

Global infectious disease surveillance technologies and data sharing protocols

Edited by

Jiaying Li and Jie Huang

Coordinated by

Xiangqi Meng

Published in

Frontiers in Public Health



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-6864-4
DOI 10.3389/978-2-8325-6864-4

Generative AI statement

Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Global infectious disease surveillance technologies and data sharing protocols

Topic editors

Jiaying Li — The University of Queensland, Australia

Jie Huang — Southern University of Science and Technology, China

Topic coordinator

Xiangqi Meng — Southern University of Science and Technology, China

Citation

Li, J., Huang, J., Meng, X., eds. (2025). *Global infectious disease surveillance technologies and data sharing protocols*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-6864-4

Table of contents

- 05 **Editorial: Global infectious disease surveillance technologies and data sharing protocols**
Jie Huang, Zhiyan Ding and Jiaying Li
- 09 **Infectious disease surveillance needs for the United States: lessons from Covid-19**
Marc Lipsitch, Mary T. Bassett, John S. Brownstein, Paul Elliott, David Eyre, M. Kate Grabowski, James A. Hay, Michael A. Johansson, Stephen M. Kissler, Daniel B. Larremore, Jennifer E. Layden, Justin Lessler, Ruth Lynfield, Duncan MacCannell, Lawrence C. Madoff, C. Jessica E. Metcalf, Lauren A. Meyers, Sylvia K. Ofori, Celia Quinn, Ana I. Bento, Nicholas G. Reich, Steven Riley, Roni Rosenfeld, Matthew H. Samore, Rangarajan Sampath, Rachel B. Slayton, David L. Swerdlow, Shaun Truelove, Jay K. Varma and Yonatan H. Grad
- 25 **Use of wastewater from passenger ships to assess the movement of COVID-19 and other pathogenic viruses across maritime international boundaries**
Davey L. Jones, Mathew Bridgman, Cameron Pellett, Andrew J. Weightman, Peter Kille, Álvaro García Delgado, Gareth Cross, Steve Cobley, Helen Howard-Jones, David R. Chadwick and Kata Farkas
- 35 **Detection of influenza virus in urban wastewater during the season 2022/2023 in Sicily, Italy**
Carmelo Massimo Maida, Walter Mazzucco, Walter Priano, Roberta Palermo, Giorgio Graziano, Claudio Costantino, Arianna Russo, Gina Andolina, Isabella Restivo, Viviana Giangreco, Francesca Rita Iaia, Arianna Santino, Rita Li Muli, Valeria Guzzetta, Francesco Vitale and Fabio Tramuto
- 45 **Implementation and performance evaluation of an integrated specimen referral system in Burkina Faso using the national courier services (2020–2022)**
Emilie Dama, Souleymane Porgo, Yahn-Cedric Ake, Issaka Yameogo, Sandrine Gampini, Aime-Gilles A. Adjami, Abdoulaye Nikiema, Mory Kamate, Felix Tarbangdo, Romial Sawadogo, Charles Sawadogo, Hamed S. Ouedraogo, Habibata Zerbo, Lila Rahalison, Isaïe Medah, Anicet G. Dahourou, Rebecca Greco-Kone and Flavien H. Ake
- 57 **A follow up report validating long term predictions of the COVID-19 epidemic in the UK using a dynamic causal model**
Cam Bowie and Karl Friston
- 71 **Exploring the influence of environmental indicators and forecasting influenza incidence using ARIMAX models**
Xiaoyan Zheng, Qingquan Chen, Mengcai Sun, Quan Zhou, Huanhuan Shi, Xiaoyang Zhang and Youqiong Xu

- 82 **Longitudinal wastewater-based surveillance of SARS-CoV-2 during 2023 in Ethiopia**
Daniel Abera Dinssa, Gebremedhin Gebremicael, Yohannes Mengistu, Noah C. Hull, Dinknesh Chalchisa, Girma Berhanu, Atsbeha Gebreegziabxier, Ashley Norberg, Sarah Snyder, Sarah Wright, Waktole Gobena, Adugna Abera, Yohannes Belay, Dawit Chala, Melaku Gizaw, Mesay Getachew, Kirubel Tesfaye, Mesfin Tefera, Mahlet Belachew, Tegegne Mulu, Solomon Ali, Abebaw Kebede, Daniel Melese, Saro Abdella, Tobias F. Rinke de Wit, Yenew Kebede, Mesay Hailu, Dawit Wolday, Masresha Tessema and Getachew Tollera
- 98 **Advanced optimal control approaches for immune boosting and clinical treatment to enhance dengue viremia models using ABC fractional-order analysis**
G. M. Vijayalakshmi, M. Ariyanatchi, Lenka Cepova and Krishnasamy Karthik
- 115 **Effect and prediction of long-term weather and pollutant exposure on hemorrhagic fever with renal syndrome: based on statistical models**
Weiming Hou



OPEN ACCESS

EDITED BY

Marc Jean Struelens,
Université libre de Bruxelles, Belgium

*CORRESPONDENCE

Jie Huang
✉ huangj@sustech.edu.cn
Jiaying Li
✉ jiaying.li@sydney.edu.au

RECEIVED 31 July 2025

ACCEPTED 11 August 2025

PUBLISHED 29 August 2025

CITATION

Huang J, Ding Z and Li J (2025) Editorial:
Global infectious disease surveillance
technologies and data sharing protocols.
Front. Public Health 13:1676987.
doi: 10.3389/fpubh.2025.1676987

COPYRIGHT

© 2025 Huang, Ding and Li. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Editorial: Global infectious disease surveillance technologies and data sharing protocols

Jie Huang^{1*}, Zhiyan Ding² and Jiaying Li^{3*}

¹School of Public Health and Emergency Management, Southern University of Science and Technology, Shenzhen, China, ²International Curriculum Center of RDFZ, Beijing, China, ³School of Civil Engineering, University of Sydney, Sydney, NSW, Australia

KEYWORDS

global health, infectious disease, surveillance, data sharing, pandemic

Editorial on the Research Topic

Global infectious disease surveillance technologies and data sharing protocols

In this Research Topic titled “*Global Infectious Disease Surveillance Technologies and Data Sharing Protocols*”, we issued a call for papers “*related to the research, practice and architectural design of surveillance technologies and data sharing protocols that empower infectious disease prevention and preparedness at a global scale*”. We specifically encouraged submissions showcasing innovative ideas and discoveries in wastewater-based surveillance, innovative contactless technologies that could be deployed in public transportation vehicles, especially those across national and regional borders. At the time of the first SARS outbreak in the beginning of this century, the concept of wastewater-based surveillance was unimaginable. Today, it has become a reality, presenting a promising component of an integrated global aircraft-based genomic surveillance network (1). We also welcomed technical contributions employing artificial intelligence (AI) and blockchain technologies to enable real-time, transparent global data sharing. As articulated in our initial call: “*Once the global pandemic situation could be monitored and checked on anyone’s smart-phone, like those for weather and air pollution, the lofty ideology of global pandemic prevention will be realized, from bottom up*”.

This Research Topic accepted a total of nine articles, with four primarily focused on sampling and experimental methodologies (i.e. “**wet**” studies), while four focused on statistical and computational modeling (i.e. “**dry**” studies). In December 2022, at the moment during China’s transition in COVID-19 control strategies, we published a commentary titled “*The World Needs a ‘Pandemic’ Solution for a Pandemic Problem*” (2). There, we introduced the term “pandemic” (pan-da-mic). There, “da” refers to data applications widely used and needed to fight against and prevent pandemics, while “mic” means microbiology and, in particular, various omics technologies. Therefore, the concept of “pandemic” stresses the essential convergence of biotechnology (wet) and information technology (dry) in modern surveillance frameworks. A particularly noteworthy contribution in this Research Topic is the review by [Lipsitch et al.](#), titled “*Infectious Disease Surveillance Needs for the United States: Lessons from COVID-19*”. The authors presented a comprehensive roadmap for improving national and global infectious disease surveillance systems. Drawing on insights from the COVID-19 pandemic,

the article identified critical data types and infrastructure elements necessary to support real-time decision-making. The authors emphasized the integration of diverse data streams, including mobility patterns, internet search trends, clinical diagnostics, and wastewater signals, into purpose-built, responsive systems. Importantly, the article highlighted the importance of equitable and locally adaptive systems, capable of informing interventions not only during acute crises but also for ongoing public health challenges.

Three of the wet studies feature the term “wastewater” in their titles, reflecting the growing importance of wastewater-based surveillance in the global infectious disease monitoring landscape. Jones et al. investigated the feasibility of using *wastewater from passenger ships* as a surveillance tool for viral pathogens crossing maritime borders. Their study demonstrated successful detection of SARS-CoV-2 and norovirus in blackwater collected from short-haul ferries operating between the United Kingdom and Ireland. These findings validated the potential of maritime wastewater-based surveillance for tracking pathogen transmission across international boundaries, offering an important monitoring tool in the context of international travel. Maida et al. presented *urban wastewater surveillance in Sicily* during the 2022/2023 influenza season. The temporal trends of influenza viral RNA in wastewater were found to mirror clinical case trends, indicating the potential of wastewater-based surveillance as a non-invasive and cost-effective complement to traditional influenza surveillance in urban European settings. Dinssa et al. conducted a longitudinal study of *SARS-CoV-2 in Ethiopian wastewater* throughout 2023. They found a high positivity rate in untreated wastewater samples and a strong correlation between viral RNA levels and COVID-19 case trends. Their work underscored the capability of wastewater-based surveillance in low-resource settings, where limited access to clinical diagnostics may lead to underestimation of infection prevalence. This work provided compelling evidence that wastewater-based surveillance can fill critical surveillance gaps in resource-limited contexts. The fourth study from Dama et al. described the implementation of an integrated specimen reference system in Burkina Faso. This system employed existing courier networks to transport human biological specimens for priority diseases including COVID-19 from district-level clinics to reference laboratories in Burkina Faso. This innovative system achieved >99% on-time delivery with preserved sample integrity, proving that scalable, cost-effective logistical infrastructure can significantly enhance disease surveillance outcomes, especially for time-sensitive or high-risk conditions like the COVID-19 pandemic. Together, these four studies exemplify diverse and pragmatic approaches to enhance the front-line data collection for infectious disease surveillance, spanning novel applications of wastewater-based surveillance to innovations in biospecimen logistics.

All four dry studies include “model(s)” in their titles and collectively reflect a broad spectrum of modeling strategies and regional applications. Bowie and Friston assessed the predictive validity of a *dynamic causal model (DCM)* for long-term outcomes of the COVID-19 pandemic. While DCM captured several key pandemic dynamics, it tended to overestimate deaths and hospitalization rates due to fixed assumptions about virulence

persistence. Their work offered a critical reflection on modeling assumptions and proposed more adaptive model frameworks that incorporate evolving population immunity. Hou applied time-series and machine learning methods to examine the epidemiology of *hemorrhagic fever with renal syndrome (HFRS)* in relation to environmental drivers. By integrating meteorological and air pollutant data using distributed lag non-linear models and support vector machines, the study provided a refined seasonal risk framework for HFRS outbreaks, highlighting the role of air quality as a significant predictor of disease outbreaks. Zheng et al. evaluated *ARIMAX models to predict influenza incidence* in Fuzhou, China, incorporating air pollutants and meteorological indicators. They found that PM₁₀ was a particularly strong predictor and demonstrated that the inclusion of environmental indicators improved model accuracy. These findings provided practical implications for real-time influenza forecasting and public health early warning systems. Vijayalakshmi et al. developed an *optimal control framework for dengue transmission* using fractional-order differential equations based on the Atangana-Baleanu Caputo (ABC) calculus. Their mathematical model accounted for both symptomatic and asymptomatic infections and demonstrated that immune boosting and clinical treatment strategies could significantly reduce disease burden when integrated into control policies. Collectively, these four modeling papers presented the richness and diversity of analytic approaches that can support infectious disease prediction, environmental risk assessment, and intervention optimization across varied geographic and epidemiological contexts.

The COVID-19 pandemic, once a defining global crisis, now feels like a distant memory. Yet today, its urgency has largely receded from public consciousness and institutional agendas. As Darwin’s theory of evolution suggests, humans are remarkably adaptive. But adaptation should not become synonymous with complacency. This moment calls for difficult, but necessary questions: Has anything fundamentally changed in the academic, operational, or policy landscape of global public health? If a COVID-like pandemic was to emerge tomorrow, would policymakers and societies respond more wisely, more swiftly, and more effectively? In China, as of July 2025, the infectious disease currently making headlines is the Chikungunya virus, transmitted by mosquitoes (3). In response, public health authorities have encouraged the public to drain stagnant water and apply insect repellent. These measures, while useful, have remained largely unchanged for over a century. Such public health intervention should reflect the leap in infrastructure, technology, or governance that reflects the lessons of COVID-19.

The echoes of “*I have a dream*” from Martin Luther King Jr. and “*we choose to go to the moon*” from President Kennedy continue to inspire visionary thinking. In the realm of global infectious disease surveillance, what are the equivalent aspirations? Do we have a unifying “dream” or a collective “moonshot” in this space? Or are we still navigating a fragmented landscape of national agendas and disconnected efforts? Public health is classically defined as “*the science and art of preventing disease, prolonging life, and promoting health through the organized efforts and informed choices of society*” (4). While biology and medicine anchor the scientific foundation, the “art” lies in policy, culture, communication,

and the complexities of human behavior. From this perspective, public health is therefore inherently interdisciplinary, but this very breadth also risks diffusion of focus and a lack of accountability. Without concrete systems and enforceable structures, the noble ideals of public health remain vulnerable to drift. In this context, the World Health Organization (WHO) should evolve from a reactive body to a proactive global leader. It should articulate a clear, realistic, and actionable strategy for global infectious disease surveillance that is able to propel nations into coordinated efforts. Much like the International Olympic Committee (IOC), which established universal anti-doping protocols and inspired a shared framework for athletic integrity, WHO should provide both the inspiration and the infrastructure to coordinate global health preparedness. It should not only be the moral authority but also the architect of scalable solutions, setting enforceable global standards and guiding strategic investments to ensure no country is left behind.

In our original call for papers, we referenced the global anti-doping protocol as an instructive model: “*An exemplary is the protocol of world doping control, where all nations are obligated by the International Olympics Committee (IOC) to be sampled at any time by a WADA accredited laboratory*”. We further developed this idea in a recent Viewpoint article, inspired by a simple yet striking observation that the headquarters of the World Anti-Doping Agency (WADA) and the International Civil Aviation Organization (ICAO) sit just 30 meters apart in Montreal (5). Though they govern vastly different domains in sports and aviation, respectively, these two organizations succeed through international cooperation, cross-border enforcement, and standardized protocols. We proposed that ICAO could adopt a system similar to WADA, integrating infectious disease surveillance into international air travel. If designed and implemented with scientific rigor, equity, and transparency, such a system could serve as the foundational architecture for real-time, scalable global infectious disease surveillance. This proposal is both concrete and feasible, and represents a meaningful step toward a coordinated, adaptive, and enforceable global response infrastructure.

Concluding remarks

The nine papers in this Research Topic collectively demonstrate the global diversity, creativity, and commitment in advancing infectious disease surveillance and preparedness. From ferry ports in the UK to wastewater plants in Ethiopia, from dengue transmission modeling in India to influenza forecasting in China, these studies reinforce the critical need for both robust frontline data collection and sophisticated analytic capabilities. Together, they reaffirm the critical importance of interdisciplinary collaboration across epidemiology, data science, and technology. As the COVID-19 pandemic has shown, a real-time, transparent, and decentralized surveillance infrastructure is no longer aspirational but a necessity. We encourage the global public health community

to continue pushing the boundaries of innovation at this intersection of technology, data science, and epidemiology, ensuring that scientific insights translate into operational readiness. We hope this Research Topic serves as both a reflection and an inspiration to harmonize science, policy, and technology in the service of global health security.

While ideals can inspire, only tangible frameworks and enforceable standards can drive meaningful change. This distinction, between dreams and actionable solutions, lies at the heart of this Research Topic. Without structures that hold governments and institutions accountable, without interoperable systems that support timely data sharing, and without enforceable global agreements that transcend national interests, even the most visionary declarations risk becoming symbolic rather than substantive. We call on researchers, policymakers, and global institutions to move from rhetoric to rigor, from ambition to architecture.

Author contributions

JH: Conceptualization, Supervision, Writing – original draft, Writing – review & editing. ZD: Data curation, Project administration, Writing – review & editing. JL: Conceptualization, Supervision, Writing – original draft, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Li J, Hosegood I, Powell D, Tschärke B, Lawler J, Thomas KV, et al. A global aircraft-based wastewater genomic surveillance network for early warning of future pandemics. *Lancet Glob Health*. (2023) 11:e791–5. doi: 10.1016/S2214-109X(23)00129-8
2. Huang J, Döring O, Liu GG. The world needs a “pandemic” solution for a pandemic problem. *China CDC Wkly*. (2022) 4:1183–4. doi: 10.46234/ccdcw2022.238
3. Yan Y, Liu X, Lu S, Chang L, Dong J, Ji H, et al. Two-year surveillance of dengue, zika, and chikungunya viruses among Chinese blood donors—Guangxi and Yunnan PLADs, China. (2025) 2022–2023. *China CDC Weekly*. 7:357. doi: 10.46234/ccdcw2025.058
4. Winslow C-E. The untitled fields of public health. *Science*. (1920) 51:23–33. doi: 10.1126/science.51.1306.23
5. Huang J, Güllüoğlu M, Döring O, Wang H, Li J, Liu Y. Global infectious disease surveillance: bridge a. (2025) 30-metre gap between the International Civil Aviation Organization and the World Anti-Doping Agency. *J Glob Health*. 15:03010. doi: 10.7189/jogh.15.03010



OPEN ACCESS

EDITED BY

Jie Huang,
Southern University of Science and
Technology, China

REVIEWED BY

Bin Zhu,
Southern University of Science and
Technology, China
Leshan Xiu,
Shanghai Jiao Tong University, China
Xinxin Han,
Southern University of Science and
Technology, China

*CORRESPONDENCE

Marc Lipsitch
✉ mlipsitc@hsph.harvard.edu

RECEIVED 27 March 2024

ACCEPTED 18 June 2024

PUBLISHED 15 July 2024

CITATION

Lipsitch M, Bassett MT, Brownstein JS,
Elliott P, Eyre D, Grabowski MK, Hay JA,
Johansson MA, Kissler SM, Larremore DB,
Layden JE, Lessler J, Lynfield R,
MacCannell D, Madoff LC, Metcalf CJE,
Meyers LA, Ofori SK, Quinn C, Bento AI,
Reich NG, Riley S, Rosenfeld R, Samore MH,
Sampath R, Slayton RB, Swerdlow DL,
Truelove S, Varma JK and Grad YH (2024)
Infectious disease surveillance needs for the
United States: lessons from Covid-19.
Front. Public Health 12:1408193.
doi: 10.3389/fpubh.2024.1408193

COPYRIGHT

© 2024 Lipsitch, Bassett, Brownstein, Elliott,
Eyre, Grabowski, Hay, Johansson, Kissler,
Larremore, Layden, Lessler, Lynfield,
MacCannell, Madoff, Metcalf, Meyers, Ofori,
Quinn, Bento, Reich, Riley, Rosenfeld,
Samore, Sampath, Slayton, Swerdlow,
Truelove, Varma and Grad. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Infectious disease surveillance needs for the United States: lessons from Covid-19

Marc Lipsitch^{1,2*}, Mary T. Bassett³, John S. Brownstein⁴,
Paul Elliott⁵, David Eyre⁶, M. Kate Grabowski⁷, James A. Hay⁶,
Michael A. Johansson⁸, Stephen M. Kissler⁹,
Daniel B. Larremore^{9,10}, Jennifer E. Layden¹¹, Justin Lessler¹²,
Ruth Lynfield¹³, Duncan MacCannell¹⁴, Lawrence C. Madoff¹⁵,
C. Jessica E. Metcalf¹⁶, Lauren A. Meyers¹⁷, Sylvia K. Ofori²,
Celia Quinn¹⁸, Ana I. Bento¹⁹, Nicholas G. Reich²⁰, Steven Riley²¹,
Roni Rosenfeld²², Matthew H. Samore²³, Rangarajan Sampath²⁴,
Rachel B. Slayton²⁵, David L. Swerdlow², Shaun Truelove¹²,
Jay K. Varma²⁶ and Yonatan H. Grad²⁷

¹Center for Forecasting and Outbreak Analytics, US Centers for Disease Control and Prevention, Atlanta, GA, United States, ²Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, United States, ³François-Xavier Bagnoud Center for Health and Human Rights, Department of Social and Behavioral Sciences, Harvard T. H. Chan School of Public Health, Boston, MA, United States, ⁴Boston Children's Hospital, Harvard Medical School, Boston, MA, United States, ⁵Department of Epidemiology and Public Health Medicine, Imperial College London, London, United Kingdom, ⁶Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom, ⁷Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States, ⁸Division of Vector-Borne Diseases, US Centers for Disease Control and Prevention, Atlanta, GA, United States, ⁹Department of Computer Science, University of Colorado Boulder, Boulder, CO, United States, ¹⁰BioFrontiers Institute, University of Colorado Boulder, Boulder, CO, United States, ¹¹Office of Public Health Data, Surveillance, and Technology, US Centers for Disease Control and Prevention, Atlanta, GA, United States, ¹²Department of Epidemiology, UNC Gillings School of Public Health, Chapel Hill, NC, United States, ¹³Minnesota Department of Health, Minneapolis, MN, United States, ¹⁴US Centers for Disease Control and Prevention, Office of Advanced Molecular Detection, Atlanta, GA, United States, ¹⁵Massachusetts Department of Public Health, Boston, MA, United States, ¹⁶Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, United States, ¹⁷Department of Integrative Biology, University of Texas at Austin, Austin, TX, United States, ¹⁸Division of Disease Control, New York City Department of Health and Mental Hygiene, New York City, NY, United States, ¹⁹Department of Public and Ecosystem Health, College of Veterinary Medicine, Cornell University, Ithaca, NY, United States, ²⁰Departments of Biostatistics and Epidemiology, University of Massachusetts Amherst, Amherst, MA, United States, ²¹United Kingdom Health Security Agency, London, United Kingdom, ²²Departments of Computer Science and Computational Biology, Carnegie Mellon University, Pittsburgh, PA, United States, ²³Division of Epidemiology, Department of Medicine, University of Utah, Salt Lake City, UT, United States, ²⁴Siemens Healthcare Diagnostics, Inc., San Diego, CA, United States, ²⁵Division of Healthcare Quality Promotion, US Centers for Disease Control and Prevention, Atlanta, GA, United States, ²⁶SIGA Technologies, New York City, NY, United States, ²⁷Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, MA, United States

The COVID-19 pandemic has highlighted the need to upgrade systems for infectious disease surveillance and forecasting and modeling of the spread of infection, both of which inform evidence-based public health guidance and policies. Here, we discuss requirements for an effective surveillance system to support decision making during a pandemic, drawing on the lessons of COVID-19 in the U.S., while looking to jurisdictions in the U.S. and beyond to learn lessons about the value of specific data types. In this report, we define the range of decisions for which surveillance data are required, the data elements needed to inform these decisions and to calibrate inputs and outputs of transmission-dynamic models, and the types of data needed to inform decisions

by state, territorial, local, and tribal health authorities. We define actions needed to ensure that such data will be available and consider the contribution of such efforts to improving health equity.

KEYWORDS

pandemic, COVID-19, surveillance and forecast system, public health, infectious diseases, mathematical model

Introduction and purpose

To monitor pandemic pathogens effectively, modern surveillance systems should make use of the growing wealth of routine data from the health sector and from a range of other sources with new applications to disease surveillance such as mobility, internet searches, and wastewater. Public health experts must integrate these data in new ways that increase their value. We need purpose-built systems to detect new and evolving threats and to provide information as quickly as possible about those threats. What are the characteristics of the new pathogens or new variants of existing pathogens? What is their incidence and prevalence? What is the vulnerability of the population to infection and disease? What is the impact of our efforts to respond to these threats?

Systems to generate, integrate, and interpret these data should be designed and built with the explicit purpose of providing timely evidence to inform decisions about disease control and mitigation. First, they will provide direct input into decision making. For example, evidence of low vaccine effectiveness may prompt efforts to boost or change formulations or doses. As another example, real-time lab order data for diagnostic tests may prompt adjustments to resource allocation. Second, these data will parameterize scenario and forecasting models (1–3). For instance, estimates of per-case severity of a new variant, incorporated into forecasts or other models of case burden, may influence planning for hospital capacity and supply stockpiling and distribution.

This document reflects the framing ideas and the discussions held at a symposium organized by Harvard T.H. Chan School of Public Health entitled “Quantitative Tools and Data Opportunities for Pandemic Surveillance and Response,” held June 29–30, 2022, involving a range of public health and public officials, surveillance experts and other epidemiologists, and epidemic modelers. We first aim to identify the most important decisions for disease control and mitigation and the evidence that is needed to inform them. We then describe a set of surveillance activities designed to provide timely, reliable, and appropriately scaled data to inform these decisions. Our focus in this report is limited to domestic detection, characterization, and estimation of the burden of a pandemic pathogen in terms of direct health effects. Although we note the importance of monitoring economic, social, and indirect public health impacts of a disease control measures, we do not offer a comprehensive treatment of this element of pandemic monitoring and response activities. This paper builds on earlier efforts (4) while incorporating both the new possibilities that technology now provides, as well as the lessons of COVID-19.

We differentiate between two related but distinct goals of surveillance, as this document will focus on only one. The first goal is

to provide early warning about a potential pandemic, and so this type of surveillance includes global monitoring and rapid identification of domestic introductions. The second goal is to provide support for decision making during an ongoing pandemic, including tracking incidence, prevalence, and the pathogen's properties. While we will briefly remark on the former—surveillance for early warning—we will primarily focus on the latter—surveillance for decision making.

Detecting a jurisdiction's first cases of a new disease

The first set of decisions faced by a domestic public health jurisdiction, following the appearance of a pandemic threat somewhere else in the world, concerns the questions of whether, how, and to what extent to scale up a response to reduce the risk of importation or, if importation has happened, to control its spread within the jurisdiction. Measures to reduce importation via restriction or testing of inbound travelers may buy limited time to prepare (5, 6), though such measures lose relevance once local transmission is established (7).

To inform decisions about how to balance scarce public health resources between preventing importation vs. controlling local spread, it is critical to assess the risk that the infection has already arrived and started spreading within the jurisdiction. Testing and sequencing of specimens from international travelers at airports and analysis of wastewater from international flights may provide evidence of pathogen importation (8). Early evidence of local spread may come from informal communications among health care providers, reporting systems such as ProMED-mail (9), and “pre-health care” data (e.g., absenteeism, internet search queries). Signals may arise from monitoring of syndromes compatible with infections, or the volume, distribution, and results from clinical laboratory tests. Increasingly, wastewater monitoring can be the site of early detection of new threats or variants (10) and from routine programs or enhanced efforts at sequencing of clinical samples (11). They may also come from anomalous findings in sentinel and research efforts [e.g., the Seattle Flu Study at the start of COVID-19 (12)].

With each type of monitoring, there is first the question of what defines the signal we are looking for and then there is a tradeoff between having a highly sensitive and timely system capable of sounding an alarm early on one hand and producing too many false alarms on the other. In most situations, a high positive predictive value for such systems will be essential because the cost of responding to frequent false alarms is high. Much more work is needed to assess how to use and combine complementary monitoring signals to identify points at which an alarm should be escalated into a response.

Surveillance for decision making during a pandemic

A comprehensive list of the decisions and guidance required in a pandemic would fill a much longer document than this. Still, based on the combined experience of the emergency response phase of COVID-19 (2020–2023) and H1N1 influenza (2009–10) pandemics, we propose a set of consistent themes that capture many of the major types of decisions arising both in the early days of each pandemic and throughout.

A brief, necessarily incomplete, list of such decisions follows, adapted from the list in (4):

- 1 Public health goals of a response (elimination, control, protection of high-risk groups, protection of health care functioning, or a combination) and overall scale of response needed to meet these goals.
- 2 Timing of scale-up and scale-down of response.
- 3 Choice of nonpharmaceutical countermeasures (individual-targeted such as quarantine, isolation, and personal protection; population-targeted such as closures). This includes decisions about the timing, magnitude, and geographic range of protective measures that may be socioeconomically costly. A related set of decisions concerns how such measures should be prioritized, i.e., who should receive protective equipment when it is scarce, and how closures should be targeted to reduce economic and social disruption.
- 4 Choice of medical countermeasures, including diagnostics, therapeutics, and vaccines. This includes decisions about development, stockpiling, procurement, expanding capacity (e.g., building alternative care sites), and more. Here too, questions of allocation and prioritization are central. This also includes planning for potential surges.
- 5 Specific policies for each of the issues above in special populations including vulnerable communities, and settings such as health care, schools, congregate settings, transport, etc. (See *Centering Equity* below)
- 6 Balance between community countermeasures to reduce severe disease or reduce transmission (e.g., allocation of resources to those at high risk of complications or high risk of transmission).
- 7 Design and implementation of staged alert systems to provide real-time risk awareness and trigger policy changes (13, 14).
- 8 Imposition and removal of international travel screening and restrictions.
- 9 Choice of public health communication strategies.

Each of these decisions requires specific data to decide how to improve health equitably, effectively, and efficiently while minimizing social and economic disruption. For example, decisions on testing, isolation, and quarantine policies require evidence on the natural history of infectiousness (or at least a proxy such as viral load), test sensitivity at different levels of viral shedding, the relationship between symptoms and infectiousness, and the potential economic and social consequences for various communities of the policies under consideration. In contrast, decisions about the timing of vaccine boosters require evidence on the effectiveness of existing vaccines against infection, transmission, and severe disease endpoints, stratified by such factors as pathogen variant, time since vaccination, and age,

as well as understanding of how vaccine protection is distributed across demographic groups.

Decisions faced by state, territorial, local, and tribal authorities

In the federal system in the United States, public health is decentralized and typically not coordinated among states. State, territorial, local, and tribal (STLT) governments are responsible for nearly all binding policy decisions in public health, with governance health structures varying by state (15). The purview of these bodies includes (16) prescribing and enforcing isolation, quarantine, mask mandates, and restrictions on businesses and gatherings; vaccine prioritization and distribution; and (to a degree) diagnostic testing. They also hold responsibility for closely related areas, such as public education. STLT governments all have a desire for similar types of data, but vary in how much they need, how quickly they need it, and how they use it.

Many decisions involve procurement and distribution of countermeasures. Because STLT authorities are making allocation decisions within their jurisdictions (e.g., for counties, cities, hospitals, schools), jurisdiction-wide measures of disease activity are rarely sufficient; instead, more geographically granular numbers are required (Table 1).

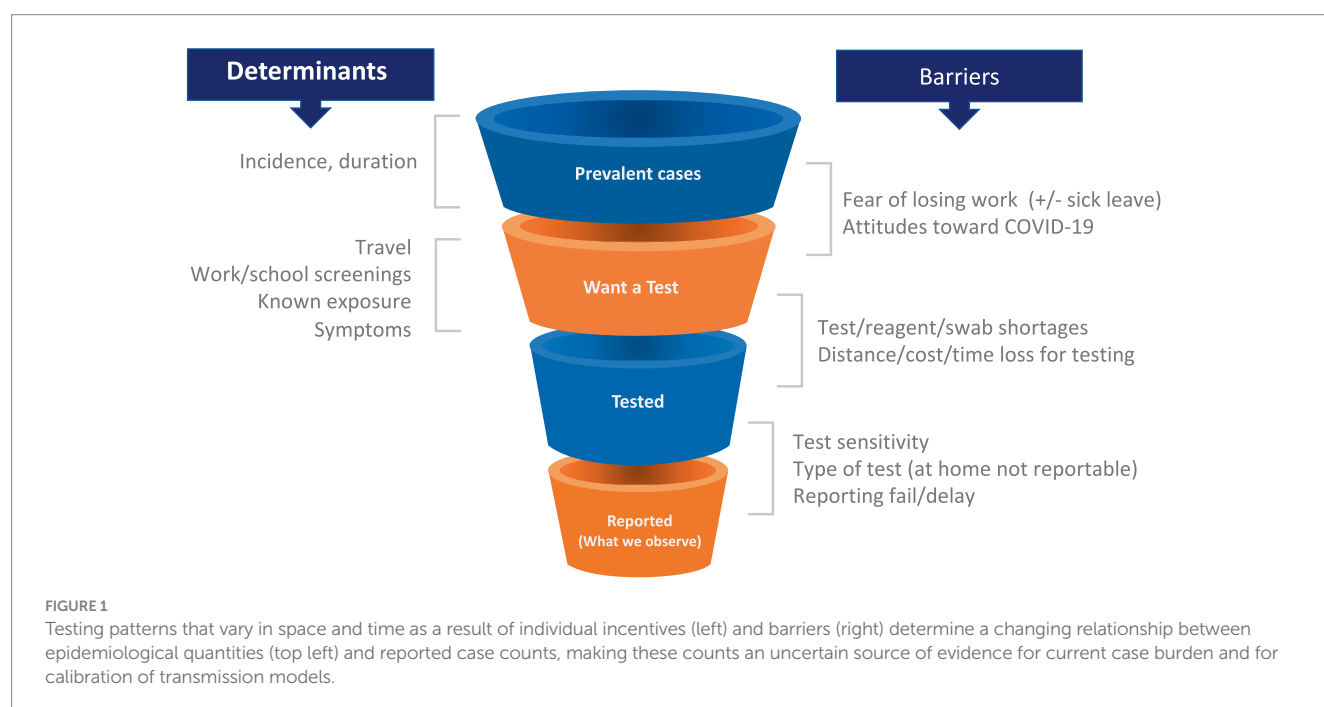
Data needs for decision support: the COVID-19 experience

A range of data sources could and, during the COVID-19 pandemic, did provide evidence to support decisions by health authorities. Following initial social media reports of clusters of pneumonia, some of the earliest specific data to characterize the COVID-19 threat came from traditional sources, such as from case reports posted on Chinese public health department websites (17). A key challenge was the repeated change in the syndromic case definition in the early days (18). But other early data came from unexpected sources, such as cruise ships (19), restaurants (20), and fishing vessels (21), where conditions allowed inference of the path of transmission and thereby provided evidence about the degree and mechanisms of spread. Specifically, these provided some of the earliest evidence of asymptomatic and aerosol spread, which, when properly interpreted, aided in the design and prioritization of testing and other control measures. As had been true in the 2009 influenza pandemic (22, 23), sampling of travelers provided early estimates of the extent of global spread, growth rates, and likely under-detection (24, 25).

As the pandemic spread, the strengths and limitations of each data source became evident. Multiple data types were required to provide even an incomplete picture of trends in incidence and prevalence and behavior (26, 27). For example, case counts were used as an important indicator of disease burden. However, the relationship between new cases and true incidence varies as a function of numerous factors, including test availability, test reporting requirements by jurisdiction (which did not always include reports of negative tests), rates of testing through clinical facilities (which declined with the growth of rapid antigen testing), and incentives to get or avoid testing (Figure 1). Some of these limitations can be mitigated by breaking out case counts by

TABLE 1 summarizes key decisions and associated needs for jurisdiction-level data and analytics in COVID-19 cited by state and local leaders during the symposium.

Decision	Data/analytics need
Size of response needed	Rapid threat characterization
Choice of community countermeasures	County-level disease burden and transmission measures
How to ensure adequate supply of hospital beds, ventilators, personal protective equipment	Forecasts of demand for these items
School and congregate setting policies (closure of schools, infection control measures in jails, prisons, nursing homes, etc.)	Understanding of rates of transmission into, within, and from each of these settings and impact of testing and infection control on these rates as well as population specific health-risks
Countermeasure deployment within a jurisdiction	Age, racial, ethnic, and geographic patterns of transmission and disease burden. Note: these are often crude proxies for social determinants of infection and outcome risk, not adequate for scientific understanding of why particular groups are at risk, but nonetheless potentially useful for focusing prevention and treatment efforts on those with high vulnerability.
Efforts to distribute and promote vaccination	Variant prevalence, vaccine coverage, and vaccine effectiveness against dominant and emerging variants



the reason why an individual was tested (symptoms, travel, surveillance), but this was not consistently done in the US. As a result of these limitations, hospitalizations and even deaths were increasingly used as the more reliable indicators of case numbers, sacrificing some timeliness for a more consistent relationship to the underlying incidence of infection. Random sampling approaches (described below) can overcome these limitations and provide more consistent and reliable estimates of incidence and prevalence and how these change over time. Only the United Kingdom and Luxembourg used random sampling on a large scale, perhaps because of the cost and logistical challenges. Notwithstanding their limitations, case counts were the major early data source in the United States and provided critical evidence especially when linked to demographic information. Syndromic surveillance—done routinely as part of monitoring

influenza trends—from emergency room visits and hospital admissions were also valuable data sources, particularly when testing was limited. However, interpretation of syndromic surveillance was complicated by changes in healthcare seeking behavior and the increased use of telemedicine.

Novel data streams provided confirmatory evidence as well as early warnings of trends that might not be evident in case counts. For example, wastewater surveillance for SARS-CoV-2 was adopted in numerous jurisdictions from 2020 to 2022 and provided evidence on local epidemic trends, although the precise relationship between wastewater abundance and the number of infected persons depends on the wastewater sampling scheme and on shedding patterns (among other issues), and thus difficult to quantify (28). Moreover, by its nature, wastewater cannot indicate who has been infected, thus

leaving the demographic profile of infected persons uncertain. Finally, wastewater surveillance as currently applied will miss infections in areas with high reliance on septic systems, which serve roughly one fifth of the households in the U.S. with heavy concentrations in certain geographies, inducing inequities in whose infections are tracked by this approach (29).

Another novel data stream is the use of population-wide distributions of viral load measured from PCR testing, which, in the aggregate, can provide information on the trajectory of the epidemic, even from a single cross-sectional analysis (30). This approach has reached proof of concept and has the advantage that it may be less sensitive to trends in testing behaviors than measures of incidence based on case counts, and unlike wastewater surveillance, it can provide some information on the demography and precise location of cases. However, further work is needed to see how a transition to non-PCR testing for many new cases, the halting of pre-procedural and asymptomatic testing, and the shifts in viral kinetics that come with immunity (from vaccination and infection) affects the nature of this signal (31–33). Moreover, the identifiability of time-since-infection from viral load, which is needed for the approach to work, depends on the asymmetry in viral load over time [fast rise, slower decline (30)], which may or may not be a feature of future infectious diseases.

Digital data can also be used for surveillance and to inform on epidemic trajectory. ProMED-mail (9) and HealthMap (34) are valuable for flagging and disseminating reporting and information on events known or suspected to be infections and outbreaks. Data from search engines, social media, and news reports data can also inform epidemic dynamics and for forecasting (35–37).

Finally, testing for antibodies in sera collected either for the purpose of serologic surveillance or in convenience samples (e.g., blood banks, discarded clinical samples) was used to characterize both the landscape of population immunity (i.e., who was and wasn't vulnerable to reinfection) and to distinguish between those who had acquired immunity via vaccination vs. infection (38). Secondary analyses of COVID-19 vaccine studies identified complexities in answering the latter question, finding that infection does not reliably induce antibodies to non-vaccine antigens in vaccinated individuals (39).

An important conclusion is that no one data source or surveillance tactic is sufficient. In a setting like the U.S., multiple surveillance approaches are needed at scale. Beyond the obvious need to combine data sources, several points stand out.

The first is the value of data completeness and of linking data types to produce evidence that is greater than the sum of the parts. For instance, while counts of cases and hospitalizations are valuable, missing race/ethnicity, geographic, and other patient characteristics have impeded efforts to improve services to groups that are underserved or experience high disease burden and to improve equity in health-related outcomes. Similarly, meticulously linking sequence data from patient isolates with demographic and clinical predictors of severe outcomes, including vaccination history, and clinical outcomes can help to evaluate the threat posed by novel variants (40). Unfortunately, despite prodigious amounts of SARS-CoV-2 sequencing in the U.S., this form of linkage has been relatively rare to date.

Second is the value of clear and accessible data dashboards with transparent data sources to make the state of the epidemic locally

evident to the public. The same data should also be available to analysts in public health departments, academia, and other sectors via application programming interfaces (APIs) to facilitate rapid data analysis. This can facilitate shared decision making and help to increase public support for control measures. For example, the city of Austin, Texas developed a COVID-19 staged alert system that guided local policy between May 2020 and March 2022 (13, 14). The public-facing dashboard featured a single graph that tracked COVID-19 hospital admissions and clearly indicated thresholds between the red, orange, yellow, green, and blue risk levels that were linked to specific actions. The county judge, city mayor, and public health authority cited the dashboard almost daily to communicate risks, explain changes in policy, and cultivate adherence via news outlets and social media. This system was only possible because local authorities required area hospitals to report daily admissions beginning in April 2020, long before such data were generally available.

Finally, discussion at the Symposium emphasized the value of metrics that could be compared across jurisdictions. Decision-makers expressed a desire for objective criteria by which their performance can be judged. Comparisons across states, for example, were hampered by differential testing rates that affected case counts in ways not reflecting actual prevalence. A CDC-supported academic effort called *covidestim* (41) used Bayesian evidence synthesis to harmonize estimates of current and cumulative infections across states and counties, providing an example of what could be done by health authorities. However, this effort was also hampered by unanticipated changes in reporting tempo, as well as 'data dumps' and data backfilling. Different definitions of COVID-19 hospitalization across states and over time impeded comparisons of outcomes that would have provided indications to elected leaders of the quality of their responses and informed improved responses.

Surveillance inputs to forecasts, scenario projections, and analytic models

As noted above, many aspects of pandemic decision-making can directly incorporate evidence from surveillance, and will also make use of

- Nowcasts: estimates of current burden of cases, hospitalizations, deaths, and other quantities that account for delays in reporting (42);
- Forecasts: relatively short-term projections using time-series and other statistical modeling techniques, sometimes supplemented with transmission-dynamic approaches to estimate future case, morbidity, and mortality burden, typically on the scale of days to weeks (1, 3);
- Scenario models: longer-term estimates of pandemic dynamics using transmission-dynamic modeling approaches to anticipate multiple possible futures under stated assumptions about behavior, viral evolution, vaccine durability, etc., typically on the scale of months to years (43);
- Results from analytic models: estimates about different characteristics of the pathogen or a population of concern that are specifically designed to inform a decision or guidance, such

as school- or nursing-home-based testing policies (44), border restrictions or contact tracing procedures (45), or quarantine approaches (46).

These categories of decision-support tools require estimates of *input quantities* that represent the assumptions of the models—for example, in scenario models, estimates of per-case severity or vaccine effectiveness. These estimates need to be timely, representative, and specific to the pathogen or variants circulating or anticipated to circulate (for example, due to importation). In addition, scenario and forecast model *output* must be calibrated to existing measurements of disease burden: incidence of infection, diagnosed cases, hospitalizations, deaths, and other relevant metrics, as well as against cumulative measures such as seroprevalence. The categories of input and output are somewhat fluid, as a model with sufficient data to calibrate outputs may be able to estimate the values of some of the quantities described here as inputs. In a fully Bayesian framework, both external estimates (as priors) and calibration to output data may contribute to posterior parameter estimates.

For forecasts, evaluation can be performed quickly due to the short-term horizon of the predictions made, with results that can provide feedback to modelers about places where models are mis-specified. Evaluating scenario projections is more complicated, as multiple sets of counterfactual projections are made under different assumptions about how a pandemic situation will evolve over the course of months or years (47). Most (or perhaps all) of the scenarios will not be realized exactly as assumed, making evaluation less straightforward.

Together the quality and timeliness of these input parameters and output calibrations are important determinants of how useful a model is for decision making. While there are techniques to adjust for incomplete or lagged information, the absence of certain ingredients—especially model output calibration targets such as numbers of cases or hospitalizations—can critically compromise the ability to generate models that reflect reality to the point of hampering basic situational awareness. Data systems that support modeling and in turn decision-making during pandemics should be considered vital national security capabilities and prioritized accordingly.

A list of the key needs for model inputs is as follows, many of which may change as a pathogen evolves (referred to below by their letters):

- a Pathogen kinetics/epidemiological parameters (e.g., incubation period, latent period, infectious period, infection fatality ratio). Estimation of these inputs may itself require simple models, particularly at the early stages of a pandemic [ref: Gostic paper].
 - b Transmissibility and efficiency of various transmission mechanisms
 - c Risk factors for infection and severity
 - d Individual and population immunity (including effects of infection, vaccination, and waning)
 - e Diagnostic test characteristics, including specificity and sensitivity for active (acute) and past infection
 - f Vaccine effectiveness and waning of effectiveness, for infection, severe disease, and mortality endpoints.
 - g Treatment effectiveness
 - h Policies, uptake, and effectiveness of nonpharmaceutical interventions
 - i Population mobility and interactions: contact networks and patterns by setting
 - j Importation risk
 - k Other co-circulating pathogens of concern (e.g., if concurrent with significant influenza transmission)
 - l Capacity and utilization of healthcare resources (including hospital beds, therapeutics, and vaccines)
- Key additional data requirements for fitting models— as well as for general situational awareness – include:
- m Geographically and demographically stratified incidence, duration and prevalence of infection, hospitalization, ICU admission, death, and other relevant metrics associated with the pathogen, ideally by variant.
 - n Strain-specific incidence.

Meeting these needs

This draft framework is a preliminary attempt to scope a system that could meet the needs listed above for situational awareness, decision support, and inputs and outputs for modeling and analytics for a new variant or a new pandemic. Capacity to achieve these would also be applicable to other pathogens, especially, but not only, respiratory ones.

A. Estimating model inputs

System 1: high-frequency sampling for pathogen kinetics and diagnostic sensitivity (quantities a,e)

Possible Mechanism: Surveillance would be established to obtain repeated samples (for COVID-19, respiratory samples) from individuals exposed to a pathogen of interest (now, SARS-CoV-2) from the time of exposure through infection to the time of clearance. High-frequency sampling will provide detailed profiles of pathogen kinetics, which could be subgrouped by prior infection history, vaccination status, pathogen variant, demographics, and other predictors. Simultaneous use of nucleic acid amplification (NAAT), culture, and antigen-based testing on these specimens would provide detailed estimates of the sensitivity of each as a function of symptoms, pathogen load, pathogen infectious capacity, variant, and time since exposure/first-positive to inform choice of diagnostics and isolation/test/quarantine policy.

Performers might be research/surveillance networks or STLT health departments (recognizing that the health departments may have limited bandwidth in the context of an outbreak). The ability to scale up is critical. While pathogen kinetics are not likely to vary from place to place, geographic diversity in sites capable of performing these investigations will increase the timeliness of results in case one region is hit much earlier than others.

Settings may include households, universities, day cares and schools; intensely monitored cohorts such as sports leagues or health

care workers, congregate settings such as homeless shelters, correctional and detention facilities, or nursing homes.

Precedents: United Kingdom Household study (48) and United States National Basketball Association studies (32, 49).

System 2: integrating routine sequencing with detailed clinical data (quantities b,c,d,f,g,n)

Possible Mechanism: A payer-provider network with diverse geographic and demographic representation (alternatively, a private sector entity or consortium of public health departments and laboratories capable of merging clinical data with sequence data) would track individuals as a cohort (not necessarily defined by long-term follow-up but perhaps with exposure or a positive test as an entry criterion) with known vaccine and prior infection history through diagnosis (outpatient or inpatient) and through the cascade of care to estimate the probability and severity of infection as a function of this history (vaccine effectiveness and infection-acquired immunity) and variant. Sequencing of positive clinical specimens would enable the variant-specific estimates. This system would provide a reliable infrastructure for assessing severity, vaccine effectiveness, and treatment effectiveness linked to infection and vaccination history for each new variant/virus. It would be crucial to link electronic health record (EHR) within the network to key external sources of data such as immunization registries. Improving completeness of such registries is also a high priority to improve the quality of these inferences. Strategies for linking pathogen genome sequencing with EHRs will depend on whether these data are from clinically validated systems and, if not, will require consideration to ensure use for research and not clinical purposes.

It would be valuable to explore to what extent such studies could be done in networks such as PCORNet (50) or the Vaccine Safety Datalink (51) that assemble EHRs from multiple health systems into a common data model; questions include how rapidly this could be done and whether sequence data could be linked to these records.

In addition to payer-providers, robust testing, reporting, and data collection capabilities should be considered for congregate settings at high-risk for transmission such as skilled nursing facilities, correctional facilities, detention facilities, and homeless shelters that can follow individuals from positive tests through outcomes.

Precedents: Cohort studies on variant-specific relative severity (52), relative vaccine effectiveness (53) and absolute vaccine effectiveness (54) have been performed during the COVID-19 pandemic. None of these included genomic sequencing or serological profiling integrated with clinical data collection, in part due to the issues of linking with EHRs as mentioned above. Integration of sequencing in particular is essential for the likely future scenario where one cannot rely on proxies for genetic variant that have been exceptionally convenient in COVID-19, notably the failure of the S-gene PCR target in certain polymerase chain reaction-based diagnostic tests.

In the US, this work could build upon or integrate with existing platforms such as VISION and Investigating Respiratory Viruses in the Acutely Ill (IVY) (55). Key additions would be sequencing and more comprehensive estimates of severity.

System 3: behavioral surveillance and other routine data collection (quantities h,i,j,l)

Goals of behavioral surveillance are to provide real-time estimates of mobility, work-from-home frequency, proportion of schools open

or closed, use of other nonpharmaceutical interventions such as masks, and vaccine behavior/hesitancy. Data useful during COVID-19 included vaccine coverage from HHS Protect (56) and Census Pulse (57) and other surveys on vaccine intentions, mask use, work-from-home, and school opening/closure. Private sector [e.g., mobility (58)] and publicly available data [e.g. (59)], exist that measure many quantities of interest. These include self-reported mask use, absenteeism data from school and work, internet search queries, and much more. Further work needs to be done in several areas to enhance the value of these data streams:

- identify cost-effective sources of such data;
- quantify the degrees of representativeness in measurement from these different data sources by such factors geography, race/ethnicity, and social determinants of health;
- improve our mechanistic understanding of how these measures of mobility relate to transmission behavior, which will likely differ by social factors, pathogen transmission routes, and epidemic stage, among other factors (27)

A particular example of one such data stream is air travel and other travel data to estimate importation risk.

Precedent: Census Pulse and other surveys exist. Many local jurisdictions have used mobility data from private providers, often via academic intermediaries¹ to assess local trends. Vaccine coverage data exist with some limitations. The Center for Disease Control and Prevention's (CDC) Division of Global Migration and Quarantine maintains access to timely estimates of air travel volume.

B. Fitting model outputs

System 4: repeated testing for infection and immunity in a random sample of the population (m)

Mechanism: An academic, government (e.g., CDC or a coalition of state health departments), or private sector entity would identify a longitudinal sample and/or repeated cross-sections representative of the U.S. population for monthly testing for infection and immunity as evidence of prior infection. In the COVID-19 case, this would be PCR testing of respiratory samples and antibody measurement in blood; testing approaches might differ for future pathogens. Samples would be obtained by home visit or mail/courier. Specimens testing positive for one or more respiratory viruses would be sequenced. The initial sample would be powered to detect US-level trends; scale-up in a pandemic would enable regional/state-level and demographic-specific (e.g., age, race, sex-specific) estimates of virus prevalence and seroprevalence irrespective of symptoms and at the level of variant/subtype/species/type (depending on the pathogen).

In pathogens with antibody-based immunity, blood samples would be tested for multiple antibodies including vaccine and nonvaccine antigens of the novel pathogen. These would provide a population-based denominator for severity estimates, enable calibration of scenario and forecast models, track trends in viral

¹ e.g., <https://www.covid19mobility.org/>

species/variants in an unbiased way, and estimate the magnitudes of health inequities to better prioritize prevention measures (60).

Addition of serologic testing of a random, representative sample of the population would supplement existing passive serosurveillance approaches such as from blood donors (38), newborn heel sticks (61), or discarded specimens (62, 63). Longitudinal sampling would enable more precise estimates of rates of waning of antibody concentrations (64, 65) and the consequences for estimation of cumulative incidence using particular assays.

An alternative approach would be to use healthcare-based testing of individuals requiring admission for conditions not directly related to the pandemic, using weighting to standardize the population seeking health care to the background population (66, 67), though the quality of such data would need continuing validation.

Other alternatives would include the use of voluntary testing results, such as those gathered by test-proctoring telehealth services, retail pharmacies, or the like. CDC/FDA requirements to ask the reason for a test would facilitate interpretation (symptomatic vs. travel vs. exposure, for example).

Precedents: The main proposal could be roughly modeled on the United Kingdom COVID-19 Infection Survey and REACT-1 studies. One of the alternative approaches—universal testing of individuals requiring admission for non-pandemic reasons—was used in New York City early in the COVID-19 pandemic (68) and has been used in Indiana with reported high value (66, 67) for both prevalence and seroprevalence.

System 5: maintain hospitalization surveillance data (l,m)

Hospitals have been required to report COVID-19 and influenza hospitalizations to HHS, and these formed the backbone of multiple forecasting and scenario modeling efforts in the US. It is critical to maintain the generation, interpretation, timeliness, and accuracy of these data to inform forecasts. In addition to the forecasting products, these data underlie hospital capacity and burden situational awareness, the ability to monitor outbreaks, and community burden indicators.

Precedents: Exists as of September 2023 but needs to be maintained at a base level outside of emergencies and be able to ramp up quickly at a time of new emergency (69).

C. Actions needed

Administrative and reporting preparedness

The response to COVID-19 required collaborations across sectors—public, private, and academic—but these collaborations were often forced to work through administrative frameworks that were not designed with speed and flexibility in mind. In turn, such mis-specified frameworks ultimately slowed or limited some critical public health projects and prevented others from being undertaken entirely. To address this class of problem, we propose six ideas below that would update, recast, or create key frameworks that establish links across sectors and that facilitate the urgent work of pandemics, while maintaining safeguards and oversight.

- 1 *Emergency data use agreements and formats.* Data use agreements (DUAs) are core elements to collaborative work across institutions, but they pose two types of challenges. First,

the process for negotiating an agreement acceptable to the institutions providing and receiving the data is often slow. The staff on each side tasked with reviewing and signing off on these agreements may have many competing priorities or be overwhelmed as an outbreak or pandemic may dramatically increase the volume of DUAs. Work on a sensitive or high-profile project, such as associated with an outbreak of infectious disease or a pandemic, generates additional scrutiny and often further lengthens the review process. Second, conflicting limitations can stall progress or even undermine a project before it starts. For example, in a partnership between academics and government public health institutions, academic institutions may deem the freedom to publish without interference to be non-negotiable. Public health institutions, however, may require veto power over what, if anything, is published, due to the sensitivity of the institution's data and ownership thereof. To address these problems, one solution is to establish Emergency Use Data Authorizations (EUDAs) for public health data with a standing framework vetted and updated regularly (e.g., annually), perhaps at the individual state level. Such EUDAs would catalyze collaborations and enable investigators at both institutions to shift the balance of effort up front from administrative to research tasks. As these are put in place, discussions about data formats can take place, ideally also in advance, to ensure that when data are delivered they are as ready-to-use as possible.

- 2 *Surveillance versus research: updating the Common Rule.*

Projects designated as human subjects research require institutional review board (IRB) review, whereas those designated as public health surveillance are deemed not to be research, and thus do not require IRB review. This surveillance-research dichotomy has substantial implications for timeliness and speed of work, because writing, reviewing, and adjudicating IRB reviews—while vitally important for protecting the rights, welfare, and well-being of human subjects—may take days to weeks. The boundaries between surveillance and research are governed by the Common Rule, which states that public health surveillance activities “include those associated with providing timely situational awareness and priority setting during the course of an event or crisis that threatens public health (including natural or man-made disasters)” [45 CFR 46.102(I)(2)] (70). Unfortunately, these boundaries lacked clarity and standardization as questions arose during the COVID-19 pandemic. For example, while case monitoring is clearly surveillance and a routine public health activity, one could make a strong argument that “situational awareness and priority setting” includes assessing vaccine effectiveness and disease severity for new variants. However, analysis of variants requires pathogen genome sequencing, which is viewed by some as constituting research, as is evaluation of vaccine effectiveness, another critical public health function which is not exclusively a research objective. Modifying the text of the Common Rule to explicitly include examples such as these or providing an interpretation of the surveillance/situational awareness exemption that includes these activities would considerably improve the ability for public health agencies to maintain situational awareness and set priorities, quite in line with the spirit of the exemption.

- 3 *Streamlined IRBs*. Where projects fall under human subjects research designation and require IRB review, generic, pathogen-agnostic study protocols for specific populations would accelerate research by decreasing the time to first data. Preapproval of a range of well-defined studies targeted at emergency response and using specific data sets would retain critical protections for human subjects, while allowing high-urgency protocols to be “on the shelf” and ready for fast rollout. As an additional feature, such preapproved protocols would also free up valuable researcher and IRB reviewer time, having converted per-submission efforts during a pandemic into fixed-cost efforts ahead of time. Moreover, designing consent processes for normal “peacetime” studies to allow use of data and specimens in public health emergencies could avoid some of the delays experienced during COVID-19 with, for example, use of the Seattle Flu Study’s specimens to understand early transmission of the virus in the United States (12).
- 4 *Case reporting standardization*. Tracking and understanding outbreaks, particularly at their beginnings, rely on case reporting. Ideally, public health efforts would follow case trends over time and across regions, compare and monitor clinical features including disease progression, resolution, and response to interventions, and track demographics of infected individuals. But lack of standardization of case report protocols, parallel or overlapping surveillance systems that result in duplication (often with varyingly completed fields for the same case), and inadequate systems for incorporating updates as further information about a case accumulates after the initial report, among other issues, result in case report data that require much time and effort to sort through. Worse, these issues may render some fraction of case reports unreliable. Improving national surveillance systems to be more uniform, timely, and flexible could serve both local and national surveillance needs would help address these issues (60).
- 5 *Dataset accessibility*. In the absence of a United States national healthcare system, research into the distribution and burden of clinical conditions depends on academic or private data streams, including surveys and surveillance systems constructed to address specific questions, and databases of insurance claims which represent utilization of the healthcare system. Insurance claims datasets include those from (i) employer-based insurance companies (e.g., MarketScan) (ii) all-payer claims databases available in some states (which, since a 2016 Supreme Court decision (71), are no longer necessarily ‘all payers’), (iii) Medicare for individuals over 65 years of age, (iv) Medicaid, which provides coverage to over 18% of the United States; and (v) data bases for other specific populations, such as those of the Veterans Affairs Health System, the Indian Health Service, and the Department of Defense. While these datasets can provide an important window into healthcare use across demographics and geography, access to these datasets can be expensive and time and labor intensive. Gaining access to Medicaid data, for example, presents a substantial burden, since this has to be acquired on a state-by-state basis. Establishing standing flexible DUAs for these datasets, with a single agreement across states for Medicaid and other state-controlled data, could enable both routine surveillance-type analysis to identify trends (such as disease outbreaks or

patterns of disease spread) and to evaluate the impact of clinical and public health interventions.

- 6 *Public health-health care partnerships*: While the United States does not have a national health system for all, it has a wealth of data in the health care sector that can inform public health decision making. Multiple studies at the Centers for Disease Control and Prevention (CDC) and other institutions harnessed such data to provide estimates of key quantities such as vaccine effectiveness (72, 73), variant severity (74), and antiviral effectiveness (75), as well as for surveillance of disease burden and its correlates (76). Building public health partnerships with the health care sector in advance to set in place the administrative, information technology, and financial arrangements to make possible high-quality analyses of this sort rapidly (and automated where possible) would greatly increase the timeliness and value of such efforts (77).

Strengthening personnel and research ties, including globally

In response to the COVID-19 emergency and the need for expertise to gather, analyze, and interpret evidence around the pandemic and the clinical and public health responses, many academics put aside their usual research programs to engage directly in public health activities and research. The close interactions between academics and local, state, and national public health officials were often productive and important for guiding the pandemic response but raised issues that should be addressed before the next pandemic. These include the *ad hoc* way in which these academic-public health collaborations came into being, the lack of uniformity of access to academics with appropriate expertise across states, and the misalignment of incentives between public health and academic work.

Ideally, academia-public health collaborations can be rapidly scaled up in times of need through established pathways. One idea is to create a “rotator” program, in which academics (and potentially those in training, including doctoral students and postdoctoral fellows) are embedded within public health agencies—and similarly public health officials are embedded within academic groups—for intervals (such as 3 or 6 months) that build familiarity, collegiality, and accessibility. The LEAP fellowship through the Infectious Disease Society of America (78) and the joint Infectious Diseases/EIS fellowship (79) programs are efforts in this direction. Another approach is to establish an academic career path in which some fraction of time and effort are based in public health activities, analogous to academic medicine paths in which researchers spend some fraction of their time doing clinical work. Cooperative agreements established in 2023 between the CDC’s Center for Forecasting and Outbreak Analytics (CFA) and academic and other groups include a surge provision whereby the performers on these agreements would provide scientific assistance in times of crisis. Relatedly, an official “public health reserve corps” of analysts and modelers could provide a workforce available to be called up to prepare for and respond to emergencies. Formal recognition of these paths as prestigious and vital, and placing value on these activities within the academic systems of rewards and incentives, will be key to success (Box 1).

Often tools developed for one public health jurisdiction solve common problems and could be readily implemented in other jurisdictions, underscoring the importance of making code open access and ideally making tools generalizable. This would have the benefits of “not reinventing the wheel” and allowing those jurisdictions lacking local expertise access to useful tools. A curated clearinghouse of such tools, organized by research problem and perhaps hosted by CFA offers one strategy for providing access. Broad efforts to create and maintain state-of-the-art tools for epidemiologic modeling, such as Epiverse (82) and Recon are encouraging developments in this space (83).

We have focused on domestic systems in the United States, but that international cooperation is essential for multiple reasons. Maintaining systems to detect new pathogens at multiple locations in a globalized world will speed detection, facilitating timely responses including development of countermeasures (25) and providing early warning to other jurisdictions (84). For pathogen characterization, as we note in Box 2, each jurisdiction can benefit from findings in other jurisdictions on quantities that are relatively similar across populations, including for example the effectiveness of vaccines and treatments, as well as certain features of infection natural history (e.g., viral load kinetics). Strengthening these capacities globally means that locations that have exceptional data and study infrastructure and/or early experience with a pathogen or variant, can contribute to the global store of knowledge of pathogen characteristics (49, 54, 85–87).

Educating the consumers

Tools are best deployed by those who understand how they work, how they are limited, and how they can be modified to improve their applicability to specific situations. Workforce development mandates to build subject matter expertise within public health departments, such as through CDC efforts via the Office of Science and the Office of Advanced Molecular Detection and through fellowships such as the CDC/Association of Public Health Laboratories (APHL) program, are critical efforts. While waiting for these workforce programs to get up and running, and since public health officials may not stay current with the frontiers of analytical and modeling methods, opportunities for regular formal trainings should be developed. For example, meetings such as the Council of State and Territorial Epidemiologists Annual Conference could provide a forum for workshops on advances in modeling, genomic epidemiology, and other fields. Relatedly, encouraging public health officials to attend field-specific meetings (e.g., Epidemics, Applied Bioinformatics in Public Health Microbiology) could provide opportunities for knowledge sharing, relationship building, and networking across sectors and disciplines.

Improving knowledge flow

Successful communication of a health agency’s current understanding of a pandemic and outlook for its future requires a combination of approaches to communicate different kinds of data and outlooks, for different audiences. It has been suggested that principles for such communications include: thematic structure related to informing key decisions, synthesis of evidence from multiple sources, quantification of uncertainty, inclusion of visualizations as well as text and tables, and inclusion of forward-looking material

(outlooks for the future) (91); another important principle is open access to the data underlying figures in these reports. CDC’s Technical Reports on the Mpox epidemic in 2022 (92) sought to put these principles into practice, explicitly emulating aspects of the United Kingdom Health Security Agency (UKHSA) Technical Briefings from COVID-19 (93). Creating a regular cadence for such reports during an emergency, as was the case in the United Kingdom during the height of the COVID-19 pandemic, may help develop an audience and facilitate knowledge flow.

Centering equity

The World Health Organization has stated that “Countries have an obligation to develop appropriate, feasible, sustainable public health surveillance systems” to ensure that the health needs of populations are quantified so that they can be addressed. While there has been a disproportionate impact of COVID-19 on racial and ethnic minorities and on socioeconomically disadvantaged populations in the United States (94) and elsewhere (95, 96), a persistent problem is that race/ethnicity data are too often missing from surveillance data. Under the plausible hypothesis that those with missing data on race/ethnicity are among the most disadvantaged, these missing data could lead to attenuated estimates of the degree of inequities; whether or not this is the case, it reduces the quality of the estimates by adding uncertainty. Improving the completeness of race-ethnicity reporting is an urgent priority to maximize the value of surveillance data to enhance health equity. Some symposium participants, while agreeing with the need for better reporting of such data, argued that in the presence of ongoing racial segregation, ZIP code or other geographic tags can be a useful proxy when such data are unavailable. Early maps of COVID-19 in New York City showed a higher prevalence of COVID-19 diagnoses in areas that were home to largely Black and Hispanic populations, as well as some areas where most residents were White and many believed to be first-responders (68). This was reflected in elevated COVID-19 mortality rates among Black, Hispanic, and Native American populations compared to White populations throughout the United States, particularly in the early waves of the pandemic (97).

The Presidential COVID-19 Health Equity Task Force final report from 2021 (98) recommends strategies for enhancing equity in data, analytics, and research. These recommendations include standardizing demographic and socioeconomic categories, supporting equity-centered data collection, tracking and reporting on health outcomes for people in congregate and high-risk settings, and research and analysis on behavioral health. In a similar spirit, for any clinical or public health intervention, one should ask in what ways the intervention exacerbates or alleviates inequities. To put this into practice, one goal is the development of real-time metrics that inform municipalities and states on the equity of interventions and health outcomes, enabling adjustments and responses to keep equity at the forefront of intervention decisions.

Other important examples of the links between surveillance and health equity were discussed during the symposium. While documenting disparate impacts is a necessary starting point, identifying appropriate measures to rectify these inequities will often require understanding where, why, when, and how they arise (99). An early example in the U.S. was a documentation that higher SARS-CoV-2 prevalence among

BOX 1 Modeling and analytic support for STLT jurisdictions: the role of academic groups.

COVID-19 stimulated numerous collaborations between STLT health authorities and university and other research institutions to support decision making with modeling, analytics, and forecasts. These took multiple forms, ranging from the establishment of advisory councils to mayors and governors, to bilateral collaborations (80, 81) and formal consortia (<https://modelingconsortium.ucsf.edu/>). However, there are stark differences among jurisdictions in the number of such academic groups within the jurisdiction and/or with existing or prior links with the jurisdiction's health department, creating inequities in access to this kind of advice. The benefits of working with academic partners can include local knowledge and the capacity to surge efforts in an emergency. Potential barriers to such collaborations that should be addressed up front where possible include academics' need and incentives to publish, which may compete for time with their role in decision support, as well as the demands of academic schedules, whereby, for example, a key analyst on a project may have to devote effort to exams at times when they are needed for decision support. Administrative preparedness in the form of preexisting data use agreements can vastly accelerate these efforts.

Establishment of trust is essential to the success of academic – STLT collaborations. Elected and health officials at the symposium noted the repeated challenges of figuring out which models and modelers to trust, both locally and nationally. Participants observed that academic collaborations were most effective when there was a pre-existing relationship between the groups and the jurisdictions, and noted the benefits to both parties of cultivating these relationships in “peacetime” through collaboration on non-pandemic activities. Academic groups' ability to speak freely can lend credibility and objectivity to their analyses; however, trust can be undermined if academic groups with access to limited, publicly available data release analyses in publications or preprints that may be inconsistent with more complete data that are available to health departments but not publicly available. Frequent contact to share tentative conclusions and compare them against the evolving understanding of health officials can enhance the quality of analyses by incorporating more complete data, if these can be shared, and can enhance the trust between the parties, improving future interactions. When such interactions work well, they do not stifle the conclusions of academic groups but rather ensure that these conclusions are based on the best current understanding and to ensure that health officials are aware of what is being published about data from their jurisdictions. Academic incentives and structures are particularly not suited for routinely repeated analyses, such as reproductive number estimation, nowcasting, and forecasting, although academic centers have played key roles in these areas for over 2 years during COVID-19. Automation, as in the California consortium's dashboard, is one solution. CDC's Center for Forecasting and Outbreak Analytics is beginning to take on some of these tasks and will increasingly serve as a focal point for such repeated, real-time analyses.

BOX 2 National insight from local evidence.

Implementation of public health policies is a state/territorial/tribal and local responsibility in the U.S., as we noted above. Infectious disease surveillance is also decentralized, often with two levels of reporting (local/county and state/territorial) below the national level. From the perspective of national decision makers seeking a clear picture of an unfolding pandemic, decentralized surveillance has obvious limitations, particularly in a setting where data systems and data use agreements vary across jurisdictions. Efforts are underway, and should be expanded, to improve the speed, completeness, and accuracy of data flowing from states, localities, and health care systems to the CDC and other federal actors. Such efforts are essential for timely situational awareness and for calibrating the outputs of scenarios and forecasts to granular (state or county-level) data to form a national picture.

While incidence, prevalence, and health care burden are intrinsically local quantities that need to be estimated everywhere and over time, many aspects of surveillance and associated epidemiology are generalizable, such that findings in one local jurisdiction can inform control measures everywhere. These include characteristics of the pathogen, such as severity and natural history; and characteristics of countermeasures, such as test sensitivity and the effectiveness of drugs and vaccines. For these purposes, local conditions can facilitate detailed characterization that may not be possible, but also may not be necessary, on a larger geographic scale.

Some of the earliest evidence of low severity for the 2009 H1N1 influenza pandemic came from a study at the University of Delaware, where a comparatively self-contained population could be studied in detail (88). We noted in a postmortem of that pandemic that the findings from that study were not widely known until months later because of limited dissemination (4), arguably prolonging the state of alarm unnecessarily during that pandemic. In COVID-19, early findings of asymptomatic/presymptomatic infection and likely transmission from studies in a nursing home and a cruise ship, respectively (19, 89), were documented very early and widely disseminated, but still did not fully inform control measures.

In many other cases, detailed surveillance and epidemiology in local jurisdictions or health systems provided evidence of national and international importance. A few examples included:

- Evidence from the Yukon-Kuskokwim (Alaska) Health Corporation about the persistence of antigen test positivity 5 or more days after initial positive test or symptom onset during the early Omicron era (90)
- Evidence from the Kaiser Permanente Southern California health system about the relative clinical severity of Omicron BA.1 variant compared to Delta before it and BA.2 after (74)
- others

Each of these provided evidence that could be generalized beyond the location where it was generated, because it concerned generalizable features of the infection or countermeasures based on its biology. The degree to which these investigations informed policy and guidelines varied, indicating a need for a systematic approach to disseminating findings of wide importance and updating guidance in a way that reflects the totality of data.

mothers admitted for labor and delivery in New York City was associated with residence in boroughs with smaller reductions in mobility, suggesting inability to work from home as a potential driver of risk (68). Subsequent

modeling work explained racial disparities in infection rates in U.S. cities as a consequence of higher exposure by minority groups not only to infection generally, but particularly to more crowded venues with higher

infection risk (58). The age distribution of mortality by race/ethnicity (100), with its skew to younger ages among Black, Hispanic, and Indigenous individuals, also pointed to an increased exposure risk. The United Kingdom Government's Race Disparity Unit published a series of reports through the first 2 years of the pandemic enumerating hypotheses for mechanisms to explain disparate impacts, stating the current evidence related to these hypotheses, and recommending actions to address these drivers of higher incidence and severity in racial and ethnic minorities (60). In the United States, such studies may require linkage of disparate data bases to identify where disparities arise during the cascade of care (101), a strategy that has long been useful in HIV/AIDS surveillance to understand loss points in the continuum of care (102, 103). For COVID, a full cascade would require an estimate of the actual number of infected individuals, the number of people who have been identified by testing (reflecting under-diagnosis), the number treated when treatment became available (reflecting under-treatment), the number hospitalized (reflecting access to care and disease severity), and the number of fatalities, jointly stratified by race and ethnicity, age group and sex. Ascertaining these would require both modeling-based estimates and data from multiple sources (e.g., clinical laboratories and vital registries). For example, an analysis from a New York City hospital suggested no racial difference in case fatality among hospitalized patients, supporting the idea that racial differences in exposure (more infections) rather than racial differences in outcome contributed to racial differences in overall mortality (104).

Other sources of inequity can affect case ascertainment and thus identification of opportunities for intervention. Geographic and temporal variation in testing effort in the U.S. was very large, resulting in difficulties in comparing incidence across jurisdictions. Rural areas were often the least able to access testing, though there were important exceptions (90). Notably, the use of random sampling stratified by geography mitigated this problem significantly in the United Kingdom (85), though it did not solve it entirely because participation was of necessity voluntary. Equity considerations may change as public health authorities rely on new data sources; for example, mobility estimates may depend on smartphone ownership, while wastewater surveillance for pathogen abundance will be unavailable in areas using septic systems (29, 105).

Expanding the range of data types

As described above, any health system, but particularly one as decentralized as that of the U.S., benefits from the ability to ingest and synthesize multiple types of data. Increased use of wastewater data (106, 107) has contributed to early warning of rising infection incidence and to surveillance for new variants. Further work to standardize collection and better define the quantitative relationships between true infection incidence and total and variant-specific concentrations of viral genomes in wastewater is needed to improve the value of such data, as well as a clear mapping of where it will not be informative, such as areas using septic systems. Likewise, mobility data from various sources (58) can be useful in informing strategies for disease monitoring and surveillance, modeling disease spread, and guiding interventions. Immune measures from serology provide a window onto past infection and a lens onto the landscape of risk (108). Here, further work is needed to ensure data standardization and accuracy as well as routine and frequent updating to capture important temporal variations. Such new forms of data may also raise

privacy considerations that have not entirely been solved (109, 110).

Crowdsourced and survey data (111, 112) can provide important insights into behaviors that affect the interpretations of other data; for example, the increasing prevalence of self-testing using antigen tests for COVID-19 reduces the utility of PCR-positive case counts.

A key to making use of this expanded range of data types is solving the problem of how to synthesize multiple data types into a single estimate of a quantity of interest, accounting for the different properties of each data type (76), including understanding the different biases that will affect each data stream. Significant further work is needed to advance the ability to do this in real time. A related but distinct problem is how to link data across data systems to understand the continuum of care and otherwise improve inference about the course of individual cases.

Expanding the range of data sources

The use of claims data from health care payers (insurers) and electronic medical records from providers has exploded in many areas of health services research. There have been some notable examples of such data for surveillance to address the questions described in this report (53, 72–74), but in the United States there remains untapped potential to expand such efforts and improve their timeliness. This will require building relationships between public health entities and health care systems in their jurisdictions, including relationships between scientific investigators in each sector with regular discussions for bidirectional learning. In the spirit of administrative preparedness above, this will require up-front planning of master agreements to move resources in a timely fashion to address pressing questions. Health providers and public health have suffered from a “two cultures” challenge that results in the need to expand public health training of investigators and other personnel in health systems, acknowledge the contributions of health systems to community benefits, and find ways to produce incentives so that contributing to public health surveillance aligns with the business interests of health systems. Medical examiners and coroners are another group that has been disconnected from public health but with whom cooperation can enhance and help to calibrate surveillance for pathogen-specific deaths, as illustrated by some examples both domestically (113) and abroad (114).

As we described above, new data sources become useful in proportion to our understanding of their “normal” behavior. As we expand the range of data types, it will be essential to monitor new data streams and continue to monitor old ones outside of epidemic periods to establish a baseline that can be used to calibrate signals of new outbreaks and estimate the exceedance caused by the ongoing transmission of novel pathogens (115).

Conclusion

Data and modeling needs change over the course of a pandemic and vary by the jurisdictional dimensions, requiring anticipatory, rapid, dynamic, and locally adapted and scaled activities to optimize pandemic management and population health. Here, we have sought to describe concepts, tools, and strategies to address those needs, building on those enacted during the COVID-19 pandemic and those that could have

facilitated this work. While not a comprehensive list, we hope that the ideas we propose and envision serve as a useful resource and guide in efforts to manage ongoing infectious diseases challenges and preparedness for the inevitable next pandemic.

Author contributions

ML: Conceptualization, Methodology, Project administration, Writing – original draft, Writing – review & editing. MB: Investigation, Writing – review & editing. JB: Investigation, Writing – review & editing. PE: Investigation, Writing – review & editing. DE: Investigation, Writing – review & editing. MG: Investigation, Writing – review & editing. JH: Investigation, Writing – review & editing. MJ: Investigation, Writing – review & editing. SK: Investigation, Writing – review & editing. DL: Investigation, Writing – review & editing. JeL: Investigation, Writing – review & editing. JuL: Investigation, Writing – review & editing. RL: Investigation, Writing – review & editing. DM: Investigation, Writing – review & editing. LCM: Investigation, Writing – review & editing. CM: Investigation, Writing – review & editing. LAM: Investigation, Writing – review & editing. SO: Investigation, Writing – review & editing. CQ: Investigation, Writing – review & editing. AB: Investigation, Writing – review & editing. NR: Investigation, Writing – review & editing. SR: Investigation, Writing – review & editing. RR: Investigation, Writing – review & editing. MS: Investigation, Writing – review & editing. RS: Investigation, Writing – review & editing. RBS: Investigation, Writing – review & editing. DS: Investigation, Writing – review & editing. ST: Investigation, Writing – review & editing. JV: Investigation, Writing – review & editing. YG: Funding acquisition, Investigation, Project administration, Writing – original draft, Writing – review & editing, Conceptualization.

References

- Sherratt K, Gruson H, Grah R, Johnson H, Niehus R, Prasse B, et al. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations. *eLife*. (2023) 12:1916. doi: 10.7554/eLife.81916
- Cramer EY, Huang Y, Wang Y, Ray EL, Cornell M, Bracher J, et al. The United States COVID-19 forecast hub dataset. *Sci Data*. (2022) 9:462. doi: 10.1038/s41597-022-01517-w
- Lopez VK, Cramer EY, Pagano R, Drake JM, O'Dea EB, Adey M, et al. Challenges of COVID-19 case forecasting in the US, 2020–2021. *PLoS Comput Biol*. (2024) 20:e1011200. doi: 10.1371/journal.pcbi.1011200
- Lipsitch M, Finelli L, Heffernan RT, Leung GM, Redd SC. Improving the evidence base for decision making during a pandemic: the example of 2009 influenza A/H1N1 [research support, N.I.H., extramural]. *Biosecur Bioterror*. (2011) 9:89–115. doi: 10.1089/bsp.2011.0007
- Scalia Tomba G, Wallinga J. A simple explanation for the low impact of border control as a countermeasure to the spread of an infectious disease. *Math Biosci*. (2008) 214:70–2. doi: 10.1016/j.mbs.2008.02.009
- Quilty BJ, Clifford S, Flasche S, Eggo RMCMMID nCoV working group. Effectiveness of airport screening at detecting travellers infected with novel coronavirus (2019-nCoV). *Euro Surveill*. (2020) 25:80. doi: 10.2807/1560-7917.ES.2020.25.5.2000080
- Cowling BJ, Lau LL, Wu P, Wong HW, Fang VJ, Riley S, et al. Entry screening to delay local transmission of 2009 pandemic influenza A (H1N1) [research support, N.I.H., extramural research support, non-U.S. Gov't]. *BMC Infect Dis*. (2010) 10:82. doi: 10.1186/1471-2334-10-82
- Li J, Hosegood I, Powell D, Tschärke B, Lawler J, Thomas KV, et al. A global aircraft-based wastewater genomic surveillance network for early warning of future pandemics. *Lancet Glob Health*. (2023) 11:e791–5. doi: 10.1016/S2214-109X(23)00129-8
- Carrión M, Madoff LC. ProMED-mail: 22 years of digital surveillance of emerging infectious diseases. *Int Health*. (2017) 9:177–83. doi: 10.1093/inthealth/ihx014
- Tisza MJ, Hanson BM, Clark JR, Wang L, Payne K, Ross MC, et al. Virome sequencing identifies H5N1 avian influenza in wastewater from nine cities. *medRxiv*. (2024) 2024:24307179. doi: 10.1101/2024.05.10.24307179
- Ma KC, Shirk P, Lambrou AS, Hassell N, Zheng XY, Payne AB, et al. Genomic surveillance for SARS-CoV-2 variants: circulation of omicron lineages - United States, January 2022–may 2023. *MMWR Morb Mortal Wkly Rep*. (2023) 72:651–6. doi: 10.15585/mmwr.mm7224a2
- Boeckh M, Chu HY, Englund JA, Lockwood CM, Nickerson DA, Shendure J, et al. The Seattle flu study: when regulations hinder pandemic surveillance. *Nat Med*. (2022) 28:7–8. doi: 10.1038/s41591-021-01587-0
- Yang H, Surer O, Duque D, Morton DP, Singh B, Fox SJ, et al. Design of COVID-19 staged alert systems to ensure healthcare capacity with minimal closures. *Nat Commun*. (2021) 12:3767. doi: 10.1038/s41467-021-23989-x
- Fox SJ, Lachmann M, Tec M, Pasco R, Woody S, Du Z, et al. Real-time pandemic surveillance using hospital admissions and mobility data. *Proc Natl Acad Sci USA*. (2022) 119:119. doi: 10.1073/pnas.2111870119
- CDC. *Health Department Governance*. (2022). Available at: <https://www.cdc.gov/publichealthgateway/sites/governance/index.html>.
- Gupta S, Nguyen T, Raman S, Lee B, Lozano-Rojas F, Bento A, et al. Tracking public and private responses to the COVID-19 epidemic: evidence from state and local government actions. *Am J Health Econ*. (2021) 7:361–404. doi: 10.1086/716197
- Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial interval of COVID-19 among publicly reported confirmed cases. *Emerg Infect Dis*. (2020) 26:1341–3. doi: 10.3201/eid2606.200357
- Tsang TK, Wu P, Lin Y, Lau EHY, Leung GM, Cowling BJ. Effect of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China: a modelling study. *Lancet Public Health*. (2020) 5:e289–96. doi: 10.1016/S2468-2667(20)30089-X

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This project has been funded (in part) by contract 200–2016–91779 with the Centers for Disease Control and Prevention (CDC).

Conflict of interest

RS was employed by company Siemens Healthcare Diagnostics. JV was employed by company SIGA Technologies.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The findings, conclusions, and views expressed are those of the author(s) and do not necessarily represent the official position of the CDC.

19. Emery JC, Russell TW, Liu Y, Hellewell J, Pearson CA, Group CC-W et al. The contribution of asymptomatic SARS-CoV-2 infections to transmission on the Diamond princess cruise ship. *eLife*. (2020) 9:9. doi: 10.7554/eLife.58699
20. Lu J, Gu J, Li K, Xu C, Su W, Lai Z, et al. COVID-19 outbreak associated with air conditioning in restaurant, Guangzhou, China, 2020. *Emerg Infect Dis*. (2020) 26:1628–31. doi: 10.3201/eid2607.200764
21. Addetia A, Crawford KHD, Dingens A, Zhu H, Roychoudhury P, Huang ML, et al. Neutralizing antibodies correlate with protection from SARS-CoV-2 in humans during a fishery vessel outbreak with a high attack rate. *J Clin Microbiol*. (2020) 58:20. doi: 10.1128/JCM.02107-20
22. Lipsitch M, Lajous M, O'Hagan JJ, Cohen T, Miller JC, Goldstein E, et al. Use of cumulative incidence of novel influenza A/H1N1 in foreign travelers to estimate lower bounds on cumulative incidence in Mexico. *PLoS One*. (2009) 4:e6895. doi: 10.1371/journal.pone.0006895
23. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science*. (2009) 324:1557–61. doi: 10.1126/science.1176062
24. Niehus R, De Salazar PM, Taylor AR, Lipsitch M. Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. *Lancet Infect Dis*. (2020) 20:803–8. doi: 10.1016/S1473-3099(20)30229-2
25. De Salazar PM, Niehus R, Taylor A, Buckee CO, Lipsitch M. Identifying locations with possible undetected imported severe acute respiratory syndrome coronavirus 2 cases by using importation predictions. *Emerg Infect Dis*. (2020) 26:1465–9. doi: 10.3201/eid2607.200250
26. Bento AI, Nguyen T, Wing C, Lozano-Rojas F, Ahn YY, Simon K. Evidence from internet search data shows information-seeking responses to news of local COVID-19 cases. *Proc Natl Acad Sci USA*. (2020) 117:11220–2. doi: 10.1073/pnas.2005335117
27. Kishore N, Taylor AR, Jacob PE, Vembar N, Cohen T, Buckee CO, et al. Evaluating the reliability of mobility metrics from aggregated mobile phone data as proxies for SARS-CoV-2 transmission in the USA: a population-based study. *Lancet Digit Health*. (2022) 4:e27–36. doi: 10.1016/S2589-7500(21)00214-4
28. National Academies of Science Engineering and Medicine. Wastewater-based disease surveillance for Public Health action. Washington, DC: National Academies of Science Engineering and Medicine (2023).
29. Yu Q, Olesen SW, Duvallet C, Grad YH. Assessment of sewer connectivity in the United States and its implications for equity in wastewater-based epidemiology. *PLOS Glob Public Health*. (2024) 4:e0003039. doi: 10.1371/journal.pgph.0003039
30. Hay JA, Kennedy-Shaffer L, Kanjilal S, Lennon NJ, Gabriel SB, Lipsitch M, et al. Estimating epidemiologic dynamics from cross-sectional viral load distributions. *Science*. (2021) 373:635. doi: 10.1126/science.abh0635
31. Kissler SM, Hay JA, Fauver JR, Mack C, Tai CG, Anderson DJ, et al. Viral kinetics of sequential SARS-CoV-2 infections. *Nat Commun*. (2023) 14:6206. doi: 10.1038/s41467-023-41941-z
32. Kissler SM, Fauver JR, Mack C, Tai CG, Breban MI, Watkins AE, et al. Viral dynamics of SARS-CoV-2 variants in vaccinated and unvaccinated persons. *N Engl J Med*. (2021) 385:2489–91. doi: 10.1056/NEJMc2102507
33. Hay JA, Kissler SM, Fauver JR, Mack C, Tai CG, Samant RM, et al. Quantifying the impact of immune history and variant on SARS-CoV-2 viral kinetics and infection rebound: a retrospective cohort study. *eLife*. (2022) 11:11. doi: 10.7554/eLife.81849
34. Healthmap Homepage. (n.d.). Available at: <http://healthmap.org/en/>.
35. McGough SF, Brownstein JS, Hawkins JB, Santillana M. Forecasting Zika incidence in the 2016 Latin America outbreak combining traditional disease surveillance with search, social media, and news report data. *PLoS Negl Trop Dis*. (2017) 11:e0005295. doi: 10.1371/journal.pntd.0005295
36. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining search, social media, and traditional data sources to improve influenza surveillance. *PLoS Comput Biol*. (2015) 11:e1004513. doi: 10.1371/journal.pcbi.1004513
37. Kogan NE, Clemente L, Liautaud P, Kaashoek J, Link NB, Nguyen AT, et al. An early warning approach to monitor COVID-19 activity with multiple digital traces in near real time. *Sci Adv*. (2021) 7:989. doi: 10.1126/sciadv.abd6989
38. Jones JM, Stone M, Sulaeman H, Fink RV, Dave H, Levy ME, et al. Estimated US infection- and vaccine-induced SARS-CoV-2 Seroprevalence based on blood donations, July 2020–May 2021. *JAMA*. (2021) 326:1400–9. doi: 10.1001/jama.2021.15161
39. Follmann D, Janes HE, Chu E, Jayashankar L, Petropoulos CJ, Serebryanny L, et al. Kinetics of the antibody response to symptomatic SARS-CoV-2 infection in vaccinated and unvaccinated individuals in the blinded phase of the mRNA-1273 COVID-19 vaccine efficacy trial. *Open Forum Infect Dis*. (2023) 10:ofad069. doi: 10.1093/ofid/ofad069
40. Luoma E, Rohrer R, Parton H, Hughes S, Omeregbe E, Taki F, et al. Notes from the Field: epidemiologic characteristics of SARS-CoV-2 recombinant variant B.1.1.5 - new York City, November 1, 2022–January 4, 2023. *MMWR Morb Mortal Wkly Rep*. (2023) 72:212–4. doi: 10.15585/mmwr.mm7208a4
41. Chitwood MH, Russi M, Gunasekera K, Havumaki J, Klaassen F, Pitzer VE, et al. Reconstructing the course of the COVID-19 epidemic over 2020 for US states and counties: results of a Bayesian evidence synthesis model. *PLoS Comput Biol*. (2022) 18:e1010465. doi: 10.1371/journal.pcbi.1010465
42. Lison A, Abbott S, Huisman J, Stadler T. Generative Bayesian modeling to nowcast the effective reproduction number from line list data with missing symptom onset dates. *PLoS Comput Biol*. (2024) 20:e1012021. doi: 10.1371/journal.pcbi.1012021
43. Howerton E, Contamin L, Mullany LC, Qin M, Reich NG, Bents S, et al. Evaluation of the US COVID-19 scenario modeling hub for informing pandemic response under uncertainty. *Nat Commun*. (2023) 14:7260. doi: 10.1038/s41467-023-42680-x
44. Kahn R, Holmdahl I, Reddy S, Jernigan J, Mina MJ, Slayton RB. Mathematical modeling to inform vaccination strategies and testing approaches for COVID-19 in nursing homes. *Clin Infect Dis*. (2021) 74:597–603. doi: 10.1093/cid/ciab517
45. Pung R, Clapham HE, Russell TW, Group CC-WLee VJ, Kucharski AJ. Relative role of border restrictions, case finding and contact tracing in controlling SARS-CoV-2 in the presence of undetected transmission: a mathematical modelling study. *BMC Med*. (2023) 21:802. doi: 10.1186/s12916-023-02802-0
46. Quilty BJ, Clifford S, Hellewell J, Russell TW, Kucharski AJ, Flasche S, et al. Centre for the Mathematical Modelling of infectious diseases C-wg. Quarantine and testing strategies in contact tracing for SARS-CoV-2: a modelling study. *Lancet Public Health*. (2021) 6:e175–83. doi: 10.1016/S2468-2667(20)30308-X
47. Howerton E, Contamin L, Mullany LC, Qin M, Reich NG, Bents S, et al. Informing pandemic response in the face of uncertainty: an evaluation of the U.S. COVID-19 scenario modeling hub. *medRxiv*. (2023) 2023:23291998. doi: 10.1101/2023.06.28.23291998
48. Singanayagam A, Hakki S, Dunning J, Madon KJ, Crone MA, Koycheva A, et al. Community transmission and viral load kinetics of the SARS-CoV-2 delta (B.1.617.2) variant in vaccinated and unvaccinated individuals in the UK: a prospective, longitudinal, cohort study. *Lancet Infect Dis*. (2022) 22:183–95. doi: 10.1016/S1473-3099(21)00648-4
49. Kissler SM, Fauver JR, Mack C, Olesen SW, Tai C, Shiue KY, et al. Viral dynamics of acute SARS-CoV-2 infection and applications to diagnostic and public health strategies. *PLoS Biol*. (2021) 19:e3001333. doi: 10.1371/journal.pbio.3001333
50. Block JP, Boehmer TK, Forrest CB, Carton TW, Lee GM, Ajani UA, et al. Cardiac complications after SARS-CoV-2 infection and mRNA COVID-19 vaccination - PCORnet, United States, January 2021–January 2022. *MMWR Morb Mortal Wkly Rep*. (2022) 71:517–23. doi: 10.15585/mmwr.mm7114e1
51. McNeil MM, Gee J, Weintraub ES, Belongia EA, Lee GM, Glanz JM, et al. The vaccine safety datalink: successes and challenges monitoring vaccine safety. *Vaccine*. (2014) 32:5390–8. doi: 10.1016/j.vaccine.2014.07.073
52. Lewnard JA, Hong VX, Patel MM, Kahn R, Lipsitch M, Tartof SY. Clinical outcomes associated with omicron (B.1.1.529) variant and BA.1/BA.1.1 or BA.2 subvariant infection in southern California. *Nat Med*. (2022) 28:1933–1943.
53. Dickerman BA, Gerlovina H, Madenci AL, Kurgansky KE, Ferolito BR, Figueroa Muniz MJ, et al. Comparative effectiveness of BNT162b2 and mRNA-1273 vaccines in U.S. *Vet N Engl J Med*. (2022) 386:105–15. doi: 10.1056/NEJMoa2115463
54. Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, et al. BNT162b2 mRNA Covid-19 vaccine in a Nationwide mass vaccination setting. *N Engl J Med*. (2021) 384:1412–23. doi: 10.1056/NEJMoa2101765
55. CDC. CDC's Influenza Vaccine Effectiveness Networks. (2021). Available at: <https://www.cdc.gov/flu/vaccines-work/vaccine-effectiveness-networks.htm>.
56. CDC. COVID-19 Vaccinations in the United States. (2024). Available at: <https://covid.cdc.gov/covid-data-tracker/#vaccination-states-jurisdictions>.
57. US Census Bureau. Household Pulse Survey COVID-19 Vaccination Tracker. (2021). Available at: <https://www.census.gov/library/visualizations/interactive/household-pulse-survey-covid-19-vaccination-tracker.html>.
58. Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*. (2021) 589:82–7. doi: 10.1038/s41586-020-2923-3
59. Carnegie-Mellon University. (2021). Available at: <https://delphi.cmu.edu/covidcast/>.
60. UK Government Equality Hub and Race Disparity Unit. Final report on progress to address COVID-19 health inequalities. (2021). Available at: <https://www.gov.uk/government/publications/final-report-on-progress-to-address-covid-19-health-inequalities/final-report-on-progress-to-address-covid-19-health-inequalities>.
61. Ma KC, Hale JE, Grad YH, Alter G, Luzuriaga K, Eaton RB, et al. Trends in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Seroprevalence in Massachusetts estimated from newborn screening specimens. *Clin Infect Dis*. (2022) 75:e105–13. doi: 10.1093/cid/ciac158
62. Wiegand RE, Deng Y, Deng X, Lee A, Meyer WA3rd, Letovsky S, et al. Estimated SARS-CoV-2 antibody seroprevalence trends and relationship to reported case prevalence from a repeated, cross-sectional study in the 50 states and the District of Columbia, United States–October 25, 2020–February 26, 2022. *Lancet Reg Health Am*. (2023) 18:100403. doi: 10.1016/j.lana.2022.100403

63. Garcia-Carreras B, Hitchings MDT, Johansson MA, Biggerstaff M, Slayton RB, Healy JM, et al. Accounting for assay performance when estimating the temporal dynamics in SARS-CoV-2 seroprevalence in the U.S. *Nat Commun.* (2023) 14:944. doi: 10.1038/s41467-023-37944-5
64. Ng DL, Goldgof GM, Shy BR, Levine AG, Balcerak J, Bapat SP, et al. SARS-CoV-2 seroprevalence and neutralizing activity in donor and patient blood. *Nat Commun.* (2020) 11:4698. doi: 10.1038/s41467-020-18468-8
65. Self WH, Tenforde MW, Stubblefield WB, Feldstein LR, Steingrub JS, Shapiro NI, et al. Decline in SARS-CoV-2 antibodies after mild infection among frontline Health care personnel in a multistate hospital network - 12 states, April-august 2020. *MMWR Morb Mortal Wkly Rep.* (2020) 69:1762–6. doi: 10.15585/mmwr.mm6947a2
66. Si Y, Covello L, Wang S, Covello T, Gelman A. Beyond vaccination rates: a synthetic random proxy metric of Total SARS-CoV-2 immunity Seroprevalence in the community. *Epidemiology.* (2022) 33:457–64. doi: 10.1097/EDE.0000000000001488
67. Covello L, Gelman A, Si Y, Wang S. Routine hospital-based SARS-CoV-2 testing outperforms state-based data in predicting clinical burden. *Epidemiology.* (2021) 32:792–9. doi: 10.1097/EDE.0000000000001396
68. Kissler SM, Kishore N, Prabhu M, Goffman D, Beilin Y, Landau R, et al. Reductions in commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in new York City. *Nat Commun.* (2020) 11:4674. doi: 10.1038/s41467-020-18271-5
69. US Department of Health and Human Services. *COVID-19 Reported Patient Impact and Hospital Capacity by State Timeseries (RAW).* (2023). Available at: <https://healthdata.gov/Hospital/COVID-19-Reported-Patient-Impact-and-Hospital-Capacity-62h-syeh>.
70. US Department of Health and Human Services. Activities Deemed Not to Be Research: Public Health Surveillance (2018). Available at: <https://www.hhs.gov/ohrp/regulations-and-policy/requests-for-comments/draft-guidance-activities-deemed-not-be-research-public-health-surveillance/index.html>.
71. Gobeille V. *Liberty mutual insurance co.* S.Ct. p. 936 (2016).
72. Accorsi EK, Britton A, Fleming-Dutra KE, Smith ZR, Shang N, Derado G, et al. Association between 3 doses of mRNA COVID-19 vaccine and symptomatic infection caused by the SARS-CoV-2 omicron and Delta variants. *JAMA.* (2022) 327:639–51. doi: 10.1001/jama.2022.0470
73. Link-Gelles R, Levy ME, Gaglani M, Irving SA, Stockwell M, Dascomb K, et al. Effectiveness of 2, 3, and 4 COVID-19 mRNA Vaccine Doses Among Immunocompetent Adults During Periods when SARS-CoV-2 Omicron BA.1 and BA.2/BA.2.12.1 Sublineages Predominated - VISION Network, 10 States, December 2021–June 2022. *MMWR Morb Mortal Wkly Rep.* (2022) 71:931–9. doi: 10.15585/mmwr.mm7129e1
74. Lewnard JA, Hong VX, Patel MM, Kahn R, Lipsitch M, Tartof SY. Clinical outcomes associated with SARS-CoV-2 omicron (B.1.1.529) variant and BA.1/BA.1.1 or BA.2 subvariant infection in Southern California. *Nat Med.* (2022) 28:1933–43. doi: 10.1038/s41591-022-01887-z
75. Shah MM, Joyce B, Plumb ID, Sahakian S, Feldstein LR, Barkley E, et al. Paxlovid associated with decreased hospitalization rate among adults with COVID-19 - United States, April–September 2022. *MMWR Morb Mortal Wkly Rep.* (2022) 71:1531–7. doi: 10.15585/mmwr.mm7148e2
76. Reinhart A, Brooks L, Jahja M, Rumack A, Tang J, Agrawal S, et al. An open repository of real-time COVID-19 indicators. *Proc Natl Acad Sci USA.* (2021) 118:2118. doi: 10.1073/pnas.2111452118
77. Covid Crisis Group. Lessons from the COVID war: An investigative report. New York: Public Affairs Books (2023). 352 p.
78. Infectious Diseases Society of America. *The Leaders in Epidemiology, Antimicrobial Stewardship and Public Health Fellowship.* (2023). Available at: www.idsociety.org/professional-development/fellows-in-training-career-education-center/leap-fellowship/leap-fellowship/.
79. Infectious Diseases Society of America. *ID/EIS joint fellowship - New Opportunity!* (2023). Available at: <https://www.idsociety.org/professional-development/fellows-in-training-career-education-center/joint-id-eis-fellowship/id-eis-fellowship/>.
80. Yang W, Kandula S, Huynh M, Greene SK, Van Wye G, Li W, et al. Estimating the infection-fatality risk of SARS-CoV-2 in new York City during the spring 2020 pandemic wave: a model-based analysis. *Lancet Infect Dis.* (2021) 21:203–12. doi: 10.1016/S1473-3099(20)30769-6
81. Greene SK, McGough SF, Culp GM, Graf LE, Lipsitch M, Menzies NA, et al. Nowcasting for real-time COVID-19 tracking in new York City: an evaluation using reportable disease data from early in the pandemic. *JMIR Public Health Surveill.* (2021) 7:e25538. doi: 10.2196/25538
82. Data. (2023). Available at: <https://data.org/initiatives/epiverse/>.
83. R Epidemiology Consortium. *RECON.* (2023). Available at: <https://www.repidemicsconsortium.org/>.
84. Gouglas D, Christodoulou M, Hatchett R. The 100 days Mission—2022 global pandemic preparedness summit. *Emerg Infect Dis.* (2023) 29:21142. doi: 10.3201/eid2903.221142
85. Elliott P, Whitaker M, Tang D, Eales O, Steyn N, Bodinier B, et al. Design and implementation of a national SARS-CoV-2 monitoring program in England: REACT-1 study. *Am J Public Health.* (2023) 113:545–54. doi: 10.2105/AJPH.2023.307230
86. Pulliam JRC, van Schalkwyk C, Govender N, von Gottberg A, Cohen C, Groome MJ, et al. Increased risk of SARS-CoV-2 reinfection associated with emergence of omicron in South Africa. *Science.* (2022) 376:4947. doi: 10.1126/science.abn4947
87. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med.* (2020) 382:1199–207. doi: 10.1056/NEJMoa2001316
88. Juliano AD, Reed C, Guh A, Desai M, Dee DL, Kutty P, et al. Notes from the field: outbreak of 2009 pandemic influenza A (H1N1) virus at a large public university in Delaware, April–may 2009. *Clin Infect Dis.* (2009) 49:1811–20. doi: 10.1086/649555
89. Kimball A, Hatfield KM, Arons M, James A, Taylor J, Spicer K, et al. Asymptomatic and Presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility - King County, Washington, march 2020. *MMWR Morb Mortal Wkly Rep.* (2020) 13) 69:377–81. doi: 10.15585/mmwr.mm6913e1
90. Lefferts B, Blake I, Bruden D, Hagen MB, Hodges E, Kirking HL, et al. Antigen test positivity after COVID-19 isolation - Yukon-Kuskokwim Delta region, Alaska, January–February 2022. *MMWR Morb Mortal Wkly Rep.* (2022) 71:293–8. doi: 10.15585/mmwr.mm7108a3
91. Lipsitch M, Santillana M. Enhancing situational awareness to prevent infectious disease outbreaks from becoming catastrophic. *Curr Top Microbiol Immunol.* (2019) 424:59–74. doi: 10.1007/82_2019_172
92. CDC. *Mpox Technical Reports.* (2021). Available at: <https://www.cdc.gov/poxvirus/mpox/cases-data/technical-report.html>.
93. UKHSA. *Investigation of SARS-CoV-2 variants: technical briefings.* (2021). Available at: <https://www.gov.uk/government/publications/investigation-of-sars-cov-2-variants-technical-briefings>.
94. Feldman JM, Bassett MT. Variation in COVID-19 mortality in the US by race and ethnicity and educational attainment. *JAMA Netw Open.* (2021) 4:e2135967. doi: 10.1001/jamanetworkopen.2021.35967
95. Department of Health and Social Care U. *Technical report on the COVID-19 pandemic in the UK: A technical report for future UK Chief Medical Officers, Government Chief Scientific Advisers, National Medical Directors and public health leaders in a pandemic.* (2023). Available at: <https://www.gov.uk/government/publications/technical-report-on-the-covid-19-pandemic-in-the-uk>.
96. Mena GE, Martinez PP, Mahmud AS, Marquet PA, Buckee CO, Santillana M. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science.* (2021) 372:5298. doi: 10.1126/science.abg5298
97. Lundberg DJ, Wrigley-Field E, Cho A, Raquib R, Nsoesie EO, Paglino E, et al. COVID-19 mortality by race and ethnicity in US metropolitan and nonmetropolitan areas, march 2020 to February 2022. *JAMA Netw Open.* (2023) 6:e2311098. doi: 10.1001/jamanetworkopen.2023.11098
98. Presidential COVID-19 Health Equity Task Force. *Final Report and Recommendations.* (2021). Available at: https://www.minorityhealth.hhs.gov/assets/pdf/HETF_Report_508_102821_9am_508Team%20WTP11-compressed.pdf
99. Jones CP. Invited commentary: "race," racism, and the practice of epidemiology. *Am J Epidemiol.* (2001) 154:299–304. doi: 10.1093/aje/154.4.299
100. Bassett MT, Chen JT, Krieger N. Variation in racial/ethnic disparities in COVID-19 mortality by age in the United States: a cross-sectional study. *PLoS Med.* (2020) 17:e1003402. doi: 10.1371/journal.pmed.1003402
101. NYC Department of Health and Mental Hygiene. *Racial Inequities in COVID-19 Hospitalizations During the Omicron Wave in NYC, p. 17.* (2022). Available at: <https://www.nyc.gov/assets/doh/downloads/pdf/covid/black-hospitalizations-omicron-wave.pdf>.
102. The White House. *National HIV/AIDS Strategy for the United States: Updated to 2020.* (2015). Available at: <https://files.hiv.gov/s3fs-public/nhas-update.pdf>.
103. Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection [research support, N.I.H., extramural review]. *Clin Infect Dis.* (2011) 52:793–800. doi: 10.1093/cid/ciq243
104. Ogedegbe G, Ravenell J, Adhikari S, Butler M, Cook T, Francois F, et al. Assessment of racial/ethnic disparities in hospitalization and mortality in patients with COVID-19 in new York City. *JAMA Netw Open.* (2020) 3:e2026881. doi: 10.1001/jamanetworkopen.2020.26881
105. Circle of Blue. *Infographic: America's Septic Systems.* (2015). Available at: <https://www.circleofblue.org/2015/world/infographic-americas-septic-systems/>.
106. Kirby AE, Walters MS, Jennings WC, Fugitt R, LaCross N, Mattioli M, et al. Using wastewater surveillance data to support the COVID-19 response - United States, 2020–2021. *MMWR Morb Mortal Wkly Rep.* (2021) 70:1242–4. doi: 10.15585/mmwr.mm7036a2
107. Diamond MB, Keshaviah A, Bento AI, Conroy-Ben O, Driver EM, Ensor KB, et al. Wastewater surveillance of pathogens can inform public health responses. *Nat Med.* (2022) 28:1992–5. doi: 10.1038/s41591-022-01940-x
108. Mina MJ, Metcalf CJE, McDermott AB, Douek DC, Farrar J, Grenfell BT. A global immunological observatory to meet a time of pandemics. *eLife.* (2020) 9:9. doi: 10.7554/eLife.58989

109. Jacobs D, McDaniel T, Varsani A, Halden RU, Forrest S, Lee H. Wastewater monitoring raises privacy and ethical considerations. *IEEE Trans Technol Soc.* (2021) 2:116–21. doi: 10.1109/TTS.2021.3073886
110. Savi MK, Yadav A, Zhang W, Vembar N, Schroeder A, Balsari S, et al. A standardised differential privacy framework for epidemiological modelling with mobile phone data. *PLOS Digit Health.* (2023) 2:e0000233. doi: 10.1371/journal.pdig.0000233
111. Rader B, Gertz A, Iuliano AD, Gilmer M, Wronski L, Astley CM, et al. Use of at-home COVID-19 tests - United States, august 23, 2021-march 12, 2022. *MMWR Morb Mortal Wkly Rep.* (2022) 71:489–94. doi: 10.15585/mmwr.mm7113e1
112. Salomon JA, Reinhart A, Bilinski A, Chua EJ, La Motte-Kerr W, Ronn MM, et al. The US COVID-19 trends and impact survey: continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination. *Proc Natl Acad Sci USA.* (2021) 118:4118. doi: 10.1073/pnas.2111454118
113. Harruff RC, De Silva M, Lubin M. Medical examiner surveillance of the COVID-19 pandemic in King County, Washington, USA. *Am J Forensic Med Pathol.* (2023) 44:157–65. doi: 10.1097/PAF.0000000000000839
114. Mwananyanda L, Gill CJ, MacLeod W, Kwenda G, Pieciak R, Mupila Z, et al. Covid-19 deaths in Africa: prospective systematic postmortem surveillance study. *BMJ.* (2021) 372:n334. doi: 10.1136/bmj.n334
115. Lu FS, Nguyen AT, Link NB, Molina M, Davis JT, Chinazzi M, et al. Estimating the cumulative incidence of COVID-19 in the United States using influenza surveillance, virologic testing, and mortality data: four complementary approaches. *PLoS Comput Biol.* (2021) 17:e1008994. doi: 10.1371/journal.pcbi.1008994



OPEN ACCESS

EDITED BY

Jie Huang,
Southern University of Science and
Technology, China

REVIEWED BY

Tin Phan,
Los Alamos National Laboratory (DOE),
United States
Robert Michael McKay,
University of Windsor, Canada
Ayse Seval Palteki,
Istanbul Medipol University, Türkiye

*CORRESPONDENCE

Davey L. Jones
✉ d.jones@bangor.ac.uk

RECEIVED 28 January 2024

ACCEPTED 20 June 2024

PUBLISHED 15 July 2024

CITATION

Jones DL, Bridgman M, Pellett C,
Weightman AJ, Kille P, García Delgado Á,
Cross G, Cobley S, Howard-Jones H,
Chadwick DR and Farkas K (2024) Use of
wastewater from passenger ships to assess
the movement of COVID-19 and other
pathogenic viruses across maritime
international boundaries.
Front. Public Health 12:1377996.
doi: 10.3389/fpubh.2024.1377996

COPYRIGHT

© 2024 Jones, Bridgman, Pellett, Weightman,
Kille, García Delgado, Cross, Cobley,
Howard-Jones, Chadwick and Farkas. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Use of wastewater from passenger ships to assess the movement of COVID-19 and other pathogenic viruses across maritime international boundaries

Davey L. Jones^{1*}, Mathew Bridgman¹, Cameron Pellett¹,
Andrew J. Weightman², Peter Kille², Álvaro García Delgado¹,
Gareth Cross³, Steve Cobley³, Helen Howard-Jones¹,
David R. Chadwick¹ and Kata Farkas¹

¹School of Environmental and Natural Sciences, Bangor University, Bangor, United Kingdom,

²Microbiomes, Microbes and Informatics Group, School of Biosciences, Cardiff University, Cardiff, United Kingdom, ³Science Evidence Advice Division, Health and Social Services Group, Welsh Government, Cardiff, United Kingdom

Objective: The worldwide spread of SARS-CoV-2 and the resulting COVID-19 pandemic has been driven by international travel. This has led to the desire to develop surveillance approaches which can estimate the rate of import of pathogenic organisms across international borders. The aim of this study was to investigate the use of wastewater-based approaches for the surveillance of viral pathogens on commercial short-haul (3.5 h transit time) roll-on/roll-off passenger/freight ferries operating between the UK and the Republic of Ireland.

Methods: Samples of toilet-derived wastewater (blackwater) were collected from two commercial ships over a 4-week period and analysed for SARS-CoV-2, influenza, enterovirus, norovirus, the faecal-marker virus crAssphage and a range of physical and chemical indicators of wastewater quality.

Results: A small proportion of the wastewater samples were positive for SARS-CoV-2 (8% of the total), consistent with theoretical predictions of detection frequency (4%–15% of the total) based on the national COVID-19 Infection Survey and defecation behaviour. In addition, norovirus was detected in wastewater at low frequency. No influenza A/B viruses, enterovirus or enterovirus D68 were detected throughout the study period.

Conclusion: We conclude that testing of wastewater from ships that cross international maritime boundaries may provide a cost-effective and relatively unbiased method to estimate the flow of infected individuals between countries. The approach is also readily applicable for the surveillance of other disease-causing agents.

KEYWORDS

wastewater-based epidemiology, international sea travel, border crossing, passenger ferry, AMR, SARS-CoV-2 infection, import rate, public health surveillance

1 Introduction

It is well established that effective surveillance and a timely response are essential to limit the social, health and economic impacts of rapidly spreading diseases, such as COVID-19 (1,

2). Wastewater-based epidemiology (WBE), which measures viral markers shed by infected individuals in faeces and urine, has been successfully used for surveillance of infectious diseases at a population level, including the multi-national surveillance of SARS-CoV-2 and poliovirus (3–6). Monitoring levels of SARS-CoV-2 in wastewater has, thus, provided an effective tool and early warning system to aid in public health decision-making and tracking the success of policy interventions (7–9).

Air and maritime travel represent key factors which have facilitated the global spread of SARS-CoV-2 and other viral diseases (10–12). International shipping is of particular interest due to the large volume of potentially infected passengers which may enter the country from overseas (>20 million year⁻¹ in the UK) (13). The dense aggregation of people within port and dockyard areas may also facilitate infection between individuals (14, 15). Further, industrial ships and changes of crew and movement of goods in areas with multiple working personnel has the potential to cause outbreak on ships and within ports (16–20). These transmission events then have the potential to spread through the local community and to additional international ports. In a long-distance shipping context, an outbreak of SARS-CoV-2 poses serious risks to crew as they may lack the medical infrastructure or are unprepared to capably deal with issues should they arise (18, 21, 22). This also represents an issue for tourism-based cruise ships where viral (e.g., norovirus) outbreaks have regularly led to the quarantining of vessels (23, 24). A recent modelling study has also confirmed that international seaports are likely to represent a significant risk to the spread of SARS-CoV-2 (14).

Even though considerable concerns have been raised about COVID-19 transfer associated with long-haul shipping and cruise ships (7, 14), limited surveillance has been undertaken on short-haul, mass-transport passenger and freight ships. These short-haul international routes, however, may pose a greater risk for pathogen entry in comparison to longer-haul routes due to (i) the greater number of passengers involved, (ii) a lack of point-of departure/entry testing procedures, (iii) no on-board testing, (iv) less rigorous ship sanitation/cleaning, (v) the presence of pre-symptomatic passengers who travel not knowing they carry the virus, and (vi) the inability of conventional testing to capture infections (e.g., lateral flow devices) in comparison to cruise ships which rely more on PCR-based testing. Whilst wastewater testing has been deployed at international airports to evaluate the frequency of entry of infected individuals (25–27), this approach has yet to be critically tested on ships or at international ferry ports. The success of the approach, however, relies on a range of factors including the toilet behaviour of individuals, particularly on short-haul crossings, access to wastewater on the vessels and the subsequent capacity to quantify and sequence viral RNA/DNA in the samples.

Depending on the age and nature of the ship, on-board sanitation systems can vary significantly between vessels (28). In some situations, the black- and grey-water streams are kept separate, which is ideal for WBE, whilst in others they are mixed, leading to dilution of the viral signal. In other cases, sewage is collected on the boat and then delivered to a port reception facility for subsequent treatment (29). Access to sewage collection tanks may also be problematic on some vessels whilst addition of different sanitation agents (e.g., disinfectants) may cause issues in viral recovery. Conversely, the presence of low-water use vacuum toilets can be expected to result in more concentrated sewage in comparison to on-shore municipal sewage which may aid viral detection (30).

Due to the current paucity of information, the aim of this pilot study was to critically evaluate whether wastewater-based surveillance on short-haul international passenger/freight ships is viable for monitoring the frequency of entry of SARS-CoV-2, alongside other pathogens (e.g., norovirus, influenza-A and B, enterovirus). The study focused on the main UK to Republic of Ireland passenger route, monitoring wastewater on two of the main commercial vessels over a one-month period. The work focused on the practicality, economic viability and usefulness of the approach within the framework of a potential national border surveillance programme for pathogens of public health concern.

2 Materials and methods

2.1 Sampling locations

The project was based on the maritime route between the Holyhead Ferry Terminal located in Gwynedd, Wales, United Kingdom (53°18′58.47″N, 04°37′24.47″W) and Dublin Port located in Dublin, Ireland (53°20′57.13″N, 06°11′50.70″W). The route represents the main maritime freight and passenger link between the UK and Ireland with an estimated 1.9 million passengers per year and *ca.* 450,360 cargo truck transfers (31). The route is *ca.* 80 km from port-to-port and takes *ca.* 3 h 15 min per crossing and is serviced by several commercial companies (Supplementary Figure S1). This study focused on two superferries, namely the Stena Estrid and the Stena Adventurer (Stena AB, Gothenburg, Sweden; Supplementary Figure S2).

The Stena Estrid was built in 2019 by AVIC Weihai, Shandong Province, China and is classified as an ‘E-Flexer’ passenger roll-on/roll-off cargo (Ro-Pax) ferry. It has a capacity of 1,000 passengers, 120 cars and 210 freight vehicles. The Stena Adventurer was built in 2003 by Hyundai Heavy Industries, South Korea and is also a Ro-Pax ferry with a capacity of 1,500 passengers and 500 cars and freight vehicles.

The ships possess different wastewater management systems and thus the sampling strategy varied slightly between ships. The Stena Estrid wastewater system is separated into 2 initial chambers: (i) blackwater (raw sewage from toilets), (ii) greywater (water from sinks, showers, and kitchen appliances). These are then combined in a mixing chamber and then transferred to a screening tank to remove large non-biodegradable solids. Once mixing had occurred, wastewater is transferred to an Evac Membrane Bioreactor treatment module (Evac Oy, Espoo, Finland). Post aerobic treatment, clean water is then discharged at sea whilst the solid waste becomes a dry powder that is offloaded at shore for disposal. Samples were initially planned to be taken from blackwater chamber, however, due to access/system constraints, samples had to be taken from the screening tank, but prior to any treatment occurring (Supplementary Figure S3). The Adventurer has an older wastewater system containing of 3 chambers involving maceration (soaking), chopping and mixing. After being mixed, the wastewater is moved to a similar treatment plant to the Stena Estrid where it is aerobically treated and filtered in a containment tank where it is stored until it reaches port and then taken to a wastewater treatment plant. Samples on the Adventurer were taken prior to the anaerobic treatment stage.

2.2 Sample collection

Wastewater sampling was undertaken on Sunday, Tuesday, and Thursday on each ship from the 27th January 2022 to the 23rd February 2022. On each day, 4 independent samples were taken representing the 4 single leg journeys between Holyhead and Dublin each day (Supplementary Table S1). The samples (500 mL) were collected by the engineering crew, placed within polycarbonate bottles and refrigerated at 4°C on the ship prior to collection from the port each day. Samples were collected directly from Holyhead port and then transported to the laboratory (40 km distance) in a refrigerated box where the samples were then stored at 4°C and analysed within 24 h of collection. Basic training was provided to the ship's staff for sample collection.

2.3 Viral concentration, nucleic acid extraction, and quantification

Viral recovery and purification were undertaken according to the polyethylene glycol (PEG)-salt precipitation of Farkas et al. (32) and Kevill et al. (33). This method was chosen as it is used in the Welsh Government national wastewater COVID-19 surveillance programme. Briefly, 200 mL of each wastewater sample was placed in a sterile polypropylene centrifuge bottle and centrifuged (10,000 g, 10 min, 4°C) to remove suspended solids. 150 mL of the clarified supernatant was then transferred to a sterile polypropylene centrifuge bottle, the pH adjusted to 7.0–7.5 and 50 mL of a PEG-8000-NaCl solution added to reach a final PEG-8000 concentration of 10% and NaCl content of 2%. An aliquot of dsRNA *Pseudomonas* phage Phi6 was then added to the sample as an extraction control and the samples incubated at 4°C overnight. Post-incubation, the samples were centrifuged (10,000 g, 30 min, 4°C). The supernatant was then discarded and the pellet resuspended in 850 µL of Nuclisens lysis buffer (BioMérieux, France). The viral RNA and DNA from the resuspended pellet was then extracted using a KingFisher 96 Flex automated purification system (Thermo Scientific, Waltham, United States) using Nuclisens extraction reagents (BioMérieux, France) as described elsewhere (33). The final volume of the RNA/DNA eluent was 100 µL.

One-step RT-qPCR for the SARS-CoV-2 N1 gene region and Phi6 targets was performed using an TaqMan™ Fast Virus 1-Step Master Mix (Applied Biosystems Inc., United States), on a Quant Studio Flex 6 (Applied Biosystems Inc., United States) using previously published primers and probes (34, 35) (Supplementary Table S2). The mastermix contained 10 pmol of the forward, 20 pmol of the reverse primers and 5 pmol probe, 16 nmol MgSO₄, 1 µg bovine serum albumin (BSA), molecular grade water and 4 µL sample/standard/control in 20 µL reaction mix. RT-qPCR settings were: Hold step 50°C 30 min for reverse transcription, 95°C 20 s for reverse transcriptase inactivation, followed by 45 amplification cycles of 95°C 13 s, 60°C 45 s.

Multiplex RT-qPCR assays were used for the detection of influenza A/B viruses (flu A and flu B) and for Enteroviruses (EV), enterovirus D68 (EV-D68) and norovirus GII (NoVGII) using previously published primers and probes (36–38) (Supplementary Table S2). The same reaction conditions as for SARS-CoV-2 quantification were used except that the mixture contained no added MgSO₄.

For crAssphage an established assay using the QuantiFast qPCR mix was used (33) with 2 µL sample added to 20 µL reaction mix.

All samples were run in duplicate, against a dilution series (1–10⁵ copies µl⁻¹ per reaction) of in house developed ssRNA standards for SARS-CoV-2 and phi6 (33), commercial ssRNA standards for flu A/B and EV-D68 (Twist Bioscience, United States) or plasmid DNA for NoVGII and crAssphage (39, 40). PCR no template controls (molecular-grade water) determined the absence of contamination during the PCR set-up.

2.4 SARS-CoV-2 sequencing

Selected RNA extracts were further purified with Mag-Bind® TotalPure NGS beads (Omega Bio-Tek) to remove potential inhibitors prior to reverse transcription into cDNA with LunaScript® RT SuperMix (NEB) prior to SARS-CoV-2 amplification and sample indexing using EasySeq™ SARS-CoV-2 kit (Nimagen). The protocol used has been customised previously for use with wastewater (41). Amplified products were quantified and quality controlled using Agilent TapeStation. Libraries were sequenced on an Illumina MiniSeq benchtop sequencer, producing 2 × 150-bp paired-end reads. Raw reads were processed using the ncov2019-artic-nf Nextflow pipeline (42). Briefly, reads were trimmed using Cutadapt v1.18 (43) and Nimagen V4 primer sequences were removed using iVar v1.3. Cleaned reads were aligned to the SARS-CoV-2 reference genome Wuhan-Hu-1 (MN908947.3) (44) using the Burrow-Wheeler Aligner (BWA) (45) and ca. 400,000 reads mapped per sample. Lineage abundances were then determined using the processed sequences using depth-weighted de-mixing of SNV frequency at each position in the genome using Freyja pipeline (46, 47).

2.5 Wastewater physical and chemical analysis

The samples were analysed for a range of key physicochemical markers of wastewater quality including pH, turbidity, electrical conductivity (EC), ammonium and orthophosphate (9). Turbidity was assessed using an Orion AQUAfast AQ3010 turbidity metre (Thermo Scientific, Waltham, MA, United States) whilst EC was measured using a Jenway 4520 conductivity metre and pH with a Hanna 209 pH probe (Hanna Instruments Ltd., Leighton Buzzard, United Kingdom). For NH₄⁺ and P analysis, the samples were first centrifuged (24,000 g, 5 min) to remove suspended solids. The supernatant was then retained for subsequent analysis. Inorganic P was measured colorimetrically using the molybdate blue reagent according to Murphy and Riley (48) whilst NH₄⁺ was determined colorimetrically using the salicylate procedure of (49) using a SpectroStar Nano microplate reader.

2.6 Data analysis

The qPCR quality control was carried out with QuantStudio real-time PCR software v1.7 (Applied Biosystems, Inc., United States). The standard curve slope, efficiency and R² met the requirements described in Bustin et al. (50). The qRT-PCR data was converted to gc l⁻¹ wastewater for statistical analysis. The assay limit of detection (LOD) and limit of quantification (LOQ) were tested using 10 replicates of low dilutions of genomic RNA for the RNA virus targets and plasmid DNA for crAssphage (40). The LOD was defined as the minimum concentration

whereby 10 replicates all return positive results and the LOQ was the lowest concentration where the coefficient of variation was lower than 0.25 (Supplementary Table S2). As such, quantities can be detected below this limit but are susceptible to false negatives. For comparison, the wastewater composition from the ships was directly compared with that collected as part of the national surveillance programme undertaken in Wales. The latter involved the analysis of wastewater collected from 44 centralised wastewater treatment plants across Wales 5 days a week.

To theoretically estimate the number of a- and pre-symptomatic passengers who were travelling on the transnational shipping route (i.e., import rate, *IR*) we used the following equation:

$$IR = PN \times PP \times ACR \times FSR \times TU \quad (1)$$

where *PN* is the total number of passengers sampled during the wastewater testing campaign ($n=6,942$), *PP* is the prevalence of COVID-19 in the population (3.1%–4.1% of the population), *ACR* is the amount of COVID-19 cases that are pre- or a-symptomatic (20%–30% of the total), *FSR* is the shedding frequency of SARS-CoV-2 in faeces (40%–60% of cases), and *TU* is the likelihood that passengers will use a toilet whilst on board the ship (13%). It was assumed that symptomatic passengers would not be travelling due to government travel restrictions in place when the study was undertaken.

3 Results

3.1 Prevalence of COVID-19 cases during the survey period

Wastewater sampling commenced towards the end of the third main COVID-19 wave in the UK which was associated with the emergence of the omicron variant of SARS-CoV-2. During this sampling period 0.1% to 0.2% of the UK and Irish population tested positive for SARS-CoV-2 (51). Overall, the patterns in COVID-19 cases were similar between countries. Based on the results of the COVID-19 Infection Survey (CIS), which is less prone to self-reporting bias, it is likely that the true prevalence of COVID-19 in the UK and Ireland populations ranged from 3.1% to 4.5% during the study period (51–54). At the time that the wastewater monitoring was undertaken, the wearing of face coverings was still mandatory and recommendations were in place for individuals not to travel if they had tested positive for SARS-CoV-2. Stena line staff were also asked to self-isolate if they tested positive for COVID-19. At the time of the study, passenger locator forms were not required to enter the UK and no quarantining procedures were in place. Due to the COVID-19 pandemic, the number of passengers per journey was lower than normal with each journey having an average of 154 passengers (range 38 to 612) on the Stena Estrid and 169 on the Stena Adventurer (range 28 to 775). Of these, 74% were crossing with cars or as foot passengers and 26% as commercial freight drivers. There were no differences in the passenger:freight ratio between the two ships.

3.2 SARS-CoV-2 and other viruses in ferry wastewater

SARS-CoV-2 was detected in four samples during the survey period (8.1% of the total samples, $n=49$, Figure 1). Of the positive wastewater

samples, the maximum concentration detected was 9.2×10^5 gc l⁻¹. Of the other human pathogenic viruses tested in the wastewater samples, only NoV GII was detected, albeit at a lower frequency (6.1% of the total samples) with a maximum concentration of 1.3×10^6 gc l⁻¹. Neither, enterovirus, enterovirus D68 or influenza A or B were detected in the samples. The faecal marker crAssphage was detected in all samples from the Stena Adventurer, however, recovery of crAssphage from the Stena Estrid was much lower (26% of the total samples). The mean recovery of crAssphage was 1.9×10^6 gc l⁻¹ on the Stena Adventurer which was lower than from the Stena Adventurer when samples tested positive (2.1×10^7 gc l⁻¹, $p=0.002$). Overall, the levels of crAssphage were lower than those reported in the national urban wastewater surveillance programme (mean $1.0 \times 10^9 \pm 3.0 \times 10^7$ gc l⁻¹; $p<0.001$).

3.3 SARS-CoV-2 sequencing

The samples that tested positive for the SARS-CoV-2 N1 gene region by RT-qPCR were subsequently sequenced. Sequence was acquired for 600–362,000 reads of which between 60% and 82% of the mapped to the viral genome. Although this yielded an average coverage > 1,500, sequences mapped to very restricted regions of the virus and therefore provided incomplete coverage for all samples. Overall, the percentage genome covered ranged from 18% to 35%. Consequently, we were able to ascribe one sample to the SARS-CoV-2 omicron variant, however, the other three positive samples remained unascribed. The success of sequencing appeared directly related to the amount of SARS-CoV-2 recovered in the sample.

3.4 Wastewater chemistry

The average orthophosphate concentration of wastewater on the two Stena ships (mean $\pm$ SEM, 211 ± 57 mg l⁻¹) was considerably higher than samples collected during the Welsh government national surveillance project (2.6 ± 0.1 mg l⁻¹; Figure 2A). Likewise, we found the median ammonium concentration of wastewater on the ships (320 ± 25 mg N l⁻¹) to be much higher than the national surveillance median (16 ± 1 mg N l⁻¹; Figure 2D). Further, the turbidity of the ships' wastewater samples ($1,172 \pm 122$ NTU) was higher than reported for urban wastewater in the national surveillance programme (90 ± 5 NTU). Similarly, the electrical conductivity and pH of the ships' wastewater (4.7 ± 0.2 mS cm⁻¹ and 7.9 ± 0.12 , respectively) were also different to the national surveillance programme samples (0.9 ± 0.1 mS cm⁻¹ and 7.5 ± 0.02 , Figures 2B,C). None of the wastewater characteristics had significant correlations with passenger data ($p>0.05$; data not presented).

4 Discussion

4.1 Potential of ship wastewater to capture the presence of infected individuals

Wastewater can potentially provide a non-invasive, ethically compliant and relatively unbiased way to evaluate levels of infection within a cohort of individuals all connected to a common sanitary system (5). To our knowledge, this is the first use of ship-based wastewater-based surveillance to assess the potential transfer of viral

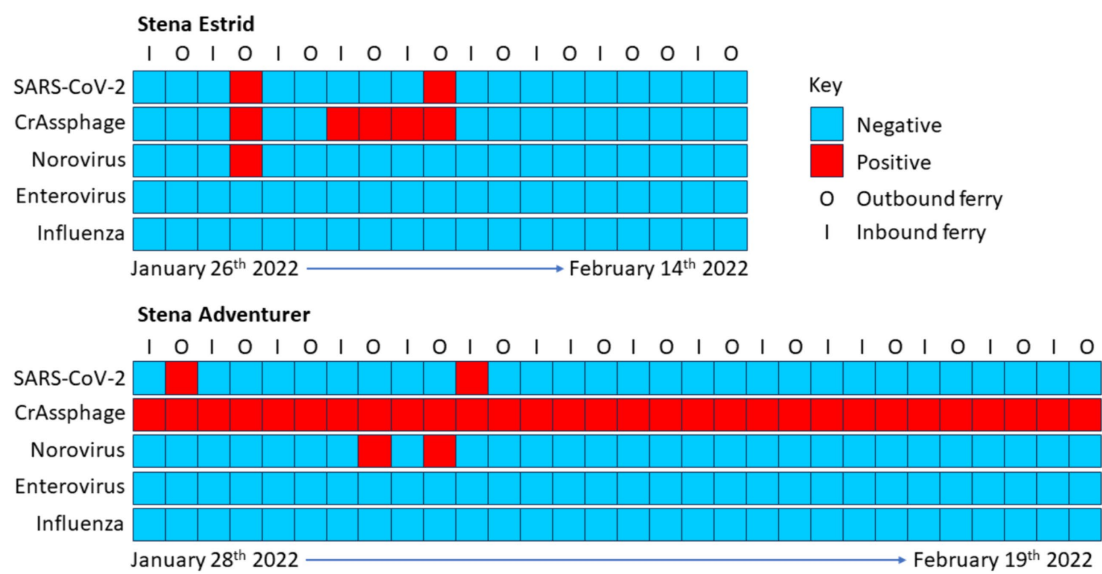


FIGURE 1 Viral detection and quantification in wastewater collected from two international short-haul ships (Stena Estrid and Stena Adventurer) taking passengers and commercial loads to and from Ireland (Dublin) and Wales (Holyhead). The outbound route is Holyhead to Dublin and the inbound route is Dublin to Holyhead. SARS-CoV-2 used the N1 gene target. The faecal-marker virus crAssphage was used as an indicator of faecal matter being present. Norovirus represents genogroup II and Influenza represents both influenza A and B. Each square represents an individual ferry crossing between Ireland and the Wales.

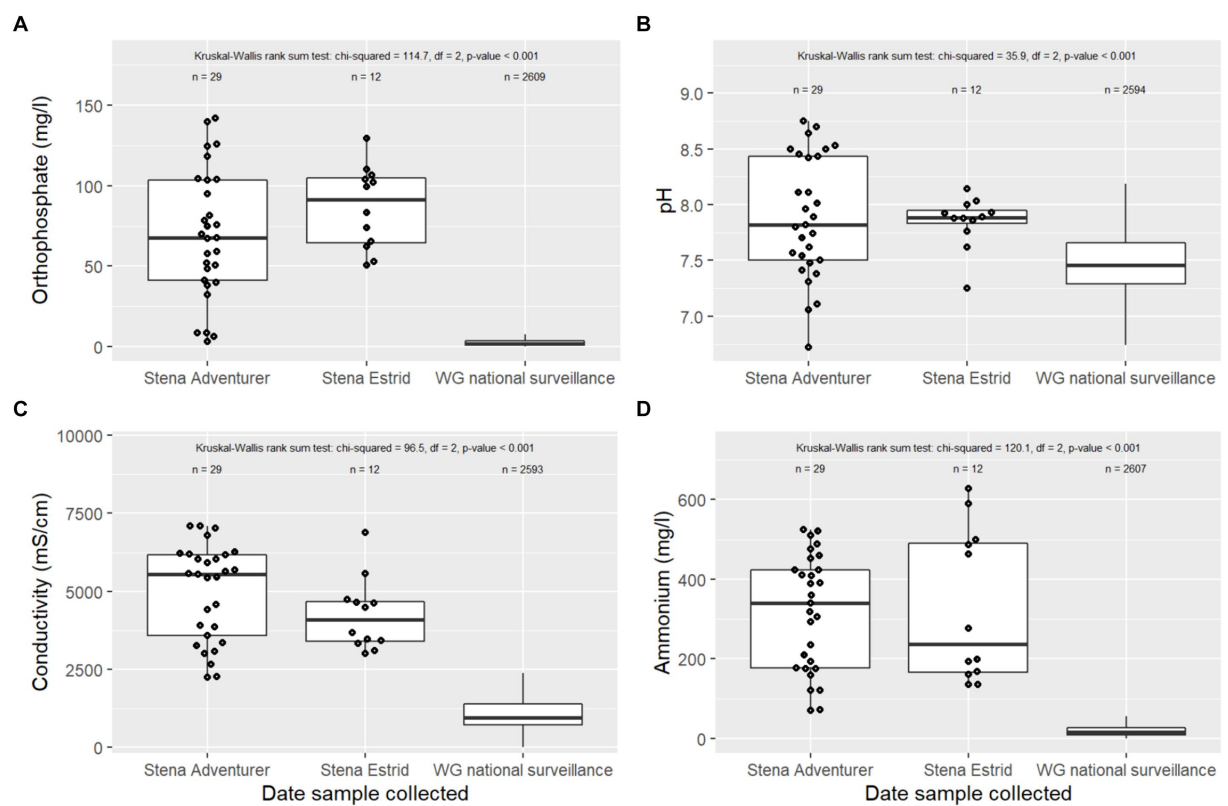


FIGURE 2 Chemical indicators of wastewater quality from two international short-haul ships (Stena Adventurer and Stena Estrid) on the international Ireland (Dublin) to Wales (Holyhead) route. (A) orthophosphate, (B) pH, (C) electrical conductivity, and (D) ammonium. For comparison, we present results for influent wastewater from 44 sites collected as part of the Welsh national COVID-19 wastewater surveillance network. The 25th, 50th, and 75th percentile ranges are depicted by the box, excluding outliers greater or lesser than 1.5 × IQR depicted by the whiskers.

pathogens across an international maritime boundary. Our results provided clear evidence that, albeit infrequent, infected individuals were crossing between the UK and Ireland during the third COVID-19 wave when guidance was still in place to prevent travel for infected individuals. Whilst wastewater analysis has previously been undertaken on ships, this has largely been from the perspective of discharging pollutants into marine waters rather than assessing the presence of infected individuals on a vessel (55–58). Further, most of this work has focused on organic pollutants (e.g., antibiotics) and faecal-derived bacteria rather than on viruses (30, 55, 59). A single study from a cruise ship carrying passengers infected with SARS-CoV-2 showed previously that viral RNA could be isolated from the ship's wastewater (60), providing the first evidence that wastewater can be used for on-board pathogen surveillance. However, long-haul cruise ships hold an isolated population where it can be guaranteed that all individuals will use the toilet facilities. Further, cruise ships are notorious for large viral outbreaks due to the close confinement of passengers over long periods of time (e.g., norovirus, influenza) (61–63). This suggests that viral titers in wastewater from cruise ships are likely to be very high and may also prove useful as a temporal indicator of outbreak progression.

In the case of short- and medium-haul passenger ferries (journey time < 6 h in duration), the frequency that individuals defecate remains unknown; however, it is expected that this will be very low in comparison to cruise ships. The continual changing of passengers (4 times daily in this study), is also likely to lead to more temporally stochastic results with lower viral titers (due to a higher urination-to-defecation ratio). Unlike cruise ships, in the context of short-haul shipping routes, it is the frequency of detection that is most important rather than the quantitative analysis of the amount of viral RNA present.

4.2 Theoretical vs. actual measured incidence of infected individuals

The success of viral surveillance using wastewater relies largely on faecal shedding and to a lesser extent vomiting and sputum, whilst very few pathogenic viruses are shed in urine (64). Previous studies have indicated that enteric and respiratory viruses are shed in faeces whether individuals are asymptomatic or symptomatic (65–67). The frequency, duration and amount of faecal shedding, however, can vary significantly between viruses, point in the infection cycle and on the nature of the individual (e.g., age, immune status etc.). Here we take a first principles approach to estimating the likely number of passengers infected with SARS-CoV-2 who can theoretically be captured using a wastewater-based approach. Although information exists for defecation frequency on long-haul ships, which suggests that most people defecate less often than on land (68), no quantitative information exists for defecation frequency on short-haul passenger ferries. Based on estimates of likely frequency of on-board defecation on short-haul flights (<13%; <3 h in duration) (69), we use this to estimate the chances of capturing infected individuals on short-haul passenger ships. Based on the total number of passengers sampled during the study period (*ca.* 6,942), a population-level COVID-19 prevalence rate of 3.1%–4.5% (51, 52), an asymptomatic carriage rate of the omicron variant of 20%–30% (70, 71) and a SARS-CoV-2 faecal shedding rate of 40%–60% (66), we estimate that theoretically the

number of infected passengers would range from 2.2 to 7.3 (Equation 1). The number of actual samples which tested positive for SARS-CoV-2 ($n=4$) directly falls within this range. One assumption we have made is that symptomatic people did not travel based on government guidance at the time of the study and that diarrhoea is not a primary symptom of omicron infections, the dominant variant in circulation at the time (72).

4.3 Use of wastewater for the surveillance of other viral pathogens

Although the main premise of this study was to evaluate the use of wastewater for COVID-19 border surveillance, we showed that the approach can also be used to evaluate the prevalence and movement of other viruses and is likely suitable for other disease-causing agents (e.g., anti-microbial resistant bacteria, protozoa). Here we also detected the RNA of norovirus in wastewater on several occasions. Indeed, wastewater may be better for the surveillance of enteric viruses as the frequency and volume of defecation is much greater (e.g., diarrhoea), viral shedding rates occurs in all infected individuals and the rates of shedding are much greater (66). Enteric viruses also represent the leading cause of illness amongst returning travellers seeking medical care (73). Previous estimates of trans-border movement of norovirus have relied on the analysis of serum or stool samples, largely provided voluntarily from symptomatic individuals (73–76). In combination with genotyping (to assess unique lineages), wastewater could provide an unbiased assessment of norovirus entry into the country, particularly as *ca.* 10% of infections are asymptomatic and shed at similar rates to symptomatic individuals (77). The levels of norovirus circulating in the population at the time of the study were atypically low due to the COVID-19 pandemic (78, 79) suggesting that more cases may be detected post-pandemic. Similarly, the prevalence of influenza A/B and enterovirus were also unseasonably low in the population at the time of sampling, due to the knock-on effect of non-pharmaceutical interventions for COVID-19 control (80, 81). It would therefore be useful to undertake a repeat survey under more representative circumstances to evaluate the use of wastewater for catching these viruses.

4.4 Limitations of using a wastewater-based approach for pathogen surveillance on ships

Whilst wastewater analysis proved successful at showing the passage of infected individuals between the UK and Ireland, the approach has some limitations and areas for refinement as follows: (i) *Sampling approach:* For logistical reasons, we relied on taking several manual spot measurements per journey rather than deploying an automated time-integrated composite sampler. Although some mixing of the wastewater will occur within the sanitary network, it is known that a grab/spot-sampling approach does not provide the most reliable estimate of viral load, particularly for near-source testing (82). The design of a refrigerated autosampler that can retrieve a wastewater sample from a pressurised sanitary network at regular intervals (*ca.* every 10 min) would therefore be useful. Further, passive sampling approaches may be appropriate to capture time integrated information

without having to rely on complex autosamplers (83); (ii) *Independent validation*: To better validate the wastewater approach, it would be useful to take nasopharyngeal swabs from a representative sample of individuals to confirm the presence/absence of SARS-CoV-2 and influenza (63). Due to ethical and social considerations, validation for enteroviruses may be more problematic; (iii) *Defecation behaviour*: As the approach relies on shedding viruses in faeces, it would be useful to gain insight into the toilet habits of individuals and whether these are influenced by demographic factors (e.g., age, gender, nationality), passenger type (e.g., commercial truck drivers vs. tourists, journey details), timing (e.g., day vs. night voyages), season (e.g., tourist season vs. off-peak) and the health status (e.g., evidence of respiratory or gastrointestinal symptoms). This could be achieved by eliciting a passenger questionnaire on departure from the port. Alternatively, the number of individuals defecating on the boat could be assessed by the unique lineages of phages present in the human gut (e.g., crAssphage) (84). The toilet use by crew should also be a factor that needs to be considered in this analysis; (iv) *Wastewater transit time*: Although the samples were taken on a daily basis, the residence time of the wastewater in the sanitary network (e.g., holding tanks) (30), and therefore the potential loss of viral RNA/DNA remains unknown. Based on previous studies on marine wastewater discharges, we therefore recommend the deployment of a rhodamine tracer for mapping residence time (85); (v) *Origin of infection*: Due to the uncertainty in wastewater transit time, we were unable to determine with certainty whether the wastewater collected was from the UK-Ireland or Ireland-UK leg of the journey (or a mixture of both). The geographical origin of SARS-CoV-2 or norovirus in our samples could therefore not be determined with certainty. More complete genetic sequencing of the viral strains and mapping the lineages to national databases will clearly aid in this. Due to the high number of clinical samples being sequenced for SARS-CoV-2 this should be effective; (vi) *Viral recovery*: A preliminary investigation in a small number of samples showed that variations on the PEG-salt based method used here may give better viral recoveries. Given the concentrated nature and high urea content of ship blackwater (86), it is likely that improved methods for viral recovery and removal of PCR inhibitors is still needed. This is evidenced by the inability to recover crAssphage from some samples, despite its high abundance in human faeces from industrialised countries (87). Given the high solids content in the wastewater, it may also be desirable to evaluate the partitioning of viruses between the solid and liquid fraction so that the most enriched fraction can be targeted for further surveillance activities; (vii) *Other shipping routes*: This study targeted short-haul journeys, however, adopting a similar approach on longer maritime crossings would provide additional value and may be less affected by some of the limitations highlighted above. For example, the UK-Spain passenger ferry (Portsmouth-Santander) has a duration of 28.5 h, whilst the UK-Belgium route (Hull-Zeebrugge) takes 13.5 h and the UK-Norway (Harwich-to-Esbjerg) passage takes 18 h.

5 Conclusion

This study has successfully demonstrated that ship blackwater can be used to isolate and identify viruses of public health concern. Further, the frequency of detection was consistent with theoretical estimates based on known infection rates within the population.

Although some refinement of the methodology is still required, we conclude that this wastewater-based approach can be readily expanded to a wide range of faecal-borne pathogens. In combination, the methodology presented here provides a non-invasive way to assessing the frequency of pathogen transfer across international maritime boundaries and thus the contribution of maritime traffic to the global spread of disease.

Data availability statement

The data presented in the study are deposited in the Zenodo repository, accession number <https://doi.org/10.5281/zenodo.1267108>.

Author contributions

DJ: Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. MB: Investigation, Writing – review & editing. CP: Data curation, Formal analysis, Software, Writing – review & editing. AW: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. PK: Data curation, Formal analysis, Funding acquisition, Methodology, Supervision, Writing – review & editing. ÁGD: Data curation, Investigation, Methodology, Writing – review & editing. GC: Conceptualization, Funding acquisition, Validation, Writing – review & editing. SC: Conceptualization, Funding acquisition, Resources, Writing – review & editing. HH-J: Investigation, Methodology, Resources, Writing – review & editing. DC: Funding acquisition, Resources, Supervision, Writing – review & editing. KF: Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. DJ, PK, and AW were supported by a grant from Welsh Government under the Welsh Wastewater Programme (C035/2021/2022) and the UKRI Natural Environment Research Council grant NE/V010441/1.

Acknowledgments

The authors thank the Welsh Government Technical Advisory Group for feedback on the draft manuscript. We thank all the staff from Stena Line, especially the chief engineers on the two vessels (James Wilkinson and Robert Adams) and the associated crew.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2024.1377996/full#supplementary-material>

References

- Akbulaev N, Mammadov I, Aliyev V. Economic impact of COVID-19. *Sylwan*. (2020) 164:813. doi: 10.2139/ssrn.3649813
- Kaye AD, Okeagu CN, Pham AD, Silva RA, Hurley JJ, Arron BL, et al. Economic impact of COVID-19 pandemic on healthcare facilities and systems: international perspectives. *Best Pract Res Clin Anaesthesiol*. (2021) 35:293–306. doi: 10.1016/j.bpa.2020.11.009
- Tambe LAM, Mathobo P, Matume ND, Munzhedzi M, Edokpayi JN, Viraragavan A, et al. Molecular epidemiology of SARS-CoV-2 in northern South Africa: wastewater surveillance from January 2021 to May 2022. *Front Public Health*. (2023) 11:1309869. doi: 10.3389/fpubh.2023.1309869
- Tlhagale M, Liphadzi S, Bhagwan J, Naidoo V, Jonas K, van Vuuren L, et al. Establishment of local wastewater-based surveillance programmes in response to the spread and infection of COVID-19—case studies from South Africa, the Netherlands, Turkey and England. *J Water Health*. (2022) 20:287–99. doi: 10.2166/wh.2022.185
- Wade MJ, Jacomo AL, Armenise E, Brown MR, Bunce JT, Cameron GJ, et al. Understanding and managing uncertainty and variability for wastewater monitoring beyond the pandemic: lessons learned from the United Kingdom national COVID-19 surveillance programmes. *J Hazard Mater*. (2022) 424:127456. doi: 10.1016/j.jhazmat.2021.127456
- Xagorarakis I. Can we predict viral outbreaks using wastewater surveillance? *J Environ Eng*. (2020) 146:01820003. doi: 10.1061/(ASCE)EE.1943-7870.0001831
- Ahmed W, Angel N, Edson J, Bibby K, Bivins A, O'Brien JW, et al. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: a proof of concept for the wastewater surveillance of COVID-19 in the community. *Sci Total Environ*. (2020a) 728:138764. doi: 10.1016/j.scitotenv.2020.138764
- Ai Y, Davis A, Jones D, Lemeshow S, Tu H, He F, et al. Wastewater SARS-CoV-2 monitoring as a community-level COVID-19 trend tracker and variants in Ohio, United States. *Sci Total Environ*. (2021) 801:149757. doi: 10.1016/j.scitotenv.2021.149757
- Hillary LS, Farkas K, Maher KH, Lucaci A, Thorpe J, Distaso MA, et al. Monitoring SARS-CoV-2 in municipal wastewater to evaluate the success of lockdown measures for controlling COVID-19 in the UK. *Water Res*. (2021) 200:117214. doi: 10.1016/j.watres.2021.117214
- Browne A, Ahmad SS, Beck CR, Nguyen-Van-Tam JS. The roles of transportation and transportation hubs in the propagation of influenza and coronaviruses: a systematic review. *J Travel Med*. (2016) 23:tav002. doi: 10.1093/jtm/tav002
- Xing Y, Wong GWK, Ni W, Hu X, Xing Q. Rapid response to an outbreak in Qingdao, China. *N Engl J Med*. (2020) 383:e129. doi: 10.1056/nejmc2032361
- Zhang L, Yang H, Wang K, Zhan Y, Bian L. Measuring imported case risk of COVID-19 from inbound international flights—a case study on China. *J Air Transp Manag*. (2020) 89:101918. doi: 10.1016/j.jairtraman.2020.101918
- DfT. Sea passenger statistics: All routes. London, UK: Department for Transport, UK Government (2019).
- Wang Z, Meng C, Yao M, Claramunt C. Modelling the risk of imported COVID-19 infections at maritime ports based on the mobility of international-going ships. *ISPRS Int J Geo Inf*. (2022) 11:60. doi: 10.3390/ijgi11010060
- Zhang N, Miao R, Huang H, Chan EY. Contact infection of infectious disease onboard a cruise ship. *Sci Rep*. (2016) 6:38790. doi: 10.1038/srep38790
- Bühler S, Busch P, Wittkamp P, Alpers K, Doerre A, Plenge-Bönig A, et al. A SARS-CoV-2 omicron outbreak among crew members on a cruise ship in Germany in early 2022. *Int Marit Health*. (2023) 74:235–42. doi: 10.5603/imh.96935
- Fanoy E, Ummels AE, Schokkenbroek V, van Dijk B, Wiegman S, Veenstra T, et al. Outbreak of COVID-19 on an industrial ship. *Int Marit Health*. (2021) 72:87–92. doi: 10.5603/imh.2021.0016
- Marimoutou C, Tufo D, Chaudet H, Abdul Samad M, Gentile G, Drancourt M. Infection burden among medical events onboard cargo ships: a four-year study. *J Travel Med*. (2017) 24:tax010. doi: 10.1093/jtm/tax010
- von Munster T, Kleine-Kampmann S, Hewelt E, Boldt M, Schlaich C, Strolin P, et al. When the ship comes in—COVID-19-Ausbrüche an Bord von Frachtschiffen im hamburger Hafen. *Tropenmedizin Reisemedizin*. (2021) 28:13–24. doi: 10.1055/a-1337-1831
- Yuan Q, Kou Z, Jiang F, Li Z, Zhang L, Liu H, et al. A nosocomial COVID-19 outbreak initiated by an infected dockworker at Qingdao City port—Shandong Province, China, 2020. *China CDC Wkly*. (2020) 2:838–40. doi: 10.46234/ccdcw2020.224
- Beust L, Lucas D, Pougnet R, Loddé B. COVID-19 on board a cruise ship: medical management. *Int Marit Health*. (2023) 74:83–8. doi: 10.5603/imh.2023.0012
- WHO. (2021). WHO public health checklist for controlling the spread of COVID-19 in ships, sea ports and inland ports. World Health Organization, WHO Regional Office for Europe, Copenhagen, Denmark.
- Rispens JR, Freeland A, Wittry B, Kramer A, Barclay L, Vinjé J, et al. Notes from the field: multiple cruise ship outbreaks of norovirus associated with frozen fruits and berries—United States, 2019. *MMWR Morb Mortal Wkly Rep*. (2020) 69:501–2. doi: 10.15585/mmwr.mm6916a3
- Vivancos R, Keenan A, Sopwith W, Smith K, Quigley C, Mutton K, et al. Norovirus outbreak in a cruise ship sailing around the British Isles: investigation and multi-agency management of an international outbreak. *J Infect*. (2010) 60:478–85. doi: 10.1016/j.jinf.2010.03.018
- Agrawal S, Orschler L, Tavazzi S, Greither R, Gawlik BM, Lackner S. Genome sequencing of wastewater confirms the arrival of the SARS-CoV-2 omicron variant at Frankfurt airport but limited spread in the city of Frankfurt, Germany, in November 2021. *Microbiol Resour Announce*. (2022) 11:e0122921. doi: 10.1128/MRA.01229-21
- Albastaki A, Naji M, Lootah R, Almeheiri R, Almulla H, Almarri I, et al. First confirmed detection of SARS-CoV-2 in untreated municipal and aircraft wastewater in Dubai, UAE: the use of wastewater based epidemiology as an early warning tool to monitor the prevalence of COVID-19. *Sci Total Environ*. (2021) 760:143350. doi: 10.1016/j.scitotenv.2020.143350
- Farkas K, Williams R, Alex-Sanders N, Grimsley JMS, Pântea I, Wade MJ, et al. Wastewater-based monitoring of SARS-CoV-2 at UK airports and its potential role in international public health surveillance. *PLOS Glob Public Health*. (2023) 3:e0001346. doi: 10.1371/journal.pgph.0001346
- Chen Q, Wu W, Guo Y, Li J, Wei F. Environmental impact, treatment technology and monitoring system of ship domestic sewage: a review. *Sci Total Environ*. (2022) 811:151410. doi: 10.1016/j.scitotenv.2021.151410
- Vaneekhaute C, Fazli A. Management of ship-generated food waste and sewage on the Baltic Sea: A review. *Waste Manag*. (2020) 102:12–20. doi: 10.1016/j.wasman.2019.10.030
- Morehouse C. Wastewater sampling and analysis for commercial passenger vessels. (2003). In: *IEEE OCEANS 2003: Celebrating the past. Teaming toward the future*, San Diego, CA. IEEE Cat. No.03CH37492, MTS/IEEE, pp. 376–385.
- Welsh Government. Sea transport 2018. Statistical bulletin SB 56/2019. Cardiff, UK: Welsh Government (2019).
- Farkas K, Hillary LS, Thorpe J, Walker DI, Lowther JA, McDonald JE, et al. Concentration and quantification of SARS-CoV-2 RNA in wastewater using polyethylene glycol-based concentration and qRT-PCR. *Methods Protocols*. (2021) 4:17. doi: 10.3390/mps4010017
- Kevill JL, Pellett C, Farkas K, Brown MR, Bassano I, Denise H, et al. A comparison of precipitation and filtration-based SARS-CoV-2 recovery methods and the influence of temperature, turbidity, and surfactant load in urban wastewater. *Sci Total Environ*. (2022) 808:151916. doi: 10.1016/j.scitotenv.2021.151916
- Centers for Disease Control and Prevention. (2020). Real-time RT-PCR Primers and Probes for COVID-19. Centers for Disease Control and Prevention. <https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html> (Retrieved 28 January 2022).
- Gendron L, Verreault D, Veillette M, Moineau S, Duchaine C. Evaluation of filters for the sampling and quantification of RNA phase aerosols. *Aerosol Sci Technol*. (2010) 44:893–901. doi: 10.1080/02786826.2010.501351

36. ISO. Microbiology of the food chain—horizontal method for determination of hepatitis A virus and norovirus using real-time RT-PCR—part 2: Method for detection. Geneva, Switzerland: International Organization for Standardization (2019).
37. Poelman R, Schölvinc EH, Borger R, Niesters HG, van Leer-Buter C. The emergence of enterovirus D68 in a Dutch university medical center and the necessity for routinely screening for respiratory viruses. *J Clin Virol.* (2015) 62:1–5. doi: 10.1016/j.jcv.2014.11.011
38. Shu B, Kirby MK, Davis WG, Warnes C, Liddell J, Liu J, et al. Multiplex real-time reverse transcription PCR for influenza A virus, influenza B virus, and severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis.* (2021) 27:1821–30. doi: 10.3201/eid2707.210462
39. Farkas K, Adriaenssens EM, Walker DI, McDonald JE, Malham SK, Jones DL. Critical evaluation of crAssphage as a molecular marker for human-derived wastewater contamination in the aquatic environment. *Food Environ Virol.* (2019) 11:113–9. doi: 10.1007/s12560-019-09369-1
40. Farkas K, Peters DE, McDonald JE, de Rougemont A, Malham SK, Jones DL. Evaluation of two triplex one-step qRT-PCR assays for the quantification of human enteric viruses in environmental samples. *Food Environ Virol.* (2017) 9:342–9. doi: 10.1007/s12560-017-9293-5
41. Jeffries A., Child H.T., Paterson S., Loose M., van Aerle R. Wastewater sequencing using the EasySeq™ RC-PCR SARS CoV-2 (Nimagen) V2.0 V2. (2021). Available at: <https://www.protocols.io/view/wastewater-sequencing-using-the-easyseq-rc-pcr-sar-81wgb7bx3vpk/v2/materials>
42. Connor T. A Nextflow pipeline for running the ARTIC network's field bioinformatics tools. GitHub. (2021). Available at: <https://github.com/connor-lab/ncov2019-artic-nf> (Accessed 3 November 2021).
43. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal.* (2011) 17:10. doi: 10.14806/ej.17.1.200
44. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature.* (2020) 579:265–9. doi: 10.1038/s41586-020-2008-3
45. Li H, Durbin R. Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics.* (2010) 26:589–95. doi: 10.1093/bioinformatics/btp698
46. Brunner FS, Payne A, Cairns E, Airey G, Gregory R, Pickwell ND, et al. Utility of wastewater genomic surveillance compared to clinical surveillance to track the spread of the SARS-CoV-2 omicron variant across England. *Water Res.* (2023) 247:120804. doi: 10.1016/j.watres.2023.120804
47. Kayikcioglu T, Amirzadegan J, Rand H, Tesfaldet B, Timme RE, Pettengill JB. Performance of methods for SARS-CoV-2 variant detection and abundance estimation within mixed population samples. *PeerJ.* (2023) 11:e14596. doi: 10.7717/peerj.14596
48. Murphy J, Riley JP. A modified single solution method for the determination of phosphate in natural waters. *Anal Chim Acta.* (1962) 27:31–6. doi: 10.1016/S0003-2670(00)88444-5
49. Mulvaney RL. (1996). Nitrogen—Inorganic Forms. In D. L. Sparks, A. L. Page, P. A. Helmke, R. H. Loeppert, P. N. Soltanpour, M. A. Tabatabai, C. T. Johnston, & Summer M.E. (Eds.), *Methods of Soil Analysis: Part 3 Chemical Methods*, 5.3 (pp. 1123–1184. doi: 10.2136/sssabooks5.3.c38
50. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem.* (2009) 55:611–22. doi: 10.1373/clinchem.2008.112797
51. ONS. Coronavirus (COVID-19) infection survey (positivity estimates): 20 to 26 February 2022. London, UK: Office for National Statistics (2022).
52. HPSC. Weekly report on the epidemiology of COVID-19 in Ireland, week 6, 2022. Dublin, Ireland: Health Protection Surveillance Centre (2022).
53. McCabe R, Danelian G, Panovska-Griffiths J, Donnelly CA. Inferring community transmission of SARS-CoV-2 in the United Kingdom using the ONS COVID-19 infection survey. *Infect Dis Model.* (2024) 9:299–313. doi: 10.1016/j.idm.2024.01.011
54. Vihta KD, Pouwels KB, Peto T, Pritchard E, Eyre DW, House T, et al. Symptoms and SARS-CoV-2 positivity in the general population in the UK. *Clin Infect Dis.* (2021) 75:e329–37. doi: 10.1093/cid/ciab945
55. Brumovský M, Bečanová J, Sáňka O, Løken KB, Baho DL, Sørensen K, et al. Line ferries and cargo ships for the monitoring of marine contaminants of emerging concern: application along a Europe-Arctic transect. *J Hazard Mater.* (2022) 424:127232. doi: 10.1016/j.jhazmat.2021.127232
56. Loehr LC, Beegle-Krause CJ, George K, McGee CD, Mearns AJ, Atkinson MJ. The significance of dilution in evaluating possible impacts of wastewater discharges from large cruise ships. *Mar Pollut Bull.* (2006) 52:681–8. doi: 10.1016/j.marpolbul.2005.10.021
57. Perić T, Komadina P, Račić N. Wastewater pollution from cruise ships in the Adriatic Sea. *Promet Traf Transport.* (2016) 28:425–33. doi: 10.7307/ptt.v28i4.2087
58. Perić T, Mihanović V, Račić N. Evaluation model of marine pollution by wastewater from cruise ships. *Brodogradnja.* (2019) 70:79–92. doi: 10.21278/brod70305
59. Vicente-Cera I, Moreno-Andrés J, Amaya-Vías D, Biel-Maeso M, Pintado-Herrera MG, Lara-Martín PA, et al. Chemical and microbiological characterization of cruise vessel wastewater discharges under repair conditions. *Ecotoxicol Environ Saf.* (2019) 169:68–75. doi: 10.1016/j.ecoenv.2018.11.008
60. Ahmed W, Bertsch PM, Angel N, Bibby K, Bivins A, Dierens L, et al. Detection of SARS-CoV-2 RNA in commercial passenger aircraft and cruise ship wastewater: a surveillance tool for assessing the presence of COVID-19 infected travellers. *J Travel Med.* (2020b) 27:taaa116. doi: 10.1093/jtm/taaa116
61. Abe H, Ushijima Y, Amano M, Sakurai Y, Yoshikawa R, Kinoshita T, et al. Unique evolution of SARS-CoV-2 in the second large cruise ship cluster in Japan. *Microorganisms.* (2022) 10:99. doi: 10.3390/microorganisms10010099
62. Bert F, Scialoi G, Gualano MR, Passi S, Specchia ML, Cadeddu C, et al. Norovirus outbreaks on commercial cruise ships: a systematic review and new targets for the public health agenda. *Food Environ Virol.* (2014) 6:67–74. doi: 10.1007/s12560-014-9145-5
63. Rogers KB, Roohi S, Uyeki TM, Montgomery D, Parker J, Fowler NH, et al. Laboratory-based respiratory virus surveillance pilot project on select cruise ships in Alaska, 2013–15. *J Travel Med.* (2017) 24:tax069. doi: 10.1093/jtm/tax069
64. Li X, Kulandaivelu J, Guo Y, Zhang S, Shi J, O'Brien J, et al. SARS-CoV-2 shedding sources in wastewater and implications for wastewater-based epidemiology. *J Hazard Mater.* (2022) 432:128667. doi: 10.1016/j.jhazmat.2022.128667
65. Arena C, Amoros JP, Vaillant V, Balay K, Chikhi-Brachet R, Varesi L, et al. Simultaneous investigation of influenza and enteric viruses in the stools of adult patients consulting in general practice for acute diarrhea. *Virol J.* (2012) 9:116. doi: 10.1186/1743-422x-9-116
66. Jones DL, Baluja MQ, Graham DW, Corbishley A, McDonald JE, Malham SK, et al. Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Sci Total Environ.* (2020) 749:141364. doi: 10.1016/j.scitotenv.2020.141364
67. Minodier L, Charrel RN, Ceccaldi P-E, Van Der Werf S, Blanchon T, Hanslik T, et al. Prevalence of gastrointestinal symptoms in patients with influenza, clinical significance, and pathophysiology of human influenza viruses in faecal samples: what do we know? *Virol J.* (2015) 12:215. doi: 10.1186/s12985-015-0448-4
68. Sweeney WB, Krafte-Jacobs B, Britton JW, Hansen W. The constipated serviceman: prevalence among deployed U.S. troops. *Mil Med.* (1993) 158:546–8. doi: 10.1093/milmed/158.8.546
69. Jones DL, Rhymes JM, Wade M, Kevill JL, Malham SK, Grimsley JMS, et al. Suitability of aircraft wastewater for pathogen detection and public health surveillance. *Sci Total Environ.* (2022) 856:159162. doi: 10.1016/j.scitotenv.2022.159162
70. Gao Z, Xu Y, Sun C, Wang X, Guo Y, Qiu S, et al. A systematic review of asymptomatic infections with COVID-19. *J Microbiol Immunol Infect.* (2021) 54:12–6. doi: 10.1016/j.jmii.2020.05.001
71. Kim M-K, Lee B, Choi YY, Um J, Lee K-S, Sung HK, et al. Clinical characteristics of 40 patients infected with the SARS-CoV-2 omicron variant in Korea. *J Korean Med Sci.* (2022) 37:e31. doi: 10.3346/jkms.2022.37.e31
72. Torabi SH, Riahi SM, Ebrahimzadeh A, Salmani F. Changes in symptoms and characteristics of COVID-19 patients across different variants: two years study using neural network analysis. *BMC Infect Dis.* (2023) 23:838. doi: 10.1186/s12879-023-08813-9
73. Lindsay L, Dupont HL, Moe CL, Alberer M, Hatz C, Kirby AE, et al. Estimating the incidence of norovirus acute gastroenteritis among US and European international travelers to areas of moderate to high risk of traveler's diarrhea: a prospective cohort study protocol. *BMC Infect Dis.* (2018) 18:605. doi: 10.1186/s12879-018-3461-6
74. Ajami NJ, Kavanagh OV, Ramani S, Crawford SE, Atmar RL, Jiang Z, et al. Seroepidemiology of norovirus-associated travelers' diarrhea. *J Travel Med.* (2014) 21:6–11. doi: 10.1111/jtm.12092
75. Chapin AR, Carpenter CM, Dudley WC, Gibson LC, Pratdesaba R, Torres O, et al. Prevalence of norovirus among visitors from the United States to Mexico and Guatemala who experience traveler's diarrhea. *J Clin Microbiol.* (2005) 43:1112–7. doi: 10.1128/jcm.43.3.1112-1117.2005
76. Ryder RW, Oquist CA, Greenberg H, Taylor DN, Orskov F, Orskov I, et al. Travelers' diarrhea in panamanian tourists in Mexico. *J Infect Dis.* (1981) 144:442–8. doi: 10.1093/infdis/144.5.442
77. Teunis PFM, Sukhrie FHA, Vennema H, Bogerman J, Beersma MFC, Koopmans MPG. Shedding of norovirus in symptomatic and asymptomatic infections. *Epidemiol Infect.* (2015) 143:1710–7. doi: 10.1017/s095026881400274x
78. Ondrikova N, Clough HE, Douglas A, Iturriza-Gomara M, Larkin L, Vivancos R, et al. Differential impact of the COVID-19 pandemic on laboratory reporting of norovirus and *Campylobacter* in England: a modelling approach. *PLoS One.* (2021) 16:e0256638. doi: 10.1371/journal.pone.0256638
79. UKHSA. National norovirus and rotavirus report, week 12 report: Data up to week 10, 12th march 2023. London, UK: UK Health Security Agency (2023a).
80. PHE. Surveillance of influenza and other seasonal respiratory viruses in the UK winter 2020 to 2021. London, UK: Public Health England (2021).
81. UKHSA. Surveillance of influenza and other seasonal respiratory viruses in the UK, winter 2022 to 2023. London, UK: UK Health Security Agency (2023b).
82. Rafiee M, Isazadeh S, Mohseni-Bandpei A, Mohebbi SR, Jahangiri-Rad M, Eslami A, et al. Moore swab performs equal to composite and outperforms grab sampling for SARS-CoV-2 monitoring in wastewater. *Sci Total Environ.* (2021) 790:148205. doi: 10.1016/j.scitotenv.2021.148205

83. Farkas K, Kevill JL, Adwan L, Garcia-Delgado A, Dzay R, Grimsley JMS, et al. Near-source passive sampling for monitoring viral outbreaks within a university residential setting. *Epidemiol Infect.* (2024) 152:e31. doi: 10.1017/S0950268824000190
84. Edwards RA, Vega AA, Norman HM, Ohaeri M, Levi K, Dinsdale EA, et al. Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. *Nat Microbiol.* (2019) 4:1727–36. doi: 10.1038/s41564-019-0494-6
85. Heinen E., Potts K., Snow L., Trulli W., Redford D. Dilution of wastewater discharges from moving cruise ships. (2003). In: *IEEE OCEANS 2003: Celebrating the past. Teaming toward the future*, San Diego, CA. IEEE Cat. No.03CH37492, MTS/IEEE, pp. 386–389.
86. Kersh J, James CA, Gough HL. Impacts of high-strength boat waste on activated sludge processes. *J Environ Eng.* (2020) 146:04020023. doi: 10.1061/(asce)ee.1943-7870.0001679
87. Honap TP, Sankaranarayanan K, Schnorr SL, Ozga AT, Warinner C, Lewis CM Jr. Biogeographic study of human gut-associated crAssphage suggests impacts from industrialization and recent expansion. *PLoS One.* (2020) 15:e0226930. doi: 10.1371/journal.pone.0226930



OPEN ACCESS

EDITED BY

Jie Huang,
Southern University of Science and
Technology, China

REVIEWED BY

Jesus L. Romalde,
University of Santiago de Compostela, Spain
Andrea E. Kirkwood,
Ontario Tech University, Canada
Tomás de Melo,
Ontario Tech University, Canada in
collaboration with reviewer AK
Ayse Seval Palteki,
Istanbul Medipol University, Türkiye

*CORRESPONDENCE

Carmelo Massimo Maida
✉ carmelo.maida@unipa.it

RECEIVED 07 February 2024

ACCEPTED 27 June 2024

PUBLISHED 23 July 2024

CITATION

Maida CM, Mazzucco W, Priano W, Palermo R,
Graziano G, Costantino C, Russo A,
Andolina G, Restivo I, Giangreco V, Iaia FR,
Santino A, Li Muli R, Guzzetta V, Vitale F and
Tramuto F (2024) Detection of influenza virus
in urban wastewater during the season
2022/2023 in Sicily, Italy.
Front. Public Health 12:1383536.
doi: 10.3389/fpubh.2024.1383536

COPYRIGHT

© 2024 Maida, Mazzucco, Priano, Palermo,
Graziano, Costantino, Russo, Andolina,
Restivo, Giangreco, Iaia, Santino, Li Muli,
Guzzetta, Vitale and Tramuto. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Detection of influenza virus in urban wastewater during the season 2022/2023 in Sicily, Italy

Carmelo Massimo Maida^{1,2*}, Walter Mazzucco^{1,2}, Walter Priano¹,
Roberta Palermo¹, Giorgio Graziano², Claudio Costantino^{1,2},
Arianna Russo², Gina Andolina², Isabella Restivo¹,
Viviana Giangreco², Francesca Rita Iaia¹, Arianna Santino²,
Rita Li Muli², Valeria Guzzetta², Francesco Vitale^{1,2} and
Fabio Tramuto^{1,2}

¹Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties "G. D'Alessandro", University of Palermo, Palermo, Italy, ²Clinical Epidemiology Unit, Regional Reference Laboratory of Western Sicily for the Emergence of COVID-19, University Hospital "P. Giaccone", Palermo, Italy

Introduction: Seasonal influenza generally represents an underestimated public health problem with significant socioeconomic implications. Monitoring and detecting influenza epidemics are important tasks that require integrated strategies. Wastewater-based epidemiology (WBE) is an emerging field that uses wastewater data to monitor the spread of disease and assess the health of a community. It can represent an integrative surveillance tool for better understanding the epidemiology of influenza and prevention strategies in public health.

Methods: We conducted a study that detected the presence of Influenza virus RNA using a wastewater-based approach. Samples were collected from five wastewater treatment plants in five different municipalities, serving a cumulative population of 555,673 Sicilian inhabitants in Italy. We used the RT-qPCR test to compare the combined weekly average of Influenza A and B viral RNA in wastewater samples with the average weekly incidence of Influenza-like illness (ILI) obtained from the Italian national Influenza surveillance system. We also compared the number of positive Influenza swabs with the viral RNA loads detected from wastewater. Our study investigated 189 wastewater samples.

Results: Cumulative ILI cases substantially overlapped with the Influenza RNA load from wastewater samples. Influenza viral RNA trends in wastewater samples were similar to the rise of ILI cases in the population. Therefore, wastewater surveillance confirmed the co-circulation of Influenza A and B viruses during the season 2022/2023, with a similar trend to that reported for the weekly clinically confirmed cases.

Conclusion: Wastewater-based epidemiology does not replace traditional epidemiological surveillance methods, such as laboratory testing of samples from infected individuals. However, it can be a valuable complement to obtaining additional information on the incidence of influenza in the population and preventing its spread.

KEYWORDS

wastewater, surveillance, wastewater-based epidemiology, influenza viruses, influenza season

1 Introduction

Influenza is a viral acute respiratory infection with high morbidity and mortality in humans, especially in specific groups such as children and older adults, posing a constant threat to global public health because of recurring seasonal epidemics and irregularly occurring pandemics (1–3). The burden of this disease can vary widely, being determined by several factors, including the characteristics of circulating viruses, the timing of the season, the environmental temperature, how well the available vaccine is working to protect against illness, and how many people got vaccinated (4, 5). The Centers for Disease Control and Prevention (CDC) estimated that influenza has resulted in 9 million–41 million illnesses, 140,000–710,000 hospitalizations, and 12,000–52,000 deaths annually between 2010 and 2020 in the United States (6). Seasonal influenza epidemics have substantially contributed to the worldwide annual mortality rate, particularly among the older adult 65 years and over. In Italy, a mortality rate of 10.7 per 1,000 inhabitants was observed in the winter season of 2014/2015 (more than 375,000 deaths in absolute terms), corresponding to an estimated 54,000 excess deaths (+9.1%), as compared to the previous season (7), representing the highest reported mortality rate since the Second World War in this country (8). Rapid population growth, climate change, natural disasters, immigration, globalization, and the corresponding sanitation and waste management challenges are expected to intensify the problem in the future (9).

Worryingly, seasonal influenza generally represents an underappreciated public health problem with significant socio-economic implications (10). Monitoring and detecting influenza outbreaks are important but challenging tasks. To accurately track the spread of influenza, reporting systems for influenza-like illness (ILI) and laboratory-confirmed influenza infections (11) can be helpful. These systems are crucial for estimating the number of people experiencing symptoms, hospitalizations, and deaths caused by influenza, addressing vaccination campaigns, and allocating treatment resources. The surveillance of seasonal influenza is possible through data collection and sharing systems, such as FluView in the United States¹ and FluNews in Europe,² which systematically collect data on seasonal influenza and publish periodic reports to inform on epidemiological trends. Influnet is the Italian nationwide sentinel surveillance system for influenza, coordinated by the Italian National Institute of Health (NIH), collecting epidemiological and virological data that are published weekly on the integrated surveillance system portal³ according to an operative protocol⁴ and uploaded into the European database coordinated by the European Centre for Disease Prevention and Control (ECDC) (12). Collaborating sentinel doctors from each region of the country report cases of ILI observed among their patients, collecting, at the same time, biological respiratory samples to identify circulating viruses. The European case definition of ILI was adopted to ensure maximum homogeneity of detection. A case of ILI was defined as a person presenting a sudden and rapid onset of at least one of the following systemic symptoms: fever or feverishness, malaise, headache,

myalgia; and at least one of the following respiratory symptoms: cough, sore throat, shortness of breath (13). Doctors take throat swabs from ILI patients tested for influenza viruses at regional Influnet laboratories.

The experience gained over the last few years indicates that the Influenza virus and Coronaviruses are the two main viruses that pose a high risk to humans. Influenza A viruses can infect various animals and humans, leading to pandemics (14, 15). Although environmental virus monitoring can be helpful, the methods are mainly based on clinical data and not validated for environmental testing (16).

Despite this, since the beginning of the COVID-19 pandemic, the utility of wastewater-based epidemiology (WBE) has emerged as a tool for researchers to monitor the circulation of SARS-CoV-2 through the design of pilot studies that highlighted the link between environmental and clinical frameworks (17–22). WBE provides quickly anonymous and aggregated data at a low cost and at a potentially large scale through the passive contributions of the community, therefore integrating the conventional surveillance programs and strengthening health emergency response systems, as occurred with the tracking of the poliovirus during the twentieth century (23). Over the past 2 years, the number of studies supporting wastewater surveillance to monitor the circulation of respiratory pathogens and Influenza viruses in communities has been increasing (9, 24–32). As an effective health assessment approach, WBE has great potential in warning of infectious disease outbreaks for public health (20), as recently demonstrated in Italy during the COVID-19 pandemic (17, 21, 22). Our study aimed to monitor the presence of the influenza virus in the wastewater of different cities on the island. The objective was to evaluate the circulation of the virus throughout an entire Influenza season and compare the results with the conventional integrated epidemiological and virology surveillance.

2 Materials and methods

2.1 Study design and sample collections

We conducted an observational study in Sicily (Italy), the largest and most populous island in the Mediterranean Sea, accounting for about 5 million resident inhabitants (33). Five wastewater treatment plants (WTPs) located in five different municipalities, serving a cumulative population of 555,673 inhabitants (ranging from 34,000 to 314,973; 11.1% of total island residents), were included in the study. Raw 24-h composite wastewater samples ($n=188$) were collected weekly for 9 months, between August 2022 (week 31/2022) and April 2023 (week 17/2023), by an automatic sampling device. Further information about the location and the characteristics of WTPs is provided in Figure 1. The collected samples were refrigerated, transferred to the laboratory, and tested for influenza viral RNA within 24 h from sampling. The wastewater samples collection period (week 31/2022) started before the national epidemiological/virological surveillance (week 42/2022) to assess the viral RNA early detection in wastewater. This evaluation determines if the WBE methodology can serve as an early warning system for influenza circulation.

2.2 Virus concentration

All samples underwent a 30-min treatment at 56°C to minimize the potential impact of bioaerosol on personnel and environmental

1 www.cdc.gov/flu/weekly

2 www.flunews-europe.org

3 <https://respirvirnet.iss.it/pagine/rapportoInflunet.aspx>

4 https://www.salute.gov.it/imgs/C_17_pubblicazioni_3267_allegato.pdf

Wastewater treatment plants	Average inflow (m ³ /day ± SD)	Served population (N)
1) Agrigento	7,673.3 ± 1,399.4	55,000
2) Bagheria	9,540.5 ± 295.3	75,000
3) Caltanissetta	13,233.4 ± 1,532.7	76,700
4) Enna	4,091.6 ± 1,270.8	34,000
5) Palermo	72,121.2 ± 8,860.2	314,973

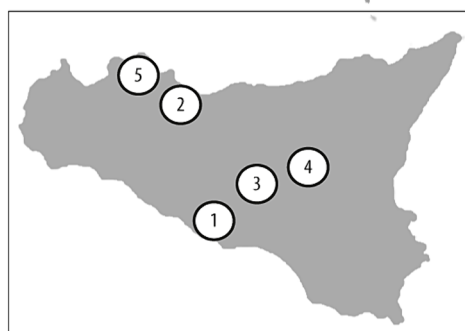


FIGURE 1

Location and the characteristics of wastewater treatment plant involved in the study.

safety (34–37). Heat-treating samples at 56°C for 30 min should cause a negligible or little effect on the sensitivity of RT-PCR (17, 38, 39). Then, each sample was concentrated using a polyethylene glycol (PEG)-based procedure, according to Wu et al. (40) protocol with minor modification. Briefly, wastewater samples (45 mL) were centrifuged at 4,500 × g for 30 min; after centrifugation, 40 mL of sample were mixed with 8% w/v polyethylene glycol 8.000 and 0.3 M NaCl (both supplied by Sigma-Aldrich, St. Louis, MO, USA), spiked with a known amount of Murine Norovirus, used as a process control. After a centrifugation step at 12,000 × g for 2 h, the viral pellet was resuspended in 2 mL of NucliSENS Lysis Buffer reagent (bioMerieux, Marcy-l'Étoile, France) for sub-sequent RNA extraction.

Viral RNA extraction was performed using a semi-automated system based on lysis and magnetic silica beads (supplied by bioMerieux, Marcy l'Etoile, France). After an incubation of 20 min at room temperature, 100 µL of magnetic silica beads were added. After further incubation for 10 min, an automated procedure was performed using the nucleic acid purification system (Auto-Pure96, All Sheng Instruments, Zhejiang, China). Before molecular tests, the extracted nucleic acids in an eluent volume of 100 µL, were purified from potential PCR inhibitors using the OneStep PCR Inhibitor Removal Kit (Zymo Research, CA, USA).

2.3 RT-qPCR

One-step real-time reverse-transcription (RT) quantitative PCR assays were used to detect the presence of Influenza A viral RNA

(IAV) and/or Influenza B viral RNA (IBV) according to the CDC protocol with minor modifications.⁵ A test was considered positive when its cycle threshold (Ct) value was <40. All q-PCR assays were performed with singleplex real-time PCR (rPCR) assays using the TaqMan technology and run on a QuantStudio™ 7 Flex Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA); primers, probes sets and reagents are described in Tables 1, 2. For the detection of viral RNA, we performed q-PCR as a single step using the Quantinova Pathogen + IC kit Polymerase (Qiagen, CA, USA). The PCR conditions were as follows: 1 cycle at 50°C for 2 min; 1 cycle at 95°C for 2 min; 45 cycles at 95°C for 15 s and 55°C for 30 s.

Viral RNA quantification was performed using 10-fold dilutions, ranging from 1.0 to 1.0 × 10⁵ Genomic Copies (GC)/µL per reaction, of a synthetic double-stranded plasmid construct carrying IAV and IBV nucleotide sequences specific for the real-time assays. qPCR standard curves were generated by linear regression of Ct values versus log₁₀ standard concentration and used to convert Ct values into influenza RNA copies/µL per reaction (Slope = − 3.385; R = 0.999; Efficiency (%) = 97.422; Y-intercept = 21.721). The influenza viral RNA's GC/L in wastewater was obtained according to the formula: (GC/µL × (100 µL/40 mL)) × 1.000 mL/L. The results were also evaluated in GC/day/inhabitant according to the following formula: flow rate of WTP in 24 h (m³) × GC (GC/L)/equivalent number of inhabitants served by the WTP. Verification of PCR inhibition was

⁵ <https://www.cdc.gov/coronavirus/2019-ncov/lab/multiplex.html>

TABLE 1 Primers and probes for detecting influenza A, influenza B and Murine Norovirus by q-PCR.

Name	Description	Oligonucleotide sequence (5'–3')
InfA-F	InfA For1	CAA GAC CAA TCY TGT CAC CTC TGA C
	InfA For2	CAA GAC CAA TYC TGT CAC CTY TGA C
InfA-R	InfA Rev1	GCA TTY TGG ACA AAV CGT CTA CG
	InfA Rev2	GCA TTT TGG ATA AAG CGT CTA CG
InfA-P	InfA Probe	FAM/TGC AGT CCT CGC TCA CTG GGC ACG/BHQ
InfB-F	InfB For	TCC TCA AYT CAC TCT TCG AGC G
InfB-R	InfB Rev	CGG TGC TCT TGA CCA AAT TGG
InfB-P	InfB Probe	FAM/CCA ATT CGA GCA GCT GAA ACT GCG GTG/BHQ
MNV orf1/2junct/F	MNV For	CAC GCC ACC GAT CTG TTC TG
MNV orf1/2junct/R	MNV Rev	GCG CTG CGC CAT CAC TC
MNV orf1/2junct/P	MNV Probe	FAM/CGC TTT GGA ACA ATG/MGBNFQ

TABLE 2 The PCR reagents.

Reagent for flu A detection	Final concentration (nM)	Volume (μl)
Quantinova Master Mix*	–	3.90
InfA For1	400	0.15
InfA For2	400	0.15
InfA Rev1	600	0.225
InfA Rev2	200	0.075
InfA Probe	300	0.45
Nuclease free water	–	5.05
Sample	–	5.00
Total volume		15.00
Reagent for flu B detection	Final concentration (nM)	Volume (μl)
Quantinova Master Mix*	–	3.90
InfB For	800	0.30
InfB Rev	800	0.30
InfB Probe	300	0.45
Nuclease free water	–	5.05
Sample	–	5.00
Total volume		15.00
Reagent for Murine Norovirus detection	Final concentration (nM)	Volume (μl)
Quantinova Master Mix*	–	3.90
InfB For	300	0.15
InfB Rev	600	0.30
InfB Probe	200	0.15
Nuclease free water	–	5.50
Sample	–	5.00
Total volume		15.00

*Quantinova Mastermix is premixed with 15 μL of ROX reference dye before use.

performed as a quality parameter of the determinations. To verify the inhibition, the PCR Ct value obtained from the sample added with 1 μL of a 1.0×10^3 GC/μL of the synthetic double-stranded plasmid was compared with the PCR Ct value of water for molecular biology

added with 1 μL of the same synthetic double-stranded plasmid, according to the following formula: $\Delta Ct = Ct \text{ (sample + control plasmid)} - Ct \text{ (water + control plasmid)}$. The sample was considered acceptable if ΔCt was ≤ 2 . Before performing sample analysis, the limit

of detection (LoD) was determined by spiking wastewater extracts with dilutions of the synthetic double-stranded plasmid solutions at concentrations of approximately 1,000, 100, 50, 20, 10, 2, and 1.0 GC/ μ L. Ten replicates of each dilution were tested. The LoD was the lowest concentration at which all ten replicates were positive. The assay had a LoD of 2.5 GC/ μ L. The concentration/extraction efficiency of the method was assessed as previously reported (22). The sample was considered acceptable if the concentration/extraction efficiency was $\geq 1\%$.

2.4 Clinical and virologic data sources

We accessed the Influnet web-based platform data (41) to obtain weekly national and regional epidemiological and virological reports, including the ILI incidence per 1,000 inhabitants for the Sicilian region and the aggregate number of influenza-positive swabs. Specifically, data were retrieved from week 42/2022 (conventionally marked as the starting week for influenza virus circulation and thus established as the onset time for the start of the national influenza circulation surveillance-system data collection) up until week 17/2023 (considered as the ending of influenza season).

2.5 Statistical analyses

The national surveillance influenza platform contains regional data regarding influenza virus surveillance. Since data collection was performed weekly, IAV and IBV viral loads (intended as viral RNA copy numbers per day/inhabitants of wastewater) detected from the five Sicilian WTPs were aggregated in weekly means and summed, thus obtaining the total IAV + IBV viral load. Moreover, new time-dependent variables (lag times) were created to assess the wastewater detection method's early-warning capacity. They were based on a method we already performed in our previous WBE study (42). Specifically, by using "WTPs sampling week" and "regional ILI incidence per 1,000 inhabitants" as key variables, the incidence was set at week 0 (intended as the week of sample collection), week 1 and 2 (respectively, 1 and 2 weeks ahead of the WTPs' sampling week).

As viral concentrations in wastewater are log-normally distributed, a log-10 transformation was applied for all the variables we analysed. Thus, although WBE data were collected from week 31/2022 to assess early virus circulation, national surveillance data were available from week 42/2022. Thus, Person's correlation test, log-linear regression analyses and significance tests, retrieving R , r^2 and p -values, were carried out through RStudio software (version 4.2.2) to compare from week 42/2022 to week 17/2023, at weeks 0, 1 and 2, the following variables:

- The mean weekly regional ILI incidence per 1,000 inhabitants with the weekly average of combined IAV and IBV viral loads derived from WTPs.
- The weekly regionally combined number of positive IAV and IBV swabs detected, with the combined IAV and IBV Regional viral load detected from WTPs.

The Shapiro–Wilk test was carried out to check for the normality of each continuous variable. A p -value < 0.05 was considered statistically significant.

3 Results

Overall, from 7 September 2022 to 30 April 2023, 189 wastewater samples were investigated every week. In particular, the following samples were collected from five municipalities and tested for IAV and IBV RNA: Agrigento ($n = 36$), Bagheria ($n = 37$), Caltanissetta ($n = 39$), Enna ($n = 39$), and Palermo ($n = 37$). Overall, IAV RNA was detected in 123/189 samples (65.1%) and IBV RNA in 37/189 samples (19.5%), while the co-presence of the two viral RNA was recorded in 22/189 (11.6%) of the analyzed samples. The recovery rate of influenza viral RNA has ranged from 1 to 100% (mean 8.72; 95% C.I. = 6.35–11.09), compared to a Murine Norovirus control of known concentration in PCR grade water. Table 3 shows the descriptive analysis of the main clinical and virological surveillance data of the flu season 2022/2023. In the entire study period, the concentration of IAV in wastewater ranged from 0.0 to 9.3×10^5 GC/day/inhabitants, while IBV ranged from 0.0 to 3.5×10^5 GC/day/inhabitants. Figure 2 depicts the weekly trends in the ILI regional incidence, reported by the national surveillance system (primary y-axis) and the influenza RNA load in sewage (secondary y-axis) per week of the year (x-axis). In week 36/2022, the first influenza-positive wastewater samples were recorded, with an average concentration of 4.4×10^4 GC/day/inhabitants. In the following weeks, there was a constantly increasing trend of viral RNA detected in the wastewater until reaching the peak of 9.3×10^5 GC/day/inhabitants in week 50/2022. From then on, the viral RNA concentration in wastewater progressively decreased until week 06/2023, after which a second lower peak occurred at week 10/2023, quantified as 3.9×10^3 GC/day/inhabitants. After that, the viral RNA concentration in wastewater regularly decreased until the absence of detection from week 14/2023. On the other hand, the epidemiological trend of ILI at a regional level showed high values starting from week 42/2022, the first surveillance week of the 2022/2023 season, and peaked in week 49/2022. Excluding small occasional increases in ILIs, the trend has been downward until the end of the surveillance season scheduled for week 17/2023. The number of cumulative ILI cases substantially overlapped with the influenza RNA load from wastewater samples, with an increasing trend of influenza viral RNA in wastewater samples comparable to the rise of ILI cases in the population. Figure 3 shows the trend of IAV and IBV circulating regionally, obtained from the virological surveillance system and the viral RNA load detected from the local wastewater samples. The wastewater analyses allowed us to record the total presence of IAV from week 36/2022 until week 51/2022. From week 52/2022 and up to week 13/2023, there was a co-circulation of the two types of viruses, and the concentration of IBV had an increasing trend until its peak recorded at week 09/2023 with a concentration of 3.5×10^5 GC/day/inhabitants. In confirmation of the co-circulation of viruses from week 52/2022 and of the subsequent predominance of IBV over IAV from week 05/2022, the ratio of IBV over IAV showed values of 0.1 in week 52/2022, of 1.5 in week 05/2023 and 12.8 in week 07/2022 and, in any case, always greater than one up to week 11/2023, the last in which the wastewater samples gave a positive result. A similar trend was shown by the regional virological surveillance of influenza-positive swabs, in which from week 46/2022 to week 50/2022, there was an exclusive circulation of the IAV, a co-circulation of both viruses up to week 17/2023 with a predominance of IBV from week 06/2023 to week 17/2023, with a ratio of type B to type A ranging from 1.2 to 4.8.

As shown in Table 3, the correlation analyses between the Influenza viral RNA load (IAV + IBV RNA concentration) detected in WTPs and the regional incidence of ILI per 1,000 inhabitants

TABLE 3 Descriptive analysis containing the total weekly mean Influenza virus load assessed in wastewater from the different WTPs, the regional weekly ILI incidence per 1,000 inhabitants, the total number of regional swabs performed and the positivity rate.

Week	Regional ILI incidence (x 1,000 inhabitants)	Total swabs performed (N)	Positivity rate (%)	Viral load (GC/day/inhabitants)
42/2022	3.7	20	10.0	4.1×10^5
43/2022	4.8	2	0.0	4.0×10^5
44/2022	4.0	1	100.0	2.9×10^5
45/2022	4.7	8	50.0	3.8×10^5
46/2022	8.7	63	38.0	3.4×10^5
47/2022	8.6	75	33.3	5.4×10^5
48/2022	12.2	92	53.3	7.8×10^5
49/2022	14.0	102	54.0	6.5×10^5
50/2022	13.6	135	58.5	9.3×10^5
51/2022	12.3	130	47.7	6.3×10^5
52/2022	12.1	99	37.3	7.0×10^5
01/2023	12.0	100	25.0	6.5×10^5
02/2023	10.9	119	5.9	4.1×10^5
03/2023	11.9	85	15.3	3.1×10^5
04/2023	9.6	51	19.6	2.8×10^5
05/2023	9.5	34	34.0	1.8×10^5
06/2023	8.1	29	29.0	2.0×10^5
07/2023	6.2	13	13.0	2.9×10^5
08/2023	8.2	30	16.6	1.8×10^5
09/2023	7.5	16	31.2	3.9×10^5
10/2023	7.2	13	13.0	3.9×10^5
11/2023	6.7	25	16.0	2.2×10^5
12/2023	6.0	21	9.5	1.5×10^5
13/2023	5.6	17	17.6	1.7×10^4
14/2023	4.3	5	60.0	0.0
15/2023	4.5	10	20.0	0.0
16/2023	4.8	1	0.0	0.0
17/2023	3.7	2	0.0	0.0

displayed a p -value < 0.001 at week 0 and < 0.0001 for weeks 1 and 2, respectively. A moderate-high correlation index (R) was retrieved, ranging from 0.55 at week 0 to 0.78 at week 2. Accordingly, a moderate-correlation index was retrieved when comparing the IAV + IBV viral RNA load detected from WTPs with total number of positive IAV + IBV regionally detected swabs at all times evaluated (Table 4: w0 $R = 0.46$, p -value < 0.01 ; w1 $R = 0.55$, p -value < 0.01 ; w2 $R = 0.63$, p -value < 0.001). In Figure 4 are showed the scatterplots describing the correlation between the RNA viral load detected in wastewater (GC/day/inhabitants) and the number of ILI detected per 1,000 inhabitants at week 2.

4 Discussion

Regardless of the influenza season's onset timing, we observed a rapid and early start of the epidemic season in our study. This resulted in the ample virus circulating in the population when the epidemiological

surveillance of the Influnet network began. This trend was also observed in the southern hemisphere, where the Australian data collection systems showed an extremely accelerated and anticipated growth concerning the normal trend (43). In Sicily (Italy), during the first week of surveillance (42/2022, 17–23 October 2022), the incidence of reported ILI, which in principle can be considered a good proxy of the incidence of flu illness (44), was 3.7 cases/1,000 inhabitants, unlike previous influenza seasons which stood at decidedly lower values (41). The anticipated presence of the circulation of influenza viruses was also recorded through the analysis of wastewater, which began in the week of 31/2022 (1–7 August 2022). In week 36/2022 (5–11 September), we simultaneously detected influenza viruses in all municipalities through wastewater analysis. This was 17 days before the start of conventional national surveillance. The values recorded ranged from 6.00×10^2 to 1.24×10^3 GC/L. Unfortunately, we cannot determine the specificity of our method due to the unavailability of sufficient clinical swabs from sentinel doctors for each municipality. Nonetheless, this early detection of pathogen circulation through WBE has the potential to benefit public health greatly. It could aid in

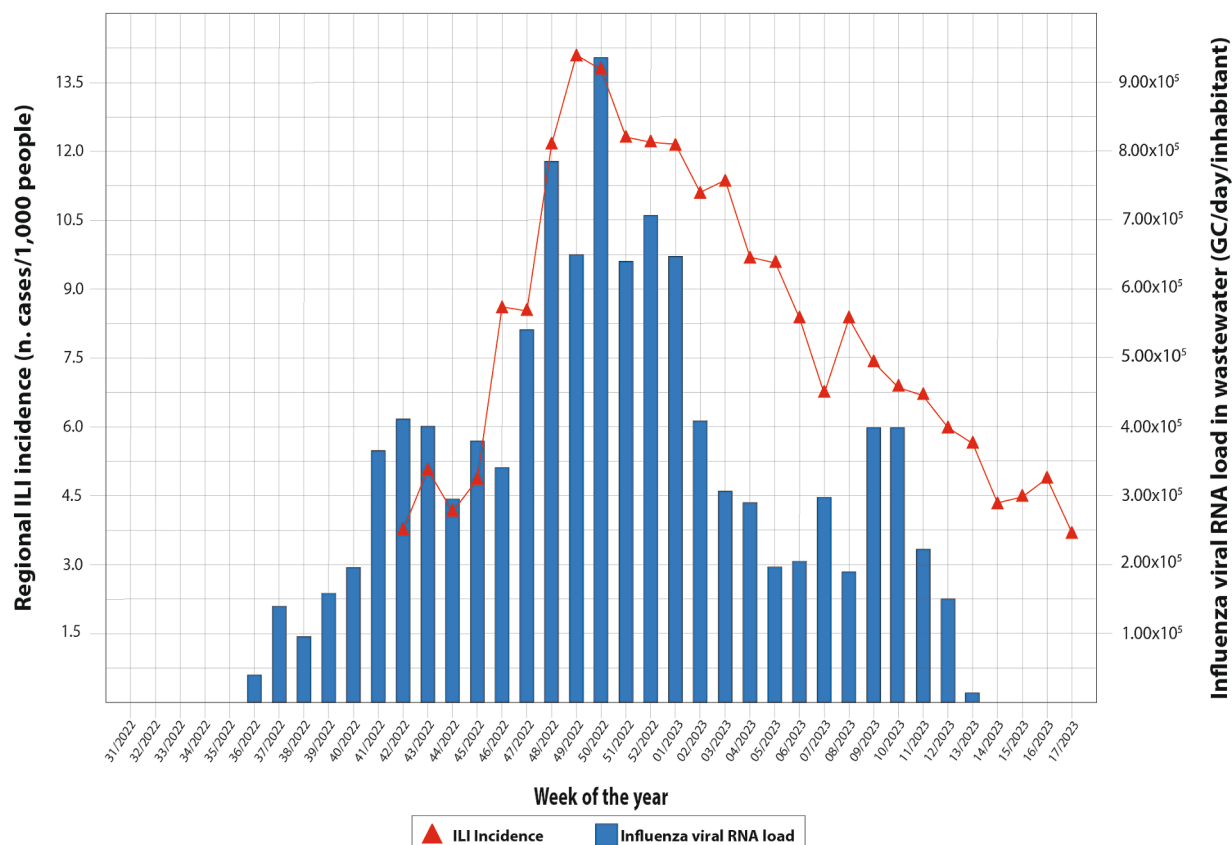


FIGURE 2

Weekly trends in the ILI regional incidence, reported by the national surveillance system and the influenza virus load in sewage per week of the year.

differentiated programming of the start of epidemiological/virological surveillance and vaccination campaigns to increase their effectiveness.

A sustained co-circulation of type A and B influenza viruses characterized Italy's 2022/2023 influenza season. Overall, IAV was prevalent (79.5% of the samples tested positive) compared to IBV (20.5%). The epidemiological data of influenza that have emerged in the southern hemisphere have attested that influenza has been spreading significantly, probably due to the reduction of distancing measures and the use of masks (43). In the five municipalities in the study, wastewater analyses showed that the majority of IAV was detected in week 50/2022 (12–18 September 2022) with 9.3×10^5 GC/day/inhabitants, while the majority of IBV was found in week 09/2023 (27 February – 05 March 2023) with 3.5×10^5 GC/day/inhabitants. The same trend, with a time lag of 7–14 days concerning wastewater, was recorded by the virological surveillance, which dated the peak circulation of the IAV in the week 49/2022, therefore 7 days earlier, and that of the IBV in week 12/2023, then 14 days later (41).

Our findings confirmed that wastewater surveillance can effectively detect influenza virus circulation and should be considered a valuable supplement to conventional influenza surveillance. More in-depth, it may be used to test influenza virus circulation in the communities for prolonged periods using a single sample approach, like the application of SARS-CoV-2 WBE used to monitor the prevalence of COVID-19. The WBE methodology could be an integrative approach to epidemiological and virological surveillance that introduces some interesting aspects to improve the estimation of influenza incidence. By monitoring various treatment plants in the

city, the percentage of subjects tested can be increased compared to the Virological Surveillance Network's target of 4% of the regional population. Additionally, collecting and transporting wastewater is more straightforward, cheaper, and potentially feasible wherever there is a sewage network, thus increasing the possibility of obtaining information even in smaller municipalities that are typically excluded from traditional surveillance systems. While there are many advantages to infectious disease wastewater monitoring, the WBE approach has some limitations, including aggregated data and the inability to perform epidemiological assessments by age groups, symptoms, or immune status for vaccinated subjects. Wastewater is a complex matrix affected by environmental factors that are not always identified, leading to inherent variability and uncertainties (45, 46). Furthermore, it's important to address the lack of standardized protocols in the various phases of the analytical process. This includes sample pre-treatment, concentration, and nucleic acid detection (47). We need to establish a testing framework that considers the different analytical sensitivities at each analysis step. For example, in the thermal pretreatment phase, some studies show negligible changes in RNA measurement (34–37), while others do not (48–50). Similarly, in the concentration phase, the PEG-supernatant may have limitations due to the nature of influenza viruses, which have an envelope. This means it may not be suitable as the reference sample for conducting an influenza-WBE study, despite successful use in other studies globally (26, 51). The direct consequence is the difficulty of determining how directly wastewater concentrations reflect the number of infected individuals (28).

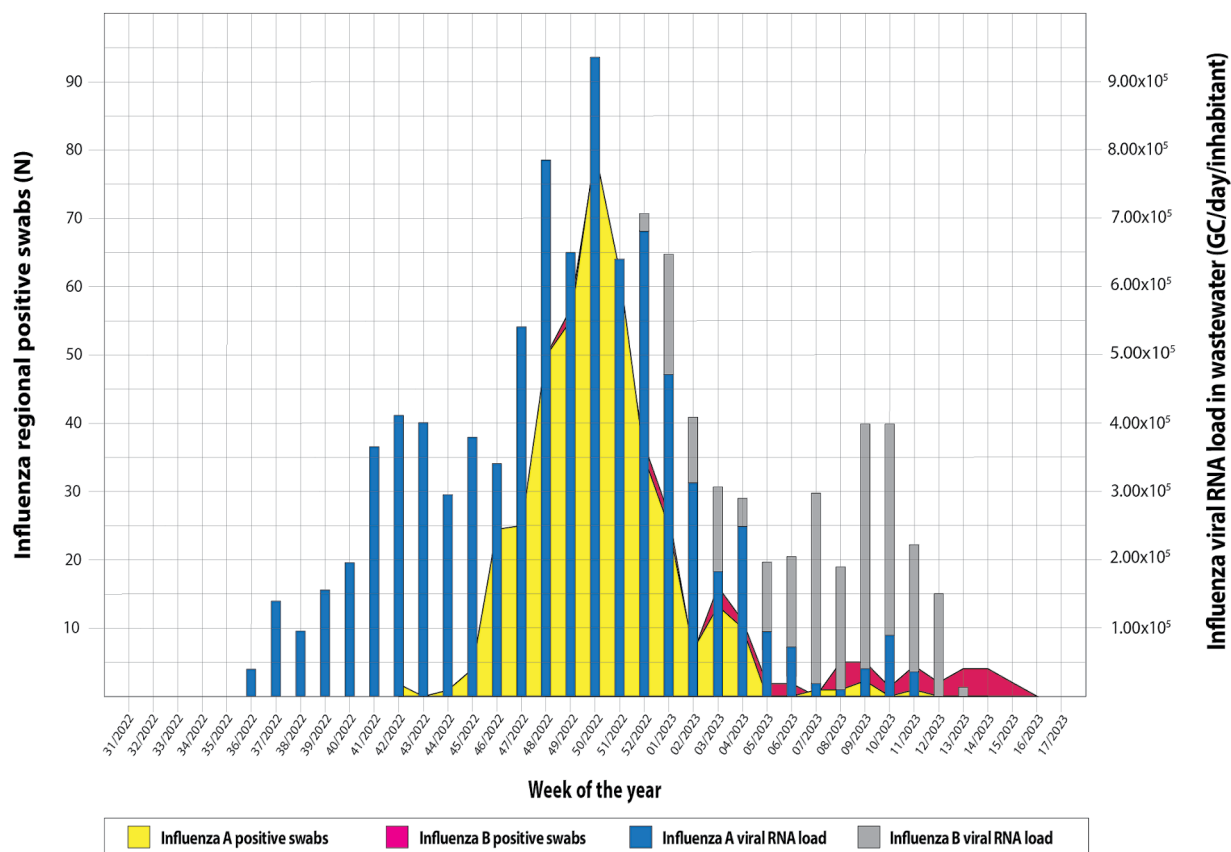


FIGURE 3

Trend of Influenza virus circulating regionally system and viral load detected from the local wastewater samples.

TABLE 4 Correlation analysis between the mean weekly RNA viral load in wastewaters and, respectively, the weekly incidence of regional ILI x 1,000 inhabitants in Sicily and the cumulative number of IAV + IBV positive swabs detected in the region at weeks 0, 1 and 2.

	Time	<i>R</i>	<i>r</i> ²	<i>p</i> -value
ILI x 1,000 inhabitants (regional) x IAV + IBV GC/day/inhabitant	w0	0.55	0.30	<0.01
	w1	0.70	0.47	<0.0001
	w2	0.78	0.61	<0.0001
<i>n</i> ° of IAV + IBV positive swabs (regional) x IAV + IBV GC/day/inhabitant	w0	0.46	0.21	<0.01
	w1	0.55	0.30	<0.01
	w2	0.63	0.40	<0.001

Wastewater-based methods can provide insight into the circulation of respiratory viruses within a specific community without testing numerous individuals. This is because a single wastewater sample represents the entire community's contribution. The results from wastewater testing can be obtained within 24 h of sample collection, providing real-time information that can be used to inform public health responses, clinical decision-making, and individual behavior modifications.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.salute.gov.it/portale/>

[influenza/dettaglioContenutiInfluenza.jsp?lingua=italiano&id=704&area=influenza&menu=vuoto&tab=5](https://www.salute.gov.it/portale/influenza/dettaglioContenutiInfluenza.jsp?lingua=italiano&id=704&area=influenza&menu=vuoto&tab=5).

Ethics statement

Ethical approval was not required for the studies involving humans because the information collected is aggregated and anonymous. The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were collected anonymously and in an aggregated way at wastewater purification plants. No information on humans was collected. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

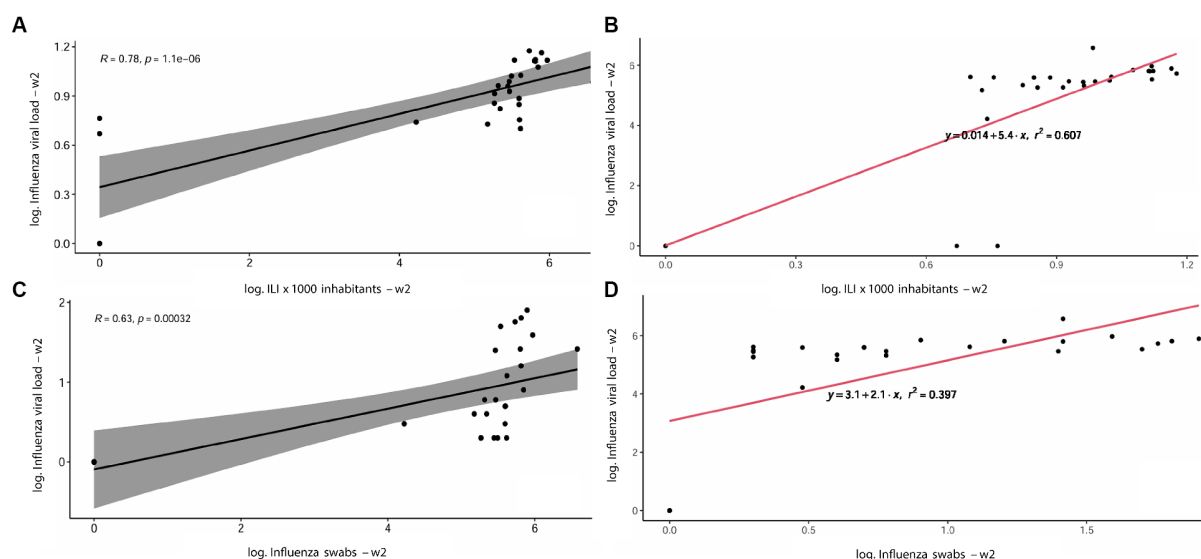


FIGURE 4

Scatterplots describing the correlation between the RNA viral load detected in wastewater (GC/day/inhabitants) and the number of ILI detected per 1,000 inhabitants at week 2 showing R and p-value (A) and r^2 (B); the correlation between Influenza viral load in wastewater and the number of positive swabs detected, with R and p-value (C) and r^2 (D).

Author contributions

CM: Conceptualization, Funding acquisition, Methodology, Validation, Writing – original draft. WM: Supervision, Writing – original draft. WP: Data curation, Formal analysis, Visualization, Writing – review & editing. RP: Data curation, Formal analysis, Visualization, Writing – review & editing. GG: Data curation, Formal analysis, Visualization, Writing – review & editing. CC: Data curation, Formal analysis, Visualization, Writing – review & editing. AR: Investigation, Resources, Writing – review & editing. GA: Investigation, Resources, Writing – review & editing. IR: Investigation, Resources, Writing – review & editing. VGi: Investigation, Resources, Writing – review & editing. FI: Investigation, Resources, Writing – review & editing. AS: Investigation, Resources, Writing – review & editing. RL: Investigation, Resources, Writing – review & editing. VGu: Investigation, Resources, Writing – review & editing. FV: Supervision, Writing – review & editing. FT: Conceptualization, Methodology, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research

was partially supported by the Sicilian Regional Government (Progetti Obiettivo PSN 2019, Linea progettuale 4, Azione 4.39, CUP: T79J21017530001).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Wang P, Xu Y, Su Z, Xie C. Impact of COVID-19 pandemic on influenza virus prevalence in children in Sichuan. *China J Med Virol.* (2022) 95:e28204. doi: 10.1002/jmv.28204
- WHO. World Health Organization guidance for surveillance during an influenza pandemic: 2017 update. Geneva: World Health Organization (2017).
- McDonald SA, Teirlinck AC, Hooiveld M, van Asten L, Meijer A, de Lange M, et al. Inference of age-dependent case-fatality ratios for seasonal influenza virus subtypes a(H3N2) and a(H1N1)pdm09 and B lineages using data from the Netherlands. *Influenza Other Respir Viruses.* (2023) 17:e13146. doi: 10.1111/irv.13146
- Vestergaard LS, Nielsen J, Krause TG, Espenhain L, Tersago K, Bustos Sierra N, et al. Excess all-cause and influenza-attributable mortality in Europe, December 2016 to February 2017. *Euro Surveill.* (2017) 22:30506. doi: 10.2807/1560-7917.ES.2017.22.14.30506
- Bonanni P, Ferro A, Guerra R, Iannazzo S, Odone A, Pompa MG, et al. Vaccine coverage in Italy and assessment of the 2012–2014 National Immunization Prevention Plan. *Epidemiol Prev.* (2015) 39:146–58.

6. Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases (NCIRD). (2022). Burden of flu. Available at: <https://www.cdc.gov/flu/about/burden/index.html>. Accessed on 09 March 2023).
7. Signorelli C, Odone A. Dramatic 2015 excess mortality in Italy: a 9.1% increase that needs to be explained. *Scand J Public Health*. (2016) 44:549–50. doi: 10.1177/1403494816649833
8. United Nations. Population division, world population prospects 2022. (2019). Available at: <https://population.un.org/wpp/>. Accessed 10 August 2023.
9. Xagorarakis I, O'Brien E. Wastewater-based epidemiology for early detection of viral outbreaks. *Women in Water Quality Mar*. (2020) 30:75–97. doi: 10.1007/978-3-030-17819-2_5
10. World Health Organization. Vaccines against influenza WHO position paper—November 2012. *Wkly Epidemiol Rec*. (2012) 87:461–76.
11. World Health Organization. (2014). Global epidemiological surveillance standards for influenza. Available at: <https://apps.who.int/iris/bitstream/handle/10665/311268/9789241506601-eng.pdf>. Accessed 02 September 2023.
12. Giacchetta I, Primieri C, Cavalieri R, Domnich A, de Waure C. The burden of seasonal influenza in Italy: a systematic review of influenza-related complications, hospitalizations, and mortality. *Influenza Other Respir Viruses*. (2022) 16:351–65. doi: 10.1111/irv.12925
13. European Centre for Disease Prevention and Control. Influenza-like illness definition. (2018). In commission implementing decision (EU) 2018/945 of 22 June 2018. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018D0945&from=EN#page=24>. Accessed on 27 August 2023.
14. Taubenberger JK, Morens DM. Influenza: the once and future pandemic. *Public Health Rep*. (2010) 125:15–26. doi: 10.1177/003335491012505305
15. Harrington WN, Kackos CM, Webby RJ. The evolution and future of influenza pandemic preparedness. *Exp Mol Med*. (2021) 53:737–49. doi: 10.1038/s12276-021-00603-0
16. Robotto A, Lembo D, Quaglino P, Brizio E, Polato D, Civra A, et al. Wastewater-based SARS-CoV-2 environmental monitoring for Piedmont. *Italy Environ Res*. (2022) 203:111901. doi: 10.1016/j.envres.2021.111901
17. La Rosa G, Iaconelli M, Mancini P, Bonanno Ferraro G, Veneri C, Bonadonna L, et al. First detection of SARS-CoV-2 in untreated wastewaters in Italy. *Sci Total Environ*. (2020) 736:139652. doi: 10.1016/j.scitotenv.2020.139652
18. Gonzalez R, Curtis K, Bivins A, Bibby K, Weir MH, Yetka K, et al. COVID-19 surveillance in southeastern Virginia using wastewater-based epidemiology. *Water Res*. (2020) 186:116296. doi: 10.1016/j.watres.2020.116296
19. Hasan SW, Ibrahim Y, Daou M, Kannout H, Jan N, Lopes A, et al. Detection and quantification of SARS-CoV-2 RNA in wastewater and treated effluents: surveillance of COVID-19 epidemic in the United Arab Emirates. *Sci Total Environ*. (2021) 764:142929. doi: 10.1016/j.scitotenv.2020.142929
20. Mao K, Zhang K, Du W, Ali W, Feng X, Zhang H. The potential of wastewater-based epidemiology as surveillance and early warning of infectious disease outbreaks. *Curr Opin Environ Sci Health*. (2020) 17:1–7. doi: 10.1016/j.coesh.2020.04.006
21. La Rosa G, Iaconelli M, Veneri C, Mancini P, Bonanno FG, Brandtner D, et al. The rapid spread of SARS-COV-2 omicron variant in Italy reflected early through wastewater surveillance. *Sci Total Environ*. (2022) 837:155767. doi: 10.1016/j.scitotenv.2022.155767
22. Maida CM, Amodio E, Mazzucco W, La Rosa G, Lucentini L, Suffredini E, et al. SARI collaboration group. Wastewater-based epidemiology for early warning of SARS-CoV-2 circulation: a pilot study conducted in Sicily, Italy. *Int J Hyg Environ Health*. (2022) 242:113948. doi: 10.1016/j.ijheh.2022.113948
23. World Health Organization. (2003). Guidelines for environmental surveillance of poliovirus circulation. Available at: www.who.int/vaccines-documents/. Accessed on 29 August 2023.
24. Mercier E, D'Aoust PM, Thakali O, Thakali O, Hegazy N, Jia JJ, et al. Municipal and neighbourhood level wastewater surveillance and subtyping of an influenza virus outbreak. *Sci Rep*. (2022) 12:15777. doi: 10.1038/s41598-022-20076-z
25. Dumke R, Geissler M, Skupin A, Helm B, Mayer R, Schubert S, et al. Simultaneous detection of SARS-CoV-2 and influenza virus in wastewater of two cities in southeastern Germany, January to May 2022. *Int J Environ Res Public Health*. (2022) 19:13374. doi: 10.3390/ijerph192013374
26. Markt R, Stillebach F, Nägele F, Kammerer A, Peer N, Payr M, et al. Expanding the pathogen panel in wastewater epidemiology to influenza and norovirus. *Viruses*. (2023) 15:263. doi: 10.3390/v15020263
27. Ahmed W, Bivins A, Stephens M, Metcalfe S, Smith WJM, Sirikanchana K, et al. Occurrence of multiple respiratory viruses in wastewater in Queensland, Australia: potential for community disease surveillance. *Sci. Total. Environ*. (2023) 864:161023. doi: 10.1016/j.scitotenv.2022.161023
28. Wolken M, Sun T, McCall C, Schneider R, Caton K, Hundley C, et al. Wastewater surveillance of SARS-CoV-2 and influenza in preK-12 schools shows school, community, and citywide infections. *Water Res*. (2023) 231:119648. doi: 10.1016/j.watres.2023.119648
29. Wolfe MK, Duong D, Bakker KM, Ammerman M, Mortenson L, Hughes B, et al. Wastewater-based detection of an influenza outbreak. *Environ Sci Technol Lett*. (2022) 9:687–92. doi: 10.1021/acs.estlett.2c00350
30. Chan M, Lee N, Chan P, Wong R, Ho W, Ngai CLK, et al. Seasonal influenza a virus in feces of hospitalized adults. *Emerg Infect Dis*. (2011) 17:2038–42. doi: 10.3201/eid1711.110205
31. Heijnen L, Medema G. Surveillance of influenza a and the pandemic influenza a(H1N1) 2009 in sewage and surface water in the Netherlands. *J Water Health*. (2011) 9:434–42. doi: 10.2166/wh.2011.019
32. Hirose R, Daidoji T, Naito Y, Watanabe Y, Arai Y, Oda T, et al. Long-term detection of seasonal influenza RNA in faeces and intestine. *Clin Microbiol Infect*. (2016) 22:813. e1–7. doi: 10.1016/j.cmi.2016.06.015
33. Italian National Statistics Institute (ISTAT). (2020). Estimated resident population—Years 2001–2019. Available at: <http://dati.istat.it/Index.aspx?QueryId=12409&lang=en>. Accessed on 8 May 2023.
34. Barrios ME, Diaz SM, Torres C, Costamagna DM, Fernandez DB, Mbayed VA. Dynamics of SARS-CoV-2 in wastewater in three districts of the Buenos Aires metropolitan region, Argentina, throughout nine months of surveillance: a pilot study. *Sci. Total. Environ*. (2021) 800:149578. doi: 10.1016/j.scitotenv.2021.149578
35. Li L, Mazurowski L, Dewan A, Carine M, Haak L, Guarini TC, et al. Longitudinal monitoring of SARS-CoV-2 in wastewater using viral genetic markers and the estimation of unconfirmed COVID-19 cases. *Sci. Total. Environ*. (2022) 817:152958. doi: 10.1016/j.scitotenv.2022.152958
36. Pillay L, Amoah ID, Deepnarain N, Pillay K, Awolusi OO, Sheena K, et al. Monitoring changes in COVID-19 infection using wastewater-based epidemiology: a south African perspective. *Sci. Total. Environ*. (2021) 786:147273. doi: 10.1016/j.scitotenv.2021.147273
37. Wu F, Xiao A, Zhang J, Moniz K, Endo N, Armas F, et al. SARS-CoV-2 RNA concentrations in wastewater foreshadow dynamics and clinical presentation of new COVID-19 cases. *Sci Total Environ*. (2022) 805:150121. doi: 10.1016/j.scitotenv.2021.150121
38. Burton J, Love H, Richards K, Burton C, Summers S, Pitman J, et al. The effect of heat-treatment on SARS-CoV-2 viability and detection. *J Virol Methods*. (2021) 290:114087. doi: 10.1016/j.jviromet.2021.114087
39. Batejat C, Grassin Q, Manuguerra J, Leclercq I. Heat inactivation of the severe acute respiratory syndrome coronavirus 2. *J Biosaf Biosecur*. (2021) 3:1–3. doi: 10.1016/j.jobb.2020.12.001
40. Wu F, Zhang J, Xiao A, Gu X, Lee WL, Armas F, et al. SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. *mSystems*. (2020) 5:e00614–20. doi: 10.1128/mSystems.00614-20
41. Ministry of Health. National Institute of Health-Integrated Influenza Surveillance System-Epidemiologic Report Influenza 2022–2023. Available at: <https://www.salute.gov.it/portale/influenza/dettaglioContenutiInfluenza.jsp?lingua=italiano&id=704&area=influenza&menu=vuoto&tab=5>. Accessed on 24 Jan 2024.
42. Maida CM, Tramuto F, Giammanco GM, Palermo R, Priano W, De Grazia S, et al. Wastewater-based epidemiology as a tool to detect SARS-CoV-2 circulation at the community level: findings from a one-year wastewater investigation conducted in Sicily. *Italy Pathogens*. (2023) 12:748. doi: 10.3390/pathogens12060748
43. Australian Government-Department of Health and Aged Care. National 2022 Influenza Season Summary. (2022). Available at: <https://www.health.gov.au/sites/default/files/2022-12/aisr-2022-national-influenza-season-summary.pdf>. Accessed on 01 July 2023.
44. Torner N, Basile L, Martinez A, Rius C, Godoy P, Jané M, et al. Assessment of two complementary influenza surveillance systems: sentinel primary care influenza-like illness versus severe hospitalized laboratory-confirmed influenza using the moving epidemic method. *BMC Public Health*. (2019) 19:1089. doi: 10.1186/s12889-019-7414-9
45. Pulicharla R, Kaur G, Brar SK. A year into the COVID-19 pandemic: rethinking of wastewater monitoring as a preemptive approach. *J Environ Chem Eng*. (2021) 9:106063. doi: 10.1016/j.jece.2021.106063
46. Sims N, Kasprzyk-Hordern B. Future perspectives of wastewater-based epidemiology: monitoring infectious disease spread and resistance to the community level. *Environ Int*. (2020) 139:105689. doi: 10.1016/j.envint.2020.105689
47. Ciannella S, Gonzalez-Fernandez C, Gomez-Pastora J. Recent progress on wastewater-based epidemiology for COVID-19 surveillance: a systematic review of analytical procedures and epidemiological modeling. *Sci. Total. Environ*. (2023) 878:162953. doi: 10.1016/j.scitotenv.2023.162953
48. Pan Y, Long L, Zhang T, Cui S, Yang P, Wang Q, et al. Potential false-negative nucleic acid testing results for severe acute respiratory syndrome coronavirus 2 from thermal inactivation of samples with low viral loads. *Clin Chem*. (2020) 66:794–801. doi: 10.1093/clinchem/hvaa091
49. Zou J, Zhi S, Su X, Kang L, Li C, Su X, et al. Heat inactivation decreases the qualitative real-time RT-PCR detection rates of clinical samples with high cycle threshold values in COVID-19. *Diagn Microbiol Infect Dis*. (2020) 98:115109. doi: 10.1016/j.diagmicrobio.2020.115109
50. Islam G, Gedge A, Lara-Jacobo L, Kirkwood A, Simmons D, Desaulniers J. Pasteurization, storage conditions and viral concentration methods influence RT-qPCR detection of SARS-CoV-2 RNA in wastewater. *Sci. Total. Environ*. (2022) 821:153228. doi: 10.1016/j.scitotenv.2022.153228
51. Zheng X, Wang M, Deng Y, Xu X, Lin D, Zhang Y, et al. A rapid, high-throughput, and sensitive PEG-precipitation method for SARS-CoV-2 wastewater surveillance. *Water Res*. (2023) 230:119560. doi: 10.1016/j.watres.2022.119560



OPEN ACCESS

EDITED BY

Jiaying Li,
The University of Queensland, Australia

REVIEWED BY

Adilya Albetkova,
Centers for Disease Control and Prevention
(CDC), United States
Ari Warokka,
Jakarta State University, Indonesia
Thomas Nyirenda,
European & Developing Countries Clinical
Trials Partnership, Netherlands

*CORRESPONDENCE

Emilie Dama
✉ ocp2@cdc.gov

RECEIVED 09 February 2024

ACCEPTED 03 June 2024

PUBLISHED 30 July 2024

CITATION

Dama E, Porgo S, Ake Y-C, Yameogo I,
Gampini S, Adjami A-GA, Nikiema A,
Kamate M, Tarbangdo F, Sawadogo R,
Sawadogo C, Ouedraogo HS, Zerbo H,
Rahalison L, Medah I, Dahourou AG,
Greco-Kone R and Ake FH (2024)
Implementation and performance evaluation
of an integrated specimen referral system in
Burkina Faso using the national courier
services (2020–2022).
Front. Public Health 12:1384382.
doi: 10.3389/fpubh.2024.1384382

COPYRIGHT

© 2024 Dama, Porgo, Ake, Yameogo,
Gampini, Adjami, Nikiema, Kamate,
Tarbangdo, Sawadogo, Sawadogo,
Ouedraogo, Zerbo, Rahalison, Medah,
Dahourou, Greco-Kone and Ake. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Implementation and performance evaluation of an integrated specimen referral system in Burkina Faso using the national courier services (2020–2022)

Emilie Dama^{1*}, Souleymane Porgo², Yahn-Cedric Ake³,
Issaka Yameogo², Sandrine Gampini⁴, Aime-Gilles A. Adjami⁵,
Abdoulaye Nikiema^{6,7}, Mory Kamate³, Felix Tarbangdo³,
Romial Sawadogo¹, Charles Sawadogo⁸, Hamed S. Ouedraogo²,
Habibata Zerbo⁹, Lila Rahalison¹⁰, Isaïe Medah²,
Anicet G. Dahourou¹, Rebecca Greco-Kone¹⁰ and Flavien H. Ake³

¹Division of Global Health Protection, Country Office, Center for Global Health, US Centers for Disease Control and Prevention, Ouagadougou, Burkina Faso, ²General Direction of Public Health, Ministry of Health and Public Hygiene, Ouagadougou, Burkina Faso, ³DAVYCAS International, Ouagadougou, Burkina Faso, ⁴World Health Organization, Ouagadougou, Burkina Faso, ⁵World Health Organization, Expanded Special Project for Elimination of NTDs (ESPEN), Ouagadougou, Burkina Faso, ⁶Integrated Quality Laboratory Services (IQLS), Ouagadougou, Burkina Faso, ⁷African Society for Laboratory Medicine (ASLM), Addis Ababa, Ethiopia, ⁸Directorate of Biomedical Laboratory, Ministry of Health and Public Hygiene, Ouagadougou, Burkina Faso, ⁹National Laboratory for Animal Health, Ministry of Agriculture, Animal Resources and Fisheries, Ouagadougou, Burkina Faso, ¹⁰Division of Global Health Protection, US Centers for Disease Control and Prevention, Atlanta, GA, United States

Introduction: In 2017, the Ministry of Health and Public Hygiene (MoH) of Burkina Faso designed and piloted a specimen transport system using the national courier services (La Poste BF) in 4 districts. Based on satisfactory performance indicators, the MoH set a vision aimed at scaling up this system to strengthen disease detection and surveillance of epidemic prone diseases across the country. This work describes the implementation process, performances, and lessons learned.

Methodology: This work describes the implementation process, performances, and lessons learned. Under the leadership of the Directorate of Population Health Protection within the MoH, a stepwise approach was used to bring together multiple partners across sectors to develop the first needed documents including a guide, an implementation plan, Standard Operating Procedures, and data collection tools. Then, the execution phase included equipment purchase, trainings, and consensus on a financing mechanism. Key indicators were defined to allow performance monitoring

Result: The integrated biological specimen referral system (SITEB) was officially launched in January 2020 to transport human biological specimens of priority diseases including COVID-19 from district level to reference laboratories nationwide. As of December 31, 2022, La Poste BF transported 168,856 packages containing 206,314 specimens from all 13 regions. 99.66% of packages were delivered in <24 h as required, and 99.68% of specimens were in good condition at reception. COVID-19 specimens represented respectively 18% and 63% of samples transported in 2020 and 2021.

Discussion: The political will combined with the experience gained during the pilot phase and the commitment and support from all stakeholders laid to the foundation of the effective implementation of this system. Collaboration

between two government entities (MoH and Minister of Transport, Urban Mobility, and Road Safety) to benefit public health has led to reasonable pricing for sustainability. Although all documents integrate the “One Health” approach, the system ensures the transport of only human samples for now. Despite security constraints, Burkina Faso has successfully set up a system using the national postal service to ensure the routine transport of specimens for all diseases under laboratory surveillance including laboratory tests for HIV and TB from the district level to reference laboratories nationwide. This system has also proved to be useful and efficient in managing public health emergency.

KEYWORDS

integrated specimen referral system, Burkina Faso, laboratory system, disease surveillance, national courier services, Ministry of Health

1 Introduction

An efficient laboratory network supported by a robust transport system for biological specimens is essential to detect, prevent, and respond effectively to public health threats (1). In low-income countries, where detection capacity is particularly low in peripheral laboratories, an efficient specimen referral system is needed to support disease surveillance and the management of public health threats. To help countries achieve this, several international organizations including the World Health Organization (WHO), the U.S. Centers for Disease Control and Prevention (US-CDC), the United Nations Dangerous Goods Programme (UN DGP), and the International Organization for Standardization (ISO) provided guidance on the implementation of specimen transport systems (2–5). Despite the many guidelines and regulations, setting up an efficient specimen transport system within a laboratory network remains challenging for many countries, particularly in resource-limited sub-Saharan Africa. The main difficulties are linked to a lack of coordination, low national funding, poor implementation of laboratory policies, poor transport services, and insecurity (6, 7). Many sub-Saharan countries have been testing various means and approaches to setting up specimen transportation systems that aim to be effective and efficient despite limited resources.

Under the auspices of the U.S. President’s Emergency Plan for AIDS Relief (PEPFAR) and the Global Health Security Agenda (GHSA), along with support from other international donors and NGOs, several countries have engaged in improving access to diagnostic services and surveillance systems using a performance specimens referral system. A hub network system based on different *ad-hoc* methods, including national postal courier services, was used in Uganda to increase access to Early HIV Infant Diagnosis (EID) services from 36% to 51%. This system also reduced transportation costs by 62% while reducing the turn-around times by 46.9% (8). With the support and technical assistance from a public-private partnership (PPP), the postal services were successfully used in Uganda and Ethiopia to strengthen the tuberculosis specimen referral system and increase referrals from presumptive multidrug resistant tuberculosis cases (9, 10). A similar increase in viral load tests, reagents used, and facilities accessing testing was noticed by Faruna et al. when a PPP was used to improve Nigeria’s national integrated specimen referral network (11). In Malawi, earlier study conducted by the National TB Control

Programme reported that peripheral units using a bus service to transport sputum to central reference laboratory for culture and sensitivity testing had a better record of specimens arriving at the CRL than those using alternative means of transport (12). While these examples have focused on transporting specimens of a specific disease, other studies have taken a more inclusive approach by integrating several diseases.

A pilot study conducted in 3 districts in Mali, included specimens from meningitis, measles, yellow fever, and polio suspected cases. This study showed that shipments of specimens from districts to the central level using the postal service was feasible and faster than public transportation. However, further analysis regarding the most efficient costing mechanism is needed (6). Inspired by the “hub” model adopted by Ethiopia and Haiti (10, 13), Guinea has developed and approved a national specimen referral policy which includes 6 diseases (Ebola, Acute flaccid paralysis, measles, yellow fever, cholera, and meningitis) using a stepwise process. The implementation of this policy has been piloted in three prefectures in Lower Guinea (14).

In 2017, a baseline assessment carried out in Burkina Faso revealed the absence of an integrated specimen transport system and highlighted the existence of fragmented disease-specific transport systems. These parallel systems were funded by different partners and used laboratory agents, increasing costs and time spent away from laboratory duties. To address this, the Ministry of Health and Public Hygiene (MoH) of Burkina Faso designed and piloted a specimen transport system using the national courier services (La Poste BF ex SONAPOST) in 4 districts under the lead of the Directorate of Population Health Protection (DPSP- Direction de la Protection de la Sante de la Population). Monitoring and evaluation of La Poste BF’s performance was deemed satisfactory, with 95% of specimens sent to the reference laboratories under the appropriate conditions in <24 h and at comparatively affordable costs (15). Based on this success, the MoH has set a vision aimed at expanding and implementing an integrated biological specimen referral system, SITEB (System Intégré de Transport des échantillons Biologiques) using La Poste BF’s services to transport all specimen types from districts to reference laboratories across the country. A stepwise approach was used to bring together multiple partners to develop a standardized specimen transport system that integrates other diseases and enhances laboratory capacity and public health infrastructure, thereby providing global health

security implementation. This paper describes the process used to implement the SITEB using La Poste BF, the system's performance after 3 years of implementation, lessons learned, and challenges.

2 Methods

2.1 Implementation process of an integrated system for specimen transport

Figure 1 summarizes the key stages in the process of implementing an integrated sample transport system, the chronology of their implementation and the main outcomes.

2.1.1 Identification of a lead department for the project implementation and set up of a technical working group

To concretize the Ministry of Health's vision, the Directorate of Population Health Protection (DPSP-Direction de la Protection de la Sante de la Population) was designated as a lead department to collaborate closely with the Directorate of Biomedical Laboratory (DLBM-Direction des Laboratoire de Biologie Medicale) on SITEB implementation. The experience gained by the DPSP during the pilot phase with the Severe Acute Respiratory Infections (SARI) sentinel surveillance with LaPoste BF (15) was an asset for this directorate in charge of the epidemiological surveillance of diseases and also the focal point for the GHSA and the International Health Regulations (IHR) in the country.

To facilitate the project's operationalization, a SITEB technical working group (SITEB-TWG) was established and formalized by the Secretary General of the MoH. This group meets quarterly or as needed and regularly invites other stakeholders.

2.1.2 Scoping meeting with partners and technical departments of interest

In low-income countries, the international partners primarily fund the transportation of specimens through several parallel systems and processes for epidemiological surveillance of most priority diseases, including meningitis, measles, influenza, dengue/arboviruses, and polio. Adopting an integrated system encompassing all specimen types across the nation necessitated the support of these partners and essential stakeholders. Some are using laboratorian technicians and other postal services or private courier services. Partners and all the MoH technical directorates involved in specimen transport were presented with the MoH's vision during this meeting. This system covers twenty-one diseases, including zoonotic, animal, and human diseases. It guarantees the transportation of around ten types of specimens from regional and health district hospitals to national and regional reference laboratories. Partners such as the US-CDC, WHO, and the African Society for Laboratory Medicine (ASLM) were represented. The key entities within the MoH involved in the specimen transport that also took part were the Directorate of Preventable Disease (DPV-Direction de la Prevention de la Maladie par la vaccination), DLBM, DPSP, and Public Health Emergency Operation Center (CORUS-Centre des Opérations de Reponse aux

Urgences Sanitaires). On behalf of the Global Fund, the Health Development Support Program (PADS-Programme d'Appui au Développement Sanitaire) represented HIV and TB programs. As the MoH did not have the capacity and dedicated staff for the overall implementation of this ambitious project, an implementing partner (Davycas International) was appointed to carry out and monitor this project, including a phasing-out plan. This partner was selected based on its expertise and its capacity to work on joint projects with multiple partners and departments within the MoH to achieve public health objectives in Burkina Faso.

2.1.3 Development of a guide and data collection tools

A national guide for implementing the SITEB was developed during workshops by the SITEB-TWG using a participatory, multisectoral, and multidisciplinary approach. It was then approved by the Ministry of Health during a validation workshop attended by the partners. This guide highlights the project context, the expected roles and responsibilities of the various stakeholders, and the requirements in terms of quality insurance, and biosafety and biosecurity associated with the specimen transport process. It also describes SOPs on the preparation, packaging, storage, shipping, and reception of packages, for each specimen type. Furthermore, the national guide for implementing the SITEB includes biosafety and biosecurity requirements on the category of specimens transported and international guidelines. Specific indicators were identified to ensure monitoring of the quality of the specimens transported and the overall performance of La Poste BF.

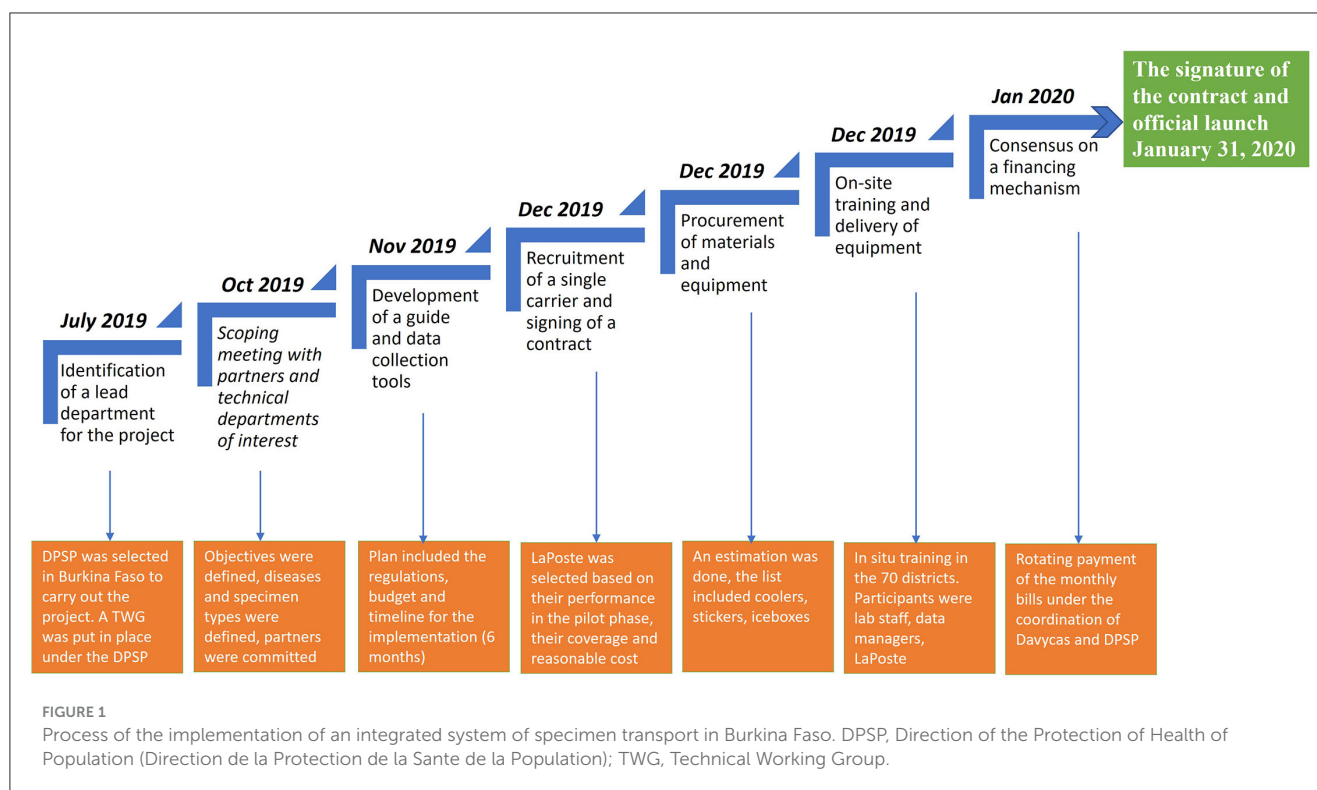
To support the implementation of this guide, job aids, and data collection tools were developed with SITEB-TWG contribution. The disease notification forms included in the SITEB have been revised to take traceability aspects into account.

2.1.4 Recruitment of a single carrier and signing of a contract

Burkina Faso chose to contract La Poste BF as a courier service to transport specimens throughout the country. This semi-private courier service had collaborated successfully with the MoH during the pilot phase and was interested in this project aiming at improving the health of the population. Another important criterion was its good geographical coverage with an office in each country province. As part of implementing the SITEB using La Poste BF, the contract of the pilot phase had been revised to include other priority diseases. However, the pricing terms did not change. Same as in the pilot phase, the shipping cost was based on the weight of the coolers (2.5 kg) and the number of packages transported with and without specimens (return of empty coolers). Collaboration between two government bodies, the MoH and the Ministry of Transport, facilitated negotiations to achieve affordable pricing.

2.1.5 Procurement of materials and equipment

A needs assessment was conducted based on the frequency of packages transported per week for each disease. Then, the quantity of each item was estimated, and the implementing partners



placed orders. Each of the 70 districts and the eight regional health districts received three packages containing the following: plastic stickers with thermic transfer printing (humidity resistant), absorbent cotton to cushion shocks and absorb liquids in the event of spills, and an infrared thermometer to record temperature at reception. To facilitate the return of the coolers, the address of each sending laboratory was printed and attached to the coolers they received. In addition to this, the address of all other possible destination laboratories was also given to each laboratory.

2.1.6 On-site training followed by delivery of equipment

The cascade training of the field agents and the handing over of the necessary equipment and support have been an important step that marked the launch of the new system. The adopted training format (region by region) gathering both field agents and those of the La Poste BF was conducted in each of the 13 regions of the country. Before the training, the SITEB-TWG developed modules covering the description of the SITEB, the role, and responsibility of the field agents, the standard operating procedures (SOPs), and related support documents. The module on the presentation of the SOPs provided details on the categorization and identification of infectious substances, the triple packaging, and the transportation and biosecurity considerations based on international resource documents such as the WHO Guidance on regulations for the transport of infectious substances 2015–2016, Laboratory Safety Manual, Third Edition, WHO 2005. In addition, guidance documents from the US-CDC and the national safety guide for medical laboratories were used. A frequently asked questions (FAQ) sheet was also developed to help trainers provide

harmonized answers. Trainers were mainly SITEB-TWG members and La Poste BF agents.

To ensure the engagement and ownership of the leaders at the national level, a briefing session was organized for regional health directors and the heads of districts. This was followed by two-day training of data managers, laboratory technicians, and La Poste BF staff in all districts. People trained were clinicians, human, animal, and environmental laboratory staff and La Poste BF's transporters.

2.1.7 Meeting with partners to define a financing mechanism

The partners' commitment was obtained from the start to implement SITEB, but an agreement on the financing strategy still needed to be established. Implementing a single mechanism was challenging because partners have various financial management requirements. The MoH and international partners provided the financing mechanism by establishing an annual commitment contract which includes the monthly payment schedule of the invoices. Based on the quantity of packages transported, a monthly bill is produced by La Poste BF and sent to the implementing partners and the lead department under the MoH. It was decided at the start of the year, that each partner would inform the DPSP of the number of monthly postal invoices it can handle in a year, regardless of specimen type or amount.

2.1.8 The signature of the contract and launch

The signing of the agreement between La Poste BF and the MoH was followed by an official launch chaired by the MoH and attended by the regional health directors and the district

chief medical officers with media coverage. Finally, a note on the implementation of SITEB signed by the MoH's General Secretary was disseminated.

2.2 Monitoring and evaluation mechanisms

Monitoring and evaluation (M&E) have been an essential component of the implementation of SITEB using both papers based (Table 1) and electronic data collection platform. M&E aspects have been integrated throughout the system from the case notification, specimen collection, transport, receipt, and biological results reporting. A unique labeling system with barcode stickers is assigned to each specimen to facilitate tracking.

An electronic System for Tracking of Epidemiological Data and Laboratory Specimens (STELab-System de tracabilite des donnees epidemiologiques et de laboratoire), is also used for the visualization, validation, reporting, and management of data. STELab is a web-interfaced electronic platform for case-based surveillance data entry (16). It allows the real-time recording of surveillance and laboratory data on priority diseases as well as the tracking information of a package. Its primary role was to track meningitis laboratory specimens (16). Because of the excellent results of this system, its new version has been extended to all specimens under the SITEB. Thus, today the STELab platform includes 24 diseases including zoonotic and vaccine preventable diseases. Key indicators were identified to monitor the performance of the SITEB using La Poste BF:

- *Percentage of packages delivered within 24 h*: this indicator is calculated from the time of pickup of the package from the site. The denominator is the total number of packages picked up.
- *Percentage of packages delivered in good condition*: the package at the reception does not present any non-conformity (correct packing and label).
- *Percentage of specimens delivered in good condition*: the specimens at the reception were not in good condition (Good temperature, packaging).

Data are regularly pulled from the STELab platform to produce a SITEB quarterly bulletin that is disseminated to all districts and stakeholders including partners. This bulletin summarizes the performance of La Poste BF during the reported period and since the implementation of the SITEB. The target was 100% for each of indicator.

2.3 Data collection and analysis

Data were collected on the STELab platform and were cross referenced with data collected by La Poste BF. Data were confirmed, and all discrepancies were corrected. Excel software was used to develop tables and conduct all analyses presented in this manuscript.

The comparison of pre- and post- SITEB data has not been possible as there was no coordinated system in place before the implementation of the SITEB allowing centralization of data

and monitoring of indicators. The performance evaluation used indicators and target percentage.

3 Results

3.1 Key outcomes from the process of SITEB implementation

The different activities implemented before the effective start of the SITEB led to key outcomes that are critical for such a system. First, there is the development of the national guide for implementing the SITEB and an operational plan of SITEB including all the SOPs. Agreement has been reached to include the following diseases in the system: Severe Acute Respiratory illness including COVID-19, dengue/arthroviruses, rotavirus, norovirus, measles, tuberculosis (TB), meningitis, and Human Immuno deficiency Virus (HIV). The specimen types that SITEB can transport include Nasopharyngeal (NP) and Oropharyngeal (OP), serum, stool, sputum, pleural fluid, bronchoalveolar puncture fluid, pus, urine, and Cerebrospinal Fluid (CSF). Based on WHO guidance on regulations for transporting infectious substances (4), all selected pathogens are categorized as class 6.2 (Infectious Substances), category B. In line with this classification, the following wording has been taped to each cooler “UN 3373, *Biological substance, Category B.*”

Figure 2 describes the specimen transport circuit in the healthcare pyramid. Specimens are transported from all the 70 districts to the national reference laboratories which include the reference laboratories for meningitis, Influenza, norovirus/rotavirus, viral hemorrhagic fevers, HIV, and TB, in addition to the immunization department that received specimens of measles and poliomyelitis.

3.2 Monitoring of key indicators of SITEB performance

Packages were transported from 70 districts to the national reference laboratories from all 13 country's regions. In addition to the national reference laboratories, HIV, and TB specimens were also sent to the national level laboratories since the viral load testing and TB testing are decentralized and some regions don't have the testing capacity. Ouagadougou and Bobo-Dioulasso are Burkina Faso's two largest cities, hosting all the national reference laboratories. National Reference Laboratories (NRL) for antimicrobial resistance and viral hemorrhagic fevers are in Bobo-Dioulasso and the remaining are in Ouagadougou (Influenza, meningitis, measles, rotavirus HIV, and TB).

From January 31, 2020 to December 31, 2022 La Poste BF transported 16,858 packages from the district level to the NRL and national level laboratories across the country. Among them, 99.66% (16,800/16,858) were delivered in <24 h as required in the contract with La Poste BF. Only 0.05% (9/16,856) of packages transported were found damaged during the transport. The breakdown per year shows that 14.72% (2,481) of packages were transported in 2020 against 42.30% (71,310) and 42.98% (7,246) in 2021 and

TABLE 1 The physical data collection tools and the levels of the system where they are available and completed.

Data collection tools	Description
Record of specimen package shipments (collection sites)	In addition to data collected outside the sampling site, it collects the: - date/time of collection of the package by La Poste BF package number, - specimen or sticker number - laboratory signature - signature of La Poste BF
Examination request form (collection sites)	- Name of the prescriber - Date and time of the specimen collection - Requesting department - Examination requested. - Reason for the request and a space for the results
Individual notification form/case investigation (collection sites)	In general, the individual notification form/case investigation includes: - socio-demographic data - clinical information - sampling data - transport data and laboratory results
Summary sheet for tracking specimens (collection sites)	The summary monitoring sheets for certain specimens (sputum, specimens of animal origin) contain the name of the sampling site, a list of all the specimens contained in the cooler, and the transport data.
Package delivery form (La Poste BF/Sender, Recipient)	Issued by La Poste BF, it collects the: - date and time of collection of the package - Name of the sender - Package number and the references of the La Poste BF agent. It is signed by the senders and recipients, including the drivers, at each change of hands in order to ensure traceability.
Delivery form (La Poste BF/Sender, Recipient)	Similar to that of La Poste BF, it is used in areas where La Poste BF's services are temporarily unavailable. It is issued to the person delivering the package upon arrival.
Incident management register (La Poste BF/Sender, Recipient)	It contains the date and time of the incident; describes the type of incident (case of spillage, loss or theft of coolers, etc.), the people to contact.
Record of receipt of specimens by the laboratory/reference site. (All receiving sites)	- date and time of arrival, - the conformity of the package, - the package number, - the signature of the laboratory and La Poste BF In case of rejection of non-compliant specimens, reasons are specified

Collection sites: Health centers, Direction of preventable diseases, Direction of Animal National Laboratory, National reference laboratories.

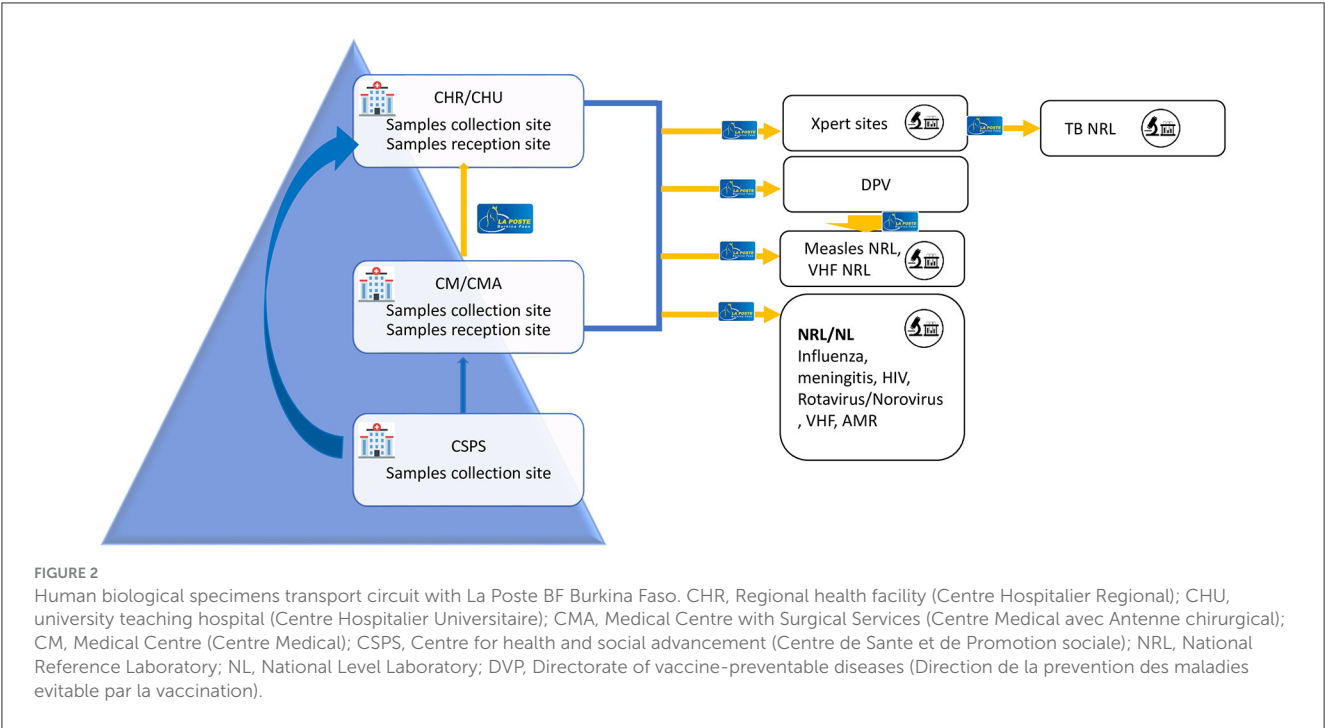


TABLE 2 Evaluation of the performance of SITEB.

Indicators	2020	2021	2022	Total
Packages delivered	2,481	7,131	7,246	16,858
Packages delivered in <24 h	2,437	7,121	7,242	16,800
	98.23%	99.86%	99.94%	99.66%
Packages delivered in good conditions	2,481	7,127	7,241	16,849
	100%	99.94%	99.93%	99.95%
Specimens delivered	29,731	117,818	58,765	206,314
Specimens delivered in good condition	29,477	117,514	58,666	205,657
	99.15%	99.74%	99.83%	99.68%

2022 respectively. No packages were reported missing or lost during transportation.

During the reporting period, 206,314 specimens were transported of which 14.41% (29,731) in 2020, 57.11% (117,818) in 2021, and 28.48% (29,731) in 2022 (Table 2). The breakdown of specimens transported by disease and year reveals that in 2020 and 2022, HIV specimens were predominant with 59% and 53% respectively, while in 2021 COVID-19 specimens accounted for 63% of all specimens transported by SITEB. No specimen of acute flaccid paralysis was transported during the first year of the implementation of the SITEB whereas, in 2021 and 2022, respectively 553 (0.4%) and 2,393 (4.07%) specimens were transported (Figure 3). We didn't find any significant differences in the number of specimens transported in on year from another.

Packages (with and without) specimens were transported from all the country's 13 regions. Overall, the number of packages transported increased between 2020 and 2022. The region with the highest number of packages is the Center region, with more than 10,000 specimens, followed by the Southwest, Hauts-Bassins, and Center West regions with more than 1,000 packages picked and delivered by La Poste BF over the reporting period. The regions where La Poste BF transported fewer packages are the Sahel (38), Plateau Central (223), Cascades (231), Center East (304), and North (348) (Figure 4).

3.3 Contract and rates

Like during the pilot phase, the SITEB contract was signed directly between the DPSP representing the MoH and La Poste BF representing the Ministry of Transport. The contract stipulates that La Poste BF is responsible for collecting the packages containing category B biological specimens from the public health district laboratories, delivering them to the recipient laboratory, and returning the empty triplicate packages to the sending health establishments. Two important elements of this contract are the description of the commitments of both parties and the pricing. The contract is structured on an escalating scale, with the unit price per package decreasing by 500F CFA (\$ 0.84) as the number of packages to be transported increases. The minimum amount per package is 2,500 FCFA (~\$ 4) and the maximum is 4,000 CFA (~\$

7) (Table 3). It is important to note that this contract is still flexible and does not provide a comprehensive list of diseases.

4 Discussion

SITEB is a disease non-specific system that harmonizes the transport of human biological specimens as part of national epidemiological surveillance and laboratory tests for HIV VL/EID, TB using the national postal system known as La Poste BF. To our knowledge, Burkina Faso is the first country in West Africa to implement such an integrated specimen transport system using postal services. This paper presents the stepwise process used to set up an integrated specimen transport system and its performance after 3 years of implementation. Monitoring key indicators over the 3 years of the SITEB implementation has shown the satisfactory performance of the transport of all types of human specimens from the district level throughout the country by the postal service.

This project was born of political will, followed by a clear vision of the MoH. Regulatory texts or policies must govern the specimen referral system in a country to enable effective intra- and inter-sectoral collaboration and optimization of support from partners. While some countries have developed specific policies to comply with this requirement (14), Burkina Faso, through its framework document for the development of biomedical laboratories and optimization of biological diagnosis, has clearly defined its vision about specimen transport, and listed in the same document the strategies to achieve this goal. The vision and the definition of the country's objectives in regard to specimen transport system prompted the development and validation of a national guide for the implementation of an integrated specimen transport system by the SITEB TWG based on the One Health approach.

The excellent country-wide coverage of La Poste BF's services enabled specimens to be transported to all 13 regions of Burkina Faso. However, the accessibility of some security-challenged areas due to terrorist attacks forced La Poste BF to limit its presence. This situation has led to population movements within the country, and the closure of health facilities, thereby limiting the population's access to healthcare (17, 18). In Burkina Faso, the regions most affected by the humanitarian crisis are the Sahel, Center-North, Nord, Est, and Boucle du Mouhoun. Although not on a continuous and systematic basis, La Poste BF has been transporting specimens from functional

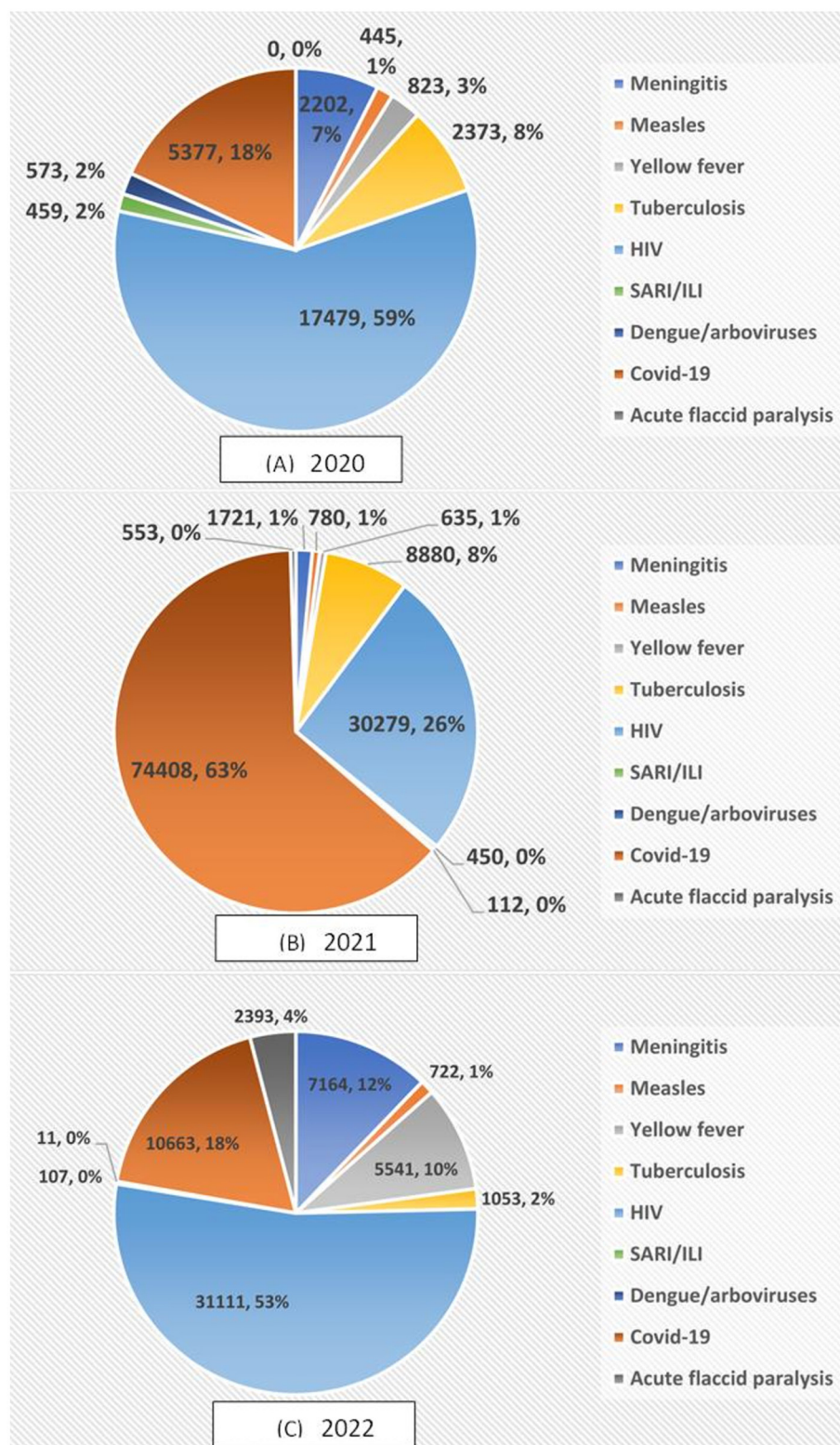


FIGURE 3

Distribution of the number of specimens transported by SITEB per disease in (A) 2020, (B) 2021, and (C) 2022.

districts in these security-challenged regions since the launch of SITEB. When this proved impossible for security reasons, other strategies were developed and deployed. An additional

factor to explain the considerable diversity in the number of packages per region is the COVID-19 crisis. More than half of the packages transported came from the central region, which

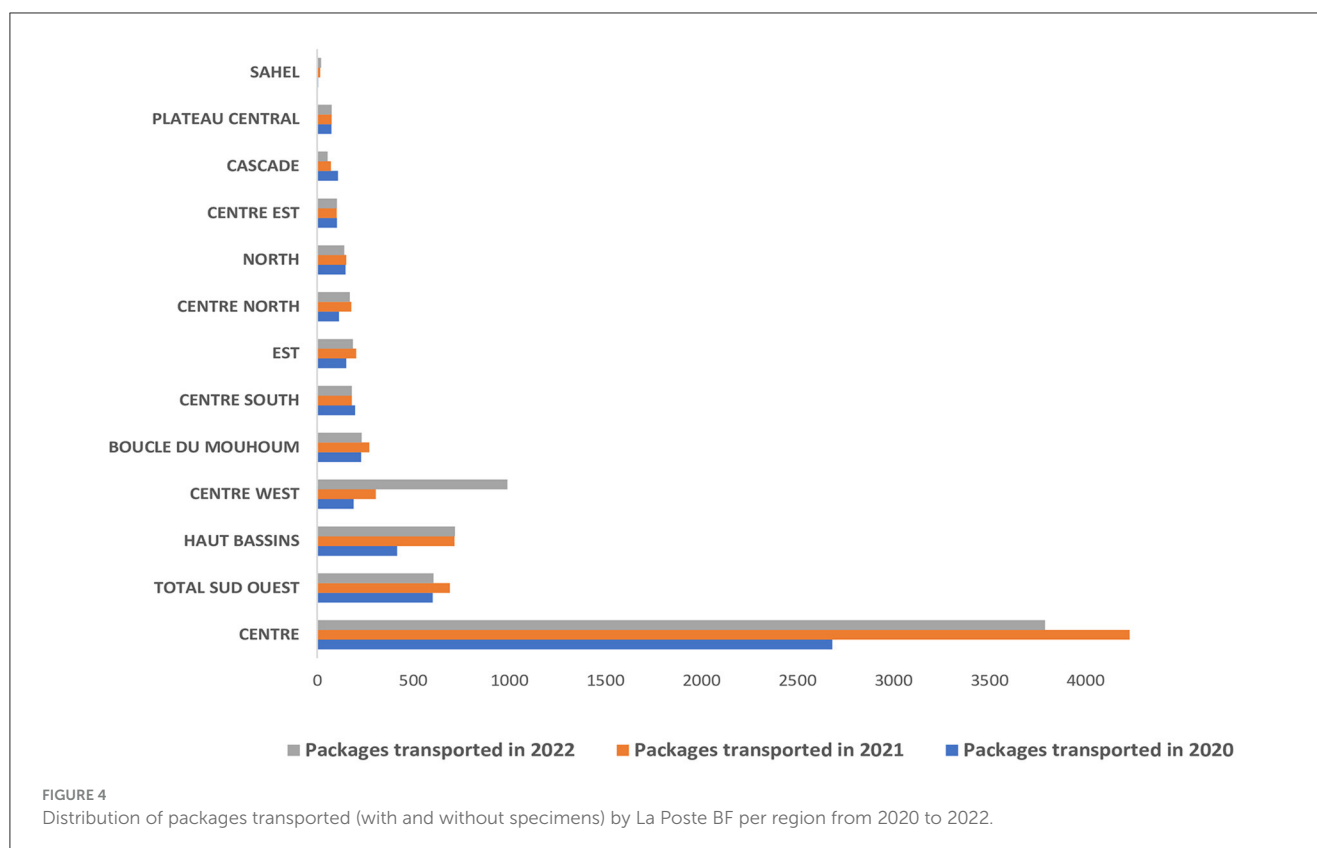


TABLE 3 Price list of LaPoste Burkina Faso for the transport of specimens.

Quantity/range	Unit amount; TTC F CFA	Minimum amount; TTC F CFA (~\$)	Maximum amount; TTC F CFA (~\$)	Observations
[001–500]	4 000	2,000,000 (3,229)	2,000,000 (3,229)	Fixed
[501–1,000]	3 500	2,003,500 (3,235)	3,750,000 (6,055)	Fixed price + Nb of packages * unit cost applied from the 501 st package
[1,001–1,500]	3 000	3,753,000 (6,060)	5,250,000 (8,477)	Fixed price + No. of packages * unit cost applied from the 1,001 st package
[1,501–3,000]	2 500	5,252,500 (8,480)	9,000,000 (14,531)	Fixed price + Nb of packages * unit cost applied from the 1,501 st package
[3,001 et +]	2 000	9,002,000 (14,545)	—	Fixed price + No. of packages * unit cost applied from the 3,001 st package

was the epicenter of COVID-19, followed by Hauts-Bassins region with 15,712 and 3,517 cases detected between 2020 and 2022 representing more than 70% and 15% of the total cases respectively (19).

From the launch of SITEB in January 2020 to December 2022, out of 16,858 packages transported, 99.66% were delivered within 24 h from pick-up time at the collection site (the required turnaround time), compared to 77% during the pilot phase (15). This result shows a significant improvement in the post office's performance, dispelling initial fears about its ability to meet this challenge. To carry out its mission by the agreed upon contract, La Poste BF has signed an agreement with several public transport

companies in the country's main cities, hired additional staff and procured logistical resources. A similar pilot study in Mali showed that only 46% of specimens transported by public transport system were delivered within the required timeframe (72 h), compared to 71% of specimens transported by Mali's postal service specific means of transportation. The same study found a comparable percentage of specimens delivered in good conditions between the two types of transport (6). Indeed, the public transport network in Burkina Faso is diversified and well organized, with regular departures to major cities. While transporting biological specimens in public transport vehicles can be perceived as a risk, the triple packaging and extra protection provided by La Poste

BF help to further minimize the risk of exposure to potentially dangerous pathogens contained in the specimens. Using drones or unmanned aircraft System technology is being explored by some studies to transport specimens, vaccines, and other laboratory supplies. However, a cost-effectiveness analysis of the use of these new technologies which integrates all considerations (including security) must be conducted (20–22).

The flexibility of the contract provisions to permit the integration of additional specimens or adjustments during their term is one of the system's features and success. In 2021, this encouraged the incorporation of COVID-19 and acute flaccid paralysis specimens. COVID-19 had not yet been declared a pandemic by the World Health Organization (WHO) at the time of the SITEB's launch, and the country intended to keep using the traditional system for transporting acute flaccid paralysis specimens because the disease was considered on its way out. During the 1st months of COVID-19 in Bobo-Dioulasso, there was only one laboratory in the country capable of performing diagnosis and it was in Bobo Dioulasso. MoH vehicles transported specimens from suspected COVID-19 in other regions to Bobo-Dioulasso. However, as the number of cases increased across the country, specimen transportation became difficult due to logistical issues. COVID-19 specimens were integrated into the SITEB without any changes to the initial contract or pricing after several meetings and briefing sessions with the post office.

There was little to no significant difference in the number of packages that La Poste BF transported in 2021 and 2022. However, the number of specimens transported was twice higher in 2021 due to COVID-19 specimens (63%) which were transported by the dozen in a single package. Indeed, the peak of COVID-19 cases was notified in 2021 between January and February with more than 69,000 COVID-19 samples tested. It is worth noting that this number includes those of COVID-19 suspected cases but also samples collected from international travelers who are required to test. The drop in COVID-19 cases and the implementation of vaccination in June 2021 explains the decrease in COVID-19 specimens collected in 2022. The decrease in the number of specimens transported by SITEB for other diseases such as meningitis, yellow fever, SARI/ILI, and dengue/arbovirus between 2020 and 2021 can be attributed to the impact of COVID-19 on the health system in general and on disease surveillance in particular, as documented in numerous studies (22, 23). Several initiatives and actions were implemented to re-energize disease surveillance, which had been slowed by COVID-19, and improvements were seen in late 2021 and 2022.

In most developing countries, disease surveillance, including specimen transport, is funded by international partners. To minimize the risk of this system collapsing due to a lack of resources, particular emphasis was placed on negotiating rates. Rates were negotiated between the Ministry of Health and La Poste BF to ensure that the country would be able to meet costs in the event of a reduction or cessation of partner support. During the pilot phase, which only involved 4 districts, the cost of transporting a package by La Poste BF was around 28 USD (17,500 CFA), when the system was extended, the cost was negotiated to ~6 USD (3,500 CFA) per package, almost 5 times cheaper. The estimated cost for implementation of the SITEB in Burkina Faso is approximately 662,000 USD (400,000,000 CFA) which

includes meetings, equipment, trainings, and document printing. The average cost of the La Poste BF's monthly bill is ~23,528 USD (14,215,762 CFA). The cost-benefit analysis of such a system using postal services or hub systems is still a gap in many studies (23).

5 Lessons learned

The experience of Burkina Faso provides important lessons and recommendations that must be considered to ensure the successful development and implementation of an efficient and integrated specimen transport system. The following are key lessons learned from this experience:

- A strong political will is essential to engage partners and stakeholders.
- It is critical to select an efficient operator (public or private) capable of providing services throughout the country.
- Contract flexibility is essential so that, in addition to surveillance and clinical diagnosis, the system can be used in response to epidemics or other public health events.
- Throughout the process, sustainability, and a multi-sector approach (One Health) must be considered.
- A good monitoring and evaluation plan must be developed to ensure that the system runs smoothly and to allow assessing performance and impact of the system.
- To ensure specimen transport in insecure areas where government offices and health facilities are not operational, an innovative strategy must be developed.
- It is important to consider an implementation and coordination partner with dedicated staff to ensure smooth implementation while ensuring a phase-out.

6 Challenges and perspectives

The main limitation of the SITEB is the non-integration of animal and environment specimens in this system. Although, the guide and all data collection tools have been revised according to the One Health approach, the implementation must still be effective. Discussions are ongoing to make this happen. In the clauses of the current contract La Poste BF picks up specimens from the district level while there is no formal system in place to transfer specimens from peripheral level to district level. A pilot phase is underway in 2 regions where La Poste BF picks specimens from the peripheral level to extend it to the entire country after an evaluation and a revision of the contract. Finally, although the partners have put in place a mechanism for paying monthly postal bills, the SITEB's operating costs (SITEB-TWG meetings, supervision, equipment replacement, document printing) still need to be included, and there needs to be a government budget line to support the operation of this system.

7 Conclusion

The involvement of stakeholders at all levels, as well as partners, contributed to the success of this innovative system. Furthermore,

the success and lessons learned from the pilot phase (15) have made a significant contribution to laying the foundations of this integrated system, which is now widely used in the surveillance of priority diseases in Burkina Faso, as well as in the management of health crises. Several countries have attempted to use national mail services to transport biological specimens to strengthen surveillance of a country's set of priority diseases or specific diseases such as tuberculosis and/or HIV (10). However, Burkina Faso is one of the countries that has successfully implemented a national mail service for a specimen transport system, which considers all the diseases under laboratory-based surveillance and covers the whole country down to the district level. The performance of the SITEB after 3 years of implementation made it a major pillar in laboratory-based surveillance of priority diseases in Burkina Faso. It ensures the transport of all specimens collected for surveillance purposes including VIH and TB from district level across the country. Thanks to its flexibility, it also plays an important role in the management of public health emergencies for an early detection and quick response. The integration of animal specimens remains a big gap, but efforts are underway to address this.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

ED: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. SP: Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Visualization, Writing – review & editing. Y-CA: Data curation, Formal analysis, Methodology, Visualization, Writing – review & editing. IY: Conceptualization, Methodology, Project administration, Supervision, Writing – review & editing. SG: Funding acquisition, Methodology, Resources, Writing – review & editing. A-GA: Conceptualization, Methodology, Supervision, Writing – review & editing. AN: Conceptualization, Funding acquisition, Methodology, Writing – review & editing. MK: Data curation, Methodology, Visualization, Writing – review & editing. FT: Data curation, Methodology, Visualization, Writing – review & editing. RS: Methodology, Validation, Writing – review & editing. CS: Funding acquisition, Methodology, Resources, Writing – review & editing. HO: Resources, Supervision, Validation, Writing – review & editing. HZ: Methodology, Validation, Writing – review & editing. LR: Funding acquisition, Methodology, Writing – review & editing. IM: Conceptualization, Methodology, Writing – review & editing. AD: Conceptualization, Methodology, Validation, Writing – review & editing. RG-K: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. FA: Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was funded by the U.S. Centers for Disease Control and Prevention (through the African Society for Laboratory Medicine Grant number 6 NU2GGH001847 and Jhpiego Grant number 6 NU2HGH000007), the West African Health Organization (WAHO), and the WHO.

Acknowledgments

We thank the SITEB-TWG members who played a critical role in this project, the DPSP for the leadership, the DLBM, the CRL, and the personnel of all laboratories at the national, regional and district level, LaPost BF staff, the Centers for Disease Control and Prevention (US-CDC), the African Society for Laboratory Medicine, Davycas International, Jhpiego, and WHO for financial and technical support.

Conflict of interest

Y-CA, MK, FT, and FA were employed by DAVYCAS International. AN was employed by Integrated Quality Laboratory Services (IQLS).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AA declared a shared affiliation with the authors ED, RS, LR, AD, and RG-K to the handling editor at the time of review.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The findings and conclusions of this article are those of the authors and do not necessarily represent the official position of the Centers For Disease Control and Prevention.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2024.1384382/full#supplementary-material>

References

- Mangal CN, Maryogo-Robinson L. Leveraging the laboratory response network model for the global health security agenda. *Biosecur Bioterror*. (2014) 12:274–83. doi: 10.1089/bsp.2014.0039
- Schneider F, Maurer C, Friedberg RC. International organization for standardization (ISO) 15189. *Ann Lab Med*. (2017) 37:365–70. doi: 10.3343/alm.2017.37.5.365
- Centers for Disease Control and Prevention. *Biosafety in Microbiological and Biomedical Laboratories, 6th edition*. (2020). Available online at: <https://www.cdc.gov/labs/BMBL.html> (accessed September 28, 2023).
- World Health Organization. *Guidance on Regulations for the Transport of Infectious Substances 2021–2022*. (2021). Available online at: <https://www.who.int/publications/i/item/9789240019720> (accessed February 25, 2021).
- International Air Transport Association. *Dangerous Goods Regulations, 64th edition*. (2022). Available online at: <https://www.iata.org/contentassets/b08040a138dc4442a4f066e6fb99fe2a/dgr64-addendum1-en-01-january-2023-addendum-1-english.pdf> (accessed December 15, 2022).
- Kassambara H, Nana ML, Samassa F, Traore MD. Sample transport optimization: mali pilot study. *Health Secur*. (2020) 18:S92–S97. doi: 10.1089/hs.2019.0061
- Kploanyi EE, Kenu J, Atsu BK, Opere DA, Asiedu-Bekoe F, Schroeder LF, et al. An assessment of the laboratory network in Ghana: a national-level ATLAS survey (2019–2020). *African J Labor Med*. (2023) 12:1844. doi: 10.4102/ajlm.v12i1.1844
- Kiyaga C, Sendagire H, Joseph E, McConnell I, Grosz J, Narayan V, et al. Uganda's new national laboratory sample transport system: a successful model for improving access to diagnostic services for Early Infant HIV Diagnosis and other programs. *PLoS ONE*. (2013) 8:e78609. doi: 10.1371/journal.pone.0078609
- Joloba M, Mwangi C, Alexander H, Nadunga D, Bwanga F, Modi N, et al. Strengthening the tuberculosis specimen referral network in Uganda: the role of public-private partnerships. *J Infect Dis*. (2016) 213 Suppl 2:S41–6. doi: 10.1093/infdis/jiw035
- Kebede Y, Fonjungo PN, Tibesso G, Shrivastava R, Nkengasong JN, Kenyon T, et al. Improved specimen-referral system and increased access to quality laboratory services in Ethiopia: the role of the public-private partnership. *J Infect Dis*. (2016) 213 Suppl 2:S59–64. doi: 10.1093/infdis/jiv576
- Faruna T, Akintunde E, Odelola B. Leveraging private sector transportation/logistics services to improve the National Integrated Specimen Referral Network in Nigeria. *Busin Manage Dynam*. (2019) 82019:08–20.
- Harries AD, Michongwe J, Nyirenda TE, Kemp JR, Squire SB, Ramsay AR, et al. Using a bus service for transporting sputum specimens to the Central Reference Laboratory: effect on the routine TB culture service in Malawi. *Int J Tuberc Lung Dis*. (2004) 8:204–10.
- Louis FJ, Osborne AJ, Elias VJ, Buteau J, Boncy J, Elong A, et al. Specimen referral network to rapidly scale-up CD4 testing: the hub and spoke model for Haiti. *J AIDS Clin Res*. (2015) 6:1000488. doi: 10.4172/2155-6113.1000488
- Standley CJ, Muhayangabo R, Bah MS, Barry AM, Bile E, Fischer JE, et al. Creating a national specimen referral system in Guinea: lessons from initial development and implementation. *Front Public Health*. (2019) 7:83. doi: 10.3389/fpubh.2019.00083
- Dama E, Nikiema A, Nichols K, Bicaba BW, Porgho S, Greco Kone R, et al. Designing and piloting a specimen transport system in Burkina Faso. *Health Secur*. (2020) 18:S98–S104. doi: 10.1089/hs.2019.0068
- Diallo AO, Kiemtore T, Bicaba BW, Medah I, Tarbangdo TF, Sanou S, et al. Development and implementation of a cloud-based meningitis surveillance and specimen tracking system in Burkina Faso, 2018. *J Infect Dis*. (2019) 220:S198–205. doi: 10.1093/infdis/jiz376
- Ozer P, Dembele A, Yameogo SS, Hut E, de Longueville F. The impact of COVID-19 on the living and survival conditions of internally displaced persons in Burkina Faso. *World Dev Perspect*. (2022) 25:100393. doi: 10.1016/j.wdp.2022.100393
- Druetz T, Browne L, Bicaba F, Mitchell MI, Bicaba A. Effects of terrorist attacks on access to maternal healthcare services: a national longitudinal study in Burkina Faso. *BMJ Global Health*. (2020) 5:e002879. doi: 10.1136/bmjgh-2020-002879
- Ministry of Health and Public Hygiene of Burkina Faso. Rapport de situation sur l'épidémie de la maladie à Coronavirus (COVID-19) au Burkina Faso. In: *Centre des Operations de Reponse aux Urgences Sanitaires*. Available online at: <https://www.sante.gov.bf/accueil>
- Whelen AC, Becker SJ, Uluiviti VR, Maddox N. Insights in Public Health: leveraging Pacific laboratories to boost global health security. *Hawai'i J Med Public Health*. (2016) 75:389–92.
- Balajee SA, Arthur R, Mounts AW. Global health security: building capacities for early event detection, epidemiologic workforce, and laboratory response. *Health Secur*. (2016) 14:424–32. doi: 10.1089/hs.2015.0062
- Global Laboratory Initiative. *Gli Guide to TB Specimen Referral Systems and Integrated Networks*. (2018). Available online at: <https://www.stoptb.org/gli-guide-to-tb-specimen-referral-systems-and-integrated-networks>
- Gebregergs GB, Sinishaw MA, Shiferaw MB, Antife T, Assefa M, Fiseha D, et al. Evaluation of the postal service for referral of specimen of drug resistance tuberculosis in Amhara region, Ethiopia; mixed method. *Afr Health Sci*. (2021) 21:619–27. doi: 10.4314/ahs.v21i2.17



OPEN ACCESS

EDITED BY

Jie Huang,
Southern University of Science and
Technology, China

REVIEWED BY

Chulwook Park,
Seoul National University, Republic of Korea
Yali Zheng,
Xiamen University, China

*CORRESPONDENCE

Cam Bowie
✉ cam.bowie1@gmail.com

RECEIVED 09 March 2024

ACCEPTED 23 August 2024

PUBLISHED 09 September 2024

CITATION

Bowie C and Friston K (2024) A follow up
report validating long term predictions of the
COVID-19 epidemic in the UK using a
dynamic causal model.
Front. Public Health 12:1398297.
doi: 10.3389/fpubh.2024.1398297

COPYRIGHT

© 2024 Bowie and Friston. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

A follow up report validating long term predictions of the COVID-19 epidemic in the UK using a dynamic causal model

Cam Bowie^{1*} and Karl Friston²

¹Retired, Somerset, United Kingdom, ²Queen Square Institute of Neurology, University College London, London, United Kingdom

Background: This paper asks whether Dynamic Causal modelling (DCM) can predict the long-term clinical impact of the COVID-19 epidemic. DCMs are designed to continually assimilate data and modify model parameters, such as transmissibility of the virus, changes in social distancing and vaccine coverage—to accommodate changes in population dynamics and virus behavior. But as a novel way to model epidemics do they produce valid predictions? We presented DCM predictions 12 months ago, which suggested an increase in viral transmission was accompanied by a reduction in pathogenicity. These changes provided plausible reasons why the model underestimated deaths, hospital admissions and acute-post COVID-19 syndrome by 20%. A further 12-month validation exercise could help to assess how useful such predictions are.

Methods: we compared DCM predictions—made in October 2022—with actual outcomes over the 12-months to October 2023. The model was then used to identify changes in COVID-19 transmissibility and the sociobehavioral responses that may explain discrepancies between predictions and outcomes over this period. The model was then used to predict future trends in infections, long-COVID, hospital admissions and deaths over 12-months to October 2024, as a prelude to future tests of predictive validity.

Findings: Unlike the previous predictions—which were an underestimate—the predictions made in October 2022 overestimated incidence, death and admission rates. This overestimation appears to have been caused by reduced infectivity of new variants, less movement of people and a higher persistence of immunity following natural infection and vaccination.

Interpretation: despite an expressive (generative) model, with time-dependent epidemiological and sociobehavioral parameters, the model overestimated morbidity and mortality. Effectively, the model failed to accommodate the “law of declining virulence” over a timescale of years. This speaks to a fundamental issue in long-term forecasting: how to model decreases in virulence over a timescale of years? A potential answer may be available in a year when the predictions for 2024—under a model with slowly accumulating T-cell like immunity—can be assessed against actual outcomes.

KEYWORDS

dynamic causal model, COVID-19 mitigation measures, acute-post COVID-19, hospital admissions, mortality incidence

Background

Dynamic causal modelling (DCM) stands apart from most modelling in epidemiology by predicting *mitigated* outcomes and quantifying the uncertainty associated with those outcomes (1–3). This contrasts with quantitative epidemiological forecasts that do not consider the effect of prevalence on sociobehavioral responses. Usually, epidemiological projections are over few weeks—and rest upon fitting curves to the recent trajectory of various data; e.g., (4). DCM considers what is most likely to happen based upon a generative model that best explains all the data available. This mandates a model of sociobehavioral responses that mitigate viral transmission, such as social distancing, lockdown, testing and tracing, etc. In turn, this requires a detailed consideration of how various sorts of data are generated. For example, it has to model fluctuations in testing capacity and sampling bias due to people self-selecting when symptomatic. The advantage of this kind of modelling is that any data generated by the model can be used to inform the model parameters that underwrite fluctuations in latent states, such as the prevalence of infection. Latent states refer to those states of the population that cannot be estimated directly and have to be inferred from observable data.

In October 2022, the predictions carried out 12 months earlier using a Dynamic Causal model were assessed and found to underestimate the waves of new COVID-19 infections in the period October 2021 to October 2022 by 43%, deaths by 20%, tests by 24%, hospital admissions by 31% and long COVID by 21% (5). This method of modelling besides predicting health outcomes can also estimate changing characteristics of the epidemic, such as the properties of viral transmission, immunity induced by vaccine or infection, and the propensity to leave home thereby increasing the risk of catching the infection. We concluded that the underestimation of predictions could be explained by the arrival of the Omicron variants and the changes in public health policies in the UK (6–8).

This paper is a sequel to the previous paper which, besides seeking to validate the previous 12-month predictions, makes predictions to October 2023. It sets out to assess the underlying properties of the epidemic during that period from October 2022 to October 2023. It also seeks to predict what will happen in the 12 months to October 2024 assuming the current properties of the epidemic remain as they are in October 2023. We take the opportunity to provide predictions under priors based upon recent empirical estimates of latent, incubation and infectious periods. In 2024, the accuracy of predictions should speak to the usefulness of constraining parameter estimates with informative (empirical) priors of this sort.

This article can be read as a technical report, following up on previous reports, in which certain predictions were made. We anticipate a follow-up report evaluating the predictions made in this article over the forthcoming year, which will also provide an overall synthesis of long-term forecasting with dynamic causal modelling. This report provides the opportunity to compare long-term forecasts with what actually happened over timescales of years. We therefore take the opportunity to compare the predictions and actual outcomes quantitatively. Crucially, this comparison is in the latent state space of the causes of epidemiological (and behavioral) measurements. In other words, because we are using a generative or forward model of the epidemic, we can revisit the predicted fluctuations in time-dependent epidemiological and behavioral parameters *in the light of post-hoc estimates using the same model*. This

effectively identifies where prior assumptions about key time-dependent parameters were not endorsed by empirical outcomes. This may be useful for future modelling initiatives along these lines.

Methods

Dynamic causal models

The dynamic causal model (DCM) used in this research has been continually updated with data as the epidemic has unfolded. It is designed to allow modification of model parameters, such as transmissibility of the virus, changes in social distancing, and vaccine coverage—to accommodate changes in population dynamics and virus behavior. A recent model (26th September 2023) was used to explore the effect of changing transmission of the various Omicron variants and the likely seasonal effect of the coming winter. One modification was tightening the constraints on changes in antibody immunity over time. The potential benefit of a successful Find, Test, Trace, Isolate and Support scheme was also incorporated into the model.

General and specific features of DCMs

The general and specific features of Dynamic Causal Models have been described in our previous publication (9). Since October 2022 our DCM COVID-19 model has been updated 20 times with the recent update on 26th September 2023 (10).

Data sources and assumptions

16 of the 24 data sources used in the model and in our previous report have been discontinued ([Supplementary Table S1](#)):

- UKHSA COVID-19 data dashboard (11)
 - Deaths within 28 days of COVID-19 infection – June 2023
 - Critical care bed admissions – May 2023
 - Hospital occupancy of COVID-19 cases – May 2023
 - COVID-19 antibody tests – October 2022
- Office of National Statistics (12)
 - Deaths by age – July 2023
 - Vaccinations by age – July 2023
- UK Government dashboard - Mobility – April 2022
- Google mobility Report (13)– October 2022
- IHME estimate of Incidence (confirmed and non-confirmed cases) – April 2023 (14)

The UK Government COVID-19 dashboard still provides eight key input variables such as confirmed cases, hospital admissions, certified deaths, tests and vaccine coverage (11). The Office of National Statistics (ONS) discontinued the Coronavirus (COVID-19) infection survey in March 2023 (15) which had provided the best estimates of incidence using routine antibody tests and symptom questionnaires on a regular basis to a random population sample.

The trend in the use of non-pharmaceutical interventions by the UK government is measured using the Oxford Tracker stringency index (7). The incidence of long COVID is calculated using the findings of a

global meta-analysis of post-acute COVID-19 syndrome (with defined clusters of self-reported symptoms occurring 3 months after initial infection) which found the risk of long COVID following symptoms in the community is 7.9%, in hospital admissions is 27.9% and ARDS (acute respiratory distress syndrome) is 41.4% (16). The image of the proportion of variants in circulation used in Figures K-T is taken from Our World in Data (17) which uses data sourced from Gisaid (18).

For the predictions to October 2024, it is assumed that mitigation efforts such as improved ventilation in schools and workplaces will not take place, that lockdown will not be re-imposed, and that no new more virulent variants will arrive.

Model priors

To predict outcomes over the next year, the model was run using the latest available data and prior estimates used by the DCM dashboard (19). To address the predictive validity of empirical priors we ran the model to furnish predictions with changes to the prior estimates of the model parameters, where recent research suggests appropriate values. These empirical priors were as follows: prior time constant for the latent period is 5.5 days and for the incubation period is 6.5 days in line with the results of a recent meta-analysis (20). The infectious period is given a prior time constant of 4.3 days in line with a recent paper (21), Table 1 [mean growth phase 1.6 days; mean decline phase 2.7 days]. [Supplementary Table S2](#) provides a comparison of the priors that maximize model evidence and the new (empirical) priors.

For completeness, three scenarios were modelled to identify the likely effect of improving the Find Test Trace Isolate Support (FTTIS) system from a baseline of 25% effective to 40 and 60% effective.

Findings

Comparing projected with actual COVID-19 deaths, cases, tests, hospital admissions and incidence of long COVID

Last year's projections overestimated incidence three-fold, confirmed cases two-fold, deaths and tests by 1.4 times, hospital

admissions by 2.2 times and long COVID by 2.7 times (Table 1). The actual estimates of incidence and long COVID are only available for the first half of the year but the overestimates will still be substantial.

The reasons for the overestimations are found in the following two sets of graphics, which compare various outcome and parameters of model results made in October 2022 with and without knowledge of the course of the epidemic over the recent 12-month period to October 2023 (Figures 1–5). In other words, we were able to compare the time course of key epidemiological parameters estimated with and without the data covering that period (from October 2022 until October 2023). The discrepancy between these predicted and post-dictive estimates provides one account of the overestimates above.

The key overestimate was the projected large spike of infections over the winter period of 2022/2023 which did not materialize (Figure 1, top graph). Instead, we had continuous spikes of infection at lower numbers than in the previous year (Figure 1, bottom graph). The winter wave was predicted to be accompanied by large numbers of deaths and hospital admissions which did not materialize (Figures 2–3). In short, the predicted winter wave was much greater than what transpired, partly due to a projected high level of mobility (i.e., contact rates) (Figure 4) and despite a sustained level of immunity (Figure 5).

To understand the overestimates, one can look at the trajectory of the time-dependent parameters used for both predictions (Figures 6–10). The post-hoc or post-dictive estimates showed a tiny reduction but starting at a much longer starting point of 4.4 as compared to 2.8 days in the latent period (Figure 6). The incubation period, however, was longer than originally anticipated, falling not to 1.94 days but only to 4.6 days from a starting point of 5.1 as compared to 2.1 days (Figure 7). Transmission strength had increased from each infected person infecting 1 in 3 contacts to infecting 80% of contacts (Figure 8). What may also be key is the change in expected antibody persistence, falling in the original from 197 to 159 days but assumed to remain constant in the late model with a posterior prior value of 105 days (Figure 9). Another key difference is the less than expected rise in the proportion of people leaving their homes, for example with only 30% of the older adult leaving home as compared to 60% in the

TABLE 1 Cumulative numbers of COVID-19 cases, deaths, tests, hospital admissions and post COVID-19 Syndrome – 1st February 2020 – 1st October 2023 and 12 month projected numbers for 1st October 2023–2024 – UK.

Scenario assuming FTTIS is 25% effective	DCM 2022 projection	Actual	Data source	DCM 2023 projection
Cumulative totals from 1st February 2020 to	1st October 2023	1st October 2023		1st October 2023 to 1st October 2024
Estimated incidence	485,603,813	131,242,140	IHME - 1 Apr 2023	40,692,662
Confirmed cases by PCR and LFT	53,409,837	24,743,787	Our World in Data - 30 Sep 2023	524,351
Deaths within 28 days of a positive PCR test	330,957	229,765	Our World in Data - 30 Sep 2023	24,100
Tests (both PCR and LFD)	821,181,901	602,512,524	UK Covid-19 dashboard - 30 Sep 2023	14,080,675
Hospital admissions	1,867,580	862,553	UK Covid-19 dashboard - 30 Sep 2023	175,303
Post COVID-19 Syndrome	4,726,602	1,734,000	ONS Infection survey - 30 Mar 2023	3,139,699

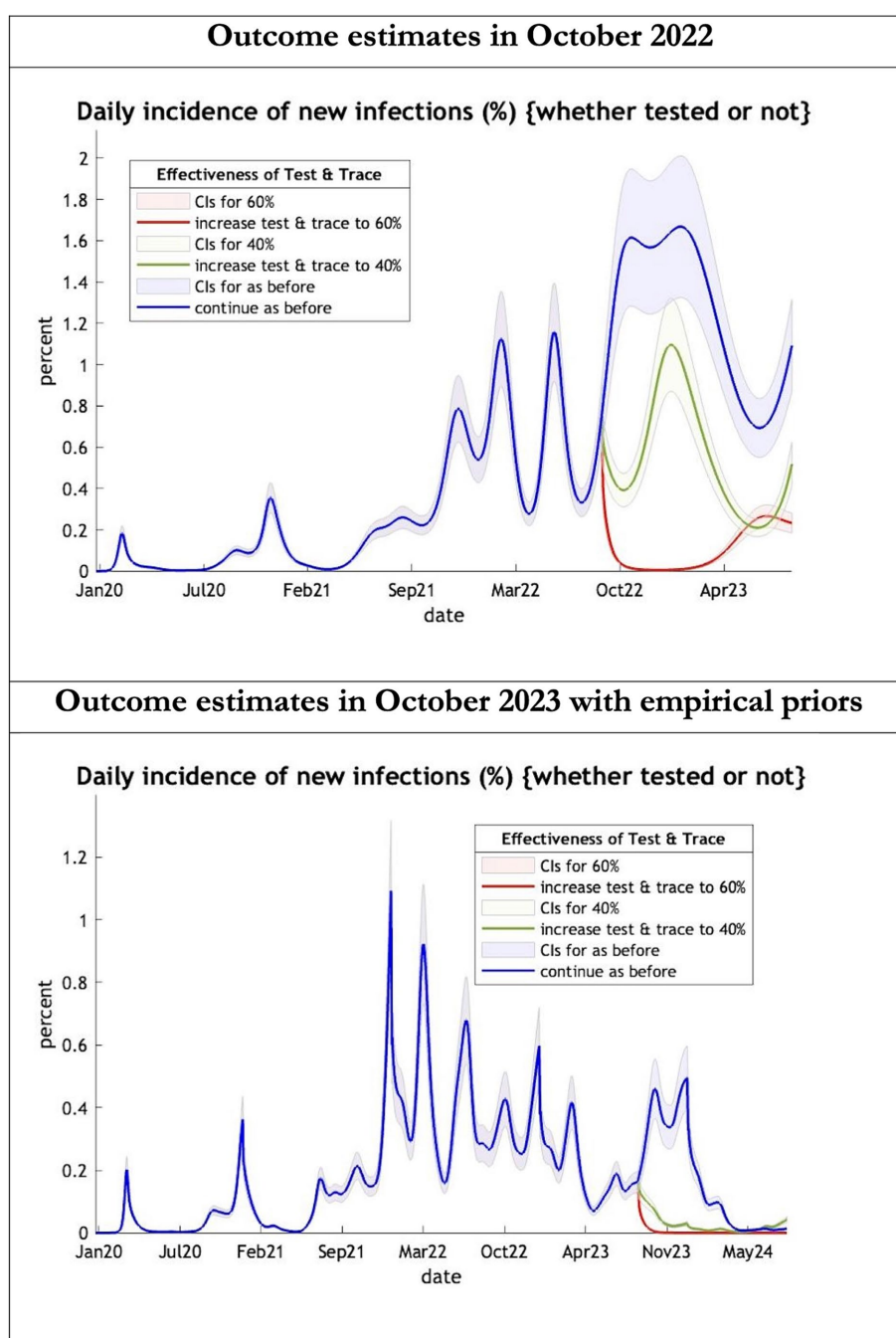


FIGURE 1

Epidemic curves of COVID-19 incidence from Jan 2020 – UK estimated by a DCM on two occasions (October 2022 and October 2023). The model can estimate incidence including cases not tested; each figure offers three projections: blue if the contact tracing system remains at 24% effective, green if it improves effectiveness to 40% and red to 60% from 1st October 2022 in the top graph and 1st October 2023 in the bottom graph: CIs, 90% credible intervals. Interpretation: The predictions with October 2022 priors are more than double the predictions using empirical priors in October 2023.

earlier model (Figure 10). Unfortunately, the empirical data stopped at the start of the 12-months under review so we cannot be sure of the actual level of movement. By March 2023 18% of people were still wearing face masks outside and 11% in public transport (22) and 14% of adults avoided contact with vulnerable people, so it is likely that mobility increased but did not return to pre-pandemic levels.

Future predictions

For the period October 2023 to October 2024 the model was used to predict the cumulative effect of the epidemic on case numbers, deaths, tests, hospital admissions and long COVID (Table 1 and Figures 1–3). The predictions using empirical priors suggest a wave this coming winter but with few deaths and tests but still plenty of

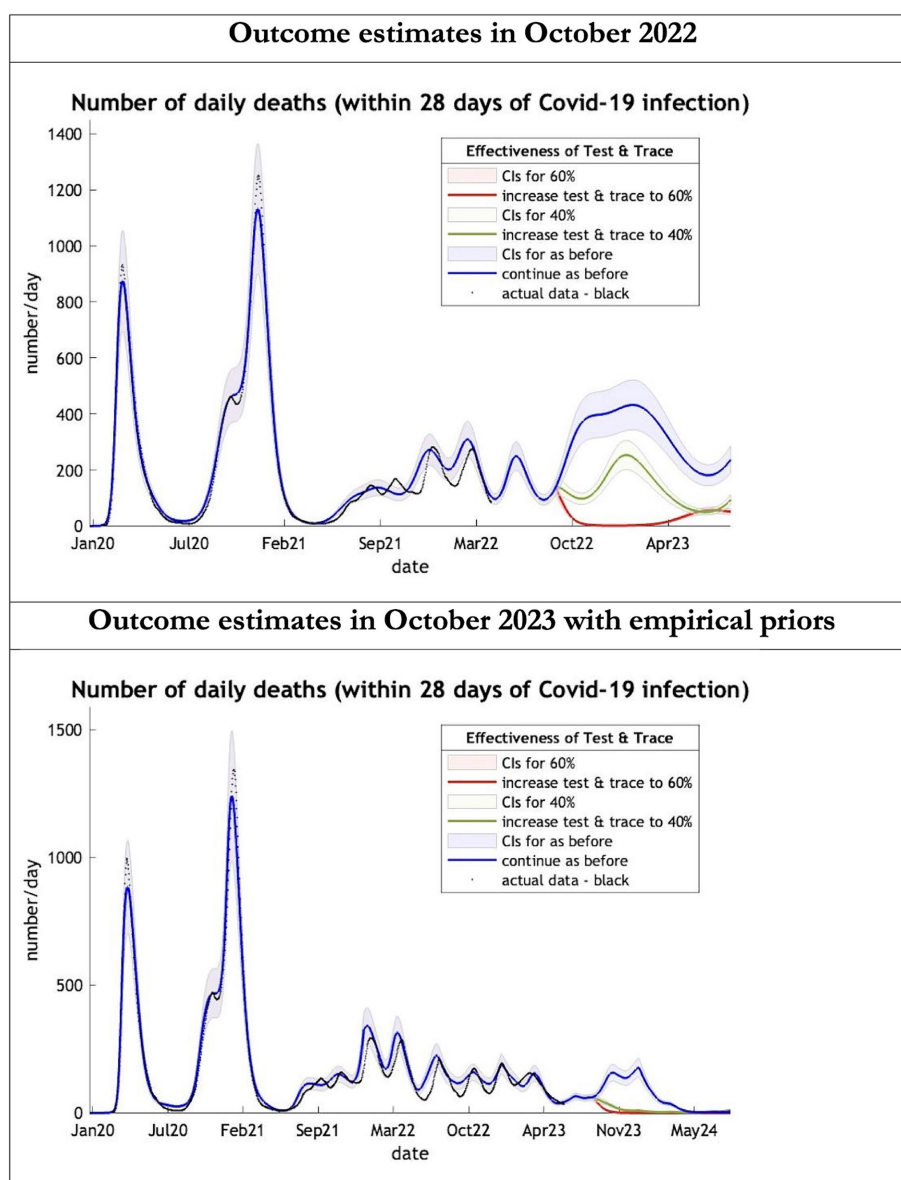


FIGURE 2

Epidemic curves of COVID-19 mortality from Jan 2020 – UK estimated by a DCM on two occasions (October 2022 and October 2023). The model can estimate projections of daily mortality certified as occurring within 28 days of a positive COVID-19 test; actual data in black is shown up till 16 June 2023—the last day of available data; each figure offers three projections: blue if the contact tracing system remains at 24% effective, green if it improves effectiveness to 40% and red to 60% from 1st October 2022 in the top graph and 1st October 2023 in the bottom graph: CIs, 90% credible intervals. Interpretation: The model is able to ape the empirical mortality series closely; the model with empirical priors offers a prediction which is half the 2022 estimates in the Oct 22 to Oct 23 period.

hospital admissions and long COVID patients. Under the empirical priors COVID-19 cases will fall but still over 40 million cases and over 3 million long-COVID cases will occur in next the 12-month period. The effect of a more efficient Test and Trace system would have little influence in reducing cases using either set of priors (Figures 2, 3).

Discussion

The overestimates of the 12-month projections to October 2023 seem to relate to better retained immunity from previous

infections and vaccines at the same time as a reduction in the trend of the new variants becoming more infectious. The reason the predicted large winter wave did not occur probably relates to these factors plus a more than anticipated caution by individuals in leaving home (i.e., exposing themselves to higher transmission risk). We have no way of assessing how many infections did actually occur because the ONS infection study was stopped and estimates from other models were discontinued. Tests became infrequent and not freely available, but many particularly older adult people still observed isolation periods when thought to be infected despite pressure to ignore such practices and the removal of legal sanctions in February 2022. The year also saw

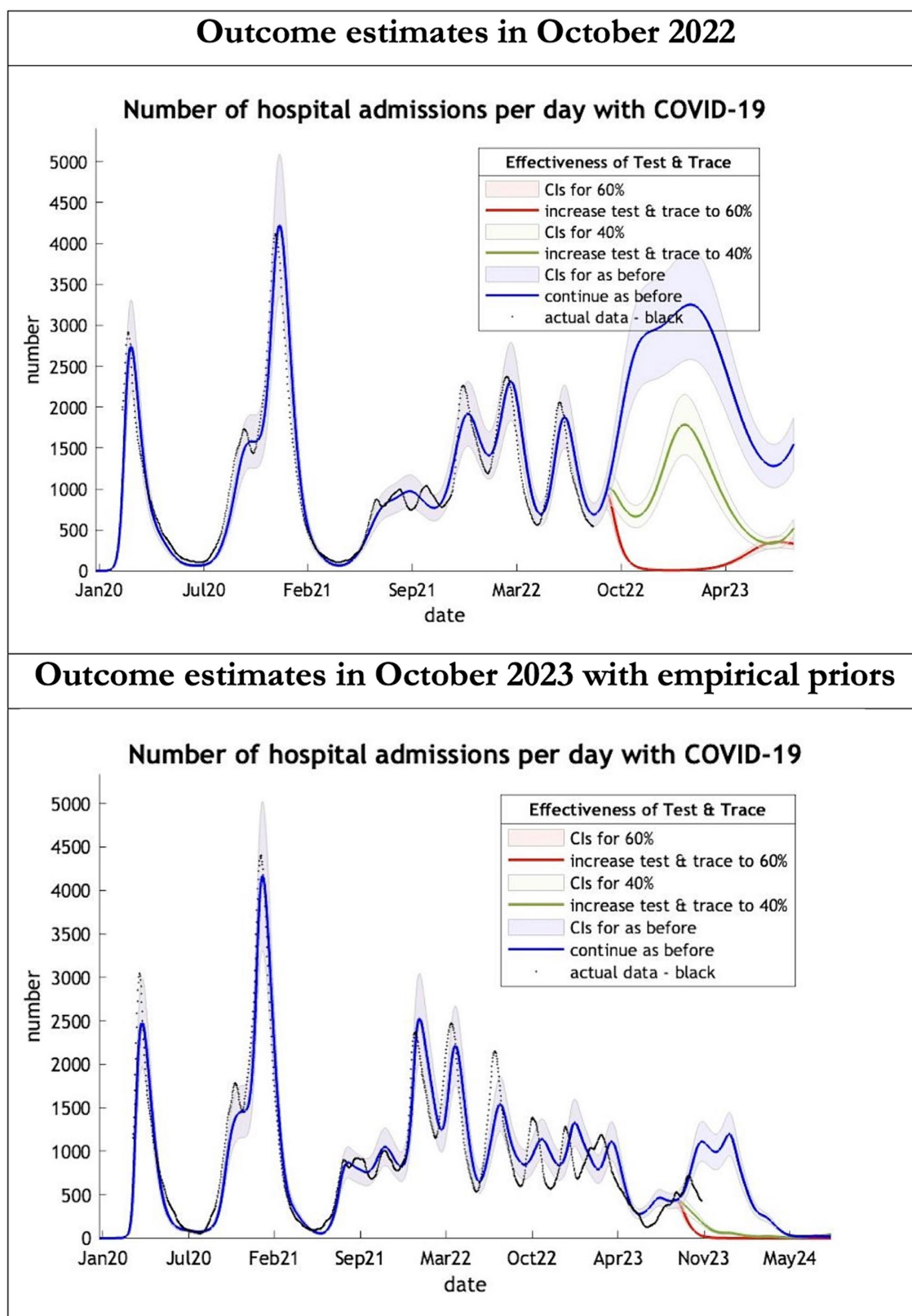


FIGURE 3

Epidemic curves of COVID-19 hospital admissions from Jan 2020 – UK estimated by a DCM on two occasions (October 2022 and October 2023). The model estimates number of hospital admissions; actual data in black; each figure offers three projections: blue if the contact tracing system remains at 24% effective, green if it improves effectiveness to 40% and red to 60% from 1st October 2022 in the top graph and 1st October 2023 in the bottom graph: CIs, 90% credible intervals. Interpretation: The 2022 estimates follow the available actual data closely until August 2022 and predicted a much larger admission rate than what occurred later. The 2023 predictions with up-to-date priors got the admission rate more or less right.

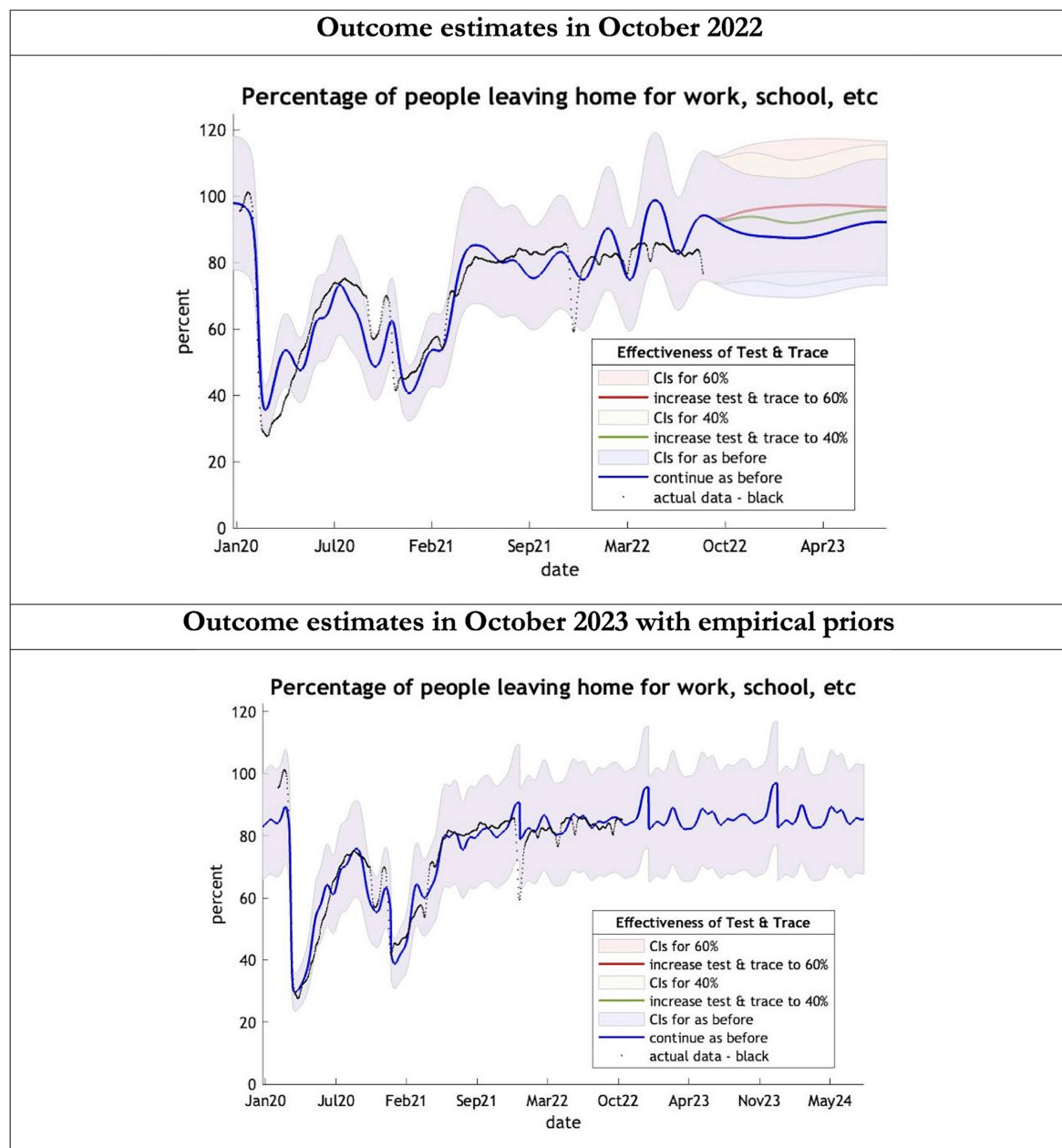


FIGURE 4

Epidemic curves of COVID-19 mobility from Jan 2020 – UK estimated by a DCM on two occasions (October 2022 and October 2023). The model estimates the number of people leaving home each day; actual data in black taken from Google Global Mobility Report; the top figure offers three projections: blue if the contact tracing system remains at 24% effective, green if it improves effectiveness to 40% and red to 60% from 1st October 2022 in the top graph and 1st October 2023 in the bottom graph: CIs, 90% credible intervals. Interpretation: The model is able to ape the actual data with exceptions in Dec 2021. The empirical 2023 priors model is able to moderate the swings in estimates seen in the model using the 2022 priors.

antiviral therapies improve associated with a drop in case fatalities.

Finally, we have specified predictions for the upcoming year, until October 2024 based on empirical priors over the successive periods of infection. It will be interesting to see whether these empirical priors improve the model's predictive validity.

In the next of these technical reports, we will use the current and previous reports as documentary evidence of predictions to assess the predictive accuracy of dynamic causal modelling over

a forecasting timescale of weeks, months and years. We anticipate doing this by adopting the final structure of the generative model but estimating epidemiological and behavioral parameters from limited timeseries—up until a certain point in time—and assessing the posterior predictive accuracy at a series of points in the future, as the pandemic evolved. This may provide a useful reference for future pandemic modelling that leverages the unprecedented amount of data and insights generated by the COVID pandemic.

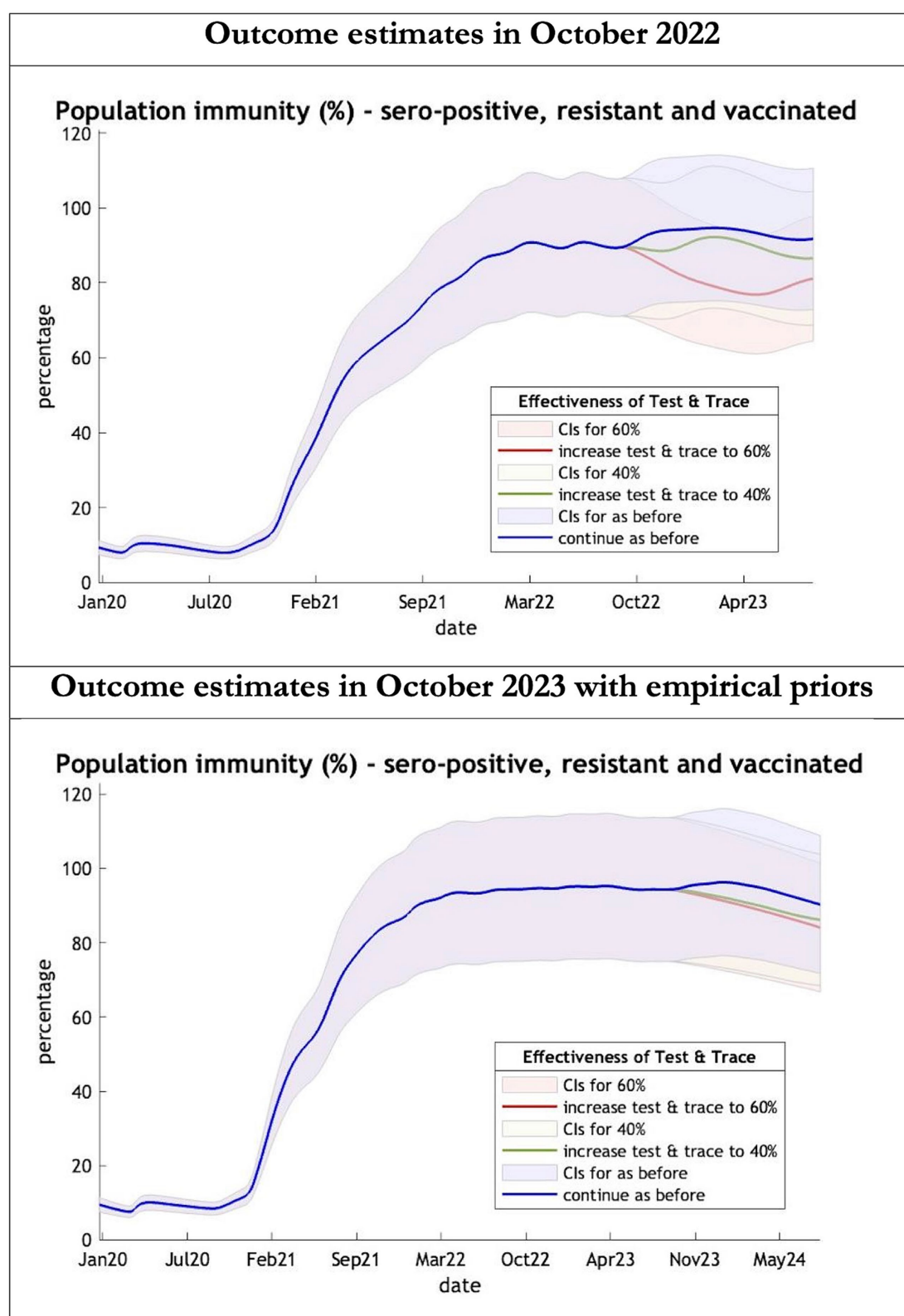


FIGURE 5

Epidemic curves of population immunity to COVID-19 from January 2020 – UK estimated by a DCM on two occasions (October 2022 and October 2023). The model estimate of population immunity to COVID-19 (% of population) including that induced by infection, natural resistance and immunization; each figure offers three projections: blue if the contact tracing system remains at 24% effective, green if it improves effectiveness to 40% and red to 60% from 1st October 2022 in the top graph and 1st October 2023 in the bottom graph: CIs, 90% credible intervals. Interpretation: Both models share similar estimates of population immunity. Neither have been able to take into account the probable declining virulence over years found in pandemics with novel viruses.

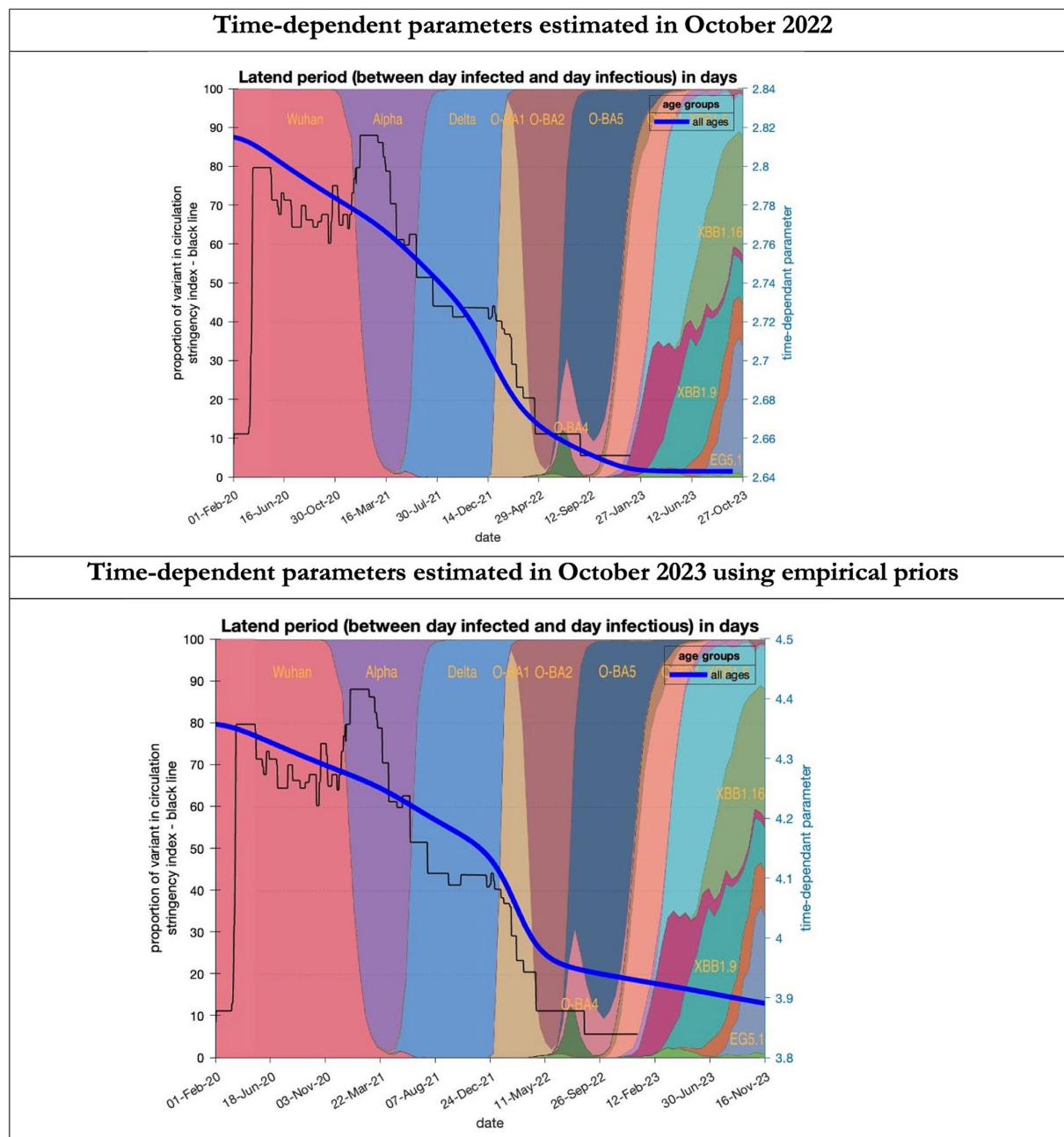


FIGURE 6

Changing estimates of latent period of COVID-19 infection in relation to the emergence of new variants and changes in response to public health policies: UK February 2020 to October 2023. Latent period (between day infected and day infectious) is measured as time constant for all age groups combined; prior in top graph of 3 days with initial model estimate of 2.8 days (infected period - Supplementary Table S2) dropping to 2.64 by October 2023; prior in the bottom graph of 5.5 days with initial model estimate of 4.36 dropping to 3.9 by November 2023; stringency index dropping from 80% in March 2020 to 5% by December 2021; proportion of variant in circulation as backdrop showing variants from the original Wuhan variant in 2020 to Omicron strains more recently (reproduced from Our World in Data). Interpretation: The variants have evolved to increase infectivity by reducing the latent period between the day infected and the day infectious. This has occurred in the both models whatever the original prior assumption used.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and

accession number(s) can be found at: <https://www.fil.ion.ucl.ac.uk/spm/covid-19/>. The figures in Figure 1 can be reproduced using annotated (MATLAB/Octave) code that is available as part of the free and open source academic software SPM (23). The

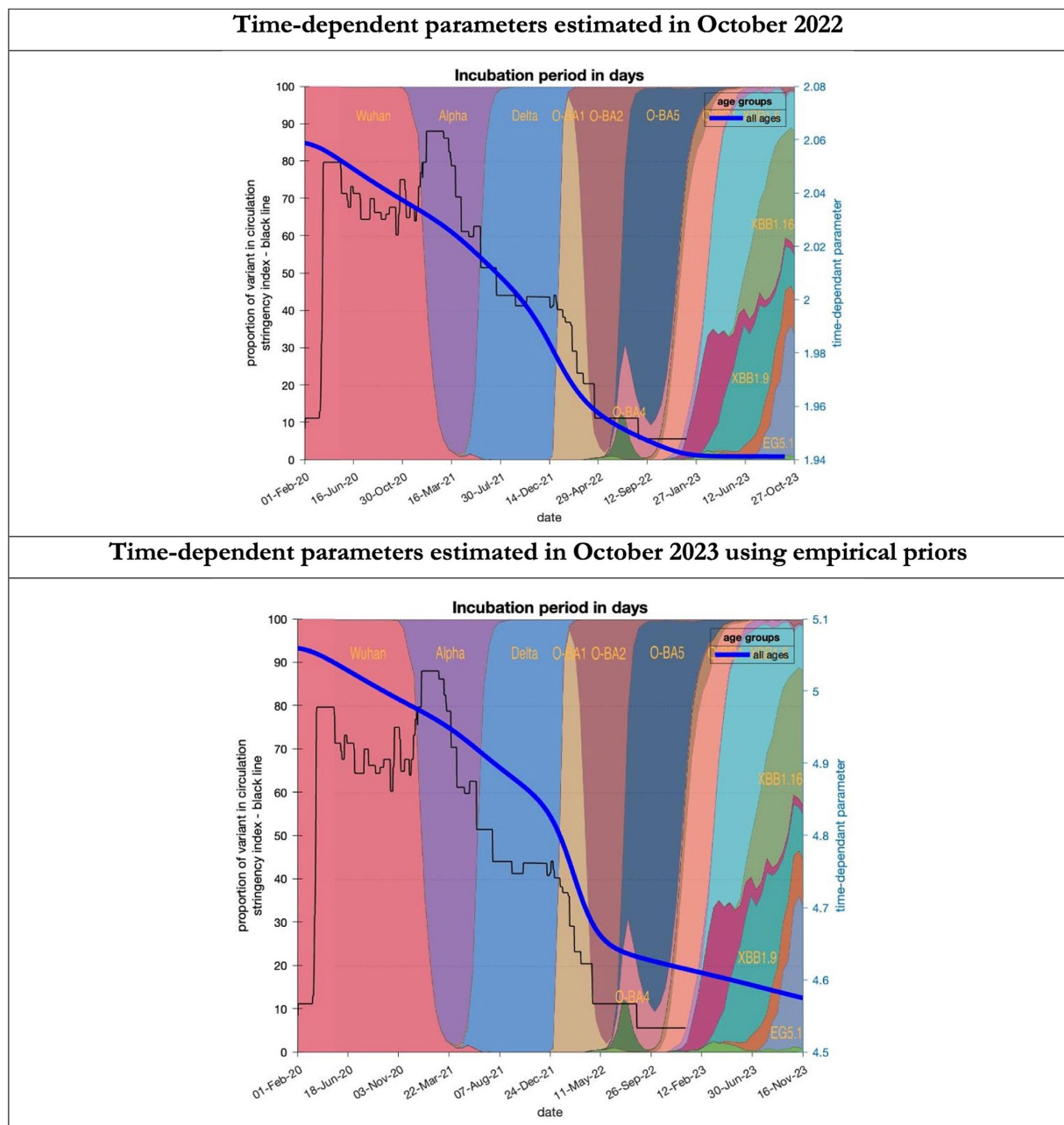


FIGURE 7

Changing estimates of incubation period of COVID-19 infection in relation to the emergence of new variants and changes in response to public health policies: UK February 2020 to October 2023. Incubation period (between day infected and start of symptoms) is measured as time constant for all age groups combined; prior in top graph of 4 days with initial model estimate of 2.06 days (asymptomatic period - Supplementary Table S2) dropping to 1.94 by October 2023; empirical prior in the bottom graph of 6.5 days with initial model estimate of 5.06 dropping to 4.6 by November 2023; stringency index dropping from 80% in March 2020 to 5% by December 2021; proportion of variant in circulation as backdrop showing variants from the original Wuhan variant in 2020 to Omicron strains more recently (reproduced from Our World in Data). Interpretation: As with the latent period the incubation period has shrunk in both models indicating the evolution of the variants which became more infectious.

routines are called by a demonstration script that can be invoked by DEM_COVID, DEM_COVID_X, DEM_COVID_T, DEM_COVID_I or DEM_COVID_LTLA at the MATLAB prompt. At the time of writing, these routines are available in the development version of the next SPM release. An archive of the relevant source code for each publication is available from figshare (<https://figshare.com/articles/>

Dynamic_Causal_Modelling_of_COVID-19/12174006). The remaining results in this paper can be reproduced using modified scripts found at <https://www.dropbox.com/scl/fo/zyv10xs8sn9ueuw7mhkis/h?rlkey=ewxlffkdiki89yzgjw6tz355g&dl=0>. The routine data used in the manuscripts are available from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University,

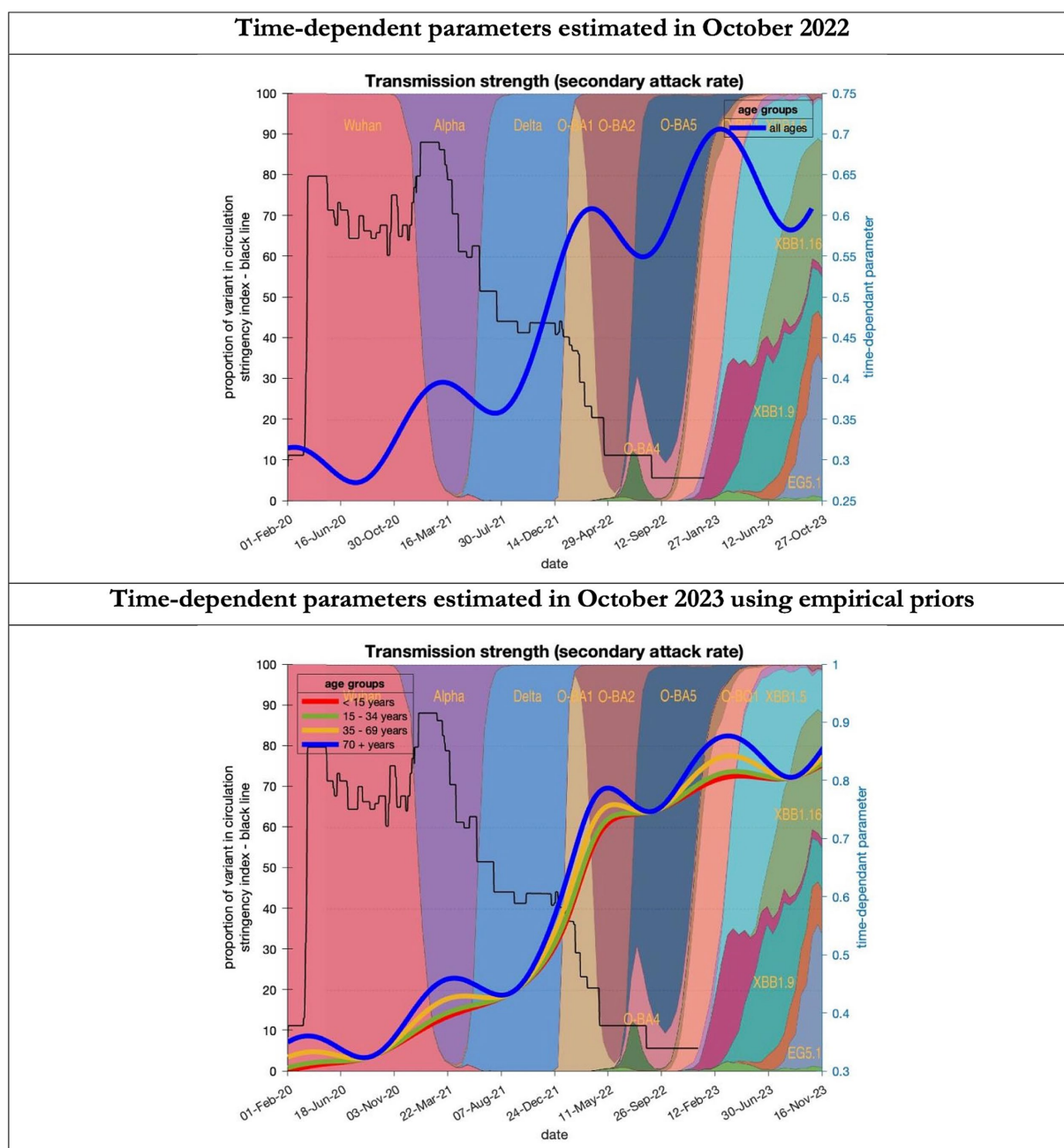


FIGURE 8

Changing estimates of transmission strength of COVID-19 infection in relation to the emergence of. Transmission strength is measured as the secondary attack rate; prior value of 0.3 (i.e. an infected person infects 1 in 3 contacts) which rises with the new variants to 0.7 (i.e. an infected person infects 70% of contacts); top graph combines all ages, bottom graph estimates transmission strength for each age group; stringency index dropping from 80% in March 2020 to 5% by December 2021; proportion of variant in circulation as backdrop showing variants from the original Wuhan variant in 2020 to Omicron strains more recently (reproduced from Our World in Data). Interpretation: Despite the different prior assumptions in the two figures the increase in transmission strength is evident in both models.

Coronavirus (COVID-19) UK Historical Data by Tom White and GOV.UK Coronavirus (COVID-19) in the UK.

from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required

Author contributions

CB: Formal analysis, Writing – original draft, Writing – review & editing. KF: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology,

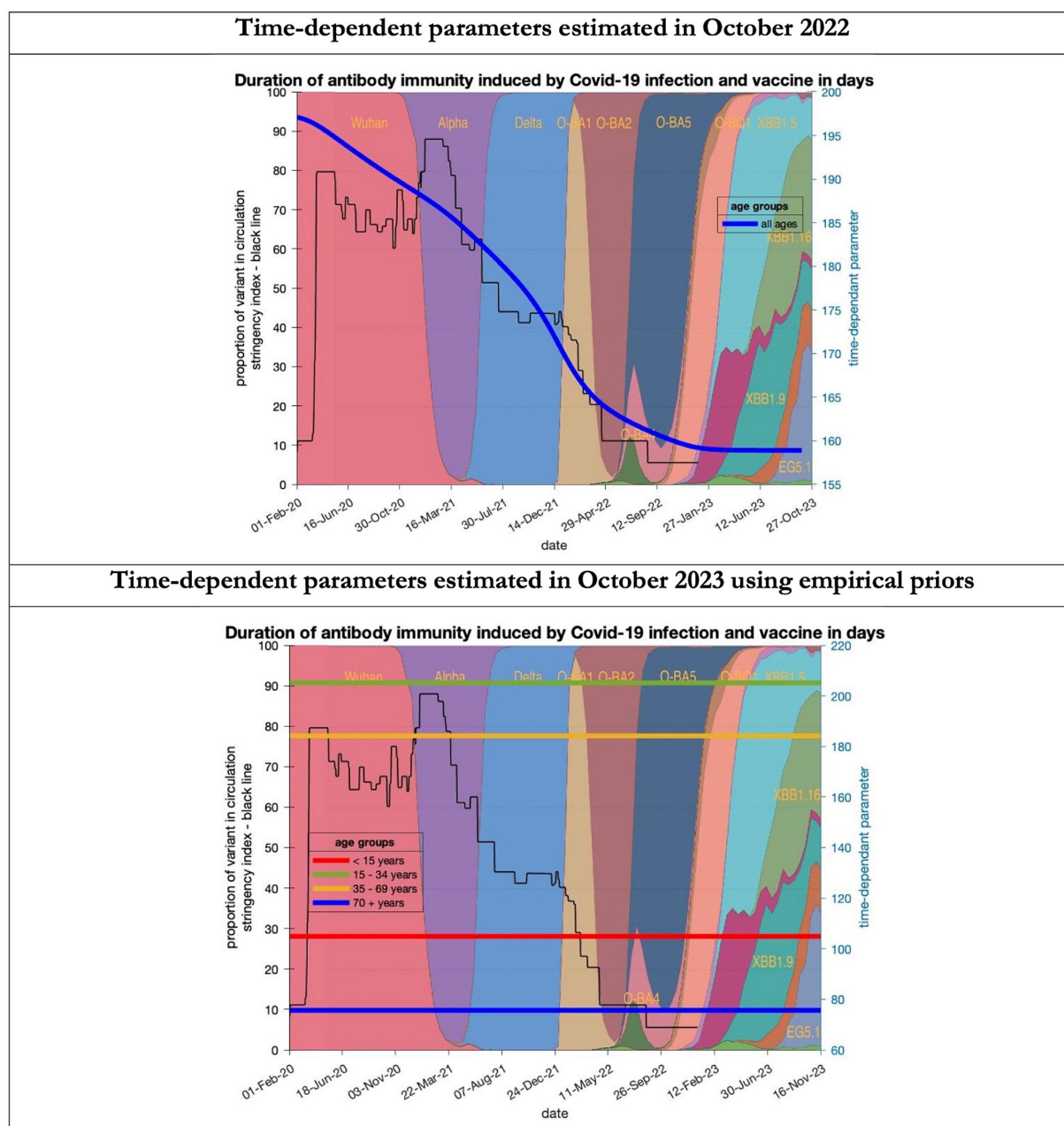


FIGURE 9

Changing estimates of duration of antibody immunity induced by COVID-19 infection and vaccine in relation to the emergence of new variants and changes in response to public health policies: UK February 2020 to October 2023. Duration of antibody immunity induced by COVID-19 infection and vaccine measured as time constant for all age groups combined in top graph and by age group in bottom graph; with initial model estimate of 196 days falling to 160 days by October 2023 in top graph; model estimates for each age group in bottom graph maintained at those values throughout the period; proportion of variant in circulation as backdrop showing variants from the original Wuhan variant in 2020 to Omicron strains more recently (reproduced from Our World in Data). Interpretation: The model used in 2022 assumed the possible time related change in the antibody immunity parameter whereas the 2023 model assumes no change. Further empirical data will be required to understand the changes in antibody immunity over time.

Supervision, Validation, Writing – original draft, Writing – review & editing.

Wellcome through core funding to the Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology [203147].

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. KF is supported by

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

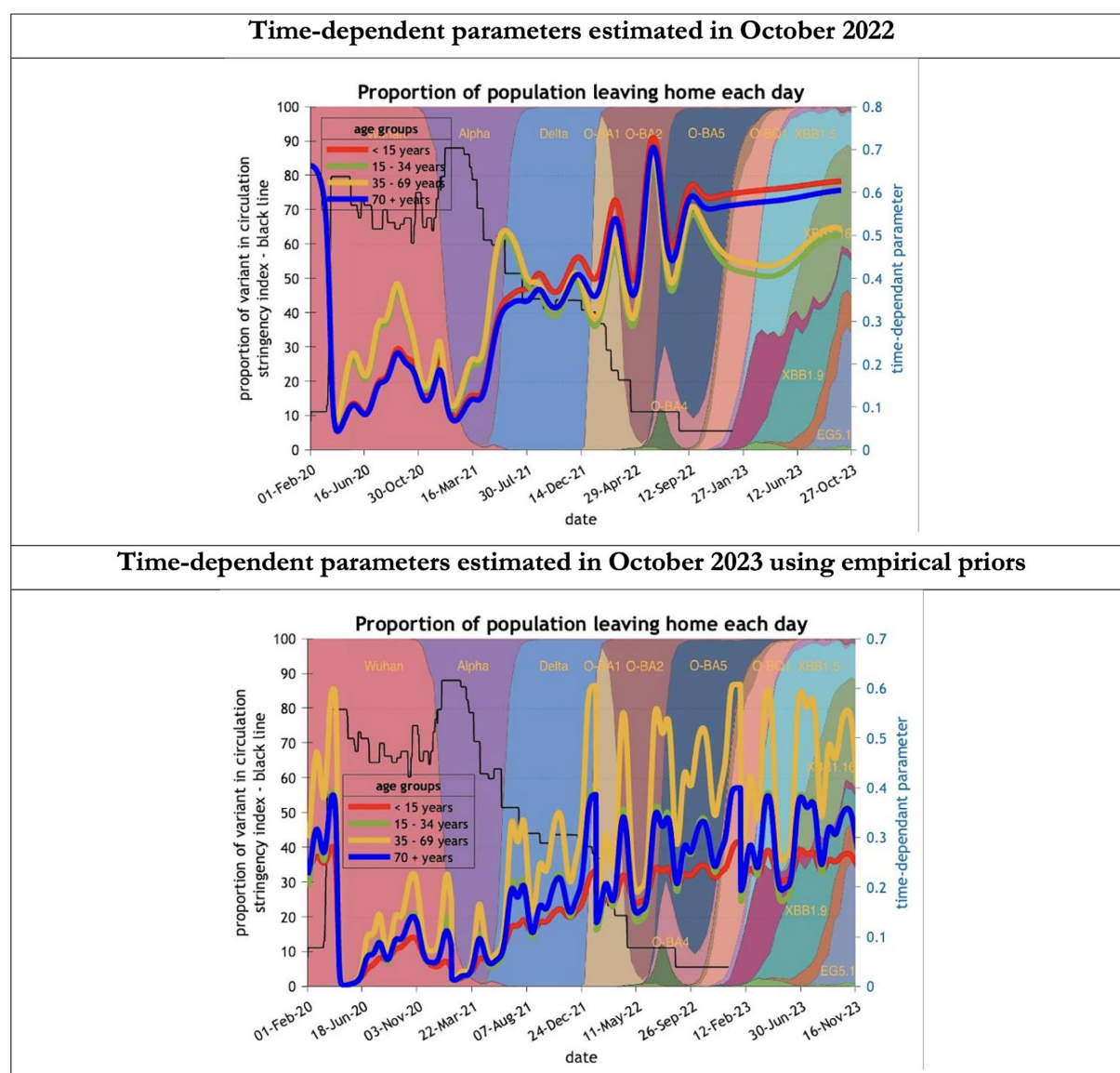


FIGURE 10

Changing estimates of the proportion of people leaving home each day in relation to the emergence of new variants and changes in response to public health policies: UK February 2020 to October 2023. The proportion of people leaving home each day for each age group; for example for those aged 70 years and above the top graph shows an estimate of 66% leaving home prior to the epidemic falling to 5% at first lockdown and rising slowly to 60% by October 2023; in the bottom graph the initial estimate for the same age group was 23% leaving home falling to 0% at the first lockdown rising to 32% by October 2023; stringency index dropping from 80% in March 2020 to 5% by December 2020; proportion of variant in circulation as backdrop showing variants from the original Wuhan variant in 2020 to Omicron strains more recently (reproduced from Our World in Data). Interpretation: The 2023 predictions estimate a much less mobile population than the 2022 model. This could partly explain the overestimate of infections identified in the earlier model.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2024.1398297/full#supplementary-material>

References

1. Friston K, Costello A, Pillay D. 'Dark matter,' second waves and epidemiological modelling. *BMJ Glob Health*. (2020) 5:e003978. doi: 10.1136/bmjgh-2020-003978
2. Friston KJ, Flandin G, Razi A. Dynamic causal modelling of COVID-19 and its mitigations. *Sci Rep*. (2022) 12:12419. doi: 10.1038/s41598-022-16799-8
3. Friston KJ, Parr T, Zeidman P, Razi A, Flandin G, Daunizeau J, et al. Second waves, social distancing, and the spread of COVID-19 across the USA. *Wellcome Open Res*. (2021) 5:103. doi: 10.12688/wellcomeopenres.15986.3
4. Birrell P, Blake J, van Leeuwen E, Gent N, De Angelis D. Real-time nowcasting and forecasting of COVID-19 dynamics in England: the first wave. *Philos Trans R Soc B Biol Sci*. (2021) 376:20200279. doi: 10.1098/rstb.2020.0279
5. Bowie C, Friston K. A 12-month projection to September 2022 of the COVID-19 epidemic in the UK using a dynamic causal model. *Frontiers. Public Health*. (2022) 10:210. doi: 10.3389/fpubh.2022.999210
6. FT Visual & Data Journalism Team. Free to read: Lockdowns compared: Tracking governments' coronavirus responses. Available at: <https://ig.ft.com/coronavirus-lockdowns> (Accessed September 27, 2022).
7. Oxford tracker. COVID-19 Government Response Tracker. Available at: <https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker> (Accessed September 27, 2022).
8. UKHSA. Investigation of SARS-CoV-2 variants: technical briefings. Available at: <https://www.gov.uk/government/publications/investigation-of-sars-cov-2-variants-technical-briefings> (Accessed September 27, 2022).
9. Bowie C, Friston K. Using a dynamic causal model to validate previous predictions and offer a 12-month forecast of the long-term effects of the COVID-19 epidemic in the UK. *Front. Public Health*. (2023) 10:1108886. doi: 10.3389/fpubh.2022.1108886
10. Friston K. Long-term forecasting of the COVID-19 epidemic. Available at: <https://www.fil.ion.ucl.ac.uk/spm/covid-19/forecasting/> (Accessed October 24, 2023).
11. England Summary | Coronavirus (COVID-19) in the UK. (2023) Available at: <https://coronavirus.data.gov.uk> (Accessed October 25, 2023).
12. Office for National Statistics. Search - Office for National Statistics. Available at: <https://www.ons.gov.uk/search?q=covid> (Accessed October 25, 2023).
13. Google. COVID-19 Community Mobility Report. Available at: <https://www.google.com/covid19/mobility?hl=en> (Accessed October 4, 2022).
14. IHME. COVID-19 estimation updates: Institute for Health Metrics and Evaluation. (2020) Available at: <http://www.healthdata.org/covid/updates> (Accessed May 18, 2020).
15. Office for National Statistics. Coronavirus (COVID-19) Infection Survey, UK - Office for National Statistics. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveysurvey/24march2023> (Accessed October 25, 2023).
16. Global Burden of Disease Long COVID Collaborators. Estimated global proportions of individuals with persistent fatigue, cognitive, and respiratory symptom clusters following symptomatic COVID-19 in 2020 and 2021. *JAMA*. (2022) 328:1604–15. doi: 10.1001/jama.2022.18931
17. Our World in Data SARS-CoV-2 variants in analyzed sequences. Available at: <https://ourworldindata.org/grapher/covid-variants-area?country=-GBR> (Accessed November 21, 2023).
18. GISAID. Editorial Board affirms Open Access designation of GISAID. *Re3data.org*. (2012) . doi: 10.17616/R3Q59F
19. Friston K. Long-term forecasting of the COVID-19 epidemic. Available at: <https://www.fil.ion.ucl.ac.uk/spm/covid-19/forecasting/> (Accessed November 6, 2023).
20. Xu X, Wu Y, Kummer AG, Zhao Y, Hu Z, Wang Y, et al. Assessing changes in incubation period, serial interval, and generation time of SARS-CoV-2 variants of concern: a systematic review and meta-analysis. *BMC Med*. (2023) 21:374. doi: 10.1186/s12916-023-03070-8
21. Hakki S, Zhou J, Jonnerby J, Singanayagam A, Barnett JL, Madon KJ, et al. Onset and window of SARS-CoV-2 infectiousness and temporal correlation with symptom onset: a prospective, longitudinal, community cohort study. *Lancet Respir Med*. (2022) 10:1061–73. doi: 10.1016/S2213-2600(22)00226-0
22. Office for National Statistics. Coronavirus (COVID-19) latest insights - Office for National Statistics. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/lifestyle> (Accessed October 30, 2023)
23. Friston Karl. Dynamic Causal Modelling of COVID-19; dashboard. Available at: <https://www.fil.ion.ucl.ac.uk/spm/covid-19/> (Accessed May 24, 2021).



OPEN ACCESS

EDITED BY

Jiaying Li,
The University of Queensland, Australia

REVIEWED BY

Tianfeng He,
Ningbo Municipal Center for Disease Control
and Prevention, China
Martyn Regan,
The University of Manchester,
United Kingdom

*CORRESPONDENCE

Youqiong Xu
✉ joancoco@126.com
Xiaoyang Zhang
✉ dawnsunz@126.com

RECEIVED 30 May 2024

ACCEPTED 29 August 2024

PUBLISHED 23 September 2024

CITATION

Zheng X, Chen Q, Sun M, Zhou Q, Shi H,
Zhang X and Xu Y (2024) Exploring the
influence of environmental indicators and
forecasting influenza incidence using ARIMAX
models.

Front. Public Health 12:1441240.
doi: 10.3389/fpubh.2024.1441240

COPYRIGHT

© 2024 Zheng, Chen, Sun, Zhou, Shi, Zhang
and Xu. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Exploring the influence of environmental indicators and forecasting influenza incidence using ARIMAX models

Xiaoyan Zheng^{1,2}, Qingquan Chen^{1,2}, Mengcai Sun^{1,2},
Quan Zhou^{1,2}, Huanhuan Shi^{1,2}, Xiaoyang Zhang^{1,2*} and
Youqiong Xu^{1,2*}

¹The Affiliated Fuzhou Center for Disease Control and Prevention of Fujian Medical University, Fuzhou, China, ²The School of Public Health, Fujian Medical University, Fuzhou, China

Background: Influenza is a respiratory infection that poses a significant health burden worldwide. Environmental indicators, such as air pollutants and meteorological factors, play a role in the onset and propagation of influenza. Accurate predictions of influenza incidence and understanding the factors influencing it are crucial for public health interventions. Our study aims to investigate the impact of various environmental indicators on influenza incidence and apply the ARIMAX model to integrate these exogenous variables to enhance the accuracy of influenza incidence predictions.

Method: Descriptive statistics and time series analysis were employed to illustrate changes in influenza incidence, air pollutants, and meteorological indicators. Cross correlation function (CCF) was used to evaluate the correlation between environmental indicators and the influenza incidence. We used ARIMA and ARIMAX models to perform predictive analysis of influenza incidence.

Results: From January 2014 to September 2023, a total of 21,573 cases of influenza were reported in Fuzhou, with a noticeable year-by-year increase in incidence. The peak of influenza typically occurred around January each year. The results of CCF analysis showed that all 10 environmental indicators had a significant impact on the incidence of influenza. The ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with PM₁₀(lag5) model exhibited the best prediction performance, as indicated by the lowest AIC, AICc, and BIC values, which were 529.740, 530.360, and 542.910, respectively. The model achieved a fitting RMSE of 2.999 and a predicting RMSE of 12.033.

Conclusion: This study provides insights into the impact of environmental indicators on influenza incidence in Fuzhou. The ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with PM₁₀(lag5) model could provide a scientific basis for formulating influenza control policies and public health interventions. Timely prediction of influenza incidence is essential for effective epidemic control strategies and minimizing disease transmission risks.

KEYWORDS

influenza incidence, air pollution, meteorological factor, time series, ARIMAX model, prediction influenza incidence, prediction

1 Introduction

Influenza, a widely prevalent respiratory infection, exerts a substantial impact on the health of millions of people worldwide each year, leading to severe morbidity and occasional deaths (1). While, like other respiratory infections, influenza is typically most prevalent during the winter and spring seasons, recent reports have illuminated a noteworthy surge in summer influenza cases (2). This emerging trend presents fresh challenges for health authorities and influenza surveillance efforts. The onset and propagation of influenza are influenced by a multitude of factors, including environmental indicators such as air pollutants (3) and meteorological factors (4). Therefore, it is of paramount importance to attain accurate predictions of influenza incidence and develop a thorough understanding of the factors that influence it.

Timely prediction of infectious diseases is essential to maintaining and improving public health (5). It helps the government to formulate and implement effective epidemic control strategies, ensuring the availability of adequate medical resources and healthcare personnel, thereby minimizing the risk of disease transmission. Currently, various methods are employed for predicting infectious diseases, encompassing infectious disease dynamics model (6), logistic regression model (7), gray prediction theory (8, 9), ARIMA model (10–12), Prophet model (13), Holt-Winters model (14), and LSTM models (15). Each of these methods possesses its own set of advantages and drawbacks. Notably, the ARIMA model stands out in its ability to accurately identify the seasonality and trends of infectious diseases. For instance, Wu et al. utilized the ARIMA method to forecast the incidence of pulmonary tuberculosis under the regular COVID-19 epidemic prevention and control measures in China (16). Ahn et al. (17) effectively applied the ARIMA model to anticipate the incidence of rheumatic diseases during the COVID-19 pandemic in Korea. While previous studies have extensively delved into the prediction of infectious diseases, researchers often overlook the potential impacts of air pollution and meteorological factors on infectious diseases. There exists a certain degree of correlation between environmental indicators and the incidence of infectious diseases (18, 19). Thus, the inclusion of environmental indicators in the predictive model for infectious diseases is anticipated to enhance the accuracy of predictions to some extent.

In recent years, the incidence of influenza in Fuzhou has been increasing year by year, adding to the challenges of disease prevention and treatment. Notably, in 2023, during a spring peak in Fuzhou, the monthly reported cases of influenza reached 2,749, marking the highest number reported in a single month over the past decade. Therefore, the analysis of factors influencing influenza incidence and the provision of corresponding predictions and early warnings are crucial for the development of effective prevention and control strategies.

Our study initiated an analysis of the impact of environmental indicators, including air pollution and meteorological factors, on influenza incidence. It then developed an optimal ARIMA model based on influenza incidence data. Subsequently, to enhance prediction accuracy, environmental indicators were systematically introduced into the optimal ARIMA model, leading to the establishment of the ARIMAX model. Finally, we selected the optimal ARIMAX model for the prediction analysis of influenza incidence in Fuzhou.

2 Materials and methods

2.1 Study area and data sources

Fuzhou, situated in the southeast coastal area of China, serves as the capital city of Fujian Province and spans an area of 11,968.53 square kilometers. As of the end of 2022, Fuzhou had a permanent resident population of 8.448 million. The monthly data on influenza cases were sourced from the Fuzhou Center for Disease Control and Prevention. The surveillance of influenza cases followed the criteria outlined by the World Health Organization and the Chinese Center for Disease Control and Prevention for influenza-like cases. Population statistics were extracted from the Fuzhou Statistical Yearbook. We utilized monthly influenza incidence (per 100,000 populations) data spanning from January 2014 to December 2022. This dataset was split into two subsets: a training set covering the period from January 2014 to December 2022, and a test set spanning from January 2023 to September 2023.

The monthly air pollution monitoring data used in this study covers the period from January 2014 to September 2023 and was provided by the Environmental Monitoring Center under the Environmental Protection Administration of Fuzhou. The air pollutants included particulate matter $2.5\mu\text{m}$ ($\text{PM}_{2.5}$), particulate matter $10\mu\text{m}$ (PM_{10}), sulfur dioxide (SO_2), carbon monoxide (CO), nitrogen dioxide (NO_2), and ozone (O_3). Simultaneously, the monthly meteorological data for the same period were procured from the Fuzhou Meteorological Bureau, encompassing meteorological factors such as monthly average temperature ($^{\circ}\text{C}$), monthly maximum temperature ($^{\circ}\text{C}$), monthly minimum temperature ($^{\circ}\text{C}$), and monthly average wind speed (m/s). The monitoring data for the above environmental indicators was obtained with authorization from the Fuzhou Environmental Protection Bureau and the Fuzhou Meteorological Bureau.

2.2 Construction of the seasonal ARIMA model

Autoregressive Integrated Moving Average Model (ARIMA) is a widely-used method for the analysis and prediction of time series data (20). It finds applications in forecasting infectious diseases like varicella (21), tuberculosis (22), and COVID-19 (23). The fundamental concept underlying ARIMA model is to utilize historical data to make future predictions. ARIMA model is primarily composed of three components: Autoregressive (AR), Integration (I), and Moving Average (MA). For time series data exhibiting periodic patterns, the Seasonal Autoregressive Integrated Moving Average Model (SARIMA) combines seasonal differencing with the standard ARIMA model, making it well-suited for modeling data with recurring characteristics.

In our study, we developed a SARIMA model denoted as $\text{ARIMA}(p, d, q)(P, D, Q)_s$, where p signifies the AR order, d stands for the differencing order and q represents the MA order. Meanwhile, s indicates the period of seasonal trend, while P , D and Q correspond to the seasonal terms within the SARIMA model. The determination of these parameters, (p, d, q) and (P, D, Q) , is achieved through an analysis of the Partial Autocorrelation Function (PACF) and the Autocorrelation Function (ACF). The choice of the parameter s

depends on the length of the seasonal cycle. The seasonal model can be mathematically represented as follows:

$$\phi_p(B)\tilde{\phi}_p(B^s)y_t^* = \theta_q(B)\tilde{\theta}_q(B^s)\varepsilon_t \quad (1)$$

In Equation 1, $\phi_p(B)$ represents a non-seasonal autoregressive lag polynomial, $\tilde{\phi}_p(B^s)$ represents seasonal moving average lag polynomial, $\theta_q(B)$ represents seasonal moving average lag polynomial. To ensure the stability of our time series, we initially applied differencing, a crucial step in the analysis. We then conducted an augmented Dickey–Fuller (ADF) test to verify the temporal stability of the series. Subsequently, we employed the corrected Akaike’s information criterion (AICc) to assess the goodness of fit of the SARIMA model, with the model associated with the lowest AICc value considered the optimal choice. Finally, we conducted the Ljung–Box test to ascertain whether the residual sequence of the model exhibited characteristics of white noise. If the p -value is greater than 0.05, the model satisfies the test’s criteria and can be employed for predictive analysis.

2.3 Construction of the ARIMAX model

ARIMAX model, which incorporates exogenous variables related to the target time series as input variables, builds upon the foundation of the ARIMA model to enhance prediction accuracy (24). The primary objective of the ARIMAX model is to capture trends and seasonal fluctuations within time series data by amalgamating autoregressive, differencing, moving average components, and exogenous variables, thereby offering precise predictions and robust analytical capabilities. In contrast to the ARIMA model, the ARIMAX model takes into account exogenous variables that are associated with the time series data. These exogenous variables can encompass other time series data or non-time series data, such as environmental indicators (25, 26) and government policies (27). The role of exogenous variables is to furnish additional information that aids in refining model fitting and prediction accuracy.

In this study, we developed an ARIMAX model for each exogenous environmental variable using data from six air pollutants and four meteorological factors. Our approach consisted of three main steps: Initially, we conducted the cross-correlation function (CCF) to assess the time-delay correlation between different variables and influenza incidence. Subsequently, we integrated significant environmental indicators as exogenous variables into the optimal ARIMA model, thereby creating alternative ARIMAX models. Finally, we selected the best-fitting ARIMAX model based on three criteria: (a) the Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICc), Bayesian Information Criterion (BIC), Root mean squared error (RMSE) values are smaller than the optimal ARIMA model; (b) the degree that the residual sequence of the model is white noise by Ljung–Box test; (c) the model’s performance in predicting influenza incidence in 2023.

The primary innovation of our study lies in the integration of environmental indicators into the ARIMAX framework. By incorporating exogenous variables related to influenza incidence, we can gain a more comprehensive understanding of the multifaceted factors influencing disease transmission. This approach not only

improves the accuracy of our predictions but also provides valuable insights for public health interventions. Furthermore, we employ advanced model selection criteria, such as the corrected AICc, to ensure optimal model fitting. Through these enhancements, our research contributes a novel perspective to the application of ARIMA models in the field of epidemiology, demonstrating their adaptability and relevance in addressing contemporary public health challenges.

2.4 Statistical methods

Descriptive statistics were employed to illustrate changes in influenza incidence, air pollutants and meteorological factors. Time series plots (line plots) were utilized to visualize their temporal distribution. The cross-correlation function (CCF) was used to evaluate the lag effect of environmental influencing factors. For the development of ARIMA and ARIMAX models, as well as data visualization, we utilized the R packages “forecast,” “stats,” and “ggplot2” in R (version 4.2.1, The R Foundation). The significance level was set at 0.05.

2.5 Ethical approval and consent to participate

We obtained ethical approval from the Ethical Review Committee of the Fuzhou Center for Disease Control and Prevention (Approval No. IRB2020008) to conduct a secondary analysis of aggregated data collected by the Fuzhou CDC, China. The informed consent requirement was waived by the Ethical Review Committee of the Fuzhou Center for Disease Control and Prevention for this study. This study was carried out following the Helsinki Declaration contents.

3 Results

From January 2014 to September 2023, a total of 21,573 cases of influenza were reported in Fuzhou, with an incidence rate of 2.228 ± 4.593 (as shown in Table 1). The highest number of cases was recorded in June 2023, with 2,749 reported cases. Analysis of the time series chart of influenza incidence reveals that the peak of influenza cases typically occurs around January each year. Overall, there is a noticeable year-by-year increase in influenza incidence (as depicted in Figure 1).

Upon reviewing the data from the past few years, it becomes evident that nearly every winter is marked by severe air pollution in Fuzhou. Simultaneously, there is a notable increase in the incidence of influenza. Overall, the concentrations of all other five air pollutants, with the exception of O_3 , exhibit a consistent downward trend, as illustrated in Figure 2. The mean concentrations of $PM_{2.5}$, PM_{10} , SO_2 , CO , NO_2 , and O_3 were 24.160, 45.620, 5.479, 0.660, 24.060, and $88.260 \mu g/m^3$, respectively.

During the study period, the time series of meteorological factors exhibited a strong cyclical and seasonal pattern overall, with peak values occurring during the summer and troughs observed in the winter (as depicted in Figure 3). The mean values of the monthly average temperature, maximum temperature, minimum temperature

TABLE 1 The descriptive statistics of the monthly influenza incidence and environmental indicators in Fuzhou, 2014–2023.

Variable	Range	Mean $\pm$ S.D.	P25	P50	P75	IQR
Incidence (/100,000)	0.236–32.540	2.228 $\pm$ 4.593	0.645	0.974	1.845	1.200
Average temperature ($^{\circ}$ C)	9.528–30.000	19.893 $\pm$ 6.114	13.926	20.000	25.831	11.905
Maximum temperature ($^{\circ}$ C)	12.050–35.000	23.600 $\pm$ 6.346	17.540	23.140	29.230	11.690
Minimum temperature ($^{\circ}$ C)	6.291–25.000	16.182 $\pm$ 5.923	10.059	16.390	21.960	11.901
Average wind speed (m/s)	4.300–9.681	6.762 $\pm$ 0.985	6.112	6.700	7.336	1.224
PM _{2.5} (μ g/m ³)	12.000–56.000	24.160 $\pm$ 7.912	18.000	23.000	29.000	11.000
PM ₁₀ (μ g/m ³)	23.000–89.000	45.620 $\pm$ 12.901	36.000	44.000	53.000	17.000
SO ₂ (μ g/m ³)	3.000–16.000	5.479 $\pm$ 1.827	4.000	5.000	6.000	2.000
CO (mg/m ³)	0.326–1.165	0.660 $\pm$ 0.142	0.577	0.668	0.735	0.158
NO ₂ (μ g/m ³)	8.000–52.000	24.060 $\pm$ 8.915	17.000	23.000	30.000	13.000
O ₃ (μ g/m ³)	45.000–130.000	88.260 $\pm$ 18.542	75.000	87.000	102.000	27.000

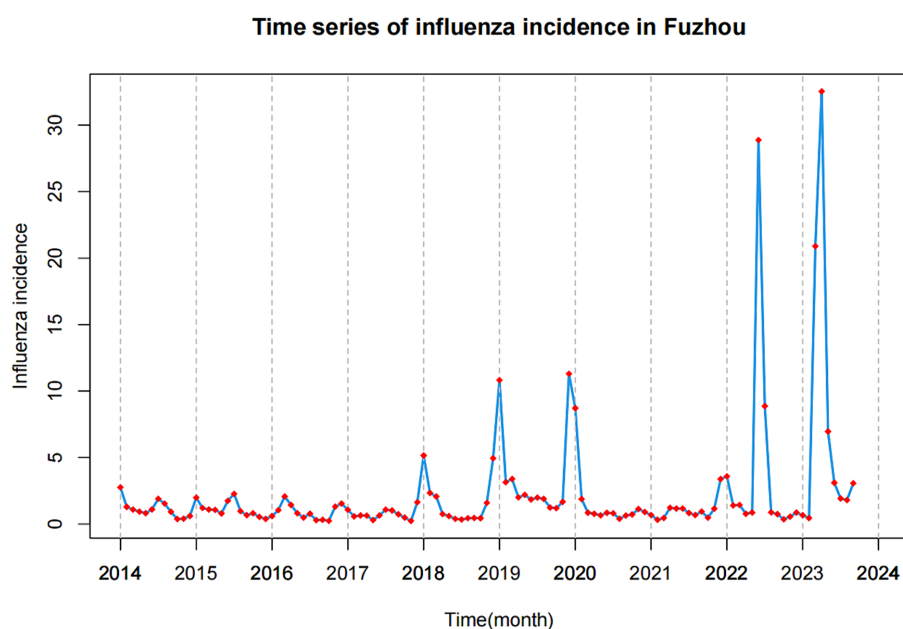


FIGURE 1
Time series of influenza incidence in Fuzhou from January 2014 to September 2023.

and average wind speed were 19.893, 23.600, 16.182, and 6.762 m/s, respectively.

We investigated the lagged relationship between 10 environmental indicators and influenza incidence using cross-correlation analysis. As illustrated in Table 2, SO₂, CO, NO₂, average temperature, maximum temperature, and minimum temperature exhibited direct and statistically significant associations with influenza incidence, while the lag variables for the other three environmental indicators also displayed significant associations with influenza incidence.

To begin with, it is imperative to establish an optimal ARIMA model for predicting influenza incidence in Fuzhou. Prior to modeling, we conducted an ADF test to assess the stability of both influenza incidence and 10 environmental indicators, aiming to ascertain if differential processing was necessary. All *p*-values from the tests were found to be less than 0.05, signifying the data were

stationary and did not need to be differential processed. Consequently, we conclude that the parameters *d* and *D* in the ARIMA(*p*, *d*, *q*) (*P*, *D*, *Q*)_s model were both 0. Given that our predictive models were constructed using influenza incidence data spanning January 2014 to December 2022, we decomposed the data into trend, season, and random items. The influenza time series showed an upward trend. Meanwhile, this analysis also revealed a pronounced seasonality in influenza incidence data, characterized by a seasonal period of 12 (refer to Figure 4). Consequently, the parameter *s* of the ARIMA model was set at 12, and the model can be expressed as ARIMA(*p*, 0, *q*) (*P*, 0, *Q*)₁₂.

We developed the model using data from the training set (January 2014 to December 2022) and assessed the prediction performance of the model using the test set data (January 2023 to September 2023). To determine the values of the remaining ARIMA model parameters

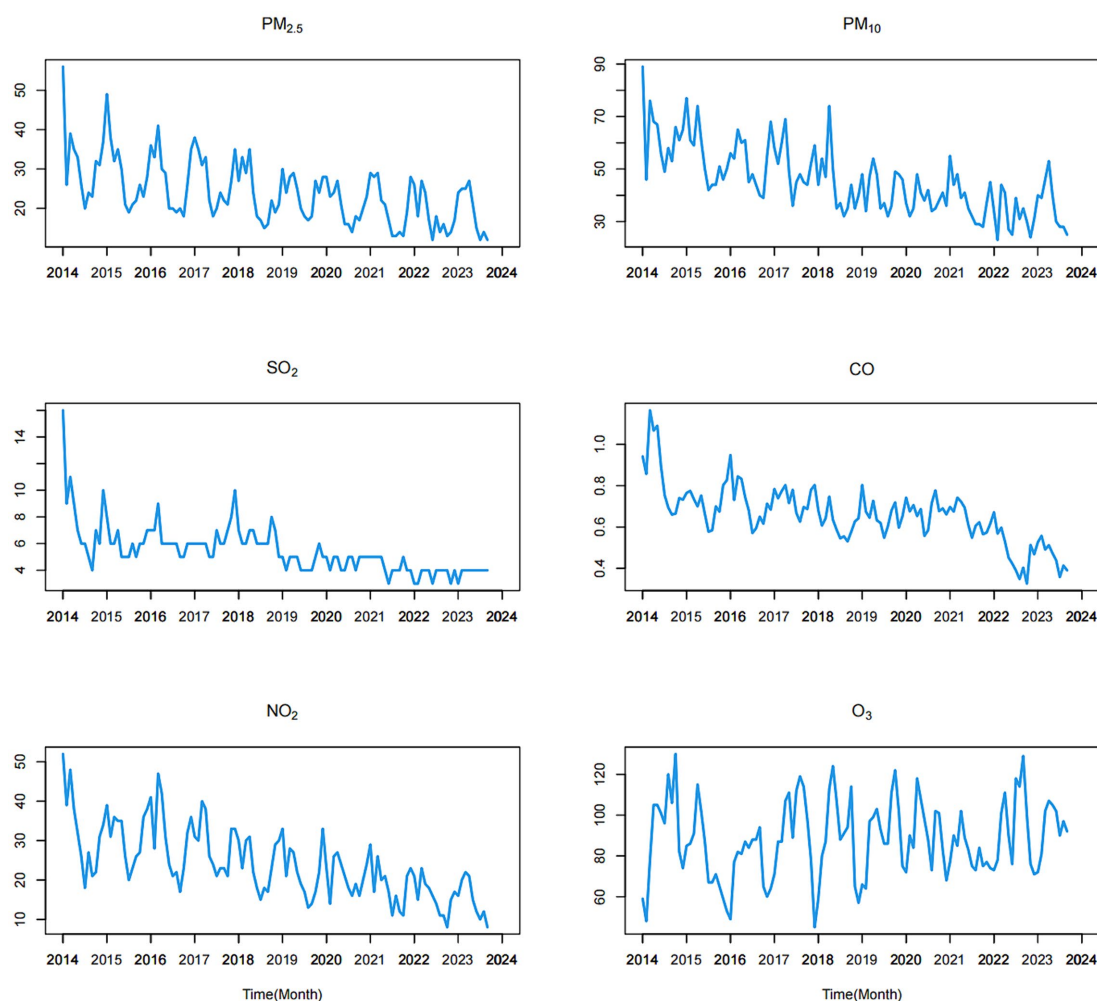


FIGURE 2
Time series of the six air pollution variables from January 2014 to September 2023.

p , q , P , and Q , we generated ACF and PACF plots based on the training set data. The plots for ACF and PACF reveal the temporal dependence of influenza incidence, with maximum autocorrelation and partial correlation coefficients observed at lags 0 (refer to Figure 5).

Through the analysis of the ACF and PACF plots of the original time series, it can be determined that the remaining parameters p , q , P , and Q of the ARIMA model should be 0, or 1. To automatically identify the model order of the ARIMA model, we used the `auto.arima` function from the “forecast” package to select a total of 13 alternative models (Table 3). Finally, the optimal model was identified as ARIMA(0, 0, 1) (1, 0, 0)₁₂, boasting the lowest AIC, AICc, and BIC values, which stood at 552.910, 553.303, and 563.640, respectively. Additionally, the Ljung–Box test confirmed that the residual sequence resembles white noise ($p > 0.05$). The ARIMA(0, 0, 1) (1, 0, 0)₁₂ model excelled in both fitting and predicting influenza incidence. When applied to the training set, the model yielded the fitting RMSE of 3.002; the model was employed to predict influenza incidence in the test set, achieving the predicting RMSE of 12.475.

To investigate the potential influence of environmental indicators, such as air pollutants and meteorological factors, on

influenza incidence, we systematically integrated these environmental indicators one by one into the ARIMA(0, 0, 1) (1, 0, 0)₁₂ model to formulate an optimal ARIMAX model. We integrated the maximum lag correlation variables for each environmental indicator into the ARIMA(0, 0, 1) (1, 0, 0)₁₂ model, thus creating 10 distinct ARIMAX models. The Ljung–Box test was employed to assess these 10 models, and results indicated that the residual sequences of the models exhibited white noise characteristics (All $p > 0.05$).

Based on the outcomes summarized in Table 4, it was determined that the ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with PM₁₀(lag5) model had the lowest AIC, AICc, and BIC values, signifying superior fitting accuracy and suitability for predicting influenza incidence in Fuzhou. During the model-fitting phase using the training set, this ARIMAX model achieved a RMSE of 2.999. When applied to forecast influenza incidence in the test set, the model had an RMSE of 12.033.

Figure 6 graphically presents the fitting and predictive results of influenza incidence rates based on the ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with PM₁₀(lag5) model. These results demonstrate the efficacy of the ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with PM₁₀(lag5) model in accurately forecasting influenza incidence in Fuzhou. Notably, the model

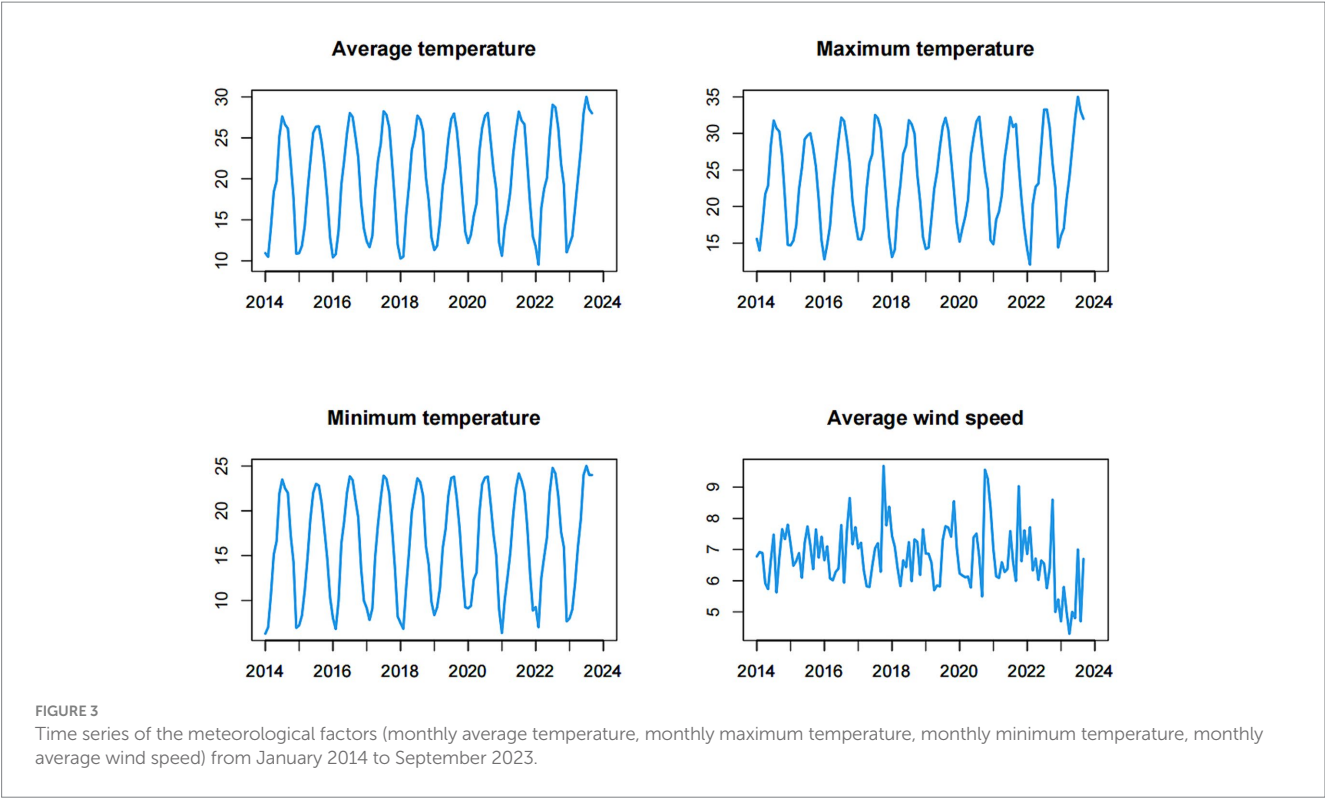


TABLE 2 The correlation coefficients and maximum lag correlation coefficients between influenza incidence and environmental indicators.

	PM _{2.5}	PM ₁₀	SO ₂	CO	NO ₂	O ₃
Corr-Coef	−0.018	−0.070	−0.183*	−0.220*	−0.183*	−0.006
Max lag Corr-Coef	−0.266*	−0.291*	−0.184*	−0.326*	−0.259*	0.238*
Its lag order (Max)	5	5	1	4	4	3

	Ave.temp	Max.temp	Min.temp	Ave.ws
Corr-Coef	−0.053*	−0.051*	−0.054*	−0.240
Max lag Corr-Coef	0.211*	0.225*	0.195*	−0.290*
Its lag order (Max)	3	3	3	2

* $p < 0.05$; Ave.temp, average temperature; Max.temp, maximum temperature; Min.temp, minimum temperature; Ave.ws, average wind speed.

displayed commendable fitting accuracy in both the training and test sets.

4 Discussion

Influenza is a respiratory viral disease caused by the influenza virus (28). It typically manifests with acute respiratory symptoms, but for individuals with weaker immune systems, such as the young, the older adults, or those with compromised immunity, influenza can lead to more severe complications even life-threatening outcomes (2). Over the past decade, Fuzhou has witnessed a notable surge in the incidence of influenza, indicating a critical influenza epidemic. Hence, investigating the factors influencing influenza incidence is crucial for the evidence-based development of influenza control policies and the implementation of timely public health interventions.

In 2022, the winter flu peak did not occur in Fuzhou, primarily attributed to the outbreak of COVID-19 and the strict epidemic prevention and control measures implemented, including the complete suspension of in-room dining and the promotion of remote work. These measures effectively reduced interpersonal contact, thereby mitigating the spread of influenza. The proactive interventions in response to the COVID-19 outbreak in Fuzhou had a positive impact on curbing the high incidence of influenza. However, China removed many restrictive COVID-19 prevention and control measures after January 8, 2023. It resulted in a rapid increase in COVID-19 infections and necessitated home-based treatments for many citizens, contributing to a partial reduction in the spread of influenza. These observations underscore the need for in-depth analysis in future studies to understand the specific mechanisms and long-term trends of various public health interventions on influenza transmission.

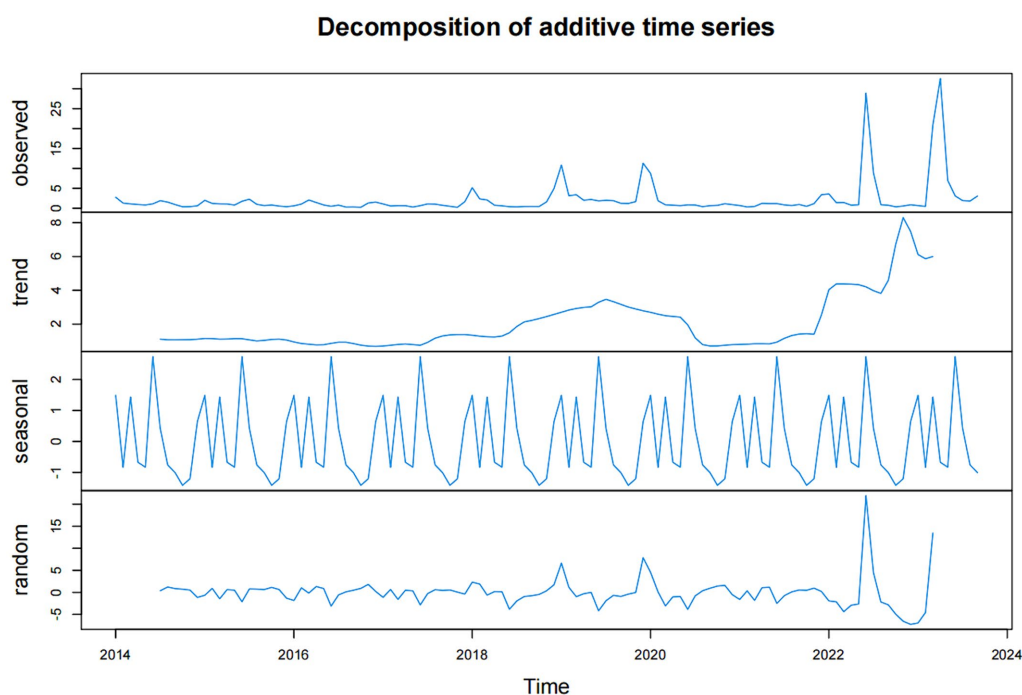


FIGURE 4

The data of influenza incidence in Fuzhou were decomposed into trend part, seasonal part and random part.

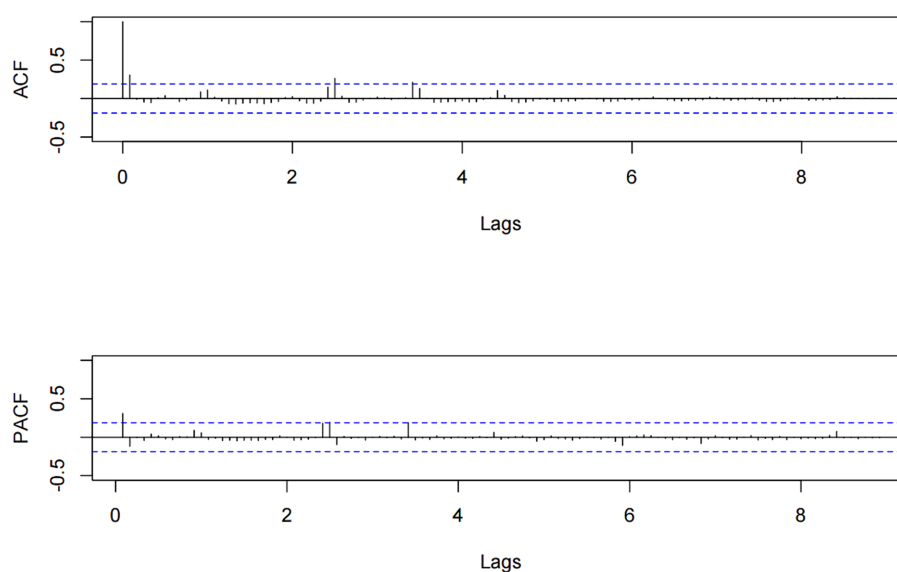


FIGURE 5

The ACF and PACF charts for influenza incidence data from January 2014 to December 2022.

During the period from 2022 to 2023, Fuzhou experienced consecutive summer influenza peaks, with a higher number of reported cases in both years. Apart from the conducive climate conditions of high temperature and humidity during summer, which potentially facilitate the transmission of the influenza virus, the reasons behind the summer influenza peaks in the 2 years might be different, contingent upon the contextual circumstances prevailing at the time.

In 2022, amidst a significant influenza pandemic, Fuzhou encountered no COVID-19 outbreak in June 2022, and residents reduced their mask-wearing behavior due to hot weather conditions. Concurrently, with medical resources extensively allocated for monitoring and treating respiratory diseases during the influenza pandemic, this likely resulted in intensified surveillance and reporting of influenza cases. During June to July 2023, the emergence of a summer influenza peak in Fuzhou may

be associated with China's relaxation of numerous restrictive COVID-19 control measures, such as mask-wearing and avoidance of crowded places, effective from January 8, 2023. Subsequently, residents' immune systems may have weakened. During the COVID-19 pandemic, heightened attention to personal protection and hygiene practices might have reduced exposure of the immune system to common viruses. Following the easing of restrictions, resumption of social activities may have diminished the immune system's resistance to the influenza virus, thereby precipitating its outbreak. Moreover, there was a notable increase in social gatherings. Post-lockdown, individuals likely resumed more social and congregational activities such as dining, gatherings, and tourism. Such congregation could have facilitated the spread of the influenza virus, contributing to the peak in influenza cases. Finally, the relaxation of healthcare resource pressures could also have played a role. During the COVID-19 pandemic, medical resources were primarily directed toward combating the outbreak, potentially leading to neglect in the prevention and control of

other diseases. Post-lockdown, while healthcare resources might have eased, reduced vigilance toward COVID-19 may have led to diminished attention and control measures for influenza, thereby fostering its transmission.

There have been many previous studies have demonstrated the association between various diseases and environmental indicators, including diseases like dengue fever (29, 30), COVID-19 (31–33), and tuberculosis (34). In the case of influenza, environmental indicators can influence the occurrence of influenza epidemics through factors such as the variation and transmission of influenza virus and the immune status of the population (35). The Cross-Correlation Function (CCF) measures the correlation between two variables at different time lags, making it particularly well-suited for analyzing lagged effects and time-delayed relationships between variables. Additionally, as the impact of environmental indicators may exhibit a time lag in disease incidence (36, 37), we investigated the lagged correlation between influenza incidence and these environmental indicators.

Our analysis revealed that most of the lagged air pollution variables exhibited a negative association with influenza incidence. This implies that as air pollution levels increase, the incidence of influenza tends to decrease. This negative correlation can, in part, be attributed to the adverse impact of severe air pollution on the human immune system, thereby increasing the risk of infectious diseases (38). However, the manifestation of this weakened immune system in terms of influenza incidence may not be immediately evident and could require some time to become apparent. This phenomenon might also be linked to public awareness of declining air quality. Following the perception of deteriorating air quality, individuals may have adopted proactive protective measures, including reducing outdoor activities and wearing face masks to mitigate their exposure to air pollution (39). These self-protective behaviors could contribute to a reduction in the likelihood of influenza virus transmission, consequently lowering the incidence of influenza. Moreover, it's essential not to overlook the impact of the COVID-19 pandemic in recent years. From 2019 to 2022, widespread mask-wearing in public to prevent COVID-19 not only effectively curtailed the spread of the novel coronavirus but also had the side effect of reducing the

TABLE 3 Parameters and AICc of the alternative ARIMA models.

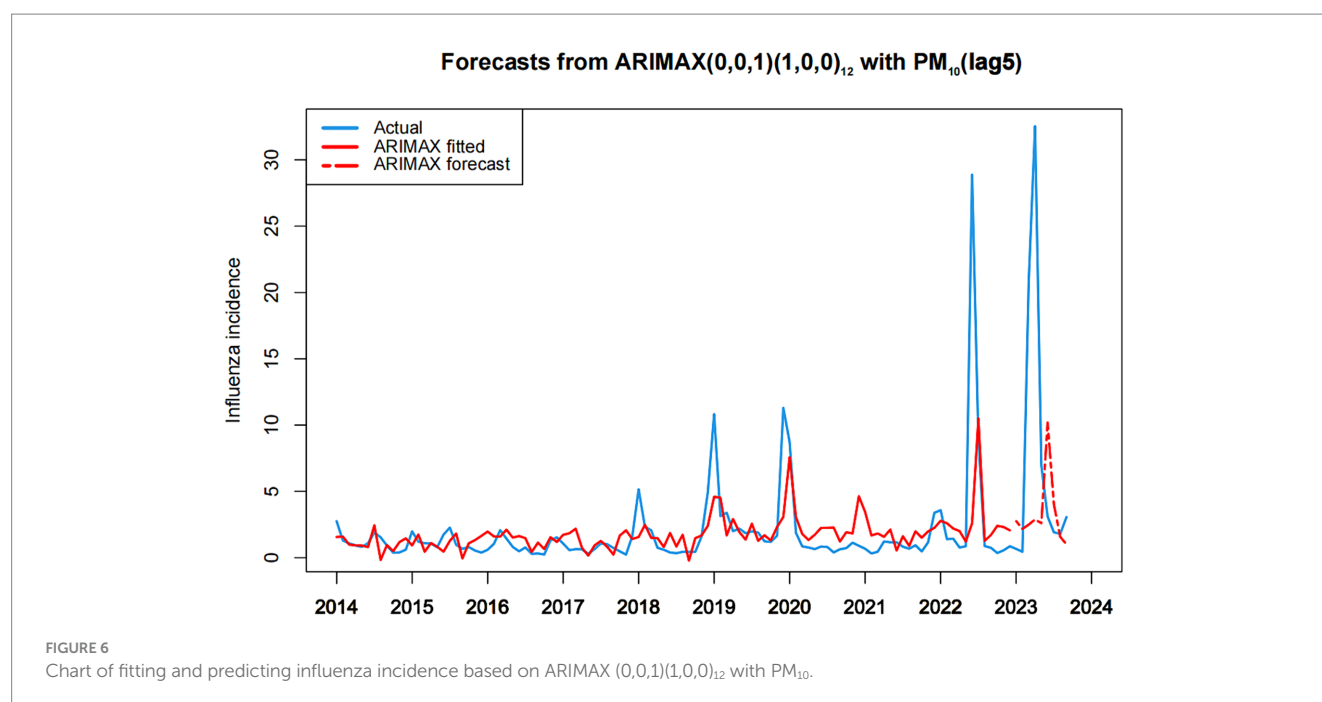
Alternative ARIMA model	AICc
ARIMA(2,0,2) (1,0,1) ₁₂	Inf
ARIMA(1,0,0) (1,0,0) ₁₂	554.387
ARIMA(0,0,1) (0,0,1) ₁₂	553.402
ARIMA(0,0,1) (1,0,1) ₁₂	555.492
ARIMA(0,0,1) (0,0,2) ₁₂	555.519
ARIMA(0,0,1) (1,0,0)₁₂	553.303
ARIMA(0,0,1) (2,0,0) ₁₂	555.493
ARIMA(0,0,1) (2,0,1) ₁₂	Inf
ARIMA(0,0,0) (1,0,0) ₁₂	562.445
ARIMA(1,0,1) (1,0,0) ₁₂	555.499
ARIMA(0,0,2) (1,0,0) ₁₂	555.499
ARIMA(1,0,2) (1,0,0) ₁₂	Inf
ARIMA(0,0,1) (1,0,0) ₁₂	560.726

The bold values represent the best performing models and parameters. Inf, Infinity.

TABLE 4 The performance of the ARIMA(0, 0, 1) (1, 0, 0)₁₂ and 10 ARIMAX models.

Model	Variable	MA(1)	SAR(1)	AIC	AICc	BIC
ARIMA(0,0,1) (1,0,0) ₁₂		0.503*	0.387*	552.910	553.300	563.640
ARIMA(0,0,1) (1,0,0) ₁₂ with PM _{2.5} (lag5)	−0.048*	0.516*	0.401*	533.560	534.170	546.730
ARIMA(0,0,1) (1,0,0) ₁₂ with PM ₁₀ (lag5)	−0.068*	0.520*	0.426*	529.740	530.360	542.910
ARIMA(0,0,1) (1,0,0) ₁₂ with SO ₂ (lag4)	−0.084*	0.500*	0.387*	550.580	551.180	563.950
ARIMA(0,0,1) (1,0,0) ₁₂ with NO ₂ (lag4)	−0.083	0.499*	0.388*	535.580	536.190	548.800
ARIMA(0,0,1) (1,0,0) ₁₂ with CO(lag4)	−4.254	0.503*	0.399*	536.560	537.170	549.780
ARIMA(0,0,1) (1,0,0) ₁₂ with O ₃ (lag3)	0.031	0.515*	0.396*	539.990	540.600	553.260
ARIMA(0,0,1) (1,0,0) ₁₂ with Ave.temp(lag3)	0.007*	0.519*	0.400*	542.630	543.240	555.900
ARIMA(0,0,1) (1,0,0) ₁₂ with Max.temp(lag3)	0.023*	0.516*	0.397*	542.550	543.160	555.820
ARIMA(0,0,1) (1,0,0) ₁₂ with Min.temp(lag3)	−0.012*	0.524*	0.405*	542.620	543.220	555.890
ARIMA(0,0,1) (1,0,0) ₁₂ with Ave.ws(lag2)	−0.067	0.521*	0.387*	546.680	547.280	560.000

*P<0.05; Ave.temp, average temperature; Max.temp, maximum temperature; Min.temp, minimum temperature; Ave.ws, average wind speed. The bold values represent the best performing models and parameters.



transmission of influenza (40). Interestingly, our analysis showed a positive association between the third-order lagged variable of O_3 and influenza incidence. This positive correlation may be attributed to high concentrations of O_3 inducing lung inflammation (41), which weakens the immune system and heightens susceptibility to infections. Furthermore, O_3 might also influence the pathogen's transmission mode, potentially rendering it more prone to airborne transmission.

The analysis revealed that influenza incidence demonstrated a negative association with three distinct temperature variables, indicating that the higher the temperature, the lower the influenza incidence. The intricacies of this relationship become more pronounced when accounting for the temperature's delayed effects. The third-order lagged temperature variable demonstrated a significant positive correlation with influenza incidence. This observed pattern could be indicative of the seasonal pattern of influenza virus transmission, further complicated by temperature's influence on human behavior and immune responses. The transmission of the influenza virus may exhibit nuanced seasonal variations, influenced by changing atmospheric temperatures (42). While increasing temperatures generally correlate with reduced influenza incidence, the full manifestation of this trend may experience delays due to the time-sensitive nature of human immune and behavioral adjustments. This suggests that people may still be at risk of spreading the flu virus for some time after the temperatures rise. Notably, behavioral patterns also shift in response to seasonal temperature changes. During warmer periods, increased outdoor activities and social interactions could inadvertently amplify influenza transmission risks, potentially leading to a spike in cases as temperatures rise. In relation to average wind speed, while the mean value demonstrated no significant correlation with influenza, the second-order lagged wind speed showed a significant negative correlation with influenza incidence, indicating that wind

speed also has a long-term lag negative correlation effect on influenza incidence.

We utilized time series analysis to examine the correlation between influenza incidence and environmental indicators in Fuzhou from January 2014 to September 2023. The environmental indicators encompassed air pollution variables (PM_{2.5}, PM₁₀, SO₂, CO, NO₂, and O_3) and meteorological factors (mean temperature, minimum temperature, maximum temperature, and wind speed). In our study, the time series data of influenza incidence in Fuzhou from January 2014 to September 2023 were found to be stationary and exhibited seasonal distribution. However, since the model used in the study was able to effectively capture the seasonal effects, there was no need to difference the time series data of influenza incidence. We also experimented with introducing seasonal differences in the time series data of influenza incidence; however, we observed that this adjustment did not lead to an improvement in the model's performance. Therefore, the data of influenza incidence were not processed by differencing in this study. First, the ARIMA(0, 0, 1) (1, 0, 0)₁₂ model was identified as the most optimal ARIMA model for forecasting influenza incidence in Fuzhou, with AIC, AICc, and BIC values of 552.910, 553.300, and 563.640, respectively. This model was employed to fit the training set, yielding a fitting RMSE of 3.002. Subsequently, the model was utilized for prediction analysis on the test set, yielding a predicting RMSE of 12.475. To enhance prediction accuracy, the maximum lag correlation variables of environmental indicators during the study period were incorporated into the optimal ARIMA model. The results demonstrated that the AIC, AICc, and BIC values of the 10 ARIMAX models, each including a single environmental index, were lower than those of the ARIMA(0, 0, 1) (1, 0, 0)₁₂ model. This suggested that considering environmental indicators could enhance the predictive performance of the model. Comparing the AIC, AICc, and BIC values of all ARIMAX models, the ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with

PM₁₀(lag5) model had the lowest AIC, AICc, and BIC values, specifically 529.740, 530.360, and 542.910, respectively. Moreover, this model exhibited a fitting RMSE of 2.999 and a predicting RMSE of 12.033, both of which were superior to the optimal ARIMA model. The ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with PM₁₀(lag5) model can be effectively employed for short-term prediction of influenza incidence in Fuzhou. This approach provides a scientifically grounded basis for formulating influenza control policies and public health interventions in Fuzhou.

The findings from our study suggest several implications for further research. Firstly, there is a need to explore the specific mechanisms through which environmental factors, such as air pollution and meteorological conditions, influence influenza transmission dynamics. Additionally, future studies could investigate the applicability of the ARIMAX model in different geographical contexts and for other infectious diseases. Expanding the dataset to include more diverse populations and environmental conditions could enhance the robustness of predictive models. Lastly, interdisciplinary research integrating public health, environmental science, and epidemiology will be essential for developing comprehensive strategies to mitigate the impact of influenza and improve public health preparedness.

In our study, we examined both ARIMA and ARIMAX modeling approaches to analyze influenza incidence in Fuzhou. The strengths of the ARIMA model include its simplicity and strong theoretical foundation, making it effective for stationary time series data. However, it does not account for external factors, which can limit its explanatory power. On the other hand, the ARIMAX model allows for the incorporation of exogenous variables, enhancing predictive accuracy and capturing lagged effects, which is crucial for understanding the impact of environmental indicators. Nevertheless, the ARIMAX model introduces complexity and relies heavily on the quality of data for the exogenous variables, which can pose challenges in interpretation and model validation. Ultimately, the ARIMAX model provided a more comprehensive analysis for our research questions. While the ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with PM₁₀(lag5) model incorporating environmental indicators provides valuable insights into the relationship between these factors and influenza incidence, it is essential to acknowledge its limitations. Firstly, the model relies heavily on historical data, which may not capture sudden changes in environmental conditions or emerging infectious disease patterns. Additionally, while environmental indicators such as air pollution and meteorological factors are significant, they are not the sole determinants of influenza occurrence. Biological factors, human behavior, and public health interventions also play crucial roles. Thus, while our statistical analysis demonstrates a correlation, it does not imply causation, and the model's predictions should be interpreted with caution. Therefore, while our findings suggest a potential relationship, further research, including controlled studies and experimental designs, is necessary to establish definitive causal links between environmental pollution factors and influenza incidence. Additional, future research should consider integrating biological and socio-economic factors to enhance the comprehensiveness of predictive models.

5 Conclusion

The incidence of influenza in Fuzhou has shown a significant increase in the past decade. Our study indicates that air pollution and meteorological factors exert an impact on influenza incidence, often exhibiting a lag effect. The ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with PM₁₀(lag5) model was developed using historical data on influenza incidence and air pollutant levels in Fuzhou, demonstrated excellent predictive performance for forecasting influenza incidence. Therefore, the ARIMAX(0, 0, 1) (1, 0, 0)₁₂ with PM₁₀(lag5) model could provide a scientific basis for the formulation of influenza control policies and public health interventions in Fuzhou.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

XZhe: Supervision, Data curation, Writing – original draft, Validation, Formal analysis. QC: Visualization, Writing – original draft, Validation, Formal analysis. MS: Writing – review & editing, Validation. QZ: Writing – review & editing, Formal analysis, Data curation. HS: Writing – review & editing, Validation. XZha: Writing – review & editing, Supervision, Resources, Conceptualization. YX: Writing – review & editing, Visualization, Supervision, Resources, Project administration, Conceptualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was financially supported by Fuzhou Science and Technology Major Project (2019-SZ-63, 2020-Z-5, and 2022-S-032), and Fujian Provincial Health and Family Planning Commission, China (2021Z01001).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Macias AE, McElhaney JE, Chaves SS, Nealon J, Nunes MC, Samson SI, et al. The disease burden of influenza beyond respiratory illness. *Vaccine*. (2021) 39:A6–A14. doi: 10.1016/j.vaccine.2020.09.048
- Moriyama M, Hugentobler WJ, Iwasaki A. Seasonality of respiratory viral infections. *Annu Rev Virol*. (2020) 7:83–101. doi: 10.1146/annurev-virology-012420-022445
- Zhang R, Li Y, Bi P, Wu S, Peng Z, Meng Y, et al. Seasonal associations between air pollutants and influenza in 10 cities of southern China. *Int J Hyg Environ Health*. (2023) 252:114200. doi: 10.1016/j.ijheh.2023.114200
- Zhou L, Yang H, Pan W, Xu J, Feng Y, Zhang W, et al. Association between meteorological factors and the epidemics of influenza (sub)types in a subtropical basin of Southwest China. *Epidemics*. (2022) 41:100650. doi: 10.1016/j.epidem.2022.100650
- Jia P, Liu S, Yang S. Innovations in public health surveillance for emerging infections. *Annu Rev Public Health*. (2023) 44:55–74. doi: 10.1146/annurev-publichealth-051920-093141
- Cobey S. Modeling infectious disease dynamics. *Science*. (2020) 368:713–4. doi: 10.1126/science.abb5659
- Maze MJ, Sharples KJ, Allan KJ, Biggs HM, Cash-Goldwasser S, Galloway RL, et al. Estimating acute human leptospirosis incidence in northern Tanzania using sentinel site and community behavioural surveillance. *Zoonoses Public Health*. (2020) 67:496–505. doi: 10.1111/zph.12712
- Guo X, Shen H, Liu S, Xie N, Yang Y, Jin J. Predicting the trend of infectious diseases using grey self-memory system model: a case study of the incidence of tuberculosis. *Public Health*. (2021) 201:108–14. doi: 10.1016/j.puhe.2021.09.025
- Zhang L, Zheng Y, Wang K, Zhang X, Zheng Y. An optimized Nash non-linear grey Bernoulli model based on particle swarm optimization and its application in prediction for the incidence of hepatitis B in Xinjiang. *China Comput Biol Med*. (2014) 49:67–73. doi: 10.1016/j.combiomed.2014.02.008
- Xu B, Li J, Wang M. Epidemiological and time series analysis on the incidence and death of AIDS and HIV in China. *BMC Public Health*. (2020) 20:1906. doi: 10.1186/s12889-020-09977-8
- Ruiz MS, O'Rourke A, Allen ST. Impact evaluation of a policy intervention for HIV prevention in Washington. *DC AIDS Behav*. (2016) 20:22–8. doi: 10.1007/s10461-015-1143-6
- Wei W, Wang G, Tao X, Luo Q, Chen L, Bao X, et al. Time series prediction for the epidemic trends of monkeypox using the ARIMA, exponential smoothing, GM (1, 1) and LSTM deep learning methods. *J Gen Virol*. (2023) 104:839. doi: 10.1099/jgv.0.001839
- Wang Y, Yan Z, Wang D, Yang M, Li Z, Gong X, et al. Prediction and analysis of COVID-19 daily new cases and cumulative cases: time series forecasting and machine learning models. *BMC Infect Dis*. (2022) 22:495. doi: 10.1186/s12879-022-07472-6
- Garcia KKS, Abrahão AA, Oliveira AFM, Henriques KMD, de Pina-Costa A, Siqueira AM, et al. Malaria time series in the extra-Amazon region of Brazil: epidemiological scenario and a two-year prediction model. *Malar J*. (2022) 21:157. doi: 10.1186/s12936-022-04162-1
- Liao Z, Lan P, Fan X, Kelly B, Innes A, Liao Z. SIRVD-DL: a COVID-19 deep learning prediction model based on time-dependent SIRVD. *Comput Biol Med*. (2021) 138:104868. doi: 10.1016/j.combiomed.2021.104868
- Wu Z, Chen Z, Long S, Wu A, Wang H. Incidence of pulmonary tuberculosis under the regular COVID-19 epidemic prevention and control in China. *BMC Infect Dis*. (2022) 22:641. doi: 10.1186/s12879-022-07620-y
- Ahn SM, Eun S, Ji S, Hong S, Lee CK, Yoo B, et al. Incidence of rheumatic diseases during the COVID-19 pandemic in South Korea. *Korean J Intern Med*. (2023) 38:248–53. doi: 10.3904/kjim.2022.135
- Peng R, Wang Y, Zhai J, Zhang J, Lu Y, Yi H, et al. Driving effect of multiplex factors on human brucellosis in high incidence region, implication for brucellosis based on one health concept. *One Health*. (2022) 15:100449. doi: 10.1016/j.onehlt.2022.100449
- Abdulsalam FI, Antunez P, Yimthiang S, Jawjit W. Influence of climate variables on dengue fever occurrence in the southern region of Thailand. *PLOS Glob Public Health*. (2022) 2:e0000188. doi: 10.1371/journal.pgph.0000188
- Nelson BK. Statistical methodology: V. Time series analysis using autoregressive integrated moving average (ARIMA) models. *Acad Emerg Med*. (1998) 5:739–44. doi: 10.1111/j.1553-2712.1998.tb02493.x
- Wang M, Jiang Z, You M, Wang T, Ma L, Li X, et al. An autoregressive integrated moving average model for predicting varicella outbreaks – China, 2019. *China CDC Wkly*. (2023) 5:698–702. doi: 10.46234/ccdcw2023.134
- Li ZQ, Pan HQ, Liu Q, Song H, Wang JM. Comparing the performance of time series models with or without meteorological factors in predicting incident pulmonary tuberculosis in eastern China. *Infect Dis Poverty*. (2020) 9:151. doi: 10.1186/s40249-020-00771-7
- Ilu SY, Prasad R. Improved autoregressive integrated moving average model for COVID-19 prediction by using statistical significance and clustering techniques. *Heliyon*. (2023) 9:e13483. doi: 10.1016/j.heliyon.2023.e13483
- Hossain MS, Ahmed S, Uddin MJ. Impact of weather on COVID-19 transmission in south Asian countries: an application of the ARIMAX model. *Sci Total Environ*. (2021) 761:143315. doi: 10.1016/j.scitotenv.2020.143315
- Zheng Y, Wang K, Zhang L, Wang L. Study on the relationship between the incidence of influenza and climate indicators and the prediction of influenza incidence. *Environ Sci Pollut Res Int*. (2021) 28:473–81. doi: 10.1007/s11356-020-10523-7
- Ruchiraset A, Tantrakarnapa K. Time series modeling of pneumonia admissions and its association with air pollution and climate variables in Chiang Mai Province, Thailand. *Environ Sci Pollut Res Int*. (2018) 25:3277–85. doi: 10.1007/s11356-018-3284-4
- Kuipers MAG, Beard E, West R, Brown J. Associations between tobacco control mass media campaign expenditure and smoking prevalence and quitting in England: a time series analysis. *Tob Control*. (2018) 27:455–62. doi: 10.1136/tobaccocontrol-2017-053662
- Uyeki TM, Hui DS, Zambon M, Wentworth DE, Monto AS. Influenza. *Lancet*. (2022) 400:693–706. doi: 10.1016/S0140-6736(22)00982-5
- Jing QL, Cheng Q, Marshall JM, Hu WB, Yang ZC, Lu JH. Imported cases and minimum temperature drive dengue transmission in Guangzhou, China: evidence from ARIMAX model. *Epidemiol Infect*. (2018) 146:1226–35. doi: 10.1017/S0950268818001176
- Bostan N, Javed S, Nabgha-E-Amen ESA, Tahir F, Bokhari H. Dengue fever virus in Pakistan: effects of seasonal pattern and temperature change on distribution of vector and virus. *Rev Med Virol*. (2017) 27:e1899. doi: 10.1002/rmv.1899
- Roy MP. Temperature and COVID-19: India. *BMJ Evid Based Med*. (2021) 26:e1. doi: 10.1136/bmjebm-2020-111459
- Bronte O, García-García F, Lee DJ, Urrutia I, Uranga A, Nieves M, et al. Impact of outdoor air pollution on severity and mortality in COVID-19 pneumonia. *Sci Total Environ*. (2023) 894:164877. doi: 10.1016/j.scitotenv.2023.164877
- Liu M, Li Z, Liu M, Zhu Y, Liu Y, Kuetche MWN, et al. Association between temperature and COVID-19 transmission in 153 countries. *Environ Sci Pollut Res Int*. (2022) 29:16017–27. doi: 10.1007/s11356-021-16666-5
- Maharjan B, Gopali RS, Zhang Y. A scoping review on climate change and tuberculosis. *Int J Biometeorol*. (2021) 65:1579–95. doi: 10.1007/s00484-021-02117-w
- Yang J, Yang Z, Qi L, Li M, Liu D, Liu X, et al. Influence of air pollution on influenza-like illness in China: a nationwide time-series analysis. *EBioMedicine*. (2023) 87:104421. doi: 10.1016/j.ebiom.2022.104421
- Kim J, Kim H, Kweon J. Hourly differences in air pollution on the risk of asthma exacerbation. *Environ Pollut*. (2015) 203:15–21. doi: 10.1016/j.envpol.2015.03.040
- Li Y, Wu J, Hao J, Dou Q, Xiang H, Liu S. Short-term impact of ambient temperature on the incidence of influenza in Wuhan, China. *Environ Sci Pollut Res Int*. (2022) 29:18116–25. doi: 10.1007/s11356-021-16948-y
- Loaiza-Ceballos MC, Marin-Palma D, Zapata W, Hernandez JC. Viral respiratory infections and air pollutants. *Air Qual Atmos Health*. (2022) 15:105–14. doi: 10.1007/s11869-021-01088-6
- Wang X, Wang X, Guan X, Xu Y, Xu K, Gao Q, et al. The impact of ambient air pollution on an influenza model with partial immunity and vaccination. *Math Biosci Eng*. (2023) 20:10284–303. doi: 10.3934/mbe.2023451
- Nielsen RT, Dalby T, Emborg HD, Larsen AR, Petersen A, Torpdahl M, et al. COVID-19 preventive measures coincided with a marked decline in other infectious diseases in Denmark, spring 2020. *Epidemiol Infect*. (2022) 150:e138. doi: 10.1017/S0950268822001145
- Kilburg-Basnyat B, Reece SW, Crouch MJ, Luo B, Boone AD, Yaeger M, et al. Specialized pro-resolving lipid mediators regulate ozone-induced pulmonary and systemic inflammation. *Toxicol Sci*. (2018) 163:466–77. doi: 10.1093/toxsci/kfy040
- Xu Z, Hu W, Williams G, Clements AC, Kan H, Tong S. Air pollution, temperature and pediatric influenza in Brisbane, Australia. *Environ Int*. (2013) 59:384–8. doi: 10.1016/j.envint.2013.06.022



OPEN ACCESS

EDITED BY

Jiaying Li,
The University of Queensland, Australia

REVIEWED BY

Pankaj Bharali,
North East Institute of Science and
Technology (CSIR), India
Mahesh Dharne,
National Chemical Laboratory (CSIR), India

*CORRESPONDENCE

Gebremedhin Gebremicael
✉ kudus12@gmail.com

[†]These authors have contributed equally to
this work

RECEIVED 02 March 2024

ACCEPTED 16 September 2024

PUBLISHED 07 October 2024

CITATION

Abera Dinssa D, Gebremicael G, Mengistu Y,
Hull NC, Chalchisa D, Berhanu G, Gebreegziabxier A, Norberg A, Snyder S, Wright S,
Gobena W, Abera A, Belay Y, Chala D,
Gizaw M, Getachew M, Tesfaye K, Tefera M,
Belachew M, Mulu T, Ali S, Kebede A,
Melese D, Abdella S, Rinke de Wit TF,
Kebede Y, Hailu M, Wolday D, Tessema M and
Tollera G (2024) Longitudinal
wastewater-based surveillance of
SARS-CoV-2 during 2023 in Ethiopia.
Front. Public Health 12:1394798.
doi: 10.3389/fpubh.2024.1394798

COPYRIGHT

© 2024 Abera Dinssa, Gebremicael, Mengistu,
Hull, Chalchisa, Berhanu, Gebreegziabxier,
Norberg, Snyder, Wright, Gobena, Abera,
Belay, Chala, Gizaw, Getachew, Tesfaye,
Tefera, Belachew, Mulu, Ali, Kebede, Melese,
Abdella, Rinke de Wit, Kebede, Hailu, Wolday,
Tessema and Tollera. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Longitudinal wastewater-based surveillance of SARS-CoV-2 during 2023 in Ethiopia

Daniel Abera Dinssa^{1†}, Gebremedhin Gebremicael^{1*†},
Yohannes Mengistu², Noah C. Hull³, Dinknesh Chalchisa¹,
Girma Berhanu¹, Atsbeha Gebreegziabxier¹, Ashley Norberg³,
Sarah Snyder³, Sarah Wright⁴, Waktale Gobena¹, Adugna Abera¹,
Yohannes Belay¹, Dawit Chala¹, Melaku Gizaw¹,
Mesay Getachew¹, Kirubel Tesfaye¹, Mesfin Tefera¹,
Mahlet Belachew¹, Tegegne Mulu¹, Solomon Ali⁵,
Abebaw Kebede⁶, Daniel Melese¹, Saro Abdella¹,
Tobias F. Rinke de Wit⁷, Yenew Kebede⁶, Mesay Hailu¹,
Dawit Wolday⁸, Masresha Tessema¹ and Getachew Tollera¹

¹Ethiopian Public Health Institute (EPHI), Addis Ababa, Ethiopia, ²Global Health, The Association of Public Health Laboratories (APHL), Addis Ababa, Ethiopia, ³Global Health and Environmental Health, The APHL, Bethesda, MD, United States, ⁴Environmental Health, The APHL, Bethesda, MD, United States, ⁵Department of Microbiology, Immunology and Parasitology, St. Paul's Hospital Millennium Medical College, Addis Ababa, Ethiopia, ⁶Africa Centres for Disease Control and Prevention (Africa CDC), Surveillance and Disease Intelligence Division, Addis Ababa, Ethiopia, ⁷Amsterdam Institute of Global Health and Development, Department of Global Health, Amsterdam University Medical Center, Amsterdam, Netherlands, ⁸Department of Biochemistry and Biomedical Sciences, Michael G. DeGroote Institute for Infectious Diseases Research and McMaster Immunology Research Center, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

Introduction: Although wastewater-based epidemiology (WBE) successfully functioned as a tool for monitoring the coronavirus disease 2019 (COVID-19) pandemic globally, relatively little is known about its utility in low-income countries. This study aimed to quantify severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA in wastewater, estimate the number of infected individuals in the catchment areas, and correlate the results with the clinically reported COVID-19 cases in Addis Ababa, Ethiopia.

Methods: A total of 323 influent and 33 effluent wastewater samples were collected from three Wastewater Treatment Plants (WWTPs) using a 24-h composite Moore swab sampling method from February to November 2023. The virus was captured using Ceres Nanotrap® Enhancement Reagent 2 and Nanotrap® Microbiome A Particles, and then nucleic acids were extracted using the Qiagen QIAamp Viral RNA Mini Kit. The ThermoFisher TaqPath™ COVID-19 kit was applied to perform real-time reverse transcriptase polymerase chain reaction (qRT-PCR) to quantify the SARS-CoV-2 RNA. Wastewater viral concentrations were normalized using flow rate and number of people served. In the sampling period, spearman correlation was used to compare the SARS-CoV-2 target gene concentration to the reported COVID-19 cases. The numbers of infected individuals under each treatment plant were calculated considering the target genes' concentration, the flow rate of treatment plants, a gram of feces per person-day, and RNA copies per gram of feces.

Results: SARS-CoV-2 was detected in 94% of untreated wastewater samples. All effluent wastewater samples ($n = 22$) from the upflow anaerobic sludge blanket

(UASB) reactor and membrane bioreactor (MBR) technology were SARS-CoV-2 RNA negative. In contrast, two out of 11 effluents from Waste Stabilization Pond were found positive. Positive correlations were observed between the weekly average SARS-CoV-2 concentration and the cumulative weekly reported COVID-19 cases in Addis Ababa. The estimated number of infected people in the Kality Treatment catchment area was 330 times the number of COVID-19 cases reported during the study period in Addis Ababa.

Discussion: This study revealed that SARS-CoV-2 was circulating in the community and confirmed previous reports of more asymptomatic COVID-19 cases in Ethiopia. Additionally, this study provides further evidence of the importance of wastewater-based surveillance in general to monitor infectious diseases in low-income settings.

Conclusion: Wastewater-based surveillance of SARS-CoV-2 can be a useful method for tracking the increment of COVID-19 cases before it spreads widely throughout the community.

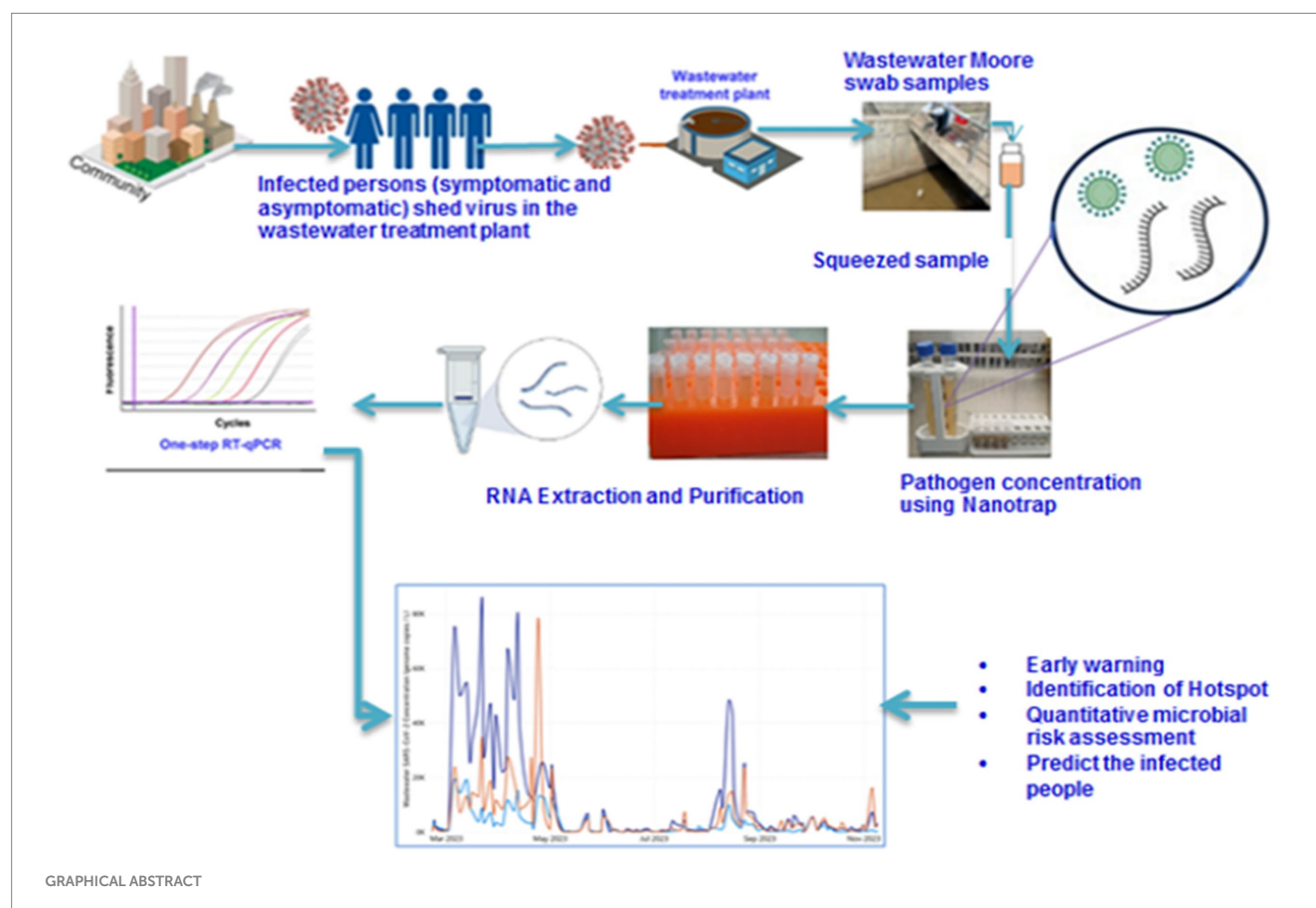
KEYWORDS

COVID-19, SARS-CoV-2, qRT-qPCR, wastewater treatment plants, wastewater-based epidemiology

Background

Economic stability and human health are considerably affected by infectious diseases as they cause one-fourth of the mortalities around the world (1). The recent outbreak of COVID-19 caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)

escalated into a global pandemic since it first appeared in Wuhan, China, in December 2019 (2). In January 2020, it led to a declaration of a Public Health Emergency of International Concern by the World Health Organization (WHO) (3). Since then, SARS-CoV-2 has been responsible for more than 773 million confirmed cases and around 7 million deaths worldwide as of December 2023 (4). In this



regard, Africa reported only around 1.2% of confirmed cases and 2.5% of deaths. The first COVID-19 case in the African continent was reported from Egypt on the 14th of February 2020 (5). On February 25, Nigeria became the second country to report a first case, and on February 27, Algeria became the third country to do so (6). The first cases in other African countries, including Ethiopia, were only detected in March 2020 (7). Most index cases originated in Europe, where the epidemic's epicenter had moved by March 13. As a result, the pandemic spread quickly to Africa (8). Consequently, this led to long-lasting collateral damage on the continent from interruptions in the initiatives for TB, HIV/AIDS, malaria, and vaccine-preventable illnesses (9). Ethiopia reported around 5 and 4.3% of the African total confirmed cases and deaths, respectively (4).

Surveillance focused on clinical and laboratory testing which has drawbacks such as excessive costs, failure to detect asymptomatic patients, and underestimating of infection prevalence (10). Current data suggest that worldwide 35–45% of all SARS-CoV-2 infections account for asymptomatic infected persons (11–13). However, the percentage of asymptomatic cases in Africa and Ethiopia is 67 and 74%, respectively (14–16). Recent study findings in Ethiopia indicated that high asymptomatic cases are associated with persistently activated immune system (17, 18). This will affect the clinical COVID-19 case detection and reporting as testing of samples was prompted mainly by symptoms (19). Hence, the community may not be prepared in terms of infection prevention and control, and management of COVID-19 infection (20).

SARS-CoV-2 RNA can be detected in feces and urine from asymptomatic and symptomatic individuals. Fecal shedding can persist for several weeks, typically longer than positivity in oropharyngeal swabs (21, 22). The extended presence of viral RNA in feces and fecal viral RNA shedding with gastrointestinal (GI) symptoms implies that SARS-CoV-2 infects the GI tract (23–25). Anyhow, virus shedding in the feces of symptomatic and asymptomatic infected individuals enables the detection of viral RNA in influent sewage or wastewater (26, 27). Wastewater-based epidemiology (WBE) for COVID-19 surveillance can be used as an alternative for early warning of COVID-19 outbreaks or as a control mechanism for potential virus transmission independent of individual healthcare-seeking behaviors. In addition, WBE can be scaled relatively easily, is less expensive than human subject testing, and, if collected at strategic points, can represent local populations (28, 29). Monitoring SARS-CoV-2 circulation in the community will remain important for reinforcing preparedness and identifying hotspots for further classical surveillance interventions, particularly in regions with inadequate health system infrastructure, human resources, and testing capacity.

Previously, numerous human infectious illnesses (such as polio and typhoid) have been the focus of research in this WBE (30, 31). In high-income countries, wastewater-based surveillance is well utilized for the monitoring of SARS-CoV-2 (32). However, few African countries have conducted wastewater-based SARS-CoV-2 surveillance (33–35). This may be partly attributable to low sewage coverage with deficient testing coverage, which limits COVID-19 surveillance through sewage monitoring (36). In Ethiopia, there is only one study in wastewater-based SARS-CoV-2 using a small sample size, and it is focused only on the qualitative test (37). This study aimed to quantify SARS-CoV-2 RNA in wastewater, estimate the number of infected individuals in the catchment area, and correlate results with clinically reported COVID-19 cases in Addis Ababa, Ethiopia.

Materials and methods

Study setting and sampling sites

Addis Ababa is Ethiopia's capital city, with an estimated 5,460,591 population in 2023 (38). Administratively, it is divided into 11 subcities. Based on the information obtained from Addis Ababa Water and Sewerage Authority (AAWSA), the wastewater treatment capacity in Addis Ababa is nearly 86%, out of which 34% are currently connected to sewer lines, and 52% rely on vacuum trucks, the remaining could be considered as illegal connection or disposal. Currently, Addis Ababa city has 4 centralized and 35 decentralized wastewater treatment plants (WWTPs). The centralized WWTPs are Kality, Kality old, Kotebe old, and Chefe (unpublished Strategic Environmental and Social Assessment [SESA] of Addis Ababa City Sanitation Master Plan, 2024).

Influent wastewater samples were collected 3 times a week from three sampling sites (Kality, Bulbula, and Mikililand) using the Moore swab method (39) (Figure 1). Kality treatment plant (KTP) is the oldest centralized system, mainly serving residents in the central, southern, and eastern parts of the city (40) with an estimated population coverage of nearly 2,000,000 (unpublished data from AAWSA). The upflow anaerobic sludge blanket reactor (UASB) technology is applied at this site. A membrane Bioreactor (MBR) wastewater treatment technology, which combines a biological-activated sludge process and membrane filtration domestic wastewater treatment, is used at the Bulbula wastewater treatment site (41). The third wastewater treatment plant included in this study was Mikililand Waste Stabilization Pond (WSP). Mikililand WSP systems comprise 7 series of different types of ponds (42). It is situated in the northwestern part of the capital city. Technical details of the wastewater treatment process at the three wastewater treatment plants are presented in Table 1.

Study design and sample collection

A longitudinal study design was conducted between February and November 2023 at three wastewater treatment plants in Addis Ababa. The Moore swab, or cotton gauze of size (120 × 15 cm), was folded to achieve an 8-ply pad and tied with a string that was long enough to immerse the swab into the influent discharge (39, 43). The prepared Moore swab was then autoclaved and sealed in Ziploc® bag. The string was attached to a solid structure and fully submersed into the wastewater. On all three wastewater collecting sites, the swab installation period was between 9:30 a.m. and 11:30 a.m. on Sundays, Tuesdays, and Wednesdays of each week. Following a 24-h period for the installation of the swab, the wastewater from the submersion was collected and placed in a Ziploc® bag. Finally, the exterior of Ziploc® bag was decontaminated with 70% ethanol and then transported using an ice-cold box to the Ethiopian Public Health Institute's laboratory. Accordingly, 323 influent and 33 effluent wastewater samples were collected. The influent samples were composed of 110 from Kality, 108 from Bulbula, and 105 from Mikililand treatment plants. Whereas, 11 effluent samples were collected from each of the three treatment plants with the same installation time of influent samples in October 2023 and November 2023.

Furthermore, to evaluate the effectiveness of the Moore swab sampling method in capturing virus particles from wastewater, influent samples were collected in parallel, covering the same 24-h period for 3 weeks from February 2023 to March 2023 using an on-site autosampler placed at KTP ($n=8$).

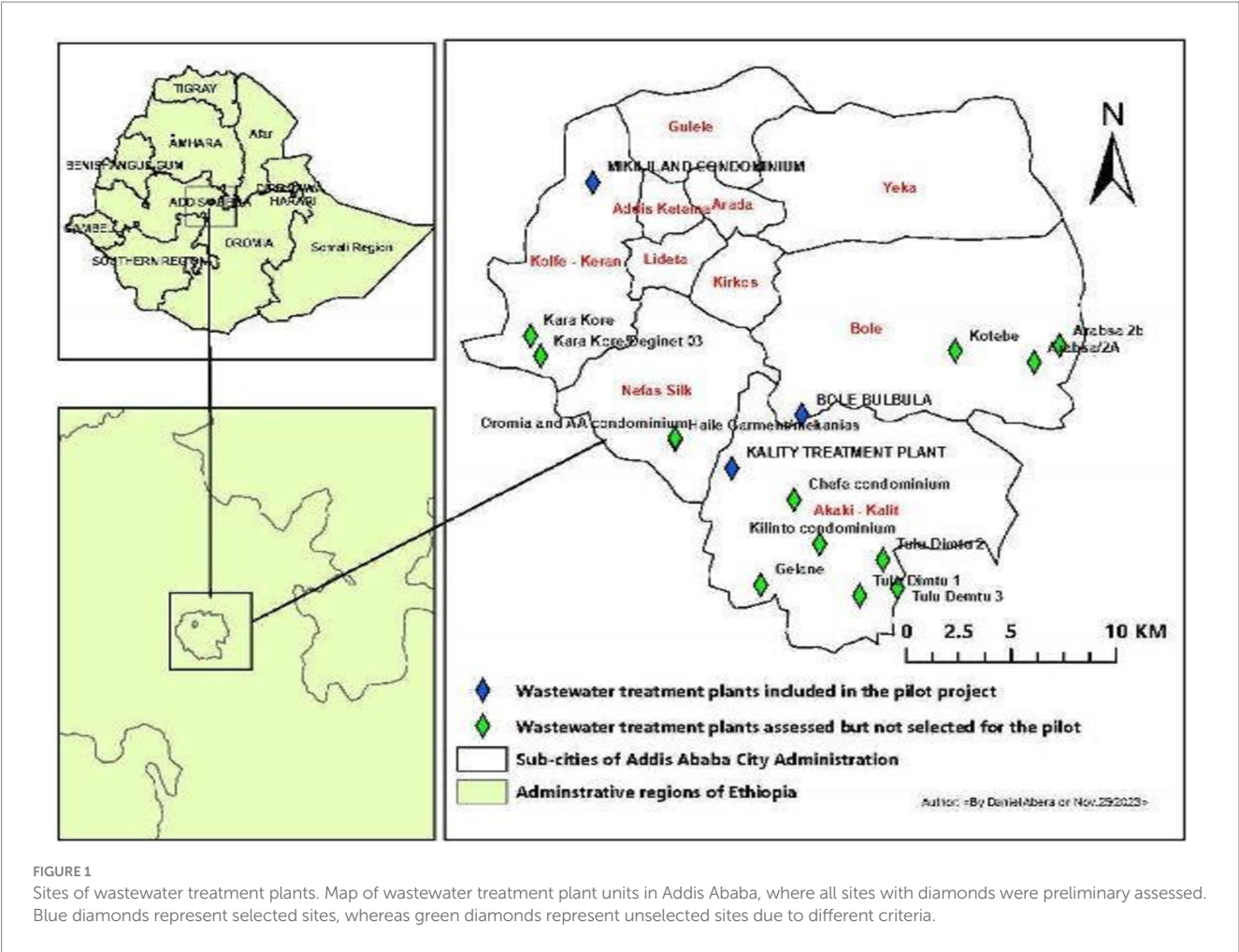


TABLE 1 Description of the selected WWTPs.

	WWTP name	Sub-city/location	Design capacity m ³ per day	Average daily flow rate in m ³ per day	Served population	Type of treatment technology in place
1	Kality treatment plant (KTP)	Akaki Kality Sub-City/ Southern Addis Ababa	100,000 ^a	65,000	2,000,000	USAB
2	Bulbula Treatment Plant	Bole Sub-City/ Central Addis Ababa	20,000 ^b	325	34,000	MBR
3	Mikililand Wastewater Stabilization Pond (WSP)	Kolfe-Keranio Sub-City/ Northwest Addis Ababa	3,000 ^c	NA	24,000	WSP

Description of the selected WWTPs for this surveillance of SARS-CoV-2 in Addis Ababa, Ethiopia. ^aInformation was collected from the AAWSA authority. ^{b,c}Information was from the literature that is cited (37) and (42). NA = Not available (we did not know the average flow rate due to lack of measuring instruments at the WSP).

RNA capture and extraction

Each Moore swab was squeezed of all liquid into a sterile container from which a 10 ml wastewater aliquot was taken using a 15 ml tube for RNA capture and extraction. For RNA concentration, 100µl of Nanotrap® Enhancement Reagent 2 (ER2; SKU# 10112, Ceres Nanoscience, Inc., Manassas, VA) and 150µl of Nanotrap® Microbiome A Particles (SKU#44202, Ceres Nanosciences, Inc., Manassas, VA, USA) were added into the 15-ml tube containing 10-ml wastewater and mixed well. After samples were incubated at room temperature for 10 min, Nanotrap® Microbiome A Particles pellet was separated using a

DynaMag™-15 Magnet (Thermo Fisher Scientific, Waltham, MA, USA). After washing the pellet, 150µl of 1× phosphate-buffered saline (PBS) for suspension and 5µl of MS2 phage control were added to each pellet, and negative control (RNase free water); then RNA extraction was executed using QIAamp Viral RNA Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer’s instruction (44). MS2 spike-in to each sample can minimize false negatives. Briefly, 560µl QIAGEN Virus Lysis Buffer was added to PBS suspended pellet to lyse the cells. Following a 10-min incubation at room temperature of the solution, the Nanotrap® Microbiome A Particles and the lysate solution sample were separated using the DynaMag™-2 magnet (Thermo Fisher Scientific,

Waltham, MA, USA). The lysate supernatants were collected in a new 1.5-ml microcentrifuge tube, and the pellet was discarded. For high nucleic acid concentration, 560 µl of 100% ethanol was added to the concentrate, and the lysate was added to the QIAamp Mini column. After washing using wash buffer, the QIAamp Mini column was placed in a clean 1.5-ml microcentrifuge tube, elution was performed using 80-µl Buffer AVE, and the eluted viral RNA was stored at −80°C.

Real-time reverse transcriptase polymerase chain reaction (qRT-PCR)

The TaqPath™ COVID-19 control was used as a quantification standard RNA control (1×10^4 copies/µl stock). A 10^4 copies/µl was diluted to 2×10^3 copies/µl using dilution buffer and then used as stock. The stock solution was then serially diluted 5-fold in low-binding 1.5-ml tubes. The limit of detection of the TaqPath™ COVID-19 is 10 genomic copy equivalents (GCE)/reaction (45), but we did not do the limit of detection in our setting.

TaqPath™ COVID-19 qRT-PCR reaction master mix was prepared according to the manufacturer's instructions (45). A total of 15-µl master mix was added to each well of the plate. Approximately 10 µl of extracted nucleic acid, quantification standard RNA, and nuclease-free water for no template control (NTC) were added to the assay wells containing the master mix. In the Plate Setup window of QuantStudio™ 5 (Thermo Fisher Scientific, Waltham, MA), FAM, VIC, ABY, and JUN dyes were used as reporter dyes for the viral targets of the primers and probes: ORF1ab, Nucleocapsid (N) gene, Spike (S) gene, and MS2 phage control, respectively (45). Thermal cycling conditions included 2-min of uracil-N-glycosylase (UNG) incubation at 25°C, 10-min of reverse transcription at 53°C, 2-min at 95°C for reverse transcription deactivation, and initial activation of Speed Star HS DNA polymerase, followed by 40 cycles of 3 s denaturation at 95°C and 30 s annealing/extension at 60°C. All samples with cycle threshold (Ct) values of ORF1ab, N gene, and S gene <37; MS2 <32 were considered positive according to the manufacturer (45).

Determination of viral concentration in wastewater

The PCR test results were interpreted as follows: when any two or more of the viral targets were reported, the sample was considered positive for SARS-CoV-2; when only one viral target was detected within repeated tests, the result was considered inconclusive; whereas when all the viral targets were not detected but the internal control (MS2) detected, the sample was considered as negative for SARS-CoV-2. Preliminary reverse transcriptase qPCR data analysis and quality control were performed using the QuantStudio Flex 5 reverse transcriptase qPCR software v1.5.1 (Applied Biosystems, Inc., USA). Viral concentrations were expressed as genome copies of RNA extract per liter (gc/L). Using Excel and the following formula, viral concentrations (gc/L) in the concentrated samples were determined:

$$\text{Concentration of viral genome in wastewater} \left(\frac{\text{copies}}{\text{L}} \right) = \frac{\text{Copies in RT - qPCR reaction (copies)}}{\text{Volume of nucleic acid extracted used for RT - qPCR (ml)} \times \text{Concentration factor}^{\text{SS}}} \times 1,000$$

*If 10 µl of the nucleic acid extract is used in RT-qPCR assay the value in ml is 0.01.

$$\text{Concentration factor}^{\text{SS}} = \frac{\text{Wastewater sample volume used (ml)}}{\text{Volume of nucleic acid extracted (ml)}}$$

Virus concentration levels (genome copies per L) were normalized by multiplying with the daily WWTP flow rate of specific WWTP and then dividing by the number of people served to get daily load/persons in sewershed [million gene copies (MGC)/person/day]. However, viral concentration levels in all samples from Mikililand WSP were only expressed as genome copies/L of RNA due to a lack of daily flow rate data.

Estimating the number of infectious individuals

The number of daily reported COVID-19 cases in Addis Ababa during the study (February to November 2023) was obtained from the Public Health Emergency Management Center at the Ethiopian Public Health Institute. The number of residents served by the WWTP was obtained from the respective Woreda offices and the Addis Ababa Water and Sewerage Authority (AAWSA) (Table 1). Using two different approaches that have been previously published, the number of infected individuals within each WWTP's service area was calculated (27, 46).

The equations used for calculation are indicated below:

Method 1 (Equation 1) (27):

$$\begin{aligned} &\text{Predicted Infected person} \\ &= \frac{\left(\frac{\text{RNA copies}}{\text{Liter wastewater}} \right) \times \left(\frac{\text{Liter of wastewater}}{\text{day}} \right)}{\left(\frac{\text{g of feces}}{\text{person - day}} \right) \times \left(\frac{\text{RNA copies}}{\text{g of feces}} \right)} \quad (1) \end{aligned}$$

A positive individual is thought to excrete 128 g of feces per person per day and shed 10^7 RNA copies per g of feces (27).

Method 2 (Equation 2) (46):

$$\begin{aligned} &\text{Predicted infected person} \\ &= \frac{\text{Number of RNA copies per liter of wastewater}}{\text{Contribution of RNA copies per person to total wastewater}} \quad (2) \end{aligned}$$

10^7 RNA copies/g of feces was multiplied by 120 ml of the volume of feces excreted by humans (considering the density of feces as 1.07 g/ml), and total wastewater (L) received at WWTP (46).

Statistical analysis

According to the Kolmogorov–Smirnov test, the viral concentration data were not normally distributed. We tested for significant differences in viral concentration (gc/ml) across sites using a Kruskal–Wallis rank sum and pairwise Wilcoxon tests. Spearman correlation was used to assess the correlation between reported cases

and viral data. All data analysis was performed using Inter cooled STATA version 14.0 (College Station, TX, USA). The graphs are presented using Power BI.

Ethical statement

Informed consent is not applicable for environmental wastewater samples as no human subject is involved. However, for the use of reported COVID-19 cases data from Addis Ababa, permission was granted from the Ministry of Health, which owns the data. For COVID-19 protection, care was taken during sample collection and analysis using personal protection equipment and a standardized method. All respective bodies (government and non-government) participated in this study adhered to the sample collection and laboratory testing protocols. In addition, this study obtained ethical clearance from the Ethiopian Public Health Institute Scientific and Ethical Review Office (Ref. EPHI 6.13/577). Official approval was obtained from AAWSA, the government body that administers Addis Ababa city's water supply and sewerage services. Access to the treatment plant and site-level information was obtained from this authority.

Results

Method optimization for wastewater-based SARS-CoV-2 detection

SARS-CoV-2 detection and quantification from wastewater samples using the Moore swab method is a relatively new method in Ethiopia, apart from its use in polio surveillance. The comparison of the on-site autosampler method in place at KTP and the Moore swab sampling technique for SARS-CoV-2 detection in wastewater is presented in Table 2.

As shown in the table, there is no significant difference in the Ct values of the target genes (ORF1ab, N, and S genes) between the autosampler and Moore swab sampling techniques. Moreover, viral concentrations (gc/L) of the target genes were not significantly different using the autosampler and Moore swab sampling technique ($p > 0.05$; Figure 2). Although the autosampler method of wastewater sampling is reliable, it has limitations that impede effective

surveillance, especially from small catchments with limited accessibility. Since Moore swab sampling is more cost-effective and requires fewer resources to process, we decided to continue our monitoring of wastewater for SARS-CoV-2 using this technique.

Wastewater-based SARS-CoV-2 qualitative test result

Wastewater samples collected from 21 February 2023 to 9 November 2023 in Addis Ababa at KTP, Bulbula WWTP, and Mikililand WSP were tested for SARS-CoV-2 by qRT-PCR. A total of 323 wastewater Moore swab samples were tested. Each run had negative controls and produced all negative results. Of these, 304 (94%) tested positive for SARS-CoV-2 by qRT-PCR, defined as a Ct value of <37 for two or more SARS-CoV-2 target genes. In addition, 14/323 (4%) of the samples tested were inconclusive for SARS-CoV-2 by qRT-PCR, defined as a Ct value of <37 for one SARS-CoV-2 target gene only in duplicate testing, and only 5/323 (2%) were negative, defined as a Ct value of ≥ 37 for three SARS-CoV-2 target genes and a Ct value of <32 for MS2 (internal control). Around 95% of samples from KTP were positive, whereas 2% were negative and 3% were inconclusive. Approximately 90% of the Bulbula samples were positive, with the remaining 3% negative and 7% inconclusive. Finally, 97% of the Mikililand samples were positive, 3% were inconclusive, and no negative results were found.

To determine the presence of SARS-CoV-2 RNA, 33 treated effluent water samples were taken from these three wastewater treatment plants. From each wastewater treatment plant, 11 treated wastewater samples were collected. All treated samples were collected in the morning from 8:00 a.m. to 12:00 p.m. by collecting 500–1,000 ml of water in sterile plastic containers. The collected samples were transported using ice and concentrated within 24 h, using the same process as influent wastewater. The SARS-CoV-2 extraction and detection procedure for treated wastewater samples was the same as for influent wastewater. Of the total 33 samples, 22 treated wastewater samples from Kaliti and Bulbula WWTP were negative, whereas two of the total treated samples from the Mikililand stabilized pond were positive. Five treated samples from the Mikililand stabilized pond were inconclusive, and the remaining four samples were negative.

TABLE 2 Comparison of Moore swab sampling technique and autosampler.

Date	ORF1ab of swab	ORF1ab of auto	N gene of swab	N gene of auto	S gene of swab	S gene of auto
23 February 2023	32.001	32.543	32.719	33.390	31.571	30.163
27 February 2023	33.361	32.085	33.638	34.101	33.663	30.564
2 March 2023	32.753	32.111	36.441	33.578	31.223	31.115
6 March 2023	30.408	29.154	28.942	28.352	30.082	29.475
9 March 2023	30.848	31.617	29.348	29.414	31.593	31.449
13 March 2023	30.443	29.882	28.884	28.241	31.170	30.266
15 March 2023	31.665	29.841	30.119	28.559	31.348	29.577
16 March 2023	31.606	31.185	30.579	29.571	30.761	31.431

The Ct-value of the target genes detected using autosampler and Moore swab sampling technique. The Ct values of the target genes between autosampler and swab sampling technique were not significantly different (ORF1ab, $p = 0.1386$; N gene, $p = 0.0858$; S gene, $p = 0.1386$).

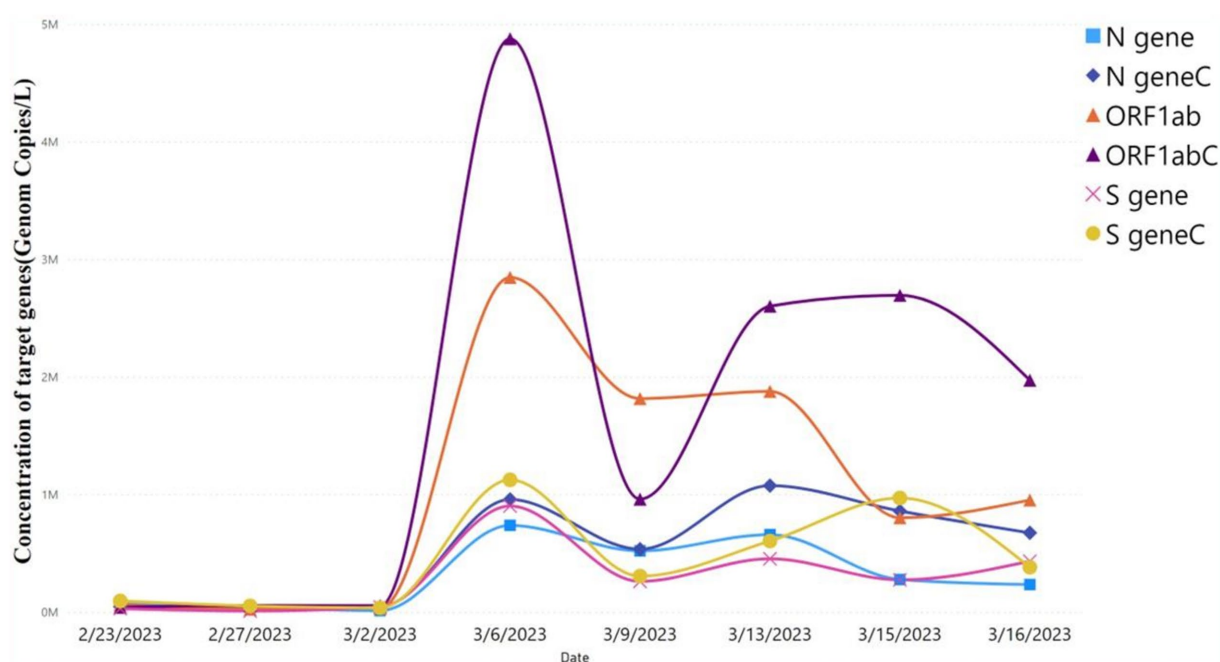


FIGURE 2

Viral concentration of autosampler and Moore swab sampling technique. Comparison of viral target genes concentration level using autosampler vs. Moore swab sampling technique. ORF1ab, N gene, and S gene were the target genes. Target genes with "C" represent the concentration of viral target genes using the autosampler, whereas target genes without "C" represent the viral target concentration using the Moore swab sampling technique. The level of concentration of the target genes predicted by the autosampler vs. swab sampling technique was not significant ($p > 0.05$).

Quantification RNA of SARS-CoV-2 in wastewater

The wastewater samples that tested positive for SARS-CoV-2 RNA by qualitative methods were subjected to quantitative PCR for three targets (ORF1ab, N gene, and S gene). The performance efficiency range of the ORF1ab, N, and S genes among the test runs was 91.8 to 105.7, 93.0 to 109.8, and 88.0 to 104.6, respectively, and the detailed results are summarized in [Supplementary Table S1](#).

The concentration of these three viral targets in the influent wastewater samples across the three wastewater treatment plants (WWTPs) is presented using a Box-Whisker plot ([Figure 3A](#)). The median viral concentration (gc/L) and (interquartile range [IQR]) obtained for ORF1ab, N gene, and S gene in positive samples from KTP was 60,388 (21,544–430,339), 26,355 (7,748–125,372) and 6,2573 (12,221–24,9,039), respectively. Similarly, the median viral concentration (gc/L) and IQR for ORF1ab, N gene, and S gene from Bulbula-positive samples was 52,780 (19,078–375,512), 38,301 (12,273–186,201), and 43,549 (10,476–240,648), respectively. Whereas the median viral concentration (gc/L) and IQR for ORF1ab, N gene, and S gene in Mikililand-positive samples was 64,762 (18,087–309,415), 45,580 (15,681–158,475), and 51,454 (11,318–184,333), respectively. Hence, there was no significant difference among the study sites in viral concentration: ORF1ab ($p = 0.7341$), N gene ($p = 0.2087$), and S gene ($p = 0.8721$). The detailed viral load of each positive sample is presented in [Supplementary Table S2_sheet 1](#).

After normalization of virus concentration levels (gc/L) using daily flow rate and number of people served by each WWTP, the

median viral concentration of daily load per person in sewershed (million genome copies [MGC/person-day]) and IQR was generated. Accordingly, the values for ORF1ab, N gene, and S gene in positive samples from KTP were 2055 (725–13,400), 861 (268–4,016), and 2,221 (436–7,607), respectively. Whereas, for Bulbula-positive samples, the results for ORF1ab, N gene, and S gene were 477 (136–2,387), 295 (871854), and 383 (69–1786), respectively. Therefore, there was a significant difference among the study sites in viral concentration of daily load per person in sewershed: ORF1ab ($p < 0.0001$), N gene ($p = 0.0008$), and S gene ($p < 0.0001$). The viral concentration of daily load per person in the sewershed of three viral targets in KTP and Bulbula WWTPs is presented using a Box-Whisker plot ([Figure 3B](#)).

Trend of viral concentration in wastewater and correlational analysis against COVID-19 daily cases

[Figure 4](#) demonstrates the dynamics of SARS-CoV-2 tests performed and the number of reported COVID-19 clinical cases for the year 2023. Daily reported COVID-19 cases of Addis Ababa were presented in [Supplementary Table S2_sheet 2](#). A significant decrease of daily cases during the months of April and May 2023 presented in line with the decrease in frequency of COVID-19 testing.

Positivity rates were in line with viral concentrations predicted by the three WWTPs ([Figures 5A–D](#)).

Virus concentration levels, as determined through WWTP testing, were normalized for the flow rate and number of people served. KTP is the oldest centralized system, mostly serving residents in the central, southern, and eastern parts of Addis Ababa. [Figure 5A](#) demonstrates the wastewater concentration of the target

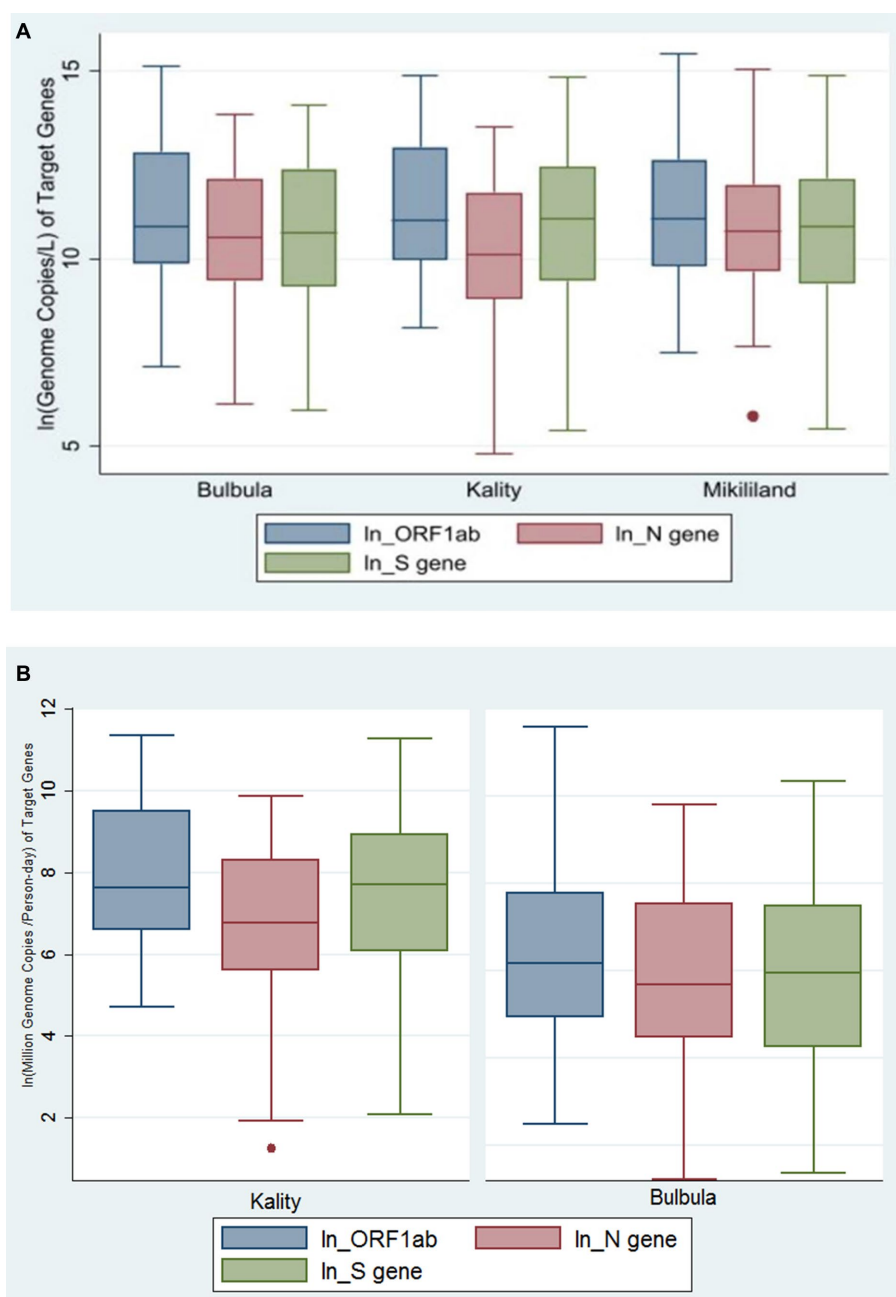


FIGURE 3

(A) Genome copies per L of SARS-CoV-2 gene targets in three wastewater treatment plants using Box-Whiskers plot. The data represents the average number of SARS-CoV-2 gene copies for ORF-1ab gene, N gene, and S gene per L of wastewater sample obtained in the influent wastewater samples from the Three WWTPs. (B) The viral concentration of daily load per person in sewershed (MGC per Person-day) of SARS-CoV-2 gene targets in two wastewater treatment plants using Box-Whiskers plot.

genes in samples obtained from KTP increasing sharply starting 2 March (ORF1ab = 1,123 MGC/person/day, N-gene = 1,234 MGC/person/day, and S gene = 742 MGC/person/day) to 13 March (ORF1ab = 60,066 MGC/person/day, N-gene = 17,707 MGC/person/day, and S gene = 14,773 MGC/person/day; Figure 5A). Then the wastewater concentration of the target genes fluctuated up to 27 April within the range of ORF1ab = 12,453–65,424 MGC/person/day, N-gene = 2,622–17,707 MGC/person/day, and S gene = 8,815–38,031 MGC/person/day. Then the concentration decreased sharply from 1 May up to 10 May and sustained less than 1,000 MGC/

person/day of each target genes up to 5 July. Subsequently, the concentration in KTP increased by 3 October and decreased again by 24 October (Figure 5A). The trend of concentration of the target genes in wastewater samples of Bulbula and Mikililand WSP was almost similar to that of the concentration trend of KTP and with a bit of difference in time of increments or decrements (Figures 5B,C).

The case-based surveillance unit in EPHI does not have a daily active cases report for the exact residents that are served by each WWTP. However, considering the large population coverage of the KTP (i.e., serving more than one-third of the population and wide geographic

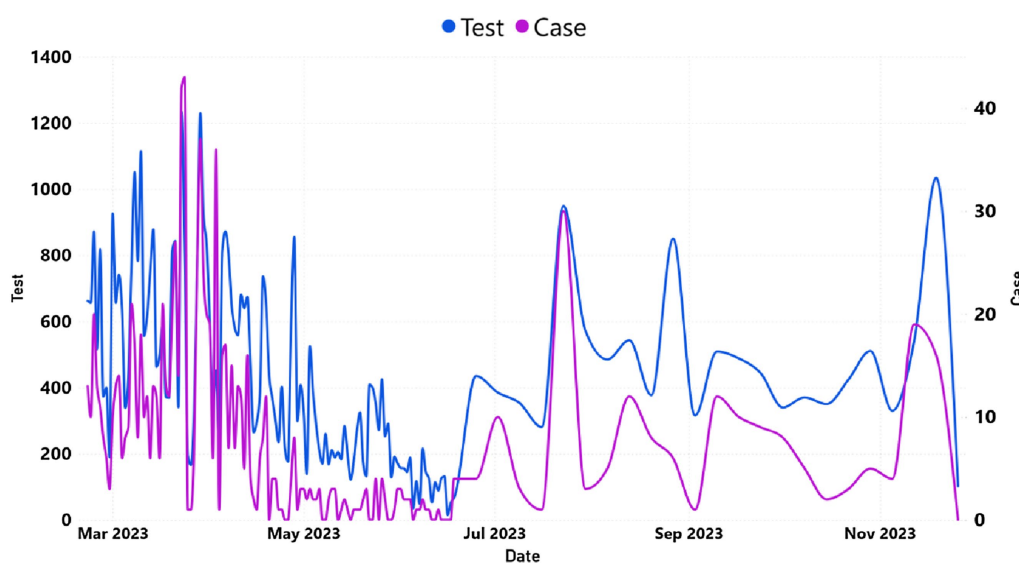


FIGURE 4

The trend of daily COVID-19 cases and tested individuals in Addis Ababa. The y-axis on the left represented the number of tested cases, whereas the y-axis on the right represented the number of reported cases.

coverage), it was found important to make a trend analysis of Addis Ababa daily cases against the trend of concentration for the target genes in wastewater samples collected from KTP and Bulbula WWTP.

As indicated in [Figure 5D](#), active clinical case counts doubled from 23 February to 23 March (in 9–21 active cases). This was reflected in a 15-fold increase in the average concentration of target genes in the wastewater (ORF1ab increased by a factor of 27; the N gene increased by a factor of 7, and the S gene increased by a factor of 13). Moreover, the increase in viral target positivity in the wastewater occurred approximately 10 days ahead of the increase in reported clinical cases. Again, at a later moment in the year, a more limited increase of active case counts from 17 July to 26 July (from zero to three active cases) was preceded by a wastewater increase starting from 5 July (ORF1ab = 356 MGC/person/day, N gene = 165 MGC/person/day, and S gene = 142 MGC/person/day), 12 days earlier. This increase lasted until 16 August (ORF1ab = 33,318 MGC/person/day, N gene = 5,285 MGC/person/day, and S gene = 10,790 MGC/person/day).

The finding indicates a positive correlation between the trend of weekly average SARS-CoV-2 MGC number in wastewater samples of WWTPs and the cumulative weekly reported COVID-19 cases in Addis Ababa. These were statistically significant for all three sites: KTP (0.5648, $p = 0.0002$), Bulbula (0.4052, $p = 0.0116$), and Mikililand (0.4247, $p = 0.0098$; [Supplementary Table S3](#)).

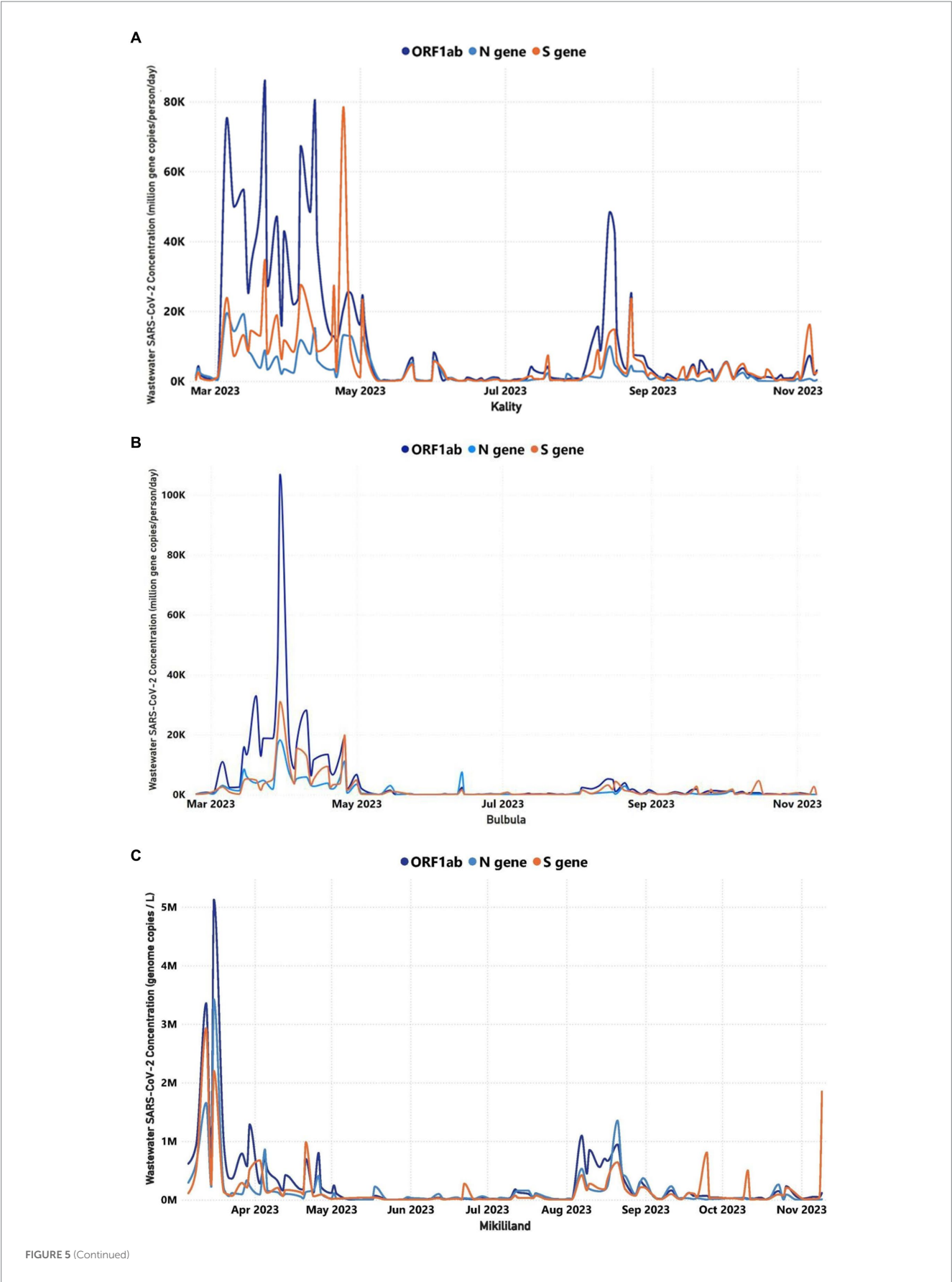
Estimated numbers of COVID-19-infected individuals and correlation with reported cases in Addis Ababa

Two methods were used to estimate the number of daily infected individuals among the population served by KTP and Bulbula WWTP based on the SARS-CoV-2 gene copy number obtained from the wastewater samples (27, 46). The numbers of daily predicted infected persons using method 1 and method 2 in KTP were similar and ranged from 10^2 to 10^4 , as represented in [Figure 6A](#). At Bulbula WWTP, these numbers were in the range of 10^0 – 10^4 . The daily predicted infected individuals from KTP were 330 times the median

value higher than the weekly cumulative reported COVID-19 cases ([Table 3](#)). The median predicted SARS-CoV-2 infected people of method 1 and method 2 from Kality was 3,303 and 3,523, respectively, whereas the median of weekly cumulative reported COVID-19 cases was 10. Correlational analyses of reported cases trend with the estimated number of infected individuals trend are shown in [Figure 6A](#) and [Supplementary Table S3](#). Similarly, the two methods resulted in higher mean values of daily predicted infected individuals from WWTPs compared to weekly cumulative reported COVID-19 cases ([Figure 6B](#)). The predicted number of infected individuals using the two methods followed a decreasing trend similar to the reported COVID-19 cases in Addis Ababa, and a statistically significant correlation was observed with data from KTP WWTP using Spearman correlation ($r = 0.5307$; $p = 0.0006$) and Bulbula WWTP ($r = 0.4816$; $p = 0.0022$). However, there is a significant difference between the number of predicted cases and reported cases for each surveillance week ($p < 0.0001$ for KTP and $p = 0.0029$) for Bulbula WWTP.

Discussion

Numerous studies conducted since the beginning of the COVID-19 pandemic have shown that WBE is a useful tool for tracking the evolution of the pandemic and providing early warning signs for the emergence or reemergence of public health threats (47, 48). The SARS-CoV-2 limit of detection in wastewater is principally determined by three laboratory procedures: virus concentration, RNA extraction, and qRT-PCR. The concentration method used here is known to preferentially bind intact virus particles but not cell-free nucleic acid. Thus, using other crude concentration methods or laboratory procedures without concentration may overestimate the intact viral burden. Using a technology that binds intact virus particles also provides greater evidence of active infection vs. cleared viral nucleic acid. Grab and autosampler sampling are the two most common wastewater sampling methods, but grab sampling has



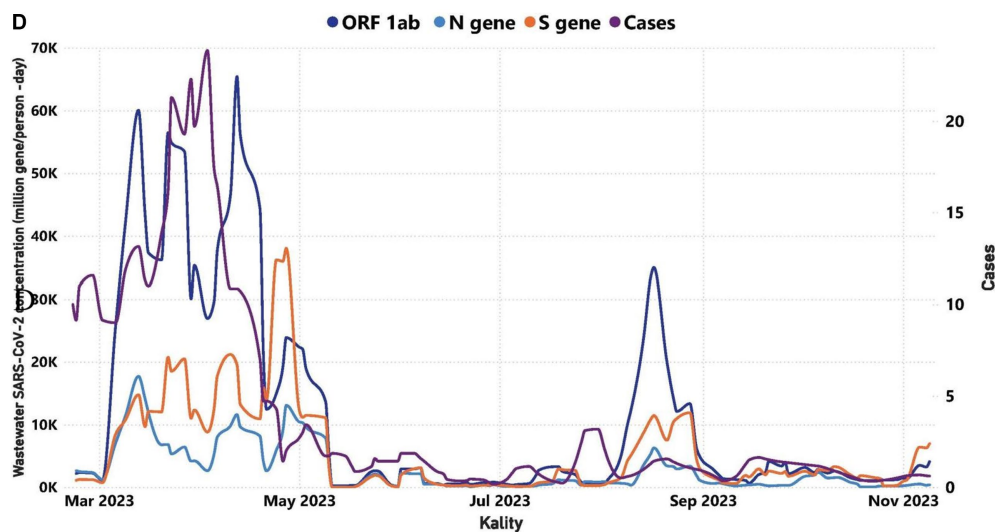


FIGURE 5

Trends in the viral target genes concentration of wastewater over time for three WWTPs in Addis Ababa (February 22, 2023–November 9, 2023): (A) Trends in the viral target genes concentration of wastewater over time for KTP, (B) Trends in the viral target genes concentration of wastewater over time for Bulbula WWTP, (C) Trends in the viral concentration of wastewater over time for Mikililand WSP, and (D) A comparison between the COVID-19 cases illustrated in purple color line that were reported in Addis Ababa and the SARS-CoV-2 target genes concentrations in KTP. The y-axis on the left represented the MGC/person-day of the target genes, whereas the y-axis on the right represented the number of reported cases. The correlation between the trend of daily reported cases and RNA concentration was significant ($p < 0.05$) in Kality. The number of cases and average SARS-CoV-2 concentration is based on the 7-day rolling average.

drawbacks in terms of missing viral shedding discharges to sewers, and autosampler has limited accessibility (49). Our result showed that the concentration of target genes was a bit higher in the autosampler compared to the Moore swab sampler (Figure 2). The primary cause of this discrepancy may be the Moore swab or gauze sampling methods' uptake rates, which could have been affected by inhibitors or virus losses after 8 hours of contact to the wastewater samples (50), which in the current study was installed for 24h. However, no significant difference in the Ct value and viral concentration was observed between autosampler and Moore swab samples for SARS-CoV-2 target genes (ORF1ab, N, and S genes), which is consistent with other studies (51). We conclude that Moore swab sampling is a more economical and resource-efficient sampling technique for the monitoring of SARS-CoV-2 in wastewater in our low-resource setting and may be extended to other pathogens of interest.

In our study, SARS-CoV-2 RNA was detected in a majority of influent wastewater samples (94%). This high rate revealed a much higher COVID-19 prevalence than actually clinically detected. A prior study using an antibody prevalence analysis showed that there was a significant underreporting of COVID-19 cases in Ethiopia (52). This can be explained by the fact that the far majority of actual COVID-19 cases in Addis Ababa are either mild or asymptomatic, with patients not seeking healthcare and testing services (19).

The positive SARS-CoV-2 detection rate in Addis Ababa was approximately identical to that of Kenya (81%) using the same technique of collection and testing (53), but higher than that of Malawi (8%) using samples taken from rivers and defunct WWTPs (33). In Malawi, samples from the defunct WWTP were found to have higher SARS-CoV-2 positive rates (21%) than river water samples (7%). Thus, the discrepancy in positive rate between our findings and Malawi might be attributed to the variance in viral shedding discharges into sewer lines of WWTP, rivers, and defunct WWTP (54). Furthermore, the variation in results may be attributed to

TABLE 3 The median Reported COVID-19 cases and predicated infectious cases.

Name	Median COVID-19 reported cases from Addis Ababa	Median predicated COVID-19 cases	
		Method 1 (27)	Method 2 (46)
Kality WWTP	10	3,303	3,523
Bulbula WWTP	10	17	18

differences in flow rate, methodology, data collection, and actual virus concentration differences since Malawi used grab sampling and polyethylene glycol (PEG) with no internal control (MS2) and potentially generated false negatives (55). The intensity of community transmission of SARS-CoV-2, the timing of the study, and the population served might also be important variables that make a difference observed for the positivity rates. The SARS-CoV-2 viral copy numbers (GC/L) of the amplification target genes were similar over the year 2023 (ORF1ab = 10^3 – 10^6 , N and S genes = 10^2 – 10^6 gc/L) in all three WWTP influent wastewater samples. This result shows the genome copies per 10 ml were not different at each treatment plant.

However, we observed a significant difference in terms of daily load per person for all target genes between KTP and Bulbula WWTPs; this is attributable to the difference in the prevalence of infected individuals that are served by each plant and the flow rate of the treatment plants.

For treated wastewater samples, the SARS-CoV-2 RNA was absent in all ($n = 11$ each) of the treated wastewater samples from two wastewater treatment plants (KTP and Bulbula). This result suggests that the UASB used in KTP and MBR technology used in Bulbula can successfully remove SARS-CoV-2 from wastewater to levels that are under the limit of detection of qPCR. However, at Mikililand WWTP (using a stabilization pond) SARS-CoV-2 RNA was still detected in 2 treated

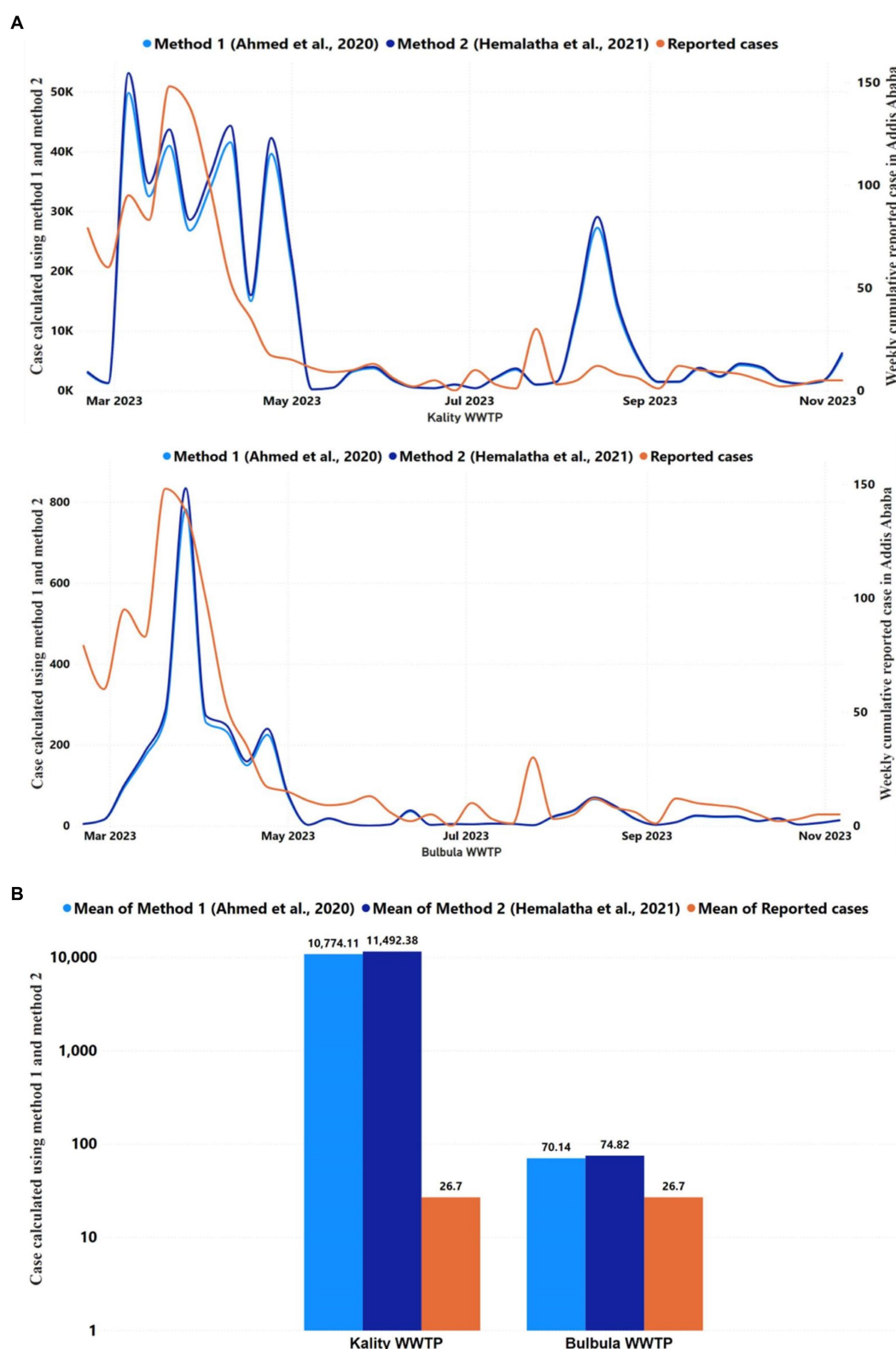


FIGURE 6

Reported and predicted COVID-19 infected cases across 2 WWTPs, (A) Trends in the COVID-19 reported cases and daily average predicted infected individuals using Ahmed et al. (27) and Hemalatha et al. (46) methods in Kaliti and Bulbula WWTP. The y-axis on the left represented the number of predicted cases using two methods, whereas the y-axis on the right represented the number of reported cases. The correlation between the trend of cumulative weekly reported and daily predicted COVID-19 cases was significant ($p < 0.05$) among the two WWTPs. (B) The figure represents the mean of COVID-19 reported cases of each WWTP during the study period and the predicted infected individuals using the Ahmed et al. (27) and Hemalatha et al. (46) methods for each WWTP for the study period.

wastewater samples ($n=11$), 5 being inconclusive. This demonstrates the limitations of the applied treatment method for viral eradication. Similar results were found in research conducted in Spain (56), where 2/18 of treated samples still tested positive for SARS-CoV-2.

In general, it is important to emphasize that the models used in this study may be crude compared to some of the more recently developed models generated. The Pepper Mild Mottle Virus (PMMoV), which is the most abundant RNA virus in human feces and occurs naturally in wastewater, has been used in recent studies (57, 58) to normalize qRT-PCR data. This approach may be more reliable in estimating the depth of infection in a community and could be used in an Ethiopian setting as well in the future.

Generally, we found fluctuating viral concentrations (MGC per person-day) over the study period. The overall change in the SARS-CoV-2 viral load in wastewater is positively correlated with reported COVID-19 clinical cases though the clinical testing frequency was low.

Thus, our result shows a significant positive correlation between trend viral loads in wastewater and reported COVID-19 clinical cases. This finding is consistent with previous studies in New York (59), India (60), and Hong Kong (61). In our setting, the increase in viral concentration started in the wastewater approximately 7–14 days ahead of the increase in COVID-19 clinical cases, as reported elsewhere (21). Furthermore, the amount of virus in wastewater did not drop off when the number of COVID-19 clinical cases significantly declined, which is consistent with a prior study that showed viral RNA might remain in fecal samples for up to 10 days (62).

The higher daily predicted infected persons from KTP, which was 330 times greater than the weekly cumulative recorded COVID-19 cases, revealed the high prevalence of asymptomatic individuals shedding SARS-CoV-2 to the sewage system in the catchment area. This is in line with previous studies in Ethiopia that have shown a significant inverse correlation between parasite infection prevalence and lack of COVID-19 symptoms due to shifts in activation status of the immune system (63, 64). Most people infected with SARS-CoV-2 in Ethiopia do not get sick (15), partly due to widespread parasitic infections (65, 66) and they may not seek medical care. Alternatively, the difference may be a result of the delay in active case reporting because qRT-PCR testing is biased as many tested individuals are not randomly undergoing diagnostic procedures, but their participation is motivated by the onset of symptoms either in themselves or in the person sharing their work or living environment, the prevalence of asymptomatic infection within the community as measured by rapid antigen tests might be underestimated due to sensitivity issues (67). On the other hand, the daily predicted infected individuals from Bulbula WWTP were merely 1.8 times the median value of the weekly cumulative reported cases in Addis Ababa (Table 3). The difference in predicted infected people in KTP and Bulbula is primarily attributable to the difference in flow rate at the treatment plants (Supplementary Table S3), which may further depend on the number of people served. Accordingly, the more the population served, the more viral shading is in the wastewater.

Conclusion

In conclusion, this study was undertaken to assess the presence of SARS-CoV-2 in the wastewater samples in three WWTPs in Addis

Ababa and evaluate its predictive value for clinical COVID-19 case reporting. Nanotrap® Microbiome A particles, Nanotrap® Enhancement Reagent 2 method, and Moore swab collection methods appeared to be effective in concentrating the virus from wastewater and can, therefore, be used in resource-limited settings. The significantly higher rate of SARS-CoV-2 detection from wastewater samples suggests a hidden high prevalence of COVID-19 disease in the population that remains overtly asymptomatic and/or underreported. Effluent wastewater treatment was only partly successful in making SARS-CoV-2 RNA undetectable at the KTP and Bulbula WWTP but not at Mikililand, indicating cautiousness is recommended. The peak in SARS-CoV-2 positivity rates in wastewater typically indicated a rise in clinical COVID-19 cases within 1–2 weeks later. The wastewater surveillance experience developed through this project can be applied to other national priority diseases in the future.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Ethics statement

The studies involving humans were approved by Ethiopian Public Health Institute Scientific and Ethical Review Office (Ref. EPHI 6.13/577). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

DA: Writing – review & editing, Writing – original draft, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Visualization. GG: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Visualization, Validation. YM: Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization, Formal analysis. NH: Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization. DiC: Writing – review & editing, Validation, Methodology, Investigation. GB: Writing – review & editing, Validation, Methodology, Investigation, Visualization. AG: Writing – review & editing, Validation, Methodology, Investigation, Data curation. AN: Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Investigation. SS: Writing – review & editing, Resources, Project administration, Funding acquisition, Methodology, Conceptualization. SW: Writing – review & editing, Project administration, Methodology, Funding acquisition, Investigation, Conceptualization. WG: Writing – review & editing, Methodology, Investigation, Validation. AA: Investigation, Conceptualization, Supervision, Writing – review &

editing, Methodology. YB: Writing – review & editing, Validation, Methodology, Investigation, Data curation. DaC: Validation, Writing – review & editing, Methodology, Investigation, Data curation. MG: Writing – review & editing, Methodology, Investigation, Validation. MG: Writing – review & editing, Methodology, Investigation, Validation. KT: Writing – review & editing, Methodology, Investigation, Validation. MeT: Validation, Supervision, Writing – review & editing, Methodology, Investigation. MB: Writing – review & editing, Validation, Methodology, Investigation. TM: Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation. SoA: Validation, Writing – review & editing, Supervision, Methodology, Investigation. AK: Validation, Writing – review & editing, Supervision, Methodology, Investigation. DM: Funding acquisition, Writing – review & editing, Supervision, Resources, Methodology, Investigation. SaA: Data curation, Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation. TR: Writing – review & editing, Validation, Supervision, Methodology, Investigation, Formal analysis. YK: Writing – review & editing, Supervision, Resources, Methodology, Investigation, Funding acquisition. MH: Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition. DW: Writing – review & editing, Validation, Supervision, Methodology, Investigation, Formal analysis, Data curation. MaT: Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation. GT: Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was supported by Project STELLAR, funded by The Global Fund to Fight AIDS, Tuberculosis, and Malaria under the COVID-19 Response Mechanism (C19RM). Its contents are solely

the responsibility of the authors and do not necessarily represent the official view of the funder.

Acknowledgments

We gratefully acknowledge the technical support provided by The Association of Public Health Laboratories (APHL), The Africa Centres for Disease Control and Prevention (CDC), Wisconsin University, and the Technical Working Group (TWG). We also gratefully acknowledge Addis Ababa Water and Sewerage Authority (AAWSA) and three WWTP site teams for supporting sample collection. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2024.1394798/full#supplementary-material>

References

- Bloom JS, Sathe L, Munugala C, Jones EM, Gasperini M, Lubock NB, et al. Swab-Seq: a high-throughput platform for massively scaled up SARS-CoV-2 testing. *MedRxiv*. (2020)
- Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. (2020) 579:270–3. doi: 10.1038/s41586-020-2012-7
- World Health Organization (2020). WHO director-General's statement on IHR emergency committee on novel coronavirus (2019-nCoV). Geneva: World Health Organization. Available at: <https://www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ih-er-emergency-committee-on-novel-coronavirus-2019-ncov>. (Accessed January 12, 2024).
- WHO COVID-19 dashboard (2024). Available at: <https://data.who.int/dashboards/covid19/cases?n=c> [Accessed February 01, 2024].
- CDC (2020). Africa identifies first case of coronavirus disease: statement by the director of Africa CDC. Available at: <https://africacdc.org/news-item/africa-identifies-first-case-of-coronavirus-disease-statement-by-the-director-of-africa-cdc/>. (Accessed August 19, 2024).
- WHO (2020). A second COVID-19 case is confirmed in Africa. Available at: <https://www.afro.who.int/news/second-covid-19-case-confirmed-africa> (Accessed August 19, 2024).
- Lone SA, Ahmad A. COVID-19 pandemic – an African perspective. *Emerg Microbes Infect.* (2020) 9:1300–8. doi: 10.1080/22221751.2020.1775132
- Nextstrain (2020). Available at: <https://nextstrain.org/ncov/africa?l=clock&p=grid> (Accessed August 20, 2024).
- Massinga Loembé M, Tshangela A, Salyer SJ, Varma JK, Ouma AEO, Nkengasong JN. COVID-19 in Africa: the spread and response. *Nat Med*. (2020) 26:999–1003. doi: 10.1038/s41591-020-0961-x
- Manirambona E, Lucero-Priso DE, Shomuyiwa DO, Denkyira SA, Okesanya OJ, Haruna UA, et al. Harnessing wastewater-based surveillance (WBS) in Africa: a historic turning point towards strengthening the pandemic control. *Discov Water*. (2024) 4:9. doi: 10.1007/s43832-024-00066-0
- Sah P, Fitzpatrick MC, Zimmer CF, Abdollahi E, Juden-Kelly L, Moghadas SM, et al. Asymptomatic SARS-CoV-2 infection: a systematic review and meta-analysis. *Proc Natl Acad Sci*. (2021) 118:e2109229118. doi: 10.1073/pnas.2109229118
- Wang Z, Wu P, Wang J, Lourenço J, Li B, Rader B, et al. Assessing the asymptomatic proportion of SARS-CoV-2 infection with age in China before mass vaccination. *J R Soc Interface*. (2022) 19:20220498. doi: 10.1098/rsif.2022.0498
- Oran DP, Topol EJ. The proportion of SARS-CoV-2 infections that are asymptomatic. *Ann Intern Med*. (2021) 174:655–62. doi: 10.7326/M20-6976
- WHO Africa (2022). Over two-thirds of Africans exposed to virus which causes COVID-19: WHO study. Available from: <https://www.afro.who.int/news/over-two-thirds-africans-exposed-virus-which-causes-covid-19-who-study#:~:text=The%20new%20analysis%20suggests%20that,the%20virus%20by%20September%202021> (Accessed February 17, 2024).

15. Abraha HE, Gessesse Z, Gebrecherkos T, Kebede Y, Weldegiargis AW, Tequare MH, et al. Clinical features and risk factors associated with morbidity and mortality among patients with COVID-19 in northern Ethiopia. *Int J Infect Dis.* (2021) 105:776–83. doi: 10.1016/j.ijid.2021.03.037
16. Muleta D, Simieneh A, Duguma T, Tekalign E, Worku T, Ayele G, et al. SARS-CoV-2 infections, clinical characteristics, and related risk factors: the first 8 months surveillance study conducted in Southwest Ethiopia. *INQUIRY J Health Care Organ Prov Financ.* (2023) 60:00469580231166794.
17. Wolday D, Tasew G, Amogne W, Urban B, Schallig HDFH, Harris V, et al. Interrogating the impact of intestinal parasite-microbiome on pathogenesis of COVID-19 in sub-Saharan Africa. *Front Microbiol.* (2021) 12:614522. doi: 10.3389/fmicb.2021.614522
18. Wolday D, Ndungu FM, Gómez-Pérez GP, de Wit TFR. Chronic immune activation and CD4(+) T cell lymphopenia in healthy African individuals: perspectives for SARS-CoV-2 vaccine efficacy. *Front Immunol.* (2021) 12:693269. doi: 10.3389/fimmu.2021.693269
19. Mohamed AM, Matar E, Isa HM, Moosa AK, Hasan WF, Mohamed AG, et al. Presence of SARS-CoV-2 virus in wastewater in the Kingdom of Bahrain during the COVID-19 pandemic. *Influenza Other Respir Viruses.* (2023) 17:e13194. doi: 10.1111/irv.13194
20. Dadi LS, Kefeni BD, Kadi H, Kifle E, Yesuf EA, Hailemariam G, et al. Household readiness to Care for Mild and Asymptomatic COVID-19 cases at home, Southwest Ethiopia: a community-based cross-sectional study. *Ethiop J Health Sci.* (2022) 32:1071–82. doi: 10.4314/ejhs.v32i6.3
21. Jones DL, Baluja MQ, Graham DW, Corbushley A, McDonald JE, Malham SK, et al. Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Sci Total Environ.* (2020) 749:141364. doi: 10.1016/j.scitotenv.2020.141364
22. Jeong HW, Kim SM, Kim HS, Kim YI, Kim JH, Cho JY, et al. Viable SARS-CoV-2 in various specimens from COVID-19 patients. *Clin Microbiol Infect.* (2020) 26:1520–4. doi: 10.1016/j.cmi.2020.07.020
23. Natarajan A, Zlitni S, Brooks EF, Vance SE, Dahlen A, Hedlin H, et al. Gastrointestinal symptoms and fecal shedding of SARS-CoV-2 RNA suggest prolonged gastrointestinal infection. *Medicine.* (2022) 3:371–87.e9. doi: 10.1016/j.medj.2022.04.001
24. Cheung KS, Hung IFN, Chan PPY, Lung KC, Tso E, Liu R, et al. Gastrointestinal manifestations of SARS-CoV-2 infection and virus load in fecal samples from a Hong Kong cohort: systematic review and Meta-analysis. *Gastroenterology.* (2020) 159:81–95. doi: 10.1053/j.gastro.2020.03.065
25. Xiao F, Sun J, Xu Y, Li F, Huang X, Li H, et al. Infectious SARS-CoV-2 in feces of patient with severe COVID-19. *Emerg Infect Dis.* (2020) 26:1920–2. doi: 10.3201/eid2608.200681
26. Sherchan SP, Shahin S, Ward LM, Tandukar S, Aw TG, Schmitz B, et al. First detection of SARS-CoV-2 RNA in wastewater in North America: a study in Louisiana, USA. *Sci Total Environ.* (2020) 743:140621. doi: 10.1016/j.scitotenv.2020.140621
27. Ahmed W, Angel N, Edson J, Bibby K, Bivins A, O'Brien JW, et al. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: a proof of concept for the wastewater surveillance of COVID-19 in the community. *Sci Total Environ.* (2020) 728:138764. doi: 10.1016/j.scitotenv.2020.138764
28. Maida CM, Amodio E, Mazzucco W, la Rosa G, Lucentini L, Suffredini E, et al. Wastewater-based epidemiology for early warning of SARS-COV-2 circulation: a pilot study conducted in Sicily, Italy. *Int J Hyg Environ Health.* (2022) 242:113948. doi: 10.1016/j.ijheh.2022.113948
29. Henriques TB, Cassini ST, de Pinho Keller R. Contribution of wastewater-based epidemiology to SARS-CoV-2 screening in Brazil and the United States. *J Water Health.* (2023) 21:343–53. doi: 10.2166/wh.2023.260
30. GPEI – Global Polio Eradication Initiative. Polio eradication strategy 2022–2026: Delivering on a promise. Geneva: World Health Organization (2021).
31. Hooda Y, Islam S, Kabiraj R, Rahman H, Sarkar H, da Silva KE, et al. Old tools, new applications: use of environmental bacteriophages for typhoid surveillance and evaluating vaccine impact. *PLoS Negl Trop Dis.* (2024) 18:e0011822. doi: 10.1371/journal.pntd.0011822
32. Committee on Community Wastewater-Based Infectious Disease Surveillance (2023). Wastewater-based disease surveillance for public health action. Washington (DC): National Academies Press (US); 2, Wastewater Surveillance for COVID-19. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK591716/>. (Accessed July 22, 2024).
33. Barnes KG, Levy JL, Gauld J, Rigby J, Kanjerwa O, Uzzell CB, et al. Utilizing river and wastewater as a SARS-CoV-2 surveillance tool in settings with limited formal sewage systems. *Nat Commun.* (2023) 14:7883. doi: 10.1038/s41467-023-43047-y
34. Iwu-Jaja C, Ndlovu NL, Rachida S, Yousif M, Taukobong S, Macheke M, et al. The role of wastewater-based epidemiology for SARS-CoV-2 in developing countries: cumulative evidence from South Africa supports sentinel site surveillance to guide public health decision-making. *Sci Total Environ.* (2023) 903:165817. doi: 10.1016/j.scitotenv.2023.165817
35. Nkambule S, Johnson R, Mathee A, Mahlangeni N, Webster C, Horn S, et al. Wastewater-based SARS-CoV-2 airport surveillance: key trends at the Cape Town international airport. *J Water Health.* (2023) 21:402–8. doi: 10.2166/wh.2023.281
36. Pandey D, Verma S, Verma P, Mahanty B, Dutta K, Daverey A, et al. SARS-CoV-2 in wastewater: challenges for developing countries. *Int J Hyg Environ Health.* (2021) 231:113634. doi: 10.1016/j.ijheh.2020.113634
37. Ali S, Gudina EK, Gize A, Aliy A, Adankie BT, Tsegaye W, et al. Community wastewater-based surveillance can be a cost-effective approach to track COVID-19 outbreak in low-resource settings: feasibility Assessment for Ethiopia context. *Context.* (2022) 19, 1–14. doi: 10.3390/ijerph19148515
38. UN World Urbanization Prospects (2024). Addis Ababa Population. Available at: <https://worldpopulationreview.com/world-cities/addis-ababa-population> (Accessed February 16, 2024).
39. Liu P, Ibaraki M, VanTassel J, Geith K, Cavallo M, Kann R, et al. A sensitive, simple, and low-cost method for COVID-19 wastewater surveillance at an institutional level. *Sci Total Environ.* (2022) 807:151047. doi: 10.1016/j.scitotenv.2021.151047
40. Ali M, Robele S. Major contributing factors to the lower level of connection to the existing sewer network in Addis Ababa: the case of the Kaliti catchment. *Front Environ Eng.* (2023) 2:1158656. doi: 10.3389/fenv.2023.1158656
41. Radjenović J, Matošić M, Mijatović I, Petrović M, Barceló D. (2008). Emerging contaminants from industrial and municipal waste. Berlin, Heidelberg: Springer Berlin Heidelberg. 37–101.
42. Zenebe BT. (2019). Assessment of quality management practices of condominium house projects in addis ababa: the case of wastewater management in mikililand site. Available at: <http://www.repository.smuc.edu.et/bitstream/123456789/4722/1/Final%20May%2025.pdf>. (Accessed January 15, 2024).
43. Rafiee M, Isazadeh S, Mohseni-Bandpei A, Mohebbi SR, Jahangiri-rad M, Eslami A, et al. Moore swab performs equal to composite and outperforms grab sampling for SARS-CoV-2 monitoring in wastewater. *Sci Total Environ.* (2021) 790:148205. doi: 10.1016/j.scitotenv.2021.148205
44. QIAGEN (2023). QIAamp® Viral RNA Mini Handbook. Available at: <https://www.qiagen.com/us/resources/download.aspx?id=c80685c0-4103-49ea-aa72-8989420c3018&lang=en> (Accessed November 30, 2022).
45. Scientific T. (2020). TaqPath™ COVID-19 CE-IVD RT-PCR kit instructions for use (publication number MAN0019215). Available at: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019215_TaqPathCOVID-19_CE-IVD_RT-PCR%20Kit_IFU.pdf. (Accessed January 15, 2024).
46. Hemalatha M, Kiran U, Kuncha SK, Kopperi H, Gokulan CG, Mohan SV, et al. Surveillance of SARS-CoV-2 spread using wastewater-based epidemiology: comprehensive study. *Sci Total Environ.* (2021) 768:144704. doi: 10.1016/j.scitotenv.2020.144704
47. Singer AC, Thompson JR, Filho CRM, Street R, Li X, Castiglioni S, et al. A world of wastewater-based epidemiology. *Nat Water.* (2023) 1:408–15. doi: 10.1038/s44221-023-00083-8
48. Sims N, Kasprzyk-Hordern B. Future perspectives of wastewater-based epidemiology: monitoring infectious disease spread and resistance to the community level. *Environ Int.* (2020) 139:105689. doi: 10.1016/j.envint.2020.105689
49. Valenzuela EF, Menezes HC, Cardeal ZL. Passive and grab sampling methods to assess pesticide residues in water. A review. *Environ Chem Lett.* (2020) 18:1019–48. doi: 10.1007/s10311-020-00998-8
50. Li J, Verhagen R, Ahmed W, Metcalfe S, Thai PK, Kaserzon SL, et al. In Situ Calibration of passive samplers for viruses in wastewater. *ACS ES&T Water.* (2022) 2:1881–90. doi: 10.1021/acsestwater.1c00406
51. Wilson M, Qiu Y, Yu J, Lee BE, McCarthy DT, Pang X. Comparison of auto sampling and passive sampling methods for SARS-CoV-2 detection in wastewater. *Pathogens (Basel, Switzerland).* (2022) 11, 1–7. doi: 10.3390/pathogens11030359
52. Merkt S, Ali S, Gudina EK, Adissu W, Gize A, Muenchhoff M, et al. Long-term monitoring of SARS-CoV-2 seroprevalence and variants in Ethiopia provides prediction for immunity and cross-immunity. *Nat Commun.* (2024) 15:3463. doi: 10.1038/s41467-024-47556-2
53. GFATM (2023). Wastewater-based surveillance as an early warning tool to monitor infectious diseases at community level. Available at: https://www.theglobalfund.org/media/13232/c19rm_wastewater-surveillance_presentation_en.pdf (Accessed February 14, 2024).
54. Tandukar S, Sthapit N, Thakali O, Malla B, Sherchan SP, Shakya BM, et al. Detection of SARS-CoV-2 RNA in wastewater, river water, and hospital wastewater of Nepal. *Sci Total Environ.* (2022) 824:153816. doi: 10.1016/j.scitotenv.2022.153816
55. West NW, Hartrick J, Alamin M, Vasquez AA, Bahmani A, Turner CL, et al. Passive swab versus grab sampling for detection of SARS-CoV-2 markers in wastewater. *Sci Total Environ.* (2023) 889:164180. doi: 10.1016/j.scitotenv.2023.164180
56. Randazzo W, Truchado P, Cuevas-Ferrando E, Simón P, Allende A, Sánchez G. SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. *Water Res.* (2020) 181:115942. doi: 10.1016/j.watres.2020.115942
57. Kim S, Kennedy LC, Wolfe MK, Criddle CS, Duong DH, Topol A, et al. SARS-CoV-2 RNA is enriched by orders of magnitude in primary settled solids relative to liquid wastewater at publicly owned treatment works. *Environ Sci Water Res Technol.* (2022) 8:757–70. doi: 10.1039/d1ew00826a
58. Sanchez Jimenez B, Sterling T, Brown A, Modica B, Gibson K, Collins H, et al. Wastewater surveillance in the COVID-19 post-emergency pandemic period: a

promising approach to monitor and predict SARS-CoV-2 surges and evolution. *Heliyon*. (2023) 9:e22356. doi: 10.1016/j.heliyon.2023.e22356

59. Hoar C, Chauvin F, Clare A, McGibbon H, Castro E, Patinella S, et al. Monitoring SARS-CoV-2 in wastewater during new York City's second wave of COVID-19: sewershed-level trends and relationships to publicly available clinical testing data. *Environ Sci Water Res Technol*. (2022) 8:1021–35. doi: 10.1039/D1EW00747E

60. Wani H, Menon S, Desai D, D'Souza N, Bhatena Z, Desai N, et al. Wastewater-based epidemiology of SARS-CoV-2: assessing prevalence and correlation with clinical cases. *J Food Environ Virol*. (2023) 15:131–43. doi: 10.1007/s12560-023-09555-2

61. Zheng X, Li S, Deng Y, Xu X, Ding J, Lau FTK, et al. Quantification of SARS-CoV-2 RNA in wastewater treatment plants mirrors the pandemic trend in Hong Kong. *Sci Total Environ*. (2022) 844:157121. doi: 10.1016/j.scitotenv.2022.157121

62. Gogoi G, Singh SD, Kalyan E, Koch D, Gogoi P, Kshattray S, et al. An interpretative review of the wastewater-based surveillance of the SARS-CoV-2: where do we stand on its presence and concern? *Front Microbiol*. (2024) 15:1338100. doi: 10.3389/fmicb.2024.1338100

63. Toru M, Atnaf A, Mengist HM, Reta A. The COVID-19 severity and its association with intestinal parasite coinfection and urine biochemical parameters among

COVID-19-confirmed patients admitted to Debre Markos university COVID-19 center, Northwest Ethiopia. *Biomed Res Int*. (2024) 2024:1–9. doi: 10.1155/2024/3064374

64. Wolday D, Gebrecherkos T, Arefaine ZG, Kiros YK, Gebreegzabher A, Tasew G, et al. Effect of co-infection with intestinal parasites on COVID-19 severity: a prospective observational cohort study. *EClinicalMedicine*. (2021) 39:101054. doi: 10.1016/j.eclinm.2021.101054

65. Wondimu H, Mihret M. Prevalence and associated factors of intestinal parasites among food handlers working in food service establishments in Northwest Ethiopia, 2022. *J Parasitol Res*. (2023) 2023:1–6. doi: 10.1155/2023/3230139

66. Girma A, Aemiro A. Prevalence and associated risk factors of intestinal parasites and enteric bacterial infections among selected region food handlers of Ethiopia during 2014–2022: a systematic review and Meta-analysis. *Sci World J*. (2022) 2022:1–14. doi: 10.1155/2022/7786036

67. Dankova Z, Novakova E, Skerenova M, Holubekova V, Lucansky V, Dvorska D, et al. Comparison of SARS-CoV-2 detection by rapid antigen and by three commercial RT-qPCR tests: a study from Martin University Hospital in Slovakia. *Int J Environ Res Public Health*. (2021) 18:7037. doi: 10.3390/ijerph18137037



OPEN ACCESS

EDITED BY

Jiaying Li,
The University of Queensland, Australia

REVIEWED BY

Soufiane Bentout,
Centre Universitaire Ain Temouchent, Algeria
Suani Tavares Rubim De Pinho,
Federal University of Bahia (UFBA), Brazil

*CORRESPONDENCE

Lenka Cepova
✉ lenka.cepova@vsb.cz
Krishnasamy Karthik
✉ karthikmeed@gmail.com

RECEIVED 09 March 2024

ACCEPTED 17 September 2024

PUBLISHED 21 October 2024

CITATION

Vijayalakshmi GM, Ariyanatchi M, Cepova L
and Karthik K (2024) Advanced optimal
control approaches for immune boosting and
clinical treatment to enhance dengue viremia
models using ABC fractional-order analysis.
Front. Public Health 12:1398325.
doi: 10.3389/fpubh.2024.1398325

COPYRIGHT

© 2024 Vijayalakshmi, Ariyanatchi, Cepova
and Karthik. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Advanced optimal control approaches for immune boosting and clinical treatment to enhance dengue viremia models using ABC fractional-order analysis

G. M. Vijayalakshmi¹, M. Ariyanatchi¹, Lenka Cepova^{2*} and Krishnasamy Karthik^{3*}

¹Department of Mathematics, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India, ²Department of Machining, Assembly and Engineering Metrology, Faculty of Mechanical Engineering, VSB-Technical University of Ostrava, Ostrava, Czechia, ³Department of Mechanical Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India

Introduction: This work focuses on the Dengue-viremia ABC (Atangana-Baleanu Caputo) fractional-order differential equations, accounting for both symptomatic and asymptomatic infected cases. Symptomatic cases are characterized by higher viremia levels, whereas asymptomatic cases exhibit lower viremia levels. The fractional-order model highlights memory effects and other advantages over traditional models, offering a more comprehensive representation of dengue dynamics.

Methods: The total population is divided into four compartments: susceptible, asymptomatic infected, symptomatic infected, and recovered. The model incorporates an immune-boosting factor for asymptomatic infected individuals and clinical treatment for symptomatic cases. Positivity and boundedness of the model are validated, and both local and global stability analyses are performed. The novel Adams-Bash numerical scheme is utilized for simulations to rigorously assess the impact of optimal control interventions.

Results: The results demonstrate the effectiveness of the proposed control strategies. The reproduction numbers must be reduced based on specific optimal control conditions to effectively mitigate disease outbreaks. Numerical simulations confirm that the optimal control measures can significantly reduce the spread of the disease.

Discussion: This research advances the understanding of Dengue-viremia dynamics and provides valuable insights into the application of ABC fractional-order analysis. By incorporating immune-boosting and clinical treatment into the model, the study offers practical guidelines for implementing successful disease control strategies. The findings highlight the potential of using optimal control techniques in public health interventions to manage disease outbreaks more effectively.

KEYWORDS

dengue fractional-order mathematical modeling, Atangana-Baleanu operator, Lyapunov stability, basic reproduction value, optimal control, Adams-Bashforth method

1 Introduction

Worldwide, thousands of dengue cases are reported every year. The world's tropical and subtropical regions are affected by dengue infection, which is a mosquito-carrying disease. A high temperature and flu-like symptoms are signs of mild illness or asymptomatic to stern disease. DHF (Dengue Hemorrhagic Fever) or DSS (Dengue Show Syndromes Syndromes) is a highly infectious form of dengue fever that causes serious bleeding, shock, and death. Generally, it was noticed that only one out of four dengue contagions is symptomatic. Dengue virus occurs in four major types (DENV types 1, 2, 3, and 4), all of which can cause serious illness. The usual signs of DENV type 1 are like a common cold and mild fever, which will not lead directly to DHF; conversely, later DENV types can lead to DHF (1–3).

To understand the dynamical behavior of dengue transmission, we formulated a mathematical model, particularly focusing on vector-borne disease transmission from mosquitoes to humans. Esteva and Vargas (4, 5) pioneered the creation of a fundamental dengue model and explored numerous fundamental mathematical concepts and their accompanying numerical simulations. Feng et al. (6) presented a two-strain dengue infection model and examined competitive exclusion. Researchers have conducted numerous studies to better understand the transmission of dengue fever (7–9).

The importance of fractional-order models lies in their ability to capture the complex dynamics and long-term dependencies within the transmission process. By incorporating fractional derivatives, these models provide a more comprehensive understanding of disease spread, which is crucial for designing effective intervention strategies. The fractional-order models can accommodate the nuanced behavior of dengue transmission, offering insights that integer-order models may overlook, thereby enhancing the accuracy and effectiveness of disease control measures.

The fractional order model has been conclusively demonstrated by a recent study to be capable of controlling the trend of complex diffusion disorder (10–14). Many have emphasized various mathematical models for Dengue transmission and prevention (15–19). All cited references explain the transmission process of Dengue infection from different perspectives, including dynamic analysis, evaluation of vaccination, and optimal control measures (20–24). The most updated studies on Dengue with real-life data are presented in (25, 26). The mathematical description of Dengue is briefly described in Deterministic and Stochastic terms. The evolution of dengue with asymptomatic carriers using optimal control measures was investigated in (27).

Therefore, motivated by the aforementioned literature, we propose a computational framework for the dissemination of dengue at a given viremia level. We investigated whether symptom-free people were markedly more susceptible to mosquitoes than clinically symptom-positive patients. The new idea of a mathematical model to analyse the immune-boosting factor for asymptomatic infected cases and the waning immunity that cases re-infect is reported. To make practical applications and simulations easier, we utilize the Adams-Bash forth numerical scheme, which is renowned for its accuracy and stability. This

choice ensures that our model reflects real-world scenarios while maintaining computational efficacy. A key highlight of this study is the incorporation of optimal control strategies into the ABC fractional order Dengue viremia model. These strategies are designed to explore how interventions, such as self-prevention and vector control, can be optimized to curtail disease spread. The analysis extends to investigating disease-free and endemic stability, providing crucial insights into the long-term behavior of the system under various control scenarios.

This article is prepared as follows: In portion 2, we review the fundamental definitions for the fractional-order operator and provide a list of mathematical properties that were used throughout the work. The dengue viral mathematical model with fractional order was presented in portion 3. Portion 4 examines the local as well as global consistency of the suggested model through the Routh-Hurwitz criteria and the Lyapunov function. An optimal control solution and discussion are present in portion 5. The final section focuses on numerical simulations and a comprehensive conclusion.

2 Fundamental results

This section introduces fractional derivation and some of its properties, which will be used in the following components.

Definition 2.1.

Consider $\psi \in \mathbb{H}'(0, \mathbb{T})$ and $\eta \in [0, t]$, then Atangana-Baleanu fraction component in Caputo case is

$${}^{ABC}_0 \mathfrak{D}_t^\eta \psi(t) = \frac{\overline{\mathcal{N}(\eta)}}{1-\eta} \int_0^t \frac{d}{dx} \psi(s) \mathcal{N}_\eta \left[-\frac{\eta}{1-\eta}(t-s) \right] ds \quad (1)$$

The method yields a variation operator Caputo-Fabrizio that replaces

$$\mathcal{N}_\eta \left[-\frac{\eta}{1-\eta}(t-s) \right] ds \text{ by } \mathcal{N}_1 = \exp \left[-\frac{\eta}{1-\eta}(t-s) \right]$$

It's noteworthy that ${}^{ABC}_0 \mathfrak{D}_t^\eta [\text{constant}] = 0$. Here $\overline{\mathcal{N}(\eta)}$ is the typical function and it is defined as $\overline{\mathcal{N}(0)} = 1$ and $\overline{\mathcal{N}(1)} = 1$. $\overline{\mathcal{N}(\eta)}$ depict the familiar Mittag-Leffler operator, it also reflects the exponential function generality.

Definition 2.2.

The fractional integral of ABC with order η given by

$${}^{ABC}_0 \mathfrak{I}_t^\eta \psi(t) = \frac{1-\eta}{\overline{\mathcal{N}(\eta)}} \psi(t) + \frac{\eta}{\overline{\mathcal{N}(\eta)} \Gamma(\eta)} \int_0^t \psi(s) (t-s)^{\eta-1} ds \quad (2)$$

Lemma 2.1.

Consider a fractional-order system

$${}^{ABC}_0 \mathfrak{D}_t^\eta x(t) = f(t, x), \quad t > t_0 \quad (3)$$

Where $\eta \in (0, 1)$ in the initial case $x(t_0)$.

If $f(t, x)$ fulfills the Lipschitz condition in relation to x , then system (Equation 3) exhibits a unique solution in the region $[t_0, +\infty) \times \varphi$ and $\varphi \subseteq \mathbb{R}^n$.

Lemma 2.2.

If $x(t) \in \mathbb{R}^+$ become an ongoing and attainable consequence. Then

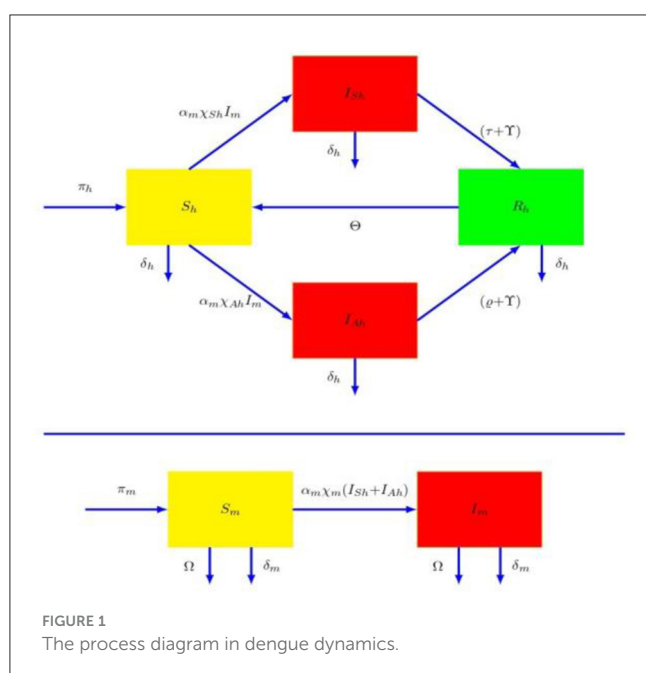
$${}^{ABC}_0\mathcal{D}_t^\eta \left(x(t) - x^* - x^* \ln \frac{x(t)}{x^*} \right) \leq \left(1 - \frac{x^*}{x(t)} \right) {}^{ABC}_0\mathcal{D}_t^\eta x(t) \quad (4)$$

Here $t > t_0$, $\eta \in (0, 1)$ and $x^* \in \mathbb{R}^+$.

3 Evaluation of dengue dynamics

In this section, we expand upon the previously described Dengue SIR-SI model (18) by incorporating additional factors and refining the classification of both human and mosquito populations. Our model includes viremia levels, an immune-boosting factor for asymptomatic infected cases, and clinical treatment for symptomatic infected cases.

To study the mode of spread of dengue sickness, the human species ($\mathcal{N}_h$) is subdivided into four classes: susceptible ($\mathcal{S}_h$), symptomatic infectious ($\mathcal{I}_{Sh}$), asymptomatic infectious ($\mathcal{I}_{Ah}$) and recovered human populations ($\mathcal{R}_h$). We classified female mosquito species ($\mathcal{N}_m$) into Susceptible ($\mathcal{S}_m$) and infective mosquitoes ($\mathcal{I}_m$). A Susceptible individual among as one who is not infected and immune, infected humans are both asymptomatic and symptomatic are those who have acquired Dengue viremia from an infected mosquito populations and are all capable of spreading dengue virus to susceptible mosquitoes. Let we examines π_h and π_m the acquisition rates of humans and mosquitoes. The proposed model, illustrated in the flowchart, demonstrates the dengue transmission dynamics. Based on Figure 1, we developed



the following differential equation.

$$\begin{aligned} \frac{d\mathcal{S}_h}{dt} &= \pi_h - \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h \mathcal{I}_m - \delta_h \mathcal{S}_h + \theta \mathcal{R}_h \\ \frac{d\mathcal{I}_{Sh}}{dt} &= \alpha_m \chi_{\mathcal{S}_h} \mathcal{S}_h \mathcal{I}_m - (\tau + \gamma + \delta_h) \mathcal{I}_{Sh} \\ \frac{d\mathcal{I}_{Ah}}{dt} &= \alpha_m \chi_{\mathcal{A}_h} \mathcal{S}_h \mathcal{I}_m - (\rho + \gamma + \delta_h) \mathcal{I}_{Ah} \\ \frac{d\mathcal{R}_h}{dt} &= \gamma (\mathcal{I}_{Sh} + \mathcal{I}_{Ah}) - \theta \mathcal{R}_h - \delta_h \mathcal{R}_h + \tau \mathcal{I}_{Sh} + \rho \mathcal{I}_{Ah} \\ \frac{d\mathcal{S}_m}{dt} &= \pi_m - \alpha_m \chi_m (\mathcal{I}_{Sh} + \mathcal{I}_{Ah}) \mathcal{S}_m - \delta_m \mathcal{S}_m \\ \frac{d\mathcal{I}_m}{dt} &= \alpha_m \chi_m (\mathcal{I}_{Sh} + \mathcal{I}_{Ah}) \mathcal{S}_m - \delta_m \mathcal{I}_m \end{aligned} \quad (5)$$

Where

α_m - individual mosquito's biting rate

$\chi_{\mathcal{S}_h}$ - Dissemination to human by mosquitoes, which leads to a symptomatic infectious in humans

$\chi_{\mathcal{A}_h}$ - Dissemination to human by mosquitoes, which leads to a asymptomatic infectious in humans

χ_m - Viremia dissemination to mosquito by human species

δ_h - Human Fatality rate

τ - Symptomatic infected human treatment rate

γ - Recovering rate.

θ - Transition rate at which a recovered person becomes defenseless due to loss of immunity

ρ - Rates of immunosuppression for asymptomatic victims

δ_m - Rate of mosquito natural mortality (an average mosquito life span)

From the basic cases $(\mathcal{S}_h, \mathcal{I}_{Sh}, \mathcal{I}_{Ah}, \mathcal{R}_h, \mathcal{S}_m, \mathcal{I}_m) \geq 0$.

In this approach, the aggregate human and mosquito population ratios are provided by

$$\mathcal{N}_h = \mathcal{S}_h + \mathcal{I}_{Sh} + \mathcal{I}_{Ah} + \mathcal{R}_h \text{ and } \mathcal{N}_m = \mathcal{S}_m + \mathcal{I}_m.$$

In addition, the area of biological significance for the aforementioned dengue model is indicated and presented by the covered set

$$\Phi = \left\{ \mathcal{S}_h, \mathcal{I}_{Sh}, \mathcal{I}_{Ah}, \mathcal{R}_h, \mathcal{S}_m, \mathcal{I}_m \in \mathbb{R}_+^6 : \mathcal{S}_h + \mathcal{I}_{Sh} + \mathcal{I}_{Ah} + \mathcal{R}_h \leq \mathcal{N}_h, \mathcal{S}_m + \mathcal{I}_m \leq \mathcal{N}_m \right\}$$

A fractional representation of the $\mathfrak{ABC}$ model as

$$\begin{aligned} {}^{ABC}_0\mathcal{D}_t^\eta \mathcal{S}_h &= \pi_h - \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h \mathcal{I}_m - \delta_h \mathcal{S}_h + \theta \mathcal{R}_h \\ {}^{ABC}_0\mathcal{D}_t^\eta \mathcal{I}_{Sh} &= \alpha_m \chi_{\mathcal{S}_h} \mathcal{S}_h \mathcal{I}_m - (\tau + \gamma + \delta_h) \mathcal{I}_{Sh} \\ {}^{ABC}_0\mathcal{D}_t^\eta \mathcal{I}_{Ah} &= \alpha_m \chi_{\mathcal{A}_h} \mathcal{S}_h \mathcal{I}_m - (\rho + \gamma + \delta_h) \mathcal{I}_{Ah} \\ {}^{ABC}_0\mathcal{D}_t^\eta \mathcal{R}_h &= \gamma (\mathcal{I}_{Sh} + \mathcal{I}_{Ah}) - \theta \mathcal{R}_h - \delta_h \mathcal{R}_h + \tau \mathcal{I}_{Sh} + \rho \mathcal{I}_{Ah} \\ {}^{ABC}_0\mathcal{D}_t^\eta \mathcal{S}_m &= \pi_m - \alpha_m \chi_m (\mathcal{I}_{Sh} + \mathcal{I}_{Ah}) \mathcal{S}_m - \delta_m \mathcal{S}_m \\ {}^{ABC}_0\mathcal{D}_t^\eta \mathcal{I}_m &= \alpha_m \chi_m (\mathcal{I}_{Sh} + \mathcal{I}_{Ah}) \mathcal{S}_m - \delta_m \mathcal{I}_m \end{aligned} \quad (6)$$

A fractional derivation of Atangana-Baleanu of order $0 < \eta < 1$ is denoted ${}^{ABC}_0\mathcal{D}_t^\eta$ in Caputo notation.

4 Model analysis

This section examines the validity, singularity and positive variance of the solution of the SIR-SI type model. Additionally, a reliability estimate for Model (Equation 6) has also been developed.

4.1 Existence and uniqueness

Theorem 4.1.

For each non-negative initial stage $(S_h(0), I_{S_h}(0), I_{A_h}(0), R_h(0), S_m(0), I_m(0)) \in \mathbb{R}_+^6$, then there survives a oneness solution of fractional order model (Equation 6).

Proof

Let $\Phi = \{S_h, I_{S_h}, I_{A_h}, R_h, S_m, I_m \in \mathbb{R}_+^6 : \max(|S_h|, |I_{S_h}|, |I_{A_h}|, |R_h|, |S_m|, |I_m|) \leq \varepsilon\}$.

Define a mapping

$M(x) = (M_1(x), M_2(x), M_3(x), M_4(x), M_5(x), M_6(x))$ and

$$M_1(x) = \pi_h - \alpha_m (\chi_{S_h} + \chi_{A_h}) S_h I_m - \delta_h S_h + \theta R_h$$

$$M_2(x) = \alpha_m \chi_{S_h} S_h I_m - (\tau + \gamma + \delta_h) I_{S_h}$$

$$M_3(x) = \alpha_m \chi_{A_h} S_h I_m - (\varrho + \gamma + \delta_h) I_{A_h}$$

$$M_4(x) = \gamma (I_{S_h} + I_{A_h}) - \theta R_h - \delta_h R_h + \tau I_{S_h} + \varrho I_{A_h}$$

$$M_5(x) = \pi_m - \alpha_m \chi_m (I_{S_h} + I_{A_h}) S_m - \delta_m S_m$$

$$M_6(x) = \alpha_m \chi_m (I_{S_h} + I_{A_h}) S_m - \delta_m I_m$$

Where $x = (S_h, I_{S_h}, I_{A_h}, R_h, S_m, I_m) \in \Phi$

For any $x, \bar{x} \in \Phi$, we have

$$\begin{aligned} \|M(x) - \overline{M(x)}\| &= |M_1(x) - \overline{M_1(x)}| + |M_2(x) - \overline{M_2(x)}| \\ &+ |M_3(x) - \overline{M_3(x)}| + |M_4(x) - \overline{M_4(x)}| \\ &+ |M_5(x) - \overline{M_5(x)}| + |M_6(x) - \overline{M_6(x)}| \\ &\leq |\pi_h - \alpha_m (\chi_{S_h} + \chi_{A_h}) S_h I_m - \delta_h S_h + \theta R_h - \pi_h| \\ &+ |\alpha_m (\chi_{S_h} + \chi_{A_h}) \overline{S_h I_m} + \delta_h \overline{S_h} \overline{R_h}| \\ &+ |\alpha_m \chi_{S_h} S_h I_m - (\tau + \gamma + \delta_h) I_{S_h} - \alpha_m \chi_{S_h} \overline{S_h I_m} \\ &+ (\tau + \gamma + \delta_h) \overline{I_{S_h}}| \\ &+ |\alpha_m \chi_{A_h} S_h I_m - (\varrho + \gamma + \delta_h) I_{A_h} - \alpha_m \chi_{A_h} \overline{S_h I_m} \\ &+ (\varrho + \gamma + \delta_h) \overline{I_{A_h}}| \\ &+ |\gamma (I_{S_h} + I_{A_h}) - \theta R_h - \delta_h R_h + \tau I_{S_h} + \varrho I_{A_h} \\ &- \gamma (\overline{I_{S_h}} + \overline{I_{A_h}}) + \theta \overline{R_h} + \delta_h \overline{R_h} - \tau \overline{I_{S_h}} - \varrho \overline{I_{A_h}}| \\ &+ |\pi_m - \alpha_m \chi_m (I_{S_h} + I_{A_h}) S_m - \delta_m S_m - \pi_m + \alpha_m \chi_m \\ &(\overline{I_{S_h}} + \overline{I_{A_h}}) \overline{S_m} + \delta_m \overline{S_m}| + |\alpha_m \chi_m \\ &(\overline{I_{S_h}} + \overline{I_{A_h}}) \overline{S_m} - \delta_m \overline{I_m} - \alpha_m \chi_m (\overline{I_{S_h}} + \overline{I_{A_h}}) \overline{S_m} + \delta_m \overline{I_m}| \\ &\leq [2\alpha_m (\chi_{S_h} + \chi_{A_h}) \mathcal{P} + \delta_h] |S_h - \overline{S_h}| + (2\tau + 2\gamma + \delta_h) \\ &|I_{S_h} - \overline{I_{S_h}}| + (2\varrho + 2\gamma + \delta_h) \end{aligned}$$

$$\begin{aligned} &|I_{A_h} - \overline{I_{A_h}}| + (2\theta + \delta_h) |R_h - \overline{R_h}| + [2\alpha_m (\mathcal{Q} + \mathcal{R}) \chi_m + \delta_h] \\ &|S_m - \overline{S_m}| + \delta_m |I_m - \overline{I_m}| \leq \mathcal{K} \|x - \bar{x}\|. \end{aligned}$$

$$\text{Where } \mathcal{K} = \max \left\{ \begin{array}{l} 2\alpha_m (\chi_{S_h} + \chi_{A_h}) \mathcal{P} + \delta_h, \\ 2(\tau + \gamma) + \delta_h, 2(\varrho + \gamma) + \delta_h, 2\theta + \delta_h, \\ 2\alpha_m (\mathcal{Q} + \mathcal{R}) \chi_m + \delta_h, \delta_m \end{array} \right\}$$

Basically, since $M(x)$ satisfies the Lipschitz requirement. Model (Equation 6) has a singular solution based on Lemma 1.

4.2 Positivity solution

Since system (Equation 6) deals with mosquitoes and populace, all components of system are positive. Following is our discussion:

Theorem 4.2.

Let $(S_h, I_{S_h}, I_{A_h}, R_h, S_m, I_m) > 0$ be represent the system (Equation 6) solution for the primary points $S_h(0), I_{S_h}(0), I_{A_h}(0), R_h(0), S_m(0), I_m(0)$ and represents an immutable set

$$\Phi = \{S_h, I_{S_h}, I_{A_h}, R_h, S_m, I_m \in \mathbb{R}_+^6 : \mathcal{N}_h = \frac{\pi_h}{\delta_h}, \mathcal{N}_m = \frac{\pi_m}{\delta_m}\},$$

then, all elements of the closed set

Φ is traveling in $\mathbb{R}_+^6$ space is positive invariant.

Proof

The given equation is used to construct the Lyapunov function:

$$\mathbb{L}(t) = (\mathbb{L}_1(t), \mathbb{L}_2(t)) = (S_h + I_{S_h} + I_{A_h} + R_h, S_m + I_m)$$

The function $\mathbb{L}(t)$ satisfies

$$\begin{aligned} \mathbb{L}(\dot{t}) &= (\mathbb{L}_1(\dot{t}), \mathbb{L}_2(\dot{t})) = (\dot{S}_h + \dot{I}_{S_h} + \dot{I}_{A_h} + \dot{R}_h, \dot{S}_m + \dot{I}_m) \\ &= (\pi_h - \delta_h S_h - I_{S_h} \delta_h - I_{A_h} \delta_h - \delta_h R_h, \pi_m - \delta_m S_m - \delta_m I_m) \\ &= (\pi_h - \delta_h \mathbb{L}_1, \pi_m - \delta_m \mathbb{L}_2) \quad (7) \end{aligned}$$

Therefore, it is simple to demonstrate Equation 7 as regards:

$$\begin{cases} \mathbb{L}_1(\dot{t}) = \pi_h - \delta_h \mathbb{L}_1 \leq 0, & \text{for } \mathbb{L}_1 \geq \frac{\pi_h}{\delta_h} \\ \mathbb{L}_2(\dot{t}) = \pi_m - \delta_m \mathbb{L}_2 \leq 0, & \text{for } \mathbb{L}_2 \geq \frac{\pi_m}{\delta_m} \end{cases} \quad (8)$$

Inferring $\mathbb{L}(\dot{t}) \leq 0$ from the above equations, which indicates that f is positively stable collection. On the other hand, by solving system (Equation 6)

$$0 \leq (\mathbb{L}_1(t), \mathbb{L}_2(t)) < \left(\frac{\pi_h}{\delta_h} + \mathbb{L}_1(0) e^{-\delta_h t}, \frac{\pi_m}{\delta_m} + \mathbb{L}_2(0) e^{-\delta_m t} \right)$$

Where $\mathbb{L}_1(0)$ and $\mathbb{L}_2(0)$ are the primary states of $\mathbb{L}_1(t)$ and $\mathbb{L}_2(t)$ respectively. Therefore, $t \rightarrow \infty, 0 \leq (\mathbb{L}_1(t), \mathbb{L}_2(t)) \leq \left(\frac{\pi_h}{\delta_h}, \frac{\pi_m}{\delta_m} \right)$ and we can conclusion that Φ is a desirable set.

This establishes the theorem.

4.3 Basic reproduction value $\mathcal{R}_0$

Let $\mathcal{C}_f = (S_h^*, I_{S_h}^*, I_{A_h}^*, R_h^*, S_m^*, I_m^*)$ be the contagious free equilibrium of Equation 6. We have $\mathcal{C}_f = \left(\frac{\pi_h}{\delta_h}, 0, 0, \frac{\pi_m}{\delta_m}, 0 \right)$. The algorithm of the next iteration matrix is utilized to estimate $\mathcal{R}_0$.

Obviously, the infected compartments are $\mathcal{I}_{S_h}$, $\mathcal{I}_{A_h}$ and $\mathcal{I}_m$ as a consequence of Equation 6. There are

$$\begin{aligned} \mathcal{A}^{BC}_0 \mathcal{D}_t^\eta \mathcal{I}_{S_h} &= \alpha_m \chi_{S_h} \mathcal{S}_h \mathcal{I}_m - (\tau + \gamma + \delta_h) \mathcal{I}_{S_h} \\ \mathcal{A}^{BC}_0 \mathcal{D}_t^\eta \mathcal{I}_{A_h} &= \alpha_m \chi_{A_h} \mathcal{S}_h \mathcal{I}_m - (\varrho + \gamma + \delta_h) \mathcal{I}_{A_h} \\ \mathcal{A}^{BC}_0 \mathcal{D}_t^\eta \mathcal{I}_m &= \alpha_m \chi_m (\mathcal{I}_{S_h} + \mathcal{I}_{A_h}) \mathcal{S}_m - \delta_m \mathcal{I}_m \end{aligned} \quad (9)$$

Then we derive

$$\mathcal{F} = \begin{pmatrix} 0 & 0 & \alpha_m \chi_{S_h} \mathcal{S}_h \\ 0 & 0 & \alpha_m \chi_{A_h} \mathcal{S}_h \\ \alpha_m \chi_m \mathcal{S}_m & \alpha_m \chi_m \mathcal{S}_m & 0 \end{pmatrix},$$

$$\mathcal{V} = \begin{pmatrix} \tau + \gamma + \delta_h & 0 & 0 \\ 0 & \varrho + \gamma + \delta_h & 0 \\ 0 & 0 & \delta_m \end{pmatrix}$$

The basic reproduction value is given by

$$\mathcal{R}_0 = \rho(\mathcal{FV}^{-1}) = \frac{\pi_h \pi_m \alpha_m^2 (\chi_{S_h} + \chi_{A_h})}{\delta_h (\delta_m)^2 (\tau + \gamma + \delta_h) (\varrho + \gamma + \delta_h)}$$

Where $\rho(\mathcal{FV}^{-1})$ denote the spectral radius. Surmise that $\mathcal{C}_p = (\mathcal{S}_h^{**}, \mathcal{I}_{S_h}^{**}, \mathcal{I}_{A_h}^{**}, \mathcal{R}_h^{**}, \mathcal{S}_m^{**}, \mathcal{I}_m^{**})$ represents the endemic equilibrium for Equation 6. So that

$$\begin{aligned} \mathcal{S}_h^{**} &= \frac{\pi_h (\theta + \delta_h) u_1 u_2}{(\alpha_m (\chi_{S_h} + \chi_{A_h}) \mathcal{I}_m + \delta_h) u_1 u_2 (\theta + \delta_h) - \theta \alpha_m \mathcal{I}_m u_4} \\ \mathcal{I}_{S_h}^{**} &= \frac{\alpha_m \chi_{S_h} \mathcal{S}_h \mathcal{I}_m}{u_1} \\ \mathcal{I}_{A_h}^{**} &= \frac{\alpha_m \chi_{A_h} \mathcal{S}_h \mathcal{I}_m}{u_2} \\ \mathcal{R}_h^{**} &= \frac{\alpha_m \mathcal{S}_h \mathcal{I}_m u_4}{(\theta + \delta_h) u_1 u_2} \\ \mathcal{S}_m^{**} &= \frac{u_1 u_2 \pi_m}{\alpha_m^2 \chi_m \mathcal{S}_h \mathcal{I}_m u_3 + \delta_m u_1 u_2} \\ \mathcal{I}_m^{**} &= \frac{(\theta + \delta_h) u_3 \alpha_m \pi_h \mathcal{S}_m - \delta_h u_1 u_2 \delta_m}{(\chi_{S_h} + \chi_{A_h}) u_5 - \theta u_4} \end{aligned}$$

Where $u_1 = \tau + \gamma + \delta_h$, $u_2 = \varrho + \gamma + \delta_h$, $u_3 = u_1 \chi_{A_h} + u_2 \chi_{S_h}$, $u_4 = u_1 \chi_{A_h} (\tau + \gamma) + u_2 \chi_{S_h} (\varrho + \gamma)$, and $u_5 = u_1 u_2 (\theta + \delta_h) \delta_m$.

4.4 Local stability

In this part, we are covering the analysis of firmness conditions of contagious free equilibrium $\mathcal{C}_f$ and contagious persistence equilibrium $\mathcal{C}_p$ points. A steady state analysis of this equilibrium results in the following Theorem 4.3 and Theorem 4.4.

The obtained Jacobian matrix is:

$$\mathcal{J} = \begin{pmatrix} -\alpha_m (\chi_{S_h} + \chi_{A_h}) \mathcal{I}_m - \delta_h & 0 & 0 & \theta & 0 & -\alpha_m (\chi_{S_h} + \chi_{A_h}) \mathcal{S}_h \\ \alpha_m \chi_{S_h} \mathcal{I}_m & -(\tau + \gamma + \delta_h) & 0 & 0 & 0 & \alpha_m \chi_{S_h} \mathcal{S}_h \\ \alpha_m \chi_{A_h} \mathcal{I}_m & 0 & -(\varrho + \gamma + \delta_h) & 0 & 0 & \alpha_m \chi_{A_h} \mathcal{S}_h \\ 0 & \tau + \gamma & \varrho + \gamma & -(\theta + \delta_h) & 0 & 0 \\ 0 & -\alpha_m \chi_m \mathcal{S}_m & -\alpha_m \chi_m \mathcal{S}_m & 0 & -(\alpha_m \chi_m (\mathcal{I}_{S_h} + \mathcal{I}_{A_h}) + \delta_m) & 0 \\ 0 & \alpha_m \chi_m \mathcal{S}_m & \alpha_m \chi_m \mathcal{S}_m & 0 & \alpha_m \chi_m (\mathcal{I}_{S_h} + \mathcal{I}_{A_h}) & -\delta_m \end{pmatrix} \quad (10)$$

Theorem 4.3.

If $\mathcal{R}_0 < 1$, the non-contagious equilibrium $\mathcal{C}_f$ is locally stable.

Proof

The structure (Equation 6) in the Jacobian matrix of $\mathcal{C}_f$ follows

$$\mathcal{J}(\mathcal{C}_f) = \begin{pmatrix} -\delta_h & 0 & 0 & \theta & 0 & -\alpha_m (\chi_{S_h} + \chi_{A_h}) \frac{\pi_h}{\delta_h} \\ 0 & -(\tau + \gamma + \delta_h) & 0 & 0 & 0 & \alpha_m \chi_{S_h} \frac{\pi_h}{\delta_h} \\ 0 & 0 & -(\varrho + \gamma + \delta_h) & 0 & 0 & \alpha_m \chi_{A_h} \frac{\pi_h}{\delta_h} \\ 0 & \tau + \gamma & \varrho + \gamma & -(\theta + \delta_h) & 0 & 0 \\ 0 & -\alpha_m \chi_m \frac{\pi_m}{\delta_m} & -\alpha_m \chi_m \frac{\pi_m}{\delta_m} & 0 & \delta_m & 0 \\ 0 & \alpha_m \chi_m \frac{\pi_m}{\delta_m} & \alpha_m \chi_m \frac{\pi_m}{\delta_m} & 0 & 0 & -\delta_m \end{pmatrix} \quad (11)$$

To determine the eigenvalue from the above-described matrix $\det(\mathcal{J}(\mathcal{C}_f) - \lambda \mathcal{I}) = 0$

We obtain the Eigen values $\lambda_1 = -\delta_h$, $\lambda_2 = -(\tau + \gamma + \delta_h)$, $\lambda_3 = \delta_m$, $\lambda_4 = \theta + \delta_h$ and the characteristic relation is

$$\lambda^2 + (\varrho + \gamma + \delta_h + \delta_m) \lambda + \delta_m (\tau + \gamma + \delta_h) (\varrho + \gamma + \delta_h) [1 - \mathcal{R}_0] = 0$$

When $\mathcal{R}_0 < 1$, it is obvious that $\lambda_5 < 1$ and $\lambda_6 < 1$, all the Eigen values satisfy the condition $|\arg(\lambda_i)| > \frac{\pi}{2}$, $i = 1, 2, \dots, 6$. the without contagious equilibrium $\mathcal{C}_f$ is locally asymptotically stable.

Theorem 4.4.

If $\mathcal{R}_0 > 1$, the equilibrium point $\mathcal{C}_p$ is locally stable, then system (Equation 6) has ubiquitous contagion.

Proof

Jacobian matrix evaluated in static equilibrium:

$$\det(\mathcal{J}(\mathcal{C}_p) - \lambda \mathcal{I}) = 0 \quad (12)$$

We obtain the Eigen values are $\lambda_1 = (\theta + \delta_h)$, $\lambda_2 = (\alpha_m \chi_m (\mathcal{I}_{S_h}^{**} + \mathcal{I}_{A_h}^{**}) + \delta_m)$, $\lambda_3 = \alpha_m (\chi_{S_h} + \chi_{A_h}) \mathcal{I}_m^{**} + \delta_h$ and the characteristic relation

$$\lambda^3 + a_1 \lambda^2 + a_2 \lambda + a_3 = 0 \quad (13)$$

Where

$$\begin{aligned} a_1 &= (\tau + \varrho + 2(\gamma + \delta_h)) \\ a_2 &= (\varrho + \gamma + \delta_h + \delta_m) [(\tau + \gamma + \delta_h) - \alpha_m^2 (\chi_{S_h} + \chi_{A_h}) \chi_m \mathcal{S}_h^{**} \mathcal{S}_m^{**}] \\ a_3 &= (\tau + \gamma + \delta_h) [(\varrho + \gamma + \delta_h) \delta_m - \alpha_m^2 (\chi_{S_h} + \chi_{A_h}) \chi_m \mathcal{S}_h^{**} \mathcal{S}_m^{**}] \end{aligned}$$

By using Routh-Hurwitz Criteria (22, 23), if the following provisions are handling

$$a_1 > 0, a_2 > 0, a_3 > 0 \text{ and } a_1 a_2 - a_3 > 0.$$

Then C_p is approximately stable locally. The evidence is conclusive.

4.5 Global stability

Theorem 4.5.

If $\mathcal{R}_0 < 1$, the point of without contagious equilibrium C_f is global stability on Φ .

Proof

Create a Lyapunov function $\mathbb{V}_1(t)$,

$$\mathbb{V}_1(t) = (\mathcal{S}_h - \mathcal{S}_h^* \ln \mathcal{S}_h) + \mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h} + \mathcal{R}_h + (\mathcal{S}_m - \mathcal{S}_m^* \ln \mathcal{S}_m) + \mathcal{I}_m \quad (14)$$

Calculating the fractional order derivatives of $\mathbb{V}_1(t)$ in the solution direction of Equation 6, from Lemma 2, we obtain

$$\begin{aligned} {}^{ABC}_0 \mathcal{D}_t^\eta \mathbb{V}_1(t) &\leq \left(1 - \frac{\mathcal{S}_h^*}{\mathcal{S}_h}\right) {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{S}_h \\ &+ {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{I}_{\mathcal{S}_h} + {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{I}_{\mathcal{A}_h} + {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{R}_h \\ &+ \left(1 - \frac{\mathcal{S}_m^*}{\mathcal{S}_m}\right) {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{S}_m + {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{I}_m \\ &\leq \left(1 - \frac{\mathcal{S}_h^*}{\mathcal{S}_h}\right) (\pi_h - \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h \mathcal{I}_m - \delta_h \mathcal{S}_h + \theta \mathcal{R}_h) \\ &+ (\alpha_m \chi_{\mathcal{S}_h} \mathcal{S}_h \mathcal{I}_m - (\tau + \gamma + \delta_h) \mathcal{I}_{\mathcal{S}_h}) \\ &+ (\alpha_m \chi_{\mathcal{A}_h} \mathcal{S}_h \mathcal{I}_m - (\varrho + \gamma + \delta_h) \mathcal{I}_{\mathcal{A}_h}) + (\gamma (\mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h}) \\ &- \theta \mathcal{R}_h - \delta_h \mathcal{R}_h + \tau \mathcal{I}_{\mathcal{S}_h} + \varrho \mathcal{I}_{\mathcal{A}_h}) \\ &+ \left(1 - \frac{\mathcal{S}_m^*}{\mathcal{S}_m}\right) (\pi_m - \alpha_m \chi_m (\mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h}) \mathcal{S}_m - \delta_m \mathcal{S}_m) \\ &+ (\alpha_m \chi_m (\mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h}) \mathcal{S}_m - \delta_m \mathcal{I}_m) \\ &\leq \pi_h \left(1 - \frac{\mathcal{S}_h^*}{\mathcal{S}_h}\right) + \alpha_m \chi_{\mathcal{S}_h} \mathcal{S}_h^* \mathcal{I}_m + \alpha_m \chi_{\mathcal{A}_h} \mathcal{S}_h^* \mathcal{I}_m \\ &- \delta_h \mathcal{S}_h + \delta_h \mathcal{S}_h^* + \theta \mathcal{R}_h \left(1 - \frac{\mathcal{S}_h^*}{\mathcal{S}_h}\right) - \delta_h \mathcal{I}_{\mathcal{S}_h} \\ &- \delta_h \mathcal{I}_{\mathcal{A}_h} - \theta \mathcal{R}_h - \delta_h \mathcal{R}_h + \pi_m \left(1 - \frac{\mathcal{S}_m^*}{\mathcal{S}_m}\right) + \alpha_m \chi_m \\ &(\mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h}) \mathcal{S}_m^* - \delta_m \left(1 - \frac{\mathcal{S}_m^*}{\mathcal{S}_m}\right) - \delta_m \mathcal{I}_m \end{aligned}$$

Substituting the reaction of without contagious free $C_f = \left(\frac{\pi_h}{\delta_h}, 0, 0, \frac{\pi_m}{\delta_m}, 0\right)$, we obtain:

$${}^{ABC}_0 \mathcal{D}_t^\eta \mathbb{V}_1(t) \leq (\pi_h + \delta_h) \left(2 - \frac{\mathcal{S}_h^*}{\mathcal{S}_h} - \frac{\mathcal{S}_h}{\mathcal{S}_h^*}\right) + (\pi_m - \delta_m) \left(2 - \frac{\mathcal{S}_m^*}{\mathcal{S}_m} - \frac{\mathcal{S}_m}{\mathcal{S}_m^*}\right) \quad (15)$$

It is clear that each term in Equation 15 must be negative. We have ${}^{ABC}_0 \mathcal{D}_t^\eta \mathbb{V}_1(t) \leq 0$ due to LaSalle's invariance principle (24), the function ${}^{ABC}_0 \mathcal{D}_t^\eta \mathbb{V}_1(t)$ is required to be negative finite.

The maximally invariant sets $\mathcal{S}_h = \mathcal{S}_h^*, \mathcal{S}_m = \mathcal{S}_m^*$ which is singleton $C_f = (\mathcal{S}_h^*, \mathcal{I}_{\mathcal{S}_h}^*, \mathcal{I}_{\mathcal{A}_h}^*, \mathcal{R}_h^*, \mathcal{S}_m^*, \mathcal{I}_m^*)$ contains the limit set for each solution. This demonstrates C_f is globally asymptotically stable on Φ .

Theorem 4.6.

When $\mathcal{R}_0 > 1$, the positive contagious equalization level of system (Equation 6) arises and is globally stable on Φ .

Proof

Let's create a lyapunov function of the following form

$$\begin{aligned} \mathbb{V}_2(t) &= (\mathcal{S}_h - \mathcal{S}_h^* \ln \mathcal{S}_h) + (\mathcal{I}_{\mathcal{S}_h} - \mathcal{I}_{\mathcal{S}_h}^* \ln \mathcal{I}_{\mathcal{S}_h}) \\ &+ (\mathcal{I}_{\mathcal{A}_h} - \mathcal{I}_{\mathcal{A}_h}^* \ln \mathcal{I}_{\mathcal{A}_h}) \\ &+ (\mathcal{R}_h - \mathcal{R}_h^* \ln \mathcal{R}_h) + (\mathcal{S}_m - \mathcal{S}_m^* \ln \mathcal{S}_m) + (\mathcal{I}_m - \mathcal{I}_m^* \ln \mathcal{I}_m) \\ {}^{ABC}_0 \mathcal{D}_t^\eta \mathbb{V}_2(t) &\leq \left(1 - \frac{\mathcal{S}_h^*}{\mathcal{S}_h}\right) {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{S}_h \\ &+ \left(1 - \frac{\mathcal{I}_{\mathcal{S}_h}^*}{\mathcal{I}_{\mathcal{S}_h}}\right) {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{I}_{\mathcal{S}_h} + \left(1 - \frac{\mathcal{I}_{\mathcal{A}_h}^*}{\mathcal{I}_{\mathcal{A}_h}}\right) {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{I}_{\mathcal{A}_h} + \\ &\left(1 - \frac{\mathcal{R}_h^*}{\mathcal{R}_h}\right) {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{R}_h \\ &+ \left(1 - \frac{\mathcal{S}_m^*}{\mathcal{S}_m}\right) {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{S}_m + \left(1 - \frac{\mathcal{I}_m^*}{\mathcal{I}_m}\right) {}^{ABC}_0 \mathcal{D}_t^\eta \mathcal{I}_m \\ &\leq \left(1 - \frac{\mathcal{S}_h^*}{\mathcal{S}_h}\right) (\pi_h - \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h \mathcal{I}_m - \delta_h \mathcal{S}_h + \theta \mathcal{R}_h) \\ &+ \left(1 - \frac{\mathcal{I}_{\mathcal{S}_h}^*}{\mathcal{I}_{\mathcal{S}_h}}\right) (\alpha_m \chi_{\mathcal{S}_h} \mathcal{S}_h \mathcal{I}_m - (\tau + \gamma + \delta_h) \mathcal{I}_{\mathcal{S}_h}) \\ &+ \left(1 - \frac{\mathcal{I}_{\mathcal{A}_h}^*}{\mathcal{I}_{\mathcal{A}_h}}\right) (\alpha_m \chi_{\mathcal{A}_h} \mathcal{S}_h \mathcal{I}_m - (\varrho + \gamma + \delta_h) \mathcal{I}_{\mathcal{A}_h}) + \left(1 - \frac{\mathcal{R}_h^*}{\mathcal{R}_h}\right) \\ &(\gamma (\mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h}) - \theta \mathcal{R}_h - \delta_h \mathcal{R}_h + \tau \mathcal{I}_{\mathcal{S}_h} + \varrho \mathcal{I}_{\mathcal{A}_h}) \\ &+ \left(1 - \frac{\mathcal{S}_m^*}{\mathcal{S}_m}\right) (\pi_m - \alpha_m \chi_m (\mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h}) \mathcal{S}_m - \delta_m \mathcal{S}_m) + \left(1 - \frac{\mathcal{I}_m^*}{\mathcal{I}_m}\right) \\ &(\alpha_m \chi_m (\mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h}) \mathcal{S}_m - \delta_m \mathcal{I}_m) \\ &\leq \pi_h \left(1 - \frac{\mathcal{S}_h^*}{\mathcal{S}_h}\right) + \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h^* \mathcal{I}_m - \delta_h \mathcal{S}_h \left(1 - \frac{\mathcal{S}_h^*}{\mathcal{S}_h}\right) \\ &+ \theta \mathcal{R}_h \frac{\mathcal{S}_h^*}{\mathcal{S}_h} + \mathcal{I}_{\mathcal{S}_h}^* (\tau + \gamma) - \delta_h \mathcal{I}_{\mathcal{S}_h} \left(1 - \frac{\mathcal{I}_{\mathcal{S}_h}^*}{\mathcal{I}_{\mathcal{S}_h}}\right) - \alpha_m \mathcal{S}_h \\ &\left(\chi_{\mathcal{S}_h} \frac{\mathcal{I}_{\mathcal{S}_h}^*}{\mathcal{I}_{\mathcal{S}_h}} + \chi_{\mathcal{A}_h} \frac{\mathcal{I}_{\mathcal{A}_h}^*}{\mathcal{I}_{\mathcal{A}_h}}\right) \mathcal{I}_m + (\varrho + \gamma) \mathcal{I}_{\mathcal{A}_h}^* \\ &- \delta_h \mathcal{I}_{\mathcal{A}_h} \left(1 - \frac{\mathcal{I}_{\mathcal{A}_h}^*}{\mathcal{I}_{\mathcal{A}_h}}\right) - \gamma (\mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h}) \\ &\frac{\mathcal{R}_h^*}{\mathcal{R}_h} + \theta \mathcal{R}_h^* - \delta_h \mathcal{R}_h \left(1 - \frac{\mathcal{R}_h^*}{\mathcal{R}_h}\right) + (\tau \mathcal{I}_{\mathcal{S}_h} + \varrho \mathcal{I}_{\mathcal{A}_h}) \frac{\mathcal{R}_h^*}{\mathcal{R}_h} \\ &+ \pi_m \left(1 - \frac{\mathcal{S}_m^*}{\mathcal{S}_m}\right) - \delta_m \left(1 - \frac{\mathcal{I}_m^*}{\mathcal{I}_m}\right) - \delta_m \left(1 - \frac{\mathcal{I}_m^*}{\mathcal{I}_m}\right) \\ &+ \alpha_m \chi_m (\mathcal{I}_{\mathcal{S}_h} + \mathcal{I}_{\mathcal{A}_h}) \left(\mathcal{S}_m^* + \mathcal{S}_m \frac{\mathcal{I}_m^*}{\mathcal{I}_m}\right) \\ {}^{ABC}_0 \mathcal{D}_t^\eta \mathbb{V}_2(t) &\leq (\pi_h - \delta_h + \theta \mathcal{R}_h) \end{aligned}$$

$$\begin{aligned} & \left(2 - \frac{\mathcal{S}_h^{**}}{\mathcal{S}_h} - \frac{\mathcal{S}_h}{\mathcal{S}_h^{**}}\right) + (\tau + \gamma - \delta_h) \left(2 - \frac{\mathcal{T}_{\mathcal{S}_h}^{**}}{\mathcal{T}_{\mathcal{S}_h}} - \frac{\mathcal{T}_{\mathcal{S}_h}}{\mathcal{T}_{\mathcal{S}_h}^{**}}\right) \\ & + (\varrho + \gamma - \delta_h) \\ & \left(2 - \frac{\mathcal{T}_{\mathcal{A}_h}^{**}}{\mathcal{T}_{\mathcal{A}_h}} - \frac{\mathcal{T}_{\mathcal{A}_h}}{\mathcal{T}_{\mathcal{A}_h}^{**}}\right) + (\gamma + \delta_h - \theta) \left(2 - \frac{\mathcal{R}_h^{**}}{\mathcal{R}_h} - \frac{\mathcal{R}_h}{\mathcal{R}_h^{**}}\right) \\ & + (\pi_m - \delta_m) \left(2 - \frac{\mathcal{S}_m^{**}}{\mathcal{S}_m} - \frac{\mathcal{S}_m}{\mathcal{S}_m^{**}}\right) + \delta_m \left(2 - \frac{\mathcal{T}_m^{**}}{\mathcal{T}_m} - \frac{\mathcal{T}_m}{\mathcal{T}_m^{**}}\right) \quad (16) \end{aligned}$$

Hence, the condition in Equation 16 ensures

$ABC_0 \mathcal{D}_t^\eta \mathbb{V}_2(t) \leq 0$ for all $(\mathcal{S}_h^{**}, \mathcal{T}_{\mathcal{S}_h}^{**}, \mathcal{T}_{\mathcal{A}_h}^{**}, \mathcal{R}_h^{**}, \mathcal{S}_m^{**}, \mathcal{T}_m^{**}) \in \Phi$ and strict the quality holds for $\mathcal{S}_h = \mathcal{S}_h = \mathcal{S}_h^{**}, \mathcal{T}_{\mathcal{S}_h} = \mathcal{T}_{\mathcal{S}_h}^{**}, \mathcal{T}_{\mathcal{A}_h} = \mathcal{T}_{\mathcal{A}_h}^{**}, \mathcal{R}_h = \mathcal{R}_h^{**}, \mathcal{S}_m = \mathcal{S}_m^{**}$ and $\mathcal{T}_m = \mathcal{T}_m^{**}$. therefore the equilibrium point $\mathcal{C}_p$ becomes globally stable on Φ .

5 Optimum control approach

In this portion, we will discuss how to optimize the problem and analyze the performance of the control function. Consolidation of optimal controlling problem a dynamics of control system can be described as system (Equation 6).

$$\begin{aligned} ABC_0 \mathcal{D}_t^\eta \mathcal{S}_h &= \pi_h - \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h \mathcal{T}_m - \delta_h \mathcal{S}_h + \theta \mathcal{R}_h - \mathbb{U}_1(t) \mathcal{S}_h \\ ABC_0 \mathcal{D}_t^\eta \mathcal{T}_{\mathcal{S}_h} &= \alpha_m \chi_{\mathcal{S}_h} \mathcal{S}_h \mathcal{T}_m - (\tau + \gamma + \delta_h) \mathcal{T}_{\mathcal{S}_h} \\ ABC_0 \mathcal{D}_t^\eta \mathcal{T}_{\mathcal{A}_h} &= \alpha_m \chi_{\mathcal{A}_h} \mathcal{S}_h \mathcal{T}_m - (\varrho + \gamma + \delta_h) \mathcal{T}_{\mathcal{A}_h} \\ ABC_0 \mathcal{D}_t^\eta \mathcal{S}_m &= \pi_m - \alpha_m \chi_m (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) \mathcal{S}_m - \delta_m \mathcal{S}_m - \mathbb{U}_2(t) \mathcal{S}_m \\ ABC_0 \mathcal{D}_t^\eta \mathcal{T}_m &= \alpha_m \chi_m (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) \mathcal{S}_m - \delta_m \mathcal{T}_m - \mathbb{U}_2(t) \mathcal{T}_m \end{aligned}$$

Where

$\mathbb{U}_1$ — Self-precaution (long sleeved pants and shorts, increase immune system, consultation at

the nearest health care) minimizes the susceptible individuals.

$\mathbb{U}_2$ — Use of chemical insecticide sprays destroying the susceptible and infected mosquito cases

The optimal solution being minimized could be expressed as:

$$\mathbb{C}(\mathbb{U}_1, \mathbb{U}_2) = \int_0^{t_f} \left(a\mathcal{S}_h + b\mathcal{T}_{\mathcal{S}_h} + c\mathcal{T}_{\mathcal{A}_h} + d\mathcal{S}_m + e\mathcal{T}_m + f\mathbb{U}_1^2 + g\mathbb{U}_2^2 \right) dt \quad (18)$$

To reduce the cost of two controls $\mathbb{U}_1$ and $\mathbb{U}_2$ the objective is reduced $\mathcal{S}_h, \mathcal{T}_{\mathcal{S}_h}, \mathcal{T}_{\mathcal{A}_h}$ and $\mathcal{S}_m, \mathcal{T}_m$.

Therefore, we need to obtain optimal controls $\mathbb{U}_1^*$ and $\mathbb{U}_2^*$

$$\mathbb{C}(\mathbb{U}_1^*, \mathbb{U}_2^*) = \min_{\mathbb{U}_1, \mathbb{U}_2} \{ \mathbb{C}(\mathbb{U}_1, \mathbb{U}_2) | \mathbb{U}_1, \mathbb{U}_2 \in \Phi \} \quad (19)$$

A set of constraints $\Phi = \{(\mathbb{U}_1, \mathbb{U}_2) | \mathbb{U}_i : [0, t_f] \rightarrow [0, \infty) \text{ lebesgue quantifiable } i = 1, 2\}$.

The expense of minimizing $\mathcal{S}_h, \mathcal{T}_{\mathcal{S}_h}, \mathcal{T}_{\mathcal{A}_h}, \mathcal{S}_m$ and $\mathcal{T}_m$ is represented by the term $a\mathcal{S}_h, b\mathcal{T}_{\mathcal{S}_h}, c\mathcal{T}_{\mathcal{A}_h}, d\mathcal{S}_m$ and $e\mathcal{T}_m$ respectively. Likewise, $f\mathbb{U}_1^2, g\mathbb{U}_2^2$ represents the cost for controls $\mathbb{U}_1, \mathbb{U}_2$. The most prevalent PMP can be used to find the

adequacy condition required for the control system to be satisfied. Equations 17, 19 can be transformed into the following point-wise Hamiltonian $\mathbb{H}$ for $(\mathbb{U}_1, \mathbb{U}_2)$ regression problem using the aforesaid principle.

$$\begin{aligned} \mathbb{H} &= \{ a\mathcal{S}_h + b\mathcal{T}_{\mathcal{S}_h} + c\mathcal{T}_{\mathcal{A}_h} + d\mathcal{S}_m + e\mathcal{T}_m + f\mathbb{U}_1^2 + g\mathbb{U}_2^2 \} \\ &+ \lambda_{\mathcal{S}_h} \{ \pi_h - \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h \mathcal{T}_m - \delta_h \mathcal{S}_h + \theta \mathcal{R}_h - \mathbb{U}_1(t) \mathcal{S}_h \} \\ &+ \lambda_{\mathcal{T}_{\mathcal{S}_h}} \{ \alpha_m \chi_{\mathcal{S}_h} \mathcal{S}_h \mathcal{T}_m - (\tau + \gamma + \delta_h) \mathcal{T}_{\mathcal{S}_h} \} \\ &+ \lambda_{\mathcal{T}_{\mathcal{A}_h}} \{ \alpha_m \chi_{\mathcal{A}_h} \mathcal{S}_h \mathcal{T}_m - (\varrho + \gamma + \delta_h) \mathcal{T}_{\mathcal{A}_h} \} \\ &+ \lambda_{\mathcal{S}_m} \{ \pi_m - \alpha_m \chi_m (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) \mathcal{S}_m - \delta_m \mathcal{S}_m - \mathbb{U}_2(t) \mathcal{S}_m \} \\ &+ \lambda_{\mathcal{T}_m} \{ \alpha_m \chi_m (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) \mathcal{S}_m - \delta_m \mathcal{T}_m - \mathbb{U}_2(t) \mathcal{T}_m \} \quad (20) \end{aligned}$$

Where $\lambda_{\mathcal{S}_h}, \lambda_{\mathcal{T}_{\mathcal{S}_h}}, \lambda_{\mathcal{T}_{\mathcal{A}_h}}, \lambda_{\mathcal{S}_m}$ and $\lambda_{\mathcal{T}_m}$ are the ad-joint variable or co-state variable.

$$\begin{aligned} \frac{d\lambda_{\mathcal{S}_h}}{dt} &= \frac{\partial \mathbb{H}}{\partial \mathcal{S}_h} = a + \lambda_{\mathcal{S}_h} \{ -\alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{T}_m - \delta_h \} \\ &- \left(\lambda_{\mathcal{T}_{\mathcal{S}_h}} + \lambda_{\mathcal{T}_{\mathcal{A}_h}} \right) \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{T}_m \\ \frac{d\lambda_{\mathcal{T}_{\mathcal{S}_h}}}{dt} &= \frac{\partial \mathbb{H}}{\partial \mathcal{T}_{\mathcal{S}_h}} = b - \lambda_{\mathcal{S}_h} \mathbb{U}_1(t) + \lambda_{\mathcal{T}_{\mathcal{S}_h}} \{ -(\tau + \gamma + \delta_h) \} \\ &+ \lambda_{\mathcal{T}_m} \alpha_m \chi_m \mathcal{S}_m \\ \frac{d\lambda_{\mathcal{T}_{\mathcal{A}_h}}}{dt} &= \frac{\partial \mathbb{H}}{\partial \mathcal{T}_{\mathcal{A}_h}} = c - \lambda_{\mathcal{S}_h} \mathbb{U}_1(t) \\ &+ \lambda_{\mathcal{T}_{\mathcal{A}_h}} \{ -(\varrho + \gamma + \delta_h) \} + \lambda_{\mathcal{T}_m} \alpha_m \chi_m \mathcal{S}_m \\ \frac{d\lambda_{\mathcal{S}_m}}{dt} &= \frac{\partial \mathbb{H}}{\partial \mathcal{S}_m} \\ (17) \quad &= d + \lambda_{\mathcal{S}_m} \{ -\alpha_m \chi_m (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) - \delta_m - \mathbb{U}_2(t) \} \\ &+ \lambda_{\mathcal{T}_m} \{ \alpha_m \chi_m (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) \} \\ \frac{d\lambda_{\mathcal{T}_m}}{dt} &= \frac{\partial \mathbb{H}}{\partial \mathcal{T}_m} = e\lambda_{\mathcal{T}_m} \{ -\delta_m - \mathbb{U}_2(t) \} \\ &- \left(\lambda_{\mathcal{T}_{\mathcal{S}_h}} + \lambda_{\mathcal{T}_{\mathcal{A}_h}} \right) \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h + \lambda_{\mathcal{S}_h} \\ &\{ -\alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h \} \quad (21) \end{aligned}$$

The conditions for transversality are

$$\lambda_{\mathcal{S}_h}(t_f) = \lambda_{\mathcal{T}_{\mathcal{S}_h}}(t_f) = \lambda_{\mathcal{T}_{\mathcal{A}_h}}(t_f) = \lambda_{\mathcal{S}_m}(t_f) = \lambda_{\mathcal{T}_m}(t_f) = 0.$$

For $0 < \mathbb{U}_i < 1, i = 1, 2$. From the interior of controls, we have:

$$\begin{aligned} \frac{\partial \mathbb{H}}{\partial \mathbb{U}_1} &= 2f\mathbb{U}_1 - \lambda_{\mathcal{S}_h} (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) \\ \frac{\partial \mathbb{H}}{\partial \mathbb{U}_2} &= 2g\mathbb{U}_2 - \lambda_{\mathcal{S}_m} \mathcal{S}_m - \lambda_{\mathcal{T}_m} \mathcal{T}_m \quad (22) \end{aligned}$$

From where:

$$\begin{aligned} \mathbb{U}_1 &= \frac{\lambda_{\mathcal{S}_h} (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h})}{2f} \\ \mathbb{U}_2 &= \frac{\lambda_{\mathcal{S}_m} \mathcal{S}_m + \lambda_{\mathcal{T}_m} \mathcal{T}_m}{2g} \quad (23) \end{aligned}$$

5.1 Utilization of optimal solutions

Theorem 5.1. $(\mathbb{U}_1^*, \mathbb{U}_2^*)$ is a control factor can reduce over $\mathbb{U}$ provided by

$$\begin{aligned}\mathbb{U}_1^* &= \max \left\{ 0, \min \left\{ 1, \frac{1}{2f} \lambda_{\mathbb{S}_h} \left(\mathfrak{T}_{\mathbb{S}_h} + \mathfrak{T}_{\mathbb{A}_h} \right) \right\} \right\} \\ \mathbb{U}_2^* &= \max \left\{ 0, \min \left\{ 1, \frac{1}{2g} \lambda_{\mathbb{S}_m} \mathbb{S}_m + \lambda_{\mathfrak{T}_m} \mathfrak{T}_m \right\} \right\}\end{aligned}\quad (24)$$

Where $\lambda_{\mathbb{S}_h}, \lambda_{\mathfrak{T}_{\mathbb{S}_h}}, \lambda_{\mathfrak{T}_{\mathbb{A}_h}}, \lambda_{\mathbb{S}_m}$ and $\lambda_{\mathfrak{T}_m}$ are co-state variable that satisfy the condition (Equations 17–24) in addition, the transversality characteristic that follows

$$\lambda_{\mathbb{S}_h}(t_f) = \lambda_{\mathfrak{T}_{\mathbb{S}_h}}(t_f) = \lambda_{\mathfrak{T}_{\mathbb{A}_h}}(t_f) = \lambda_{\mathbb{S}_m}(t_f) = \lambda_{\mathfrak{T}_m}(t_f) = 0.$$

$$\mathbb{U}_1^* = \begin{cases} 0 & \text{if } \mathbb{U}_1 \leq 0, \\ \mathbb{U}_1 & \text{if } 0 < \mathbb{U}_1 < 1, \\ 1 & \text{if } \mathbb{U}_1 \geq 0. \end{cases}$$

And

$$\mathbb{U}_2^* = \begin{cases} 0 & \text{if } \mathbb{U}_2 \leq 0, \\ \mathbb{U}_2 & \text{if } 0 < \mathbb{U}_2 < 1, \\ 1 & \text{if } \mathbb{U}_2 \geq 0. \end{cases}\quad (25)$$

Proof

To demonstrate the survival of optimal control solutions, the configuration of the Lipschitz criterion of the system and the convexity of the integral in Equation 21 are related and state variable that constrains $\mathbb{U}_1$ and $\mathbb{U}_2$ to the boundary of the state solution. So we employ PMP and get the following:

$$\begin{aligned}\mathcal{ABC}_0 \mathcal{D}_t^\eta \lambda_{\mathbb{S}_h}(t) &= \frac{\partial \mathbb{H}}{\partial \mathbb{S}_h}; \\ \mathcal{ABC}_0 \mathcal{D}_t^\eta \lambda_{\mathfrak{T}_{\mathbb{S}_h}}(t) &= \frac{\partial \mathbb{H}}{\partial \mathfrak{T}_{\mathbb{S}_h}}; \mathcal{ABC}_0 \mathcal{D}_t^\eta \lambda_{\mathfrak{T}_{\mathbb{A}_h}}(t) = \frac{\partial \mathbb{H}}{\partial \mathfrak{T}_{\mathbb{A}_h}}; \\ \mathcal{ABC}_0 \mathcal{D}_t^\eta \lambda_{\mathbb{S}_m}(t) &= \frac{\partial \mathbb{H}}{\partial \mathbb{S}_m}; \mathcal{ABC}_0 \mathcal{D}_t^\eta \lambda_{\mathfrak{T}_m}(t) = \frac{\partial \mathbb{H}}{\partial \mathfrak{T}_m};\end{aligned}\quad (26)$$

with,

$$\lambda_{\mathbb{S}_h}(t_f) = \lambda_{\mathfrak{T}_{\mathbb{S}_h}}(t_f) = \lambda_{\mathfrak{T}_{\mathbb{A}_h}}(t_f) = \lambda_{\mathbb{S}_m}(t_f) = \lambda_{\mathfrak{T}_m}(t_f) = 0.$$

The Hamilton can be differentiated with regard to achieve the conditional optimum:

$$\frac{\partial \mathbb{H}}{\partial \mathbb{U}_1} = 0, \frac{\partial \mathbb{H}}{\partial \mathbb{U}_2} = 0. \quad (27)$$

The ad-joint system (Equations 20, 21) derived from Equation 17, the optimum system (Equation 23) is accessible from Equation 24. The optimal method is the constrained system (Equation 17) and its initial state is ad-joint the system includes (Equation 20), and condition for intersection.

6 Adams-Bash forth method

Here, we formulate the system of Equation 6 a recently invented numerical approach, the Adams-Bash forth method (24). The framework (Equation 6) can be used to test the essential theorem from fractional calculus,

$$\begin{aligned}\mathbb{S}_h(t) &= \mathbb{S}_h(0) + \frac{1-\eta}{\mathcal{ABC}(\eta)} \mathbb{K}_1 \left(t, \mathbb{S}_h(t), \mathfrak{T}_{\mathbb{S}_h}(t), \mathfrak{T}_{\mathbb{A}_h}(t), \right. \\ &\quad \left. \mathfrak{R}_h(t), \mathbb{S}_m(t), \mathfrak{T}_m(t) \right) + \\ &\quad \frac{\eta}{\mathcal{ABC}(\eta) \Gamma(\eta)} \int_0^t \mathbb{K}_1(\varpi, \mathbb{S}_h(\varpi), \mathfrak{T}_{\mathbb{S}_h}(\varpi), \\ &\quad \mathfrak{T}_{\mathbb{A}_h}(\varpi), \\ &\quad \mathfrak{R}_h(\varpi), \mathbb{S}_m(\varpi), \mathfrak{T}_m(\varpi)) (t-\varpi)^{\eta-1} d\varpi\end{aligned}\quad (28)$$

$$\begin{aligned}\mathfrak{T}_{\mathbb{S}_h}(t) &= \mathfrak{T}_{\mathbb{S}_h}(0) + \frac{1-\eta}{\mathcal{ABC}(\eta)} \mathbb{K}_2 \left(t, \mathbb{S}_h(t), \mathfrak{T}_{\mathbb{S}_h}(t), \mathfrak{T}_{\mathbb{A}_h}(t), \right. \\ &\quad \left. \mathfrak{R}_h(t), \mathbb{S}_m(t), \mathfrak{T}_m(t) \right) + \\ &\quad \frac{\eta}{\mathcal{ABC}(\eta) \Gamma(\eta)} \int_0^t \mathbb{K}_2(\varpi, \mathbb{S}_h(\varpi), \mathfrak{T}_{\mathbb{S}_h}(\varpi), \mathfrak{T}_{\mathbb{A}_h}(\varpi), \mathfrak{R}_h \\ &\quad (\varpi), \mathbb{S}_m(\varpi), \mathfrak{T}_m(\varpi)) (t-\varpi)^{\eta-1} d\varpi\end{aligned}\quad (29)$$

$$\begin{aligned}\mathfrak{T}_{\mathbb{A}_h}(t) &= \mathfrak{T}_{\mathbb{A}_h}(0) + \frac{1-\eta}{\mathcal{ABC}(\eta)} \mathbb{K}_3 \left(t, \mathbb{S}_h(t), \mathfrak{T}_{\mathbb{S}_h}(t), \mathfrak{T}_{\mathbb{A}_h}(t), \right. \\ &\quad \left. \mathfrak{R}_h(t), \mathbb{S}_m(t), \mathfrak{T}_m(t) \right) + \\ &\quad \frac{\eta}{\mathcal{ABC}(\eta) \Gamma(\eta)} \int_0^t \mathbb{K}_3(\varpi, \mathbb{S}_h(\varpi), \\ &\quad \mathfrak{T}_{\mathbb{S}_h}(\varpi), \mathfrak{T}_{\mathbb{A}_h}(\varpi), \mathfrak{R}_h(\varpi), \\ &\quad \mathbb{S}_m(\varpi), \mathfrak{T}_m(\varpi)) (t-\varpi)^{\eta-1} d\varpi\end{aligned}\quad (30)$$

$$\begin{aligned}\mathfrak{R}_h(t) &= \mathfrak{R}_h(0) + \frac{1-\eta}{\mathcal{ABC}(\eta)} \mathbb{K}_4 \left(t, \mathbb{S}_h(t), \mathfrak{T}_{\mathbb{S}_h}(t), \right. \\ &\quad \left. \mathfrak{T}_{\mathbb{A}_h}(t), \mathfrak{R}_h(t), \mathbb{S}_m(t), \mathfrak{T}_m(t) \right) + \\ &\quad \frac{\eta}{\mathcal{ABC}(\eta) \Gamma(\eta)} \int_0^t \mathbb{K}_4(\varpi, \mathbb{S}_h(\varpi), \mathfrak{T}_{\mathbb{S}_h}(\varpi), \\ &\quad \mathfrak{T}_{\mathbb{A}_h}(\varpi), \\ &\quad \mathfrak{R}_h(\varpi), \mathbb{S}_m(\varpi), \mathfrak{T}_m(\varpi)) (t-\varpi)^{\eta-1} d\varpi\end{aligned}\quad (31)$$

$$\begin{aligned}\mathbb{S}_m(t) &= \mathbb{S}_m(0) + \frac{1-\eta}{\mathcal{ABC}(\eta)} \mathbb{K}_5 \left(t, \mathbb{S}_h(t), \right. \\ &\quad \left. \mathfrak{T}_{\mathbb{S}_h}(t), \mathfrak{T}_{\mathbb{A}_h}(t), \mathfrak{R}_h(t), \mathbb{S}_m(t), \mathfrak{T}_m(t) \right) + \\ &\quad \frac{\eta}{\mathcal{ABC}(\eta) \Gamma(\eta)} \int_0^t \mathbb{K}_5(\varpi, \mathbb{S}_h(\varpi), \mathfrak{T}_{\mathbb{S}_h}(\varpi), \\ &\quad \mathfrak{T}_{\mathbb{A}_h}(\varpi), \\ &\quad \mathfrak{R}_h(\varpi), \mathbb{S}_m(\varpi), \mathfrak{T}_m(\varpi)) (t-\varpi)^{\eta-1} d\varpi\end{aligned}\quad (32)$$

$$\begin{aligned} \mathcal{T}_m(t) &= \mathcal{T}_m(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_6(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \\ &\mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \mathcal{S}_m(t), \mathcal{T}_m(t)) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^t \mathbb{K}_6(\omega, \mathcal{S}_h(\omega), \mathcal{T}_{\mathcal{S}_h}(\omega), \\ &\mathcal{T}_{\mathcal{A}_h}(\omega), \mathcal{R}_h(\omega), \mathcal{S}_m(\omega), \mathcal{T}_m(\omega)) (t-\omega)^{\eta-1} d\omega \quad (33) \end{aligned}$$

Where,

$$\left\{ \begin{aligned} \mathbb{K}_1(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \mathcal{S}_m(t), \mathcal{T}_m(t)) \\ &= \pi_h - \alpha_m (\chi_{\mathcal{S}_h} + \chi_{\mathcal{A}_h}) \mathcal{S}_h \mathcal{T}_m - \delta_h \mathcal{S}_h + \theta \mathcal{R}_h \\ \mathbb{K}_2(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \mathcal{S}_m(t), \mathcal{T}_m(t)) \\ &= \alpha_m \chi_{\mathcal{S}_h} \mathcal{S}_h \mathcal{T}_m - (\tau + \gamma + \delta_h) \mathcal{T}_{\mathcal{S}_h} \\ \mathbb{K}_3(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \mathcal{S}_m(t), \mathcal{T}_m(t)) \\ &= \alpha_m \chi_{\mathcal{A}_h} \mathcal{S}_h \mathcal{T}_m - (\varrho + \gamma + \delta_h) \mathcal{T}_{\mathcal{A}_h} \\ \mathbb{K}_4(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \mathcal{S}_m(t), \mathcal{T}_m(t)) \\ &= \gamma (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) - \theta \mathcal{R}_h - \delta_h \mathcal{R}_h + \tau \mathcal{T}_{\mathcal{S}_h} + \varrho \mathcal{T}_{\mathcal{A}_h} \\ \mathbb{K}_5(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \mathcal{S}_m(t), \mathcal{T}_m(t)) \\ &= \pi_m - \alpha_m \chi_m (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) \mathcal{S}_m - \Omega \mathcal{S}_m - \delta_m \mathcal{S}_m \\ \mathbb{K}_6(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \mathcal{S}_m(t), \mathcal{T}_m(t)) \\ &= \alpha_m \chi_m (\mathcal{T}_{\mathcal{S}_h} + \mathcal{T}_{\mathcal{A}_h}) \mathcal{S}_m - (\Omega + \delta_m) \mathcal{T}_m \end{aligned} \right. \quad (34)$$

The following structure is obtained at time t_{n+1} ,

$$\begin{aligned} \mathcal{S}_h(t_{n+1}) &= \mathcal{S}_h(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_1 \\ &(t_n, \mathcal{S}_{h_n}, \mathcal{T}_{\mathcal{S}_{h_n}}, \mathcal{T}_{\mathcal{A}_{h_n}}, \mathcal{R}_{h_n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_{n+1}} \mathbb{K}_1(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_{n+1}-t)^{\eta-1} dt \quad (35) \end{aligned}$$

$$\begin{aligned} \mathcal{T}_{\mathcal{S}_h}(t_{n+1}) &= \mathcal{T}_{\mathcal{S}_h}(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_2 \\ &(t_n, \mathcal{S}_{h_n}, \mathcal{T}_{\mathcal{S}_{h_n}}, \mathcal{T}_{\mathcal{A}_{h_n}}, \mathcal{R}_{h_n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_{n+1}} \mathbb{K}_2(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_{n+1}-t)^{\eta-1} dt \quad (36) \end{aligned}$$

$$\begin{aligned} \mathcal{T}_{\mathcal{A}_h}(t_{n+1}) &= \mathcal{T}_{\mathcal{A}_h}(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_3 \\ &(t_n, \mathcal{S}_{h_n}, \mathcal{T}_{\mathcal{S}_{h_n}}, \mathcal{T}_{\mathcal{A}_{h_n}}, \mathcal{R}_{h_n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_{n+1}} \mathbb{K}_3(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_{n+1}-t)^{\eta-1} dt \quad (37) \end{aligned}$$

$$\begin{aligned} \mathcal{R}_h(t_{n+1}) &= \mathcal{R}_h(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_4 \\ &(t_n, \mathcal{S}_{h_n}, \mathcal{T}_{\mathcal{S}_{h_n}}, \mathcal{T}_{\mathcal{A}_{h_n}}, \mathcal{R}_{h_n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_{n+1}} \mathbb{K}_4(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_{n+1}-t)^{\eta-1} dt \quad (38) \end{aligned}$$

$$\begin{aligned} \mathcal{S}_m(t_{n+1}) &= \mathcal{S}_m(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_5 \\ &(t_n, \mathcal{S}_{h_n}, \mathcal{T}_{\mathcal{S}_{h_n}}, \mathcal{T}_{\mathcal{A}_{h_n}}, \mathcal{R}_{h_n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_{n+1}} \mathbb{K}_5(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_{n+1}-t)^{\eta-1} dt \quad (39) \end{aligned}$$

$$\begin{aligned} \mathcal{T}_m(t_{n+1}) &= \mathcal{T}_m(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_6 \\ &(t_n, \mathcal{S}_{h_n}, \mathcal{T}_{\mathcal{S}_{h_n}}, \mathcal{T}_{\mathcal{A}_{h_n}}, \mathcal{R}_{h_n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_{n+1}} \mathbb{K}_6(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_{n+1}-t)^{\eta-1} dt \quad (40) \end{aligned}$$

While, at t_n we have

$$\begin{aligned} \mathcal{S}_h(t_n) &= \mathcal{S}_h(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_1(t_{n-1}, \mathcal{S}_{h_{n-1}}, \mathcal{T}_{\mathcal{S}_{h_{n-1}}}, \mathcal{T}_{\mathcal{A}_{h_{n-1}}}, \\ &\mathcal{R}_{h_{n-1}}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_n} \mathbb{K}_1(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_n-t)^{\eta-1} dt \quad (41) \end{aligned}$$

$$\begin{aligned} \mathcal{T}_{\mathcal{S}_h}(t_n) &= \mathcal{T}_{\mathcal{S}_h}(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_2(t_{n-1}, \mathcal{S}_{h_{n-1}}, \mathcal{T}_{\mathcal{S}_{h_{n-1}}}, \mathcal{T}_{\mathcal{A}_{h_{n-1}}}, \\ &\mathcal{R}_{h_{n-1}}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_n} \mathbb{K}_2(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_n-t)^{\eta-1} dt \quad (42) \end{aligned}$$

$$\begin{aligned} \mathcal{T}_{\mathcal{A}_h}(t_n) &= \mathcal{T}_{\mathcal{A}_h}(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_3(t_{n-1}, \mathcal{S}_{h_{n-1}}, \mathcal{T}_{\mathcal{S}_{h_{n-1}}}, \\ &\mathcal{T}_{\mathcal{A}_{h_{n-1}}}, \mathcal{R}_{h_{n-1}}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_n} \mathbb{K}_3(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_n-t)^{\eta-1} dt \quad (43) \end{aligned}$$

$$\begin{aligned} \mathcal{R}_h(t_n) &= \mathcal{R}_h(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_4(t_{n-1}, \mathcal{S}_{h_{n-1}}, \mathcal{T}_{\mathcal{S}_{h_{n-1}}}, \mathcal{T}_{\mathcal{A}_{h_{n-1}}}, \\ &\mathcal{R}_{h_{n-1}}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_n} \mathbb{K}_4(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_n-t)^{\eta-1} dt \quad (44) \end{aligned}$$

$$\begin{aligned} \mathcal{S}_m(t_n) &= \mathcal{S}_m(0) + \frac{1-\eta}{ABC(\eta)} \mathbb{K}_5(t_{n-1}, \mathcal{S}_{h_{n-1}}, \mathcal{T}_{\mathcal{S}_{h_{n-1}}}, \mathcal{T}_{\mathcal{A}_{h_{n-1}}}, \\ &\mathcal{R}_{h_{n-1}}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}}) + \\ &\frac{\eta}{ABC(\eta) \lceil \eta} \int_0^{t_n} \mathbb{K}_5(t, \mathcal{S}_h(t), \mathcal{T}_{\mathcal{S}_h}(t), \mathcal{T}_{\mathcal{A}_h}(t), \mathcal{R}_h(t), \\ &\mathcal{S}_m(t), \mathcal{T}_m(t)) (t_n-t)^{\eta-1} dt \end{aligned}$$

$$+\mathbb{A}_{\eta,1}^1-\mathbb{A}_{\eta,2}^1 \tag{53}$$

$$\begin{aligned} \mathfrak{T}_{\mathcal{S}_\eta}(t_{n+1}) &= \mathfrak{T}_{\mathcal{S}_\eta}(t_n) + \frac{1-\eta}{ABC(\eta)} \\ &\left[\mathbb{K}_2 \left(t_n, \mathcal{S}_{\theta_n}, \mathfrak{T}_{\mathcal{S}_{\theta_n}}, \mathfrak{T}_{\mathcal{A}_{\theta_n}}, \mathfrak{R}_{\theta_n}, \right. \right. \\ &\quad \left. \left. \mathfrak{S}_{m_n}, \mathfrak{T}_{m_n} \right) - \right. \\ &\quad \left. \mathbb{K}_2 \left(t_{n-1}, \mathcal{S}_{\theta_{n-1}}, \mathfrak{T}_{\mathcal{S}_{\theta_{n-1}}}, \mathfrak{T}_{\mathcal{A}_{\theta_{n-1}}}, \mathfrak{R}_{\theta_{n-1}}, \right. \right. \\ &\quad \left. \left. \mathfrak{S}_{m_{n-1}}, \mathfrak{T}_{m_{n-1}} \right) \right] \\ &+ \mathbb{A}_{n,1}^2 - \mathbb{A}_{n,2}^2 \end{aligned} \quad (54)$$

$$\begin{aligned} & \mathfrak{T}_{\mathcal{A}_h}(\mathfrak{k}_{n+1}) + \mathfrak{T}_{\mathcal{A}_h}(\mathfrak{k}_n) + \frac{1-\eta}{ABC(\eta)} \\ & \left[\begin{array}{c} \mathbb{K}_3\left(\mathfrak{k}_n, \mathfrak{S}_{\mathfrak{h}_n}, \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}_n}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_n}}, \mathfrak{R}_{\mathfrak{h}_n}, \mathfrak{S}_{\mathfrak{m}_n}, \mathfrak{T}_{\mathfrak{m}_n}\right) - \\ \mathbb{K}_3\left(\mathfrak{k}_{n-1}, \mathfrak{S}_{\mathfrak{h}_{n-1}}, \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}_{n-1}}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_{n-1}}}, \mathfrak{R}_{\mathfrak{h}_{n-1}}, \mathfrak{S}_{\mathfrak{m}_{n-1}}, \mathfrak{T}_{\mathfrak{m}_{n-1}}\right) \end{array} \right] \\ & + \mathbb{A}_{n,1}^3 - \mathbb{A}_{n,2}^3 \end{aligned} \quad (55)$$

$$\begin{aligned} \mathfrak{R}_h(t_{n+1}) &= \mathfrak{R}_h(t_n) + \frac{1-\eta}{ABC(\eta)} \\ &\left[\mathbb{K}_4 \left(t_n, \mathfrak{S}_{h_n}, \mathfrak{T}_{\mathfrak{S}_{h_n}}, \mathfrak{T}_{\mathcal{A}_{h_n}}, \mathfrak{R}_{h_n}, \mathfrak{S}_{m_n}, \mathfrak{T}_{m_n} \right) - \right. \\ &\left. \mathbb{K}_4 \left(t_{n-1}, \mathfrak{S}_{h_{n-1}}, \mathfrak{T}_{\mathfrak{S}_{h_{n-1}}}, \mathfrak{T}_{\mathcal{A}_{h_{n-1}}}, \mathfrak{R}_{h_{n-1}}, \mathfrak{S}_{m_{n-1}}, \mathfrak{T}_{m_{n-1}} \right) \right] \\ &+ \mathbb{A}_{n,1}^4 - \mathbb{A}_{n,2}^4 \end{aligned} \quad (56)$$

$$\begin{aligned} \mathbb{S}_m(t_{n+1}) &= \mathbb{S}_m(t_n) + \frac{1-\eta}{ABC(\eta)} \\ &\left[\mathbb{K}_5 \left(t_n, \mathbb{S}_{\eta_n}, \mathbb{T}_{\mathbb{S}_{\eta_n}}, \mathbb{T}_{\mathbb{A}_{\eta_n}}, \mathbb{R}_{\eta_n}, \mathbb{S}_{m_n}, \mathbb{T}_{m_n} \right) - \right. \\ &\left. \mathbb{K}_5 \left(t_{n-1}, \mathbb{S}_{\eta_{n-1}}, \mathbb{T}_{\mathbb{S}_{\eta_{n-1}}}, \mathbb{T}_{\mathbb{A}_{\eta_{n-1}}}, \mathbb{R}_{\eta_{n-1}}, \mathbb{S}_{m_{n-1}}, \mathbb{T}_{m_{n-1}} \right) \right] \\ &+ \mathbb{A}_{n,1}^5 - \mathbb{A}_{n,2}^5 \end{aligned} \quad (57)$$

$$\begin{aligned} \mathfrak{T}_m(\mathfrak{k}_{n+1}) &= \mathfrak{T}_m(\mathfrak{k}_n) + \frac{1-\eta}{\mathcal{ABC}(\eta)} \\ &\left[\mathbb{K}_6 \left(\mathfrak{k}_n, \mathfrak{S}_{\mathfrak{h}_n}, \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}_n}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_n}}, \mathfrak{R}_{\mathfrak{h}_n}, \mathfrak{S}_{m_n}, \mathfrak{T}_{m_n} \right) - \right. \\ &\left. \mathbb{K}_6 \left(\mathfrak{k}_{n-1}, \mathfrak{S}_{\mathfrak{h}_{n-1}}, \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}_{n-1}}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_{n-1}}}, \mathfrak{R}_{\mathfrak{h}_{n-1}}, \mathfrak{S}_{m_{n-1}}, \mathfrak{T}_{m_{n-1}} \right) \right] \\ &+ \mathbb{A}_{n,1}^6 - \mathbb{A}_{n,2}^6 \end{aligned} \quad (58)$$

Where

$$\mathbb{A}_{\eta,1}^1 = \frac{\eta}{ABC(\eta) \Gamma \eta} \int_0^{t_{n+1}} \mathbb{K}_1 \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{R}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_{n+1} - t)^{\eta-1} dt \quad (59)$$

$$\mathbb{A}_{\eta,1}^2 = \frac{\eta}{\mathcal{ABC}(\eta) \lceil \eta \rceil} \int_0^{t_{n+1}} \mathbb{K}_2 \left(t, \mathcal{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathcal{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{K}_{\mathfrak{h}}(t), \mathcal{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_{n+1} - t)^{\eta-1} dt \quad (60)$$

$$\mathbb{A}_{\eta,1}^3 = \frac{\eta}{ABC(\eta)|\eta} \int_0^{t_{n+1}} \mathbb{K}_3 \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{R}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_{n+1} - t)^{\eta-1} dt \quad (61)$$

$$\mathbb{A}_{\eta,1}^4 = \frac{\eta}{ABC(\eta)|\eta|} \int_0^{t_{n+1}} \mathbb{K}_4 \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{R}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_{n+1} - t)^{\eta-1} dt \quad (62)$$

$$\mathbb{A}_{\eta,1}^5 = \frac{\eta}{ABC(\eta)|\eta|} \int_0^{t_{n+1}} \mathbb{K}_S \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{K}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_{n+1} - t)^{\eta-1} dt \quad (63)$$

$$\mathbb{A}_{\eta,1}^6 = \frac{\eta}{ABC(\eta) \Gamma \eta} \int_0^{t_{n+1}} \mathbb{K}_6 \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{R}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_{n+1} - t)^{\eta-1} dt \quad (64)$$

and

$$\mathbb{A}_{\eta,2}^1 = \frac{\eta}{\mathcal{ABC}(\eta) \lceil \eta} \int_0^{t_n} \mathbb{K}_1 \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{K}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_n - t)^{\eta-1} dt \quad (65)$$

$$\mathbb{A}_{\eta,2}^2 = \frac{\eta}{\mathcal{ABC}(\eta) \lceil \eta} \int_0^{t_n} \mathbb{K}_2 \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{K}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_n - t)^{\eta-1} dt \quad (66)$$

$$\mathbb{A}_{\eta,2}^3 = \frac{\eta}{\mathcal{ABC}(\eta) \upharpoonright \eta} \int_0^{t_n} \mathbb{K}_3 \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{K}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_n - t)^{\eta-1} dt \quad (67)$$

$$\mathbb{A}_{\eta,2}^4 = \frac{\eta}{\mathcal{ABC}(\eta) \lceil \eta} \int_0^{t_n} \mathbb{K}_4 \left(t, \mathbb{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathbb{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(tL), \right. \\ \left. \mathfrak{K}_{\mathfrak{h}}(t), \mathbb{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_n - t)^{\eta-1} dt \quad (68)$$

$$\mathbb{A}_{\eta,2}^5 = \frac{\eta}{\mathcal{ABC}(\eta) \upharpoonright \eta} \int_0^{t_n} \mathbb{K}_5 \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{K}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_n - t)^{\eta-1} dt \quad (69)$$

$$\mathbb{A}_{\eta,2}^6 = \frac{\eta}{\mathcal{ABC}(\eta) \upharpoonright \eta} \int_0^{t_n} \mathbb{K}_6 \left(t, \mathfrak{S}_{\mathfrak{h}}(t), \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}}}(t), \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}}}(t), \right. \\ \left. \mathfrak{K}_{\mathfrak{h}}(t), \mathfrak{S}_{\mathfrak{m}}(t), \mathfrak{T}_{\mathfrak{m}}(t) \right) (t_n - t)^{\eta-1} dt \quad (70)$$

Now, approximating $\mathbb{A}_{\eta,1}^1$, $\mathbb{A}_{\eta,1}^2$, $\mathbb{A}_{\eta,1}^3$, $\mathbb{A}_{\eta,1}^4$, $\mathbb{A}_{\eta,1}^5$, $\mathbb{A}_{\eta,1}^6$ and $\mathbb{A}_{\eta,2}^1$, $\mathbb{A}_{\eta,2}^2$, $\mathbb{A}_{\eta,2}^3$, $\mathbb{A}_{\eta,2}^4$, $\mathbb{A}_{\eta,2}^5$, $\mathbb{A}_{\eta,2}^6$ with the help of Lagrange's polynomials

$$\mathbb{P}(t) \approx \frac{t - t_{n-1}}{t_n - t_{n-1}} f(t_n, y_n) + \frac{t - t_n}{t_{n-1} - t_n} f(t_{n-1}, y_{n-1}),$$

$$= \frac{t - t_{n-1}}{h} f(t_n, y_{n-1}) + \frac{t - t_n}{h} f(t_{n-1}, y_{n-1}) \quad (71)$$

Now, only consider the Equation 59 to evaluate under the Equation 71, that is given as

$$\begin{aligned} \mathbb{A}_{\eta,1}^1 &= \frac{\eta}{\mathcal{ABC}(\eta) \lceil \eta} \int_0^{k_{n+1}} (k_{n+1} - t)^{\eta-1} \\ &\left[\frac{k_{n+1} - t_{n-1}}{h} \mathbb{K}_1 \left(k_n, \mathcal{S}_{\eta n}, \mathcal{T}_{\mathcal{S}_{\eta n}}, \mathcal{T}_{\mathcal{A}_{\eta n}}, \mathcal{R}_{\eta n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n} \right) + \right. \\ &\left. \left[\frac{k_{n+1} - t_n}{h} \mathbb{K}_1 \left(k_{n+1}, \mathcal{S}_{\eta n-1}, \mathcal{T}_{\mathcal{S}_{\eta n-1}}, \mathcal{T}_{\mathcal{A}_{\eta n-1}}, \mathcal{R}_{\eta n-1}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}} \right) \right] d k_{n+1} \right. \\ &\left. \frac{\eta \mathbb{K}_1 \left(k_n, \mathcal{S}_{\eta n}, \mathcal{T}_{\mathcal{S}_{\eta n}}, \mathcal{T}_{\mathcal{A}_{\eta n}}, \mathcal{R}_{\eta n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n} \right)}{\mathcal{ABC}(\eta) \lceil \eta} \right. \\ &\left. \left[\int_0^{k_{n+1}} (k_{n+1} - t)^{\eta-1} \mathbb{K}_1(k_{n+1} - t) d k_{n+1} \right] d k_{n+1} \right. \\ &\left. \mathbb{K}_1 \left(k_{n+1}, \mathcal{S}_{\eta n-1}, \mathcal{T}_{\mathcal{S}_{\eta n-1}}, \mathcal{T}_{\mathcal{A}_{\eta n-1}}, \mathcal{R}_{\eta n-1}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}} \right) \right. \\ &\left. \times \frac{\eta}{\mathcal{ABC}(\eta) \lceil \eta} \right. \\ &\left. \left[\int_0^{k_{n+1}} (k_{n+1} - t)^{\eta-1} \mathbb{K}_1(k_{n+1} - t) d k_{n+1} \right] d k_{n+1} \right. \\ &\left. \mathbb{A}_{\eta,1}^1 = \frac{\eta \mathbb{K}_1 \left(k_n, \mathcal{S}_{\eta n}, \mathcal{T}_{\mathcal{S}_{\eta n}}, \mathcal{T}_{\mathcal{A}_{\eta n}}, \mathcal{R}_{\eta n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n} \right)}{\mathcal{ABC}(\eta) \lceil \eta} \left[\frac{2 h t_{n+1}^{\eta}}{\eta} - \frac{k_{n+1}^{\eta+1}}{\eta+1} \right] - \right. \\ &\left. \mathbb{K}_1 \left(k_{n+1}, \mathcal{S}_{\eta n-1}, \mathcal{T}_{\mathcal{S}_{\eta n-1}}, \mathcal{T}_{\mathcal{A}_{\eta n-1}}, \mathcal{R}_{\eta n-1}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}} \right) \right. \\ &\left. \times \frac{\eta}{\mathcal{ABC}(\eta) \lceil \eta} \right] \left[\frac{h t_{n+1}^{\eta}}{\eta} - \frac{k_{n+1}^{\eta+1}}{\eta+1} \right] \end{aligned} \quad (72)$$

Similarly, for $\mathbb{A}_{n,1}^2 \dots \mathbb{A}_{n,1}^6$

$$\begin{aligned} \mathbb{A}_{\eta,1}^2 &= \frac{\eta^{\mathbb{K}_2} \left(t_{\eta,n}, \mathfrak{S}_{\mathfrak{h}_{\eta,n}}, \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}_{\eta,n}}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_{\eta,n}}}, \mathfrak{R}_{\mathfrak{h}_{\eta,n}}, \mathfrak{S}_{\mathfrak{m}_{\eta,n}}, \mathfrak{T}_{\mathfrak{m}_{\eta,n}} \right)}{\mathcal{ABC}(\eta) \upharpoonright \eta \downarrow \mathfrak{h}} \\ &\quad \left[\frac{2\mathfrak{h} t_{\eta,n+1}^{\eta}}{\eta} - \frac{t_{\eta,n+1}^{\eta+1}}{\eta+1} \right] - \\ \mathbb{K}_2 \left(t_{\eta,n-1}, \mathfrak{S}_{\mathfrak{h}_{\eta,n-1}}, \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}_{\eta,n-1}}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_{\eta,n-1}}}, \mathfrak{R}_{\mathfrak{h}_{\eta,n-1}}, \mathfrak{S}_{\mathfrak{m}_{\eta,n-1}}, \mathfrak{T}_{\mathfrak{m}_{\eta,n-1}} \right) \\ &\quad \times \frac{\eta}{\mathcal{ABC}(\eta) \upharpoonright \eta \downarrow \mathfrak{h}} \left[\frac{\mathfrak{h} t_{\eta,n+1}^{\eta}}{\eta} - \frac{t_{\eta,n+1}^{\eta+1}}{\eta+1} \right] \end{aligned} \quad (73)$$

$$\begin{aligned} \mathbb{A}_{\eta,1}^3 &= \frac{\eta \mathbb{K}_3 \left(t_n, \mathfrak{S}_{\mathfrak{h}_n}, \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}_n}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_n}}, \mathfrak{R}_{\mathfrak{h}_n}, \mathfrak{S}_{\mathfrak{m}_n}, \mathfrak{T}_{\mathfrak{m}_n} \right)}{\overline{ABC}(\eta) \overline{\eta} \, \mathfrak{h}} \\ &\quad \left[\frac{2\mathfrak{h} \eta_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] - \\ \mathbb{K}_3 \left(t_{n-1}, \mathfrak{S}_{\mathfrak{h}_{n-1}}, \mathfrak{T}_{\mathfrak{S}_{\mathfrak{h}_{n-1}}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_{n-1}}}, \mathfrak{R}_{\mathfrak{h}_{n-1}}, \mathfrak{S}_{\mathfrak{m}_{n-1}}, \mathfrak{T}_{\mathfrak{m}_{n-1}} \right) \\ &\quad \times \frac{\eta}{\overline{ABC}(\eta) \overline{\eta} \, \mathfrak{h}} \left[\frac{\mathfrak{h} \eta_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \end{aligned} \quad (74)$$

$$\mathbb{A}_{\eta,1}^4 = \frac{\eta \mathbb{K}_4 \left(t_n, \mathfrak{S}_{0n}, \mathfrak{T}_{\mathfrak{S}_{0n}}, \mathfrak{T}_{\mathcal{A}_{0n}}, \mathfrak{R}_{0n}, \mathfrak{S}_{mn}, \mathfrak{T}_{mn} \right)}{\mathcal{AB}(\eta) [\eta \text{ } t]} \\ \left[\frac{2\eta t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] -$$

$$\mathbb{K}_4 \left(t_{n-1}, s_{\mathfrak{h}_{n-1}}, \mathfrak{T}_{S_{\mathfrak{h}_{n-1}}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_{n-1}}}, \mathfrak{K}_{\mathfrak{h}_{n-1}}, s_{m_{n-1}}, \mathfrak{T}_{m_{n-1}} \right) \\ \times \frac{\eta}{ABC(\eta) \Gamma(\eta)} \left[\frac{\mathfrak{h} t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \quad (75)$$

$$\begin{aligned} \mathbb{A}_{\eta,1}^5 &= \frac{\eta \mathbb{K}_5 \left(t_n, \mathbb{S}_{h_n}, \mathbb{T}_{\mathbb{S}_{h_n}}, \mathbb{T}_{\mathcal{A}_{h_n}}, \mathbb{R}_{h_n}, \mathbb{S}_{m_n}, \mathbb{T}_{m_n} \right)}{\mathcal{ABC}(\eta) \Gamma \eta \, \mathfrak{h}} \\ &\quad \left[\frac{2\mathfrak{h} t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] - \\ \mathbb{K}_5 \left(t_{n-1}, \mathbb{S}_{h_{n-1}}, \mathbb{T}_{\mathbb{S}_{h_{n-1}}}, \mathbb{T}_{\mathcal{A}_{h_{n-1}}}, \mathbb{R}_{h_{n-1}}, \mathbb{S}_{m_{n-1}}, \mathbb{T}_{m_{n-1}} \right) \\ &\quad \times \frac{\eta}{\mathcal{ABC}(\eta) \Gamma \eta \, \mathfrak{h}} \left[\frac{\mathfrak{h} t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \end{aligned} \quad (76)$$

$$\begin{aligned} \mathbb{A}_{\eta,1}^6 &= \frac{\eta \mathbb{K}_6 \left(t_n, \mathcal{S}_{h_n}, \mathcal{T}_{\mathcal{S}_{h_n}}, \mathcal{T}_{\mathcal{A}_{h_n}}, \mathcal{R}_{h_n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n} \right)}{\mathcal{ABC}(\eta) \Gamma \eta \, \mathfrak{h}} \\ &\quad \left[\frac{2\mathfrak{h} t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] - \\ \mathbb{K}_6 &\left(t_{n-1}, \mathcal{S}_{h_{n-1}}, \mathcal{T}_{\mathcal{S}_{h_{n-1}}}, \mathcal{T}_{\mathcal{A}_{h_{n-1}}}, \mathcal{R}_{h_{n-1}}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}} \right) \\ &\quad \times \frac{\eta}{\mathcal{ABC}(\eta) \Gamma \eta \, \mathfrak{h}} \left[\frac{\mathfrak{h} t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \end{aligned} \quad (77)$$

And, from $\mathbb{A}_{\eta,2}^1$ to $\mathbb{A}_{\eta,2}^6$ are given as

$$A_{\eta,2}^1 = \frac{\eta \mathbb{K}_1 \left(t_{\eta n}, \mathfrak{S}_{\eta n}, \mathfrak{T}_{\mathfrak{S}_{\eta n}}, \mathfrak{T}_{\mathcal{A}_{\eta n}}, \mathfrak{R}_{\eta n}, \mathfrak{S}_{m_n}, \mathfrak{T}_{m_n} \right)}{\mathcal{ABC}(\eta) \lceil \eta \text{ } \mathfrak{h} \text{ } \left[\frac{\mathfrak{h} t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] - \mathbb{K}_1 \left(t_{n-1}, \mathfrak{S}_{\eta n-1}, \mathfrak{T}_{\mathfrak{S}_{\eta n-1}}, \mathfrak{T}_{\mathcal{A}_{\eta n-1}}, \mathfrak{R}_{\eta n-1}, \mathfrak{S}_{m_{n-1}}, \mathfrak{T}_{m_{n-1}} \right)} \times \frac{1}{\mathcal{ABC}(\eta) \lceil \eta \text{ } \mathfrak{h} \text{ } } \quad (78)$$

$$A_{\eta,2}^2 = \frac{\eta \mathbb{K}_2 \left(t_n, \mathbb{S}_{\mathfrak{h}_n}, \mathfrak{T}_{\mathbb{S}_{\mathfrak{h}_n}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_n}}, \mathfrak{K}_{\mathfrak{h}_n}, \mathbb{S}_{\mathfrak{m}_n}, \mathfrak{T}_{\mathfrak{m}_n} \right)}{\mathcal{ABC}(\eta) \lceil \eta \ \mathfrak{h} \left[\frac{\mathfrak{h} t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] - \mathbb{K}_2 \left(t_{n-1}, \mathbb{S}_{\mathfrak{h}_{n-1}}, \mathfrak{T}_{\mathbb{S}_{\mathfrak{h}_{n-1}}}, \mathfrak{T}_{\mathcal{A}_{\mathfrak{h}_{n-1}}}, \mathfrak{K}_{\mathfrak{h}_{n-1}}, \mathbb{S}_{\mathfrak{m}_{n-1}}, \mathfrak{T}_{\mathfrak{m}_{n-1}} \right) \times \frac{1}{\mathcal{ABC}(\eta) \lceil \eta \ \mathfrak{h}}}$$

$$\mathbb{A}_{\eta,2}^3 = \frac{\eta \mathbb{K}_3 \left(t_n, \mathcal{S}_{\eta n}, \mathcal{T}_{\mathcal{S}_{\eta n}}, \mathcal{T}_{\mathcal{A}_{\eta n}}, \mathcal{R}_{\eta n}, \mathcal{S}_{m_n}, \mathcal{T}_{m_n} \right)}{\mathcal{ABC} \left(\eta \right) \left[\eta \right]} - \left[\frac{\eta t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] - \mathbb{K}_3 \left(t_{n-1}, \mathcal{S}_{\eta n-1}, \mathcal{T}_{\mathcal{S}_{\eta n-1}}, \mathcal{T}_{\mathcal{A}_{\eta n-1}}, \mathcal{R}_{\eta n-1}, \mathcal{S}_{m_{n-1}}, \mathcal{T}_{m_{n-1}} \right)$$

$$\times \frac{1}{ABC(\eta) \lceil \eta \rceil h} \quad (80)$$

$$\begin{aligned} A_{\eta,2}^4 &= \frac{\eta \mathbb{K}_4(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n})}{ABC(\eta) \lceil \eta \rceil h} \\ &\quad \left[\frac{h t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] - \\ \mathbb{K}_4(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) & \\ \times \frac{1}{ABC(\eta) \lceil \eta \rceil h} & \quad (81) \end{aligned}$$

$$\begin{aligned} A_{\eta,2}^5 &= \frac{\eta \mathbb{K}_5(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n})}{ABC(\eta) \lceil \eta \rceil h} \\ &\quad \left[\frac{h t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] - \\ \mathbb{K}_5(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) & \\ \times \frac{1}{ABC(\eta) \lceil \eta \rceil h} & \quad (82) \end{aligned}$$

$$\begin{aligned} A_{\eta,2}^6 &= \frac{\eta \mathbb{K}_6(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n})}{ABC(\eta) \lceil \eta \rceil h} \\ &\quad \left[\frac{h t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] - \\ \mathbb{K}_6(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) & \\ \times \frac{1}{ABC(\eta) \lceil \eta \rceil h} & \quad (83) \end{aligned}$$

Finally, using Equations 72–77, and Equations 78–82 in Equations 53–58 therefore, we obtain the numerical solution of model (Equation 6), as a result

$$\begin{aligned} S_h(t_{n+1}) &= S_h(t_n) + \frac{1-\eta}{ABC(\eta)} \\ &\quad \left[\mathbb{K}_1(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n}) - \right. \\ &\quad \left. \mathbb{K}_1(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \right] \\ &\quad + \frac{\eta \mathbb{K}_1(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n})}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{2h t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \\ &\quad - \mathbb{K}_1(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \\ &\quad \times \frac{\eta}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{h t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \\ &\quad - \frac{\eta \mathbb{K}_1(t_n, S_h(t_n), \mathcal{I}_{S_h(t_n)}, \mathcal{I}_{A_h(t_n)}, R_h(t_n), S_m(t_n), \mathcal{I}_m(t_n))}{ABC(\eta) \lceil \eta \rceil h} \\ &\quad \left[\frac{h t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] - \end{aligned}$$

$$\begin{aligned} &\mathbb{K}_1(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \\ &\times \frac{1}{ABC(\eta) \lceil \eta \rceil h} \quad (84) \end{aligned}$$

$$\begin{aligned} \mathcal{I}_{S_h}(t_{n+1}) &= \mathcal{I}_{S_h}(t_n) + \frac{1-\eta}{ABC(\eta)} \\ &\quad \left[\mathbb{K}_2(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n}) - \right. \\ &\quad \left. \mathbb{K}_2(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \right] \\ &\quad + \frac{\eta \mathbb{K}_2(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n})}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{2h t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \\ &\quad - \mathbb{K}_2(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \\ &\quad \times \frac{\eta}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{h t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \\ &\quad - \frac{\eta \mathbb{K}_2(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n})}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{h t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] - \\ &\quad \mathbb{K}_2(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \\ &\quad \times \frac{1}{ABC(\eta) \lceil \eta \rceil h} \quad (85) \end{aligned}$$

$$\begin{aligned} \mathcal{I}_{A_h}(t_{n+1}) &= \mathcal{I}_{A_h}(t_n) + \frac{1-\eta}{ABC(\eta)} \\ &\quad \left[\mathbb{K}_3(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n}) \right. \\ &\quad \left. - \mathbb{K}_3(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \right] \\ &\quad + \frac{\eta \mathbb{K}_3(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n})}{ABC(\eta) \lceil \eta \rceil h} \end{aligned}$$

$$\begin{aligned} &\left[\frac{2h t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] - \\ &\mathbb{K}_3(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \\ &\times \frac{\eta}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{h t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \\ &- \frac{\eta \mathbb{K}_3(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n})}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{h t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] - \\ &\mathbb{K}_3(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \\ &\times \frac{1}{ABC(\eta) \lceil \eta \rceil h} \quad (86) \end{aligned}$$

$$\begin{aligned} R_h(t_{n+1}) &= R_h(t_n) + \frac{1-\eta}{ABC(\eta)} \\ &\quad \left[\mathbb{K}_4(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n}) \right. \\ &\quad \left. - \mathbb{K}_4(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}}) \right] \end{aligned}$$

$$\begin{aligned}
& + \frac{\eta \mathbb{K}_4 \left(t_n, \mathfrak{S}_{h_n}, \mathfrak{I}_{S_{h_n}}, \mathfrak{I}_{A_{h_n}}, \mathfrak{R}_{h_n}, \mathfrak{S}_{m_n}, \mathfrak{I}_{m_n} \right)}{\mathcal{ABC}(\eta) \lceil \eta \rceil h} \left[\frac{2h t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] - \frac{\eta \mathbb{K}_4 \left(t_n, \mathfrak{S}_{h_n}, \mathfrak{I}_{S_{h_n}}, \mathfrak{I}_{A_{h_n}}, \mathfrak{R}_{h_n}, \mathfrak{S}_{m_n}, \mathfrak{I}_{m_n} \right)}{\mathcal{ABC}(\eta) \lceil \eta \rceil h} \left[\frac{h t_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] \\
& - \mathbb{K}_4 \left(t_{n-1}, \mathfrak{S}_{h_{n-1}}, \mathfrak{I}_{S_{h_{n-1}}}, \mathfrak{I}_{A_{h_{n-1}}}, \mathfrak{R}_{h_{n-1}}, \mathfrak{S}_{m_{n-1}}, \mathfrak{I}_{m_{n-1}}} \right) \\
& \times \frac{\eta}{\mathcal{ABC}(\eta) \lceil \eta \rceil h} \left[\frac{h t_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \times \frac{1}{\mathcal{ABC}(\eta) \lceil \eta \rceil h} \quad (87)
\end{aligned}$$

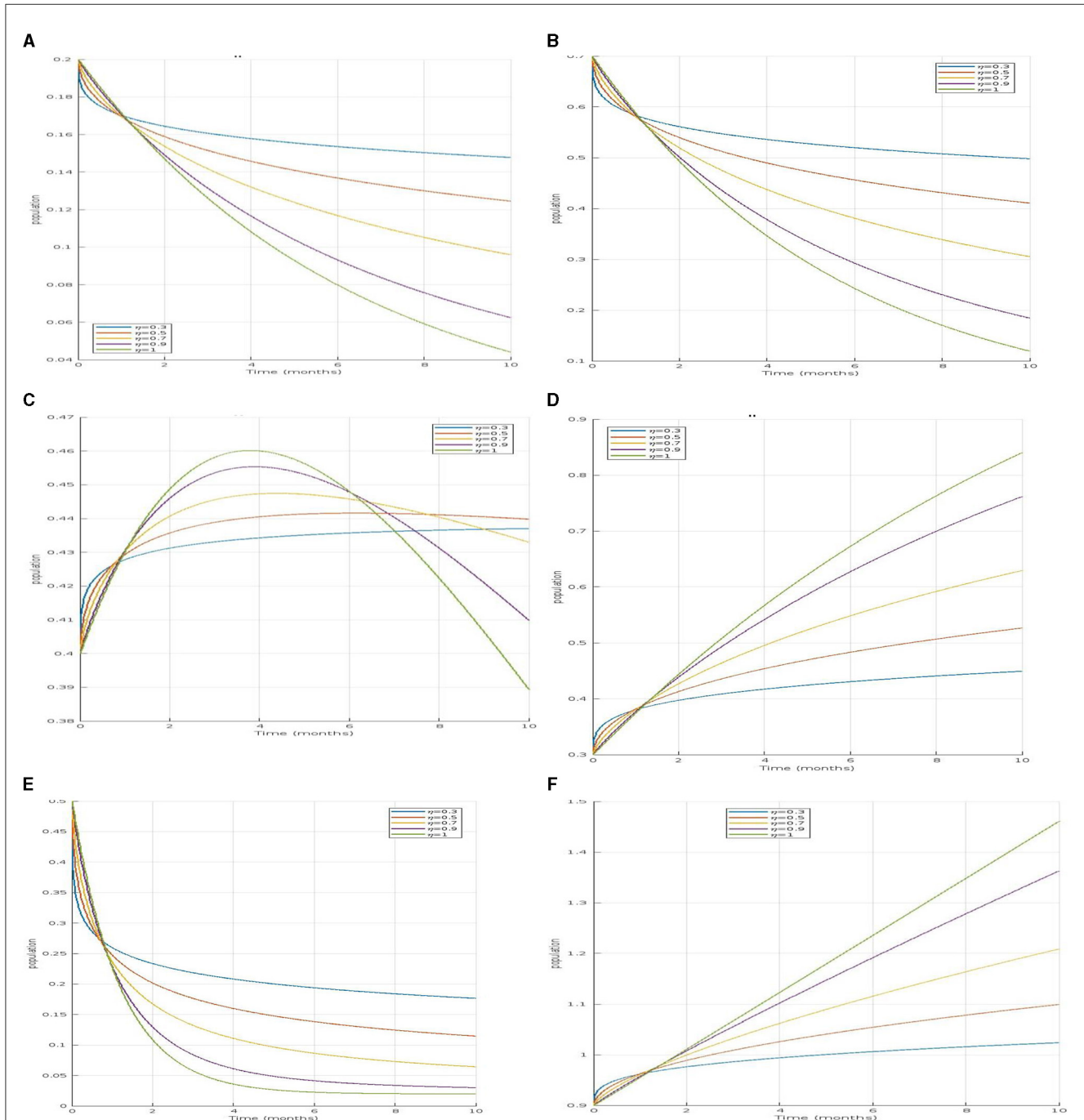


FIGURE 2

(A–D) Illustrate the time series of susceptible, infected, and recovered human populations for fractional orders and respectively. (E, F) Depict the time series of susceptible and infected mosquito populations for the same fractional orders. (A) Simulation of $\mathfrak{S}_h$ for different fractional order. (B) Simulation of $\mathfrak{I}_{S_h}$ for different fractional order. (C) Simulation of $\mathfrak{I}_{S_m}$ for different fractional order. (D) Simulation of $\mathfrak{R}_h$ for different fractional order. (E) Simulation of $\mathfrak{S}_m$ for different fractional order. (F) Simulation of $\mathfrak{I}_{S_m}$ for different fractional order.

$$\begin{aligned}
S_m(t_{n+1}) = S_m(t_n) &+ \frac{1-\eta}{ABC(\eta)} \\
&\left[\mathbb{K}_5 \left(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n} \right) \right. \\
&\left. - \mathbb{K}_5 \left(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}} \right) \right] \\
&+ \frac{\eta \mathbb{K}_5 \left(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n} \right)}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{2ht_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \\
&- \mathbb{K}_5 \left(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}} \right) \\
&\times \frac{\eta}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{ht_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \\
&- \frac{\eta \mathbb{K}_5 \left(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n} \right)}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{ht_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] \\
&- \mathbb{K}_5 \left(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}} \right) \\
&\times \frac{1}{ABC(\eta) \lceil \eta \rceil h} \quad (88)
\end{aligned}$$

$$\begin{aligned}
\mathcal{I}_m(t_{n+1}) = \mathcal{I}_m(t_n) &+ \frac{1-\eta}{ABC(\eta)} \\
&\left[\mathbb{K}_6 \left(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n} \right) \right. \\
&\left. - \mathbb{K}_6 \left(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}} \right) \right] \\
&+ \frac{\eta \mathbb{K}_6 \left(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n} \right)}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{2ht_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \\
&- \mathbb{K}_6 \left(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}} \right) \\
&\times \frac{\eta}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{ht_{n+1}^\eta}{\eta} - \frac{t_{n+1}^{\eta+1}}{\eta+1} \right] \\
&- \frac{\eta \mathbb{K}_6 \left(t_n, S_{h_n}, \mathcal{I}_{S_{h_n}}, \mathcal{I}_{A_{h_n}}, R_{h_n}, S_{m_n}, \mathcal{I}_{m_n} \right)}{ABC(\eta) \lceil \eta \rceil h} \left[\frac{ht_n^\eta}{\eta} - \frac{t_n^{\eta+1}}{\eta+1} \right] \\
&- \mathbb{K}_6 \left(t_{n-1}, S_{h_{n-1}}, \mathcal{I}_{S_{h_{n-1}}}, \mathcal{I}_{A_{h_{n-1}}}, R_{h_{n-1}}, S_{m_{n-1}}, \mathcal{I}_{m_{n-1}} \right) \\
&\times \frac{1}{ABC(\eta) \lceil \eta \rceil h} \quad (89)
\end{aligned}$$

6.1 Outcomes and results of simulation

In this section, we analyzed the dynamics of dengue disease spread using both fractional and non-fractional models across different compartments of the human and mosquito populations. Simulations based on dengue case data collected from Karnataka by NVBDC from August 2023 to May 2024, as detailed in reference (28), revealed notable differences in the behavior of each population compartment under fractional versus non-fractional conditions.

In Figure 2A, the susceptible population decreases with values of $\eta = 0.3, 0.5, 0.7$, and 0.9 , indicating a more realistic and

variable decline due to the complex interactions and memory effects incorporated. In contrast, the non-fractional model shows a constant value of $\eta = 1$, reflecting a simpler and less dynamic decrease. This suggests that the fractional model captures a more nuanced reduction in the susceptible population over time compared to the non-fractional approach. This suggests that the fractional model captures a more nuanced reduction in the susceptible population over time compared to the non-fractional approach.

For infected asymptomatic humans $\mathcal{I}_{A_h}$ and infected symptomatic humans $\mathcal{I}_{S_h}$, both populations will initially increase as the infection spreads but will eventually decrease as individuals recover or move between compartments in the Figures 2B, C.

Similarly, for the recovered human population R_h in Figure 2D, the fractional model reflects a slower recovery rate, acknowledging the variability in recovery times, while the non-fractional model suggests a quicker recovery that might not align with real-world scenarios. In the mosquito populations, the susceptible mosquito population S_m decrease in Figure 2E, more slowly in the fractional model, indicating that mosquitoes remain susceptible for longer periods. The infected mosquito population $\mathcal{I}_m$ in Figure 2F, also rises gradually in the fractional model, unlike the rapid increase seen in the non-fractional model. Overall, the fractional models provide a more realistic representation of the disease dynamics by incorporating memory effects and delays, which better reflect the natural progression and spread of dengue compared to the more immediate transitions observed in non-fractional models.

The comparison clearly shows that fractional-order models provide a more nuanced understanding of how diseases like dengue evolve over time, influencing both human and mosquito populations. The ability of these models to incorporate memory effects allows them to better simulate the slow and cumulative impacts of disease control measures and environmental changes, offering a more realistic depiction of disease dynamics and aiding in the development of more effective intervention strategies.

TABLE 1 Description of parameter values.

Symbols	Baseline values
π_m	0.0071
π_{f_i}	0.057
δ_{f_i}	0.00042
δ_m	0.02
χ_{S_h}	0.00567
χ_{A_h}	0.01691
θ	0.025
α_m	0.5
τ	0.03436
ϱ	0.40
γ	0.0947
χ_m	0.0113

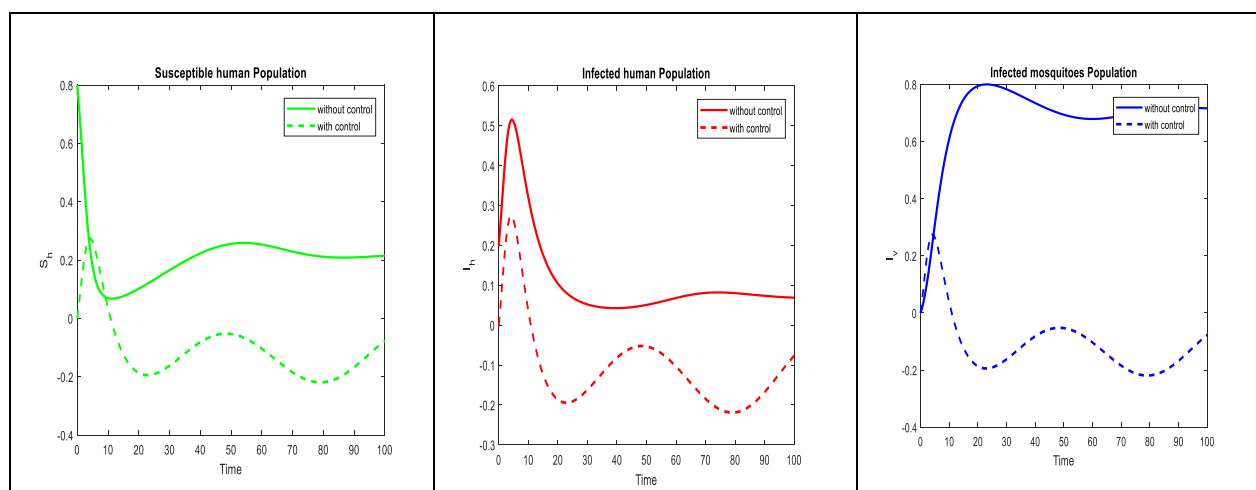


FIGURE 3
The influence of optimal control on the proposed model dynamics is illustrated.

Under this section, simulation results are performed, and values of specification are stated in Table 1.

The disease trajectory can be seen in Figure 3, when optimal control strategies are implemented and their effectiveness in reducing infection rates is highlighted. Comparing the two control strategies, it is evident that self-precautionary measures have a more immediate and direct effect on reducing human infection rates. This suggests that public education campaigns and community involvement can be impactful tools in controlling dengue viremia. Controlling both susceptible and infected mosquito populations is crucial for interrupting the disease transmission cycle. The impact of this strategy on reducing mosquito populations can be observed in the control diagram, illustrating the importance of vector management. Timing is critical for control strategies. The effect effectiveness of vector control may be contingent on seasonal variation in mosquito populations, while self-precaution can be promoted consistently. To maximize their impact, it is crucial to assess the optimal timing and deployment of these strategies.

7 Conclusions

The aim of this study is to explore the effect of dengue viremia on the occurrence of different illnesses. We have presented a comprehensive exploration of ABC fractional order Dengue viremia, a novel mathematical model that incorporates critical factors such as relapse and temporary immunity. After the model is created, the positivity and range of solution is evaluated, and the system survival and originality are verified. The basic reproduction value $\mathcal{R}_0$ is determined by evaluating the equilibrium points. The Rough Hurwitz technique is commonly used to estimate local stability, while lyapunov functions are used to estimate global stability. Specifically, when $\mathcal{R}_0 < 1$ in $\mathcal{C}_f$, it indicates that the disease is unlikely to establish itself. If $\mathcal{R}_0 > 1$ at $\mathcal{C}_p$, it indicates that the disease is likely to continue to spread. Through the utilization of the Adams-Bash forth numerical scheme, we have successfully simulated disease dynamics, achieving a balance

between computational efficiency and accuracy. In addition, we have developed the optimum measures by eradicating the population of mosquitoes and reducing the number of victims. The numerical simulation findings show the behavior of Dengue sickness model affected by different fractional orders, and they can serve as Dengue prevention and control recommendation. The research underscores the importance of mathematical modeling and optimal control techniques in addressing complex infectious disease like Dengue viremia. To develop interventions that reduce and control dengue, it is important to ensure that $\mathcal{R}_0$ is below as a guideline. For future studies, our model can refine control strategies and adapt them to specific regions and epidemics, which is a promising way to treat infectious diseases and safeguard public health on a global scale.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

GMV: Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. MA: Conceptualization, Data curation, Methodology, Resources, Validation, Writing – original draft, Writing – review & editing. LC: Data curation, Formal analysis, Funding acquisition, Investigation, Writing – review & editing. KK: Data curation, Investigation, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We are thankful to the editors and the reviewers for many valuable suggestions to improve this paper.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. World Health Organization. *Epidemiology (WHO)*. (2021). <https://www.who.int/denguecontrol/disease/en/2021>
2. Kautner I, Robinson MJ, Kuhnle U. Dengue virus infection: epidemiology, pathogenesis, clinical presentation, diagnosis, and prevention. *J Pediatr*. (1997) 131:516–24. doi: 10.1016/S0022-3476(97)70054-4
3. Gubler DJ. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends Microbiol*. (2002) 10:100–3. doi: 10.1016/S0966-842X(01)02288-0
4. Esteve L, Vargas C. Analysis of a dengue disease transmission model. *Math Biosci*. (1998) 150:131–51. doi: 10.1016/S0025-5564(98)10003-2
5. Esteve L, Vargas C. A model for dengue disease with variable human population. *J Math Biol*. (1999) 38:220–40. doi: 10.1007/s002850050147
6. Feng Z, Velasco-Hernández JX. Competitive exclusion in a vector-host model for the dengue fever. *J Math Biol*. (1997) 35:523–44. doi: 10.1007/s002850050064
7. Owolabi KM, Atangana A. Analysis and application of new fractional Adams–Bashforth scheme with Caputo–Fabrizio derivative. *Chaos, Solitons Fractals*. (2017) 105:111–9. doi: 10.1016/j.chaos.2017.10.020
8. Pongsumpun P. Mathematical model of dengue disease with the incubation period of virus. *World Acad Sci Eng Technol*. (2008) 44:328–32.
9. Pinho STRD, Ferreira CP, Esteve L, Barreto FR, Morato e Silva VC, and Teixeira MGL. Modelling the dynamics of dengue real epidemics. *Philosoph Trans Royal Soc A: Mathem Phys Eng Sci*. (2010) 368:5679–93. doi: 10.1098/rsta.2010.0278
10. Kongnuy R. Mathematical modeling for dengue transmission with the effect of season. *Int J Biol Med Sci*. (2010) 5:74–8. doi: 10.5281/zenodo.1327917
11. Side S, Noorani SM. A SIR model for spread of dengue fever disease (simulation for South Sulawesi, Indonesia and Selangor, Malaysia). *World J Model Simul*. (2013) 9:96–105. doi: 10.13140/RG.2.1.5042.6721
12. Gakkhar S, Chavda NC. Impact of awareness on the spread of dengue infection in human population. *Appl Math*. (2013) 4:142–7. doi: 10.4236/am.2013.48A020
13. Diethelm K, Ford NJ. Analysis of fractional differential equations. *J Math Anal Appl*. (2002) 265:229–48. doi: 10.1006/jmaa.2000.7194
14. Soewono E, Supriatna AK. A two-dimensional model for the transmission of dengue fever disease. *Bullet Malaysian Mathem Sci Soc*. (2001) 24:1.
15. Jan R, Khan MA, Kumam P, Thounthong P. Modeling the transmission of dengue infection through fractional derivatives. *Chaos, Solitons Fractals*. (2019) 127:189–216. doi: 10.1016/j.chaos.2019.07.002
16. Caputo M, Fabrizio M. A new definition of fractional derivative without singular kernel. *Prog Fract Different Appl*. (2015) 1:73–85. doi: 10.12785/pfda/010201
17. Fatmawati Khan MA, Alfiniyah C, Alzahrani E. Analysis of dengue model with fractal-fractional Caputo–Fabrizio operator. *Adv. Diff. Equat*. (2020) 2020:422. doi: 10.1186/s13662-020-02881-w
18. Boulaaras S, Kan R. Dynamical analysis of the transmission of dengue fever via caputo – fabrizio fractional derivative. *Chaos, Solitons Fractals*. (2022) 2022:100072. doi: 10.1016/j.chaos.2022.100072
19. Sanusi W, Badwi N, Zaki A, Sidjara S, Sari N, Pratama MI, et al. Analysis and simulation of SIRS model for dengue fever transmission in South Sulawesi, Indonesia. *J Appl Mathem*. (2021) 2021:1–8. doi: 10.1155/2021/2918080
20. Ahmad S, Javeed S, Ahmad H, Khushi J, Elagan SK, Khames A. Analysis numerical solution of novel fractional model for dengue. *Results Physics*. (2021) 28:104669. doi: 10.1016/j.rinp.2021.104669
21. Nur W, Rachman H, Abdal NM, Abdy M, Side S. SIR model analysis for transmission of dengue fever disease with climate factors using lyapunov function. *J Phys.* (2018) 1028:012117. doi: 10.1088/1742-6596/1028/1/012117
22. Khan MA. Dengue infection modeling and its optimal control analysis in East Java, Indonesia. *Heliyon*. (2021) 7:1. doi: 10.1016/j.heliyon.2021.e06023
23. Bonyah E, Juga ML, Chukwu CW. A fractional order dengue fever model in the context of protected travelers. *Alexandria Eng J*. (2022) 61:927–36. doi: 10.1016/j.aej.2021.04.070
24. Khan FM, Khan ZU, Lv YP, Yusuf A, Din A. Investigating of fractional order dengue epidemic model with ABC operator. *Results Physics*. (2021) 24:104075. doi: 10.1016/j.rinp.2021.104075
25. Agarwal P, Singh R, ul Rehman. A. Numerical solution of hybrid mathematical model of dengue transmission with relapse and memory via Adam–Bashforth–Moulton predictor-corrector scheme. *Chaos, Solitons and Fractals*. (2021) 143:110564. doi: 10.1016/j.chaos.2020.110564
26. Anggriani N, Supriatna AK, Soewono E. A critical protection level derived from dengue infection mathematical model considering asymptomatic and symptomatic classes. *J Phys*. (2013) 423:012056. doi: 10.1088/1742-6596/423/1/012056
27. Jan R, Khan MA, Gómez-Aguilar JF. Asymptomatic carriers in transmission dynamics of dengue with control interventions. *Optimal Cont Appl Meth*. (2020) 41:430–47. doi: 10.1002/oca.2551
28. Vijayalakshmi GM, Ariyanatchi M. Adams–Bashforth Moulton numerical approach on dengue fractional atangana baleanu caputo model and stability analysis. *Int J Appl Comput Mathem*. (2024) 10:32. doi: 10.1007/s40819-023-01652-x



OPEN ACCESS

EDITED BY

Jiaying Li,
The University of Queensland, Australia

REVIEWED BY

Surjit Singh,
Sister Nivedita University, India
Rita Yi Man Li,
Hong Kong Shue Yan University,
Hong Kong SAR, China

*CORRESPONDENCE

Weiming Hou
✉ hwm100908@163.com

RECEIVED 29 February 2024

ACCEPTED 15 January 2025

PUBLISHED 31 January 2025

CITATION

Hou W (2025) Effect and prediction of long-term weather and pollutant exposure on hemorrhagic fever with renal syndrome: based on statistical models.
Front. Public Health 13:1393763.
doi: 10.3389/fpubh.2025.1393763

COPYRIGHT

© 2025 Hou. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Effect and prediction of long-term weather and pollutant exposure on hemorrhagic fever with renal syndrome: based on statistical models

Weiming Hou*

Department of Medical Engineering, Air Force Medical Center, PLA, Air Force Medical University, Beijing, China

Background: Previous studies have typically explored daily lagged relationships between hemorrhagic fever with renal syndrome (HFRS) and meteorology, with a limited seasonal exploration of monthly lagged relationships, interactions, and the role of pollutants in multiple predictions of hemorrhagic fever.

Methods: Our researchers collected data on HFRS cases from 2005 to 2018 and meteorological and contaminative factors from 2015 to 2018 for the northeastern region. First, we applied the moving epidemic method (MEM) to estimate the epidemic threshold and intensity level. Then, we used a distributed lag non-linear model (DLNM) and a generalized additive model (GAM) with a maximum lag of 6 months to evaluate the lagged and interaction effects of meteorological and pollution factors on the HFRS cases. Multiple machine learning models were then applied after Spearman's rank correlation coefficient analysis was performed to screen for environmental factors in the Northeastern region.

Results: There was a yearly downward trend in the incidence of HFRS in the northeastern region. High prevalence threshold years occurred from 2005 to 2007 and from 2012 to 2014, and the epidemic months were mainly concentrated in November. During the low prevalence threshold period, the main lag factor was low wind direction. In addition, the meteorological lag effect was pronounced during the high prevalence threshold period, where the main lag factors were cold air and hot dew point. Low levels of the AQI and PM₁₀ and high levels of PM_{2.5} showed a dangerous lag effect on the onset of HFRS, while extremely high levels of PM_{2.5} appeared to have a protective effect. High levels of the AQI and PM₁₀, as well as low levels of PM_{2.5}, showed a protective lag effect. The model of PM_{2.5} and the AQI interaction pollution is better. The support vector machine (SVM)-radial algorithm outperformed other algorithms when pollutants are used as predictor variables.

Conclusion: This is the first mathematically based study of the seasonal threshold of HFRS in northeastern China, allowing for accurate estimation of the epidemic level. Our findings suggest that long-term exposure to air pollution is a risk factor for HFRS. Therefore, we should focus on monitoring pollutants in cold conditions and developing HFRS prediction models.

KEYWORDS

hemorrhagic fever with renal syndrome, moving epidemic method, pollutants, time series models, machine learning

1 Introduction

Hemorrhagic fever with renal syndrome (HFRS), also known as epidemic hemorrhagic fever, is a rodent-borne disease caused by various strains of the hantavirus or Seoul virus, characterized by fever, hemorrhage, and acute renal dysfunction (1). As one of the countries most affected by the HFRS epidemic, China has seen a significant decrease in the incidence of HFRS in most regions since 2000. Although preventive measures such as rodent eradication and vaccination have been implemented (2), transient epidemics still occur at certain times and in specific regions.

Early assessments of epidemic thresholds and risk classification focused on influenza and respiratory infections (3, 4), which have proven novel in application and effective for infectious diseases in China. However, there is a lack of relevant studies on HFRS. Earlier studies have suggested that climatic factors may contribute to the incidence of HFRS. According to an epidemiological survey in 2002, rainfall was identified as a predictor of HFRS transmission in the epidemic source ($r = -0.63$) (5). Furthermore, several studies have gradually refined the understanding of the relationship between meteorological factors and HFRS, highlighting varying effects in terms of lag and dose–response relationships. For example, in Nei Menggu province, Wen-Yi Zhang et al. found that rainfall, land temperature, and humidity were associated with HFRS onset at a lag of 3–5 months, after controlling for autocorrelation, seasonality, and long-term trends (6). Recent studies have also shown that wet and warm climatic conditions in the northeastern favor the occurrence and growth of HFRS (7). However, there is limited variability in climatic factors across different epidemic risk classifications. In addition, HFRS may be associated with air pollutants in terms of incidence because it is partly transmitted via the aerosol route. However, although several studies have confirmed the lag and correlation with air pollution in infectious diseases, few studies have been conducted on HFRS (8, 9).

The overall goal of this study was to explore the epidemiological characteristics of HFRS, the graded warning system, the lag and interaction effects of climate and pollutants, and the subsequent development of models for predicting HFRS outbreaks. Our specific objectives were to (a) calculate the epidemic thresholds and assess the risk levels, (b) explore the effects of lags and interactions of meteorological and pollution factors, and (c) construct stratified models for HFRS onset, selecting appropriate models for different populations.

2 Materials and methods

2.1 Setting

Supplementary Figure S1 shows the geographical location of the study area—Heilongjiang, Jilin, and Liaoning provinces. The three provinces are located in the northeastern of China and have medium levels of economic development and population size.

2.2 Data collection

We obtained HFRS case surveillance data from the National Public Health Data Center of China¹ for the study area covering the period from 2005 to 2018. All patients were diagnosed according to the HFRS management criteria issued by the Ministry of Health of the People's Republic of China. We obtained the corresponding daily weather data, including air temperature and dew point temperature, from the China Meteorological Data Sharing Service (data.cma.cn). Pollutant information, including CO, NO₂, and O₃, was originally sourced from the National Oceanic and Atmospheric Administration (NOAA).

2.3 Estimation of the epidemic threshold and intensity level

We used the R language implementation of the moving epidemic method (MEM) (package “mem”), which is available online for free. The method is based on a complex mathematical algorithm that can be summarized in three steps. The first step is the division of the pre-epidemic, epidemic, and post-epidemic periods. In the second step, the pre- and post-epidemic values of the historical seasons are used to calculate the baseline and epidemic thresholds. In the third step, the maximum values of n surveillance indicators during the epidemic period are selected separately to calculate different epidemic intensity thresholds. The unilateral 50%CI upper limit of the geometric mean of the n maximum surveillance indicators during the epidemic period is defined as the medium intensity threshold, the unilateral 90%CI upper limit as the high-intensity threshold, and the unilateral 95%CI upper limit as the very high-intensity threshold.

2.4 The lagging and interaction effect of DLNM and GAM

Distributed lag non-linear models (DLNM) have been widely used to assess the exposure–lag–response relationship between environmental factors and human diseases such as congenital heart disease, hand, foot, and mouth disease, and chronic sinusitis (8, 10–12). The model can be written as follows:

$$\log[E(Y_t)] = \alpha_1 + NS(M, df, lag, df) + \sum NS(X_t) + \sum (X_t) + NS(Time, df) + \beta Month_t$$

To analyze the lag and extreme effects of climate factors, air temperature, dew point temperature, wind direction, and wind speed were considered and applied to the cross-basis functions of a

¹ <https://www.phsciencedata.cn/>

DLNM. Here, Y_t is the number of the HFRS cases in monthly t ; α_1 is the intercept of the entire model; NS is a natural cubic spline that acts as a smooth function of the model; M represents the estimated climate or pollutants variable related to HFRS; and X_t represents other climate and pollutant variables involved in the pathogenesis of HFRS, for which non-linear confounding effects are adjusted. When constructing the meteorological factor model, $\Sigma(X_t)$ does not exist, whereas in the pollution model, meteorological factors are used as confounding factors to construct $\Sigma(X_t)$. The NS was applied to adjust for the monthly confounding effects in the model. *Month* is a binary variable used to control the effect of time, and β represents regression coefficients. The optimal degrees of freedom (df) for the spline function were estimated using the Akaike information criterion for quasi-Poisson (Q-AIC) and minimum partial regression coefficient ($PACF_{min}$) criteria. The NS with 4 df was used for the climate factors, except for wind direction, which used 5 df during the period of low epidemic intensity. For both the high epidemic intensity period and the overall model, the NS with 4 df was applied to the climate and pollutant factors. The lag space was set to 3 df . The NS with 2–3 df /year was applied to the time variable in both pollutant and climate models. The climate model was constructed using the `glm()` function, while the pollution model was constructed using the `gam()` function.

Subsequently, a generalized additive model (GAM) was used to explore the interaction between the pollutants and the prevalence of HFRS. The model formula can be written as follows:

$$\log[E(Y_t)] = \alpha_2 + s(X_1, X_2) + s(X_3) + \Sigma(X_t)$$

α_2 is the intercept; X_t represents the AQI, whereas X_2 and X_3 denote the other two pollutants. $s()$ indicates a penalized spline function. $s(X_1, X_2)$ represents the spline function for the interaction between the parameters X_1 and X_2 . X_1 , X_2 , and X_3 represent 6-month lagged variables. $\Sigma(X_t)$ represents the factors of climate.

2.5 Construction of a prediction model in GPR and SVM

A Gaussian process (GP) can be regarded as an extended function of a multivariate Gaussian distribution, which can be applied to a wide range of variables. In a Gaussian process (GP), it is assumed that any finite set of data follows a multivariate Gaussian distribution. Prior beliefs concerning the relationships between variables are incorporated into these (an infinite number of) multivariate Gaussian distributions to create a model that represents the observational variance. The combination of multiple Gaussian distributions in a GP can effectively model non-linear relationships and is more versatile than traditional parametric models, which depend on fitting a global model. This is because multivariate Gaussians can represent local covariance patterns between individual sites (13).

Support vector machines (SVMs) are a non-probabilistic binary linear regression method. Given a set of training data labeled as belonging to one of two classes, the algorithm maps the data into a space and defines a hyperplane that maximizes the margin between the two classes to separate them. This plane is called the “maximal marginal hyperplane.” An algorithm uses a kernel approach to acquire non-linear mapping to the feature space if linear integration is

impossible. Thus, the hyperplane of the feature space stands for the non-linear boundary of the determination in the input space (14). All model metrics are compared using traditional machine learning metrics such as RMSE, R^2 , and MAE (15–17). A total of 75% of the dataset is used as the training set, while the remaining 25% is used as the test set. All analyses in our study were performed using R software (version 4.1.3).

3 Results

3.1 HFRS surveillance in northeastern China

A total of 59,431 HFRS cases were reported in the three eastern provinces of China from 2005 to 2018, showing a decreasing trend each year (Table 1). This was followed by the main epidemic area in Heilongjiang province, with a total of 28,074 cases until 2018. The incidence of influenza was primarily observed in the individuals aged 15–39 and 40–59 years, accounting for 86.42% of all cases.

Based on Table 1, however, there was a short-term rise in the cases from 2012 to 2014. We also performed a calculation of the prevalence threshold and determined from Supplementary Table S1 that the optimal parameter δ was 7.0 after the calculation of the popular threshold model. As shown in Table 2, the years with a high prevalence threshold were 2005–2007 and 2012–2014, while the years with a low prevalence threshold were 2008–2011 and 2015–2018. Based on the threshold model prediction shown in Table 2 and Figure 1, it was concluded that the epidemic months were primarily concentrated in November.

3.2 Exposure–response relationships and lagging effect for the climate factors

The summary statistics for all HFRS cases and environmental variables in northeastern China are shown in Supplementary Table S2. The Spearman's rank correlation coefficient analysis showed that HFRS was significantly correlated with air temperature ($r = -0.18$, $p < 0.05$), dew point temperature ($r = -0.23$, $p < 0.01$), wind direction ($r = 0.22$, $p < 0.01$), and wind speed ($r = 0.29$, $p < 0.01$) (Supplementary Table S3). As shown in Supplementary Figure S2, these climate factors were associated with high relative risk at the lags above moderate levels, except for air temperature.

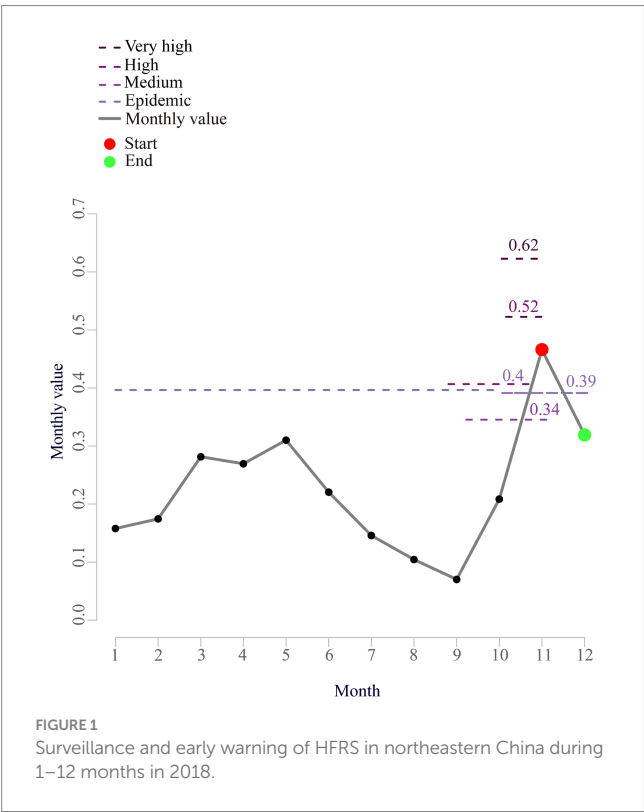
From the dose–response relationship shown in Supplementary Figure S3, air temperature showed mostly a U-shaped relationship with the risk of HFRS, both in general and across the different regions and age groups, while the other factors mostly showed an arch bridge-shaped relationship. In Liaoning province, air temperature, dew point temperature, and wind speed all showed a parabolic decreasing trend in their relationship with HFRS risk. As shown in Supplementary Table S5, the climate lag effect was weak during the low prevalence threshold period, with sensitivity mainly concentrated in the high prevalence areas of Heilongjiang province and the 0–14 years age group, where the main lag factor was low wind direction. As shown in Supplementary Table S6, the meteorological lag effect was higher during the high prevalence threshold period, with sensitivity mainly concentrated in the 0–14 years and 60 years and above age groups, where the main lag factors were cold air and hot dew point. When comparing

TABLE 1 Distribution of the HFRS cases by age groups, region, and season in northeastern China, 2005–2018.

Characteristic		0–14	15–39	40–59	≥60	Total	Population (10 ⁴)	Incidence (10 ⁻² %)
		No. of the HFRS cases (%)						
Year	2005	245(2.26%)	5,148(47.54%)	4,586(42.35%)	850(7.85%)	10,829	10,757	1.01
	2006	138(1.8%)	3,680(47.98%)	3,310(43.16%)	542(7.07%)	7,670	10,917	0.7
	2007	60(1.17%)	2,386(46.64%)	2,268(44.33%)	402(7.86%)	5,116	10,952	0.47
	2008	30(0.85%)	1,519(43.29%)	1,637(46.65%)	323(9.2%)	3,509	10,874	0.32
	2009	31(0.93%)	1,313(39.42%)	1,651(49.56%)	336(10.09%)	3,331	10,907	0.31
	2010	36(1.2%)	1,178(39.21%)	1,432(47.67%)	358(11.92%)	3,004	10,955	0.27
	2011	38(1.17%)	1,162(35.91%)	1,630(50.37%)	406(12.55%)	3,236	10,966	0.3
	2012	54(1.51%)	1,283(35.76%)	1737(48.41%)	514(14.33%)	3,588	10,973	0.33
	2013	52(1.33%)	1,311(33.5%)	1973(50.41%)	578(14.77%)	3,914	10,976	0.36
	2014	45(1.15%)	1,228(31.36%)	1992(50.87%)	651(16.62%)	3,916	10,976	0.36
	2015	28(0.93%)	895(29.7%)	1,538(51.05%)	552(18.32%)	3,013	10,947	0.28
	2016	17(0.66%)	699(27.11%)	1,384(53.69%)	478(18.54%)	2,578	10,910	0.24
	2017	36(1.3%)	743(26.81%)	1,432(51.68%)	560(20.21%)	2,771	10,875	0.25
	2018	32(1.08%)	784(26.52%)	1,478(50%)	662(22.4%)	2,956	10,836	0.27
Region	Heilongjiang	344(1.23%)	11,459(40.82%)	13,018(46.37%)	3,253(11.59%)	28,074	3,819	7.35
	Jilin	176(1.33%)	5,388(40.71%)	6,252(47.24%)	1,418(10.71%)	13,234	2,736	4.84
	Liaoning	322(1.78%)	6,482(35.77%)	8,778(48.44%)	2,541(14.02%)	18,123	4,362	4.15
Season	Spring (March–May)	270(1.69%)	6,660(41.63%)	7,322(45.77%)	1745(10.91%)	15,997		
	Summer (June–August)	126(1.01%)	4,824(38.68%)	6,004(48.14%)	1,518(12.17%)	12,472		
	Autumn (September–November)	230(1.37%)	6,213(36.93%)	8,139(48.38%)	2,241(13.32%)	16,823		
	Winter (December–February)	216(1.53%)	5,632(39.83%)	6,583(46.56%)	1708(12.08%)	14,139		
Total		842(1.42%)	23,329(39.25%)	28,048(47.19%)	7,212(12.14%)	59,431	10,917	5.44

TABLE 2 Characteristics of the peak values in each year used in the model.

Year	Peak (per 10^{-5})	Peak month	Epidemic threshold	Threshold intensity			Level	Series
				Medium	High	Very high		
2005	1.57	11	0.46	0.46	0.64	0.80	Very high	High
2006	1.03	11	0.46	0.46	0.74	0.98	Very high	High
2007	0.81	11	0.40	0.41	0.77	1.02	High	High
2008	0.47	11	0.46	0.46	0.81	1.08	Medium	Baseline
2009	0.44	6	0.47	0.47	0.81	1.07	Baseline	Baseline
2010	0.57	11	0.47	0.47	0.80	1.07	Medium	Baseline
2011	0.43	11	0.46	0.46	0.65	0.83	Baseline	Baseline
2012	0.65	11	0.46	0.46	0.55	0.68	High	High
2013	0.60	11	0.38	0.38	0.50	0.60	High	High
2014	0.55	11	0.38	0.38	0.51	0.62	High	High
2015	0.42	11	0.39	0.39	0.52	0.64	Medium	Baseline
2016	0.41	11	0.40	0.40	0.52	0.62	Medium	Baseline
2017	0.35	11	0.39	0.39	0.53	0.64	Baseline	Baseline
2018	0.47	11	0.40	0.40	0.52	0.62	Medium	Baseline



the climatic lags during the low and high prevalence threshold periods (Supplementary Tables S5, S6), we found that low wind direction and windy conditions showed a dangerous lag effect on HFRS onset ($OR > 0$), while high wind direction and windless conditions showed a protective lag effect ($OR < 0$). In addition, air temperature showed protective effects at both low and high levels, while cold air showed a dangerous effect in the 0–14 years age group during the high prevalence threshold period (OR (95% CI): $3.2e+17(8.4e+08, 1.2e+26)$). Cold dew point had a little

lag effect, while hot dew point showed a protective effect during the low prevalence threshold period. However, this effect was reversed during the high prevalence period.

3.3 Exposure–response relationships and lagging effect for the pollutants

The Spearman’s rank correlation coefficient analysis showed that HFRS was significantly correlated with the AQI ($r = 0.40, p < 0.05$), $PM_{2.5}$ ($r = 0.37, p < 0.05$), and PM_{10} ($r = 0.40, p < 0.01$) (Supplementary Table S4). As shown in Supplementary Figure S4, these factors were associated with high relative risk at the lags above high levels, except for PM_{10} . From the dose–response relationship shown in Figure 2, $PM_{2.5}$ mostly showed an arch bridge-shaped relationship, while the AQI and PM_{10} mostly showed a U-shaped relationship with the risk of HFRS, both in general and across the different regions and age groups. In Jilin province and the 0–14 years age group, the AQI exhibited a parabolic decreasing trend, while $PM_{2.5}$ showed a parabolic increasing trend. As shown in Figure 3, in terms of the total pollution lags, the effects of the low-level pollutants were mainly concentrated in the long-term lag conditions (3–6 months), while the effects of the high-level pollutants were mainly concentrated in the short-term lag conditions (1–2 months). In terms of the lagging trend, $PM_{2.5}$ differed from the other pollution factors. As shown in Table 3, except for high-level PM_{10} , the lag effect of the other pollution factors was more pronounced, and the sensitivity was mainly concentrated in Liaoning province and the age group of 40–59 years. Among these, we found that low levels of the AQI and PM_{10} and high levels of $PM_{2.5}$ showed a dangerous lag effect on the onset of HFRS ($OR > 0$), while extremely high levels of $PM_{2.5}$ (P95) showed a protective effect. In addition, high levels of the AQI and PM_{10} and low levels of $PM_{2.5}$ showed a protective lag effect ($OR < 0$). However, at extremely high levels of the AQI (P95), a dangerous effect was observed.

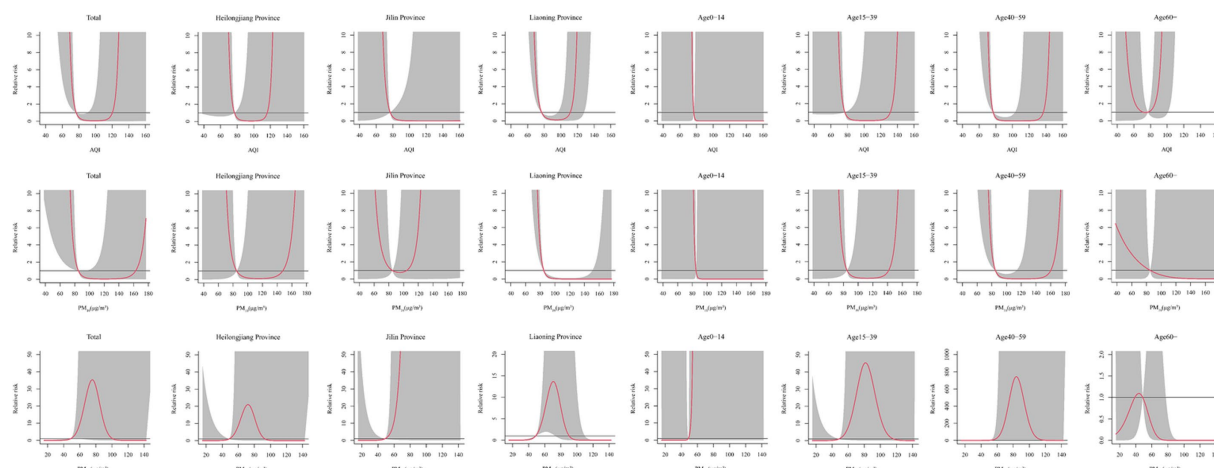


FIGURE 2

Effect of the different pollutants on the incidence of HFRS across the different months for total, regions, and age groups.

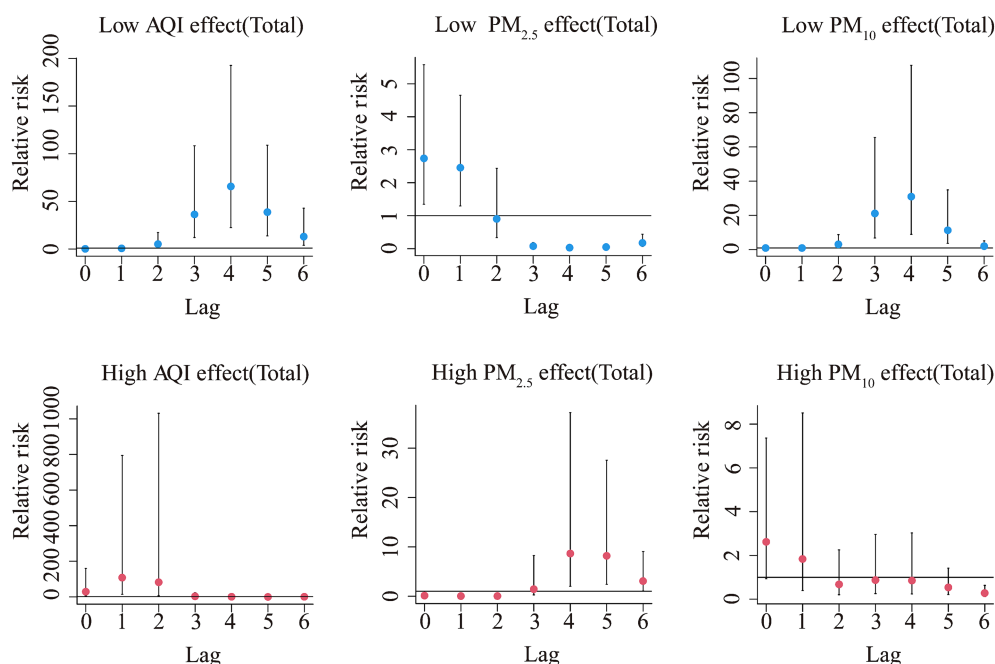


FIGURE 3

Summary of the estimated extreme effects at the 5th and the 95th percentile of the pollutants on the HFRS cases for the total during the different lag months. The median value of each pollutant (AQI: 76.79, $PM_{2.5}$: $48.7 \mu g/m^3$, PM_{10} : $84.89 \mu g/m^3$) serves as a reference level.

3.4 Interaction and comparison of the multiple pollutant models

From [Supplementary Figure S5](#), we can see that the AQI interacted with $PM_{2.5}$ and PM_{10} in relation to HFRS incidence. PM_{10} was weakly positively correlated with the risk of HFRS, while $PM_{2.5}$ showed the opposite relationship. From the interaction effect shown in [Figure 4](#), we found that low AQI combined with high levels of $PM_{2.5}$ and PM_{10} had the greatest impact on HFRS onset. The results from the test in [Supplementary Table S7](#) indicate that the model involving the interaction between $PM_{2.5}$ and the AQI performed better ($R^2 = 44.1\%$).

From [Supplementary Table S8](#) and [Table 4](#), the model fit was best in Liaoning province among the different regions ($R^2 > 70\%$) and in the 15–39 age group. In addition, the GPR model showed the same fit as that of the SVM model. In the GPR model, the prediction results were good, except for the polydot kernel function. In the SVM model, good prediction results were observed with the radial and sigmoid kernel functions. Based on the SVM-radial model for exploring the importance of the variables related to HFRS, the priority order was the pollutant factors (in the order of the AQI, PM_{10} , and $PM_{2.5}$), followed by the climatic factors (in the order of windspeed, dew point temperature, and air temperature).

TABLE 3 The cumulative effects of the extreme pollutant factors on the HFERS cases by region and age group.

Series	Variables	Cumulative effects (95%CI)						
		Low AQI effect	High AQI effect	Low PM _{2.5} effect	High PM _{2.5} effect	Low PM ₁₀ effect	High PM ₁₀ effect	
	Total cases	7.8e+05(78.501, 7.7e+09)	0.051(0.002, 1.323)	1.1e−04(7.2e−08, 0.182)	22.119(0.692, 707.182)	4e+04(7.046, 2.3e+08)	0.070(0.004, 1.248)	
		7.5e+03(15.225, 3.7e+06)	1e+04(0.003, 3.3e+10)	0.004(0.000, 0.384)	0.067(0.000, 1.8e+04)	1.1e+03(3.266, 3.7e+05)	0.364(0.000, 1.3e+04)	
Region	Heilongjiang	4.8e+06(0.808, 2.8e+13)	0.046(0.000, 13.048)	0.000(0.000, 29.414)	17.086(0.041, 7110.833)	4743.176(0.003, 7.3e+09)	0.147(0.001, 16.933)	
		2.4e+04(0.628, 8.9e+08)	7.5e+05(0.000, 8.5e+16)	0.003(0.000, 10.229)	0.002(0.000, 5.1e+06)	257.344(0.018, 3.8e+06)	2.665(0.000, 8.4e+07)	
	Jilin	3.2e+04(0.031, 3.3e+10)	0.029(0.000, 3.540)	0.001(0.000, 44.948)	38.629(0.245, 6.1e+03)	185.119(0.002, 1.7e+07)	0.948(0.020, 44.050)	
		1.1e+03(0.101, 1.3e+07)	0.005(0.000, 1.5e+07)	0.010(0.000, 10.270)	9.2e+03(0.000, 6.5e+11)	24.452(0.011, 5.3e+04)	2.9e+04(0.028, 3e+10)	
	Liaoning	1.6e+05(844.167, 3.1e+07)	0.128(0.021, 0.802)	1.6e−04(2.5e−06, 0.010)	12.214(1.754, 85.054)	4.7e+05(693.362, 3.2e+08)	0.018(0.002, 0.160)	
		2.4e+03(70.763, 8.3e+04)	1.1(221.522, 5.4e+09)	0.005(0.000, 0.064)	0.001(0.000, 0.988)	6478.497(80.564, 5.2e+05)	0.001(0.000, 1.786)	
	Age group	0–14 years	3.9e+17(0.000, 1.4e+57)	0.000(0.000, 5.3e+09)	0.000(0.000, 3.9e+24)	6.2e+05(0.000, 2.3e+23)	4.4e+14(0.000, 1e+61)	0.000(0.000, 1.4e+11)
			0.415(0.000, 7.7e+38)	0.000(0.000, 3.8e+81)	0.000(0.000, 2.7e+15)	1.3e+21(0.000, 1.1e+103)	7.2e+09(0.000, 1.2e+41)	0.000(0.000, 5.9e+52)
15–39 years		4.9e+05(0.775, 3.1e+11)	0.034(0.000, 3.569)	0.001(0.000, 24.237)	18.395(0.130, 2.6e+03)	5.1e+04(0.009, 2.9e+11)	0.099(0.001, 17.719)	
		6.1e+03(0.748, 4.9e+07)	51.177(0.000, 7.9e+10)	0.009(0.000, 7.870)	3.181(0.000, 1.4e+08)	1192.604(0.034, 4.2e+07)	11.978(0.000, 1.2e+09)	
40–59 years		1e+07(336.722, 3.1e+11)	0.012(0.000, 0.460)	6.6e−06(1.7e−09, 0.026)	124.231(2.633, 5860.926)	2.7e+05(107.071, 6.6e+08)	0.045(0.003, 0.614)	
		4.7e+04(45.058, 4.9e+07)	14.229(0.000, 2e+08)	0.001(0.000, 0.104)	29.291(0.000, 2.7e+07)	3812.795(19.754, 7.4e+05)	0.448(0.000, 5.8e+03)	
60 years and above		50.471(4e−03, 5.9e+05)	11.547(0.368, 362.732)	0.223(0.000, 427.895)	0.105(0.003, 4.041)	5.465(0.000, 1.1e+09)	0.376(0.001, 223.061)	
		7.647(0.014, 4.3e+03)	1.8e+15(1.6e+08, 2e+22)	0.576(0.005, 69.983)	5.8e−12(7.6e−18, 4.4e−06)	3.431(0.000, 1.3e+06)	0.008(0.000, 1.8e+08)	

Bold font indicates statistical significance at the 0.05 level.

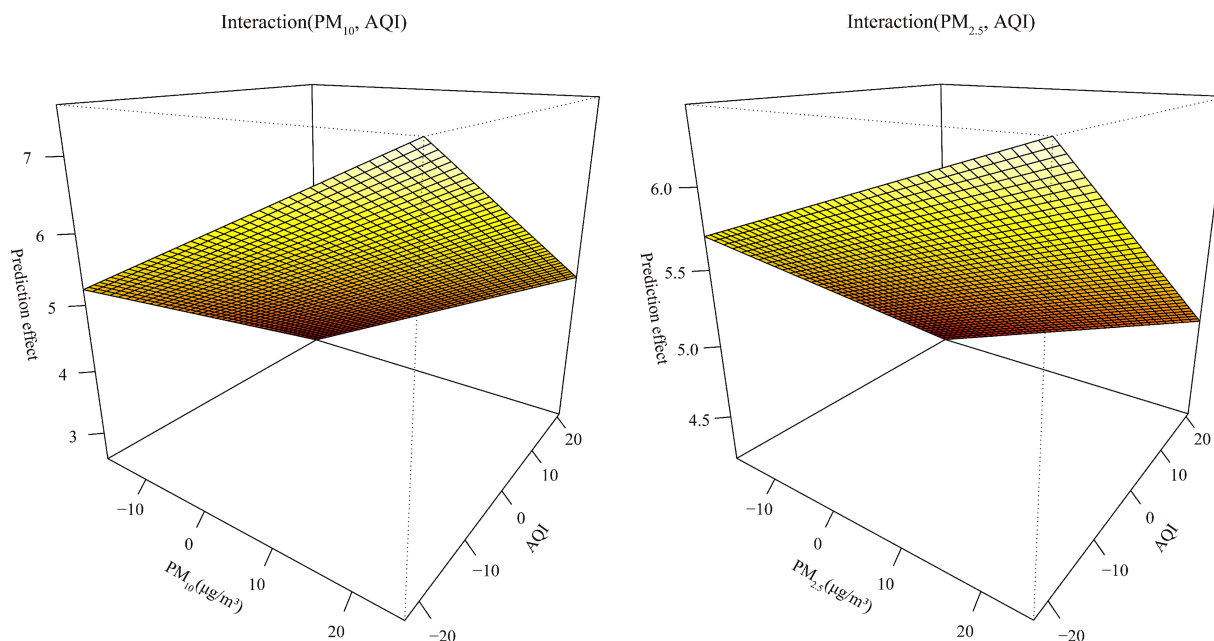


FIGURE 4

The fitting interactions of the association between the pollutants and HFRS cases in northeastern China during 2015–2018 based on the generalized additive model (GAM).

4 Discussion

In the European Centre for Disease Prevention and Control (ECDC), the MEM is a standardized approach for epidemiological classification and early warning of infectious diseases (18). However, the application is limited to diseases with a yearly upward trend, such as influenza and hand, foot, and mouth disease. The better-controlled infectious diseases, such as HFRS, have limited application in epidemic grading. Based on recent global environmental pollution and the short-term annual rise in hemorrhagic fever cases, this study applied the MEM to classify and issue warnings regarding its epidemic status. As the MEM was originally applied to weekly cases, monthly data were used in this study. The selection range for the δ parameter was adjusted from 2.5–5.0 to 4.0–8.0, and the adjustment was made based on the criteria developed after testing with reference to the data.

The prediction of HFRS is widespread both domestically and internationally, with models ranging from ARIMA (19) to Holt–Winters (20) using time series analysis for the univariate prediction of HFRS, achieving good results. However, since HFRS is a natural epidemic, environmental factors greatly influence the transmission of the pathogen and the host. Therefore, this study examined the impact of meteorological factors with lag effects during different periods, classified into high and low epidemic phases using the MEM. This will help future disease control departments implement targeted preventive measures and strategies under different climatic conditions based on the epidemic intensity. We found that Liaoning province exhibited different susceptibility compared to the other regions. This finding is in agreement with the findings of several studies, which indicated that the HFRS epidemic in Liaoning province follows a bimodal pattern (21, 22). During the high epidemic period, HFRS was mainly affected by cold air, with the most susceptible population being in the 0–14-years age group. This finding

is consistent with the findings of studies conducted in other regions of China (23, 24). The main reason may be that cold air increases indoor activity among young, immunocompromised individuals. Since rodents are the primary hosts of the HFRS virus, cold air also raises the likelihood of rodents entering indoor spaces, which significantly exacerbates the incidence of HFRS. Research on the impact of pollutants on diseases dates back to a survey conducted in the United States in 1964 (25). A subsequent study in the U.S. found that long-term exposure to fine particle pollution was linked to death from ischemic heart disease and stroke, highlighting the need for continued improvements in air quality to prevent cardiovascular disease (26). In the field of infectious diseases, air pollution research has primarily focused on respiratory diseases, with little attention given to natural epidemic diseases such as HFRS. A survey in Tianjin found that air pollution control efforts were primarily focused on fulfilling local responsibilities (27), highlighting the impact of air pollution on local health and diseases. Therefore, this study first explored the lagged relationship between air pollution and HFRS, identifying particulate matter (PM) as the main environmental factor. Specifically, low levels of PM₁₀ and high levels of PM_{2.5} were significant at a maximum lag of 6 months, with sensitivity concentrated in the age group of 40–59 years. The reason for this may be that middle-aged individuals are more likely to overlook pollution issues during periods of high air pollution, increasing their time and chances of being exposed to environmental hazards. This, in turn, can significantly enhance exposure to pathogens and host animals. Moreover, for a transmission pathway as unique as aerosols, particulate matter may contribute to the transmission rate, although the exact mechanism remains unknown. This study also conducted a multiple regression analysis of environmental factors to explore the predictive power of machine learning. Although time variables were not included in the prediction model, as in the study by Chao Zhang

TABLE 4 Comparison of the prediction results with the different kernels of the support vector machine (SVM) models.

Model	Series		Parameters	cv.fold	Training set			Test set		
					RMSE	R ²	MAE	RMSE	R ²	MAE
SVM (Linear)		Total cases	cost = 10,gamma = 0.143	10	60.657	0.687	32.258	84.242	0.086	70.695
	Region	Heilongjiang	cost = 10,gamma = 0.143	10	39.826	0.684	18.846	54.348	0.007	41.752
		Jilin	cost = 5,gamma = 0.143	10	10.986	0.712	6.546	16.699	0.022	12.518
		Liaoning	cost = 10,gamma = 0.143	10	14.752	0.756	9.796	30.271	0.335	24.729
	Age group	0–14 years	cost = 0.1,gamma = 0.143	10	2.097	0.193	1.510	1.892	0.085	1.325
		15–39 years	cost = 10,gamma = 0.143	10	14.014	0.790	8.570	17.932	0.386	14.841
		40–59 years	cost = 10,gamma = 0.143	10	30.523	0.696	17.297	44.173	0.105	35.462
		60 years and above	cost = 0.1,gamma = 0.143	10	22.295	0.197	14.206	17.911	0.092	13.482
SVM (Polynomial)		Total cases	degree = 3,cost = 4,gamma = 0.143	10	69.488	0.585	38.329	76.208	0.114	64.232
	Region	Heilongjiang	degree = 3,cost = 1,gamma = 0.143	10	58.073	0.429	32.649	43.421	0.077	30.560
		Jilin	degree = 3,cost = 4,gamma = 0.143	10	11.554	0.680	6.786	16.653	0.020	12.378
		Liaoning	degree = 3,cost = 4,gamma = 0.143	10	17.466	0.655	11.811	30.125	0.354	23.966
	Age group	0–14 years	degree = 3,cost = 0.1,gamma = 0.143	10	2.097	0.193	1.510	1.892	0.085	1.325
		15–39 years	degree = 3,cost = 3,gamma = 0.143	10	17.890	0.656	11.781	15.327	0.449	13.408
		40–59 years	degree = 3,cost = 2,gamma = 0.143	10	39.535	0.499	24.128	38.274	0.182	29.505
		60 years and above	degree = 3,cost = 0.1,gamma = 0.143	10	22.295	0.197	14.206	17.911	0.092	13.482
SVM (Radial)		Total cases	cost = 1,gamma = 0.5	10	66.263	0.705	39.039	75.872	0.103	59.891
	Region	Heilongjiang	cost = 1,gamma = 0.5	10	49.024	0.695	25.009	48.785	0.007	34.382
		Jilin	cost = 1,gamma = 0.5	10	10.968	0.767	6.570	15.864	0.009	11.854
		Liaoning	cost = 1,gamma = 1	10	11.380	0.897	7.870	31.249	0.404	26.547
	Age	0–14 years	cost = 1,gamma = 4	10	1.135	0.800	0.567	2.017	0.000	1.524
		15–39 years	cost = 1,gamma = 1	10	13.255	0.879	7.932	15.210	0.453	13.080
		40–59 years	cost = 1,gamma = 1	10	26.272	0.856	15.701	44.812	0.043	32.347
		60 years and above	cost = 1,gamma = 2	10	12.728	0.808	5.716	19.829	0.004	15.141
SVM (Sigmoid)		Total cases	coef0 = 0.1,gamma = 0.5	10	66.263	0.705	39.039	75.872	0.103	59.891
	Region	Heilongjiang	coef0 = 0.1,gamma = 0.5	10	49.024	0.695	25.009	48.785	0.007	34.382
		Jilin	coef0 = 0.1,gamma = 0.5	10	10.968	0.767	6.570	15.864	0.009	11.854
		Liaoning	coef0 = 0.1,gamma = 1	10	11.380	0.897	7.870	31.249	0.404	26.547
	Age	0–14 years	coef0 = 0.1,gamma = 4	10	1.135	0.800	0.567	2.017	0.000	1.524
		15–39 years	coef0 = 0.1,gamma = 1	10	13.255	0.879	7.932	15.210	0.453	13.080
		40–59 years	coef0 = 0.1,gamma = 1	10	26.272	0.856	15.701	44.812	0.043	32.347
		60 years and above	coef0 = 0.1,gamma = 2	10	12.728	0.808	5.716	19.829	0.004	15.141

et al. (24), the application of different models with varying parameters for hierarchical exploration helped reduce errors from omitted variables and increased confidence in the predictive power. The results showed better prediction accuracy in Liaoning province, which is consistent with previous findings regarding the lagged sensitivity of environmental factors. The SVM model proved to be more stable than the GPR. This also confirmed the advantage of combining the traditional ARIMA time series model with the SVM algorithm to enhance the time series model for HFRS disease prediction, as demonstrated by Chao Zhang et al. (24). However, this study focused more specifically on the northeastern region of China and did not explore the southern regions, which limited the ability to extrapolate the effects of HFRS and natural environmental factors across the country.

5 Conclusion

This is the first mathematically based study on the seasonal threshold of HFRS in northeastern China, enabling accurate estimation of the epidemic levels. Our findings support that long-term exposure to air pollution is a risk factor for HFRS. Therefore, we should focus on monitoring pollutants in cold conditions and developing HFRS prediction models.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: the National Public Health Data Centre of China (<https://www.phsciencedata.cn/>) and the National Oceanic and Atmospheric Administration (NOAA) (<https://www.noaa.gov/>).

Ethics statement

The studies involving humans were approved by China Center for Disease Control and Prevention. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin because Hemorrhagic fever is a Class B infectious disease under China's

Infectious Disease Prevention and Control Law, and each case reported by a medical institution is reported through the direct reporting system of the infectious disease network and requires epidemiological investigation and surveillance testing to further clarify the source of the pathogen and infection.

Author contributions

WH: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2025.1393763/full#supplementary-material>

References

- Bai Y, Xu Z, Lu B, Sun Q, Tang W, Liu X, et al. Effects of climate and rodent factors on hemorrhagic fever with renal syndrome in Chongqing, China, 1997–2008. *PLoS One*. (2015) 10:e0133218. doi: 10.1371/journal.pone.0133218
- Xiao H, Tian HY, Gao LD, Liu HN, Duan LS, Basta N, et al. Animal reservoir, natural and socioeconomic variations and the transmission of hemorrhagic fever with renal syndrome in Chenzhou, China, 2006–2010. *PLoS Negl Trop Dis*. (2014) 8:e2615. doi: 10.1371/journal.pntd.0002615
- Kang M, Tan X, Ye M, Liao Y, Song T, Tang S. The moving epidemic method applied to influenza surveillance in Guangdong, China. *Int J Infect Dis*. (2021) 104:594–600. doi: 10.1016/j.ijid.2021.01.058
- Teeluck M, Samura A. Assessing the appropriateness of the moving epidemic method and WHO average curve method for the syndromic surveillance of acute respiratory infection in Mauritius. *PLoS One*. (2021) 16:e0252703. doi: 10.1371/journal.pone.0252703
- Bi P, Tong S, Donald K, Parton K, Ni J. Climatic, reservoir and occupational variables and the transmission of haemorrhagic fever with renal syndrome in China. *Int J Epidemiol*. (2002) 31:189–93. doi: 10.1093/ije/31.1.189
- Zhang WY, Guo WD, Fang LQ, Li CP, Bi P, Glass GE, et al. Climate variability and hemorrhagic fever with renal syndrome transmission in northeastern China. *Environ Health Perspect*. (2010) 118:915–20. doi: 10.1289/ehp.0901504
- Wang Y, Wei X, Xiao X, Yin W, He J, Ren Z, et al. Climate and socio-economic factors drive the spatio-temporal dynamics of HFRS in northeastern China. *One Health*. (2022) 15:100466. doi: 10.1016/j.onehlt.2022.100466
- Luo C, Qian J, Liu Y, Lv Q, Ma Y, Yin F. Long-term air pollution levels modify the relationships between short-term exposure to meteorological factors, air pollution and the incidence of hand, foot and mouth disease in children: a DLNM-based multicity time series study in Sichuan Province, China. *BMC Public Health*. (2022) 22:1484. doi: 10.1186/s12889-022-13890-7
- Yang L, Yang J, Liu M, Sun X, Li T, Guo Y, et al. Nonlinear effect of air pollution on adult pneumonia hospital visits in the coastal city of Qingdao, China: a time-series analysis. *Environ Res*. (2022) 209:112754. doi: 10.1016/j.envres.2022.112754
- du R, Jiao W, Ma J, Zhou Q, Liang ZS, Sun S, et al. Association between ambient temperature and chronic rhinosinusitis. *Int Forum Allergy Rhinol*. (2023) 13:1906–14. doi: 10.1002/alr.23152

11. Zhang Q, Sun S, Sui X, Ding L, Yang M, Li C, et al. Associations between weekly air pollution exposure and congenital heart disease. *Sci Total Environ.* (2021) 757:143821. doi: 10.1016/j.scitotenv.2020.143821
12. Chen Y, Hou W, Dong J. Time series analyses based on the joint lagged effect analysis of pollution and meteorological factors of hemorrhagic fever with renal syndrome and the construction of prediction model. *PLoS Negl Trop Dis.* (2023) 17:e0010806. doi: 10.1371/journal.pntd.0010806
13. Cole JH, Poudel RPK, Tsagkrasoulis D, Caan MWA, Steves C, Spector TD, et al. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage.* (2017) 163:115–24. doi: 10.1016/j.neuroimage.2017.07.059
14. Lynch CM, Abdollahi B, Fuqua JD, de Carlo AR, Bartholomai JA, Balgemann RN, et al. Prediction of lung cancer patient survival via supervised machine learning classification techniques. *Int J Med Inform.* (2017) 108:1–8. doi: 10.1016/j.ijmedinf.2017.09.013
15. Li RYM, Song L, Li B, Crabbe MJC, Yue X-G. Predicting carpark prices indices in Hong Kong using AutoML. *Comput Model Eng Sci.* (2022) 134:2247–82. doi: 10.32604/cmcs.2022.020930
16. Lei T, Li RYM, Jotikastira N, Fu H, Wang C. Prediction for the inventory management chaotic complexity system based on the deep neural network algorithm. *Complexity.* (2023) 2023:1–11. doi: 10.1155/2023/9369888
17. Li RYM, Fong S, Chong KWS. Forecasting the REITs and stock indices: group method of data handling neural network approach. *Pac Rim Prop Res J.* (2017) 23:123–60. doi: 10.1080/14445921.2016.1225149
18. Green H, Charlett A, Moran-Gilad J, Fleming D, Durnall H, Thomas D, et al. Harmonizing influenza primary-care surveillance in the United Kingdom: piloting two methods to assess the timing and intensity of the seasonal epidemic across several general practice-based surveillance schemes. *Epidemiol Infect.* (2015) 143:1–12. doi: 10.1017/S0950268814001757
19. Qi C, Zhang D, Zhu Y, Liu L, Li C, Wang Z, et al. SARFIMA model prediction for infectious diseases: application to hemorrhagic fever with renal syndrome and comparing with SARIMA. *BMC Med Res Methodol.* (2020) 20:243. doi: 10.1186/s12874-020-01130-8
20. Ye GH, Alim M, Guan P, Huang DS, Zhou BS, Wu W. Improving the precision of modeling the incidence of hemorrhagic fever with renal syndrome in mainland China with an ensemble machine learning approach. *PLoS One.* (2021) 16:e0248597. doi: 10.1371/journal.pone.0248597
21. Lin H, Liu Q, Guo J, Zhang J, Wang J, Chen H. Analysis of the geographic distribution of HFRS in Liaoning Province between 2000 and 2005. *BMC Public Health.* (2007) 7:207. doi: 10.1186/1471-2458-7-207
22. Lv CX, An SY, Qiao BJ, Wu W. Time series analysis of hemorrhagic fever with renal syndrome in mainland China by using an XGBoost forecasting model. *BMC Infect Dis.* (2021) 21:839. doi: 10.1186/s12879-021-06503-y
23. Xu Q, Li R, Rutherford S, Luo C, Liu Y, Wang Z, et al. Using a distributed lag non-linear model to identify impact of temperature variables on haemorrhagic fever with renal syndrome in Shandong Province. *Epidemiol Infect.* (2018) 146:1671–9. doi: 10.1017/S095026881800184X
24. Zhang C, Fu X, Zhang Y, Nie C, Li L, Cao H, et al. Epidemiological and time series analysis of haemorrhagic fever with renal syndrome from 2004 to 2017 in Shandong Province, China. *Sci Rep.* (2019) 9:14644. doi: 10.1038/s41598-019-50878-7
25. Ferris BG Jr, Frank NR. Air pollution and disease. *Anesthesiology.* (1964) 25:470–8. doi: 10.1097/00000542-196407000-00010
26. Hayes RB, Lim C, Zhang Y, Cromar K, Shao Y, Reynolds HR, et al. PM2.5 air pollution and cause-specific cardiovascular disease mortality. *Int J Epidemiol.* (2020) 49:25–35. doi: 10.1093/ije/dyz114
27. Shan M, Wang Y, Lu Y, Liang C, Wang T, Li L, et al. Uncovering PM2.5 transport trajectories and sources at district within city scale. *J Clean Prod.* (2023) 423:138608. doi: 10.1016/j.jclepro.2023.138608

Frontiers in Public Health

Explores and addresses today's fast-moving healthcare challenges

One of the most cited journals in its field, which promotes discussion around inter-sectoral public health challenges spanning health promotion to climate change, transportation, environmental change and even species diversity.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact



Frontiers in Public Health

