.; Mansuri, S.; Soni, N.; Singh, V.; Bhat, K.H.; Singh, N.; Arora, G.; Mansuri, M.S. COVID-19 Diagnosis: A Comprehensive Review of the RT-qPCR Method for Detection of SARS-CoV-2. *Diagnostics* **2022**, *12*, 1503. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Shankaran, D.; Arumugam, P.; Vasanthakumar, R.P.; Singh, A.; Bothra, A.; Gandotra, S.; Rao, V. Modern Clinical Mycobacterium tuberculosis Strains Leverage Type I IFN Pathway for a Proinflammatory Response in the Host. *J. Immunol.* **2022**, *209*, 1736–1745. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Bothra, A.; Arumugam, P.; Panchal, V.; Menon, D.; Srivastava, S.; Shankaran, D.; Nandy, A.; Jaisinghani, N.; Singh, A.; Gokhale, R.S.; et al. Phospholipid homeostasis, membrane tenacity and survival of Mtb in lipid rich conditions is determined by MmpL11 function. *Sci. Rep.* **2018**, *8*, 8317. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Arora, G.; Bothra, A.; Prosser, G.; Arora, K.; Sajid, A. Role of post-translational modifications in the acquisition of drug resistance in Mycobacterium tuberculosis. *FEBS J.* **2021**, *288*, 3375–3393. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Arora, G.; Joshi, J.; Mandal, R.S.; Shrivastava, N.; Virmani, R.; Sethi, T. Artificial Intelligence in Surveillance, Diagnosis, Drug Discovery and Vaccine Development against COVID-19. *Pathogens* **2021**, *10*, 1048. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Thiebaut, R.; Thiessard, F.; Section Editors for the, I.Y.S.o.P.H.; Epidemiology, I. Artificial Intelligence in Public Health and Epidemiology. *Yearb. Med. Inform.* **2018**, *27*, 207–210. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Cuperlovic-Culf, M.; Nguyen-Tran, T.; Bennett, S.A.L. Machine Learning and Hybrid Methods for Metabolic Pathway Modeling. *Methods Mol. Biol.* **2023**, *2553*, 417–439. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Jang, W.D.; Kim, G.B.; Kim, Y.; Lee, S.Y. Applications of artificial intelligence to enzyme and pathway design for metabolic engineering. *Curr. Opin. Biotechnol.* **2022**, *73*, 101–107. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Chakraborty, R.; Hasija, Y. Predicting MicroRNA Sequence Using CNN and LSTM Stacked in Seq2Seq Architecture. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2020**, *17*, 2183–2188. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Awasthi, R.; Guliani, K.K.; Khan, S.A.; Vashishtha, A.; Gill, M.S.; Bhatt, A.; Nagori, A.; Gupta, A.; Kumaraguru, P.; Sethi, T. VacSIM: Learning effective strategies for COVID-19 vaccine distribution using reinforcement learning. *Intell. Based Med.* **2022**, *6*, 100060. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Nagpal, S.; Pal, R.; Ashima; Tyagi, A.; Tripathi, S.; Nagori, A.; Ahmad, S.; Mishra, H.P.; Malhotra, R.; Kutum, R.; et al. Genomic Surveillance of COVID-19 Variants With Language Models and Machine Learning. *Front. Genet.* **2022**, *13*, 858252. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Pandey, R.; Gautam, V.; Pal, R.; Bandhey, H.; Dhingra, L.S.; Misra, V.; Sharma, H.; Jain, C.; Bhagat, K.; Arushi; et al. A machine learning application for raising WASH awareness in the times of COVID-19 pandemic. *Sci. Rep.* **2022**, *12*, 810. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Paul, D.; Sanap, G.; Shenoy, S.; Kalyane, D.; Kalia, K.; Tekade, R.K. Artificial intelligence in drug discovery and development. *Drug Discov. Today* **2021**, *26*, 80–93. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Chan, H.C.S.; Shan, H.; Dahoun, T.; Vogel, H.; Yuan, S. Advancing Drug Discovery via Artificial Intelligence. *Trends Pharmacol. Sci.* **2019**, *40*, 592–604. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. [\[CrossRef\]](#) [\[PubMed\]](#)

17. Bou-Nader, C.; Bothra, A.; Garboczi, D.N.; Leppla, S.H.; Zhang, J. Structural basis of R-loop recognition by the S9.6 monoclonal antibody. *Nat. Commun.* **2022**, *13*, 1641. [[CrossRef](#)] [[PubMed](#)]
18. Strecker, J.; Demircioglu, F.E.; Li, D.; Faure, G.; Wilkinson, M.E.; Gootenberg, J.S.; Abudayyeh, O.O.; Nishimasu, H.; Macrae, R.K.; Zhang, F. RNA-activated protein cleavage with a CRISPR-associated endopeptidase. *Science* **2022**, *378*, 874–881. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.