frontiers | Research Topics

Collection of COVID-19 induced biases in medical research

Edited by
Reza Rastmanesh, Vasiliki Epameinondas Georgakopoulou and Martin Wolkewitz

Published in
Frontiers in Medicine

ISSN 1664-8714
ISBN 978-2-8325-6402-8
DOI 10.3389/978-2-8325-6402-8

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

Collection of COVID-19 induced biases in medical research

Topic editors

Reza Rastmanesh — American Physical Society, United States

Vasiliki Epameinondas Georgakopoulou — Laiko General Hospital of Athens, Greece

Martin Wolkewitz — University Medical Center Freiburg, Germany

Citation

Rastmanesh, R., Georgakopoulou, V. E., Wolkewitz, M., eds. (2025). *Collection of COVID-19 induced biases in medical research*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6402-8

Table of contents

frontiers | Frontiers in Medicine

TYPE Editorial
PUBLISHED 19 May 2025
DOI 10.3389/fmed.2025.1605703

OPEN ACCESS

EDITED AND REVIEWED BY
Beatriz S. Lima,
Research Institute for Medicines (iMed.ULisboa), Portugal

*CORRESPONDENCE
Reza Rastmanesh
✉ r.rastmanesh@gmail.com

RECEIVED 03 April 2025
ACCEPTED 28 April 2025
PUBLISHED 19 May 2025

CITATION
Rastmanesh R, Georgakopoulou VE and Wolkewitz M (2025) Editorial: Collection of COVID-19 induced biases in medical research. *Front. Med.* 12:1605703. doi: 10.3389/fmed.2025.1605703

Editorial: Collection of COVID-19 induced biases in medical research

Reza Rastmanesh[1]*, Vasiliki Epameinondas Georgakopoulou[2] and Martin Wolkewitz[3]

[1]Department of Biomedical Sciences, American Physical Society, College Park, MD, United States, [2]Department of Pathophysiology, Laiko General Hospital, National and Kapodistrian University of Athens, Athens, Greece, [3]Institute of Medical Biometry and Statistics, Division Methods in Clinical Epidemiology, Faculty of Medicine and Medical Center - University of Freiburg, Freiburg, Germany

KEYWORDS

COVID-19, bias, methodology, vaccination, face mask, study design

Editorial on the Research Topic
Collection of COVID-19 induced biases in medical research

The COVID-19 pandemic has significantly impacted various aspects of social interactions, individual behaviors, and healthcare practices. It has also altered many physiological responses, leading to the expectation that numerous medical studies may be affected by hidden biases related to the pandemic, either directly or indirectly linked to the use of face masks or the virus itself. For example, wearing face masks has been shown to create substantial biases in fields such as endocrinology, ophthalmology (especially concerning dry eye and ocular conditions), sleep research, cognitive biases (including studies on emotion recognition), and gender differences, among others. It is likely that many of these biases remain unrecognized in other medical fields.

This Research Topic encompasses submissions that address previously unreported biases arising from the COVID-19 pandemic and/or the use of face masks. Our objective was to compile manuscripts that identify novel biases, thereby facilitating a more accurate and impartial interpretation of clinical findings, methodological advancements, registered clinical trials, cohort studies, and comparative studies conducted both before and after the pandemic.

It is not surprising to state that the COVID-19 lockdown significantly affected many healthcare systems. Turati et al. demonstrated that the orthopedic and trauma departments in Italy encountered major difficulties, leading to a notable decrease in all services, such as emergency consultations, outpatient visits, and surgical procedures. This situation provides important lessons for the future, but tackling a future pandemic will necessitate collaboration across multiple disciplines.

This subsequent finding may be referred to as a "policy bias" indirectly imposed by the COVID-19 pandemic.

To improve healthcare comprehensively, public health initiatives have shifted from focusing solely on pandemic response to gaining a deeper understanding of the aftermath, which includes mental health challenges arising from societal restrictions and safety measures. The lasting impacts of the COVID-19 pandemic depend on the health system's ability to foster healthier communities, enhance individual resilience, and reduce environmental stressors moving forward. In this context, the pandemic's consequences have been examined in connection with the public health crisis and the physical isolation caused by the SARS-CoV-2 virus.

Marsico and Russo state that in addition to a person's willingness to embrace positive change, the pandemic has led to emotional instability, created lasting memories, and caused social upheaval in both private and public spheres. These groups, which are often more socially disadvantaged than others, may undermine their own confrontational behavior and be less capable of demonstrating collective resilience over time.

This confrontational behavior could inadvertently exacerbate systemic biases in medical research and policy.

A key concern for the scientific community is the rate of retractions that occurred during the COVID-19 pandemic.

Furuse showed that retraction rates generally increased until at least 2019, with the highest rates observed in the category of "Neoplasms". During the COVID-19 pandemic, there was a significant surge in publications related to "Infections" and "Respiratory Tract Diseases"; however, the retraction rates for these categories and for COVID-19-related papers were not particularly high compared to other diseases. Most disease categories showed a stronger association with retractions in China, while for COVID-19 papers, other countries exhibited higher retraction rates than China. In recent years, papers that have been retracted are less likely to appear in high-impact journals.

This phenomenon can be classified as publication bias.

Numerous research efforts have sought to assess the severity and patterns of COVID-19. Initially, during the pandemic, the complex trajectories of patients were described only in general terms, and many studies were significantly impacted by biases related to time, selection, and competing risks.

Lucke et al. demonstrated that multi-state models help mitigate these biases by simultaneously analyzing various clinical outcomes while considering their time-related nature, including ongoing cases, and accounting for competing events. A group of researchers utilized a publicly accessible dataset from COVID-19 first wave to illustrate the advantages of employing multi-state methodology in the analysis of hospital data.

They evaluated the results of the data analysis conducted with multi-state models against the results obtained when different types of bias were overlooked. Additionally, Cox regression was employed to analyze the transitions between states in the multi-state model, enabling a comparison of how covariates affect transition rates between the two states. Finally, they computed the anticipated lengths of stay and state probabilities derived from the multi-state model and represented this information through stacked probability plots. Utilizing multi-state models on real-time data enables quick identification of changes in disease progression when new variants emerge. This information is crucial for guiding medical and political leaders, as well as the general public.

Another three common methodological biases need to be addressed: competing risks, immortal-time bias, and confounding bias in real-world observational studies that assess treatment effectiveness. A team of researchers utilized a specific observational data example involving COVID-19 patients to evaluate the effects of these biases and suggest possible solutions. Indeed, neglecting competing risks, immortal-time bias, and confounding bias can distort treatment effect estimates.

According to Martinuka et al., using the basic Kaplan-Meier method produced the most inaccurate results, leading to inflated probabilities for the primary outcome in studies involving COVID-19 hospital data. This inflation could misguide clinical decisions. Therefore, it is essential to tackle both immortal-time bias and confounding bias when evaluating treatment effectiveness. The trial emulation framework presents a possible approach to mitigate all three of these methodological biases.

This was only a part of the issue. Tackling bias in how SARS-CoV-2 reinfection is defined is another key challenge. Traditionally, reinfection is identified as a positive test result that happens at least 90 days after a prior infection has been diagnosed. However, this lengthy timeframe might result in an undercount of reinfection cases. Chemaitelly et al. explored the possibility of using a different, shorter timeframe to define reinfection. The 40-day time window was appropriate for defining reinfection, irrespective of whether it was the first, second, third, or fourth occurrence. The sensitivity analysis, confined to high testers exclusively, replicated similar patterns and results. These findings will significantly impact the issue of underestimation.

The comparison between immunity gained from previous natural infections and that obtained through vaccination against SARS-CoV-2 is a significant topic. In this context, we required a statistical clarification to prevent any misinterpretation. To achieve this goal, we need access to real-world data from a large population. Weber et al. analyzed data from over 52,000 individuals. The group that was infected tended to be younger, had a higher proportion of men, and exhibited lower morbidity compared to the vaccinated group. After the initial 90 days, these differences became more pronounced. The analysis conducted during the second 90 days revealed variations in results based on different analytical methods and age groups. There were also age-related differences in mortality rates. When considering the outcome of SARS-CoV-2 infection, the impact of vaccination compared to infection differs by age, showing a disadvantage for vaccinated individuals in the younger demographic, while no significant difference was observed in older adults. It is important to analyze two observation periods: the first and second 90-day spans after infection or vaccination. Furthermore, it is necessary to implement methods to correct any imbalances. This strategy facilitates equitable comparisons, enables more thorough conclusions, and helps avoid biased interpretations. It is crucial not to mix up these results with the 40-day time frame that was proposed as suitable for identifying reinfection (Chemaitelly et al.).

As for the observational studies on the effectiveness of COVID-19 vaccines, these designs have provided crucial real-world insights that have influenced global public health policies. These studies, which mainly utilize existing data sources, have been crucial for evaluating vaccine effectiveness across various populations and for creating effective vaccination strategies. Cohort designs are commonly used in this research. The swift rollout of vaccination campaigns during the pandemic led to variations in vaccination rates influenced by socio-demographic factors, public policies, perceived risks, health-promoting behaviors, and overall health status. This may have resulted in biases such as healthy user bias, healthy vaccine effect, frailty bias, differential depletion of susceptibility bias, and confounding by indication. The pressure to publish findings rapidly may have exacerbated these biases or led to their oversight, thereby affecting the reliability of the results. The

extent of these biases can vary greatly depending on the context, data sources, and analytical techniques used, and they are likely to be more pronounced in low- and middle-income countries due to weaker data infrastructure. It is crucial to address and reduce these biases to obtain accurate estimates of vaccine effectiveness, inform public health strategies, and maintain public confidence in vaccination efforts. Agampodi et al. in their brilliant article state that clear communication about these biases and a commitment to improving the design of future observational studies are vital.

Another type of neglected bias that may obscure data analysis during the COVID-19 pandemic arises from treatment-induced differences. Prosty et al. demonstrated that during the pandemic, many patients received concomitant corticosteroids, which are known to broadly suppress inflammatory cytokines, including those associated with type II inflammation. This may have obscured any differences induced by omalizumab and biased the results toward the null hypothesis, while others did not receive corticosteroid therapy. Results from one of the articles submitted to our Research Topic suggested that the potential benefits of omalizumab in COVID-19 may be mediated independently of the modulation of the measured serum biomarkers. This finding, in itself, impacts the interpretation of many clinical trials conducted during the pandemic.

Given the numerous issues addressed in this brief editorial, the significance of interdisciplinary collaboration in mitigating biases exacerbated by the pandemic must be emphasized.

Author contributions

RR: Funding acquisition, Writing – review & editing, Writing – original draft, Formal analysis, Project administration, Resources, Visualization, Software, Supervision, Conceptualization, Methodology, Validation, Investigation, Data curation. VG: Writing – original draft. MW: Writing – original draft.

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

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 10 January 2024
DOI 10.3389/fmed.2023.1288014

OPEN ACCESS

EDITED BY
Martin Wolkewitz,
University Medical Center Freiburg, Germany

REVIEWED BY
Alessandro Miani,
University of Bristol, United Kingdom
Tommaso Filippini,
University of Modena and Reggio Emilia, Italy

*CORRESPONDENCE
Yuki Furuse
furusey.nagasaki@gmail.com

RECEIVED 03 September 2023
ACCEPTED 15 December 2023
PUBLISHED 10 January 2024

CITATION
Furuse Y (2024) Characteristics of retracted research papers before and during the COVID-19 pandemic.
Front. Med. 10:1288014.
doi: 10.3389/fmed.2023.1288014

Characteristics of retracted research papers before and during the COVID-19 pandemic

Yuki Furuse*

Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

Objectives: During the COVID-19 pandemic, a large number of research papers were published, and some of them were retracted. The present study aims to reveal the characteristics of retracted papers before and during the pandemic.

Methods: The study investigated 24,542,394 publications from 1999 to 2022 and analyzed the profiles of retracted papers from the perspectives of year, disease category, country, and journal.

Results: Retraction rates were generally increasing at least until 2019, and were the highest for "Neoplasms." The number of publications for "Infections" and "Respiratory Tract Diseases" dramatically rose during the COVID-19 pandemic; however, the retraction rates in the two categories or of COVID-19-related papers were not especially high compared to other diseases. The association with retraction was strongest for China in most disease categories, whereas for COVID-19 papers, other countries showed higher retraction rates than China. In recent years, retracted papers have become less likely to be published in high-impact journals.

Conclusion: The COVID-19 pandemic does not seem to affect the retractions of research papers much. We should keep monitoring retractions and analyze the effects of pandemics for better science.

KEYWORDS

COVID-19, SARS-CoV-2, publication, retraction, research integrity

1 Introduction

Medical research plays an important role in advancing knowledge and in improving health. The hypotheses, findings, and reviews of biomedical research are widely shared among research communities and the general public through publications. Yet, "to err is human" (1). Some publications were retracted due to honest errors or misconduct.

During the COVID-19 pandemic, research communities have been eager to study this disease in order to mitigate its impact on public health. Thanks to these efforts, we successfully identified the causative agent, SARS-CoV-2, characterized its clinical and epidemiological features, and developed remedies and vaccines. However, unfortunately, some studies were later retracted, causing chaos. Such eventually retracted studies might result in mistakes in patient care that could cause possible harm. For example, several studies have reported the effectiveness of ivermectin in treating COVID-19, but were eventually retracted (2). These retractions affected the results of a meta-analysis that initially supported the use of the medicine. After excluding data from retracted and

questionable studies, the effect did not reach statistical significance (3). In actuality, the administration of the drug aiming to prevent or treat COVID-19 has caused severe adverse events in some people (4, 5).

Researchers have rushed to conduct studies on COVID-19 and publish their results. There might have been pressure to publish results quickly to combat the pandemic, possibly leading to careless or poor research that was later revealed to be incorrect. Furthermore, a flood of submissions and acceleration of the review process during the pandemic might affect the quality of research papers regarding not only COVID-19 but also other diseases (6). As such, the pandemic may have inflated the number of retractions and also made an impact on various aspects of retractions (7). This study analyzed the profiles of retracted research papers and compared them before and during the COVID-19 pandemic. Understanding them would help conduct better research activities, minimizing the chance of future retractions even under crisis situations such as a pandemic.

2 Methods

2.1 Publication records

The number of publications by retraction status, publication year, disease category, authors' affiliated country, and journal was obtained via PubMed, accessed between 1 June 2023 and 15 October 2023.[1] Data were retrieved using the Rentrez package in R (8).

The disease category and microbial etiology of research papers were searched using MeSH terms (9). Information on the Journal Impact Factor 2021 (published in 2022), Journal Impact Factor 2018 (published in 2019), and Scimago Journal Rank (SJR) indicator 2022 was used to identify high-impact and "good" journals. All query terms used in the present study can be found in Supplementary Table S1.

When a research paper could be assigned to multiple disease categories or authors' affiliations, that paper was counted multiple times in different classifications. Thereby, affiliations were analyzed for all coauthors, not exclusively for the first or corresponding author.

2.2 Statistical analysis

The numbers of total publications, retracted papers, and retraction rates were counted and calculated by publication year and disease category. Disease categories were ranked based on the retraction rates by 3-year-periods from 1999 to 2022.

The association of affiliated countries with retracted papers was evaluated using odds ratio by the 3-year-periods. The odds ratio was calculated as follows: ("number of retracted papers from a country of interest"/"number of retracted papers from other countries")/("number of non-retracted publications from a country of interest"/"number of non-retracted publications from other countries").

The increasing trend of retraction rates was tested using the Kendall rank correlation coefficient. A multivariable regression analysis, assuming a quasi-Poisson distribution, was performed to analyze the associations of publication year, disease category, and affiliated country with retraction counts, using the number of total publications as an offset. Statistical significance was set at $p < 0.05$.

3 Results

3.1 Disease category of retracted papers

Between 1999 and 2022, 24,542,394 publications were indexed in PubMed; of these, 14,717 were retracted (6.0 in 10,000 papers). Among the 18 disease categories, "Neoplasms" had the highest number of retracted papers (no. = 3,234), followed by "Digestive System Diseases" (1,187), "Urogenital Diseases" (959), "Nervous System Diseases" (866), and "Cardiovascular Diseases" (834) (Figure 1; Supplementary Figures S1–S3). In terms of retraction rates, which were calculated as the number of retracted papers per total publications, the highest rate was observed for "Neoplasms" (13.1 in 10,000 papers), followed by "Digestive System Diseases" (10.4), "Endocrine System Diseases" (9.3), "Urogenital Diseases" (7.1), and "Stomatognathic Diseases" (7.0).

Both the total number of publications and the number of retracted papers increased from 1999 to 2019. During this period, statistically significant increases in the retraction rate were found for 16/18 disease categories. Although the total number of publications continued to increase after 2019, the retraction counts and rates declined (Figures 1B,C; Supplementary Figures S1–S3). The yearly number of retractions and retraction rates were analyzed based on not their retraction date but the publication date. The decrease of retractions after 2019 must be because this study was conducted in 2023; it still takes time to realize concerns and investigate them for papers published in 2020 or later. Therefore, comparing retraction rates in recent years to past years is difficult.

To address this issue, we compared the retraction rates among different disease categories using contemporary data. Figure 2A illustrates the trend of the contemporary ranking of retraction rates by disease category over 24 years by 3-year-periods. "Neoplasms" ranked the top in 6 out of 8 periods. "Respiratory Tract Diseases" was situated in the top 4 or 5 in the rank in 2011–2019, but it dropped to 10th in 2020–2022. "Infections" kept low positions, 13th or below, throughout the study periods, even during the COVID-19 pandemic.

3.2 Infectious diseases in retracted papers

I then investigated retracted papers related to the top five intensively studied infectious diseases identified in a previous study (10): hepatitis C, HIV infection, influenza, malaria, and tuberculosis, along with five trendy infectious diseases over the past 20 years: COVID-19, Ebola virus disease, poliomyelitis, SARS, and Zika fever. Those diseases were declared a Public Health Emergency of International Concern by the World Health Organization (11). The number of publications on such infectious diseases rose after big outbreaks, that is, around 2003 for SARS, 2009 for influenza, 2014 for Ebola, 2016 for Zika, and 2020 for COVID-19 (Figures 3A,B; Supplementary Figure S4).

1 https://pubmed.ncbi.nlm.nih.gov/

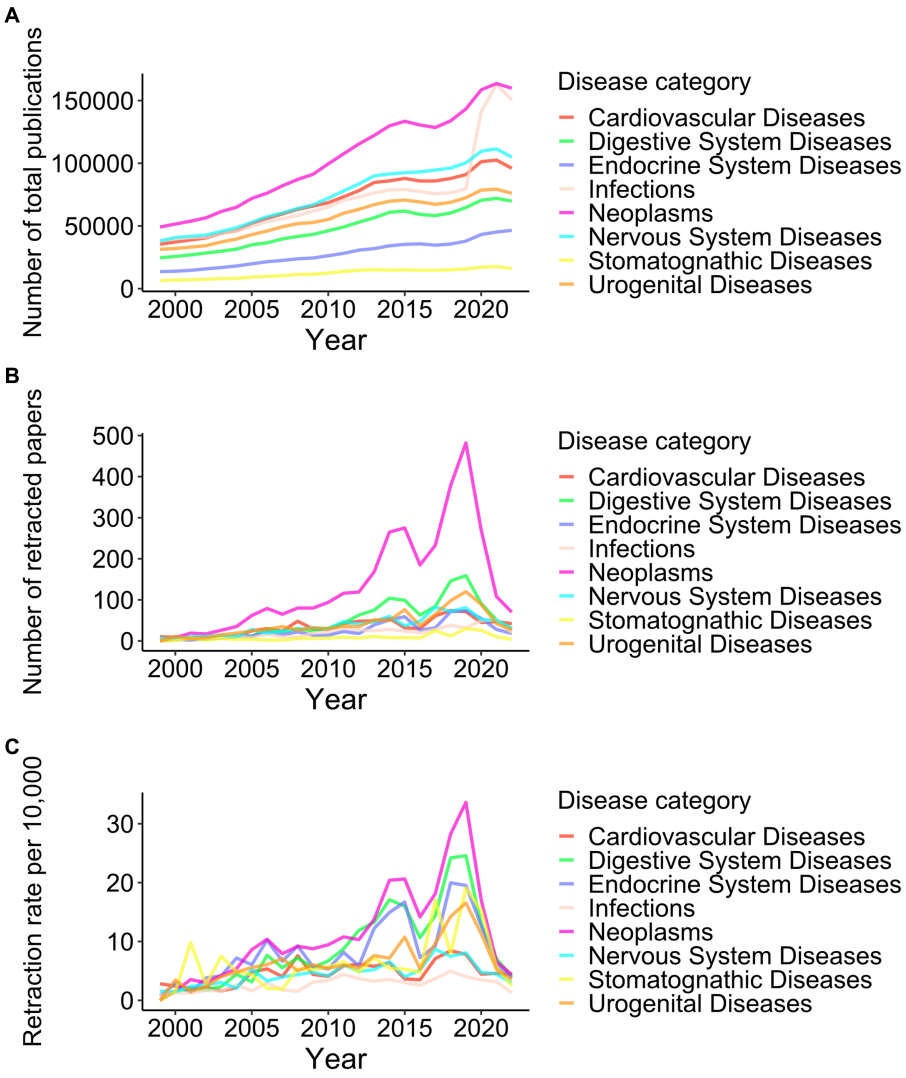

FIGURE 1
Trend of retractions among eight disease categories from 1999 to 2022. Yearly numbers of total publications **(A)**, retracted papers **(B)**, and retraction rates per 10,000 publications **(C)** in eight disease categories, which are among the top 5 diseases in publication counts or retraction rates, are shown. The results of all 18 disease categories can be found in Supplementary Figure S1.

From 2020 to 2022, 40 papers regarding COVID-19 were retracted (Figure 3C). Ten retracted papers on SARS in the corresponding period must be associated with studies comparing COVID-19 with SARS. Disease outbreaks did not lead to an increase in the retractions of research papers about influenza, Ebola, or Zika (Figures 2B, 3C,D). The peaks of SARS-related retractions in 2004 and 2009 were formed by one retracted paper each year. During the 2020–2022 COVID-19 pandemic, the retraction rate of COVID-19 papers was lower than that of publications on tuberculosis, hepatitis C, or influenza (Figure 2B).

3.3 Retracted papers and authors' affiliated country

Next, I analyzed data on countries affiliated with the authors of retracted papers in the top 20 countries with the largest number of publications. The association between retracted papers and affiliated countries was assessed by calculating the odds ratio in 3-year-periods (Figure 4; Supplementary Figure S5). During 2020–2022, a positive association for retractions was observed for China, India, and Iran. The odds ratios were high before, but recently decreased in Germany, Japan, and South Korea. The United Kingdom, France, the Netherlands, Brazil, Turkey, Switzerland, and Belgium have always shown a negative association with retractions from 1999 to 2022.

Figure 5 and Supplementary Figures S6–S8 show the retraction rates in 2020–2022 by the 18 disease categories plus COVID-19 and affiliated countries. The retraction rates of COVID-19-related papers were comparable with other diseases in most countries. While retraction rates were generally high in China, India, Iran, Turkey, and Japan (>4.0 per 10,000 publications), the retraction rates of COVID-19 papers from those countries were lower than other diseases except in India. China had the highest retraction rates in 16 out of 18 disease categories and the second highest in the remaining two categories.

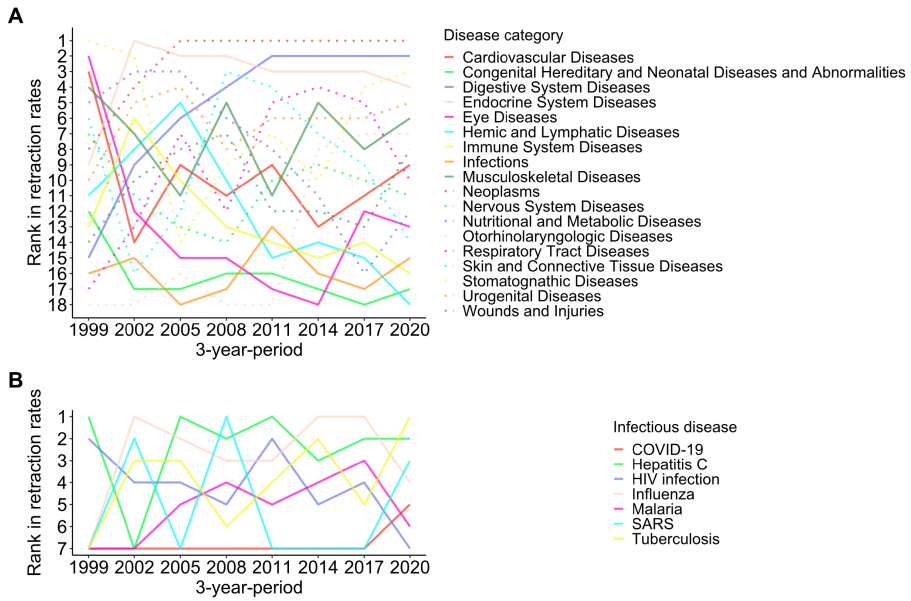

FIGURE 2
Ranking of retraction rates among different disease categories from 1999 to 2022. Ranking of retraction rates in 18 disease categories **(A)** and seven infectious diseases **(B)** in 3-year-periods are shown. Because there were no retractions for research papers on Ebola, Polio, and Zika, the three diseases were not shown in panel **(B)**.

However, when it comes to COVID-19, seven countries (India, South Korea, Australia, Belgium, Sweden, France, and Iran) showed higher retraction rates than China.

3.4 Multivariable analysis for retraction counts

Factors significantly associated with retraction count by a multivariable regression analysis are listed in Table 1. As indicated by descriptive observations so far, recent publications, particular disease categories (e.g., "Neoplasms," "Endocrine System Diseases," and "Digestive System Diseases"), and specific countries (e.g., China, Iran, India, and Japan) showed significant associations with retraction after statistical adjustments among those variables. The analysis found that publications in "Infections" are significantly less likely to be retracted compared with "Cardiovascular diseases," which is situated in the middle in retraction rates among 18 disease categories (incidence rate ratio, 0.56; p value <0.001).

3.5 Retracted papers in high-impact journals

Finally, the impact of the retracted papers was explored by checking the journals in which they were published. The proportion of all publications in high-impact journals (Journal Impact Factor 2021 > 10) has gradually decreased in most disease categories since 1999, except for the bounce of "Infections" and "Respiratory Tract Diseases" in 2020–2022 (Figure 6A; Supplementary Figure S9A).

The proportion of publications in high-impact journals was high for retracted papers before 2011 (Figure 6B; Supplementary Figure S9B). They have been decreasing recently, and in 2020–2022, retracted papers were less likely to be published in high-impact journals than the total publications in 17/18 disease categories (Figure 7; Supplementary Figure S10). The proportion of the total publications on COVID-19 published in high-impact journals was 18.0%, whereas 5.6% of the retracted COVID-19 papers were published in such journals. That proportion of retracted COVID-19 papers in high-impact journals was as low as other disease categories.

Because the COVID-19 pandemic drastically influenced the Journal Impact Factor (12), the same analysis was performed using the metric in 2018, i.e., the pre-COVID-19 period. Although the proportions of both total publications and retracted papers in high-impact journals slightly lessened when using the information of Journal Impact Factor 2018, their trend and relationship did not differ substantially (Supplementary Figures S9C,D, S10).

Similar sensitivity analyses using a different threshold (Impact Factor > 5) and indicator (SJR Q1 rank) also confirmed the decreasing trend of the proportion of retracted papers in so-called "good" journals (Supplementary Figures S9E–H). Still, there are differences in the results between analyses of high-impact journals and "good" journals. The proportion of retracted papers published in such "good" journals was higher than that of total publications in some disease categories (Supplementary Figure S10). For example, 75.0% of retracted papers about COVID-19 were published in Q1 journals, while the proportion in Q1 journals for total COVID-19 publications was 63.4%.

4 Discussion

This study found that (1) the retraction rates of medical research papers were increasing; (2) the retraction rates differed by year, disease

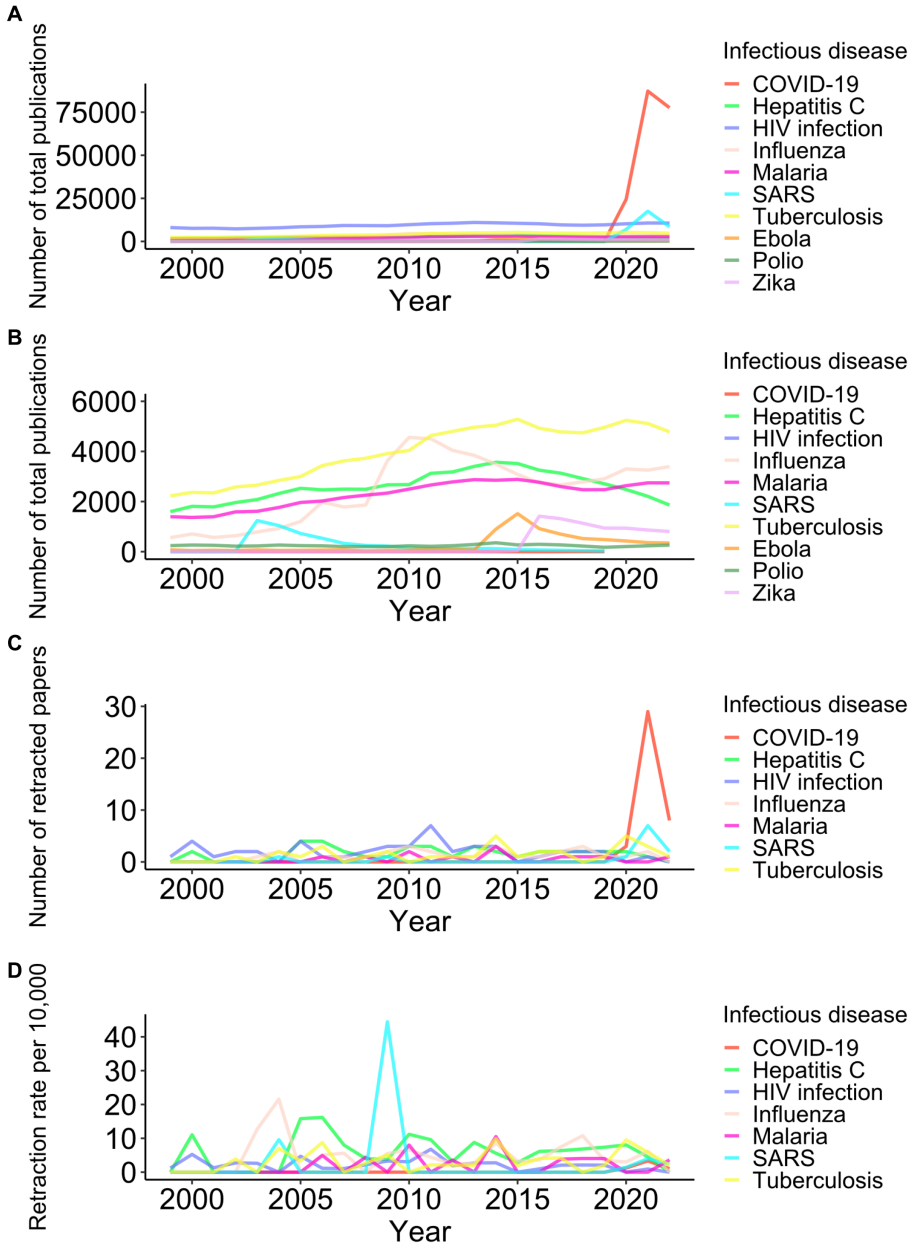

FIGURE 3
Trend of retractions among 10 infectious diseases from 1999 to 2022. Yearly numbers of total publications (**A** and **B** in different scales on the *y*-axis), retracted papers **(C)**, and retraction rates per 10,000 publications **(D)** for ten infectious diseases are shown. Because there were no retractions for research papers on Ebola, Polio, and Zika, the three diseases were excluded in panels **(C, D)**.

category, country, and journal; and (3) the surge in research on COVID-19 led to a large number of retractions, but did not result in a high retraction rate. Although some previous studies investigated retracted publications, they did not focus on the disease categories of retracted papers (13–15). The skewed retraction rates in particular research areas and countries found in the present study suggest the existence of systemic problems in specific environments, such as the lack of research ethics education, insufficient research capability, or high pressure to publish research outcomes. However, this study does not mean to blame particular research areas or countries.

Retractions of COVID-19 studies were conspicuous during the pandemic period. Yet, this study did not find a high retraction rate for COVID-19 papers. The retraction rate of publications in "Infections" did not increase during the pandemic period either. The number of publications on COVID-19 was simply enormous. Notably, China showed a low retraction rate for COVID-19-related papers unlike other diseases. Further investigation on its mechanisms could provide clues for reducing retractions in the country.

The reasons for retraction vary. Honest errors can occur because of mistakes in handling samples or data, skewed statistical analyses, inaccuracies or unverifiable information, and irreproducibility. Misconduct includes plagiarism, data fabrication or manipulation, lack of adherence to ethical protocols, undisclosed conflicts of interest, and duplicate submissions. Although the present study did

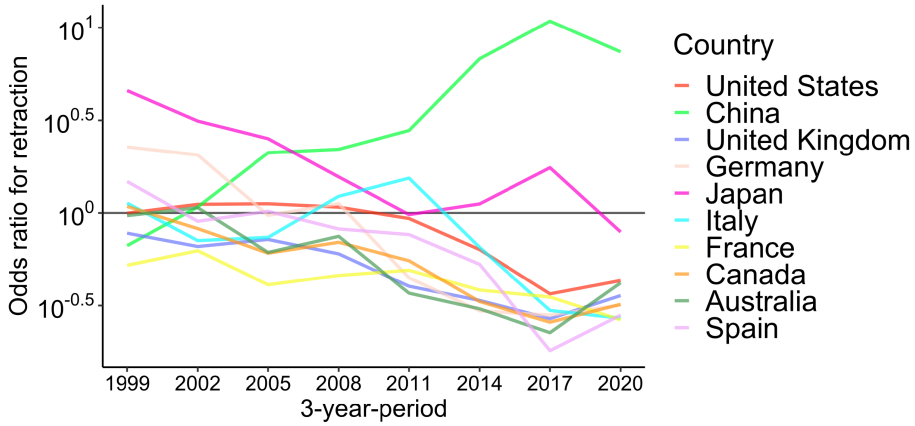

FIGURE 4
Association of affiliated countries with retraction from 1999 to 2022 in top 10 publishing countries. Odds ratios for retraction in 3-year-periods are shown for the top 10 countries with the highest publication counts. The results of the top 20 publishing countries can be found in Supplementary Figure S5.

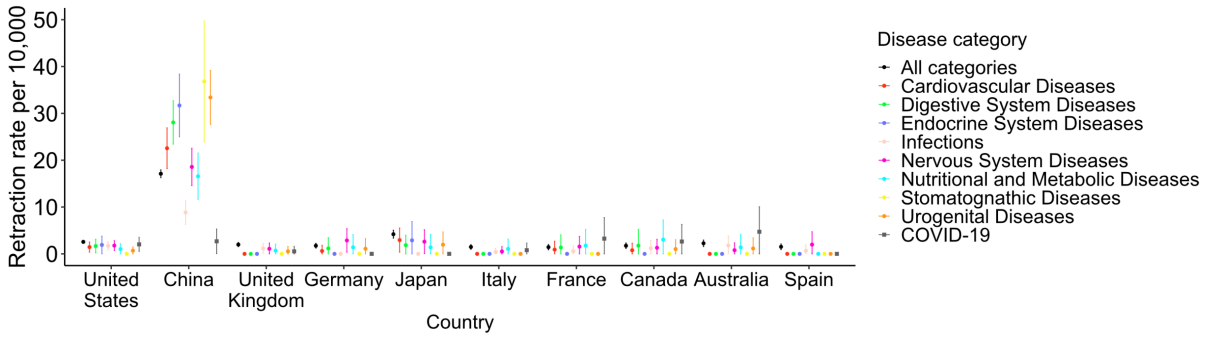

FIGURE 5
Retraction rates by country and disease category in 2020–2022, in top 10 publishing countries for eight disease categories plus COVID-19. The retraction rates per 10,000 papers in 2020–2022 are shown for eight disease categories, as determined in Figure 1, and COVID-19-related papers in the top 10 countries with the highest publication counts. Vertical lines indicate 95% confidence intervals. The results of the top 20 countries for all 18 disease categories can be found in Supplementary Figure S6.

not analyze the reasons for retractions because such data were unavailable in the PubMed database, Shi et al. reported that COVID-19 papers were retracted more often without detailed explanations or due to "non-misconduct-related concerns" than non-COVID-19 studies (7).

One of the limitations of this study is that it relied solely on the data from PubMed. The database does not always include information on the time of submission or retraction. Changes in the time difference between submission and publication and between publication and retraction during the pandemic should be of interest and explored in future research. Furthermore, PubMed does not index some preprints that played a significant role in rapidly sharing research findings during the COVID-19 pandemic. More retractions can be found in other resources, for example, the Retraction Watch Database (6, 7).

The retraction counts and rates in recent years analyzed in this study must be underestimated. Because new publications with possible concerns are still under investigation or unrealized, some more papers may be retracted in the future. Consequently, it is difficult to directly compare the retraction rates in recent years to past years and determine if the COVID-19 pandemic has affected the retraction rates of non-COVID-19 research papers. Still, the relative changes in the retraction rates among different disease categories can be discussed (Figure 2). The retraction rates of "Neoplasms," "Digestive System Diseases," and "Endocrine System Diseases" remained high in both before and during the COVID-19 pandemic periods compared with other disease categories. There was no evident increase in the (relative) retraction rates of "Infections" or "Respiratory Tract Diseases" during the pandemic. Those findings imply marginal, if any, effects of the pandemic on retractions.

The importance of awareness, education, and compliance with research integrity increases as retraction rates continue to grow. We should attempt to reduce errors and misconduct in research activities. However, it is virtually impossible to completely eliminate errors or misconduct. Additionally, an increase in retraction rates in recent years may indicate improvement in research transparency, information sharing, and constructive criticism in research communities. The decreasing trend of retracted papers in high-impact journals also suggests a rigorous peer-review process.

TABLE 1 Multivariable regression analysis for retraction count.

Variable	Category*[1]	Adjusted incidence rate ratio (95% confidence interval)*[2]	*p* value*[2]
Publication year	2017–2019	3.59 (2.82–4.67)	<0.001
	2014–2016	2.78 (2.17–3.63)	<0.001
	2011–2013	2.48 (1.92–3.24)	<0.001
	2008–2010	2.46 (1.90–3.24)	<0.001
	2005–2007	2.57 (1.98–3.40)	<0.001
	2002–2004	1.47 (1.09–2.00)	0.013
	1999–2001	Reference	
Disease category	Neoplasms	1.97 (1.73–2.25)	<0.001
	Endocrine system diseases	1.63 (1.37–1.95)	<0.001
	Digestive system diseases	1.51 (1.30–1.76)	<0.001
	Musculoskeletal diseases	1.34 (1.10–1.62)	0.003
	Urogenital diseases	1.28 (1.09–1.51)	0.003
	Skin and connective tissue diseases	1.23 (1.01–1.48)	0.035
	Cardiovascular diseases	Reference	
	Immune system diseases	0.79 (0.63–0.97)	0.031
	Eye diseases	0.71 (0.51–0.97)	0.037
	Infections	0.56 (0.46–0.68)	<0.001
	Congenital, hereditary, and neonatal diseases and abnormalities	0.47 (0.34–0.63)	<0.001
Affiliated country	China	17.57 (14.23–22.03)	<0.001
	Iran	7.98 (5.87–10.86)	<0.001
	India	5.15 (3.93–6.78)	<0.001
	Japan	4.95 (3.92–6.32)	<0.001
	United States	2.45 (1.96–3.10)	<0.001
	South Korea	1.77 (1.13–2.68)	0.009
	Italy	1.67 (1.25–2.24)	0.001
	Canada	1.47 (1.05–2.06)	0.024
	United Kingdom	Reference	

*[1]Only items with statistical significance are shown. *[2]Adjusted incidence rate ratios and *p* values were calculated using a quasi-Poisson regression model.

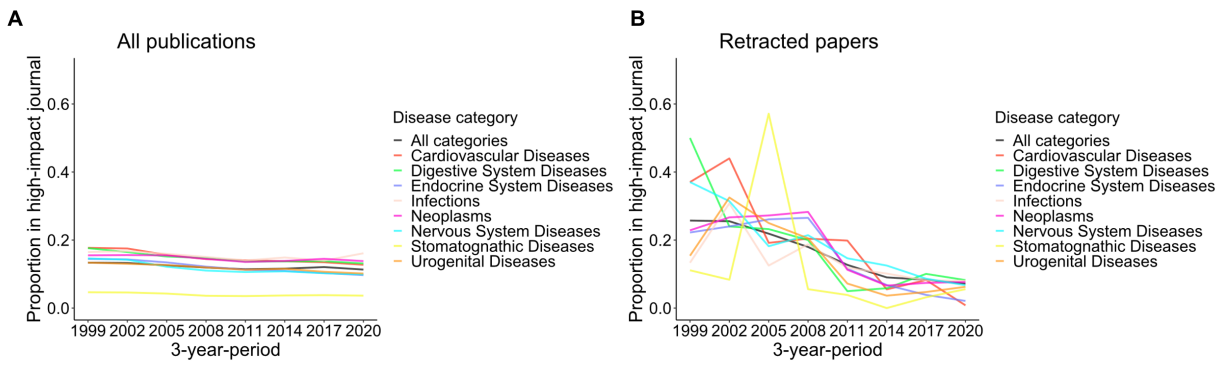

FIGURE 6
Proportion of papers published in high-impact journals among eight disease categories from 1999 to 2022. The proportions of papers published in high-impact journals (Impact Factor 2021 > 10) for total publications **(A)** and retracted papers **(B)** in eight disease categories from Figure 1 are shown in 3-year-periods. The results of all 18 disease categories can be found in Supplementary Figure S9.

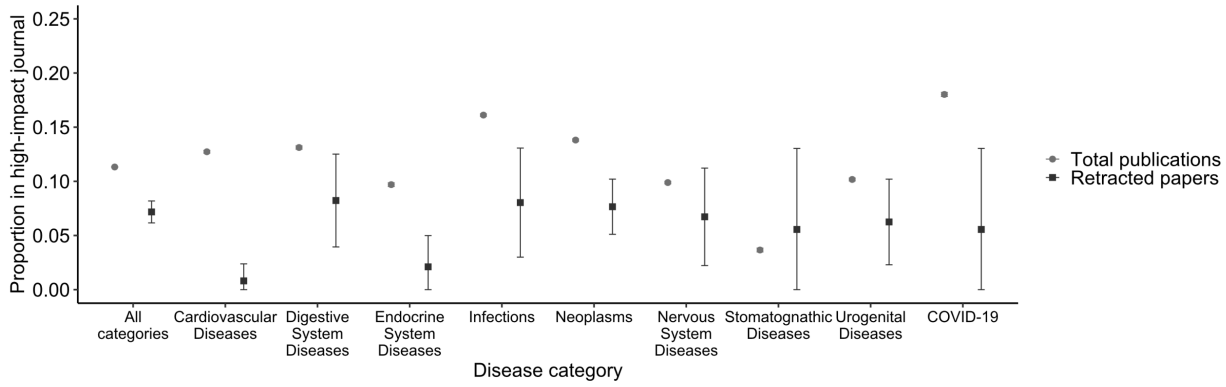

FIGURE 7
Proportion of papers published in high-impact journals in 2020–2022 among eight disease categories plus COVID-19. The proportions of papers published in high-impact journals (Impact Factor 2021 > 10) for total publications and retracted papers in 2020–2022 are depicted for eight disease categories and COVID-19. Vertical lines indicate 95% confidence intervals. The results of all 18 disease categories using different metrics can be found in Supplementary Figure S10.

We have to keep an eye on retractions and analyze how pandemics affect them to find and address issues for creating better science. Infectious disease pandemics could cause errors or misconduct in research due to pressure to rapidly publish research findings in a high-profile field. However, this study found that outbreaks of past emerging infectious diseases did not lead to an increase of retraction probability. In a crisis situation such as a pandemic, we should keep conducting research carefully and honestly to confront adversity.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://pubmed.ncbi.nlm.nih.gov.

Author contributions

YF: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Strategic Center of Biomedical Advanced Vaccine Research and Development for Preparedness and Response (JP233fa627004) and the Japan Initiative for Global Research Network on Infectious Diseases (23wm0125011h0001) from the Japan Agency for Medical Research and Development, and by the Nagasaki University State of the Art Research Program (grant number not available) from Nagasaki University. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2023.1288014/full#supplementary-material

References

1. Croskerry P. To err is human--and let's not forget it. *CMAJ*. (2010) 182:524. doi: 10.1503/CMAJ.100270

2. Ostrovsky AM, Parikh C. Impact of misinformation on ivermectin internet searches and prescribing trends during COVID-19. *J Public Health (Oxf)*. (2022) 45:631–5. doi: 10.1093/PUBMED/FDAC152

3. Hill A, Mirchandani M, Pilkington V. Ivermectin for COVID-19: addressing potential Bias and medical fraud. *Open Forum Infect Dis*. (2022) 9:ofab645. doi: 10.1093/OFID/OFAB645

4. Farah R, Kazzi Z, Brent J, Burkhart K, Wax P, Aldy K. Ivermectin associated adverse events in the treatment and prevention of COVID-19 reported to the

FACT pharmacovigilance project. *Clin Toxicol*. (2022) 60:942–6. doi: 10.1080/15563650.2022.2070187

5. Hoang R, Temple C, Correia MS, Clemons J, Hendrickson RG. Characteristics of ivermectin toxicity in patients taking veterinary and human formulations for the prevention and treatment of COVID-19. *Clin Toxicol*. (2022) 60:1350–5. doi: 10.1080/15563650.2022.2134788

6. Clark J. How covid-19 bolstered an already perverse publishing system. *BMJ*. (2023) 380:689. doi: 10.1136/BMJ.P689

7. Shi X, Abritis A, Patel RP, Grewal M, Oransky I, Ross JS, et al. Characteristics of retracted research articles about COVID-19 vs other topics. *JAMA Netw Open*. (2022) 5:E2234585. doi: 10.1001/JAMANETWORKOPEN.2022.34585

8. Winter DJ. Rentrez: an R package for the NCBI eUtils API. *R J*. (2017) 9:520. doi: 10.32614/RJ-2017-058

9. Rogers FB. Medical subject headings. *Bull Med Libr Assoc*. (1963) 51:114–6.

10. Furuse Y. Analysis of research intensity on infectious disease by disease burden reveals which infectious diseases are neglected by researchers. *Proc Natl Acad Sci U S A*. (2019) 116:478–83. doi: 10.1073/pnas.1814484116

11. Mullen L, Mullen L, Potter C, Potter C, Gostin LO, Cicero A, et al. An analysis of international health regulations emergency committees and public health emergency of international concern designations. *BMJ Glob Health*. (2020) 5:e002502. doi: 10.1136/BMJGH-2020-002502

12. Delardas O, Giannos P. How COVID-19 affected the journal impact factor of high impact medical journals: bibliometric analysis. *J Med Internet Res*. (2022) 24:e43089. doi: 10.2196/43089

13. Gaudino M, Robinson NB, Audisio K, Rahouma M, Benedetto U, Kurlansky P, et al. Trends and characteristics of retracted articles in the biomedical literature, 1971 to 2020. *JAMA Intern Med*. (2021) 181:1118–21. doi: 10.1001/JAMAINTERNMED.2021.1807

14. Vuong QH, La VP, Ho MT, Vuong TT, Ho MT. Characteristics of retracted articles based on retraction data from online sources through February 2019. *Sci Ed*. (2020) 7:34–44. doi: 10.6087/KCSE.187

15. Elango B. Characteristics of retracted editorial articles in the biomedical literature. *Scientometrics*. (2022) 127:1431–8. doi: 10.1007/S11192-021-04263-9

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 11 March 2024
DOI 10.3389/fmed.2024.1363045

OPEN ACCESS

EDITED BY
Reza Rastmanesh,
American Physical Society, United States

REVIEWED BY
Francesca Torriani,
University of California, San Diego, United States
Stefan Pilz,
Medical University of Graz, Austria

*CORRESPONDENCE
Hiam Chemaitelly
✉ hsc2001@qatar-med.cornell.edu
Laith J. Abu-Raddad
✉ lja2002@qatar-med.cornell.edu

RECEIVED 29 December 2023
ACCEPTED 22 February 2024
PUBLISHED 11 March 2024

CITATION
Chemaitelly H, Ayoub HH, Tang P, Yassine HM, Al Thani AA, Hasan MR, Coyle P, Al-Kanaani Z, Al-Kuwari E, Jeremijenko A, Kaleeckal AH, Latif AN, Shaik RM, Abdul-Rahim HF, Nasrallah GK, Al-Kuwari MG, Butt AA, Al-Romaihi HE, Al-Thani MH, Al-Khal A, Bertollini R and Abu-Raddad LJ (2024) Addressing bias in the definition of SARS-CoV-2 reinfection: implications for underestimation. *Front. Med.* 11:1363045. doi: 10.3389/fmed.2024.1363045

Addressing bias in the definition of SARS-CoV-2 reinfection: implications for underestimation

Hiam Chemaitelly[1,2,3]*, Houssein H. Ayoub[4], Patrick Tang[5], Hadi M. Yassine[6,7], Asmaa A. Al Thani[6,7], Mohammad R. Hasan[8], Peter Coyle[6,9,10], Zaina Al-Kanaani[9], Einas Al-Kuwari[9], Andrew Jeremijenko[9], Anvar Hassan Kaleeckal[9], Ali Nizar Latif[9], Riyazuddin Mohammad Shaik[9], Hanan F. Abdul-Rahim[11], Gheyath K. Nasrallah[6,7], Mohamed Ghaith Al-Kuwari[12], Adeel A. Butt[3,9,13], Hamad Eid Al-Romaihi[14], Mohamed H. Al-Thani[14], Abdullatif Al-Khal[9], Roberto Bertollini[14] and Laith J. Abu-Raddad[1,2,3,11,15]*

[1]Infectious Disease Epidemiology Group, Weill Cornell Medicine-Qatar, Cornell University, Doha, Qatar, [2]World Health Organization Collaborating Centre for Disease Epidemiology Analytics on HIV/AIDS, Sexually Transmitted Infections, and Viral Hepatitis, Weill Cornell Medicine–Qatar, Cornell University, Doha, Qatar, [3]Department of Population Health Sciences, Weill Cornell Medicine, Cornell University, New York, NY, United States, [4]Mathematics Program, Department of Mathematics, Statistics, and Physics, College of Arts and Sciences, Qatar University, Doha, Qatar, [5]Department of Pathology, Sidra Medicine, Doha, Qatar, [6]Biomedical Research Center, QU Health, Qatar University, Doha, Qatar, [7]Department of Biomedical Science, College of Health Sciences, QU Health, Qatar University, Doha, Qatar, [8]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada, [9]Hamad Medical Corporation, Doha, Qatar, [10]Wellcome-Wolfson Institute for Experimental Medicine, Queens University, Belfast, United Kingdom, [11]Department of Public Health, College of Health Sciences, QU Health, Qatar University, Doha, Qatar, [12]Primary Health Care Corporation, Doha, Qatar, [13]Department of Medicine, Weill Cornell Medicine, Cornell University, New York, NY, United States, [14]Ministry of Public Health, Doha, Qatar, [15]College of Health and Life Sciences, Hamad bin Khalifa University, Doha, Qatar

Introduction: Reinfections are increasingly becoming a feature in the epidemiology of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. However, accurately defining reinfection poses methodological challenges. Conventionally, reinfection is defined as a positive test occurring at least 90 days after a previous infection diagnosis. Yet, this extended time window may lead to an underestimation of reinfection occurrences. This study investigated the prospect of adopting an alternative, shorter time window for defining reinfection.

Methods: A longitudinal study was conducted to assess the incidence of reinfections in the total population of Qatar, from February 28, 2020 to November 20, 2023. The assessment considered a range of time windows for defining reinfection, spanning from 1 day to 180 days. Subgroup analyses comparing first versus repeat reinfections and a sensitivity analysis, focusing exclusively on individuals who underwent frequent testing, were performed.

Results: The relationship between the number of reinfections in the population and the duration of the time window used to define reinfection revealed two distinct dynamical domains. Within the initial 15 days post-infection diagnosis, almost all positive tests for SARS-CoV-2 were attributed to the original infection. However, surpassing the 30-day post-infection threshold, nearly all positive tests were attributed to reinfections. A 40-day time window emerged as a sufficiently conservative definition for reinfection. By setting the time window

at 40 days, the estimated number of reinfections in the population increased from 84,565 to 88,384, compared to the 90-day time window. The maximum observed reinfections were 6 and 4 for the 40-day and 90-day time windows, respectively. The 40-day time window was appropriate for defining reinfection, irrespective of whether it was the first, second, third, or fourth occurrence. The sensitivity analysis, confined to high testers exclusively, replicated similar patterns and results.

Discussion: A 40-day time window is optimal for defining reinfection, providing an informed alternative to the conventional 90-day time window. Reinfections are prevalent, with some individuals experiencing multiple instances since the onset of the pandemic.

KEYWORDS

reinfection, bias, time window, immunity, COVID-19, epidemiology

1 Introduction

Reinfections with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) increased as the protection conferred by natural infection waned over time (1, 2). Importantly, this increase was amplified by the emergence of the immune-evasive omicron variant and its subvariants (1–4). The occurrence of reinfections is becoming a regular feature in the epidemiology of SARS-CoV-2, resembling patterns observed in other respiratory infections such as common-cold coronaviruses (5, 6) and influenza (7–10). Gaining insight into the epidemiology of SARS-CoV-2 reinfections is an essential prerequisite for understanding the broader landscape of SARS-CoV-2 epidemiology.

However, defining a SARS-CoV-2 reinfection presents methodological challenges. The most suitable definition, in theory, entails genome sequencing of the virus in every SARS-CoV-2-positive test and evaluating whether the identified virus in a given positive test differs from that detected in the previous positive test (11–13). Implementing this approach is resource-intensive and impractical, especially at this stage of the pandemic.

A pragmatic methodological approach to defining a SARS-CoV-2 reinfection involves applying a time window, allowing for the clearance of an earlier infection to classify a new positive test as a reinfection. Consequently, reinfection is commonly defined as a SARS-CoV-2-positive test occurring at least 90 days after a previous SARS-CoV-2-positive test (3, 14, 15). Despite the fact that the vast majority of SARS-CoV-2 infections resolve within a few days (16, 17), the adoption of this 90-day time window aimed to prevent the misclassification of prolonged infections as reinfections (3, 14, 15), recognizing the persistence of some infections for weeks or even months, albeit rarely (18–20). This choice also accounted for the situation earlier in the pandemic when reinfections were rare (11, 13, 21), emphasizing the importance of distinguishing between two rare events: reinfection versus prolonged infection.

While this definition offers a practical alternative for defining reinfection, it underestimates the occurrence of reinfections, as any true reinfection within 90 days of an earlier infection is not classified as such. The inherent bias in this definition compounds over time, given that this 90-day time window is applied to every subsequent reinfection, precisely when repeat reinfections are becoming increasingly common (22, 23). SARS-CoV-2 waves have been occurring within only a few months of each other, or even occasionally within weeks (24). Therefore, a 90-day threshold may miss many true reinfections in consecutive waves if the time difference between waves is less or comparable to this set 90-day time window.

With the continual evolution of this virus and the emergence of more immune-evasive subvariants (25), this conventional 90-day time window may introduce serious bias in studies of reinfections, potentially leading to incorrect inferences drawn from studies with inaccurately estimated occurrences of reinfections. Importantly, the caution needed to distinguish the rare events of reinfection from prolonged infection early in the pandemic is no longer warranted, as reinfections are no longer rare, while prolonged infections remain as rare as they were before.

To address this challenge, this study explored the possibility of implementing an alternative, shorter time window for defining reinfection. The investigation aims to enhance the methodologies used in studying reinfections and the immune protection of natural infection while mitigating the inherent bias present in the current definition of a 90-day time window.

2 Methods

2.1 Study population and data sources

The study was conducted on the population of Qatar from February 28, 2020, the date of the first documented SARS-CoV-2 infection, up to November 20, 2023, the date of the end of the study. The analysis utilized the national, federated databases for coronavirus disease 2019 (COVID-19) laboratory testing, vaccination, hospitalization, and death retrieved from the integrated, nationwide digital health information platform (Supplementary Section S1). The platform has captured all SARS-CoV-2-related data with no missing information since the onset of the pandemic, including all polymerase chain reaction (PCR) tests and medically supervised rapid antigen tests (Supplementary Section S2).

All SARS-CoV-2 testing in any facility in Qatar is tracked nationally in one database, the national testing database. This database covers all testing in all locations and facilities throughout the country, whether public or private. SARS-CoV-2 tests are classified on the basis of symptoms and the reason for testing (clinical symptoms, contact tracing, surveys or random testing campaigns, individual requests, routine healthcare testing, pre-travel, at port of entry, or other). Testing is offered free of charge or at heavily subsidized costs depending on the reason for testing within Qatar's public healthcare system, accessible to all residents irrespective of nationality. These services are available at healthcare centers distributed across the country, catering to the diverse demographic and socio-economic segments of the population.

Up until November 1, 2022, nearly 5% of the population underwent SARS-CoV-2 testing each week, primarily for routine and non-clinical purposes (26). Based on the distribution of the reason for testing up to October 30, 2022, most of the tests in Qatar were conducted for routine reasons, such as being travel-related, and about 75% of documented infections were diagnosed not because of appearance of symptoms, but because of routine testing (26, 27). However, starting from November 1, 2022, the testing rate was reduced to less than 1% per week (28).

The first omicron wave, reaching its peak in January 2022, was of very large magnitude and placed substantial strain on the country's testing capacity (3, 27, 29). Consequently, rapid antigen testing was introduced and implemented as a substitute for PCR testing, employing identical testing protocols.

The extensive testing approach in Qatar enabled the tracking of reinfections, irrespective of symptomatic presentation, facilitating an opportunity to investigate potential biases in defining reinfection cases.

Qatar initiated its COVID-19 vaccination program in December 2020, utilizing mRNA vaccines and prioritizing individuals based on coexisting conditions and age criteria (26, 30). The vaccination is administered free of charge to all residents, irrespective of nationality, and is centrally tracked at a national level (26, 30).

Qatar has young, diverse demographics; only 9% of its residents are aged 50 or above, with 89% being expatriates from over 150 countries (31). Migrant craft and manual workers constitute about 60% of the population (32, 33), mainly single men aged 20 to 49, hailing predominantly from countries like Bangladesh, India, and Nepal, and working in development projects (34). Consequently, nationality, age, and sex serve as proxies for socio-economic status in this context (31, 33, 35, 36). Further descriptions of Qatar's population and the national databases have been reported previously (26, 27, 29, 31, 34, 37, 38).

2.2 Study design

A longitudinal study was undertaken to assess the incidence of reinfections within the population of Qatar, considering varying time windows for defining reinfection, ranging from 1 day to 180 days. The study cohort encompassed all individuals with a documented SARS-CoV-2-positive test during the study period. In this cohort, which comprised 935,192 individuals, a total of 6,170,451 tests (positive or negative) were conducted from the onset of the pandemic until the conclusion of the study on November 20, 2023. The average testing rate stood at 6.6 tests per person.

Primary infection was defined as the first documented instance of a SARS-CoV-2-positive test for an individual. Reinfection was defined as the first documented SARS-CoV-2-positive test occurring after the completion of the time window used to define reinfection, starting from the last previous SARS-CoV-2 infection diagnosis. The primary outcomes of the study included the total number of documented reinfections in the population and the maximum number of observed reinfections experienced by any given individual in the population, both examined across the various time window definitions investigated in this study.

In essence, the concept of the present study is that the relationship between the total number of reinfections in the population and the duration of the time window used to define reinfection may reveal clearly distinct dynamical domains, enabling an informed decision on setting the time window to define reinfection. The existence of two distinct dynamical domains is a reflection of the existence of two different population distributions influencing this relationship. The first is the distribution of clearing the infection, and the second is the distribution of the incidence of reinfection.

2.3 Statistical analysis

Frequency distributions and measures of central tendency were employed to characterize measures within the study cohort. Statistical analyses calculated the total number of documented reinfections in the population and the maximum number of observed reinfections experienced by any given individual, considering varying time windows for defining reinfection, ranging from 1 day to 180 days. The total number of documented first, second, third, and fourth reinfections in the population were also computed for the different time windows. This latter investigation aimed to assess whether distinct time window definitions are warranted for repeat reinfections compared to the first reinfection.

Documented reinfections constitute only a subset of all potential reinfections in a population, as many may go undocumented through a SARS-CoV-2 test. Patterns for undocumented infections may deviate from those that are documented. To address this, the study analyses were repeated in a sensitivity analysis including only high testers in the study cohort, a subset of the population less impacted by undocumented infections due to frequent repeat testing, often for routine reasons such as employment or travel (26, 27).

High testers were defined as individuals in the top 10th percentile of the real-world testing frequency distribution, encompassing all reasons for testing. According to this distribution, high testers are individuals with a testing rate of $\geq$3.4 tests per person-year. The consistency of patterns and results among high testers with those in the full cohort would support the conclusion that the proposed time window in this study may not have been influenced by the occurrence of undocumented infections. Statistical analyses were performed using Stata/SE version 18.0 (Stata Corporation, College Station, TX, USA).

A second sensitivity analysis was undertaken to examine the consistency of reinfection patterns observed in the main analysis, encompassing all times during the pandemic, when the analysis is restricted to the four largest SARS-CoV-2 infection waves, each dominated by a distinct variant (4, 21, 31, 39).

2.4 Oversight

The institutional review boards at Hamad Medical Corporation and Weill Cornell Medicine–Qatar approved this retrospective study with a waiver of informed consent. The study was reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (Supplementary Table S1).

3 Results

3.1 Optimizing the time window for defining reinfection

Table 1 shows the characteristics of the study cohort, comprising all individuals in Qatar with a documented SARS-CoV-2 infection during the study period. The study population is representative of the internationally diverse, yet predominantly young and male demographic of the country.

Figure 1 illustrates the total estimated number of reinfections in the population versus the time-window duration used for defining reinfection. Initially, the total number of reinfections declines rapidly, from 235,660 when the time window is set at only 1 day to 113,649 when the time window is extended to 15 days after the previous infection diagnosis. This swift decline supports the notion that many individuals are testing positive within the first 15 days of a positive test. At the same time, it also indicates that more and more individuals are clearing the infection, leading to a progressive increase in the number of people testing negative with each passing day within this 15-day duration. This trend is consistent with the vast majority of infected persons clearing their infection within 15 days of diagnosis.

Following this initial rapid decline, a pronounced shift in the curve's shape becomes evident within the time window spanning from 16 days to 30 days (Figure 1). This transition to a new shape of the curve indicates the presence of two distinct dynamical domains: one where nearly all positive tests are attributable to the same original infection before day 15 after the previous infection diagnosis, and another, following the transition, where nearly all positive tests after day 30 are attributed to reinfections, with negligible contribution from prolonged infections. These two domains are labeled thereafter as the "infection clearance" and "reinfection plateau" domains, respectively.

These two clearly distinct domains emerge because two different population distributions dominate the relationship between the total number of reinfections and the duration of the time window at *different* times post-infection (Figure 1). The distribution of clearing the infection predominates in the infection clearance domain, while the distribution of the incidence of reinfection dominates in the reinfection plateau domain. The transition between these domains, occurring from 16 days to 30 days post-infection, is influenced by both of these distributions.

Given the existence of these two clearly distinct dynamics, choosing a time window for defining reinfection at 40 days strikes a balance. It is adequately conservative in defining a reinfection (as opposed to a prolonged infection) while not missing many reinfections compared to a longer time window. Setting the time window at 40 days, instead of the conventional 90-day window, increases the total estimated number of reinfections in the population from 84,565 to 88,384 reinfections, capturing an additional 4.3% of reinfections that would have been missed by applying the conventional 90-day time window.

TABLE 1 Baseline characteristics of the study population.

Characteristics	Study cohort *N* (%)
Total *N*	935,192
Median age (IQR)—years	33.0 (24.0–41.0)
Age—years	
0–9 years	90,028 (9.6)
10–19 years	87,116 (9.3)
20–29 years	192,565 (20.6)
30–39 years	296,309 (31.7)
40–49 years	167,292 (17.9)
50–59 years	70,175 (7.5)
60–69 years	23,214 (2.5)
≥70 years	8,493 (0.9)
Sex	
Male	584,687 (62.5)
Female	350,505 (37.5)
Nationality*	
Bangladeshi	52,762 (5.6)
Egyptian	48,945 (5.2)
Filipino	93,921 (10.0)
Indian	210,315 (22.5)
Nepalese	65,128 (7.0)
Pakistani	39,568 (4.2)
Qatari	176,951 (18.9)
Sri Lankan	24,255 (2.6)
Sudanese	24,085 (2.6)
Other nationalities†	199,262 (21.3)
Number of coexisting conditions	
None	724,513 (77.5)
1	115,788 (12.4)
2	48,373 (5.2)
3	20,639 (2.2)
4	11,533 (1.2)
5	6,918 (0.7)
≥6	7,428 (0.8)
Vaccination‡	
Unvaccinated	564,670 (60.4)
1 dose	19,657 (2.1)
2 doses	270,227 (28.9)
3 doses	79,666 (8.5)
≥4 doses	972 (0.1)

IQR, interquartile range; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.
*Nationalities were chosen to represent the most populous groups in Qatar.
†These comprise 173 other nationalities.
‡Ascertained at time of primary infection.

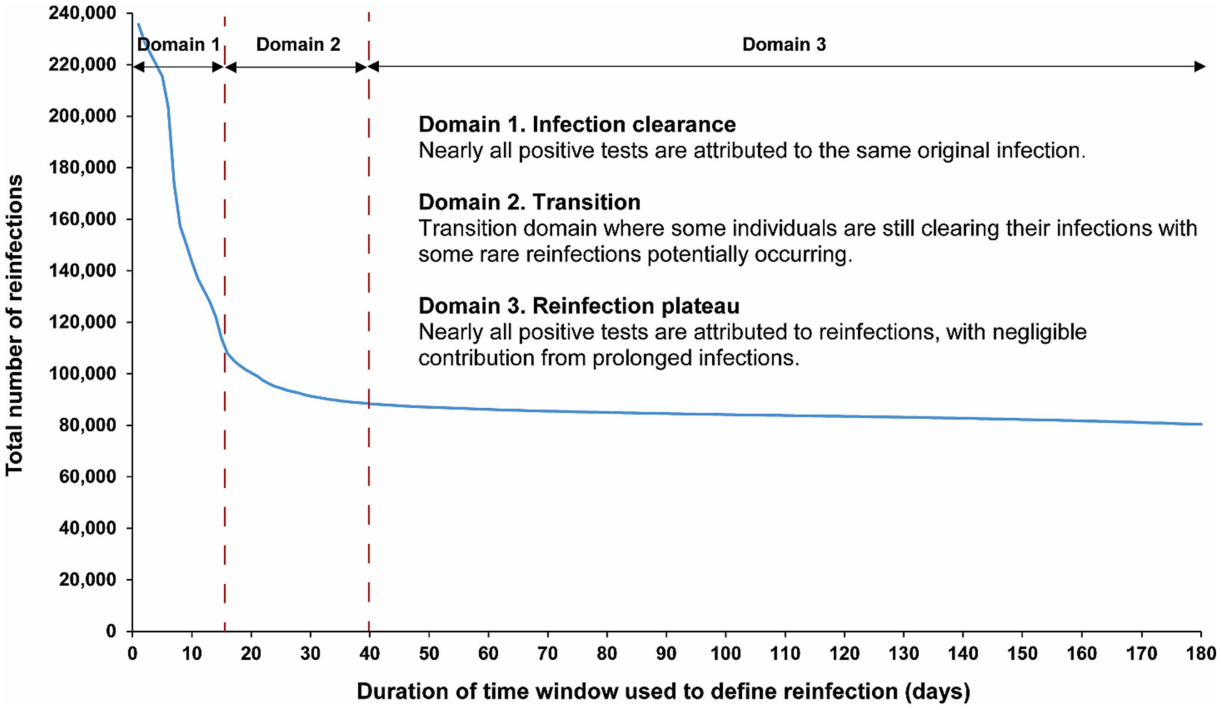

FIGURE 1
Total number of reinfections in the population versus the time-window duration used for defining reinfection. The dashed line at day 40 represents the proposed optimal time window for defining a reinfection.

3.2 Time window for first and repeat reinfections

The above analysis indicates an optimal choice of a 40-day time window for defining reinfection. Figure 2 explores the applicability of this time window individually for each of the first, second, third, and fourth reinfections. The results demonstrate that a time window of 40 days is appropriate for defining reinfection, regardless of whether it is a first, second, third, or fourth occurrence. The number of reinfections versus the time-window duration followed a largely consistent pattern, irrespective of whether the reinfection was a first reinfection or a repeat reinfection.

Notably, in the analyses of third and fourth reinfections, the transition in the shape of the number of reinfections versus the time-window duration appears to occur more rapidly, reaching the reinfection plateau sooner. This supports the possibility of an even shorter time window than 40 days for defining reinfection. This may be attributed to the fact that, by the time individuals in this population experienced their third or fourth reinfections during the pandemic, there was limited testing in the initial days after the infection to assess clearance, unlike in earlier stages. It could also be a result of faster clearance of reinfections, especially repeat occurrences (40–42).

3.3 Maximum number of reinfections in the population

Figure 3 illustrates the maximum number of observed reinfections experienced by any given individual in the population versus the time-window duration used for defining reinfection. This figure highlights the relevance of an appropriate definition for the time window in capturing the phenomenon of repeat reinfections. For example, if the time window is set at 15 days, at least one individual in the population would have been estimated to have experienced 14 reinfections throughout the pandemic. Meanwhile, the maximum number of observed reinfections is 8, 6, and 4 if the time window was set at 30, 40, and 90 days, respectively.

3.4 Sensitivity analysis: results for only high testers

The sensitivity analysis, restricted to only high testers in the population, reproduced the same patterns and results as those observed for the entire population. This held true for all study outcomes, including the total number of reinfections (Figure 4A), the number of each of the first, second, third, and fourth reinfections (Figures 4B,C), and the maximum number of observed reinfections (Figure 5). The analysis affirmed the 40-day time window as an optimal choice, suggesting that the conclusions drawn above regarding setting the time window are unlikely to have been altered by the occurrence of infections that were never documented.

3.5 Sensitivity analysis: reinfection patterns in distinct waves

The sensitivity analysis, restricting the analysis to each of the four largest SARS-CoV-2 infection waves, each dominated by a distinct variant, showed the same reinfection patterns observed in the main

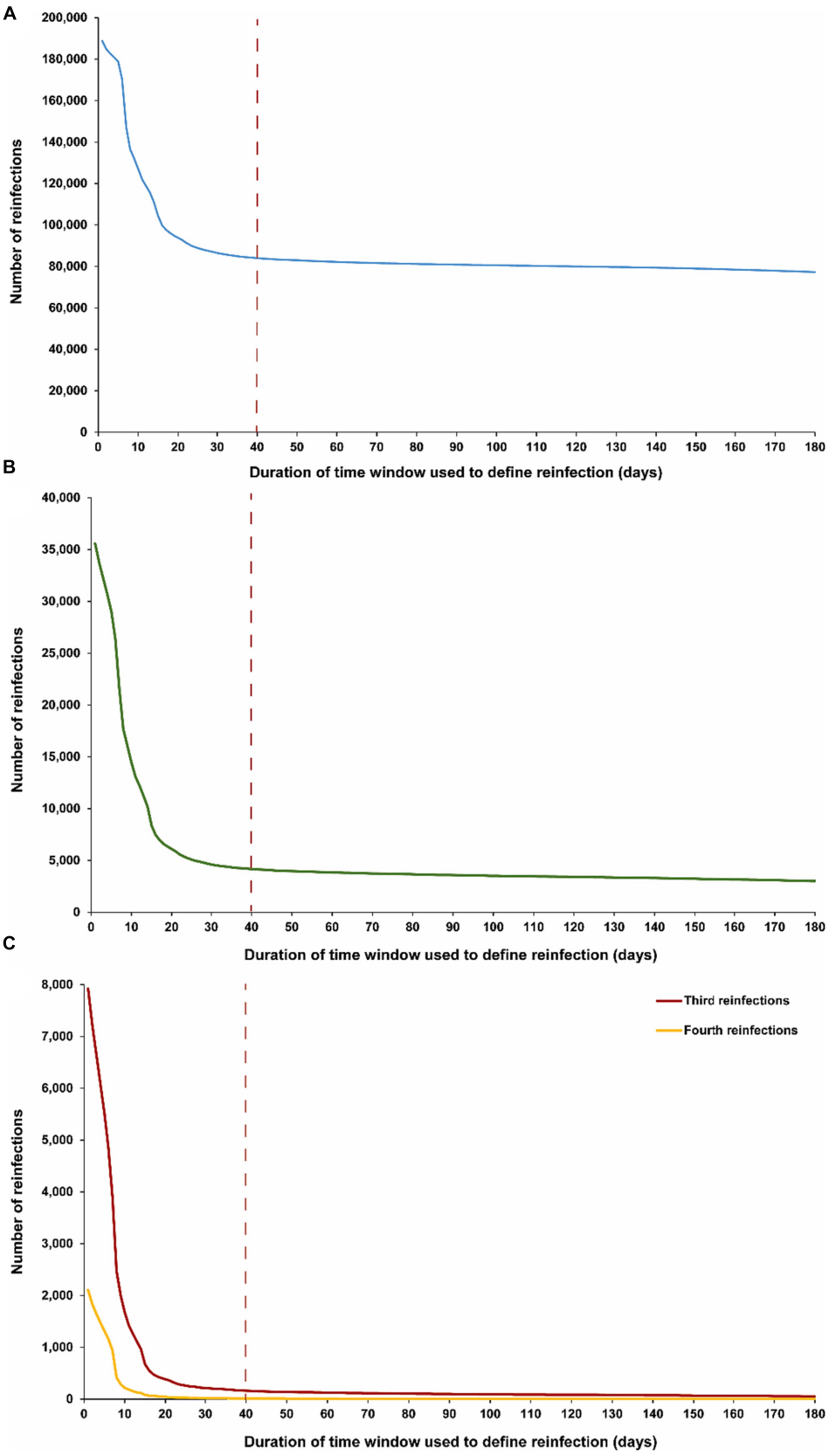

FIGURE 2
Number of **(A)** first reinfections, **(B)** second reinfections and **(C)** third, and fourth reinfections in the population versus the time-window duration used for defining reinfection. The dashed line represents the proposed optimal time window for defining a reinfection.

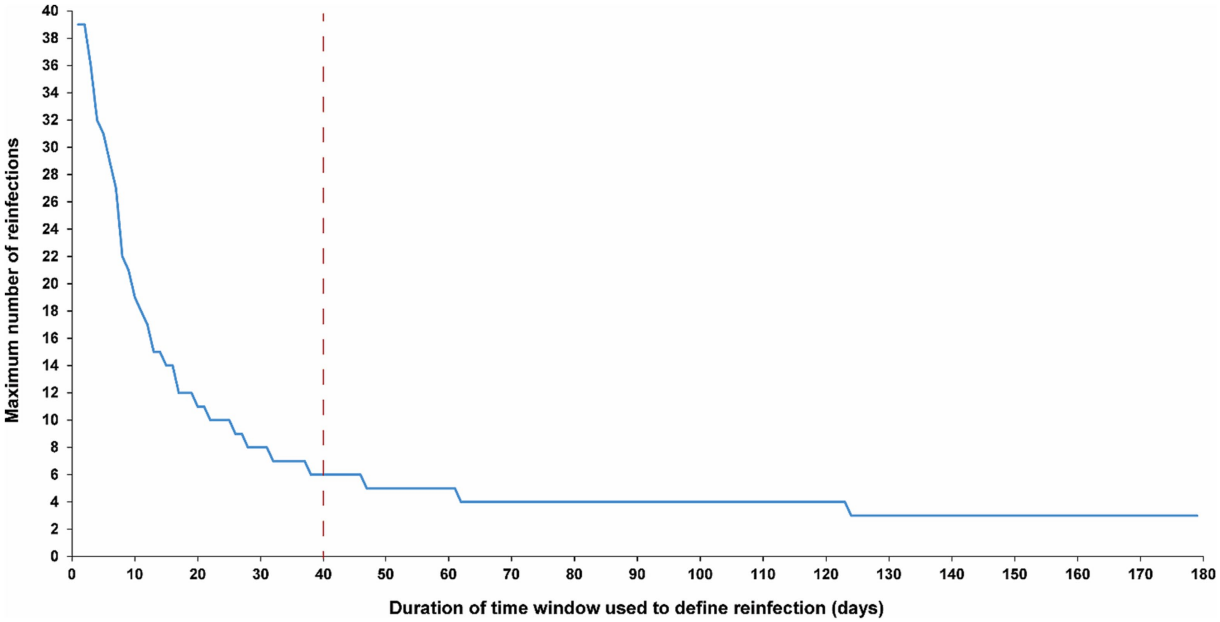

FIGURE 3
Maximum number of observed reinfections experienced by any given individual in the population versus the time-window duration used for defining reinfection. The dashed line represents the proposed optimal time window for defining a reinfection.

analysis encompassing all times during the pandemic (Figure 6). This analysis affirmed the 40-day time window as an optimal choice, independent of the wave's size or the variant that dominated the wave.

The analysis also indicated an increasingly steeper slope in viral clearance over time, especially during the first omicron wave. This trend may have been due to progressive changes in retesting requirements for individuals in isolation following documented infection, as well as the introduction of rapid antigen testing during the first omicron wave.

4 Discussion

Investigating the empirical dependence of the estimated number of reinfections in the population on the time-window duration used for defining reinfection has revealed the existence of two distinct dynamical domains, providing insights for a more effective definition of reinfection that is less susceptible to potential bias. In the first 15 days after an infection is diagnosed, labeled here as the infection clearance domain, nearly all SARS-CoV-2 positive tests are attributable to the same original infection. However, beyond the 30-day mark post-infection, within the reinfection plateau domain, nearly all positive tests are attributed to reinfections, with a negligible contribution from prolonged infections. These findings underscore that a time window of 40 days serves as an adequately conservative definition for reinfection, superseding the current conventional definition of a 90-day time window.

This conclusion emphasizes that the conventional 90-day time window is overly restrictive, resulting in significant bias in capturing reinfections and potentially leading to inaccurate or imprecise estimates in studies of reinfections, including those examining the immune protection of natural infection against reinfection. This limitation is particularly critical in the current stage of the pandemic when reinfections are common, and accurately capturing repeat reinfections is essential for a meaningful understanding of the current epidemiology of SARS-CoV-2 infection.

These findings enable the estimation of protection of natural immunity within the time window spanning from 40 to 90 days after a first infection, which is not possible under the conventional definition. Moreover, they indicate that studies assessing the protective effects of natural infection against reinfection might have overestimated this protection, especially when relying on short follow-up periods after the initial infection. It is recommended that future studies present results using both the conventional 90-day window and the proposed 40-day window to assess potential biases and elucidate the implications of adopting the new proposed time window.

An important finding from this study is the higher incidence of reinfections compared to common perception. The 90-day time window missed a proportion of reinfections relative to the 40-day time window. Instances were identified where individuals experienced up to 6 documented reinfections over the nearly 4 years of the pandemic. Given that documented reinfections represent only a fraction of the total, which includes also undocumented reinfections, this implies that reinfections are substantially underestimated. This finding suggests a resemblance in reinfection patterns across various respiratory infections, encompassing SARS-CoV-2, common-cold coronaviruses (5, 6), and influenza (7–10). Furthermore, it aligns with experimental observations derived from sequential influenza challenge studies (10). This underscores the imperative for enhanced understanding of the epidemiology of reinfections to unravel the factors contributing to the "leaky" immune protection enabling their occurrence. The incidence of reinfections increases the risk of virus mutation and evolution due to increased transmissions in the population.

This study has limitations. The definition of reinfection was deduced through the observation of infection patterns, departing

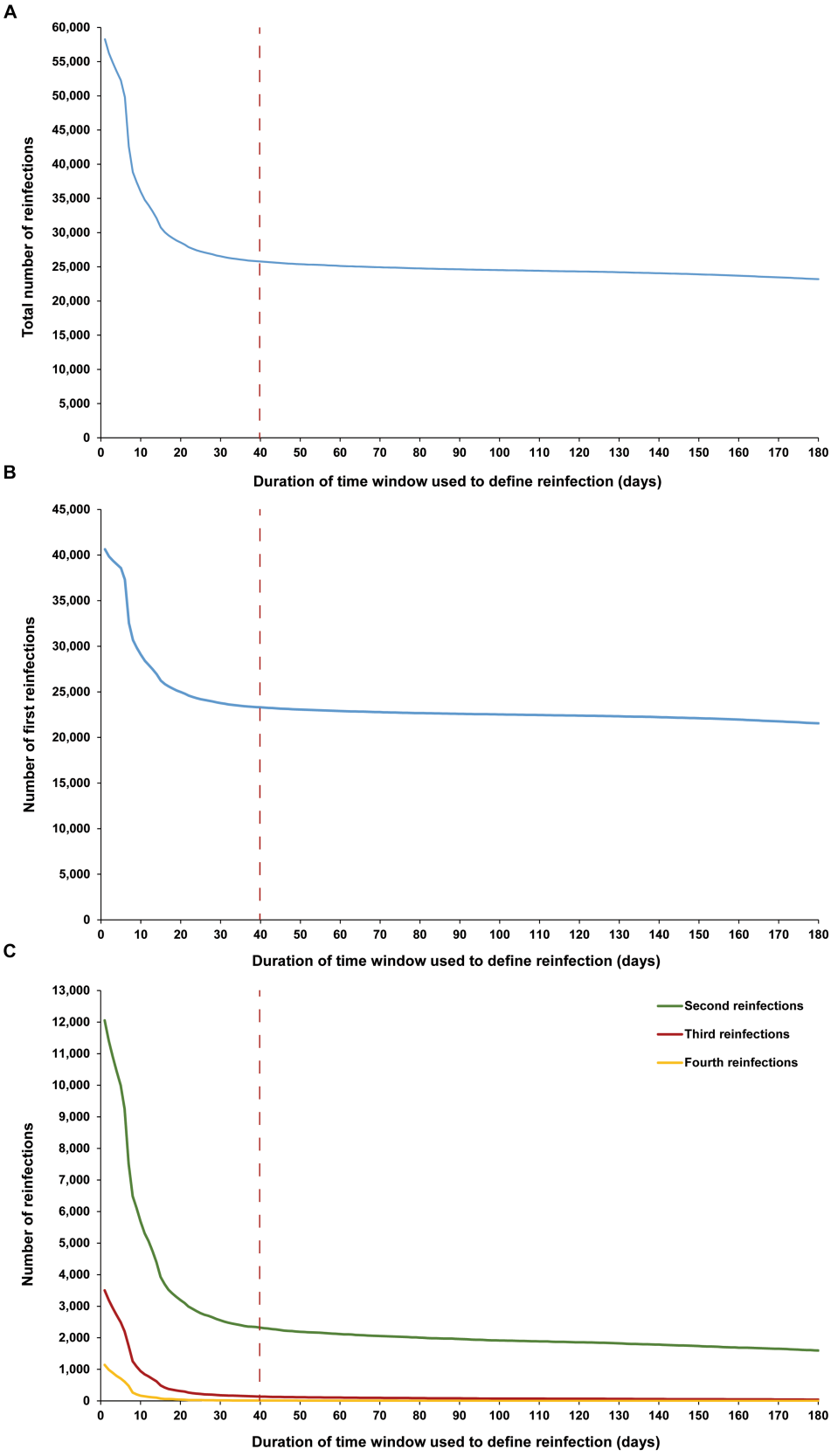

FIGURE 4
Number of **(A)** all reinfections, **(B)** first reinfections, and **(C)** second, third, and fourth reinfections among high testers in the population versus the time-window duration used for defining reinfection. High testers were defined as individuals in the top 10th percentile of the testing frequency distribution. The dashed line represents the proposed optimal time window for defining a reinfection.

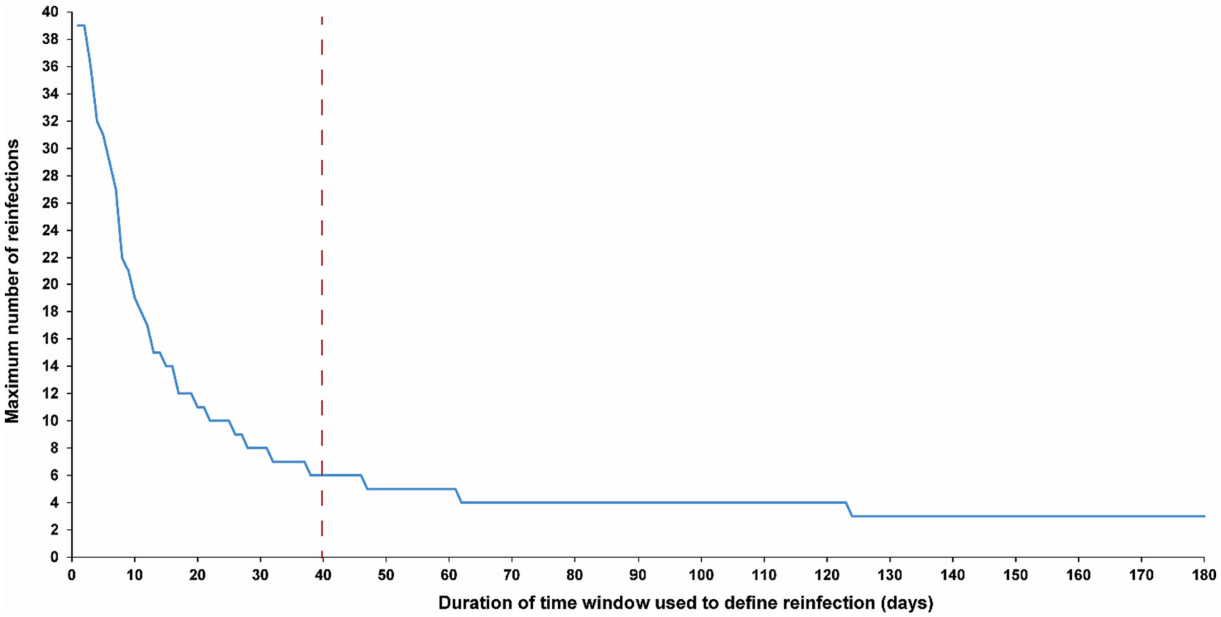

FIGURE 5
Maximum number of observed reinfections experienced by any given individual among high testers in the population versus the time-window duration used for defining reinfection. High testers were defined as individuals in the top 10th percentile of the testing frequency distribution. The dashed line represents the proposed optimal time window for defining a reinfection.

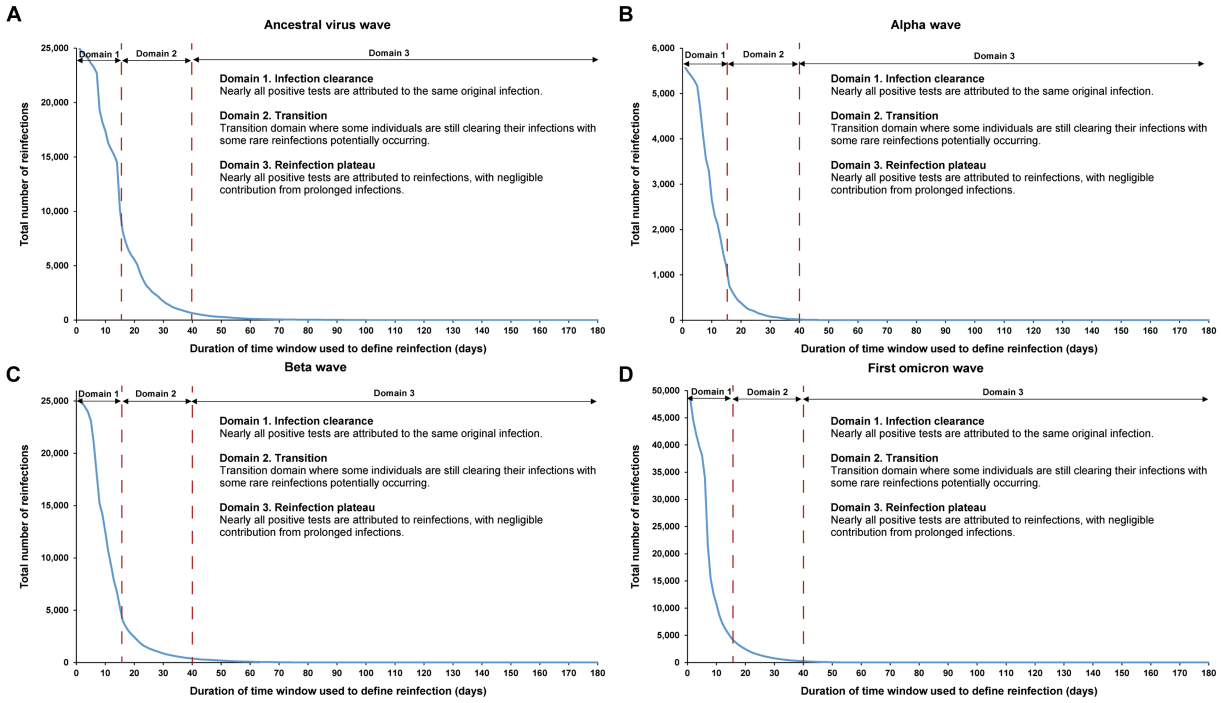

FIGURE 6
Reinfection patterns in distinct waves. Total number of reinfections in the population versus the time-window duration used for defining reinfection during the **(A)** ancestral virus wave (February 28, 2020-July 31, 2020), **(B)** alpha wave (January 18, 2021-March 7, 2021), **(C)** beta wave (March 8, 2021-May 31, 2021), and **(D)** first omicron wave (December 19, 2021-February 28, 2022). The dashed line at day 40 represents the proposed optimal time window for defining a reinfection. All curves converge to zero at large durations of the time window due to the relatively short duration of each wave in comparison to the total study duration.

from the conventional methodology of genome sequencing for every SARS-CoV-2-positive test (11–13). The latter approach entails evaluating whether the identified virus in a specific positive test differs from that detected in the preceding positive test (11–13). However, the practical implementation of the conventional approach, especially in the current stage of the pandemic, is not feasible. The outcomes of the application of the conventional approach during the early stages of the pandemic are, on the whole, consistent with the findings of the present analysis (11, 13). Notably, the conventional method has also proven to be intricate and often inconclusive when distinguishing reinfections from prolonged infections (11, 13). This complexity is exemplified in cases where only a few changes in allele frequency are observed (11, 13).

In lieu of the conventional method, we presented a novel approach, which, to the best of our knowledge, has not been previously utilized for either SARS-CoV-2 infection or any other infection. The conceptual foundation of this approach stems from recognizing two specific population distributions behind the relationship between the estimated number of reinfections and the duration of the time window used for defining reinfection. The first distribution pertains to the clearance of the infection, while the second relates to the incidence of reinfection. The observation of two discernible dynamical domains, along with a transition region between them, strongly implies that the clearance of infection dominates the first domain, while the distribution of reinfection incidence dominates the second domain.

The present analysis was conducted on the population of Qatar, characterized by a predominantly young demographic composition. Consequently, the findings may lack generalizability to other countries where elderly citizens constitute a more substantial proportion of the population. The reliance on documented infections may introduce bias, as patterns for undocumented infections may differ from those documented. Furthermore, variations in testing frequency across different population segments and over time can lead to fluctuations in the likelihood of documenting infections. High testers may not be representative of the broader population due to self-selection influenced by factors such as perceived risk associated with occupation, living conditions, vaccination status, or underlying health conditions.

However, a strength of this study lies in its comprehensive scope, encompassing the entire population of a country with high testing rates. This approach enhances the capture of infections and reinfections, contributing to the robustness of the study's findings. Additionally, the sensitivity analysis, focusing exclusively on high testers less impacted by undocumented infections, generated similar results, suggesting that the study findings are less likely to be influenced by undocumented infections.

5 Conclusion

A 40-day time window serves as an appropriately conservative definition for reinfection, offering an informed alternative to the current conventional 90-day time window. The latter, shown to be unnecessarily restrictive, introduces bias in reinfection capture that may jeopardize estimates in reinfection studies. Contrary to common perception, reinfections are more prevalent, with some individuals experiencing multiple instances since the onset of the pandemic. A nuanced understanding of the factors contributing to the "leaky" immune protection allowing for this heightened incidence of reinfections is warranted.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary materials, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by Hamad Medical Corporation and Weill Cornell Medicine–Qatar Institutional Review Boards. 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

HC: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. HA: Data curation, Writing – review & editing. PT: Data curation, Writing – review & editing. HMY: Data curation, Writing – review & editing. AAAT: Data curation, Writing – review & editing. MRH: Data curation, Writing – review & editing. PC: Data curation, Writing – review & editing. ZA-K: Data curation, Writing – review & editing. EA-K: Data curation, Writing – review & editing. AJ: Data curation, Writing – review & editing. AHK: Data curation, Writing – review & editing. ANL: Data curation, Writing – review & editing. RS: Data curation, Writing – review & editing. HFA-R: Data curation, Writing – review & editing. GKN: Data curation, Writing – review & editing. MGA-K: Data curation, Writing – review & editing. AAB: Data curation, Writing – review & editing. HEA-R: Data curation, Writing – review & editing. MHA-T: Data curation, Writing – review & editing. AA-K: Data curation, Writing – review & editing. RB: Data curation, Writing – review & editing. LJA-R: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The authors are grateful for institutional salary support from the Biomedical Research Program and the Biostatistics, Epidemiology, and Biomathematics

Research Core, both at Weill Cornell Medicine-Qatar, as well as for institutional salary support provided by the Ministry of Public Health, Hamad Medical Corporation, and Sidra Medicine. The authors are also grateful for the Qatar Genome Programme and Qatar University Biomedical Research Center for institutional support for the reagents needed for the viral genome sequencing. HA acknowledges the support of Qatar University Internal Grant ID QUCG-CAS-23/24–114. The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the article. Statements made herein are solely the responsibility of the authors. AAB has received institutional grant funding from Gilead Sciences unrelated to the work presented in this paper.

Acknowledgments

We acknowledge the many dedicated individuals at Hamad Medical Corporation, the Ministry of Public Health, the Primary Health Care Corporation, Qatar Biobank, Sidra Medicine, and Weill Cornell Medicine-Qatar for their diligent efforts and contributions to make this study possible.

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2024.1363045/full#supplementary-material

References

1. Chemaitelly H, Nagelkerke N, Ayoub HH, Coyle P, Tang P, Yassine HM, et al. Duration of immune protection of SARS-CoV-2 natural infection against reinfection. *J Travel Med*. (2022) 29:taac109. doi: 10.1093/jtm/taac109

2. Chemaitelly H, Tang P, Coyle P, Yassine HM, Al-Khatib HA, Smatti MK, et al. Protection against reinfection with the omicron BA.2.75 subvariant. *N Engl J Med*. (2023) 388:665–7. doi: 10.1056/NEJMc2214114

3. Altarawneh HN, Chemaitelly H, Hasan MR, Ayoub HH, Qassim S, AlMukdad S, et al. Protection against the omicron variant from previous SARS-CoV-2 infection. *N Engl J Med*. (2022) 386:1288–90. doi: 10.1056/NEJMc2200133

4. Chemaitelly H, Ayoub HH, Coyle P, Tang P, Yassine HM, Al-Khatib HA, et al. Protection of omicron sub-lineage infection against reinfection with another omicron sub-lineage. *Nat Commun*. (2022) 13:4675. doi: 10.1038/s41467-022-32363-4

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

6. Lavine JS, Bjornstad ON, Antia R. Immunological characteristics govern the transition of COVID-19 to endemicity. *Science*. (2021) 371:741–5. doi: 10.1126/science.abe6522

7. Ferguson NM, Galvani AP, Bush RM. Ecological and immunological determinants of influenza evolution. *Nature*. (2003) 422:428–33. doi: 10.1038/nature01509

8. Frank AL, Taber LH. Variation in frequency of natural reinfection with influenza a viruses. *J Med Virol*. (1983) 12:17–23. doi: 10.1002/jmv.1890120103

9. Wang J, Jiang L, Xu Y, He W, Zhang C, Bi F, et al. Epidemiology of influenza virus reinfection in Guangxi, China: a retrospective analysis of a nine-year influenza surveillance data: characteristics of influenza virus reinfection. *Int J Infect Dis*. (2022) 120:135–41. doi: 10.1016/j.ijid.2022.04.045

10. Memoli MJ, Han A, Walters KA, Czajkowski L, Reed S, Athota R, et al. Influenza a reinfection in sequential human challenge: implications for protective immunity and "universal" vaccine development. *Clin Infect Dis*. (2020) 70:748–53. doi: 10.1093/cid/ciz281

11. Abu-Raddad LJ, Chemaitelly H, Malek JA, Ahmed AA, Mohamoud YA, Younuskunju S, et al. Assessment of the risk of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) reinfection in an intense reexposure setting. *Clin Infect Dis*. (2021) 73:e1830–40. doi: 10.1093/cid/ciaa1846

12. Choudhary MC, Crain CR, Qiu X, Hanage W, Li JZ. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) sequence characteristics of coronavirus disease 2019 (COVID-19) persistence and reinfection. *Clin Infect Dis*. (2022) 74:237–45. doi: 10.1093/cid/ciab380

13. Abu-Raddad LJ, Chemaitelly H, Coyle P, Malek JA, Ahmed AA, Mohamoud YA, et al. SARS-CoV-2 antibody-positivity protects against reinfection for at least seven months with 95% efficacy. *EClinicalMedicine*. (2021) 35:100861. doi: 10.1016/j.eclinm.2021.100861

14. Kojima N, Shrestha NK, Klausner JD. A systematic review of the protective effect of prior SARS-CoV-2 infection on repeat infection. *Eval Health Prof*. (2021) 44:327–32. doi: 10.1177/01632787211047932

15. Pilz S, Theiler-Schwetz V, Trummer C, Krause R, Ioannidis JPA. SARS-CoV-2 reinfections: overview of efficacy and duration of natural and hybrid immunity. *Environ Res*. (2022) 209:112911. doi: 10.1016/j.envres.2022.112911

16. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N Engl J Med*. (2020) 382:1177–9. doi: 10.1056/NEJMc2001737

17. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med*. (2020) 26:672–5. doi: 10.1038/s41591-020-0869-5

18. Abu-Raddad LJ, Chemaitelly H, Malek JA, Ahmed AA, Mohamoud YA, Younuskunju S, et al. Two prolonged viremic SARS-CoV-2 infections with conserved viral genome for two months. *Infect Genet Evol*. (2021) 88:104684. doi: 10.1016/j.meegid.2020.104684

19. Corey L, Beyrer C, Cohen MS, Michael NL, Bedford T, Rolland M. SARS-CoV-2 variants in patients with immunosuppression. *N Engl J Med*. (2021) 385:562–6. doi: 10.1056/NEJMsb2104756

20. Cele S, Karim F, Lustig G, San JE, Hermanus T, Tegally H, et al. SARS-CoV-2 prolonged infection during advanced HIV disease evolves extensive immune escape. *Cell Host Microbe*. (2022) 30:154–162.e5. doi: 10.1016/j.chom.2022.01.005

21. Abu-Raddad LJ, Chemaitelly H, Ayoub HH, Coyle P, Malek JA, Ahmed AA, et al. Introduction and expansion of the SARS-CoV-2 B.1.1.7 variant and reinfections in Qatar: a nationally representative cohort study. *PLoS Med*. (2021) 18:e1003879. doi: 10.1371/journal.pmed.1003879

22. Chemaitelly H, Ayoub HH, Tang P, Coyle P, Yassine HM, Thani AAA, et al. patterns in repeat reinfections: pre and post omicron emergence. *medRxiv*. (2023) 29:23292041. doi: 10.1101/2023.06.29.23292041

23. Chemaitelly H, Ayoub HH, Tang P, Hasan MR, Coyle P, Yassine HM, et al. Immune imprinting and protection against repeat reinfection with SARS-CoV-2. *N Engl J Med*. (2022) 387:1716–8. doi: 10.1056/NEJMc2211055

24. Mahmoud MA, Ayoub HH, Coyle P, Tang P, Hasan MR, Yassine HM, et al. SARS-CoV-2 infection and effects of age, sex, comorbidity, and vaccination among older individuals: a national cohort study. *Influenza Other Respir Viruses*. (2023) 17:e13224. doi: 10.1111/irv.13224

25. Subissi L, von Gottberg A, Thukral L, Worp N, Oude Munnink BB, Rathore S, et al. An early warning system for emerging SARS-CoV-2 variants. *Nat Med*. (2022) 28:1110–5. doi: 10.1038/s41591-022-01836-w

26. Chemaitelly H, Tang P, Hasan MR, AlMukdad S, Yassine HM, Benslimane FM, et al. Waning of BNT162b2 vaccine protection against SARS-CoV-2 infection in Qatar. *N Engl J Med*. (2021) 385:e83. doi: 10.1056/NEJMoa2114114

27. Altarawneh HN, Chemaitelly H, Ayoub HH, Tang P, Hasan MR, Yassine HM, et al. Effects of previous infection and vaccination on symptomatic omicron infections. *N Engl J Med*. (2022) 387:21–34. doi: 10.1056/NEJMoa2203965

28. Chemaitelly H, Ayoub HH, AlMukdad S, Faust JS, Tang P, Coyle P, et al. Bivalent mRNA-1273.214 vaccine effectiveness against SARS-CoV-2 omicron XBB* infections. *J Travel Med*. (2023) 30:taad106. doi: 10.1093/jtm/taad106

29. Chemaitelly H, Ayoub HH, Tang P, Coyle P, Yassine HM, Al Thani AA, et al. Long-term COVID-19 booster effectiveness by infection history and clinical vulnerability and immune imprinting: a retrospective population-based cohort study. *Lancet Infect Dis*. (2023) 23:816–27. doi: 10.1016/S1473-3099(23)00058-0

30. Abu-Raddad LJ, Chemaitelly H, Bertollini R. National Study Group for Covid vaccination. Effectiveness of mRNA-1273 and BNT162b2 vaccines in Qatar. *N Engl J Med*. (2022) 386:799–800. doi: 10.1056/NEJMc2117933

31. Abu-Raddad LJ, Chemaitelly H, Ayoub HH, Al Kanaani Z, Al Khal A, Al Kuwari E, et al. Characterizing the Qatar advanced-phase SARS-CoV-2 epidemic. *Sci Rep*. (2021) 11:6233. doi: 10.1038/s41598-021-85428-7

32. Al-Thani MH, Farag E, Bertollini R, Al Romaihi HE, Abdeen S, Abdelkarim A, et al. SARS-CoV-2 infection is at herd immunity in the majority segment of the population of Qatar. *Infect Dis*. (2021) 8:ofab221. doi: 10.1093/ofid/ofab221

33. Jeremijenko A, Chemaitelly H, Ayoub HH, Alishaq M, Abou-Samra AB, Al Ajmi J, et al. Herd immunity against severe acute respiratory syndrome coronavirus 2 infection in 10 communities. *Emerg Infect Dis*. (2021) 27:1343–52. doi: 10.3201/eid2705.204365

34. AlNuaimi AA, Chemaitelly H, Semaan S, AlMukdad S, Al-Kanaani Z, Kaleeckal AH, et al. All-cause and COVID-19 mortality in Qatar during the COVID-19 pandemic. *BMJ Glob Health*. (2023) 8:e012291. doi: 10.1136/bmjgh-2023-012291

35. Ayoub HH, Chemaitelly H, Seedat S, Makhoul M, Al Kanaani Z, Al Khal A, et al. Mathematical modeling of the SARS-CoV-2 epidemic in Qatar and its impact on the national response to COVID-19. *J Glob Health*. (2021) 11:05005. doi: 10.7189/jogh.11.05005

36. Coyle PV, Chemaitelly H, Ben Hadj Kacem MA, Abdulla Al Molawi NH, El Kahlout RA, Gilliani I, et al. SARS-CoV-2 seroprevalence in the urban population of Qatar: an analysis of antibody testing on a sample of 112,941 individuals. *iScience*. (2021) 24:102646. doi: 10.1016/j.isci.2021.102646

37. Abu-Raddad LJ, Chemaitelly H, Ayoub HH, AlMukdad S, Yassine HM, Al-Khatib HA, et al. Effect of mRNA vaccine boosters against SARS-CoV-2 omicron infection in Qatar. *N Engl J Med*. (2022) 386:1804–16. doi: 10.1056/NEJMoa2200797

38. Chemaitelly H, Faust JS, Krumholz HM, Ayoub HH, Tang P, Coyle P, et al. Short -and longer-term all-cause mortality among SARS-CoV-2 -infected individuals and the pull-forward phenomenon in Qatar: a national cohort study. *Int J Infect Dis*. (2023) 136:81–90. doi: 10.1016/j.ijid.2023.09.005

39. Chemaitelly H, Bertollini R, Abu-Raddad LJ. National Study Group for Covid epidemiology. Efficacy of natural immunity against SARS-CoV-2 reinfection with the Beta variant. *N Engl J Med*. (2021) 385:2585–6. doi: 10.1056/NEJMc2110300

40. Abu-Raddad LJ, Chemaitelly H, Ayoub HH, Tang P, Coyle P, Hasan MR, et al. Relative infectiousness of SARS-CoV-2 vaccine breakthrough infections, reinfections, and primary infections. *Nat Commun*. (2022) 13:532. doi: 10.1038/s41467-022-28199-7

41. Qassim SH, Hasan MR, Tang P, Chemaitelly H, Ayoub HH, Yassine HM, et al. Effects of SARS-CoV-2 alpha, Beta, and Delta variants, age, vaccination, and prior infection on infectiousness of SARS-CoV-2 infections. *Front Immunol*. (2022) 13:984784. doi: 10.3389/fimmu.2022.984784

42. Qassim SH, Chemaitelly H, Ayoub HH, AlMukdad S, Tang P, Hasan MR, et al. Effects of BA.1/BA.2 subvariant, vaccination and prior infection on infectiousness of SARS-CoV-2 omicron infections. *J Travel Med*. (2022) 29:taac068. doi: 10.1093/jtm/taac068

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 21 March 2024
DOI 10.3389/fmed.2024.1362192

OPEN ACCESS

EDITED BY
Zhongshan Cheng,
St. Jude Children's Research Hospital,
United States

REVIEWED BY
Jinyong Pang,
University of South Florida, United States
Haisheng Zhang,
University of East-West Medicine,
United States
Xing Wei,
Sichuan University, China

*CORRESPONDENCE
Martin Wolkewitz
✉ martin.wolkewitz@uniklinik-freiburg.de

RECEIVED 27 December 2023
ACCEPTED 20 February 2024
PUBLISHED 21 March 2024

CITATION
Martinuka O, Hazard D, Marateb HR, Mansourian M, Mañanas MÁ, Romero S, Rubio-Rivas M and Wolkewitz M (2024) Methodological biases in observational hospital studies of COVID-19 treatment effectiveness: pitfalls and potential. *Front. Med.* 11:1362192. 10.3389/fmed.2024.1362192

Methodological biases in observational hospital studies of COVID-19 treatment effectiveness: pitfalls and potential

Oksana Martinuka[1], Derek Hazard[1], Hamid Reza Marateb[2,3], Marjan Mansourian[2,4], Miguel Ángel Mañanas[2,5], Sergio Romero[2,5], Manuel Rubio-Rivas[6] and Martin Wolkewitz[1]*

[1]Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center - University of Freiburg, Freiburg, Germany, [2]Biomedical Engineering Research Center (CREB), Automatic Control Department (ESAII), Universitat Politècnica de Catalunya-Barcelona Tech (UPC), Barcelona, Spain, [3]Department of Artificial Intelligence, Smart University of Medical Sciences, Tehran, Iran, [4]Department of Epidemiology and Biostatistics, School of Health, Isfahan University of Medical Sciences, Isfahan, Iran, [5]CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, Spain, [6]Department of Internal Medicine, Bellvitge University Hospital, Hospitalet de Llobregat, Barcelona, Spain

Introduction: This study aims to discuss and assess the impact of three prevalent methodological biases: competing risks, immortal-time bias, and confounding bias in real-world observational studies evaluating treatment effectiveness. We use a demonstrative observational data example of COVID-19 patients to assess the impact of these biases and propose potential solutions.

Methods: We describe competing risks, immortal-time bias, and time-fixed confounding bias by evaluating treatment effectiveness in hospitalized patients with COVID-19. For our demonstrative analysis, we use observational data from the registry of patients with COVID-19 who were admitted to the Bellvitge University Hospital in Spain from March 2020 to February 2021 and met our predefined inclusion criteria. We compare estimates of a single-dose, time-dependent treatment with the standard of care. We analyze the treatment effectiveness using common statistical approaches, either by ignoring or only partially accounting for the methodological biases. To address these challenges, we emulate a target trial through the clone-censor-weight approach.

Results: Overlooking competing risk bias and employing the naïve Kaplan-Meier estimator led to increased in-hospital death probabilities in patients with COVID-19. Specifically, in the treatment effectiveness analysis, the Kaplan-Meier estimator resulted in an in-hospital mortality of 45.6% for treated patients and 59.0% for untreated patients. In contrast, employing an emulated trial framework with the weighted Aalen-Johansen estimator, we observed that in-hospital death probabilities were reduced to 27.9% in the "X"-treated arm and 40.1% in the non-"X"-treated arm. Immortal-time bias led to an underestimated hazard ratio of treatment.

Conclusion: Overlooking competing risks, immortal-time bias, and confounding bias leads to shifted estimates of treatment effects. Applying the naïve Kaplan-Meier method resulted in the most biased results and overestimated probabilities for the primary outcome in analyses of hospital data from COVID-19 patients. This overestimation could mislead clinical decision-making. Both immortal-

time bias and confounding bias must be addressed in assessments of treatment effectiveness. The trial emulation framework offers a potential solution to address all three methodological biases.

KEYWORDS

competing risks, confounding, COVID-19, emulated trial, immortal-time bias, methodological bias, treatment effectiveness

Introduction

During the coronavirus disease 2019 (COVID-19) pandemic, routinely collected observational data has become crucial for comparative treatment effectiveness research and for identifying potential therapeutic options (1, 2). Real-world observational data was increasingly used during the pandemic's first waves when results from randomized clinical trials were either unavailable or used to complement trial findings. Observational studies can yield biased results when they are not appropriately designed and analyzed because of their type of data and potential methodological challenges (1, 3–5). While the methodological limitations of observational data have been extensively discussed, a review of early observational studies on the effectiveness of repurposed or novel treatments for COVID-19 patients indicated that fundamental methodological biases such as competing risks, immortal-time bias, and confounding bias, either alone or in combination, were still often overlooked (2). Failure to address these methodological biases can result in skewed estimates of treatment effects and, consequently, incorrect conclusions (2, 5).

A competing risk is an event that precludes the observation of the primary event of interest (6, 7). In COVID-19 studies, when in-hospital mortality is the primary outcome, discharge becomes a competing event because it hinders the observation of death in hospital (8). Conventional survival analysis techniques, such as the naïve Kaplan-Meier estimator, treat competing events as right-censored observations. This approach assumes that censored individuals will have the same probability of experiencing the event of interest as those who remain in the risk set, leading to a positive event probability instead of zero probability after the occurrence of a competing event (7, 9–11). For comprehensive mathematical proofs, we refer to the studies conducted by Zhang (11) and Coemans et al. (10). In the context of COVID-19 and analyzing in-hospital death, this assumption would imply that discharged patients have a similar risk of death as those still hospitalized, which is not clinically meaningful (7, 12). Hence, the independent censoring assumption is violated for hospital discharge because discharged patients are usually in better health conditions than those still hospitalized (13). In the presence of competing events, the naïve Kaplan-Meier method can lead to biased estimates and erroneous conclusions (13). Notably, the issue of competing risks can arise in analyzing time-to-event survival data in randomized clinical trials, observational studies, and target trial emulations (6).

Observational studies often evaluate the effectiveness of time-dependent treatments, meaning patients may initiate treatment at different times after their study entry (14). Immortal time occurs when there is a delay between cohort entry and treatment initiation, during which patients are precluded from experiencing the outcome. Misclassifying or excluding this pre-treatment period can introduce immortal-time bias, thereby biasing the estimated treatment effects (15–17). Previous studies have demonstrated that the most severe form of immortal-time bias occurs when studies incorrectly include immortal time, assuming that treated patients are at risk from the baseline. This is in contrast to methods designed to mitigate this bias, such as landmark analysis, the exposure density sampling method, and the time-dependent Cox model with time-varying treatment status (18–20). When immortal time is mistakenly included, it leads to an artificially reduced observed event rate for the treatment group and an artificially inflated event rate for the control group (14, 21). As a result, the hazard ratio (HR) for comparing the treatment vs. the control group may be underestimated (20). For negative outcomes like death, such underestimation misleadingly suggests a greater treatment effectiveness. In contrast, for positive outcomes like discharge, the underestimation of the treatment effect can make the treatment appear less effective. For a comprehensive review of the mathematical proofs, we refer to the studies conducted by Suissa (20), Beyersmann et al. (22), and the simulation study by Wang et al. (19).

Confounding represents another well-known and significant challenge in observational studies, arising from an unequal distribution of patient characteristics between treatment and control groups, which affect both treatment decision and outcome (23, 24). Therefore, simply comparing outcomes between the treatment and control groups without any adjustment can lead to biased estimates of treatment effects (25, 26). In causal analyses, common approaches to adjust for baseline characteristics include inverse probability weighting, standardization, and stratification-based adjustment methods such as stratification and matching methods (27, 28).

Throughout the COVID-19 pandemic, the target trial emulation framework was widely used to assess the effectiveness of treatments and vaccines using real-world data, particularly in the pandemic's early stages (29–32). This framework applies the principles of randomized clinical trials to emulate a hypothetical trial using observational data, thereby answering specific causal questions (24, 33). It has become crucial to explore treatment effects and address common methodological biases (34). While previous research has demonstrated that target trial emulation can handle both immortal-time bias and confounding bias, our study further confirms the importance of considering competing risks within observational data (19, 34).

The aim of this study is 3-fold: (i) to provide an overview of the three most common methodological biases in observational hospital data; (ii) to evaluate the impact of each bias using a typical example of observational hospital data and applying various analytical methodologies; and (iii) to describe the target trial emulation framework that addresses these potential methodological challenges.

For illustrative purposes, we analyzed observational hospital data from patients with COVID-19. This article provides an explanation of the potential methodological pitfalls in a descriptive manner and proposes alternative strategies for mitigating these challenges.

Methods

The Methods section is organized as follows: we introduce challenges associated with competing risks through a typical example of observational hospital data of COVID-19 patients and conduct a time-to-event analysis without accounting for the patient's treatment status. We then describe a cohort of patients used for our demonstrative analyses and introduce the concept of target trial emulation. Next, we discuss immortal-time and confounding biases, outline standard analysis methods prone to bias, and explain how these challenges can be mitigated within the emulated trial framework. We define the five models used for comparison to determine the impact of immortal-time bias and confounding bias. We emphasize that all analyses conducted, including the emulated trial, were demonstrative, and an assessment of clinical treatment effects was beyond the scope of this study.

Motivating example: competing risks in a COVID-19 hospital setting

To illustrate the concept of competing risks in time-to-event analysis of hospital data, we conducted an analysis using longitudinal patient-level data from a cohort of COVID-19 patients ($n = 478$) hospitalized at the Bellvitge University Hospital in Barcelona, Spain, from March 2020 to February 2021. These patients experienced various endpoints, including in-hospital death, discharge home, or transfer to another healthcare facilities. In this analysis, we defined in-hospital death as the primary outcome of interest and estimated the cumulative probabilities without considering the patient's treatment status. Information on patient survival status beyond the follow-up period was not available.

In the naïve analysis, we calculated the cumulative probabilities using the one minus Kaplan-Meier estimator. We compared these results with those from the Fine-Gray analysis approach, which accounts for competing events like hospital discharge by keeping patients in the risk set until the end of follow-up. The Fine-Gray model is a direct model for cumulative incidence functions in the presence of competing risks (35). We conducted two Fine-Gray analyses. In the first analysis, we treated patients discharged home as a competing event and considered patients transferred to other facilities as censored observations, thus implementing the Fine-Gray model with two events. In the second analysis, we distinguished between reasons for hospital discharge, categorizing discharge to home and transfer to another healthcare facility as separate competing events. This approach allowed us to maintain both outcomes in the risk set, corresponding to the Fine-Gray model with three events.

Using the naïve Kaplan-Meier estimator resulted in an overestimated in-hospital death probability of 55.3% (Figure 1). By recognizing discharge home as the only competing event and by censoring transferred patients, the probability of in-hospital death dropped to 43.3% (Figure 1). Finally, by considering both reasons for hospital discharge, the in-hospital death probability substantially decreased to 38.1% (Figure 1). These findings underscore the importance of recognizing and addressing competing risks in the hospital data and have also motivated us to explore the future extensions of emulated target trial methodologies.

Illustrative study population: patients with COVID-19

For this case study, we analyzed longitudinal data from hospitalized patients with COVID-19 as described above. A total of 478 patients with moderate-to-severe COVID-19 were included. Inclusion criteria for these patients were based on the Horowitz index, a ratio of the partial pressure of oxygen to the fraction of inspired oxygen (PaO_2/FiO_2) of less than 300 mmHg measured at hospital admission and the presence of at least one inflammation-related high-risk factor: C-reactive protein (>102 mg/L), lactate dehydrogenase (>394 U/L), D-dimer (>1,580 ng/mL), total lymphocyte count ($<760 \times 10^6$/L), and ferritin (>1,360 mcg/L) at the time of admission. The high-risk categories were determined following the criteria and classification established by Rubio-Rivas et al. (36). For all patients, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection was confirmed via PCR testing. The study follow-up period was 45 days post-hospital admission. Patients with no outcomes who were still alive at the end of this period were administratively censored ($n = 59$, 12.3%).

Trial emulation: study question and protocol components

To emulate a target trial, we defined our clinical aim as follows: *to evaluate the effectiveness of treatment "X" compared to the standard-of-care, which does not involve the administration of treatment "X," on the risk of in-hospital death while acknowledging its effects on hospital discharge outcomes in COVID-19 patients.* This question of interest could be subdivided into three distinct components: *assessing the impact of treatment on (i) in-hospital death, (ii) discharge alive home, and (iii) transfer to another healthcare facility*. We designed a hypothetical study protocol, specifying its components including eligibility criteria, treatment strategies and assignment, start and end of follow-up, endpoints, and causal contrast (Supplementary material 1).

Immortal-time bias

In studies evaluating time-varying or time-dependent treatments addressing immortal-time bias is crucial, for which several options are available. Two commonly used approaches that can lead to severe immortal-time bias and result in flawed estimates of treatment effect: (i) including person-time and classifying patients as treated from time zero, even if they receive treatment later during follow-up, and (ii) excluding person-time, which is the time from baseline to treatment initiation for the exposure group (16, 19, 37, 38). The landmark analysis is a design-based method involving setting fixed time as the landmark time and classifying patients according to their treatment status at the landmark (17). Patients are then followed from

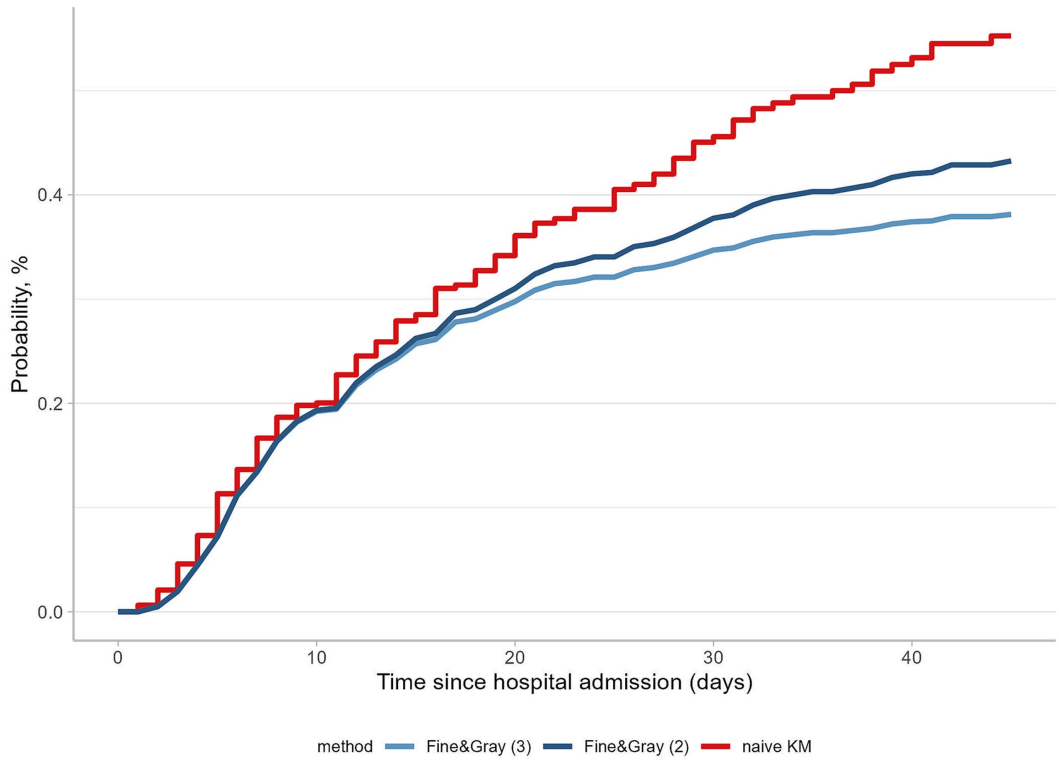

FIGURE 1
Probabilities of in-hospital death with and without accounting for competing events. Probabilities of in-hospital death are calculated taking different analytical approaches: the Fine-Gray (3) model, considering three outcomes; the Fine-Gray (2) model, considering two outcomes, and the naïve analysis using the one minus the Kaplan-Meier estimator.

the landmark time regardless of subsequent changes in their treatment status (17, 37). However, this approach has two principal limitations: (i) the choice of the landmark time and (ii) the exclusion of patients who had an outcome before the landmark time from the analysis (15, 28). To overcome these drawbacks, considering multiple landmarks and a pooled analysis via a supermodel is recommended (39). In the exposure density sampling method, unexposed patients are matched to exposed patients with respect to a time-dependent exposure. Specifically, for each exposed patient, one or more unexposed patients who have survived for a duration equivalent to that of the exposed patient are selected (40). This approach allows for the possibility that an unexposed patient may change their exposure status after matching. A simulation study demonstrated that the exposure density sampling method fully addressed immortal-time bias (40), in contrast to the simpler method of prescription time-distribution matching (18, 41). Another common approach to account for immortal person-time is to use a time-dependent model (16, 18, 19). It involves modeling time-varying treatment status and includes it as a time-dependent covariate in a proportional hazards or another regression model (19). This approach enables the classification of patients as "treated" or "untreated" on each follow-up day, allowing for the reclassification of patients from "untreated" to "treated" status upon the treatment's initiation. Alternatively, clone-censor-weight and the sequential trial approaches allow for the incorporation of time-dependent treatment status through duplication or a nested design, and can be applied within the framework of trial emulation. The cloning approach creates two exact copies of each patient, assigning one clone to the treatment and the other to the control arm. Subsequently, a clone in each arm is censored when the actual treatment received deviates from the treatment strategy of the arm to which it was initially assigned (34). This usually requires defining a clinically meaningful grace period (33, 34). In the sequential trial approach, a sequence of multiple nested trials with all potential time zeros is modeled (37). Each method has its own assumptions and limitations, which should be considered when interpreting study results. Our study focuses on three approaches: analysis that includes immortal time, modeling time-varying treatment status and using time-dependent Cox regression model, and the clone-censor-weight approach.

In our illustrative observational data example, time zero, or the baseline, was defined as hospital admission, with the possibility of administering treatment at a later follow-up time. Consequently, patients' treatment status depended on their presence in the risk set until a specific time. To evaluate the impact of included immortal time, we initially conducted a naïve analysis, mistakenly categorizing patients who received treatment during follow-up as having been treated since hospital admission (Model 1, Table 1 in the Results). In this instance, the time period between hospital admission and "X" treatment administration is immortal, as patients must be outcome-free to be categorized as treated (16). We also performed a time-dependent Cox regression analysis by modeling a time-varying treatment status using start-stop notation (Models 2–4). We used the

clone-censor-weight approach for the target trial emulation, defining the grace period as treatment administration within 2 days of hospital admission (Model 5), as elaborated in Supplementary material 2. The length of this period was based on clinical relevance. We defined two treatment strategies: (1) administration of "X" treatment during the first 2 days of hospital admission, referred to as the "X"-treated arm, and (2) no administration of "X" treatment during the first 2 days, referred to the non-"X"-treated arm. Patients who experienced outcome events within 2 days were included in both treatment arms, avoiding immortal-time bias (34).

Confounding bias

After identifying and collecting all important variables—potential confounders, several statistical approaches can be considered to mitigate confounding bias. We included the following patient baseline covariates in our study: age, sex, Charlson Comorbidity Index, levels of C-reactive protein, lactate dehydrogenase, D-dimer, total lymphocyte count, ferritin, and calendar time of hospital admission, categorized according to the pandemic waves. After examining the distribution of inflammatory variables, we applied the log and square root transformations to these variables to reduce the influence of extreme values. We assumed all these measured covariates were sufficient for controlling baseline confounding.

We first performed a univariable analysis without adjusting for baseline covariates to demonstrate the impact of ignoring time-fixed confounding (Models 1 and 2, Table 1). We then included the baseline covariates into a Cox regression model and performed multivariable analysis (Model 3). We also employed an inverse probability of treatment weighting model based on propensity scores to balance baseline covariates in the treatment and control groups (Model 4) (42). To balance the patient's characteristics and prognostic covariates between treated and untreated groups, we re-weighted the outcome variables of these patients by the inverse probability of the treatment received (28, 43). As a result, we re-weighted the patients and created a pseudo-population free of confounding (42). We used the ipw package and calculated robust standard errors (44). In emulated trial analysis, we applied the clone-censor-weight approach (Model 5). Cloning patients into two arms ensured that the two arms were balanced regarding baseline covariates, addressing time-fixed confounding bias (34, 45). Additionally, to correct for selection bias resulting from artificial censoring, we estimated inverse probability of censoring weights (34). We applied the code as presented by Maringe et al. (34) for the target trial emulation analysis. Standardized differences were assessed before and after applying inverse probability of censoring weighting (Supplementary material 3). For this model, nonparametric bootstrap was used to compute 95% normal-based confidence intervals (CI) with 500 bootstrap replications. Multiple imputations were performed to replace missing values for inflammatory covariates measured at baseline. All analysis steps were applied to the five copies of the imputed datasets. Further details on the multiple imputation analysis are found in Supplementary material 4. All statistical analyses were performed in RStudio (2022.07.1) software (46).

Results

Patients characteristics

Overall, among the 478 patients with COVID-19 included in our initial data analysis, 183 (38.3%) experienced in-hospital death, 237 (49.6%) were discharged from the hospital, and 59 (12.3%) were administratively censored at the end of the 45-day follow-up period. Among the 237 discharged patients, 139 (58.6%) were discharged to their homes, while 98 (41.4%) were transferred to other healthcare facilities. In total, 143 (29.9%) patients were treated with "X" treatment at any time during the follow-up period. In the emulated trial analysis, 73 (15.3%) patients received the "X" treatment within 2 days. Among those who received the treatment, 20 died, 26 were discharged home, and 19 were transferred to other healthcare facilities. The cohort's characteristics are detailed in Supplementary material 5.

Assessing the impact of treatment on in-hospital death rates

We calculated the cumulative incidence probabilities for in-hospital death by ignoring or accounting for competing events. Probabilities were derived using the conventional, naïve Kaplan-Meier estimator applied to the crude dataset, which was susceptible to all three biases. These results were compared to probabilities estimated from the weighted version of the Aalen-Johansen estimator used in the emulated analysis with the clone-censor-weight approach. The cumulative probabilities of in-hospital death using the naïve Kaplan-Meier estimator were 45.6% for the treated and 59.0% for the untreated group at the end of the 45-day follow-up period. In contrast, the Aalen-Johansen estimator revealed cumulative probabilities of 27.9% for the "X"-treated arm and 40.1% for the non-"X"-treated arm (Figure 2).

Estimating treatment effects with and without addressing immortal time and confounding biases

We estimated the treatment effect while either ignoring or acknowledging immortal time and confounding biases, taking different approaches for three endpoints (Table 1). Model 1, which ignored both immortal-time and confounding biases, showed a significant decrease in in-hospital death with a resulting HR of 0.66 (95% CI, 0.47–0.93). In Model 1, the estimated effect for competing events was 0.84 (95% CI, 0.59–1.21) for discharge home and 1.30 (95% CI, 0.86–1.94) for transfer to another healthcare facility. By accounting for a delay in treatment administration time through modeling a time-varying treatment status in Model 2, the HRs increased for all outcomes: 0.79 (95% CI, 0.59–1.06) for in-hospital death, 0.91 (95% CI, 0.66–1.25) for discharge home, and 1.38 (95% CI, 0.96–1.97) for transfer. In addition, after adjusting for baseline covariates in Models 3 and 4, by fitting a multivariable Cox regression (Model 3) or using the inverse probability of treatment weighting (Model 4), we observed for all outcomes shifts toward higher HRs compared to the fully crude analysis (Model 1). Most of the findings did not yield statistically

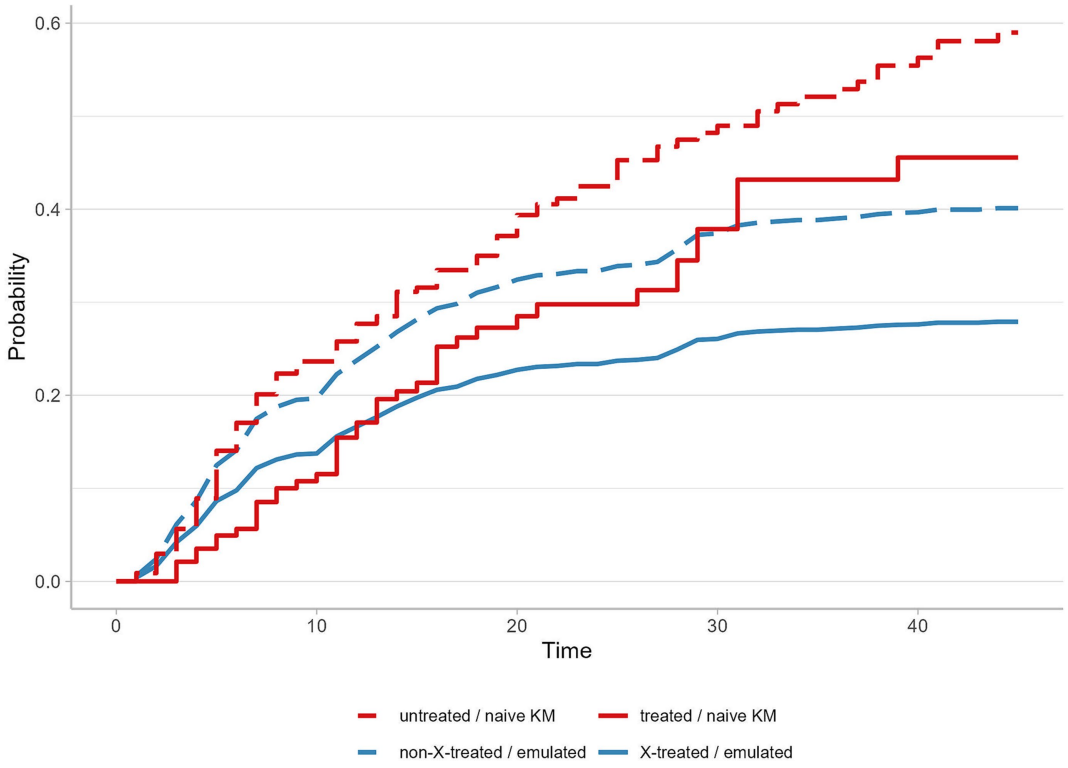

FIGURE 2
Cumulative in-hospital death probabilities by treatment group, comparing results from the naïve Kaplan-Meier estimator applied to initial data with the weighted Aalen-Johansen in emulated trial. Abbreviations: KM, the Kaplan-Meier estimator; Emulated, emulated target trial analysis using the weighted Aalen-Johansen estimator.

significant results, except in Model 1 for the in-hospital death outcome and in Model 3 for the transfer outcome.

In the emulated trial (Model 5) with defining a hypothetical protocol and a reliable 2-day treatment administration period, the resulting HRs were 0.68 (95% CI, 0.46–1.02) for in-hospital death, 1.22 (95% CI, 0.82–1.81) for discharge home, and 1.26 (95% CI, 0.77–2.07) for transfer. The trial emulation analysis allowed to model a hypothetical trial in which the treatment was administered within the first 2 days of hospital admission. This analysis showed that the treatment effect on both discharge home and transfer is toward a beneficial direction, and suggests a reduction in in-hospital death, however none of these results were statistically significant.

Discussion

This paper provides an overview of the methodological limitations of competing risks, immortal-time bias, and confounding bias when evaluating treatment effectiveness using observational hospital data from COVID-19 patients. This article demonstrates how biases may be mistakenly introduced and discusses the limitations of standard approaches that may lead to biased estimates of treatment effects. Observational studies evaluating treatment effectiveness are often complex, and have the potential for various types of biases. These combinations of biases can result in shifted effects of different magnitudes and directions, making it difficult to accurately estimate treatment effectiveness (14, 19). Our study aims to raise awareness of the common biases and the importance of addressing these limitations. This knowledge is essential for researchers assessing treatment effectiveness, particularly during the emergence or re-emergence of infectious diseases, when investigators face significant time constraints to obtain high-quality evidence of treatment effectiveness when relying on observational data, as was the case during the COVID-19 pandemic.

In our study, we illustrate the competing risk issue using a typical example of observational hospital data. Our results show that the naïve Kaplan-Meier estimator leads to biased cumulative incidence probabilities for the primary event of interest. Censoring discharged patients violated the independent censoring assumption, thus overestimating the probabilities of in-hospital death (47). Various methodologies and analytical techniques are available for analyses in the presence of competing events (48). In our emulated trial study, we used the Aalen-Johansen estimator to account for competing risks (49). This technique determined the proportion of patients who experienced a primary event of interest within the given time, considering the presence of competing events (50). Our previous studies elaborated on implementing competing risk analyses within the target trial emulation framework (51, 52). Another method to account for dependent censoring is to use the inverse probability of censoring weighting, which weights patients by the inverse probability of not yet having the competing event (48, 49). These weights can then be implemented in the Kaplan-Meier estimator (48). In fact, we agree with prior research that the choice of statistical analysis method in the

TABLE 1 Overview of statistical methods and results while addressing vs. neglecting immortal time and confounding biases.

Model	Approach	Statistical analysis method for outcome models	Hazard ratio (HR, [95% CI])			Immortal-time bias		Baseline confounding bias	
			In-hospital death	Discharge home	Transfer	Occurrence	Description	Occurrence	Description
1	Conventional	Univariable Cox regression model with treatment status incorrectly assigned at baseline	0.66 [0.47–0.93]	0.84 [0.59–1.21]	1.30 [0.86–1.94]	Yes	Ever-treated patients misclassified as treated from admission; never-treated as untreated	Yes	Baseline covariates not included in regression model
2	Conventional	Univariable, time-dependent Cox regression model with time-varying treatment status	0.79 [0.59–1.06]	0.91 [0.66–1.25]	1.38 [0.96–1.97]	No	Treated patients time classified to "untreated / "treated" periods using start-stop notation; pre-treatment time classified as "untreated"	Yes	
3	Conventional	Multivariable, time-dependent Cox regression model with baseline covariates and time-varying treatment status	0.76 [0.58–1.00]	0.92 [0.68–1.24]	1.41 [1.01–1.99]	No		No	Baseline covariates included in regression model
4	Inverse probability treatment weighting	Weighted, time-dependent Cox regression model with weights as a covariate and time-varying treatment status	0.76 [0.52–1.08]	0.98 [0.67–1.42]	1.50 [1.00–2.24]	No		No	Baseline covariates included in inverse probability treatment weights via propensity scores
5	Target trial emulation with clone-censor-weight approach	Weighted cause-specific Cox regression with censoring weights as a covariate and treatment arm	0.68 [0.46–1.02]	1.22 [0.82–1.81]	1.26 [0.77–2.07]	No	Two clones: one in 'X'-treated arm and one in non-'X'-treated arm	No	Cloning results in balanced covariates between two arms at baseline, inverse probability censoring weights applied to correct for selection bias

HR, Hazard ratio; CI, Confidence interval.

10.3389/fmed.2024.1362192

presence of competing events depends on the specific causal research question and the type of event (48).

A competing risk analysis that reports cumulative incidence for heterogeneous outcomes could be particularly beneficial. Acknowledging all clinically important endpoints can provide researchers with a more comprehensive understanding of disease progression and enhance the assessment of therapy-associated benefits and risks. In a target trial emulation study conducted by Urner et al. evaluating the effectiveness of venovenous extracorporeal membrane oxygenation (ECMO) in COVID-19 patients, the study reported results for the primary outcome of in-hospital death and for the competing event of hospital discharge (53). Their study defined discharge alive as a competing event for in-hospital death rather than a censoring event. Such an approach provides a more comprehensive understanding of ECMO's impact on various clinical outcomes (53).

Previous studies evaluated the impact of immortal-time bias and confounding bias on treatment effect estimates by comparing standard analytical approaches with emulated trials (54, 55). Hoffman et al. (54) reported that immortal time can lead to biased treatment effect estimates. The common "model-first" approaches failed to achieve the randomized controlled trial (RCT) benchmark using the same data source compared to the target trial emulation framework (54). The study conducted by Kuehne et al. evaluated the effectiveness of ovarian cancer treatment in terms of overall survival. The study found that ignoring methodological biases and using crude (univariable) analysis methods led to significant variation in effect measures, with immortal-time bias contributing more substantially to the shifted effects than confounding (55). That study also demonstrated that various methodological biases can significantly shift the treatment effect measure in different directions. Our analysis led to similar conclusions. The magnitude of immortal-time bias can be influenced by factors such as the length of the immortal time period, the proportion of exposed patients, the event rate, and the length of a study's follow-up (15, 56).

Our study also highlights the impact of baseline confounding bias and the importance of addressing it properly. To prevent confounding bias, it is essential to identify and account for all potential, clinically important confounders, and to apply appropriate statistical methods (27). The evaluation of time-dependent treatments necessitates the inclusion of post-baseline (time-dependent) confounders (54, 57, 58). High-quality, time-dependent data are crucial for drawing causal conclusions from observational data (27, 57). In our analysis, data on time-updated prognostic covariates were not available, which makes our study susceptible to time-dependent confounding bias. This is because treatment administration after baseline often depends on changing prognostic characteristics. To adjust for time-updated covariates, time-dependent clinical characteristics could be incorporated into the weights models (45, 57).

Our examination aligns with the existing literature recommending the target trial emulation framework as a beneficial approach for analyzing real-world data (24, 33, 54). This framework increases transparency in both the design and analysis stages by explicitly defining the research question, outcome, time zero, treatment strategies and assignment, and the analysis plan (24, 33). This approach facilitates the early identification and mitigation of potential biases by applying of design and/or analytical strategies (33). While the target trial emulation framework offers advantages, we acknowledge its methodological complexities and the need to address frequent challenges associated with observational data (24, 59). For more detailed introductions and tutorials on the emulated target trial framework, we refer to the articles by Hernan et al. (33), Fu (24), and Maringe et al. (34).

Our study has several potential limitations. First, it is a demonstrative study that uses a common data example from a single center, restricting the generalizability of our results regarding the magnitude of biases on the treatment effect. Therefore, our findings on the magnitude of each bias cannot be extrapolated to other settings. Second, we developed a simplified version of a hypothetical trial protocol, and additional criteria could be included in real treatment assessment studies. Third, while we accounted for numerous baseline clinical covariates to control for confounding, we admit that unmeasured confounding is probable in our study. Data on time-updated prognostic covariates were not available. Fourth, we reported HRs as a summary measure to facilitate comparisons among the various regression models. Such summary effect measures as risk differences and risk ratios are preferable to hazards and are easier to interpret clinically (47, 60). Lastly, we did not discuss additional limitations of observational studies, such as selection bias, data quality, and missing data issues, all of which can impact the accuracy of their results (4, 61). However, it is important to emphasize that our findings were not interpreted clinically.

Data availability statement

The datasets presented in this article are not readily available because data are not accessible for public use. Statistical code is available from the corresponding author upon request. Requests to access the code should be directed to martin.wolkewitz@uniklinik-freiburg.de.

Ethics statement

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Bellvitge University Hospital (PR 128/20). The studies were conducted in accordance with the local legislation and institutional requirements. Informed consent was waived after assessment by the Research Ethics Committee.

Author contributions

OM: Conceptualization, Formal Analysis, Methodology, Software, Visualization, Writing – original draft. DH: Writing – review & editing. HM: Data curation, Funding acquisition, Writing – review & editing. MM: Data curation, Funding acquisition, Writing – review & editing. MAM: Data curation, Funding acquisition, Writing – review & editing. SR: Data curation, Funding acquisition, Writing – review & editing. MR-R: Data curation, Funding acquisition, Investigation, Writing – review & editing. MW: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study has been funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) with project grant WO 1746/5-1 (MW), and Beatriu de Pinós post-doctoral programme from the Office of the Secretary of Universities and Research from the Ministry of Business and Knowledge of the Government of Catalonia programme: 2020 BP 00261 (HM); the Ministry of Science and Innovation [Ministerio de Ciencia e Innovación (MICINN)], Spain, under contract PID2020-117751RB-I00 (MAM, SR). CIBER-BBN is an initiative of the Instituto de Salud Carlos III, Spain. The funders had no role in study design, data collection and analysis, decision to publish, or manuscript preparation.

Acknowledgments

We acknowledge support by the Open Access Publication Fund of the University of Freiburg.

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/fmed.2024.1362192/full#supplementary-material

References

1. Read SH, Khachatryan A, Chandak A, Casciano R, Hodgkins P, Haubrich R, et al. Comparative effectiveness research in COVID-19 using real-world data: methodological considerations. *J Comp Eff Res*. (2021) 10:1259–64. doi: 10.2217/cer-2021-0179

2. Martinuka O, Von CM, Wolkewitz M. Methodological evaluation of bias in observational coronavirus disease 2019 studies on drug effectiveness. *Clin Microbiol Infect*. (2021) 27:949–57. doi: 10.1016/j.cmi.2021.03.003

3. Cohen JB, D'Agostino McGowan L, Jensen ET, Rigdon J, South AM. Evaluating sources of bias in observational studies of angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker use during COVID-19: beyond confounding. *J Hypertens*. (2021) 39:795–805. doi: 10.1097/HJH.0000000000002706

4. van Nguyen T, Engleton M, Davison M, Ravaud P, Porcher R, Boutron I. Risk of bias in observational studies using routinely collected data of comparative effectiveness research: a meta-research study. *BMC Med*. (2021) 19:279. doi: 10.1186/s12916-021-02151-w

5. Hempenius M, Bots SH, Groenwold RHH, de BA, Klungel OH, Gardarsdottir H. Bias in observational studies on the effectiveness of in hospital use of hydroxychloroquine in COVID-19. *Pharmacoepidemiol Drug Saf*. (2023) 32:1001–11. doi: 10.1002/pds.5632

6. Austin PC, Fine JP. Accounting for competing risks in randomized controlled trials: a review and recommendations for improvement. *Stat Med*. (2017) 36:1203–9. doi: 10.1002/sim.7215

7. Noordzij M, Leffondré K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ. When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant*. (2013) 28:2670–7. doi: 10.1093/ndt/gft355

8. Oulhaj A, Ahmed LA, Prattes J, Suliman A, Alsuwaidi AR, Al-Rifai RH, et al. (2020). The competing risk between in-hospital mortality and recovery: A pitfall in COVID-19 survival analysis research.

9. Andersen PK, Geskus RB, De WT, Putter H. Competing risks in epidemiology: possibilities and pitfalls. *Int J Epidemiol*. (2012) 41:861–70. doi: 10.1093/ije/dyr213

10. Coemans M, Verbeke G, Döhler B, Süsal C, Naesens M. Bias by censoring for competing events in survival analysis. *BMJ*. (2022) 378:e071349. doi: 10.1136/bmj-2022-071349

11. Zhang Z. Survival analysis in the presence of competing risks. *Ann Transl Med*. (2017) 5:47. doi: 10.21037/atm.2016.08.62

12. Wolkewitz M, Lambert J, Von CM, Bugiera L, Grodd M, Hazard D, et al. Statistical analysis of clinical COVID-19 data: a concise overview of lessons learned, common errors and how to avoid them. *Clin Epidemiol*. (2020) 12:925–8. doi: 10.2147/CLEP.S256735

13. Wolkewitz M, Schumacher M. Survival biases lead to flawed conclusions in observational treatment studies of influenza patients. *J Clin Epidemiol*. (2017) 84:121–9. doi: 10.1016/j.jclinepi.2017.01.008

14. Liu J, Weinhandl ED, Gilbertson DT, Collins AJ, St Peter WL. Issues regarding 'immortal time' in the analysis of the treatment effects in observational studies. *Kidney Int*. (2012) 81:341–50. doi: 10.1038/ki.2011.388

15. Tyrer F, Bhaskaran K, Rutherford MJ. Immortal time bias for life-long conditions in retrospective observational studies using electronic health records. *BMC Med Res Methodol*. (2022) 22:86. doi: 10.1186/s12874-022-01581-1

16. Renoux C, Azoulay L, Suissa S. Biases in evaluating the safety and effectiveness of drugs for the treatment of COVID-19: designing real-world evidence studies. *Am J Epidemiol*. (2021) 190:1452–6. doi: 10.1093/aje/kwab028

17. Mi X, Hammill BG, Curtis LH, Lai EC-C, Setoguchi S. Use of the landmark method to address immortal person-time bias in comparative effectiveness research: a simulation study. *Stat Med*. (2016) 35:4824–36. doi: 10.1002/sim.7019

18. Karim ME, Gustafson P, Petkau J, Tremlett H. Comparison of statistical approaches for Dealing with immortal time Bias in drug effectiveness studies. *Am J Epidemiol*. (2016) 184:325–35. doi: 10.1093/aje/kwv445

19. Wang J, Peduzzi P, Wininger M, Ma S. Statistical methods for accommodating immortal time: a selective review and comparison. (2022). arXiv [Preprint]. doi: 10.48550/arXiv.2202.02369

20. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. (2008) 167:492–9. doi: 10.1093/aje/kwm324

21. Dekkers OM, Groenwold RHH. When observational studies can give wrong answers: the potential of immortal time bias. *Eur J Endocrinol*. (2021) 184:E1–4. doi: 10.1530/EJE-20-1124

22. Beyersmann J, Gastmeier P, Wolkewitz M, Schumacher M. An easy mathematical proof showed that time-dependent bias inevitably leads to biased effect estimation. *J Clin Epidemiol*. (2008) 61:1216–21. doi: 10.1016/j.jclinepi.2008.02.008

23. Pierce M, Dunn G, Millar T. Confounding in longitudinal studies in addiction treatment research. *Addict Res Theory*. (2017) 25:236–42. doi: 10.1080/16066359.2016.1247812

24. Fu EL. Target trial emulation to improve causal inference from observational data: what, why, and how? *J Am Soc Nephrol*. (2023) 34:1305–14. doi: 10.1681/ASN.0000000000000152

25. Schuster NA, Rijnhart JJM, Bosman LC, Twisk JWR, Klausch T, Heymans MW. Misspecification of confounder-exposure and confounder-outcome associations leads to bias in effect estimates. *BMC Med Res Methodol*. (2023) 23:11. doi: 10.1186/s12874-022-01817-0

26. Assimon MM. Confounding in observational studies evaluating the safety and effectiveness of medical treatments. *Kidney360*. (2021) 2:1156–9. doi: 10.34067/KID.0007022020

27. Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC (2020).

28. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Res*. (2011) 46:399–424. doi: 10.1080/00273171.2011.568786

29. Cho K, Keithly SC, Kurgansky KE, Madenci AL, Gerlovin H, Marucci-Wellman H, et al. Early convalescent plasma therapy and mortality among US veterans

hospitalized with nonsevere COVID-19: an observational analysis emulating a target trial. *J Infect Dis*. (2021) 224:967–75. doi: 10.1093/infdis/jiab330

30. Gupta S, Wang W, Hayek SS, Chan L, Mathews KS, Melamed ML, et al. Association between early treatment with tocilizumab and mortality among critically ill patients with COVID-19. *JAMA Intern Med*. (2021) 181:41–51. doi: 10.1001/jamainternmed.2020.6252

31. Hajage D, Combes A, Guervilly C, Lebreton G, Mercat A, Pavot A, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome associated with COVID-19: an emulated target trial analysis. *Am J Respir Crit Care Med*. (2022) 206:281–94. doi: 10.1164/rccm.202111-2495OC

32. Martínez-Alés G, Domingo-Relloso A, Quintana-Díaz M, Fernández-Capitán C, Hernán MA. Thromboprophylaxis with standard-dose vs. flexible-dose heparin for hospitalized COVID-19 patients: a target trial emulation. *J Clin Epidemiol*. (2022) 151:96–103. doi: 10.1016/j.jclinepi.2022.08.006

33. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. (2016) 183:758–64. doi: 10.1093/aje/kwv254

34. Maringe C, Benitez Majano S, Exarchakou A, Smith M, Rachet B, Belot A, et al. Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. *Int J Epidemiol*. (2020) 49:1719–29. doi: 10.1093/ije/dyaa057

35. Zhang M-J, Zhang X, Scheike TH. Modeling cumulative incidence function for competing risks data. *Expert Rev Clin Pharmacol*. (2008) 1:391–400. doi: 10.1586/17512433.1.3.391

36. Rubio-Rivas M, Corbella X, Formiga F, Menéndez Fernández E, Martín Escalante MD, Baños Fernández I, et al. Risk categories in COVID-19 based on degrees of inflammation: data on more than 17,000 patients from the Spanish SEMI-COVID-19 registry. *J Clin Med*. (2021) 10:10. doi: 10.3390/jcm10102214

37. Zheng Q, Otahal P, Cox IA, de Graaff B, Campbell JA, Ahmad H, et al. The influence of immortal time bias in observational studies examining associations of antifibrotic therapy with survival in idiopathic pulmonary fibrosis: a simulation study. *Front Med*. (2023) 10:1157706. doi: 10.3389/fmed.2023.1157706

38. Mansournia MA, Nazemipour M, Etminan M. Causal diagrams for immortal time bias. *Int J Epidemiol*. (2021) 50:1405–9. doi: 10.1093/ije/dyab157

39. van Houwelingen HCPH. *Dynamic Prediction in Clinical Survival Analysis*. Boca Raton: CRC Press/Chapman and Hall (2012).

40. Wolkewitz M, Beyersmann J, Gastmeier P, Schumacher M. Efficient risk set sampling when a time-dependent exposure is present: matching for time to exposure versus exposure density sampling. *Methods Inf Med*. (2009) 48:438–43. doi: 10.3414/ME9241

41. Wolkewitz M, Beyersmann J, Ohneberg K, Schumacher M. Comparison of statistical approaches for dealing with immortal time bias in drug effectiveness studies. *Am J Epidemiol*. (2016) 184:856–8. doi: 10.1093/aje/kww156

42. Chesnaye NC, Stel VS, Tripepi G, Dekker FW, Fu EL, Zoccali C, et al. An introduction to inverse probability of treatment weighting in observational research. *Clin Kidney J*. (2022) 15:14–20. doi: 10.1093/ckj/sfab158

43. Smith MJ, Mansournia MA, Maringe C, Zivich PN, Cole SR, Leyrat C, et al. Introduction to computational causal inference using reproducible Stata, R, and Python code: a tutorial. *Stat Med*. (2022) 41:407–32. doi: 10.1002/sim.9234

44. Van der Wal W.M., Geskus R.B. (2011). "ipw"

45. Breskin A, Wiener C, Adimora AA, Brown RS, Landis C, Reddy KR, et al. Effectiveness of Remdesivir treatment protocols among patients hospitalized with COVID-19: a target trial emulation. *Epidemiology*. (2023) 34:365–75. doi: 10.1097/EDE.0000000000001598

46. RStudio Team. *RStudio: Integrated Development Environment for R* RStudio, PBC (2022). Available at: https://www.R-project.org/

47. Piovani D, Nikolopoulos GK, Bonovas S. Escollos y peligros del análisis de supervivencia bajo supuestos incorrectos: el caso de los datos de COVID-19. *Biomedica*. (2021) 41:21–8. doi: 10.7705/biomedica.5987

48. Rojas-Saunero LP, Young JG, Didelez V, Ikram MA, Swanson SA. Considering questions before methods in dementia research with competing events and causal goals. *Am J Epidemiol*. (2023) 192:1415–23. doi: 10.1093/aje/kwad090

49. van Geloven N, Le Cessie S, Dekker FW, Putter H. Transplant as a competing risk in the analysis of dialysis patients. *Nephrol Dial Transplant*. (2017) 32:ii53–9. doi: 10.1093/ndt/gfx012

50. Genet A, Bogner K, Goertz R, Böhme S, Leverkus F. Safety analysis of new medications in clinical trials: a simulation study to assess the differences between cause-specific and subdistribution frameworks in the presence of competing events. *BMC Med Res Methodol*. (2023) 23:168. doi: 10.1186/s12874-023-01985-7

51. Martinuka O, Hazard D, Marateb HR, Maringe C, Mansourian M, Rubio-Rivas M, et al. Target trial emulation with multi-state model analysis to assess treatment effectiveness using clinical COVID-19 data. *BMC Med Res Methodol*. (2023) 23:197. doi: 10.1186/s12874-023-02001-8

52. Martinuka O, Cube M v, Hazard D, Marateb HR, Mansourian M, Sami R, et al. Target trial emulation using hospital-based observational data: demonstration and application in COVID-19. *Life*. (2023) 13:777. doi: 10.3390/life13030777

53. Urner M, Barnett AG, Bassi GL, Brodie D, Dalton HJ, Ferguson ND, et al. Venovenous extracorporeal membrane oxygenation in patients with acute covid-19 associated respiratory failure: comparative effectiveness study. *BMJ*. (2022) 377:e068723. doi: 10.1136/bmj-2021-068723

54. Hoffman KL, Schenck EJ, Satlin MJ, Whalen W, Pan D, Williams N, et al. Comparison of a target trial emulation framework vs Cox regression to estimate the Association of Corticosteroids with COVID-19 mortality. *JAMA Netw Open*. (2022) 5:e2234425. doi: 10.1001/jamanetworkopen.2022.34425

55. Kuehne F, Arvandi M, Hess LM, Faries DE, Matteucci Gothe R, Gothe H, et al. Causal analyses with target trial emulation for real-world evidence removed large self-inflicted biases: systematic bias assessment of ovarian cancer treatment effectiveness. *J Clin Epidemiol*. (2022) 152:269–80. doi: 10.1016/j.jclinepi.2022.10.005

56. Harding BN, Weiss NS. Point: immortal time Bias-what are the determinants of its magnitude? *Am J Epidemiol*. (2019) 188:1013–5. doi: 10.1093/aje/kwz067

57. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *BMJ*. (2017) 359:j4587. doi: 10.1136/bmj.j4587

58. Martínez-Sanz J, Muriel A, Ron R, Herrera S, Pérez-Molina JA, Moreno S, et al. Effects of tocilizumab on mortality in hospitalized patients with COVID-19: a multicentre cohort study. *Clin Microbiol Infect*. (2021) 27:238–43. doi: 10.1016/j.cmi.2020.09.021

59. Hansford HJ, Cashin AG, Jones MD, Swanson SA, Islam N, Douglas SRG, et al. Reporting of observational studies explicitly aiming to emulate randomized trials: a systematic review. *JAMA Netw Open*. (2023) 6:e2336023. doi: 10.1001/jamanetworkopen.2023.36023

60. Tripepi G, Jager KJ, Dekker FW, Wanner C, Zoccali C. Measures of effect: relative risks, odds ratios, risk difference, and 'number needed to treat'. *Kidney Int*. (2007) 72:789–91. doi: 10.1038/sj.ki.5002432

61. Tompsett D, Zylbersztejn A, Hardelid P, de Stavola B. Target trial emulation and Bias through missing eligibility data: an application to a study of Palivizumab for the prevention of hospitalization due to infant respiratory illness. *Am J Epidemiol*. (2023) 192:600–11. doi: 10.1093/aje/kwac202

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 12 June 2024
DOI 10.3389/fmed.2024.1376275

OPEN ACCESS

EDITED BY
Cristiana Sessa,
Oncology Institute of Southern Switzerland (IOSI), Switzerland

REVIEWED BY
Ferdi Tanir,
Cukurova University, Türkiye
Stefan Pilz,
Medical University of Graz, Austria

*CORRESPONDENCE
Susanne Weber
✉ Susanne.weber@uniklinik-freiburg.de

RECEIVED 25 January 2024
ACCEPTED 23 May 2024
PUBLISHED 12 June 2024

CITATION
Weber S, Hedberg P, Naucler P and Wolkewitz M (2024) Protection from prior natural infection vs. vaccination against SARS-CoV-2—a statistical note to avoid biased interpretation.
Front. Med. 11:1376275.
doi: 10.3389/fmed.2024.1376275

Protection from prior natural infection vs. vaccination against SARS-CoV-2—a statistical note to avoid biased interpretation

Susanne Weber[1,2]*, Pontus Hedberg[3], Pontus Naucler[4] and Martin Wolkewitz[1,2]

[1]Institute of Medical Biometry and Statistics, Division Methods in Clinical Epidemiology, Faculty of Medicine and Medical Center - University of Freiburg, Freiburg, Germany, [2]Freiburg Center for Data Analysis and Modeling, University of Freiburg, Freiburg, Germany, [3]Department of Medicine, Huddinge, Karolinska Institutet, Stockholm, Sweden, [4]Department of Medicine, Solna, Karolinska Institutet and Department of Infectious Diseases, Karolinska University Hospital, Stockholm, Sweden

Introduction: The fight against SARS-CoV-2 has been a major task worldwide since it was first identified in December 2019. An imperative preventive measure is the availability of efficacious vaccines while there is also a significant interest in the protective effect of a previous SARS-CoV-2 infection on a subsequent infection (natural protection rate).

Methods: In order to compare protection rates after infection and vaccination, researchers consider different effect measures such as 1 minus hazard ratio, 1 minus odds ratio, or 1 minus risk ratio. These measures differ in a setting with competing risks. Nevertheless, as there is no unique definition, these metrics are frequently used in studies examining protection rate. Comparison of protection rates via vaccination and natural infection poses several challenges. For instance many publications consider the epidemiological definition, that a reinfection after a SARS-CoV-2 infection is only possible after 90 days, whereas there is no such constraint after vaccination. Furthermore, death is more prominent as a competing event during the first 90 days after infection compared to vaccination. In this work we discuss the statistical issues that arise when investigating protection rates comparing vaccination with infection. We explore different aspects of effect measures and provide insights drawn from different analyses, distinguishing between the first and the second 90 days post-infection or vaccination.

Results: In this study, we have access to real-world data of almost two million people from Stockholm County, Sweden. For the main analysis, data of over 52.000 people is considered. The infected group is younger, includes more men, and is less morbid compared to the vaccinated group. After the first 90 days, these differences increased. Analysis of the second 90 days shows differences between analysis approaches and between age groups. There are age-related differences in mortality. Considering the outcome SARS-CoV-2 infection, the effect of vaccination versus infection varies by age, showing a disadvantage for the vaccinated in the younger population, while no significant difference was found in the elderly.

Discussion: To compare the effects of immunization through infection or vaccination, we emphasize consideration of several investigations. It is crucial to examine two observation periods: The first and second 90-day intervals following infection or vaccination. Additionally, methods to address imbalances are essential and need to be used. This approach supports fair comparisons, allows for more comprehensive conclusions and helps prevent biased interpretations.

KEYWORDS

protection rate, vaccination, conditional survival, competing risks, survival of the fittest

Introduction

The development of vaccines is a critical and ongoing task in the fight against coronavirus disease 2019 (COVID-19). Numerous vaccines are currently under development and some are tailored to the currently circulating Omicron sublineages. Additionally, scientists are examining the extent of protection provided by a previous severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection against the risk of a subsequent infection (natural protection rate). One major interest lies in the protection rates after infection compared with those after COVID-19 vaccination, e.g., (1, 2).

According to Gail et al. (3) there are different ways to measure protection rate. Vaccine effectiveness is defined as the percentage reduction in the attack rate that can be attributed to the vaccine. The attack rate is determined as the proportion of individuals infected within the designated risk group during a specified time period. This corresponds to the risk of acquiring an infection and thus the effect measure considered corresponds to 1-relative risk (RR), with RR being the relative risk of the vaccinated compared to the unvaccinated group. Additionally, Gail et al. (3) states that, when evaluating vaccine effectiveness using data from a case control study, 1-odds ratio (OR) is an appropriate effect measure, with OR being the odds ratio between the vaccinated and unvaccinated groups. In addition to these two effect measures, one minus the hazard ratio (HR) can also be considered. There are two considerations we want to point out when comparing RR, OR and HR. Considering the comparison of RR and OR a common rule is that if the event is rare (<10%) the estimates are similar, see (4). In a survival setting 1-RR and 1-HR do not differ if the hazard is small, see (3). However, in this work we focus on a competing risk setting, where the interest is in the comparison of a measure on the risk scale (RR or OR) and a measure on the rate scale (HR). In competing risk settings, the effect measures are only comparable if in addition there is no effect on the competing hazard.

When evaluating protection rates after vaccination or infection, each of these effect measures is considered. However, they address different scales. While the HR is a measure on the rate scale, the RR and the OR are measures on the risk scale. Consequently, HRs give information about direct effects on the cause-specific hazard for the event of interest and about indirect effects via influence on possible competing event hazards. For instance, death is a competing risk for the event of interest, which is infection after vaccination or infection. In contrast, ORs and RRs are summaries of direct and indirect effects, allowing for conclusions about the probability of the occurrence of the event of interest.

Although these measures differ when facing competing risks, they are all used for investigation, as there is no unique definition of the protection rate. For instance, Letizia et al. (5) used data of an observational study for investigation of the natural protection rate. Analysis was done via Poisson regression and 1-HR is reported. Dagan et al. (6) investigated vaccine effectiveness using Kaplan–Meier estimators and used the corresponding risk estimates in order to obtain the vaccine effectiveness via 1-RR. Powell et al. (7) considered 1- odds ratio (OR) in order to compare protection rate after infection and vaccination for different variants.

It should be noted, that when considering former SARS-CoV-2 infection as an exposure and its effects, the competing risk of death is more prominent than after COVID-19 vaccination. The infection affects the mortality hazard, which has an impact on the time at risk for developing a further infection and is hence indirectly affecting the infection risk. Furthermore, in publications a reinfection after a SARS-CoV-2 infection is only possible after 90 days per epidemiological definition (1, 2). Hence, analysis of the protection rate of an infection starts after 90 days and only individuals surviving the first 90 days are at risk of a reinfection. In contrast, there is no such constraint for the analysis of protection rate after vaccination. Consequently, when comparing protection rate after infection or vaccination, analysis should start after 90 days in order to avoid immortal time bias. However, this is not a fair comparison, as it is prone to selection bias. The mortality hazard increases in the initial period after an infection and subsequently decreases until 90 days post-infection. Thus, as elderly and more morbid patients are at a higher risk of dying due to infection during the first 90 days, the population is overall healthier during the second 90 days. For the vaccinated group there is no such selection during the first 90 days, as the vaccination does not affect the mortality hazard.

Thus, comparison of the protection rate of vaccination versus natural infection poses several challenges due to significant differences in reinfection and death rates among groups within the first 90 days.

To examine the statistical challenges that arise from assessing protection rates, we have access to population-based observational data from several databases from Stockholm County, Sweden. Information on almost two million people is available from 2020 to 2022. Information about SARS-CoV-2 infections and vaccinations is provided, alongside other patient-related characteristics, all derived from population-based data sources with high coverage. Thus, when comparing the natural protection with protection after a vaccination, we can examine imbalances between groups at baseline and during follow-up and discuss solutions to address them. Furthermore, we estimate HRs and ORs for comparison, in order to obtain one measure on the rate scale and one measure on the risk scale.

For a comparison between immunization via infection or vaccination, we consider several investigations. We highlight different aspects of effect measures and insights drawn from different analyses. The aim is to promote an awareness of the differences between the causes of protection in order to create a fair comparison.

Materials and methods

Study population

Among the entire population of Stockholm County, Sweden, we identified all individuals born 2001 or earlier, thus being 18 years or older during the entire COVID-19 pandemic period. We included all individuals alive and residing in Stockholm County on the 15th of March 2020. Individuals with a PCR test positive for SARS-CoV-2 before the 16th March 2020 were excluded.

Data sources

Data were linked from three population-based data sources using personal identification numbers, unique for each Swedish resident, from the Stockholm regional healthcare data warehouse (VAL), SmiNet, and the National Vaccination Register (NVR). VAL contains data from administrative healthcare databases within the Stockholm

Region, including demographics, migration, drug prescriptions, and data on all inpatient stays and outpatient visits reimbursed by Region Stockholm (8). This includes near complete coverage of specialist care and 94% of primary care (8). SmiNet contains all PCR SARS-CoV-2 positive test results reported in accordance with the Communicable Diseases Act (9). The data from NVR included all COVID-19 vaccinations administered in Sweden to the Stockholm County population.[1]

Analysis

The aim of this work is the comparison of protection of a SARS-CoV-2 infection after a first vaccination without being infected before, or a first infection without being vaccinated before. For simplicity, we will call the first vaccination or infection "*time of first immunization*."

We perform several analyses. First, we determine the inclusion window in order to define the study population for the main analysis of this work. The main analysis addresses the comparison of protection rate after vaccination with the natural protection rate. Observation for this main analysis starts at time of first immunization. Thus, the focus of the first analysis is to investigate the time to first immunization and in the main analysis, we focus on challenges concerning group imbalances and different effect measures. We distinguish between the first and the second 90 days after first immunization and investigate both observation periods.

Determination of study cohort for main analysis (inclusion window)

As we want to compare protection rates after first vaccination and first SARS-CoV-2 infection, we have to define an inclusion window in order to define a study cohort with reasonable groups for comparison. Thus, we first have a look at the competing risk model considering time to first immunization (first vaccination or first SARS-CoV-2 infection separately). Death is a competing risk, see Supplementary Figure S1. The aim is to define an inclusion window, so that the vaccinated group and the infected group are both big enough and facing the same pandemic situation.

Follow-up for this analysis starts on 2020-03-15 when a more extensive transmission of the virus started. Note, that vaccination first was possible on December 27th 2020 in Sweden.

Main analysis

In order to investigate protection rate after immunization via SARS-CoV-2 infection or vaccination, we distinguish between two observation periods. The first observation period represents the first 90 days after the time of first immunization. The second observation period represents the second 90 days, starting at day 91. In general a reinfection after SARS-CoV-2 infection is per definition only possible after 90 days, thus investigation of the protection rate starts after the first 90 days. The first 90 days represent the selection process, selecting individuals who are surviving the first 90 days and are thus available for the main analysis. Hence, the first 90 days after immunization, as well as the second 90 days should be considered and investigated.

Time zero is the time of the first immunization. In the following, the groups for comparison of the protection rate after SARS-CoV-2 or vaccination are called the infected and the vaccinated group, respectively. The infected group will be considered as the reference group. Possible confounders measured at time zero are age (continuous), sex (binary), and comorbidity count (categorical). Comorbidity count has categories $0,1,2,3,\geq 4$ and considers the following comorbidities: cancer, cardiac disease, cerebrovascular disease, chronic kidney failure, chronic liver disease, chronic lung disease, dementia, diabetes, dialysis, down syndrome, hypertension, mental health disorder, mental retardation, neurological disease, obesity, other immunocompromising conditions and treatments, pregnancy, transplantation (solid organ or stem cell), living in nursing home, and receiving home help services.

During the first 90 days, people in the vaccinated group can get an infection or they can die without an infection. In contrast to that, people in the infected group can die during the first 90 days and they can be vaccinated for the first time, but they cannot get infected [per definition; see (1, 7)]. Thus, groups are not comparable concerning reinfection during first 90 days.

Selection for the second analysis occurs at the end of the first observation period. Available for the analysis of the second 90 days are those people still alive and without an infection at the end of the first observation period, and without first vaccination after having the first SARS-CoV-2 infection. The outcome of interest during the second observation period is the occurrence of first SARS-CoV-2 after immunization during the second 90 days.

The second analysis is a conditional analysis. The first analysis provides information about the selection process for this conditional analysis.

Analysis of the first 90 days: selection for conditional survival

The chosen model is a competing risks model with three possible events during a follow-up of 90 days: first SARS-CoV-2 after immunization via first vaccination, Death, and first Vaccination after Immunization via first SARS-CoV-2 (see Supplementary Figure S2).

Note that we do not handle second vaccination as competing risk in the vaccination group. For simplicity, we decided that a further vaccination is no reason for exclusion of analysis of the second 90 days. Due to the methodological character of this work, we think that this is a reasonable choice.

The estimated transition probabilities (via the etm package in R) are illustrated via stacked probabilities plots. Death (death without first Covid19 nor first vaccination, and death overall) during the first 90 days is investigated via Cox regression and logistic regression. Continuous baseline characteristics are given by mean, standard deviation (SD), median, first quartile (Q1), and third quartile (Q3). Categorical baseline characteristics are given by percentages. A comparison has been done of the baseline characteristics for the baseline population and the population selected for the conditional survival analysis.

Analysis of the second 90 days: conditional survival

For the conditional survival analysis, follow-up starts 90 days after first immunization, see the competing risks model as depicted in Supplementary Figure S3. Follow-up is 90 days and the possible

1 https://www.folkhalsomyndigheten.se/the-public-health-agency-of-sweden/communicable-disease-control/vaccinations/vaccination-register-and-vaccination-coverage/variable-list-for-the-national-vaccination-register/

TABLE 1 Different analysis approaches for investigating protection rate and how imbalances between groups are addressed.

Approach	How?	What is addressed?
Crude		- crude group comparison - ignoring imbalances
Adjusted		- conditional effect in selected population - addressing imbalances after selection
Matched	Generalized Full Matching (ATE in selected population)	- addressing imbalances after selection - (Generalized full matching is often faster than even nearest neighbor matching, especially for large datasets)
Weighted	ATE in selected population	- addressing imbalances after selection

ATE, average treatment effect.

events are first SARS-CoV-2 infection after immunization or death. Excluded from the analysis are individuals who had a SARS-CoV-2 infection, who are first vaccinated, or who died within the first 90 days.

Conditional survival is actually what is being done in the literature when considering natural protection rate. In these examples, the first 90 days after infection are not considered (at the most briefly in the discussion, e.g., (1, 10)).

We estimate two effect measures considered in the literature when investigating the protection rate: HR and OR. Note that usually 1 minus the respective effect measure is given. The infected group is considered as the reference group. HRs are estimated via the survival package in R, ORs are estimated via the glm function using logit as a link function.

Focus of this analysis are differences between the effect measures and imbalances between the groups.

Both a crude analysis, i.e., without any adjustment, and several approaches, addressing the imbalance between groups were done: regressions with adjustment for baseline covariates (age (continuous), sex (binary), comorbidity count (categorical)), analysis of matched cohorts (matching of selected population via the same baseline covariates using the matchit package in R), and one weighted analysis.

Matched logistic regression is done using a mixed model with the matching group as a random effect (via lme4 package in R). The matching approach performs generalized full matching in the selected population. This is a faster alternative to the full matching approach and thus applicable to a large dataset. This matching approach estimates the average treatment effect (ATE) in a population compared to the selected population available after 90 days.

The weighting approach considers the selected population and uses inverse probability weights obtained via the weightit package considering the ATE option and each of the covariates mentioned above.

An overview of the different analysis approaches is listed in Table 1.

Analysis is done in R (Version 4.1.0).

Results

Study population

On March 15, 2020, 1,860,797 subjects were available in the dataset according to the inclusion criteria. A small proportion of those ($N = 11{,}203$, 0.6%) were excluded as there were some inconsistencies with the data during the follow up, for example earlier death date than first SARS-CoV-2 infection, second SARS-CoV-2 infection, or first vaccination. Thus, for the first analysis, the determination of the inclusion window for the time-to-event analysis resulted in 1,849,594 available subjects. In the case where the death date equaled the infection or vaccination date, 0.001 was added to the time of death.

Figure 1 shows the stacked plots in age groups (10 year steps) with the cumulative transition probabilities corresponding to the competing risk model addressing the time to first immunization.

It can be seen that vaccination is first possible at the end of 2020 (27 December 2020). Furthermore, in the orange area, which represents the proportion of people being first infected without being vaccinated before, there is only a minor increase after the beginning of 2021.

According to the National Board of Health and Welfare (11) the first four surges in Sweden are as follows: 1. from March to September 2020, 2. from October 2020 to January 2021, 3. from February to June 2021, 4. from July to December 2021. These boundaries can be seen via the violet dotted lines in Figure 1.

With this information and the development of the curves in Figure 1, the inclusion window for the study population is defined as first immunization between 2020-12-27 and 2021-01-31 (dates included). Using these dates implies that for the analysis of the first 90 days the follow-up of 90 days might fall into two surges (period 2 and 3). For the analysis of the second 90 days the follow-up of 90 days lies completely in period 3.

Analysis of first 90 days

According to the inclusion window, the time of first immunization lies between 2020-12-27 and 2021-01-31 for 56,201 subjects.

The baseline characteristics (as available in December 2020 for each subject) are given in Table 2.

The infected group is younger, has more men and is less morbid compared to the vaccinated group.

In Figure 2 the cumulative incidences correspond to the competing risks model addressing the first 90 days after first immunization for the overall population and in age groups (<60 and ≥ 60).

Note that the ranges of the y-axis differ depending on the considered population (overall or in age groups). Obviously, one difference is that during the first 90 days there are no SARS-CoV-2 infections in the infected group, i.e., there are no orange areas in the left column. In contrast, there are no green areas on the right column.

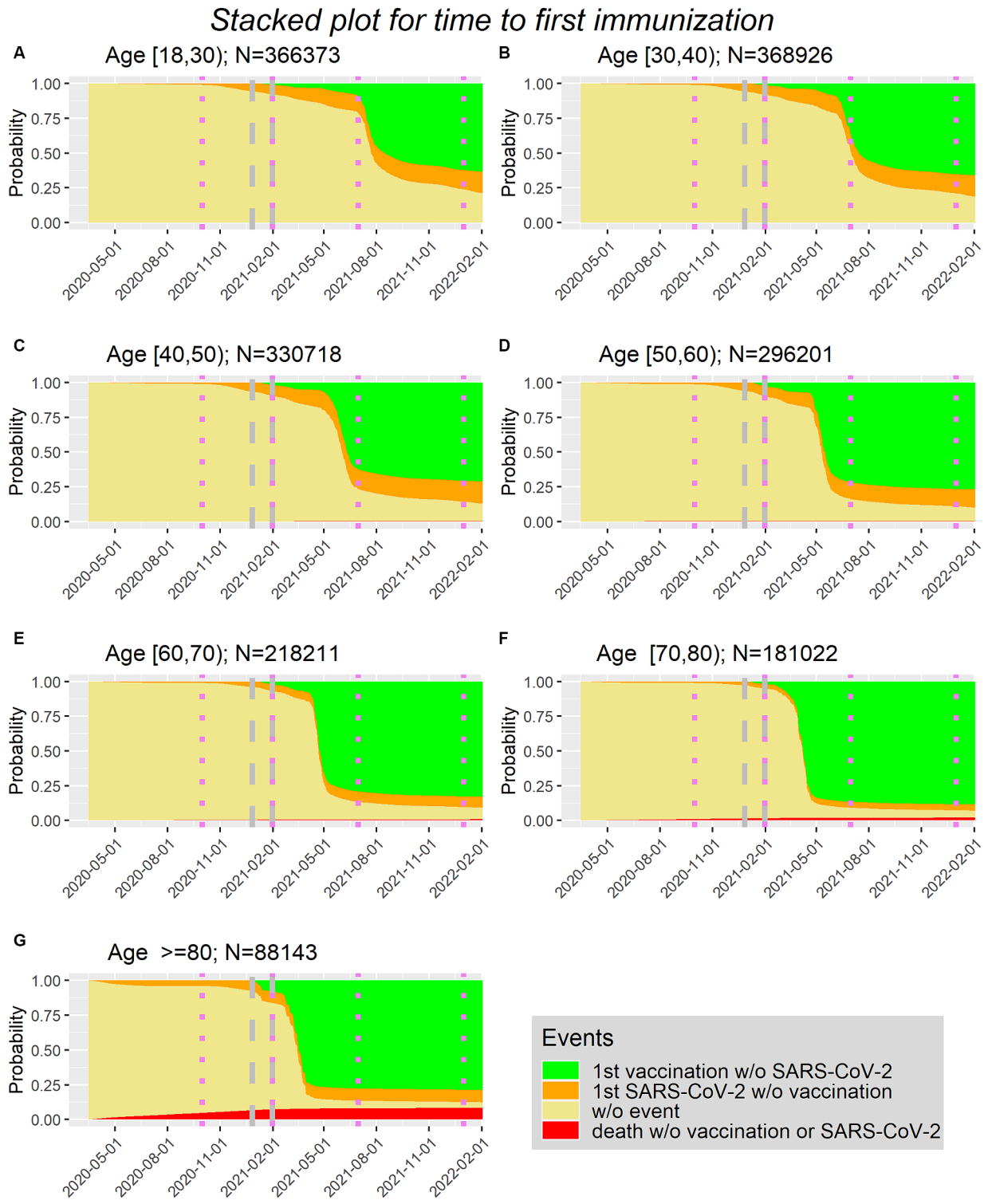

FIGURE 1
(A–G) Time to first immunization via first vaccination or first SARS-lines show borders of surges in Sweden in age groups (10 year steps, with **(A)** being the youngest group (Age between 18 and 30) and **(G)** being the oldest group (Age>=80)).

Recall that vaccination is not considered as competing event in the vaccinated group as we do not distinguish between different vaccination states. It can be seen, that the selected proportions for the conditional analysis in the overall population are similar between groups. The probability of having an event during the first 90 days is between 3.5 and 4% in both groups. In contrast to that, in the elderly population it is more likely to survive the first 90 days without an event in the vaccinated group, compared to the infected group. There is a difference between the two groups concerning the competing event death, especially in the elderly.

TABLE 2 Baseline characteristics for the population at baseline and for the selected population available after the first 90 days.

	Population at baseline		Population after 90 days	
	Infected group (*N* = 19,335)	Vaccinated group (*N* = 36,866)	Infected group (*N* = 16,924)	Vaccinated group (*N* = 35,525)
Age				
Mean (SD)	44.4 (16.9)	56.7 (20.4)	41.7 (14.3)	56.0 (20.1)
Median [Q1,Q3]	43.0 [31.0,55.0]	55.0 [41.0,73.0]	40.0 [30.0,52.0]	54.0 [41.0,71.0]
Sex				
Female	9,990 (51.7%)	26,106 (70.8%)	8,570 (50.6%)	25,223 (71.0%)
Male	9,345 (48.3%)	10,760 (29.2%)	8,354 (49.4%)	10,302 (29.0%)
Comorbidity count				
0	14,376 (74.4%)	19,623 (53.2%)	13,443 (79.4%)	19,381 (54.6%)
1	2,891 (15.0%)	5,886 (16.0%)	2,406 (14.2%)	5,762 (16.2%)
2	1,077 (5.6%)	4,018 (10.9%)	701 (4.1%)	3,769 (10.6%)
3	530 (2.7%)	3,561 (9.7%)	251 (1.5%)	3,261 (9.2%)
>=4	461 (2.4%)	3,778 (10.2%)	123 (0.7%)	3,352 (9.4%)

We performed regression analysis of death without other events during the first 90 days via Cox and logistic regression in the overall population and in age groups. In Table 3, the results are presented for crude regression and with adjustment for sex, age (continuous), and comorbidity count (categorical).

It is strikes that the unadjusted HRs and ORs are smaller than one in both age groups, but greater than one in the overall population. This situation is known as the Simpson Paradox (12). The reason therefore is that the sample size of the two groups in the age groups differ. The elderly infected group is only a small proportion of the overall SARS-CoV-2 population (3,447 of 19,335, 17.8%), whereas this is not the case for the vaccinated group (14,807 of 36,866, 40.2%). But, most of the deaths occur in the elderly population (in both immunization groups).

Adjusting or at least considering age groups leads to estimated effect sizes clearly apart from one, i.e., the unadjusted ORs by age groups are 0.225 for the younger population and 0.509 in the elderly.

In order to see how this selection process affects the study population the baseline characteristics for the selected population are also listed in Table 2. It can be seen that the differences in age and comorbidity count even increased in the available population. While the vaccinated group has only minor changes in these two variables, the infected group, notably, became less morbid from the first 90 days to the second 90 days, as more morbid patients passing away. Out of 461 subjects with a comorbidity count of ≥4 only 123 survived the first 90 days without an event. Note that in the vaccination group 3,352 of 3,778 with ≥4 comorbidities survived the first 90 days without an event. This difference is illustrated in Figure 3.

Note that the different movements in immunization groups mainly occurs in the older population, see Supplementary Figure S4 and Supplementary Table S1.

In conclusion, the population is changing within 90 days in a different way in the both groups. For the analysis of the second 90 days addressing the protection rate, the groups need to be made comparable in order to make a fair comparison.

Analysis of second 90 days

In order to see how the matching and weighting mechanisms perform, love plots are shown in Supplementary Figure S5.

In Figure 4, the results of the different regression analyses are presented. The exact values are listed in the Supplementary Table S2.

Note that adjustment is for sex, age (continuously), and comorbidity count (categorical).

The unadjusted analysis considering death shows HRs and ORs of with values greater than 12 in the overall population and values greater than 5 in the elderly. However, this is not a fair comparison. Recall the imbalances in the considered population. The vaccinated group is older and has more comorbidities. These are two factors with an impact on death. The estimated effect can be mainly explained by these group differences. This can be seen by the effect estimates for the different approaches addressing imbalances.

It strikes, that there is a difference in age groups. In the younger population there is no difference in the approaches and the confidence intervals are wide. Note, that there are only very few deaths in this population, similar in both groups. However, looking at the older population there are some minor differences between the approaches. Most of the deaths occur in this population. The increase in imbalances between groups after selection mainly comes from the elderly. In this population the infected group is more robust due to frailer patients dying in the initial 90 days after infection. While there is no such effect in the vaccinated group (see Supplementary Figure S4). Hence, the huge unadjusted effect in the overall population mainly comes from the older population. Addressing the imbalances via the different approaches reduce the estimated effect.

Considering the outcome SARS-CoV-2 infection, there is mainly a difference between age groups. In the overall population there is an effect of vaccination vs. infection throughout the approaches (not always statistically significant). This effect comes from the younger population. In this population the vaccinated group has a disadvantage compared to the infected group. This is in contrast to no difference being found in the elderly population.

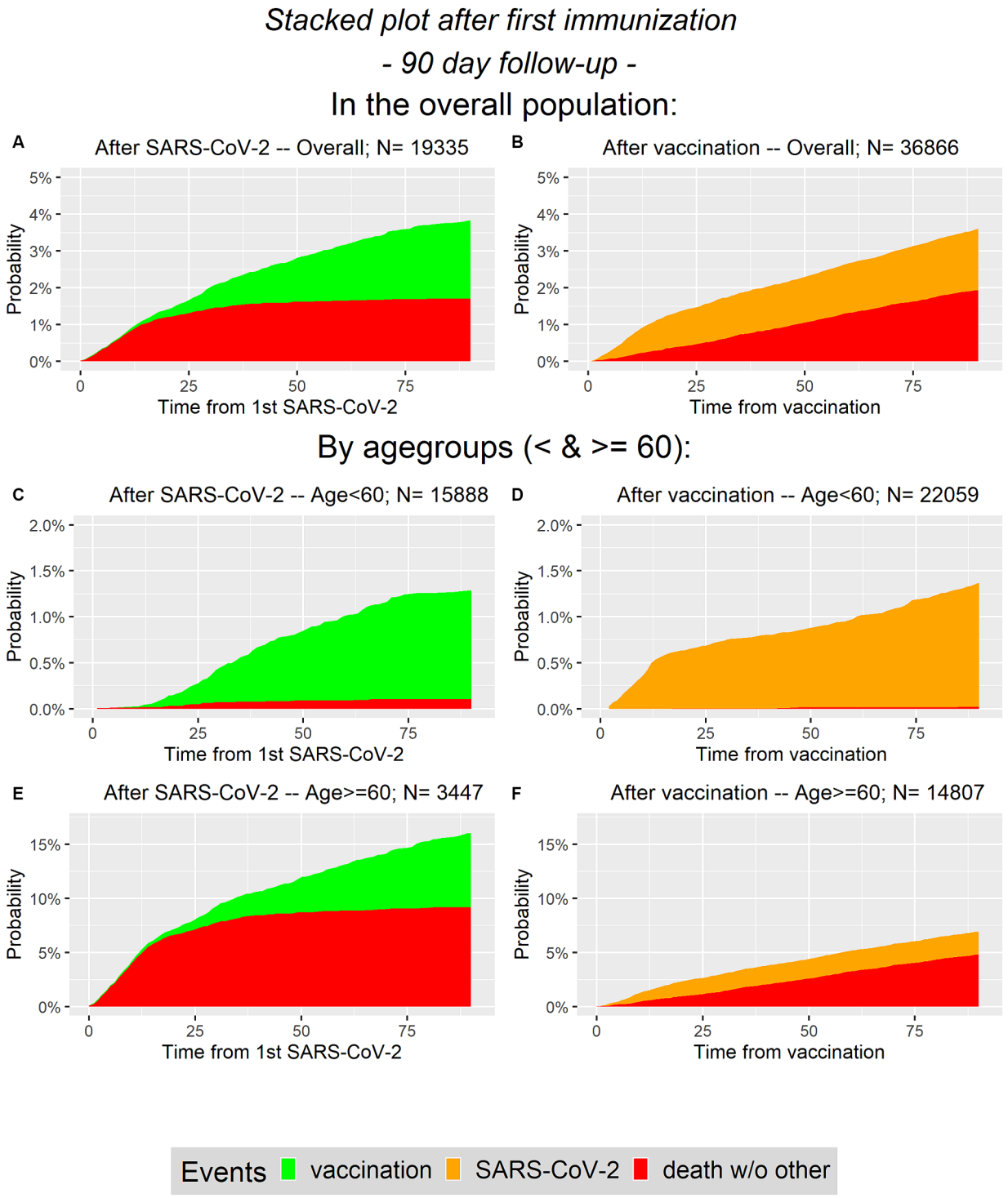

FIGURE 2
Cumulative probability plots for analysis of first 90 days. On the left panel **(A,C,E)** the infected group is presented and on the right panel **(B,D,F)** the vaccinated group is considered. In the first line **(A,B)** the overall population is presented, while the second and the third line represent the age groups (**C–F**, respectively).

Concerning the issue of competing risks there are no big differences between the estimates of the HR and the OR. Hence, the topic of competing risks is not a big issue (at least in this data example). Note that the event rates during the second 90 days are low in each age group and for each event.

Discussion

In this work we investigated challenges when comparing the protection rate for SARS-CoV-2 infection after an infection or after vaccination. We divided the time after first immunization into the

TABLE 3 Results of regression analysis for death without other event during the first 90 days after first immunization in the overall population and in age groups.

		Death without other event (vaccination vs. infection)	
Population		HR	OR
Overall	Unadjusted	1.118 [0.981;1.274]	1.153 [1.011;1.316]
	Adjusted	0.166 [0.144;0.192]	0.216 [0.184;0.253]
<60	Unadjusted	0.222 [0.081;0.607]	0.225 [0.074;0.574]
	Adjusted	0.153 [0.055;0.429]	0.185 [0.059;0.482]
≥60	Unadjusted	0.47 [0.411;0.537]	0.509 [0.444;0.585]
	Adjusted	0.171 [0.148;0.198]	0.31 [0.267;0.361]

The infection is considered as the reference group. HR, hazard ratio; OR, odds ratio.

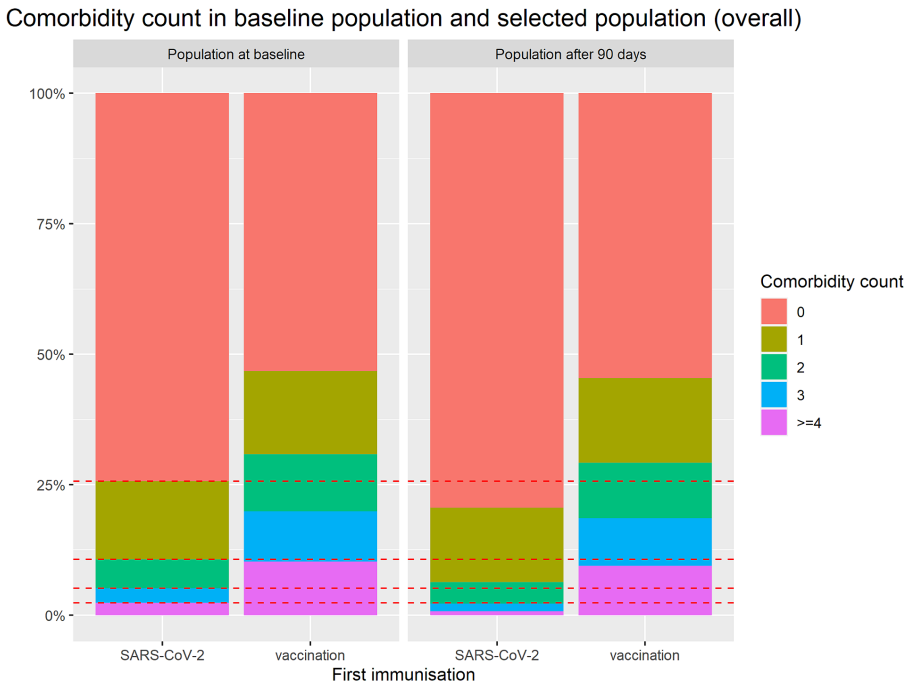

FIGURE 3
Comparison of comorbidity count between groups for baseline population and selected population available for conditional analysis after first 90 days.

first and the second 90 days of follow-up. Due to the epidemiological definition, a comparison of protection rates after infection or vaccination has to start after the first 90 days. This results in a selection process for the analysis. We have seen that this selection process differs between groups. While in the vaccinated group there were not many changes concerning the distribution of the baseline covariates in the selected population compared to the baseline population, in the infected group there were substantial differences. The selection process in the infected population resembles a "survival of the fittest" scenario. The two groups (vaccinated and infected) already had imbalances at baseline and the selection process intensifies this.

The considered approaches - adjustment, matching and weighting - obviously cannot address the selection process itself in an explicit way, as only the selected population is considered. However, it is important to consider the differences in mortality during the first 90 days when reporting the effects after 90 days. This will enable decision-makers to make a careful risk assessment.

We strongly recommend not to only start the analysis after 90 days and ignore the first 90 days, but rather investigate the selection process itself.

There are numerous approaches to address the imbalances between groups. In this work we presented adjustment, general full matching, and weighting addressing ATE as an estimand. Of course there are more possibilities to address imbalances in regression analyses. For instance, there are already several ways of performing matching or weighting, see (13). When choosing a method, one should be aware of what estimand is being addressed, in order to interpret the resulting estimation. It is crucial to be aware of differences between groups and addressing them in the analysis. Especially if the group comparison considers a vaccinated group versus a non-vaccinated group, it is possible that the data is prone to the

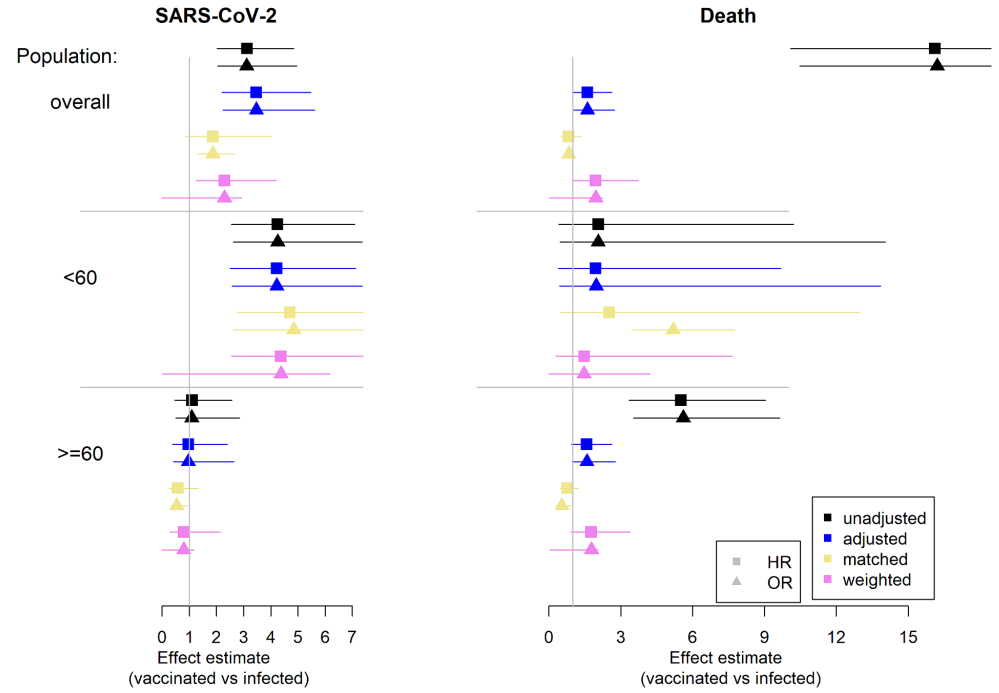

FIGURE 4
Estimation of effect estimates (HR and OR): results from regression analysis. Infection is considered as reference group. On the left column the outcome SARS-CoV-2 infection is considered and on the right column the competing event death is considered (HR, hazard ratio, OR, odds ratio).

healthy vaccine bias, see (14, 15). This bias occurs if the vaccinated group is in general more healthy than the comparison group. One way to investigate the healthy vaccine bias is comparing the non-COVID mortality in both groups. However, in order to do this, the cause of death needs to be known. Note, that in our data, the infected group, which is the non-vaccinated group, is less morbid than the vaccinated group, based on age and comorbidity count at baseline.

Furthermore, the competing risk has to be taken into account. Using data from Stockholm, the impact of the competing risk did not lead to a large discrepancy between ORs and HRs. An explanation might be that there are only few infections during the second 90 days, and even fewer deaths. Hence the event rates are low for each of the competing events. However, from a patient's point of view it is important to get information about both events. Therefore, we advise to always report both, i.e., measures on rate and on the risk scale, for the outcome and the competing event death in order to get a complete picture of the risk dynamic.

For our investigation we were able to use an extensive dataset from Stockholm County with information on over 1.8 million people. This allowed us to consider only a small inclusion window for this study and still have a sample size of over 52.000. With this inclusion window we ensured that there was no notable variation in the virus at time of immunization and in the time considered for infection after immunization (i.e., the second 90 days after immunization).

A limitation of this investigation is that we did not distinguish between different levels of vaccination. For simplicity we only considered vaccinated as being at least vaccinated once and without previous infection. For the illustrative purpose of this work, this is a justifiable simplification. Allowing for more complexity in the determination of immunization groups requires more thought concerning the time at risk considered in the analysis comparing the groups. However, since the analysis is limited to short time frames (the first and second 90-day periods after initial vaccination), the consideration of extra doses is of minor importance.

Furthermore it needs to be noted, that we did not distinguish between SARS-CoV-2 related deaths and non-related deaths. While this differentiation is not important for the purpose of this work, it is quite important for clinicians and patients and should be incorporated when investigating related research questions.

It is important to note that infections in this dataset are only identified when individuals are tested. Unfortunately, we do not have data on testing frequencies. Hence we cannot compare them between the infected and the vaccination group in order to see whether this is similar.

Even though the progress of the pandemic has led to changes in the underlying populations, the topic of this work remains relevant. Over time, the number of people with numerous infections and vaccinations has increased. Hence the analysis has become more complex. Nevertheless, the initial problem is still present. If the main analysis starts after 90 days, and thus there is a selection process during the first 90 days, it is crucial to take this first period into account as there might be differences between groups. Hence it is necessary to evaluate this problem in a simple setting in order to get a better understanding.

In conclusion, for a comparison between immunization via infection or vaccination, we strongly emphasize to consider several investigations in order to make fair comparisons and to draw comprehensive conclusions. Information on the selection process for the main analysis should be investigated and reported in the publication, namely the first 90 days. It is essential to present both in order to avoid biased interpretation.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: the individual participant data underlying this article were subject to ethical approval and cannot be shared publicly. Data from the deidentified administrative health registries are not freely available due to protection of the personal integrity of the participants. Requests to access these datasets should be directed to PN, pontus.naucler@ki.se.

Ethics statement

The studies involving humans were approved by Swedish Ethical Review Authority (Dnr 2018/1030–31, COVID-19 research amendments Dnr 2020–01385, 2020–02145, 2020–04069 and 2022–02127-02). The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because the need for consent was waived by the Swedish Ethical Review Authority (Dnr 2018/1030–31, COVID-19 research amendments Dnr 2020–01385, 2020–02145, 2020–04069 and 2022–02127-02) since analyses are based on retrospectively collected data from the administrative health registry.

Author contributions

SW: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. PH: Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. PN: Conceptualization, Supervision, Writing – original draft, Writing – review & editing. MW: Conceptualization, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The work was supported by a grant from the Swedish Research Council (Dnr 2021–04809). This study has been funded by the German Research Foundation (original: Deutsche Forschungsgemeinschaft, DFG) with project grant WO 1746/5–1.

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/fmed.2024.1376275/full#supplementary-material

References

1. Chemaitelly H, Ayoub HH, AlMukdad S, Coyle P, Tang P, Yassine HM. Protection from previous natural infection compared with mRNA vaccination against SARS-CoV-2 infection and severe COVID-19 in Qatar: a retrospective cohort study. *Lancet Microbe*. (2022) 3:e944–55. doi: 10.1016/S2666-5247(22)00287-7

2. Lind ML, Dorion M, Houde AJ, Lansing M, Lapidus S, Thomas R. Evidence of leaky protection following COVID-19 vaccination and SARS-CoV-2 infection in an incarcerated population. *Nat Commun*. (2023) 14:5055. doi: 10.1038/s41467-023-40750-8

3. Gail MH, Benichou J, Herausgeber C. *Encyclopedia of epidemiologic methods*. New York: Wiley (2000). 978 p.

4. Ranganathan P, Aggarwal R, Pramesh C. Common pitfalls in statistical analysis: odds versus risk. *Perspect Clin Res*. (2015) 6:222–4. doi: 10.4103/2229-3485.167092

5. Letizia AG, Ge Y, Vangeti S, Goforth C, Weir DL, Kuzmina NA. SARS-CoV-2 seropositivity and subsequent infection risk in healthy young adults: a prospective cohort study. *Lancet Respir Med*. (2021) 9:712–20. doi: 10.1016/S2213-2600(21)00158-2

6. Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MAUA. BNT162b2 mRNA Covid-19 vaccine in a Nationwide mass vaccination setting. *N Engl J Med*. (2021) 384:1412–23. doi: 10.1056/NEJMoa2101765

7. Powell AA, Kirsebom F, Stowe J, Ramsay ME, Lopez-Bernal J, Andrews N. Protection against symptomatic infection with delta (B.1.617.2) and omicron (B.1.1.529) BA.1 and BA.2 SARS-CoV-2 variants after previous infection and vaccination in adolescents in England, august, 2021–march, 2022: a national, observational, test-negative, case-control study. *Lancet Infect Dis*. (2023) 23:435–44. doi: 10.1016/S1473-3099(22)00729-0

8. Hedberg P, Granath F, Bruchfeld J, Askling J, Sjöholm D, Fored M. Post COVID-19 condition diagnosis: a population-based cohort study of occurrence, associated factors, and healthcare use by severity of acute infection. *J Intern Med Februar*. (2023) 293:246–58. doi: 10.1111/joim.13584

9. Rolfhamre P, Jansson A, Arneborn M, Ekdahl K. SmiNet-2: description of an internet-based surveillance system for communicable diseases in Sweden. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull*. (2006) 11:15–6. doi: 10.2807/esm.11.05.00626-en

10. Huang L, Lai FTT, Yan VKC, Cheng FWT, Cheung CL, UA CCSL. Comparing hybrid and regular COVID-19 vaccine-induced immunity against the omicron epidemic. *Npj Vaccines*. (2022) 7:162. doi: 10.1038/s41541-022-00594-7

11. Socialstyrelsen. Analys om pandemins effekter på vården (2020-2021). Available at: https://www.socialstyrelsen.se/statistik-och-data/statistik/pandemins-effekter-pa-varden/analys-uppdamda-vardbehov-efter-pandemin/

12. Hernán MA, Clayton D, Keiding N. The Simpson's paradox unraveled. *Int J Epidemiol Juni*. (2011) 40:780–5. doi: 10.1093/ije/dyr041

13. Greifer N, Stuart EA. Choosing the Estimand when matching or weighting in observational studies. (2021); Available at: https://arxiv.org/abs/2106.10577

14. Høeg TB, Ram D, Vinay P. Potential "healthy Vaccinee Bias" in a study of BNT162b2 vaccine against Covid-19. *N Engl J Med*. (2023) 389:284–6. doi: 10.1056/NEJMc2306683

15. Fürst T, Bazalová A, Fryčák T, Janošek J. Does the healthy vaccinee bias rule them all? Association of COVID-19 vaccination status and all-cause mortality from an analysis of data from 2.2 million individual health records. *Int J Infect Dis Mai*. (2024) 142:106976. doi: 10.1016/j.ijid.2024.02.019

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 17 June 2024
DOI 10.3389/fmed.2024.1390549

OPEN ACCESS

EDITED BY
Reza Lashgari,
Shahid Beheshti University, Iran

REVIEWED BY
Seetha Harilal,
Kerala University of Health Sciences, India
Javad Hassannataj Joloudari,
University of Birjand, Iran

*CORRESPONDENCE
Elisabeth Lucke
✉ anne.elisabeth.lucke@itmp.fraunhofer.de

RECEIVED 23 February 2024
ACCEPTED 30 May 2024
PUBLISHED 17 June 2024

CITATION
Lucke E, Hazard D, Grodd M, Weber S and Wolkewitz M (2024) Lessons learned: avoiding bias via multi-state analysis of patients' trajectories in real-time.
Front. Med. 11:1390549.
doi: 10.3389/fmed.2024.1390549

Lessons learned: avoiding bias via multi-state analysis of patients' trajectories in real-time

Elisabeth Lucke*, Derek Hazard, Marlon Grodd, Susanne Weber and Martin Wolkewitz

Institute of Medical Biometry and Statistics, University Hospital Freiburg, Freiburg, Germany

Objectives: Many studies have attempted to determine the disease severity and patterns of COVID-19. However, at the beginning of the pandemic, the complex patients' trajectories were only descriptively reported, and many analyses were worryingly prone to time-dependent-, selection-, and competing risk biases. Multi-state models avoid these biases by jointly analysing multiple clinical outcomes while taking into account their time dependency, including current cases, and modelling competing events. This paper uses a publicly available data set from the first wave in Israel as an example to demonstrate the benefits of analysing hospital data via multi-state methodology.

Methods: We compared the outcome of the data analysis using multi-state models with the outcome obtained when various forms of bias are ignored. Furthermore, we used Cox regression to model the transitions among the states in a multi-state model. This allowed for the comparison of the covariates' influence on transition rates between the two states. Lastly, we calculated expected lengths of stay and state probabilities based on the multi-state model and visualised it using stacked probability plots.

Results: Compared to standard methods, multi-state models avoid many biases in the analysis of real-time disease developments. The utility of multi-state models is further highlighted through the use of stacked probability plots, which visualise the results. In addition, by stratification of disease patterns by subgroups and visualisation of the distribution of possible outcomes, these models bring the data into an interpretable form.

Conclusion: To accurately guide the provision of medical resources, this paper recommends the real-time collection of hospital data and its analysis using multi-state models, as this method eliminates many potential biases. By applying multi-state models to real-time data, the gained knowledge allows rapid detection of altered disease courses when new variants arise, which is essential when informing medical and political decision-makers as well as the general population.

KEYWORDS

lessons learned, avoiding bias, pandemic preparedness, multi-state models, real time, analysis strategies

1 Introduction

1.1 Background

Having emerged in December of 2019, the SARS-CoV-2 virus has brought with it a variety of challenges. Due to its diverse clinical courses and surging waves of patients, it has impeded the provision of appropriate resources for hospitalised patients. At the beginning of the pandemic, studies attempting to understand the characteristics of COVID-19 suffered from severe time-related types of bias due to length bias, immortal-time bias, competing risk bias, and selection bias (1). For example, Zhou et al. carried out a study in Wuhan, China, in which 613 cases out of 813 were excluded because these patients had not yet experienced an outcome (2). Similarly, in Chen et al. 525 out of 799 patients were excluded from the analysis because they were still hospitalised (3). However, as Bajaj et al. argue, those in a moderate condition may stay in the hospital for longer than those in a poor condition, because the latter may succumb more quickly (4). At the same time, those in a good condition stay in the hospital for a shorter time than the ones in a moderate condition, as the former are likely to be discharged sooner. Thus, by ignoring all active cases, a selection bias arises, in which patients in a moderate condition are excluded, as their stay in the hospital is likely to be longest.

In addition, various studies suffered from competing risk bias. Competing risks refer to situations where an individual is subject to multiple possible events, and the occurrence of one event precludes the occurrence of the others (5). Assuming death is the event of interest, for example, the possibility of in-hospital death is eliminated when an individual is discharged from the hospital. Hence, being discharged is a competing event to dying in the hospital. In survival analysis, disregarding the presence of competing events can lead to a severe bias in the results.

To address the problem of poor data quality and biassed samples, this paper shows how statistical analyses can be used to avoid these biases in the context of the COVID-19 pandemic, following the example of Hazard et al. with multi-state models (6). Multi-state models entail defining certain states and the transitions among them. Multi-state models have the advantage that they are very flexible. For example, depending on the desired complexity, states can easily be consolidated. This increases the comprehensibility of the plots and at the same time can simplify the analysis (7).

1.2 Research in context

Multi-state models have been used in a variety of research contexts. For example, in modern ecology, it is used in capture-recapture experiments because the multi-state models allow for simple incorporation of temporal variation in the transition rates by modelling the rates as a parametric function over time (8). Furthermore, multi-state models are commonly used in cancer clinical trials, where patients usually experience multiple disease stages (9). The complex transitions between these stages can be comprehensively analysed using multi-state models. In addition, multi-state models can be used for predictions. More specifically in the context of hospital data, Roimi et al. used a multi-state model to predict individual patients' hospital states based on their characteristics, such as age and gender. Furthermore, they also carried out analyses to predict the total hospital utilisation (10). Similarly, Keogh et al. also predicted the length of stay in hospital wards during the COVID-19 pandemic based on the patients' characteristics. However, they extended their work by introducing the concept of "conditional expected length of stay," which is defined as the expected length of stay in a certain state, conditional on the complete pathway taken through the states (11). In yet another paper, the multi-state model is analysed with parametric methods, which has the advantage that these parameters can be used to carry out simulations (12). The advantages and disadvantages of a variety of multi-state modelling approaches are reviewed in (13).

The implication of disregarding competing events in statistical analyses has been discussed frequently among the research community. In (14), McCaw et al. outline the problem of competing risks based on two papers: in (15), Beigel et al. carried out a clinical trial evaluating the effect of remdesivir versus a placebo in hospitalised COVID-19 patients. Similarly, Li et al. conducted a trial to detect the effect of convalescent plasma as compared to the effect of the standard of care on hospitalised COVID-19 patients (16). In both studies, death is a competing event. In addition, Wolkewitz et al. carried out an analysis to determine the impact of the duration of mechanical ventilation on the development of pneumonia while considering extubation as a competing event (17). An unbiased result could only be obtained when the competing event was accounted for. Supplementary Table 2 in (18) shows an overview of papers published in high-impact journals with a competing risk problem. In all cases, being discharged alive was the competing event that should have been accounted for in the analysis. Ignoring this competing event led to an overestimation of the cumulative incidence of the event of interest which, in the cases of the papers mentioned in Supplementary Table 2, was death or a composite outcome of intubation or death. Furthermore, a systematic literature review of observational studies that evaluated drug effectiveness in patients with COVID-19, carried out by Martinuka et al. (19), assessed the studies on three common methodological pitfalls in time-to-event analyses, one of them being competing risk bias. Their results showed that only one paper out of 11 accounted for the competing risk of being discharged alive by extending the follow-up period for discharged patients. All the others suffered from a competing risk bias. This highlights the scope of the problem.

Whilst it is evident that the topic of competing risk and selection bias as well as the use of multi-state models for the analysis of hospital data is not new to the research community, the topics are rarely taken into consideration by clinicians. Hence, this work aims to illustrate the biases that early COVID-19 analyses were subject to and provide a simple and easily applicable solution to overcoming these sources of bias by using the multi-state methodology. In addition, the paper aims to show how the continuous use of multi-state models in hospital data analysis facilitates hospital planning in disease outbreak scenarios by using comprehensive data visualisation techniques, thereby enhancing pandemic preparedness.

2 Methods

2.1 Data

The data used to demonstrate the advantages of multi-state models was collected in the form of a nationwide Israeli COVID-19

Abbreviations: ELOS, Expected length of stay; M/S, Moderate/severe; ICU, Intensive care unit.

registry. It was previously used by Roimi et al., who conducted a multi-state analysis to predict hospital capacity utilisation in Israel. The data was collected in real-time and includes the day-to-day clinical course of patients hospitalised for at least 1 day between March 1st and May 2nd of 2020. It furthermore includes information on the patients' age, sex, and initial admission date.

2.2 Multi-state models

In the model, four states are defined: moderate/severe (M/S), critical, discharge, and death. A patient starts in M/S or critical and can change between these two states an unlimited number of times before either dying in the hospital or being discharged. Death and discharge were defined as absorbent states, forbidding any transition away from these states. Supplementary Figure 1 illustrates the multi-state model. One important characteristic of multi-state models is that they allow the inclusion of competing events. In this model, discharge is the competing event of death. It is important to classify discharge as a competing event because being discharged alters the probability of death, i.e., persons discharged from the hospital are likely to be healthier and therefore have a lower probability of dying than those hospitalised (5). Disregarding this in the analysis would bias the results. Supplementary Table 1 shows example clinical courses of 3 patients through this multi-state model.

As von Cube et al. described, the analysis of multi-state models implies the calculation of transition probabilities and transition-specific hazard rates (20). In this paper, transition probabilities are calculated using transition hazards, which are defined as the instantaneous risk of moving between two states. Moreover, in our model, the calculation of the transition probabilities is dependent on all hazard rates of the transitions. Further mathematical details of multi-state models are explained in von Cube et al. (20) and Wolkewitz et al. (21).

2.3 Statistical analyses

We used two different approaches to highlight the advantages of multi-state models over standard analysis techniques. In the first approach, we illustrated the bias which arises by excluding all active cases as was done in Zhou et al. and Chen et al. (2, 3). To do so, we excluded all active cases from our data and carried out a logistic regression for the outcome of "death." Based on this regression, we predicted the probability for the event "death" to have occurred by May 2nd, 2020 for the different age groups. We stratified this analysis for the state at initial admittance. We then compared the results to the probability of dying using the multi-state model to highlight the discrepancies in the results if active cases are excluded. In the second approach, we demonstrated the bias that arises if competing events are censored. This means that only the event of interest, death, was considered. We created cumulative incidence curves for this model and compared them to the cumulative incidence curves obtained when considering the competing event of being discharged. This analysis was also stratified by age groups and the initial state of admittance. The methodology is displayed in Figure 1.

In addition, we analysed the multi-state model using stacked probability plots. Transition probabilities were calculated using the *mstate* package in R (Version 4.3.1) and the code created by Hazard et al. (6). First, cause/transition-specific Cox regressions were calculated. Cox models are a popular regression method in survival analysis. They are used when the effect of covariates on censored survival times is analysed. Cox regressions are calculated to compare how covariates affect the instantaneous risk of a transition between two states, i.e., the hazard ratio. For simplicity, the transitions M/S to critical and M/S to death were merged into one transition, to model how covariates affect the risk of clinical decline from the M/S state. Similarly, the transitions critical to M/S and critical to discharge were merged into one transition to model the effect of the covariates on an improvement from the critical state. To measure the effect of the

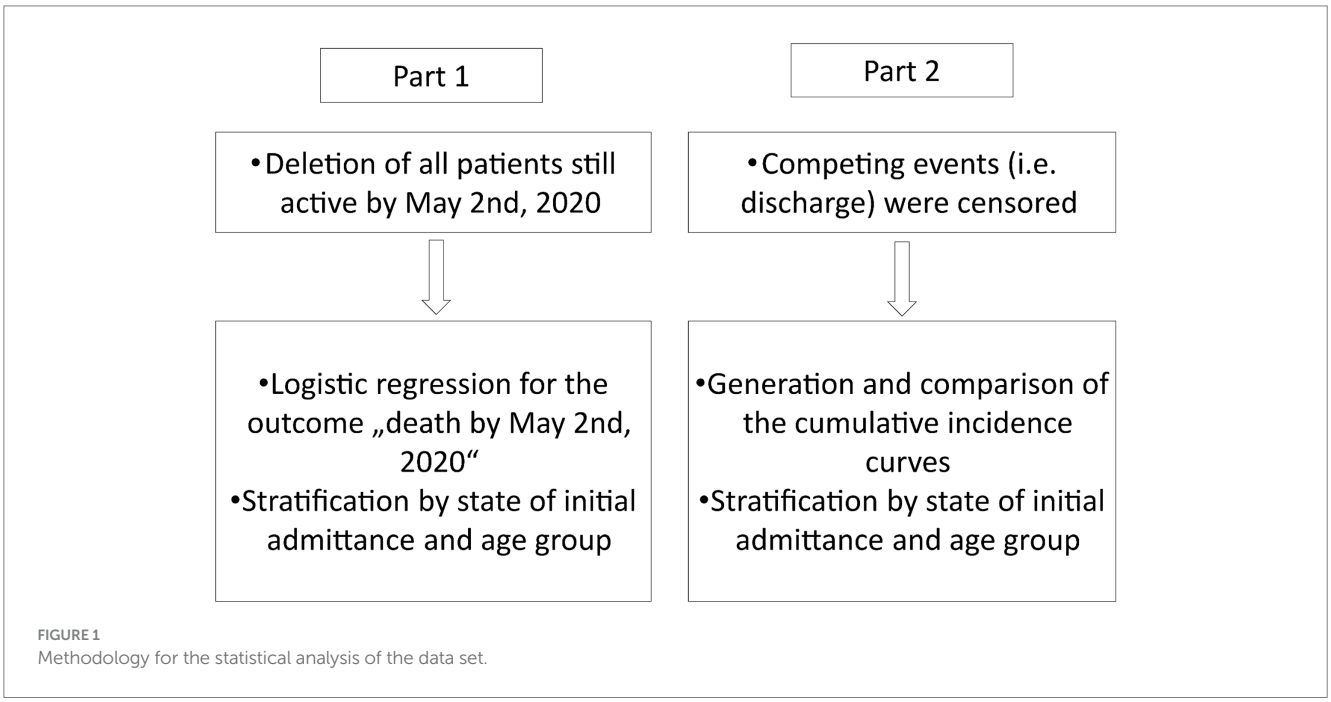

FIGURE 1
Methodology for the statistical analysis of the data set.

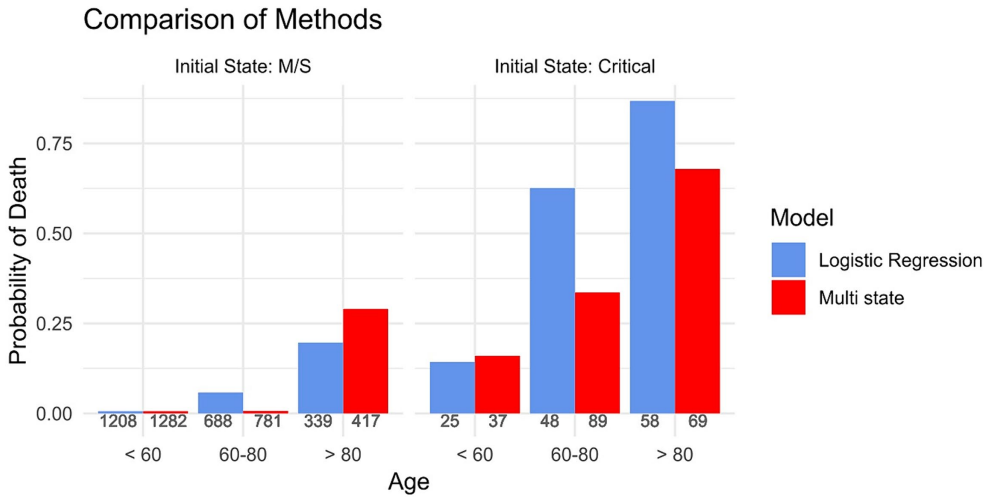

FIGURE 2
The probability of dying stratified by the initial state of admittance and age group, calculated once based on a subsample that excluded all active cases (Logistic Regression) and once based on the full cohort using multi-state methods (Multi-state). The numbers underneath each bar indicate the amount of people in the respective group.

month of admittance, a binary variable was used differentiating between admission in March and admission in April. The few cases that were admitted in May were included in April. After performing Cox regression, the baseline hazard was calculated, which is the hazard of the event of interest occurring at a certain point in time if the effect of all other covariates is zero. Then, the transition probabilities were calculated based on the baseline hazards. Finally, the transition probabilities were used to calculate the expected length of stay (ELOS) as shown in Hazard et al.

3 Results

The dataset included 2,675 patients of which 1,319 were younger than 60, 870 were between 60 and 80, and 486 were over the age of 80. 2,480 patients were originally admitted in an M/S state and only 195 patients were in a critical state at the time of admittance. By May 2nd, 2020, 198 patients had died and 311 patients (11.6%) were still active. Hence, by excluding all active cases, the data was reduced to 2,364 patients with 1,233 below 60, 734 between 60 and 80, and 397 over the age of 80. Of the 2,364 patients, 2,233 were initially in an M/S state and 131 in a critical state.

3.1 Bias in research during the COVID-19 pandemic

3.1.1 Selection bias

Figure 2 depicts the predicted 30-day hospital mortality obtained using a logistic regression model when excluding all active cases from the data. As a comparison, it also shows the predicted 30-day hospital mortality obtained when the entire cohort was analysed using multi-state models. The graph shows that when excluding active cases from the analysis the probability of death is overestimated in the groups of patients where more deaths occurred, such as in the older age groups and in those who were initially admitted in a critical state. For example, when analysing the biassed cohort, the probability of death of the patients between 60 and 80 who were initially admitted in a critical state is 0.63. In comparison, when including all individuals initially admitted in a critical state, the probability of death is 0.34. The same pattern is seen for patients above 80 when initially admitted in a critical state. All numeric values of the 30-day hospital mortality can be found in Supplementary Table 6.

3.1.2 Competing risk bias

Figure 3 shows the cumulative incidence curves of two different models. Whilst the event of interest in both models was "death," the event "discharged" was only classified as a competing event in one model ("Accounting for competing risks"). In the other, this competing event was ignored ("Ignoring competing risks"). The results show that especially for the patients who were initially admitted into a moderate/severe state, there are large discrepancies in the cumulative incidence curves between these two models. The method where competing risks are ignored overestimates the cumulative incidence of the event "death." For example, the cumulative incidence of death after 30 days when ignoring the presence of competing risks is 0.37 whereas the cumulative incidence when considering the competing risk is 0.24.

3.2 Advantages of multi-state methods to avoid bias

Having demonstrated the bias that arises through standard methods of analysis that were used at the beginning of the COVID-19 pandemic, the following results highlight the advantages of using multi-state models.

3.2.1 Planning bed capacity

Figure 4 shows the estimated probabilities in each state over time stratified for age groups and initial state of admittance. The

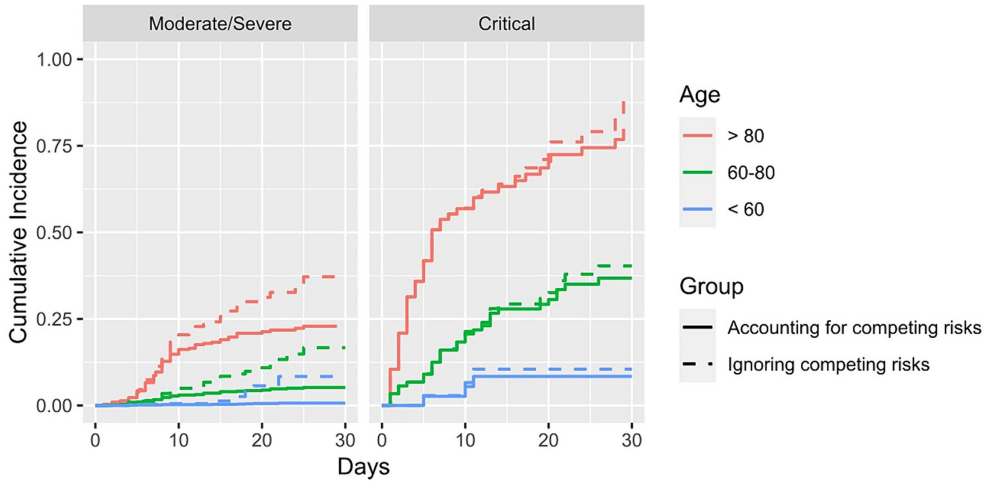

FIGURE 3
Cumulative incidence of dying stratified for age group and initial state of admittance. Solid lines show the results in which the competing risk of being discharged is accounted for, whereas the dashed lines show the results where the competing risk was not accounted for.

estimated probability of each state on any day can be determined. Figures 4A–C shows the predicted proportions of patients admitted in a moderate/severe condition in each state over time stratified by age group. For patients between 60 and 79, 30 days after hospitalisation 3.5% are in M/S, 4.7% are in critical, 85.9% are discharged and 5.9% are dead. This information is valuable for the standard care units. Similarly, Figures 4D–F shows the estimated probabilities in each state over time of those patients being admitted in a critical state. This information is relevant for the intensive care wards as it indicates how long patients stay in a particular condition and more importantly, in which condition the patients leave the ward. For example, for patients between 60 and 79, 30 days after hospitalisation 8.5% are in M/S, 16.4% are in critical, 45.5% are discharged, and 29.6% are dead. Thus, not only are these graphs useful when it comes to the planning of the different wards in the hospital, but by comparing the wards (standard care and ICU) with one another they can be useful in identifying disease patterns.

In addition, four different Cox regressions were constructed based on the model. As explained, for simplification purposes transitions one and three (as depicted in Supplementary Figure 1) were merged into one model as well as transitions four and five. The age group, sex, and initial admission date were included in the model as covariates. Overall, the age group and the binary covariate indicating the admission date (March versus April) were significant in most regressions. An increased age was associated with an increased hazard rate of transitions from M/S to critical or death, and from critical to death. The full result of the regressions is included in the Supplementary Tables 2–5 and it shows how multivariable Cox regressions can be used as an outlook to analyse how certain characteristics are risk factors for a specific transition.

Furthermore, to show how the multi-state models can be used to prepare healthcare providers for future COVID-19 waves, the ELOS in the two non-absorbent states was calculated as an example for patients between 60 and 80. For those admitted in a M/S condition, the ELOS in M/S is 9.28 days and the ELOS in the critical state is 2.75 days. In contrast, those admitted in a critical condition have an ELOS of 4.96 days in the M/S state and an ELOS of 15.91 days in the critical state, when estimating from the first day of hospitalisation. Supplementary Figure 2 shows the days in the M/S and critical state for each age group of the entire cohort.

3.2.2 Analysis to study the most recent developments in real-time

In addition, multi-state models allow the study of the most recent disease developments in real time. As an example, Figure 5 shows the stacked probability plots for the whole population stratified by the admittance date and initial admittance state. By stratifying for the hospital admission date, we show real-time changes in the clinical patterns of COVID-19. For those admitted in the M/S state, the mortality is higher when admitted in April/May than in March. This, however, is different when starting in the critical state. Here, the mortality seems to be lower for those admitted in April/May and the estimated probabilities over time in the critical condition are lower than for those admitted in March. Such information helps to clarify any observed differences in severity between the regular and the intensive care ward.

4 Discussion

In this paper, we compared standard survival analysis methods used for the analysis of COVID-19 hospital data in the beginning of the pandemic with advanced multi-state models. Supplementary Figure 2 and Supplementary Table 6 showed that selection bias led to an overestimation of death in the groups where many deaths occurred. As explained above, Bajaj et al. pointed out that excluding active cases biases the cohort towards the very ill and those that are only very lightly diseased. In our analysis, we manage to separate the severely ill from those that have a light course of the disease by stratifying for the initial state of admittance. Hereby, we show that our results support the claim of Bajaj et al., as Figure 2 shows that the probability of dying is overestimated in the patients admitted in a critical state, thus in the severely ill patients. These results are only partly reproducible in the opposite sense, i.e., showing that

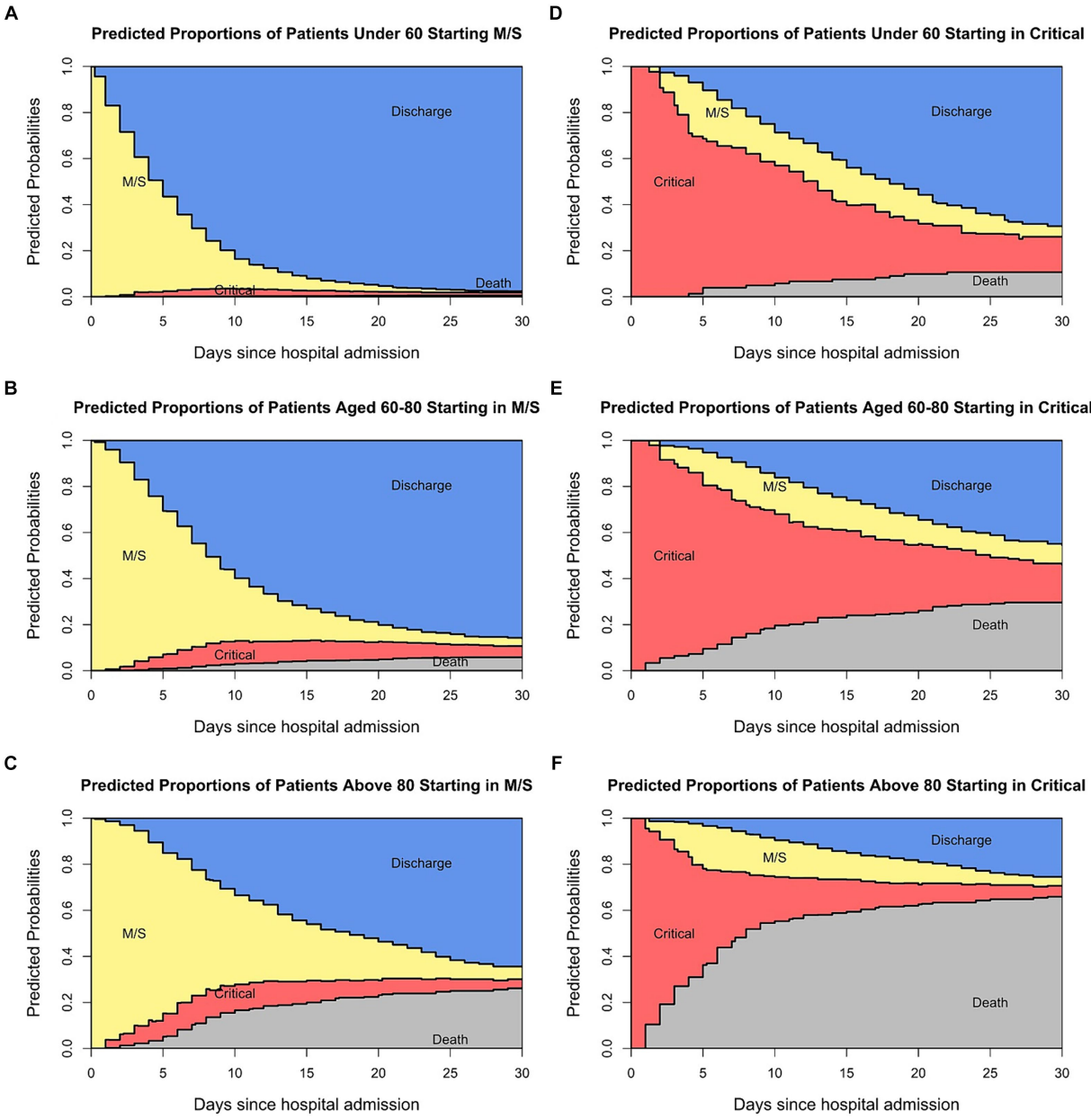

FIGURE 4
Predicted probabilities of being in each state at specific times for the patients admitted in a moderate state **(A–C)** and those admitted in a critical state **(D–F)**.

the probability of death is underestimated in the patients initially admitted in a M/S state. Whilst we see an underestimation for patients above 80 admitted in an M/S state, this cannot be observed for the other age groups. This could perhaps be explained by the small percentage of patients admitted in M/S that died, leading to an imprecise prediction of the probability of dying. In addition, Figure 3 shows that competing risk bias also leads to an overestimation of death because when ignoring competing risks, the model does not differentiate between being discharged and being hospitalised. Thus, those discharged are assumed to have the same risk of the event of interest as those who are still hospitalised. Consequently, the cumulative incidence is overestimated. In contrast, the cumulative incidence curves for the patients initially admitted in the critical state are very similar, as fewer patients are discharged. These results suggest that numerous analyses carried out at the beginning of the pandemic overestimated the severity of the disease. This is relevant as the overestimation may have led to some of the harsh public health measures, such as the closure of schools, which have, in retrospect, faced criticism for having been disproportionate (22). Hence, this example highlights the importance of obtaining unbiased information on disease severity, in which, as outlined above, multi-state models prove to be very useful.

In addition to highlighting the benefits of multi-state models in eliminating sources of bias, this paper also described further advantages of using multi-state models and corresponding stacked probability plots. One advantage is the potential these models have to assist the resource organisation of the hospital. By integrating stacked probability plots into the analysis, vital information on patients' clinical courses over time can be displayed comprehensively and concisely. Furthermore, through the use of Cox regressions and the

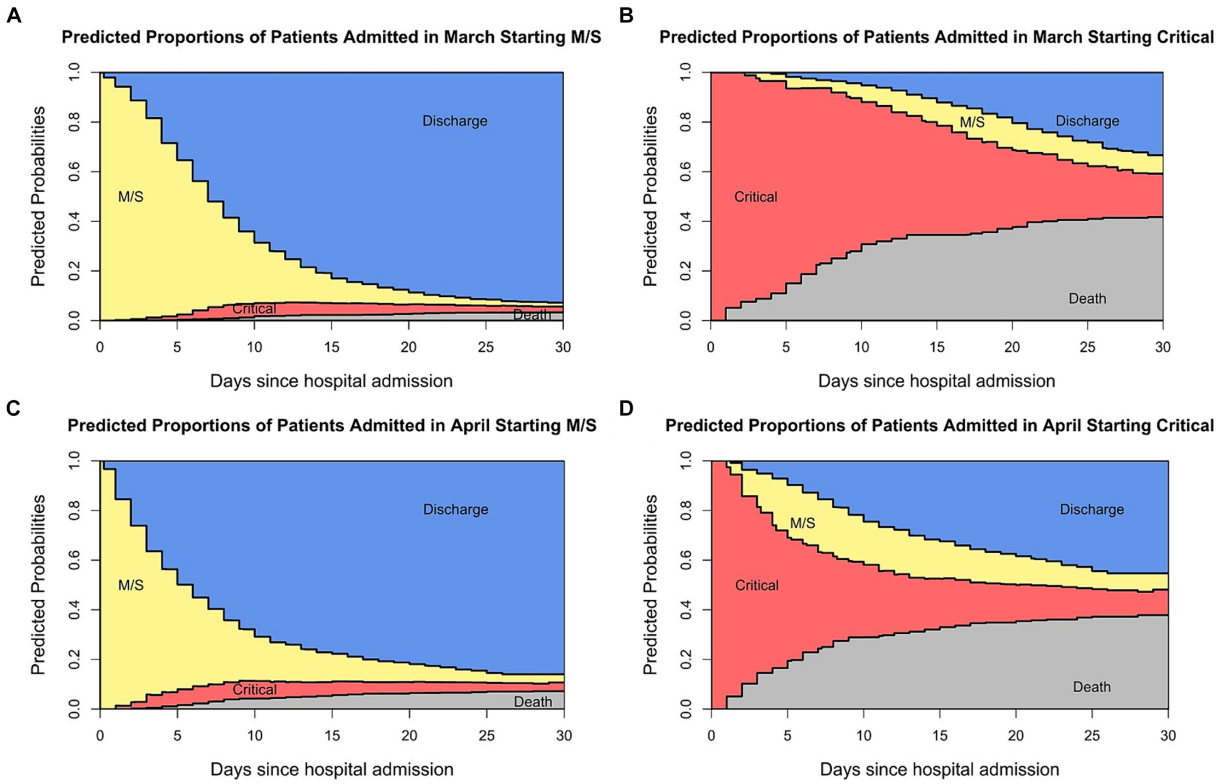

FIGURE 5
Stacked probability plots stratified by admittance date and initial state.

calculation of the ELOS, risk factors can be identified and the disease courses of individual patients can be predicted, thereby facilitating the planning of the hospital wards.

Additionally, stacked probability plots are easily interpretable and thereby facilitate communication with the general public. As van Schalkwyk et al. writes, the COVID-19 pandemic came at a time when distrust in institutions among the population was growing, for example through a change of government (23). This distrust worsened during the pandemic, as information from the media or research community was misunderstood by the population. However, stacked probability plots facilitate the interpretation of information conveyed by professionals. For example, Berger et al. carried out a country-level analysis of hospital capacity and utilisation. Besides measuring how different countries increased their ICUs as a response to COVID-19, they compared how long patients stayed in the ICU (24). However, the length of stay in the ICU can only be compared among countries if the mortality rate is the same. Otherwise, the comparison of ICU stations would not be meaningful because patients may leave the ICU due to death or due to discharge. The stacked probability plots manage to depict this idea by showing that the length of stay in the critical ward is determined by the occurrence of other states.

The utility of the plots to analyse real-time clinical patterns is especially highlighted in Figure 5, where the plots are stratified by admittance date. This is particularly clinically informative when new variants arise. Throughout the pandemic, various SARS-CoV-2 variants have emerged. The Delta variant was termed a variant of concern after its identification in India in May of 2021. It has increased transmissibility and virulence, as seen by elevated death and hospitalisation rates (25). However, there is no difference in characteristics between the wild-type virus and the Delta variant when compared by age and sex (26). In November of 2021, the Omicron variant was labelled a variant of concern. Whilst this variant showed a reduced severity overall, it led to increased hospitalizations in children under the age of 1 year (27). It is crucial to have this information in real-time because, based on such knowledge, policymakers could implement rules protecting small children and their parents, e.g., by allowing home office. Thus, it can be seen how multi-state models and stacked probability plots facilitate the communication of the disease and its real-time developments to the general public. This highlights that in the context of pandemic preparedness, it is imperative to collect high-quality hospital data continuously and promptly to understand the characteristics of the virus and to plan health care provision accordingly.

The strengths of this study are that it uses examples from the first pandemic wave to illustrate the extent to which the two forms of bias, selection bias, and competing risk bias, impact the results obtained from COVID-19 hospital data. Furthermore, the study provides an alternative approach that solves the shortfalls of standard methods and is therefore ideal for use in survival analysis in settings with more than one possible event. The limitation of the study is that it only includes data collected during the first pandemic wave. Our research would greatly benefit from analysing the clinical course of SARS-CoV-2 variants, as this would further illustrate the potential to detect differences in clinical characteristics using multi-state methods. However, whilst other data sets may exist of the time when variants were circulating, they could only be used if they are comparable to the

Israeli data set of the first wave in terms of the characteristics of the population and the hospital service. Otherwise, false conclusions would be drawn. As such comparable data was unavailable, further variants were not included in the analysis.

5 Future works and conclusion

With enhanced data collection across Europe, future works could aim to demonstrate the effectiveness of multi-state models in detecting differences in clinical courses over longer periods. In addition, in future disease outbreak scenarios, e.g., influenza outbreaks, additional covariates can be incorporated into the analysis. These variables could aim at capturing differences in the risk profiles between the patients. Examples include demographic factors and health access disparities. Upon integration of such factors, the prediction of hospital capacity utilisation will become more accurate and thus, more personalised care can be provided to the patients.

In summary, this paper shows that in the context of pandemic preparedness, it is crucial to collect the right type of data to carry out appropriate, unbiased analyses, and thus aid efforts to overcome further pandemic waves. Hence, by showing the simple but detailed analyses that can be carried out with routine hospital registries as collected in Israel, this paper aims to improve the statistical analysis techniques used, thus obtaining unbiased information on the disease of interest in a timely manner so that public health measures can be implemented accordingly.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

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

Author contributions

EL: Conceptualization, Formal analysis, Methodology, Visualization, Writing – original draft. DH: Data curation, Formal analysis, Methodology, Software, Visualization, Writing – review & editing. MG: Conceptualization, Data curation, Supervision, Writing – review & editing. SW: Supervision, Writing – review & editing. MW: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, 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

I would like to thank Roimi et al. for making their dataset publicly available.

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2024.1390549/full#supplementary-material

References

1. Wolkewitz M, Lambert J, von Cube M, Bugiera L, Grodd M, Hazard D, et al. Statistical analysis of clinical COVID-19 data: a concise overview of lessons learned, common errors and how to avoid them. *Clin Epidemiol*. (2020) 12:925–8. doi: 10.2147/CLEP.S256735

2. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. (2020) 395:1054–62. doi: 10.1016/S0140-6736(20)30566-3

3. Chen T, Wu D, Chen H, Yan W, Yang D, Chen G, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ*. (2020) 368:m1091. doi: 10.1136/bmj.m1091

4. Bajaj V, Sinha GR. *Computer-aided design and diagnosis methods for biomedical applications. Chapter 14 reliable diagnosis and prognosis of COVID-19. First* ed. Boca Raton: CRC Press (2021).

5. Wolkewitz M, Cooper BS, Bonten MJ, Barnett AG, Schumacher M. Interpreting and comparing risks in the presence of competing events. *BMJ*. (2014) 349:g5060. doi: 10.1136/bmj.g5060

6. Hazard D, Kaier K, von Cube M, Grodd M, Bugiera L, Lambert J, et al. Joint analysis of duration of ventilation, length of intensive care, and mortality of COVID-19 patients: a multistate approach. *BMC Med Res Methodol*. (2020) 20:206. doi: 10.1186/s12874-020-01082-z

7. von Cube M, Wolkewitz M, Schumacher M, Hazard D. Re: "the clinical course of coronavirus disease 2019 in a us hospital system: a multistate analysis". *Am J Epidemiol*. (2021) 190:1699–700. doi: 10.1093/aje/kwab044

8. Rushing CS. An ecologist's introduction to continuous-time multi-state models for capture-recapture data. *J Anim Ecol*. (2023) 92:936–44. doi: 10.1111/1365-2656.13902

9. Le-Rademacher JG, Peterson RA, Therneau TM, Sanford BL, Stone RM, Mandrekar SJ. Application of multi-state models in cancer clinical trials. *Clin Trials*. (2018) 15:489–98. doi: 10.1177/1740774518789098

10. Roimi M, Gutman R, Somer J, Ben Arie A, Calman I, Bar-Lavie Y, et al. Development and validation of a machine learning model predicting illness trajectory and hospital utilization of COVID-19 patients: a nationwide study. *J Am Med Inform Assoc*. (2021) 28:1188–96. doi: 10.1093/jamia/ocab005

11. Keogh RH, Diaz-Ordaz K, Jewell NP, Semple MG, de Wreede LC, Putter H, et al. Estimating distribution of length of stay in a multi-state model conditional on the pathway, with an application to patients hospitalised with COVID-19. *Lifetime Data Anal*. (2023) 29:288–317. doi: 10.1007/s10985-022-09586-0

12. Jackson CH, Tom BD, Kirwan PD, Mandal S, Seaman SR, Kunzmann K, et al. A comparison of two frameworks for multi-state modelling, applied to outcomes after hospital admissions with COVID-19. *Stat Methods Med Res*. (2022) 31:1656–74. doi: 10.1177/09622802221106720

13. Meira-Machado L, de Una-Alvarez J, Cadarso-Suarez C, Andersen PK. Multi-state models for the analysis of time-to-event data. *Stat Methods Med Res*. (2009) 18:195–222. doi: 10.1177/0962280208092301

14. McCaw ZR, Tian L, Vassy JL, Ritchie CS, Lee CC, Kim DH, et al. How to quantify and interpret treatment effects in comparative clinical studies of COVID-19. *Ann Intern Med*. (2020) 173:632–7. doi: 10.7326/M20-4044

15. Beigel JH, Tomashek KM, Dodd LE. Remdesivir for the treatment of COVID-19 – preliminary report. *Reply N Engl J Med*. (2020) 383:994. doi: 10.1056/NEJMc2022236

16. Li L, Zhang W, Hu Y, Tong X, Zheng S, Yang J, et al. Effect of convalescent plasma therapy on time to clinical improvement in patients with severe and life-threatening COVID-19: a randomized clinical trial. *JAMA*. (2020) 324:460–70. doi: 10.1001/jama.2020.10044

17. Wolkewitz M, Palomar-Martinez M, Alvarez-Lerma F, Olaechea-Astigarraga P, Schumacher M. Analyzing the impact of duration of ventilation, hospitalization, and ventilation episodes on the risk of pneumonia. *Infect Control Hosp Epidemiol*. (2019) 40:301–6. doi: 10.1017/ice.2018.360

18. Tleyjeh IM, Kashour T, Mandrekar J, Petitti DB. Overlooked shortcomings of observational studies of interventions in coronavirus disease 2019: an illustrated review for the clinician. Open forum. *Infect Dis*. (2021) 8:ofab317. doi: 10.1093/ofid/ofab317

19. Martinuka O, von Cube M, Wolkewitz M. Methodological evaluation of bias in observational coronavirus disease 2019 studies on drug effectiveness. *Clin Microbiol Infect*. (2021) 27:949–57. doi: 10.1016/j.cmi.2021.03.003

20. von Cube M, Schumacher M, Wolkewitz M. Basic parametric analysis for a multi-state model in hospital epidemiology. *BMC Med Res Methodol*. (2017) 17:111. doi: 10.1186/s12874-017-0379-4

21. Wolkewitz M, von Cube M, Schumacher M. Multistate modeling to analyze nosocomial infection data: an introduction and demonstration. *Infect Control Hosp Epidemiol*. (2017) 38:953–9. doi: 10.1017/ice.2017.107

22. Hammerstein S, König C, Dreisörner T, Frey A. Effects of COVID-19-related school closures on student achievement-a systematic review. *Front Psychol*. (2021) 12:746289. doi: 10.3389/fpsyg.2021.746289

23. van Schalkwyk MCI, McKee M. Research into policy: lessons from the COVID-19 pandemic. *Eur J Pub Health*. (2021) 31:iv3–8. doi: 10.1093/eurpub/ckab155

24. Berger E, Winkelmann J, Eckhardt H, Nimptsch U, Panteli D, Reichebner C, et al. A country-level analysis comparing hospital capacity and utilisation during the first COVID-19 wave across Europe. *Health Policy*. (2022) 126:373–81. doi: 10.1016/j.healthpol.2021.11.009

25. Robert-Koch-Institute. SARS-CoV-2: Virologische Basisdaten sowie Virusvarianten im Zeitraum von 2020–2022: Virus Varianten. (2022). https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Virologische_Basisdaten.html.

26. Hu Z, Huang X, Zhang J, Fu S, Ding D, Tao Z. Differences in clinical characteristics between Delta variant and wild-type SARS-CoV-2 infected patients. *Front Med*. (2021) 8:792135. doi: 10.3389/fmed.2021.792135

27. Guo Y, Han J, Zhang Y, He J, Yu W, Zhang X, et al. SARS-CoV-2 omicron variant: epidemiological features, biological characteristics, and clinical significance. *Front Immunol*. (2022) 13:877101. doi: 10.3389/fimmu.2022.877101

frontiers | Frontiers in Medicine

TYPE Review
PUBLISHED 30 October 2024
DOI 10.3389/fmed.2024.1474045

OPEN ACCESS

EDITED BY
Reza Rastmanesh,
American Physical Society, United States

REVIEWED BY
Sankha Shubhra Chakrabarti,
Banaras Hindu University, India
Paolo Eusebi,
University of Perugia, Italy

*CORRESPONDENCE
Suneth Agampodi
✉ suneth.agampodi@ivi.int

RECEIVED 31 July 2024
ACCEPTED 14 October 2024
PUBLISHED 30 October 2024

CITATION
Agampodi S, Tadesse BT, Sahastrabuddhe S, Excler J-L and Kim JH (2024) Biases in COVID-19 vaccine effectiveness studies using cohort design. *Front. Med.* 11:1474045. doi: 10.3389/fmed.2024.1474045

Biases in COVID-19 vaccine effectiveness studies using cohort design

Suneth Agampodi[1,2]*, Birkneh Tilahun Tadesse[3], Sushant Sahastrabuddhe[1], Jean-Louis Excler[4] and Jerome Han Kim[4,5]

[1]Innovation, Initiatives and Enterprise Development Unit, International Vaccine Institute, Seoul, Republic of Korea, [2]Section of Infectious Diseases, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, United States, [3]Epidemiology, Public Health, and Impact Unit, International Vaccine Institute, Seoul, Republic of Korea, [4]International Vaccine Institute, Seoul, Republic of Korea, [5]College of Natural Sciences, Seoul National University, Seoul, Republic of Korea

Observational studies on COVID-19 vaccine effectiveness (VE) have provided critical real-world data, informing public health policy globally. These studies, primarily using pre-existing data sources, have been indispensable in assessing VE across diverse populations and developing sustainable vaccination strategies. Cohort design is frequently employed in VE research. The rapid implementation of vaccination campaigns during the COVID-19 pandemic introduced differential vaccination influenced by sociodemographic disparities, public policies, perceived risks, health-promoting behaviors, and health status, potentially resulting in biases such as healthy user bias, healthy vaccinee effect, frailty bias, differential depletion of susceptibility bias, and confounding by indication. The overwhelming burden on healthcare systems has escalated the risk of data inaccuracies, leading to outcome misclassifications. Additionally, the extensive array of diagnostic tests used during the pandemic has also contributed to misclassification biases. The urgency to publish quickly may have further influenced these biases or led to their oversight, affecting the validity of the findings. These biases in studies vary considerably depending on the setting, data sources, and analytical methods and are likely more pronounced in low- and middle-income country (LMIC) settings due to inadequate data infrastructure. Addressing and mitigating these biases is essential for accurate VE estimates, guiding public health strategies, and sustaining public trust in vaccination programs. Transparent communication about these biases and rigorous improvement in the design of future observational studies are essential.

KEYWORDS

COVID-19, vaccine effectiveness, cohort studies, biases, misclassification bias, healthy user bias, healthy vaccinee effect, differential depletion of susceptibility bias

1 Introduction

The COVID-19 pandemic has triggered numerous vaccine studies, with 1,263 vaccine trials (all phases) registered in the World Health Organization's International Clinical Trials Registry Platform (WHO ICTRP) for more than 250 vaccine candidates (1). These vaccine trials involved more than 80 countries within the first 2 years of the pandemic (99). Due to the urgency of the pandemic, authorities were searching for rapid and reliable data on vaccine efficacy and robust real-world data on vaccine effectiveness (VE) to develop evidence-based vaccination strategies (2).

While randomized, placebo-controlled clinical trials provide robust efficacy and safety data, their value is limited by stringent inclusion criteria, lack of representativeness of diverse populations, and controlled environments that do not reflect local epidemiology

or programmatic challenges. Thus, observational studies became crucial due to their rapid implementation, cost-effectiveness, and flexibility in integrating pre-existing data sources (3). These studies can capture the performance of vaccines across diverse populations, address operational issues, and inform sustainable, contextually appropriate vaccination policies using real-world evidence (4).

The landscape of observational studies on the effectiveness of COVID-19 vaccination shows that more than 2200 observational studies were published directly or indirectly referring to VE by the end of 2022 (5). Despite their importance, observational studies are prone to different biases, some being exemplified by COVID-19 pandemic-related factors (6). Recognizing the importance of these errors in VE studies, the WHO prepared interim guidelines for evaluating COVID-19 VE in 2021 (7) and a revision in 2022 (8). These documents have comprehensively outlined studies, errors, and measures to overcome those possible challenges. Despite these guidelines, the biases in observational studies on VE prevailed. An evaluation of those biases and a closer look into how these predicted biases in VE were observed in COVID-19 offers a unique opportunity to design better vaccine effectiveness studies in future pandemics (9). Test-Negative Design (TND), a case-control approach for estimating VE, involves comparing the odds of vaccination among individuals who test positive or negative for the disease. It was widely used during the early pandemic and extensively discussed (10–16). Hence, this paper explores common biases in cohort studies (Table 1), the most common type of studies used for COVID-19 VE assessments (17), to guide future researchers in planning VE studies.

TABLE 1 Biases in cohort studies investigating COVID-19 vaccine effectiveness (VE).

Bias	Description
Socioeconomic and demographic bias	Unequal access to vaccines due to socioeconomic and demographic disparities can inflate VE estimates, as vaccinated individuals may have better overall health outcomes.
Healthy user bias	This bias occurs when individuals who get vaccinated are also more likely to engage in other health-promoting behaviors, leading to an overestimation of VE.
Healthy vaccinee effect	The tendency for vaccinated individuals to be generally healthier can result in an overestimation of VE, as they are less likely to experience severe disease outcomes.
Frailty bias	Frailty bias occurs when individuals with poorer health or more comorbidities are more likely to get vaccinated, leading to an underestimation of VE.
Confounding by indication	Confounding by indication arises when patients with a higher perceived or known risk of severe disease are more likely to get vaccinated, affecting VE estimates due to differences in baseline health risks.
Differential depletion of susceptibles bias	This bias happens when highly susceptible individuals are disproportionately removed from the population over time, making it seem as if VE is declining.
Attrition bias	Attrition bias occurs when there are systematic differences between individuals who included in the final analysis and those who loss to follow-up, leading to non-representative samples and potentially biased VE estimates.
Immortal time bias	Immortal time bias arises when an "immortal" period during which the outcome cannot occur is incorrectly included in the vaccinated group, falsely enhancing VE.
Misclassification bias	Misclassification bias occurs when individuals or events are incorrectly categorized regarding vaccination status or disease outcome, leading to inaccurate VE estimates.
Non-differential misclassification of outcomes	This bias happens when the misclassification rate is similar across groups, generally biasing results toward the null and diluting the observed effect size.
Differential misclassification of outcomes	Differential misclassification occurs when the likelihood of misclassification differs between groups, leading to either an overestimation or underestimation of VE.
Case counting window bias	Case counting window bias occurs when the time frame for counting cases is misaligned between vaccinated and unvaccinated groups, leading to biased VE estimates.
Waning immunity bias	Waning immunity bias arises when the natural decrease in immune response over time is not accounted for, creating the perception of declining VE in long-term studies.

2 Biases due to differential vaccination in populations

In COVID-19 VE studies employing a cohort design, individuals vaccinated through mass immunization programs are enrolled as cohort participants. If the vaccinated population systematically differs from the unvaccinated population, the estimated VE does not reflect the true VE. Biased estimates of VE occur when the participants selected for the study do not represent the general population, leading to systematic differences in the association observed between exposure and outcome among those selected and those eligible. These disparities lead to a bias often called "selection bias." However, the type of biases in VE studies using cohort design that occurs due to differential vaccinations are not due to conditioning on common effect but resulting from the existence of common causes of exposure and outcome, which could be classified as confounding (18). Despite the classification used (as confounding or bias), a thorough assessment of how these systematic differences occur is crucial to understanding biases in COVID-19 VE studies (Figure 1).

2.1 Coverage-dependent bias in COVID-19 VE estimates

During the initial phase of vaccination campaigns, when coverage was low, the vaccinated group systematically differed from the unvaccinated group. These differences can cause the VE estimate to be underestimated or overestimated, depending on underlying reasons such as country-specific priorities and selective vaccination practices. This bias is sometimes referred to as "early vaccine bias" (8). Similarly, when vaccine coverage is high, unvaccinated individuals systematically differ from the vaccinated or general populations. Personal beliefs, health promotion practices, contraindications, or access issues can all contribute to inaccurate VE estimates among unvaccinated people

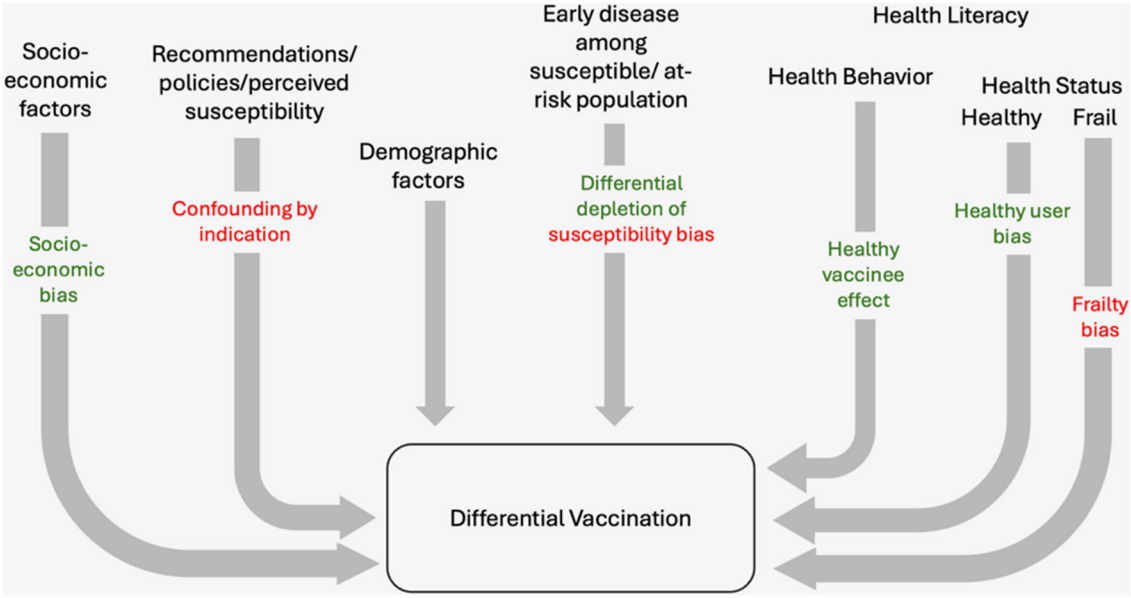

FIGURE 1
Differential COVID-19 vaccination leading to inaccurate estimates of vaccine effectiveness in cohort studies. The colored labels show the specific names given to the biases that resulted in specific differential vaccinations. Green, overestimation of VE; Red, underestimation of VE.

in high-coverage settings. The Weekly COVID-19 Vaccination Dashboard of the Center for Disease Control (CDC) clearly showed these differences since 2021, when coverage was low and high (19). Any VE study conducted during these phases may be biased. This bias likely influenced many early COVID-19 VE studies when vaccination coverage was low, and it continues to affect studies as coverage increases beyond a certain level. Many biases described under "selection" biases could be part of this broader bias.

2.2 Biased selection due to socioeconomic status and demographic factors

Despite the massive production of COVID-19 vaccines, data worldwide show socioeconomic and demographic disparities leading to unequal availability, access, and affordability of vaccination (20). These factors separate those vaccinated from those who are not. The observation of this disparity is universal (21–24) and operates at global as well as national and subnational levels (25). Early in COVID-19 immunization programs, US data reported that Black and Hispanic ethnic groups and those who live in rural areas were less likely to be vaccinated (26), and these categories also have lower access and affordability to general healthcare, resulting in worse health outcomes. Data from more than 35 million people in the UK shows that the lower socioeconomic groups were less likely to get vaccinated (19). These socioeconomic disparities and inequalities systematically differentiate the vaccinated group, which typically has a lower infection risk and better disease outcomes. This usually leads to inflation of VE estimates, which may not apply to the general population.

A similar bias may operate in the opposite direction. Certain demographics, especially the young population, reported higher levels of vaccine hesitancy and declining to get vaccinated in some settings (27, 28). Our World in Data shows that in many countries, the lowest proportion of vaccination is among those between 18 and 25 years of age (29). Even though not vaccinated, this group is generally young, healthy, and less susceptible to severe complications and deaths. Combined data from Europe and the US shows that among the adult population, the lowest case fatality rate is among this age group (30). If this occurs in a specific community where VE studies are conducted, the estimates could be lower than expected due to demographic differences in the vaccinated and unvaccinated groups. In contrast, having elderly and more vulnerable populations in vaccinated cohorts might reduce the VE estimate. Thus, this bias could affect the VE estimate in both directions.

2.3 Healthy user bias

Healthy user bias is a broader concept where individuals who engage in one health-promoting behavior (such as vaccination) are more likely to engage in other health-promoting behaviors. This bias can lead to an overestimation of VE. If vaccinated individuals are also more likely to follow other preventive measures such as mask-wearing and social distancing, the observed benefits may be due to these combined behaviors rather than the vaccine alone. Healthy user bias typically affects prospective cohort designs on VE, where the prospective vaccine campaign participants primarily consist of those with healthy behaviors. One of the early COVID-19 VE studies from Hungary showed that the crude mortality rate was 5.3 per 100,000 person days among the unvaccinated group

compared to 4.0 among the vaccinated group even during the non-endemic period, showing that the two groups may be systematically different and the vaccinated group is generally healthier (31). A retrospective cohort study using 6,974,817 individuals from December to June 2020 shows that recipients of three COVID-19 vaccines had lower non-COVID-19 mortality risk than their comparator groups. For mRNA vaccines, the adjusted hazard ratios of dose 1 and dose 2 ranged from 0.38 to 0.48 (32). Another early analysis from Milwaukee showed that unvaccinated persons had over twice the risk of non-COVID-19 natural death than the vaccinated (33). Many subsequent studies show reduced hospitalizations and deaths due to other causes in the vaccinated group, confirming the presence of healthy user bias (34, 101). In studies where all-cause mortality is considered as the outcome for VE calculation, this could have a more profound effect, where healthy behaviors may affect the risk of deaths due to other causes as well.

2.4 Healthy vaccinee effect

The "healthy vaccinee effect" is a related but distinct concept from "healthy user bias" in observational studies on VE. In these studies, individuals who choose to get vaccinated are generally healthier than those who do not (35). This can lead to overestimating VE since healthier individuals are naturally less likely to experience severe disease outcomes independent of the VE. Part of this effect could be observed in COVID-19 studies where the overall mortality and hospital admission rates are lower during the first few weeks after vaccination (36). The age-standardized all-cause mortality rate from the UK in 2021 shows that within the first 3 weeks following the first dose of vaccination, the death rate was 795.2 per 100,000 people. This rate increased to 1,232.2 per 100,000 people after the initial 3 weeks. A similar pattern was observed with the second dose, where the death rate was 471 per 100,000 people within the first 3 weeks, rising to 850 per 100,000 people after 3 weeks (37). The lower all-cause death rate within the first 3 weeks of vaccination is likely due to the fact that those vaccinated are not acutely ill at the time of vaccination, suggesting the indirect presence of the Healthy Vaccine Effect. Xu et al. reported that age, sex, and race/ethnicity groups adjusted non-COVID-19 related mortality rates among COVID-19 vaccinees were lower than those among comparators for the first three COVID-19 vaccines licensed in the USA. After the first dose, the adjusted hazard ratios (aHRs) were 0.46 for BNT162b2, 0.41 for the mRNA-1273 vaccine, and 0.55 for Ad26.COV2.S (32). Ostropolets and Hripcsak extensively analyzed this effect using a retrospective cohort design based on electronic health records. It showed that even after adjustment for many health-related variables, the vaccinated group had low overall health-seeking and hospital admission within the first few weeks of vaccination (38).

2.5 Frailty bias

The opposite of the healthy vaccine effect is known as the "frailty bias" or "unhealthy vaccine effect." The impact of frailty on the outcome of COVID-19 is well-known, and low vaccine effectiveness among frail vaccine recipients is reported in many studies. For example, Meeraus et al. analyzed over 4.5 million AZD1222 vaccine recipients and noted that VE against hospitalization was >90% in the lowest multimorbidity quartile compared to 80% in the highest quartile. Further, among the elderly who are 'fit,' the VE was 86.2%, whereas it was 72% among the frail. VE against hospitalization was lowest in immunosuppressed individuals (65%) (39). However, frailty could be a selection bias as well. This bias occurs when individuals who are more likely to get vaccinated are those with poorer health or more comorbidities, leading to an underestimation of VE. In this scenario, the vaccinated group appears to have worse health outcomes not due to the vaccine's ineffectiveness but because they were already in poorer health than the unvaccinated group.

2.6 Confounding by indication

Confounding by indication occurs when the reason for vaccination is related to the patient's health status or risk of the outcome being studied. Patients with a higher perceived or known risk of severe COVID-19 might be more likely to get vaccinated, such as older adults or those with comorbidities. An analysis of vaccine policies from 185 countries shows that, except in the Western Pacific region, all other regions prioritized clinically vulnerable and elderly populations in their vaccine rollout (39). These subgroups have a higher baseline risk for unfavorable outcomes, which can confound observed VE because their health outcomes might differ from those not vaccinated regardless of the vaccine's true effectiveness. This can lead to overestimating or underestimating VE due to differences in baseline risk between treated and untreated groups. The bias was well-known to researchers yet common in published literature (40).

While both frailty bias and confounding by indication involve differential selection for vaccination, confounding by indication is driven by the perceived risk of COVID-19 outcomes and general health status, while frailty bias is more about the inherent vulnerability of individuals being vaccinated.

2.7 Differential depletion of susceptible bias

Differential depletion of susceptible (DDS) bias occurs when the most susceptible individuals are disproportionately removed from the at-risk population over time through infection or vaccination. Initially, highly susceptible individuals are more likely to contract COVID-19, making VE appear high. Over time, as susceptible individuals gain immunity, both vaccinated and unvaccinated groups show lower infection rates. This results in an apparent waning of VE, not due to the vaccine losing effectiveness but because the overall population has become less susceptible. In COVID-19 VE studies, this bias can underestimate VE over time, falsely suggesting waning VE. This bias is a form of selection bias because the composition of the population changes, affecting the observed effectiveness. DDS could affect all observational study

designs used for VE studies. During the COVID-19 pandemic, many studies have shown waning effectiveness (41). For example, a study by Chemaitelly et al. from Egypt estimated the long-term effectiveness of COVID-19 mRNA boosters and reported negative relative effectiveness 6 months after boosting, attributing it to "negative immune imprinting" (42). However, Noam argued that in this study, it is possible that the use of discrete-time hazards conditioned on survival for at least 6 months after vaccination results in selection bias due to the depletion of susceptible from the cohort that did not receive a booster (43). Using a simulation model, Kahn et al. demonstrated that if baseline VE is high, the effect of depletion of susceptible bias is low. However, for 'leaky vaccines' (low baseline VE), the impact of this bias is higher (44). It is essential to consider the differential depletion of susceptible in vaccine effectiveness studies to avoid misleading conclusions.

3 Attrition bias

Attrition bias occurs when there are systematic differences between individuals who remain in a study and those who drop out over time (45). For instance, individuals from lower socioeconomic backgrounds, who often have different health behaviors, poorer health outcomes, and limited access to healthcare, are more likely to leave the study or not visit health facilities for diagnosis. This selective dropout can result in a non-representative sample, as these individuals are frequently at a higher risk of infection due to their circumstances. If they drop out disproportionately, the remaining participants may exhibit a lower overall risk of infection, potentially leading to an overestimation of effect size.

Attrition bias in observational VE studies is unique because these studies often involve large cohorts using secondary data. In many countries, COVID-19 vaccination was offered universally, and vaccination records are uniquely maintained in immunization registries, making the data on vaccination status better than other health data (46). However, outcome data for the cohort studies come from different registries, including hospitals, insurance, or similar databases, which can be linked to vaccination registries (47–49, 103). Access, affordability, and coverage of diagnosis and treatment may differ from vaccination, leading to a 'loss to follow-up' from the original cohort. Unlike prospective cohort studies, where investigators are aware of loss to follow-up, data linkage studies have no definitive way to determine whether there was loss to follow-up. Consequently, lost follow-up is often not reported in cohort studies on VE. For example, in the systematic review of Law et al. none of the cohort studies investigating the VE of inactivated vaccines reported the attrition rate (50). Discussing the impact of attrition on VE is specifically important based on the context and policies, health-seeking behaviors, and vulnerable groups.

4 Information biases

In VE studies using cohort design, information biases can significantly distort the VE estimates. These biases arise from misclassification of exposure or outcome, often due to inconsistent or inaccurate data collection methods. Once the cohorts are identified, it is crucial to obtain accurate data; however, measurement errors can lead to bias in the analysis, known as information bias (51). For example, variations in diagnostic testing practices, inaccuracies in vaccination records, or differential recall between vaccinated and unvaccinated individuals can result in either overestimation or underestimation of VE. The direction and magnitude of this bias depend on whether the distribution of errors for a specific variable, such as vaccination status or disease incidence, is influenced by the true values of these variables or by errors in measuring other variables.

4.1 Immortal time bias

Immortal time bias occurs in observational studies when an "immortal" time period is incorrectly classified or excluded from the analysis (52). This period is immortal because, during this time, the outcome of interest (such as infection or death) cannot occur. In COVID-19 vaccine studies, this bias can arise if the time between the initiation of the study and when individuals receive the vaccine is not correctly accounted for. If this period is mistakenly included in the vaccinated group, it can falsely enhance the vaccine's apparent effectiveness because individuals cannot experience the outcome during this "immortal" period. As an example, Flacco et al. (53) reported that the monthly mean death rates were 0.97 per 1,000 for vaccinated individuals and 2.26 per 1,000 for unvaccinated individuals, showing a significant difference in overall deaths. Their data indicates that the mean follow-up time was 561 days for those never vaccinated, compared to 399 days for those vaccinated, due to the time gap between the start of the study and vaccination. The vaccinated group had a mean of 162 days of immortal time before vaccination (only those still alive received the vaccine), while all deaths during this period were allocated to the unvaccinated group. Since this was a single cohort from a province, the vaccinated group contributed to the denominator of the unvaccinated group before receiving their vaccines. Using the same data, Berrino et al. (54) recalculated the mortality rates and concluded that the rates were almost similar when accounting for the immortal time of the vaccinated group. Immortal time bias is not due to differential vaccination but rather an error in the denominator, which may happen in the analysis process or retrospective cohort studies.

4.2 Misclassification bias

Misclassification bias occurs when individuals or events are incorrectly categorized regarding exposure or outcome status. This bias can arise if vaccination status or disease outcome is inaccurately recorded. While the misclassification of exposure status (vaccination) might seem theoretical in prospective cohort studies, most VE studies use retrospective designs. Due to incomplete or inaccurate historical data and errors in data linkage, there is a risk of misclassifying vaccination status. This is particularly relevant if vaccination records are not up-to-date or come from different healthcare providers with varying record-keeping practices. One of the first systematic reviews on the VE of COVID-19 vaccines showed that of the 42 studies published within the first 6 months, 31 used vaccination registries, five included

self-reporting, and 6 did not report the source of vaccination information, highlighting the risk of misclassification bias in these studies (17).

Misclassification of disease or outcomes is expected due to variability in diagnostic test accuracy, differences in test administration timing, and potential misinterpretation of results. The COVID-19 pandemic led to rapid advancement and large-scale production of various diagnostic test kits, none of which possess 100% sensitivity or specificity. This inevitably results in false positives and false negatives, contributing to misclassification. Several systematic reviews on COVID-19 diagnostic tests indicate significant variability in sensitivity and specificity across different platforms, tests, timing of testing relative to exposure, and result interpretation (55–57). A meta-analysis of 24 commercially available antigen kits demonstrated a pooled sensitivity of 77% and a pooled specificity of 98%, highlighting a substantial risk of misclassification (58). Despite WHO's recommendation to use RT-PCR for VE studies (8), the reliance on two-stage screening procedures increases the likelihood of misclassification. These variations can lead to incorrect disease status classification, further complicating VE estimates.

During the pandemic, changes in diagnostic criteria (59–61), referral procedures, breakthrough infections, and even COVID-19 death classifications occurred over time. Since VE studies rely on real-world data, these variations in disease diagnosis and death classifications can differ not only over time but also across different sites, hospitals, or provinces, influenced by policy changes or subjective human factors. Many studies across the globe show that a large number of COVID-19 deaths are unaccounted for and reported as "excess mortality," and there are time and space variations of these numbers (62). Surveillance and reporting biases further complicate the scenario; inconsistent case reporting practices and variability in public health surveillance intensity can lead to uneven detection and reporting of cases. Understanding the changes and events in study areas, regions, and countries over the study period is essential to comprehending misclassification biases.

4.2.1 Non-differential misclassification of outcomes

Non-differential misclassification occurs when the misclassification rate is similar across groups. This can happen if the outcome measured has a low specificity, leading to false positive cases in both groups. It generally biases the results toward the null, diluting the observed effect size and making it harder to detect a true association between vaccination and outcomes (63). Sometimes, it can happen due to the way the "outcome" is recorded. This has been shown in studies where the allocation of outcomes (COVID-19-related hospital admission) uses data from the billing process. These include patients without symptoms admitted for other reasons but tested positive for COVID-19 (38). Because the outcome documentation is from billing data, those without clinical disease but positive test results were included as those with outcomes (COVID-19 hospital admissions). The misclassification is similar for both groups; thus, the VE estimates are diluted. Similarly, a study from the Netherlands showed that 42% of cases included in the hospital register on COVID-19 patients are missing the reason for admission, and whether the positive COVID-19 test results were associated with COVID-19 clinical disease is not known (103).

4.2.2 Differential misclassification of outcome

Differential misclassification occurs when the likelihood of misclassification differs between groups (vaccinated vs. unvaccinated). Depending on the direction of the misclassification, it can lead to overestimation or underestimation of VE. In studies using secondary data for VE studies, the allocated diagnosis could be systematically different in the two groups. Some studies show that vaccinated people usually attribute mild to moderate symptoms to vaccine side effects and do not seek care, thus less likely to be diagnosed as having COVID-19 (38). On the other hand, healthcare workers may suspect COVID-19 more among unvaccinated groups during hospital visits and perform testing, leading to higher detection or "diagnostic bias." This could be partly due to policies where test results were mandated for many institutions and procedures if the individual is not vaccinated, thus leading to more cases of asymptomatic or mild cases of COVID-19 among unvaccinated groups. This will lead to differential misclassification of outcome status among vaccinated and unvaccinated groups, inflating VE estimates. However, it will not always overestimate VE. A study in Australia reported that fully vaccinated participants were twice as likely as those who were unvaccinated to report positive COVID testing intentions, which may lead to overdiagnosis of COVID-19 in the vaccinated arm, underestimating VE (64). The context-specific factors could play a major role in deciding in which direction the bias operates.

4.3 Case counting window bias

The case counting window bias occurs when the time frame for counting cases in a study is not properly aligned with the period during which a vaccine is expected to be effective. This can happen if the cases are counted from the start of follow-up for the unvaccinated group but only after vaccination for the vaccinated group; the difference in these windows can lead to biased VE estimates. So, the cases in the unvaccinated group are counted during a period when they were at higher risk. In contrast, cases in the vaccinated group are only counted after they might have already benefited from some protection. Fung et al. showed that even a vaccine with almost zero VE could be presented as having a VE of 48% if the case counting window bias is ignored (65).

4.4 Waning immunity bias

Waning immunity bias occurs when the observed effectiveness of a vaccine appears to decline over time due to a natural decrease in immune response, notably the reduction in neutralizing antibody levels. Initially, vaccines demonstrate high effectiveness due to a robust immune response. However, as antibody levels wane, individuals may become more susceptible to infection, creating the perception of declining vaccine effectiveness (VE). Numerous

studies, including several systematic reviews (45, 66, 67), have reported this phenomenon, particularly those involving mRNA vaccines against COVID-19. A systematic review of 40 studies estimated that VE against omicron infection and symptomatic disease decreased by 20% at 6 months post-primary vaccination cycle and by 30% at 9 months post-booster dose administration (68). This waning immunity is a natural biological process; failing to account for it in VE studies properly can introduce bias, leading to inaccurate interpretations. Long-term VE studies must incorporate original or short-term VE data and clearly state the duration post-vaccination to avoid biased interpretation of VE estimates. Additionally, considering the date of vaccination in analysis and data presentation is essential. Addressing this bias is crucial, particularly in the ongoing "infodemic" (69) and vaccine hesitancy.

5 Confounding

Confounding occurs when an extraneous factor influences both the likelihood of vaccination and the health outcome, leading to a distorted estimate of vaccine effectiveness (18). This manuscript has described factors such as health-seeking behaviors, demographic differences, and underlying health conditions leading to differential vaccination. While these factors were labeled as biases—such as selection bias, healthy user bias, and frailty bias—these are, in fact, confounding factors. They influence both vaccination status and outcomes, distorting the true effect of the vaccine. For instance, individuals with prior exposure to SARS-CoV-2 or underlying comorbidities may have different risks of severe COVID-19, which could distort VE estimates if not properly adjusted for (70, 71, 100). In addition to these specific factors, researchers must consider other confounders that may further affect VE estimates, such as geographic differences, occupational exposures, and previous infection history. Failure to control for such confounders can lead to inflated or underestimated VE estimates, particularly when assessing protection against severe disease. To ensure accurate and reliable VE estimates, thorough identification and adjustment of these confounding factors are essential in observational studies.

6 Biases and challenges of VE studies in LMICs

The assessment of VE in low and lower-middle-income countries (LMICs) faces significant challenges due to various biases and unique factors. Disparities in vaccine deployment, reliance on less effective vaccines, supply constraints, and dosing challenges necessitate evaluations of mixed and suboptimal regimens. Additionally, distinct health and demographic profiles can lead to differential vaccination coverage, altering estimated vaccine protection (72). High seroprevalence of naturally acquired immunity and the risk of new variant emergence due to uncontrolled transmission further complicate VE estimates. The absence of robust systems for maintaining vaccination records, integrating data across platforms, and ensuring accurate digitalization, along with limited diagnostic capacity and barriers to accessing and affording healthcare, exacerbate the risk of misclassification of both exposure (vaccination status) and outcome (disease occurrence). Such misclassification is a critical source of information bias in these studies, significantly impacting the accuracy of VE estimates. Therefore, targeted VE studies are essential to develop accurate and effective vaccination strategies tailored to LMICs. However, the availability of vaccine studies in these settings is limited. A systematic review by Petráš et al. included 761 published VE studies, but only two were from low-income countries, both from Zambia (73). One was from a prison outbreak using a case-control design with self-reporting/rapid tests (74), and the other was a hospital-based study with significant missing data (75). Among lower-middle-income countries, Indian authors published 28 studies. Except for India, only Bangladesh (76), Egypt (77), Morocco (78), and Pakistan (79–83) had estimated VE in their settings. Real-world evidence and the effect of biases on those estimates from LMICs are missing in global literature. A comprehensive analysis of these available studies is critical to understanding how specific biases affect VE estimates in future academic preparations.

7 Effects of biased estimates

Biases significantly impact the internal validity of vaccine efficacy (VE) studies, with far-reaching societal and policy implications. Early studies often reported VE as high as 95% (84), fostering overconfidence in vaccine protection and prompting global policy changes. Health agencies described vaccines as "extremely protective," creating a misconception that vaccines could prevent infection (85). This led to policy changes in some places prioritizing vaccination for transmission interruption, sometimes neglecting vulnerable populations at higher risk for severe COVID-19 outcomes. Overestimating VE altered public perception, with increased infection rates post-vaccination due to risk behaviors (86). Variations in VE estimates influence vaccination willingness (87); one study showed that 51.3% would accept a COVID-19 vaccine that is 50% effective, and 77.1% would accept a vaccine that is 95% effective (84). Differences in VE estimates can significantly fuel vaccine hesitancy and erode public trust in vaccination programs. Media coverage and anti-vaccine movements can exploit these inconsistencies, spreading misinformation and increasing hesitancy. Transparent communication about the limitations and strengths of VE studies is crucial to maintaining public trust and encouraging vaccine acceptance.

8 Minimizing biases in cohort studies on VE

Target trial emulation (TTE) is often used to minimize biases in observational studies on VE by applying design principles from randomized trials to the analysis of observational data, thereby explicitly tying the analysis to the trial it is emulating. The TTE approach was extensively used in COVID-19 VE studies when applying cohort design (14, 88–94, 102). Still, the challenges and errors in properly executing this approach can introduce the abovementioned biases, resulting in inaccuracies in estimating VE

if the emulation is improperly planned and executed. Although referred to as "TTE bias" in literature, it is not technically a specific type of bias but could be due to any other type detailed above.

Biases, once introduced, are challenging to control in observational studies. Minimizing bias during the design stage is paramount. The WHO guidelines on observational studies for COVID-19 vaccine effectiveness (8) outlined various strategies to mitigate these biases. Bayesian modeling approaches could address specific biases, such as misclassification bias due to imperfect tests (95). A thorough evaluation of all potential biases is essential when reporting VE studies but reporting and assessing the direction of biases may be more challenging than it seems. Brookmeyer and Morrison (96) demonstrated this complexity using data published in a VE study (97) to simulate different biases occur in linked registry studies. They showed that if the bias is due to a single source, the direction of the bias is predictable. However, if multiple sources of biases are present, then the direction of the bias can be either way. Often, the biases are multiple; thus, predicting direction can be difficult. It is crucial to discuss these biases' probable impact and direction on VE study results and report them comprehensively in all observational studies.

9 Conclusion

These well-known epidemiological biases may occur more frequently during a pandemic. Rapid vaccine development and distribution can lead to differential vaccination practices, prioritizing high-risk populations and exacerbating frailty bias and confounding by indication. Socioeconomic disparities and diverse health behaviors become more pronounced, while the overburdened health systems increase the risk of data inaccuracies and outcome misclassification. The frequency of occurrence of these biases varies widely based on the setting, data sources, and analysis. These biases could be higher in VE in LMIC settings due to a lack of proper data sources. Understanding and mitigating these biases are crucial for accurate VE estimates, informing public health strategies, and maintaining public trust in vaccination programs. With the peak pandemic now behind us and less urgency for rapid publication, a comprehensive investigation into the long-term vaccine effectiveness is warranted, ensuring that all previously discussed biases are thoroughly addressed. Additionally, future studies must explicitly account for whether individuals who were infected with COVID-19 subsequently received vaccination, as this may affect long-term outcomes such as mortality and PASC. While studies like Cai et al. (98) demonstrate that the risk of death declines over time but remains elevated in previously hospitalized patients, they do not account for whether subsequent vaccination modifies this risk. Incorporating this factor in future research will provide more accurate assessments of vaccine effectiveness and long-term health risks, thereby guiding public health policies more effectively.

Author contributions

SA: Conceptualization, Data curation, Validation, Visualization, Writing – original draft, Writing – review & editing. BT: Conceptualization, Writing – review & editing. SS: Supervision, Writing – review & editing. J-LE: Conceptualization, Supervision, Writing – review & editing. JK: Conceptualization, Supervision, 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 would like to thank John D. Clemens for his support during the conceptualization of this manuscript. Chat GTP 40 was used for language improvement.

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

References

1. World Health Organization. *International Clinical Trials Registry Platform*. (2024). Available at: https://trialsearch.who.int/Default.aspx (accessed March 8,2024).

2. Excler J-L, Saville M, Berkley S, Kim JH. Vaccine development for emerging infectious diseases. *Nat Med.* (2021) 27:591–600. doi: 10.1038/s41591-021-01301-0

3. Sheldrick RC. Randomized trials vs. real-world evidence: how can both inform decision-making? *J Am Med Assoc.* (2023) 329:1352–3. doi: 10.1001/jama.2023.4855

4. Bollaerts K, Wyndham-Thomas C, Miller E, Izurieta HS, Black S, Andrews N, et al. The role of real-world evidence for regulatory and public health decision-making for Accelerated Vaccine Deployment- a meeting report. *Biologicals.* (2024) 85:101750. doi: 10.1016/j.biologicals.2024.101750

5. World Health Organization. *Landscape of Observational Studies on the Effectiveness of COVID-19 Vaccination*. (2023). Available at: https://www.who.int/publications/m/item/draft-landscape-of-observational-study-designs-on-the-effectiveness-of-covid-19-vaccination (accessed March 8,2024).

6. Lipsitch M, Jha A, Simonsen L. Observational studies and the difficult quest for causality: lessons from vaccine effectiveness and impact studies. *Int J Epidemiol.* (2016) 45:2060–74. doi: 10.1093/ije/dyw124

7. Patel MK, Bergeri I, Bresee JS, Cowling BJ, Crowcroft NS, Fahmy K, et al. Evaluation of post-introduction COVID-19 vaccine effectiveness: summary of interim guidance of the World Health Organization. *Vaccine.* (2021) 39:4013–24. doi: 10.1016/j.vaccine.2021.05.099

8. World Health Organization. *Evaluation of COVID-19 Vaccine Effectiveness in a Changing Landscape of COVID-19 Epidemiology and Vaccination: Interim Guidance, 1 October 2022: Second Addendum to Evaluation of COVID-19 Vaccine Effectiveness: Interim Guidance.* Geneva: World Health Organization (2022).

9. Ricotta EE, Rid A, Cohen IG, Evans NG. Observational studies must be reformed before the next pandemic. *Nat Med.* (2023) 29:1903–5. doi: 10.1038/s41591-023-02375-8

10. Hitchings MDT, Lewnard JA, Dean NE, Ko AI, Ranzani OT, Andrews JR, et al. Use of recently vaccinated individuals to detect bias in test-negative case-control studies of COVID-19 vaccine effectiveness. *Epidemiology.* (2022) 33:450–6. doi: 10.1097/EDE.0000000000001484

11. Schnitzer ME. Estimands and estimation of COVID-19 vaccine effectiveness under the test-negative design: connections to causal inference. *Epidemiology.* (2022) 33:325–33. doi: 10.1097/EDE.0000000000001470

12. Graham S, Tessier E, Stowe J, Bernal JL, Parker EPK, Nitsch D, et al. Bias assessment of a test-negative design study of COVID-19 vaccine effectiveness used in national policymaking. *Nat Commun.* (2023) 14:3984. doi: 10.1038/s41467-023-39674-0

13. Shi X, Li KQ, Mukherjee B. Current challenges with the use of test-negative designs for modeling COVID-19 vaccination and outcomes. *Am J Epidemiol.* (2023) 192:328–33. doi: 10.1093/aje/kwac203

14. Li G, Gerlovin H, Figueroa Muñiz MJ, Wise JK, Madenci AL, Robins JM, et al. Comparison of the test-negative design and cohort design with explicit target trial emulation for evaluating COVID-19 vaccine effectiveness. *Epidemiology.* (2024) 35:137–49. doi: 10.1097/EDE.0000000000001709

15. Mésidor M, Liu Y, Talbot D, Skowronski DM, De Serres G, Merckx J, et al. Test negative design for vaccine effectiveness estimation in the context of the COVID-19 pandemic: a systematic methodology review. *Vaccine.* (2024) 42:995–1003. doi: 10.1016/j.vaccine.2023.12.013

16. Ortiz-Brizuela E, Carabali M, Jiang C, Merckx J, Talbot D, Schnitzer ME. Potential biases in test-negative design studies of COVID-19 vaccine effectiveness arising from the inclusion of asymptomatic individuals. *Am J Epidemiol.* (2024) 2024:kwae288. doi: 10.1093/aje/kwae288

17. Teerawattananon Y, Anothaisintawee T, Pheerapanyawaranun C, Botwright S, Akksilp K, Sirichumroonwit N, et al. A systematic review of methodological approaches for evaluating real-world effectiveness of COVID-19 vaccines: advising resource-constrained settings. *PLoS ONE.* (2022) 17:e0261930. doi: 10.1371/journal.pone.0261930

18. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology.* (2004) 15:615–25. doi: 10.1097/01.ede.0000135174.63482.43

19. Center for Disease Control. *Weekly COVID-19 Vaccination Dashboard* (2021). Available at: https://www.cdc.gov/covidvaxview/weekly-dashboard/index.html

20. Privor-Dumm L, Excler J-L, Gilbert S, Abdool Karim SS, Hotez PJ, Thompson D, et al. Vaccine access, equity and justice: COVID-19 vaccines and vaccination. *Br Med J Glob Health.* (2023) 8:e011881. doi: 10.1136/bmjgh-2023-011881

21. Benoni R, Sartorello A, Moretti F, Marchiori F, Accordini L, Postiglione C, et al. Disparities in access to COVID-19 vaccine in Verona, Italy: a cohort study using local health immunization data. *Front Public Health.* (2023) 11:1167414. doi: 10.3389/fpubh.2023.1167414

22. Cavillot L, Van Loenhout JaF, Devleesschauwer B, Wyndham-Thomas C, Van Oyen H, et al. Sociodemographic and socioeconomic disparities in COVID-19 vaccine uptake in Belgium: a nationwide record linkage study. *J Epidemiol Community Health.* (2023) 78:176–83. doi: 10.1136/jech-2023-220751

23. Sabatini S, Kaufmann M, Fadda M, Tancredi S, Noor N, Van Der Linden BWA, et al. Factors associated with COVID-19 non-vaccination in Switzerland: a nationwide study. *Int J Public Health.* (2023) 68:1605852. doi: 10.3389/ijph.2023.1605852

24. Welsh J, Biddle N, Butler DC, Korda RJ. Discretion in decision to receive COVID-19 vaccines and associated socio-economic inequalities in rates of uptake: a whole-of-population data linkage study from Australia. *Publ Health.* (2023) 224:82–9. doi: 10.1016/j.puhe.2023.08.020

25. Bayati M, Noroozi R, Ghanbari-Jahromi M, Jalali FS. Inequality in the distribution of COVID-19 vaccine: a systematic review. *Int J Equity Health.* (2022) 21:122. doi: 10.1186/s12939-022-01729-x

26. Murthy BP. COVID-19 vaccination coverage and demographic characteristics of infants and children aged 6 months-4 years-United States, June 20-December 31, 2022. *Morbid Mortal Week Rep.* (2023) 72:7a4. doi: 10.15585/mmwr.mm7207a4

27. Fazel M, Puntis S, White SR, Townsend A, Mansfield KL, Viner R, et al. Willingness of children and adolescents to have a COVID-19 vaccination: results of a large whole schools survey in England. *eClinicalMedicine.* (2021) 40:101144. doi: 10.1016/j.eclinm.2021.101144

28. Bergen N, Johns NE, Chang Blanc D, Hosseinpoor AR. Within-country inequality in COVID-19 vaccination coverage: a scoping review of academic literature. *Vaccines.* (2023) 11:517. doi: 10.3390/vaccines11030517

29. Havers FP, Pham H, Taylor CA, Whitaker M, Patel K, Anglin O, et al. COVID-19-associated hospitalizations among vaccinated and unvaccinated adults 18 years or older in 13 US States, January 2021 to April 2022. *J Am Med Assoc Intern Med.* (2022) 182:1071–81. doi: 10.1001/jamainternmed.2022.4299

30. Goldstein JR, Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc Nat Acad Sci USA.* (2020) 117:22035–41. doi: 10.1073/pnas.2006392117

31. Pálinkás A, Sándor J. Effectiveness of COVID-19 vaccination in preventing all-cause mortality among adults during the third wave of the epidemic in Hungary: nationwide retrospective cohort study. *Vaccines.* (2022) 10:1009. doi: 10.3390/vaccines10071009

32. Xu S, Huang R, Sy LS, Hong V, Glenn SC, Ryan DS, et al. A safety study evaluating non-COVID-19 mortality risk following COVID-19 vaccination. *Vaccine.* (2023) 41:844–54. doi: 10.1016/j.vaccine.2022.12.036

33. Atanasov VA, Barreto Parra PN, Franchi L. Selection bias and COVID-19 vaccine effectiveness against death: evidence from linked mortality and vaccination records. *SSRN Electron J.* (2022) 2022:460. doi: 10.2139/ssrn.4250460

34. Arbel R, Peretz A, Sergienko R, Friger M, Beckenstein T, Duskin-Bitan H, et al. Effectiveness of a bivalent mRNA vaccine booster dose to prevent severe COVID-19 outcomes: a retrospective cohort study. *Lancet Infect Dis.* (2023) 23:914–21. doi: 10.1016/S1473-3099(23)00122-6

35. Furst T, Straka R, Janosek J. Healthy vaccinee effect: a bias not to be forgotten in observational studies on COVID-19 vaccine effectiveness. *Pol Arch Intern Med.* (2024) 134:16634. doi: 10.20452/pamw.16634

36. Prunas O, Weinberger DM, Pitzer VE, Gazit S, Patalon T. Waning effectiveness of the BNT162b2 vaccine against infection in adolescents in Israel. *Clin Infect Dis.* (2022) 76:113–8. doi: 10.1093/cid/ciac315

37. Office for National Statistics. *Deaths Involving COVID-19 By Vaccination Status, England: Deaths Occurring Between 1 January and 31 October 2021.* O'ffice for National Statistics (2021). Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19byvaccinationstatusengland/deathsoccurringbetween1januaryand31october2021 (accessed March 8,2024).

38. Ostropolets A, Hripcsak G. COVID-19 vaccination effectiveness rates by week and sources of bias: a retrospective cohort study. *Br Med J Open.* (2022) 12:e061126. doi: 10.1136/bmjopen-2022-061126

39. Meeraus W, Joy M, Ouwens M, Taylor KS, Venkatesan S, Dennis J, et al. AZD1222 effectiveness against severe COVID-19 in individuals with comorbidity or frailty: the RAVEN cohort study. *J Infect.* (2024) 88:106129. doi: 10.1016/j.jinf.2024.106129

40. Nguyen VG, Yavlinsky A, Beale S, Hoskins S, Byrne TE, Lampos V, et al. Comparative effectiveness of different primary vaccination courses on mRNA-based booster vaccines against SARS-CoV-2 infections: a time-varying cohort analysis using trial emulation in the Virus Watch community cohort. *Int J Epidemiol.* (2023) 52:342–54. doi: 10.1093/ije/dyad002

41. Bobrovitz N, Ware H, Ma X, Li Z, Hosseini R, Cao C, et al. Protective effectiveness of previous SARS-CoV-2 infection and hybrid immunity against the omicron variant and severe disease: a systematic review and meta-regression. *Lancet Infect Dis.* (2023) 23:556–67. doi: 10.1016/S1473-3099(22)00801-5

42. Chemaitelly H, Ayoub HH, Tang P, Coyle P, Yassine HM, Al Thani AA, et al. Long-term COVID-19 booster effectiveness by infection history and clinical vulnerability and immune imprinting: a retrospective population-based cohort study. *Lancet Infect Dis.* (2023) 23:816–27. doi: 10.1016/S1473-3099(23)00058-0

43. Barda N. The ups and downs of observational vaccine research. *Lancet Infect Dis.* (2023) 23:767–8. doi: 10.1016/S1473-3099(23)00119-6

44. Kahn R, Schrag SJ, Verani JR, Lipsitch M. Identifying and alleviating bias due to differential depletion of susceptible people in postmarketing evaluations of COVID-19 vaccines. *Am J Epidemiol.* (2022) 191:800–11. doi: 10.1093/aje/kwac015

45. Booker CL, Harding S, Benzeval M. A systematic review of the effect of retention methods in population-based cohort studies. *BMC Publ Health.* (2011) 11:249. doi: 10.1186/1471-2458-11-249

46. Cameron-Blake E, Tatlow H, Andretti B, Boby T, Green K, Hale T, et al. A panel dataset of COVID-19 vaccination policies in 185 countries. *Nat Hum Behav.* (2023) 7:1402–13. doi: 10.1038/s41562-023-01615-8

47. Meijerink H, Veneti L, Kristoffersen AB, Danielsen AS, Stecher M, Starrfelt J. Estimating vaccine effectiveness against COVID-19 using cause-specific sick leave as an indicator: a nationwide population-based cohort study, Norway, July 2021—December 2022. *BMC Publ Health.* (2024) 24:1861. doi: 10.1186/s12889-024-19374-0

48. Uemura K, Ono S, Michihata N, Yamana H, Yasunaga H. Duration of effectiveness of the COVID-19 vaccine in Japan: a retrospective cohort

study using large-scale population-based registry data. *BMC Infect Dis.* (2024) 24:648. doi: 10.1186/s12879-024-09488-6

49. Urquidi C, Sepúlveda-Peñaloza A, Valenzuela MT, Ponce A, Menares V, Cortes CP, et al. Vaccine effectiveness in reducing COVID-19-related hospitalization after a risk-age-based mass vaccination program in a Chilean municipality: a comparison of observational study designs. *Vaccine.* (2024) 42:3851–6. doi: 10.1016/j.vaccine.2024.05.002

50. Law M, Ho SSH, Tsang GKC, Ho CMY, Kwan CM, Yan VKC, et al. Efficacy and effectiveness of inactivated vaccines against symptomatic COVID-19, severe COVID-19, and COVID-19 clinical outcomes in the general population: a systematic review and meta-analysis. *Lancet Region Health Western Pacific.* (2023) 37:100788. doi: 10.1016/j.lanwpc.2023.100788

51. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins Philadelphia (2008).

52. Alessandria M, Malatesta GM, Berrino F, Donzelli A. A critical analysis of all-cause deaths during COVID-19 vaccination in an Italian Province. *Microorganisms.* (2024) 12:1343. doi: 10.3390/microorganisms12071343

53. Flacco ME, Acuti Martellucci C, Soldato G, Di Martino G, Carota R, De Benedictis M, et al. COVID-19 vaccination did not increase the risk of potentially related serious adverse events: 18-month cohort study in an Italian Province. *Vaccines.* (2022) 11:10031. doi: 10.3390/vaccines11010031

54. Berrino F, Donzelli A, Bellavite P, Malatesta G. COVID-19 vaccination and all-cause and non-COVID-19 mortality. A revaluation of a study carried out in an Italian Province. *Epidemiol Prev.* (2023) 47:374–8. doi: 10.19191/EP23.6.A643.075

55. Borges LP, Martins AF, Silva BM, Dias BP, Gonçalves RL, Souza DRV, et al. Rapid diagnosis of COVID-19 in the first year of the pandemic: a systematic review. *Int Immunopharmacol.* (2021) 101:108144. doi: 10.1016/j.intimp.2021.108144

56. Dinnes J, Sharma P, Berhane S, Van Wyk SS, Nyaaba N, Domen J, et al. Rapid, point-of-care antigen tests for diagnosis of SARS-CoV-2 infection. *Cochr Database Syst Rev.* (2022) 7:CD013705. doi: 10.1002/14651858.CD013705.pub3

57. Fox T, Geppert J, Dinnes J, Scandrett K, Bigio J, Sulis G, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochr Database Syst Rev.* (2022) 11:CD013652. doi: 10.1002/14651858.CD013652.pub2

58. Manten K, Katzenschlager S, Brümmer LE, Schmitz S, Gaeddert M, Erdmann C, et al. Clinical accuracy of instrument-based SARS-CoV-2 antigen diagnostic tests: a systematic review and meta-analysis. *Virol J.* (2024) 21:99. doi: 10.1186/s12985-024-02371-5

59. Wang YY, Jin YH, Ren XQ, Li YR, Zhang XC, Zeng XT, et al. Updating the diagnostic criteria of COVID-19 "suspected case" and "confirmed case" is necessary. *Mil Med Res.* (2020) 7:17. doi: 10.1186/s40779-020-00245-9

60. Hong KH, Kim GJ, Roh KH, Sung H, Lee J, Kim SY, et al. Update of guidelines for laboratory diagnosis of COVID-19 in Korea. *Ann Lab Med.* (2022) 42:391–7. doi: 10.3343/alm.2022.42.4.391

61. Wu Y, Feng X, Gong M, Han J, Jiao Y, Li S, et al. Evolution and major changes of the diagnosis and treatment protocol for COVID-19 patients in China 2020-2023. *Health Care Sci.* (2023) 2:135–52. doi: 10.1002/hcs2.45

62. Giattino C, Ritchie H, Roser M, Ortiz-Ospina E, Hasell J. *Excess Mortality During the Coronavirus Pandemic (COVID-19)*. Our World in Data (2021). Available at: https://ourworldindata.org/excess-mortality-covid

63. Hansen CH. Bias in vaccine effectiveness studies of clinically severe outcomes that are measured with low specificity: the example of COVID-19-related hospitalisation. *Euro Surveill.* (2024) 29:259. doi: 10.2807/1560-7917.ES.2024.29.7.2300259

64. Glasziou P, Mccaffery K, Cvejic E, Batcup C, Ayre J, Pickles K, et al. Testing behaviour may bias observational studies of vaccine effectiveness. *J Assoc Med Microbiol Infect Dis Can.* (2022) 7:242–6. doi: 10.3138/jammi-2022-0002

65. Fung K, Jones M, Doshi P. Sources of bias in observational studies of COVID-19 vaccine effectiveness. *J Eval Clin Pract.* (2024) 30:30–6. doi: 10.1111/jep.13839

66. Feikin DR, Higdon MM, Abu-Raddad LJ, Andrews N, Araos R, Goldberg Y, et al. Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. *Lancet.* (2022) 399:924–44. doi: 10.1016/S0140-6736(22)00152-0

67. Higdon MM, Baidya A, Walter KK, Patel MK, Issa H, Espié E, et al. Duration of effectiveness of vaccination against COVID-19 caused by the omicron variant. *Lancet Infect Dis.* (2022) 22:1114–6. doi: 10.1016/S1473-3099(22)00409-1

68. Menegale F, Manica M, Zardini A, Guzzetta G, Marziano V, D'andrea V, et al. Evaluation of waning of SARS-CoV-2 vaccine-induced immunity: a systematic review and meta-analysis. *J Am Med Assoc Netw Open.* (2023) 6:e2310650. doi: 10.1001/jamanetworkopen.2023.10650

69. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M. Assessing the risks of "infodemics" in response to COVID-19 epidemics. *Nat Hum Behav.* (2020) 4:1285–93. doi: 10.1038/s41562-020-00994-6

70. Rennert L, Ma Z, Mcmahan CS, Dean D. Effectiveness and protection duration of COVID-19 vaccines and previous infection against any SARS-CoV-2 infection in young adults. *Nat Commun.* (2022) 13:3946. doi: 10.1038/s41467-022-31469-z

71. Altarawneh HN, Chemaitelly H, Ayoub HH, Tang P, Hasan MR, Yassine HM, et al. Effects of previous infection, vaccination, and hybrid immunity against symptomatic Alpha, Beta, and Delta SARS-CoV-2 infections: an observational study. *eBioMedicine.* (2023) 95:104734. doi: 10.1016/j.ebiom.2023.104734

72. Clemens J, Aziz AB, Tadesse BT, Kang S, Marks F, Kim J. Evaluation of protection by COVID-19 vaccines after deployment in low and lower-middle income countries. *eClinicalMedicine.* (2022) 43:101253. doi: 10.1016/j.eclinm.2021.101253

73. Petráš M, Janovská D, Lomozová D, Franklová M, Dlouhý P, Rosina J, et al. Understanding the time-driven shifts of vaccine effectiveness against any and severe COVID-19 before and after the surge of Omicron variants within 2.5 years of vaccination: a meta-regression. *Int J Infect Dis.* (2024) 142:106986. doi: 10.1016/j.ijid.2024.106986

74. Simwanza J, Hines JZ, Sinyange D, Sinyange N, Mulenga C, Hanyinza S, et al. COVID-19 vaccine effectiveness during a prison outbreak when omicron was the dominant circulating variant-Zambia, December 2021. *Am J Trop Med Hyg.* (2022) 107:1055–9. doi: 10.4269/ajtmh.22-0368

75. Chanda D, Hines JZ, Itoh M, Fwoloshi S, Minchella PA, Zyambo KD, et al. COVID-19 vaccine effectiveness against progression to in-hospital mortality in Zambia, 2021–2022. *Open For Infect Dis.* (2022) 9:ofac469. doi: 10.1093/ofid/ofac469

76. Khanam F, Islam MT, Ahmmed F, Ahmed SU, Hossen MI, Rajib MH, et al. Measuring the effectiveness of COVID-19 vaccines used during a surge of the delta variant of SARS-CoV-2 in Bangladesh: a test-negative design evaluation. *Vaccines.* (2022) 10:122069. doi: 10.3390/vaccines10122069

77. Ashmawy R, Kamal E, Amin W, Sharaf S, Kabeel S, Albiheyri R, et al. Effectiveness and safety of inactivated SARS-CoV-2 vaccine (BBIBP-CorV) among healthcare workers: a seven-month follow-up study at fifteen central hospitals. *Vaccines.* (2023) 11:892. doi: 10.3390/vaccines11050892

78. Belayachi J, Obtel M, Mhayi A, Razine R, Abouqal R. Long term effectiveness of inactivated vaccine BBIBP-CorV (Vero Cells) against COVID-19 associated severe and critical hospitalization in Morocco. *PLoS ONE.* (2022) 17:e0278546. doi: 10.1371/journal.pone.0278546

79. Ul Munamm SA, Nadeem I, Mahdi N, Saqlain M, Rana ZK, Khatana UF, et al. Comparative analysis of mRNA and inactivated COVID-19 vaccines: a study from Faisalabad district of Pakistan. *J R Coll Physicians Edinb.* (2022) 52:240–6. doi: 10.1177/14782715221131409

80. Khan UI, Niaz M, Azam I, Hasan Z, Hassan I, Mahmood SF, et al. Effectiveness of inactivated COVID-19 vaccines against SARS-CoV-2 infections among healthcare personnel in Pakistan: a test-negative case-control study. *Br Med J Open.* (2023) 13:e071789. doi: 10.1136/bmjopen-2023-071789

81. Mushtaq MZ, Nasir N, Mahmood SF, Khan S, Kanji A, Nasir A, et al. Exploring the relationship between SARS-CoV-2 variants, illness severity at presentation, in-hospital mortality and COVID-19 vaccination in a low middle-income country: a retrospective cross-sectional study. *Health Sci Rep.* (2023) 6:e1703. doi: 10.1002/hsr2.1703

82. Nadeem I, Ul Munamm SA, Ur Rasool M, Fatimah M, Abu Bakar M, Rana ZK, et al. Safety and efficacy of Sinopharm vaccine (BBIBP-CorV) in elderly population of Faisalabad district of Pakistan. *Postgrad Med J.* (2023) 99:463–9. doi: 10.1136/postgradmedj-2022-141649

83. Nisar MI, Ansari N, Malik AA, Shahid S, Lalani KRA, Chandna MA, et al. Assessing the effectiveness of COVID-19 vaccines in Pakistan: a test-negative case-control study. *J Infect.* (2023) 86:e144–7. doi: 10.1016/j.jinf.2023.01.016

84. Wagner AL, Sheinfeld Gorin S, Boulton ML, Glover BA, Morenoff JD. Effect of vaccine effectiveness and safety on COVID-19 vaccine acceptance in Detroit, Michigan, July 2020. *Hum Vaccin Immunother.* (2021) 17:2940–5. doi: 10.1080/21645515.2021.1917233

85. World Health Organization. Episode #49—can I get infected after vaccination? In: World Health Organization, editors. *Science in 5*. Geneva: World Health Organization (2021).

86. Day M. COVID-19: stronger warnings are needed to curb socialising after vaccination, say doctors and behavioural scientists. *Br Med J.* (2021) 372:n783. doi: 10.1136/bmj.n783

87. Wagner AL, Huang Z, Ren J, Laffoon M, Ji M, Pinckney LC, et al. Vaccine hesitancy and concerns about vaccine safety and effectiveness in Shanghai, China. *Am J Prev Med.* (2021) 60:S77–86. doi: 10.1016/j.amepre.2020.09.003

88. Ioannou GN, Bohnert ASB, O'hare AM, Boyko EJ, Maciejewski ML, Smith VA, et al. Effectiveness of mRNA COVID-19 vaccine boosters against infection, hospitalization, and death: a target trial emulation in the omicron (B.1.1.529) variant era. *Ann Intern Med.* (2022) 175:1693–706. doi: 10.7326/M22-1856

89. Andersson NW, Thiesson EM, Baum U, Pihlström N, Starrfelt J, Faksová K, et al. Comparative effectiveness of bivalent BA.4-5 and BA.1 mRNA booster vaccines among adults aged ≥50 years in Nordic countries: nationwide cohort study. *Br Med J.* (2023) 382:e075286. doi: 10.1136/bmj-2022-075286

90. Bajema KL, Berry K, Streja E, Rajeevan N, Li Y, Mutalik P, et al. Effectiveness of COVID-19 treatment with nirmatrelvir-ritonavir or molnupiravir among U.S. veterans: target trial emulation studies with one-month and six-month outcomes. *Ann Intern Med.* (2023) 176:807–16. doi: 10.7326/M22-3565

91. Gazit S, Saciuk Y, Perez G, Peretz A, Ben-Tov A, Stuart EA, et al. Hybrid immunity against reinfection with SARS-CoV-2 following a previous SARS-CoV-2 infection and single dose of the BNT162b2 vaccine in children and adolescents: a target trial emulation. *Lancet Microbe.* (2023) 4:e495–505. doi: 10.1016/S2666-5247(23)00103-9

92. Monteiro HS, Lima Neto AS, Kahn R, Sousa GS, Carmona HA, Andrade JSJr, et al. Impact of CoronaVac on COVID-19 outcomes of elderly adults in a large and socially unequal Brazilian city: a target trial emulation study. *Vaccine.* (2023) 41:5742–51. doi: 10.1016/j.vaccine.2023.07.065

93. Tran VT, Perrodeau E, Saldanha J, Pane I, Ravaud P. Efficacy of first dose of COVID-19 vaccine versus no vaccination on symptoms of patients with long covid: target trial emulation based on ComPaRe e-cohort. *Br Med J Med.* (2023) 2:e000229. doi: 10.1136/bmjmed-2022-000229

94. Català M, Burn E, Rathod-Mistry T, Xie J, Delmestri A, Prieto-Alhambra D, et al. Observational methods for COVID-19 vaccine effectiveness research: an empirical evaluation and target trial emulation. *Int J Epidemiol.* (2024) 53:dyad138. doi: 10.1093/ije/dyad138

95. Eusebi P, Speybroeck N, Hartnack S, Stærk-Østergaard J, Denwood MJ, Kostoulas P. Addressing misclassification bias in vaccine effectiveness studies with an application to COVID-19. *BMC Med Res Methodol.* (2023) 23:55. doi: 10.1186/s12874-023-01853-4

96. Brookmeyer R, Morrison DE. Estimating vaccine effectiveness by linking population-based health registries: some sources of bias. *Am J Epidemiol.* (2022) 191:1975–80. doi: 10.1093/aje/kwac145

97. Rosenberg ES. New COVID-19 cases and hospitalizations among adults, by vaccination status-New York, May 3-July 25, 2021. *Morbid Mortal Week Rep.* (2021) 70:37a6. doi: 10.15585/mmwr.mm7037a6

98. Cai M, Xie Y, Topol EJ, Al-Aly Z. Three-year outcomes of post-acute sequelae of COVID-19. *Nat Med.* (2024) 30:1564–73. doi: 10.1038/s41591-024-02987-8

99. Viper Group. *COVID-19 Vaccine Tracker*. (2022). Available at: https://covid19.trackvaccines.org/vaccines/ (accessed November 6, 2023).

100. De Gier B, Huiberts AJ, Hoeve CE, Den Hartog G, Van Werkhoven H, Van Binnendijk R, et al. Effects of COVID-19 vaccination and previous infection on Omicron SARS-CoV-2 infection and relation with serology. *Nat Commun.* (2023) 14:4793. doi: 10.1038/s41467-023-40195-z

101. De Gier B, Van Asten L, Boere TM, Van Roon A, Van Roekel C, Pijpers J, et al. Effect of COVID-19 vaccination on mortality by COVID-19 and on mortality by other causes, the Netherlands, January 2021-January 2022. *Vaccine.* (2023) 41:4488–96. doi: 10.1016/j.vaccine.2023.06.005

102. Mcconeghy KW, Hur K, Dahabreh IJ, Jiang R, Pandey L, Gellad WF, et al. Early mortality after the first dose of COVID-19 vaccination: a target trial emulation. *Clin Infect Dis.* (2024) 78:625–32. doi: 10.1093/cid/ciad604

103. Van Werkhoven CH, Valk AW, Smagge B, De Melker HE, Knol MJ, Hahné SJ, et al. Early COVID-19 vaccine effectiveness of XBB.1.5 vaccine against hospitalisation and admission to intensive care, the Netherlands, 9 October to 5 December 2023. *Euro Surveill.* (2024) 29:703. doi: 10.2807/1560-7917.ES.2024.29.1.2300703

frontiers | Frontiers in Medicine

TYPE Brief Research Report
PUBLISHED 19 November 2024
DOI 10.3389/fmed.2024.1437322

OPEN ACCESS

EDITED BY
Shisan (Bob) Bao,
The University of Sydney, Australia

REVIEWED BY
Ata Nevzat Yalcin,
Akdeniz University, Türkiye
Muhammed Bekçibaşı,
University of Health Sciences, Türkiye

*CORRESPONDENCE
Elena Netchiporouk
✉ elena.netchiporouk@mcgill.ca

†These authors have contributed equally to this work and share first authorship

RECEIVED 23 May 2024
ACCEPTED 30 October 2024
PUBLISHED 19 November 2024

CITATION
Prosty C, Le M, Lu Y, Khoury L, Cormier M, Cheng MP, Fowler R, Murthy S, Tsang JL, Lejtenyi D, Ben-Shoshan M, Rahme E, Golchi S, Dendukuri N, Lee TC and Netchiporouk E (2024) Effect of omalizumab on inflammatory markers in COVID-19: an exploratory analysis of the COVID-19 immunologic antiviral therapy with omalizumab (CIAO) trial.
Front. Med. 11:1437322.
doi: 10.3389/fmed.2024.1437322

Effect of omalizumab on inflammatory markers in COVID-19: an exploratory analysis of the COVID-19 immunologic antiviral therapy with omalizumab (CIAO) trial

Connor Prosty[1†], Michelle Le[2†], Yang Lu[3], Lauren Khoury[1], Maxime Cormier[4], Mathew P. Cheng[5], Robert Fowler[6], Srinivas Murthy[7], Jennifer LY Tsang[8], Duncan Lejtenyi[9], Moshe Ben-Shoshan[9], Elham Rahme[3], Shirin Golchi[3], Nandini Dendukuri[3], Todd C. Lee[5] and Elena Netchiporouk[1*]

[1]Faculty of Medicine, McGill University, Montreal, QC, Canada, [2]Division of Dermatology, Department of Medicine, McGill University, Montreal, QC, Canada, [3]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada, [4]Division of Respiratory Medicine, Department of Medicine, McGill University, Montreal, QC, Canada, [5]Division of Infectious Diseases, Department of Medicine, McGill University, Montreal, QC, Canada, [6]Department of Critical Care Medicine, Sunnybrook Health Sciences Centre, Toronto, ON, Canada, [7]Department of Pediatrics, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada, [8]Niagara Health Knowledge Institute, Niagara Health, St. Catharines, ON, Canada, [9]Division of Allergy, Immunology and Dermatology, Department of Pediatrics, Montreal Children's Hospital, McGill University, Montreal, QC, Canada

Background: The CIAO trial recently demonstrated a probable clinical benefit of omalizumab in the treatment of severe COVID-19; however, the mechanism underlying this benefit remains unclear. Therefore, we sought to longitudinally assess the impact of omalizumab on serum cytokines in CIAO trial patients to determine its mechanism of action.

Methods: Blood samples were collected on days 0, 2, 7, and 14 from patients recruited into the CIAO trial and who consented to this substudy. Blood samples were tested by a panel of 25 inflammatory cytokines, as well as for markers of mast cell activation. Levels of inflammatory biomarkers were compared over time between omalizumab- and placebo-treated patients by generalized linear mixed-effects model. Associations between biomarkers and clinical outcomes were investigated by mixed-effects logistic regression.

Results: Nineteen patients were recruited into this substudy; 10 were assigned to placebo and 9 to omalizumab. Monokine induced by gamma interferon was significantly positively associated with severe COVID-19 (Odds Ratio [OR]=1.06, 95%CI=1.00–1.11, p =0.043). Further, omalizumab significantly reduced interleukin-15 (Coefficient=–0.95, p =0.048) and macrophage inflammatory protein-1 (Coefficient=–1.31, p =0.010) levels. However, neither was significant in analyses adjusting for multiple hypothesis testing.

Conclusion: Although limited by a small sample size, these results suggest that omalizumab's potential benefit in COVID-19 may be mediated independently of modulation of the measured serum biomarkers. Further studies are needed to investigate omalizumab's mechanism of action in COVID-19.

KEYWORDS
omalizumab, coronavirus, COVID-19, clinical trial, cytokine

Introduction

Coronavirus disease 2019 (COVID-19) has had significant health and economic repercussions globally (1). The pathogenesis of COVID-19 is biphasic involving viral replication and dissemination followed by immune activation (2). Severe COVID-19 is postulated to involve pathological hyperinflammation (2, 3), which has been associated with impaired type I and III interferon (IFN) signaling (4–6). Type II inflammation has also been associated with severe disease (7). A small randomized controlled trial (RCT) studying dupilumab, a monoclonal antibody targeting IL4R-α, suggested a possible mortality benefit from inhibition of the type II inflammatory response in severe COVID-19 (8). Therefore, IFN and type II inflammation-targeting agents demonstrate potential as COVID-19 therapies.

Omalizumab is a monoclonal antibody that binds IgE and inhibits its interaction with the FceRI receptor on immune cells (9), used for refractory chronic spontaneous urticaria (10), asthma (11), and nasal polyps (12). Recent data suggest that omalizumab may have antiviral activity (13–15). In asthmatic patients, omalizumab significantly decreased the incidence of upper respiratory tract infections and the duration of viral shedding (15, 16). The postulated mechanism of action was inhibition of IgE-FceRI interaction augmenting type I IFN signaling (17, 18). Omalizumab has also been shown to attenuate type II inflammation/mast cell activation in asthmatic patients (13). In light of this evidence, the COVID-19 Immunologic Antiviral therapy with Omalizumab (CIAO) trial was conducted to assess the efficacy of omalizumab in moderate to severe COVID-19 (19). Despite early termination due to waning recruitment, among the 40 patients enrolled, omalizumab was associated with a 93% probability of reduction in death or mechanical ventilation on day 14 (19). The mechanism by which omalizumab conferred this benefit remains unclear. Thus, we sought to investigate omalizumab's mechanism of action in COVID-19 by analyzing its impact on longitudinal levels of various biomarkers.

Methods

Trial design

This substudy focused on the exploratory endpoints of the CIAO trial (19). Ethics approval was obtained from all participating sites and informed consent was acquired for each subject. Briefly, the CIAO trial was a multi-center randomized double-blind trial comparing a single dose of omalizumab 375 mg plus standard of care vs. placebo with standard of care in patients hospitalized for COVID-19 respiratory illness. The primary outcome was a composite of mechanical ventilation or all-cause mortality at 14 days with secondary endpoints including the time to clinical improvement of COVID-19 evaluated using the WHO 9-point ordinal scale (20). Data for this nested study were obtained from the subset of patients recruited at the McGill University Health Centre (MUHC), who consented to have blood samples drawn during their hospitalization on days 0, 2, 7, and 14 (±2 days).

Sample processing

Samples were centrifuged and the serum was isolated and stored at −80°C for subsequent batched analysis. The total IgE (IgE) (day 0), tryptase (day 0), and C-reactive protein (CRP) (days 0, 2, 7, 14) assays were performed by the MUHC central laboratory. CD63 expression, a marker of basophil activation, was measured using the CD63 flocast flow cytometry assay (Buhlmann; Schonenbuch, Switzerland), as previously described, on days 0, 2, 7, and 14 (21). A panel of 25 different cytokines was measured for all samples (days 0, 2, 7, and 14) using the cytokine 25-plex human panel immunoassay according to manufacturer's instructions (Invitrogen; Vienna, Austria).

Statistical analyses

All statistical analyses were performed using R (4.3.2, Foundation for Statistical Computing, Vienna, Austria). Demographic and clinical data were summarized with descriptive statistics. Mixed-effects logistic regression was used to determine whether measured biomarkers were associated with severe COVID-19 (WHO ordinal COVID-19 score 5–7). In this model, patient identification and study day were random-effect and fixed-effect variables, respectively. Generalized linear mixed-effects models (GLMM) with a gamma distribution were used to model biomarkers over time using the *lme4* package (22). Fixed-effect terms were used for study day and omalizumab assignment and a random-effect term for patient identification. A likelihood ratio test was performed to determine omalizumab's effect on biomarkers over time. *p*-values in each model were adjusted for multiple hypothesis testing separately using the Benjamini-Hochberg method (Q-value). Line graphs of the cytokine levels over time were plotted using the *ggplot2* package (23).

Results

Patient demographics

Of the 40 patients recruited into the CIAO trial, 19 were recruited at the MUHC and consented to blood tests for this substudy. Nine patients were assigned to omalizumab and 10 to placebo. The patient demographics, comorbidities, vaccination status, receipt of concomitant interventions, and WHO COVID-19 ordinal scores are presented in Supplementary Table 1. The median age of the study population was 65.0 years (interquartile range 59.5–80.0) and 31.6% were female.

Cardiovascular disease (63.2%), diabetes (52.6%), chronic kidney disease (26.3%), and cancer (21.1%) were the most common comorbidities. The majority (72.2%) of patients were vaccinated against COVID-19. At enrollment WHO ordinal severity scores for COVID-19 were most commonly 4 (36.8%), 3 (31.6%), 5 (21.1%), and 2 (10.5%). Concomitant COVID-19 therapies were received by all patients, including dexamethasone (100.0%), remdesivir (57.9%), tocilizumab (5.3%), and baricitinib (5.3%). The primary outcome, death or mechanical ventilation at day 14, occurred in 4 patients (21.1%).

Biomarker comparisons

Blood samples were obtained for 100% of participants on day 0, 73.7% on day 2, 52.6% on day 7, and 31.6% on day 14.

In unadjusted analyses, monokine induced by gamma interferon (MIG; OR = 1.06, 95%CI = 1.00–1.11, p = 0.043, Q = 0.77) was significantly positively associated with severe COVID-19, but was not significant when adjusting for multiple hypothesis testing (Table 1). None of the remaining biomarkers were statistically significantly associated with the outcome.

Mean biomarker values over time are plotted in Figure 1 for omalizumab vs. placebo. In unadjusted analyses, omalizumab significantly decreased interleukin (IL)-15 (Coefficient = −0.95, p = 0.048, Q = 0.41) and macrophage inflammatory protein-1 (MIP-1; Coefficient = −1.31, p = 0.010, Q = 0.27); however, these were not statistically significant after adjustment. Receipt of omalizumab did not significantly affect any of the other 25 biomarkers Supplementary Table 2.

Discussion

After adjustment, we did not find a statistically significant association between the biomarkers studied and disease severity or any significant differences in biomarkers over time as a function of omalizumab treatment.

Previous studies have noted that SARS-CoV2-specific IgE are correlated with COVID-19 severity (24, 25). While these results, combined with omalizumab's possible efficacy in COVID-19 may suggest a specific role for IgE or type II immunity in the pathogenesis of COVID-19, they may also merely be indicative of global immune dysregulation (7). Indeed, we did not identify any significant associations between disease severity and type II inflammatory cytokines, IFN-α, and markers of mast cell activation, in this study with limited sample size. Other studies on the subject revealed conflicting results with one study suggesting a positive association between type II inflammation and COVID-19 mortality (26), whereas another failed to detect a significant association between type II/type I inflammatory cytokine imbalance and death (27).

Our findings suggest that omalizumab's possible clinical benefit in COVID-19 is mediated by mechanisms other than modulation of serum levels of IFN-α or type II inflammation. These results contrast those for dupilumab in COVID-19, which demonstrated a probable clinical benefit of dupilumab and an associated reduction in type II inflammatory cytokines (e.g., eotaxin-3 and YKL-40) compared to placebo (8). Notwithstanding our study's limitations, there are several alternative explanations for these findings. All patients received concomitant corticosteroids which are known to broadly suppress inflammatory cytokines, including type II inflammation (28), and this may have obscured any omalizumab-induced differences and biased results toward the null. Alternatively, Djukanović et al. demonstrated that, compared to placebo, omalizumab treatment significantly reduced IL-4+ staining cells and eosinophils in bronchial biopsies in asthmatic patients (13). Therefore, omalizumab's effect may be mediated by immune modulation at the tissue rather than serum level. Additionally, the assay used for cytokine detection in this study does not analyze the levels of certain type II inflammatory cytokines (e.g., IL-9) and type I interferons (e.g., IFN-β), which have been shown in other studies to be modulated by omalizumab (18, 29). Consequently, we cannot exclude the possibility that omalizumab affected unmeasured cytokines.

TABLE 1 Mixed-effects logistic regression of biomarkers associated with severe COVID-19 (WHO ordinal COVID-19 score 5–7).

Biomarker	OR (95%CI)	p-value	Q-value
CD63+ (%)	1.59 (0.91–2.78)	0.10	0.77
CRP (mg/L)	1.03 (0.97–1.10)	0.28	0.95
Eotaxin (pg/ml)	0.72 (0.33–1.57)	0.40	0.95
GM-CSF (pg/ml)	NA	NA	NA
IFN-alpha (pg/ml)	0.98 (0.86–1.13)	0.82	0.95
IFN-gamma (pg/ml)	NA	NA	NA
IL-1 beta (pg/ml)	NA	NA	NA
IL-10 (pg/ml)	NA	NA	NA
IL-12/IL-23 (pg/ml)	1.05 (1.00–1.09)	0.072	0.77
IL-13 (pg/ml)	1.17 (0.70–1.93)	0.55	0.95
IL-15 (pg/ml)	0.97 (0.88–1.08)	0.58	0.95
IL-17A (pg/ml)	0.80 (0.49–1.31)	0.37	0.95
IL-1RA (pg/ml)	1.00 (0.99–1.00)	0.83	0.95
IL-2 (pg/ml)	0.99 (0.93–1.06)	0.82	0.95
IL-2R (pg/ml)	1.00 (1.00–1.00)	0.81	0.95
IL-4 (pg/ml)	1.00 (0.97–1.03)	0.94	0.95
IL-5 (pg/ml)	NA	NA	NA
IL-6 (pg/ml)	0.96 (0.90–1.03)	0.28	0.95
IL-7 (pg/ml)	0.89 (0.71–1.11)	0.29	0.95
IL-8 (pg/ml)	1.00 (0.98–1.02)	0.79	0.95
IP-10 (pg/ml)	1.00 (0.86–1.12)	0.95	0.95
MCP-1 (pg/ml)	1.00 (0.99–1.01)	0.54	0.95
MIG (pg/ml)	1.06 (1.00–1.11)	0.043	0.77
MIP-1 alpha (pg/ml)	1.00 (0.94–1.06)	0.87	0.95
MIP-1 beta (pg/ml)	1.00 (0.98–1.02)	0.72	0.95
RANTES (pg/ml)	1.00 (1.00–1.00)	0.93	0.95
TNF-alpha (pg/ml)	NA	NA	NA
Tryptase (ug/mL)*	0.90 (0.67–1.20)	0.47	0.95
IgE (ug/L)*	1.00 (1.00–1.01)	0.14	0.81

*IgE and tryptase were only performed on day 0; therefore, only a conventional logistic regression was performed for these laboratories.
NA: Not applicable. Too few values above the threshold of detection to perform a logistic regression.

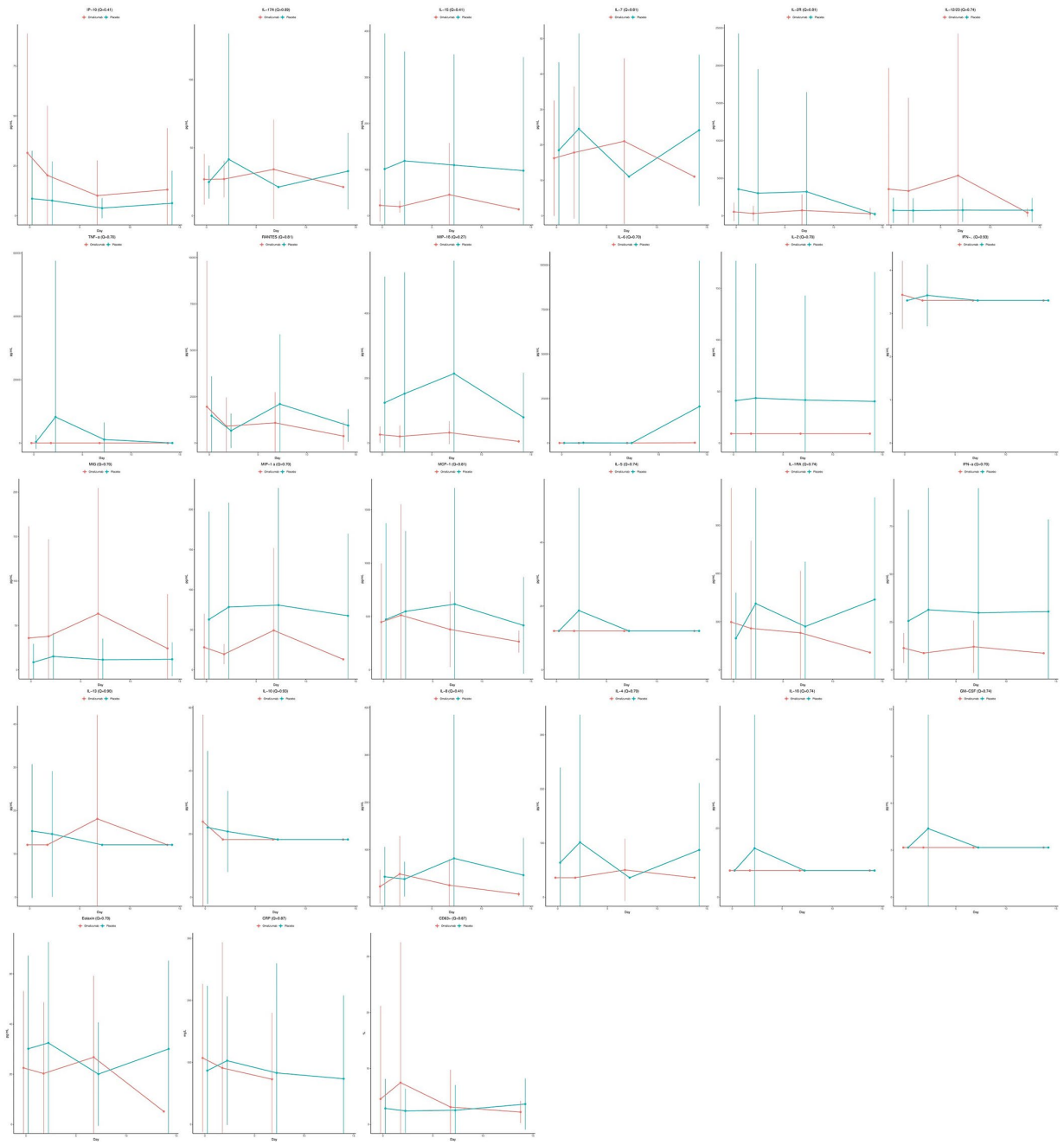

FIGURE 1
Mean cytokine values over time stratified by treatment assignment, with omalizumab in red and placebo in blue. Solid dots represent the mean and the error bars represent the standard deviations.

This study has multiple strengths including being conducted within the context of a randomized double-blind RCT, which minimized confounding by severity. Analyses were adjusted for multiple hypothesis testing to reduce the risk of type I error. Despite these strengths, this study is subject to limitations. First, the study population was small which may increase the risk of type II error. Second, the substudy population was further reduced by a loss to follow-up for blood sample collection on later study days, due to death or discharge. Third, many of the cytokine levels fell beneath the limit of detection of the assay, possibly due to concomitant receipt of corticosteroids.

Conclusion

This study suggests that the possible mortality benefit demonstrated by omalizumab in the CIAO trial could be mediated independently of modulation of serum IFN-α and type II inflammatory cytokines. Further

studies are necessary to elucidate the mechanism of action of omalizumab in COVID-19 and other viral illnesses.

Data availability statement

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

Ethics statement

The studies involving humans were approved by McGill Research Ethics Board, McGill University, Montreal, Canada. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. ML: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. YL: Data curation, Formal analysis, Methodology, Validation, Writing – original draft, Writing – review & editing. LK: Data curation, Investigation, Project administration, Writing – original draft, Writing – review & editing. MC: Data curation, Investigation, Writing – original draft, Writing – review & editing. MPC: Investigation, Project administration, Writing – original draft, Writing – review & editing. RF: Data curation, Investigation, Writing – original draft, Writing – review & editing. SM: Investigation, Writing – original draft, Writing – review & editing. JT: Investigation, Writing – original draft, Writing – review & editing. DL: Investigation, Validation, Writing – original draft, Writing – review & editing. MB-S: Investigation, Writing – original draft, Writing – review & editing. ER: Methodology, Validation, Writing – original draft, Writing – review & editing. SG: Formal analysis, Writing – original draft, Writing – review & editing. ND: Formal analysis, Writing – original draft, Writing – review & editing. TL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing. EN: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The study was funded by the Ministre de l'Économie et Innovation/Fonds de Recherche Santé Québec and the MUHC Foundation. Novartis donated the study drug.

Conflict of interest

ML reports funding support from the Canadian Institutes of Health Research (CIHR) Vanier Doctoral Scholarship. MC has been an Advisory Board/Speaker/Consultant for GSK, AstraZeneca, Sanofi/Regeneron, Novartis, Boehringer-Ingelheim, Trudell, Valeo, and COVIS. MPC reports grants from the McGill Interdisciplinary Initiative in Infection and Immunity and personal fees from GEn1E Lifesciences (as a member of the scientific advisory board) and personal fees from nplex biosciences (as a member of the scientific advisory board). JT reports a Physicians' Services Incorporated Foundation grant to her institution, outside the submitted work. She is co-chair of the Canadian Community ICU Research Network (CCIRNet) and vice-chair of the Quest Community Health Centre board of directors. SM reports a grant from the Health Research Foundation, Innovative Medicines Canada. MB-S reports salary support from Fonds de recherche du Québec – Santé. MB-S is part of the advisory Board or equivalent: Bausch, Stallergenes, Novartis, and Sanofi. Participating or participated in a clinical trial: Novartis, Aimmune, & Sanofi. ER reports funds and consulting fees from Janssen in the course of an unrelated study. TL reports salary support from Fonds de recherche du Québec – Santé. EN has been an Advisory Board/Speaker/Consultant and/or received Investigator Initiated Educational and/or Research funding from AbbVie Inc., Bausch Health, Beiersdorf, Boehringer Ingelheim International, Bristol Myers Squibb, Eli Lilly, Galderma SA., Janssen Inc., LEO Pharma, Medexus, Novartis Pharmaceuticals, Pfizer Inc., Sanofi Genzyme, Sun Pharmaceuticals, and UCB.

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2024.1437322/full#supplementary-material

References

1. WHO. Coronavirus (COVID-19) Dashboard. (2023). Available at: https://covid19.who.int (Accessed August 27)

2. Lamers MM, Haagmans BL. SARS-CoV-2 pathogenesis. *Nat Rev Microbiol.* (2022) 20:270–84. doi: 10.1038/s41579-022-00713-0

3. del Valle DM, Kim-Schulze S, Huang HH, Beckmann ND, Nirenberg S, Wang B, et al. An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat Med*. (2020) 26:1636–43. doi: 10.1038/s41591-020-1051-9

4. Bastard P, Rosen LB, Zhang Q, Michailidis E, Hoffmann HH, Zhang Y, et al. Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science*. (2020) 370. doi: 10.1126/science.abd4585

5. Zhang Q, Bastard P, Liu Z, le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science*. (2020) 370. doi: 10.1126/science.abd4570

6. Hadjadj J, Yatim N, Barnabei L, Corneau A, Boussier J, Smith N, et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science*. (2020) 369:718–24. doi: 10.1126/science.abc6027

7. Lucas C, Wong P, Klein J, Castro TBR, Silva J, Sundaram M, et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature*. (2020) 584:463–9. doi: 10.1038/s41586-020-2588-y

8. Sasson J, Donlan AN, Ma JZ, Haughey HM, Coleman R, Nayak U, et al. Safety and efficacy of Dupilumab for the treatment of hospitalized patients with moderate to severe coronavirus disease 2019: a phase 2a trial. *Open Forum Infect Dis*. (2022) 9. doi: 10.1093/ofid/ofac343

9. Easthope S, Jarvis B. Omalizumab. *Drugs*. (2001) 61:253–60. doi: 10.2165/00003495-200161020-00008

10. Maurer M, Rosén K, Hsieh HJ, Saini S, Grattan C, Gimenéz-Arnau A, et al. Omalizumab for the treatment of chronic idiopathic or spontaneous Urticaria. *N Engl J Med*. (2013) 368:924–35. doi: 10.1056/NEJMoa1215372

11. Busse W, Corren J, Lanier BQ, McAlary M, Fowler-Taylor A, Cioppa GD, et al. Omalizumab, anti-IgE recombinant humanized monoclonal antibody, for the treatment of severe allergic asthma. *J Allergy Clin Immunol*. (2001) 108:184–90. doi: 10.1067/mai.2001.117880

12. Gevaert P, Omachi TA, Corren J, Mullol J, Han J, Lee SE, et al. Efficacy and safety of omalizumab in nasal polyposis: 2 randomized phase 3 trials. *J Allergy Clin Immunol*. (2020) 146:595–605. doi: 10.1016/j.jaci.2020.05.032

13. Djukanović R, Wilson SJ, Kraft M, Jarjour NN, Steel M, Chung KF, et al. Effects of treatment with anti-immunoglobulin E antibody Omalizumab on airway inflammation in allergic asthma. *Am J Respir Crit Care Med*. (2004) 170:583–93. doi: 10.1164/rccm.200312-1651OC

14. Cardet JC, Casale TB. New insights into the utility of omalizumab. *J Allergy Clin Immunol*. (2019) 143:923–926.e1. doi: 10.1016/j.jaci.2019.01.016

15. Konstantinou GN, Podder I, Karapiperis D. Omalizumab prevents respiratory illnesses in non-atopic chronic spontaneous urticaria patients: a prospective, parallel-group, pilot pragmatic trial. *Clin Transl Allergy*. (2023) 13:e12279. doi: 10.1002/clt2.12279

16. Esquivel A, Busse WW, Calatroni A, Togias AG, Grindle KG, Bochkov YA, et al. Effects of Omalizumab on rhinovirus infections, illnesses, and exacerbations of asthma. *Am J Respir Crit Care Med*. (2017) 196:985–92. doi: 10.1164/rccm.201701-0120OC

17. Gill MA, Bajwa G, George TA, Dong CC, Dougherty II, Jiang N, et al. Counterregulation between the FcepsilonRI pathway and antiviral responses in human plasmacytoid dendritic cells. *J Immunol Baltim Md 1950*. (2010) 184:5999–6006. doi: 10.4049/jimmunol.0901194

18. Gill MA, Liu AH, Calatroni A, Krouse RZ, Shao B, Schiltz A, et al. Enhanced plasmacytoid dendritic cell antiviral responses after omalizumab. *J Allergy Clin Immunol*. (2018) 141:1735–1743.e9. doi: 10.1016/j.jaci.2017.07.035

19. Le M, Khoury L, Lu Y. COVID-19 immunologic antiviral therapy with Omalizumab (CIAO) – a randomized-controlled clinical trial. *Open Forum Infect Dis*. (2024). 11. doi: 10.1093/ofid/ofae102

20. Marshall JC, Murthy S, Diaz J, Adhikari NK, Angus DC, Arabi YM, et al. A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis*. (2020) 20:e192–7. doi: 10.1016/S1473-3099(20)30483-7

21. Netchiporouk E, Moreau L, Rahme E, Maurer M, Lejtenyi D, Ben-Shoshan M. Positive CD63 basophil activation tests are common in children with chronic spontaneous Urticaria and linked to high disease activity. *Int Arch Allergy Immunol*. (2016) 171:81–8. doi: 10.1159/000451084

22. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. (2015) 67:1–48. doi: 10.18637/jss.v067.i01

23. Wickham H. ggplot2: Elegant Graphics for Data Analysis | SpringerLink. (2009). Available at: https://link.springer.com/book/10.1007/978-3-319-24277-4 (Accessed January 18, 2024).

24. Plūme J, Galvanovskis A, Šmite S, Romanchikova N, Zayakin P, Linē A. Early and strong antibody responses to SARS-CoV-2 predict disease severity in COVID-19 patients. *J Transl Med*. (2022) 20:176. doi: 10.1186/s12967-022-03382-y

25. Tan C, Zheng X, Sun F, He J, Shi H, Chen M, et al. Hypersensitivity may be involved in severe COVID-19. *Clin Exp Allergy*. (2022) 52:324–33. doi: 10.1111/cea.14023

26. Baker JR, Mahdi M, Nicolau DV, Ramakrishnan S, Barnes PJ, Simpson JL, et al. Early Th2 inflammation in the upper respiratory mucosa as a predictor of severe COVID-19 and modulation by early treatment with inhaled corticosteroids: a mechanistic analysis. *Lancet Respir Med*. (2022) 10:545–56. doi: 10.1016/S2213-2600(22)00002-9

27. Pavel AB, Glickman JW, Michels JR, Kim-Schulze S, Miller RL, Guttman-Yassky E. Th2/Th1 cytokine imbalance is associated with higher COVID-19 risk mortality. *Front Genet*. (2021) 12:12. doi: 10.3389/fgene.2021.706902

28. Braun CM, Huang SK, Bashian GG, Kagey-Sobotka A, Lichtenstein LM, Essayan DM. Corticosteroid modulation of human, antigen-specific Th1 and Th2 responses. *J Allergy Clin Immunol*. (1997) 100:400–7. doi: 10.1016/S0091-6749(97)70255-0

29. Takaku Y, Soma T, Nishihara F, Nakagome K, Kobayashi T, Hagiwara K, et al. Omalizumab attenuates airway inflammation and interleukin-5 production by mononuclear cells in patients with severe allergic asthma. *Int Arch Allergy Immunol*. (2013) 161:107–17. doi: 10.1159/000350852

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 21 November 2024
DOI 10.3389/fmed.2024.1454863

OPEN ACCESS

EDITED BY
Vasiliki Epameinondas Georgakopoulou,
Laiko General Hospital of Athens, Greece

REVIEWED BY
Giuseppe Basile,
IRCCS Istituto Ortopedico Galeazzi, Italy
Carl Johan Drott,
Uppsala University, Sweden

*CORRESPONDENCE
Marco Crippa
✉ m.crippa51@campus.unimib.it

RECEIVED 25 June 2024
ACCEPTED 07 November 2024
PUBLISHED 21 November 2024

CITATION
Turati M, Gatti S, Rigamonti L, Zatti G, Munegato D, Crippa M, Benedettini E, Piscitelli D, Bigoni M and Turati R (2024) Impact of COVID-19 lock-down period on orthopedic and trauma surgical activity in a northern Italian hospital.
Front. Med. 11:1454863.
doi: 10.3389/fmed.2024.1454863

Impact of COVID-19 lock-down period on orthopedic and trauma surgical activity in a northern Italian hospital

Marco Turati[1,2,3], Simone Gatti[1], Luca Rigamonti[4], Giovanni Zatti[1,2], Daniele Munegato[1], Marco Crippa[1,2,3]*, Erik Benedettini[1,2,3], Daniele Piscitelli[2,5], Marco Bigoni[2,3,4] and Riccardo Turati[6]

[1]Orthopedic Department, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy, [2]School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy, [3]Transalpine Center of Pediatric Sports Medicine and Surgery, Monza, Italy, [4]Orthopedic Department, Policlinico San Pietro, Ponte San Pietro, Italy, [5]Department of Kinesiology, University of Connecticut, Storrs, CT, United States, [6]Department of Applied Economics, Universitat Autonoma de Barcelona, Barcelona, Spain

Objectives: This study aims to describe the impact of COVID-19 on Orthopedic and Trauma surgical activity in a single level-I trauma center in Northern Italy during the lockdown period. We proposed comparing surgical procedures performed during the outbreak and in the same period the previous year.

Methods: In this single-center retrospective epidemiological cohort study, the "lockdown cohort" of patients who were treated from March 1st to May 24th, 2020, was compared to the "control cohort" who received treatment during the same period in 2019. The primary outcome was to evaluate the differences between the lockdown and control cohorts regarding surgical volumes. The secondary outcome was to evaluate any differences in the type of surgical procedures performed in the two cohorts in the elective and emergency setting.

Results: Orthopedic surgical activity has suffered a global reduction of 72.4% during the lockdown period (from 36 ± 6.1 to 10.7 ± 8.4 per week; $p < 0.01$), with the ratio of emergency to elective operations increasing from 0.7:1 in 2019 to 3.3:1 in 2020. Elective surgery has in fact been almost completely suspended and was affected with a reduction of 88.9% (from 20.8 ± 5.2 to 4.3 ± 2.8 cases per week; $p < 0.01$), while emergency trauma surgery suffered a 49.7% reduction (from 15.1 ± 3.2 to 8.2 ± 6.1 cases per week; $p < 0.01$).

Conclusion: The COVID-19 outbreak severely impacted Italy, particularly the Lombardy region, and affected the national health system. The 2020 COVID-19 lockdown has heavily conditioned our Orthopedic and Trauma department surgical activity.

KEYWORDS

COVID-19, coronavirus, orthopedic, traumatology, lock-down

1 Introduction

Following the outbreak of the epidemic in China, Italy—particularly the Lombardy region—emerged as the second epicenter of the novel coronavirus disease 2019 (COVID-19). The first Italian case was recorded in Codogno on February 21st, 2020, and from there, the virus spread rapidly throughout the region and the entire country (1). In response to this national emergency, the Italian government adopted two main strategies: health policies focused on strengthening hospital system capacity, and preventive measures like lockdowns and social distancing to reduce the risk of virus transmission (2).

On March 9th, the Italian government announced a national lockdown (hereinafter DPCM-1), and the Lombardy Regional Council issued a decree to guide the regional response to the outbreak, reorganizing the health system. To control the spread of the virus, approximately 90% of hospital departments were dedicated to treating COVID-19 patients. This reallocation of resources led to a reduction in the number of functioning operating rooms, allowing hospital staff (such as operating room nurses and anesthesiologists) and medical equipment, including ventilators, to be redeployed for COVID-19 care. Additionally, all licensed physicians, including those from the national army, were recruited to assist in the crisis (3).

The orthopedics department, which handles both elective and emergency services, also contributed to the pandemic response. Therefore, the measures implemented in the orthopedic field had to be adjusted to reorganize staff and equipment to address the challenges posed by COVID-19, while still maintaining emergency orthopedic services (4).

In this study, we aimed to investigate and analyze how the SARS-CoV-2 pandemic affected orthopedic services, particularly in terms of surgical activity in both elective and emergency settings. Our primary objective was to evaluate the differences in surgical volumes between the lockdown period and a control cohort. The secondary objective was to assess any changes in the types of surgical procedures performed during these periods (2019 vs. 2020) in both the elective and emergency settings.

2 Materials and methods

In this single-center retrospective epidemiological cohort study we reviewed the operating room activities of an Orthopedic and Trauma Department performed during the first Italian "wave" of COVID-19 pandemic in a single level-I trauma center in northern Italy (5).

Orthopedic operating room (OR) activities are divided into two categories: Trauma and Elective surgeries. Elective surgeries include scheduled orthopedic procedures such as arthroplasties, foot and ankle surgeries, hardware removal, nonunion treatments, pediatric elective procedures, and Sports Medicine interventions. Before surgery, all elective patients were evaluated in an outpatient setting, after which an orthopedic surgeon scheduled their procedure. Trauma surgeries, on the other hand, involve patients presenting through the emergency department who are indicated for surgery. Importantly, no changes to surgical guidelines occurred during this period.

All patients who underwent orthopedic surgery were included in the analysis, regardless of the department to which they were admitted, with no age restrictions applied. Only patients with incomplete data were excluded.

Data on patient characteristics (including age and gender), wait times from hospitalization to surgery, principal diagnoses, and the surgical procedures performed were retrieved from our computerized medical records. All data collection adhered to current Italian and European privacy laws and followed ethical guidelines for study conduct.

Both diagnoses and procedures are registered according to the International Classification of Diseases ICD 9th edition (ICD-9). Collected data are divided into two cohorts:

- The "lockdown cohort" included patients who underwent surgery at our hospital during the 12-week period from March 1st to May 24th, 2020, which coincided with the full lockdown in Lombardy.
- The "control cohort" consisted of patients who were surgically treated during the same 12-week period in the previous year, 2019. To account for variations in surgical activity on different days of the week, we used 1-week time spans (7 days of activity) as the unit of analysis. This approach allowed us to evaluate evolving trends during the pandemic and make comparisons with the previous year's data.

Focusing on the total volume of surgical procedures, Figure 1 shows a significant decline in the weekly number of surgeries in 2020, with no procedures performed during week 3, which coincided with the first peak of COVID-19 cases at the end of March. Figure 2 highlights stark differences in the distribution of total surgeries over the entire period, with the median number of procedures in 2019 being nearly four times higher than in 2020. These findings highlight the relevant impact of the pandemic on surgical activities.

To understand whether such drop in surgical procedure influenced homogeneously the activity of the Orthopedic and Trauma center, we decompose the evolution and the distribution of surgical procedures between elective (Figures 3, 4) and trauma (Figures 5, 6) procedures. The results show stark differences in elective surgeries between the lockdown and control cohorts, whereas the differences in trauma procedures between 2019 and 2020 are less pronounced.

The COVID pandemic influenced not only the volume of surgical procedures, but also the type of procedures. To illustrate this, we analyzed the five most performed (i.e., "Top 5") procedures in 2019 and 2020, distinguishing between elective and trauma surgeries. In the elective field, there was a drastic reduction in both the volume and variety of procedures. Comparing Figures 7, 8 the number of joint replacement surgeries and arthroscopic procedures dropped by 90%. For instance, the meniscectomy, the most common elective surgery in 2019, was performed only three times in 2020. In contrast, as shown in Figures 9, 10 the types of trauma procedures remained relatively consistent between 2019 and 2020. The most frequently performed surgery in both cohorts was open reduction and internal fixation (ORIF) for femoral fractures (26.4% of all procedures in 2019 and 19.8% in 2020). In the lockdown cohort, hemiarthroplasty ranked second (12.1%), followed by ORIF for tibia

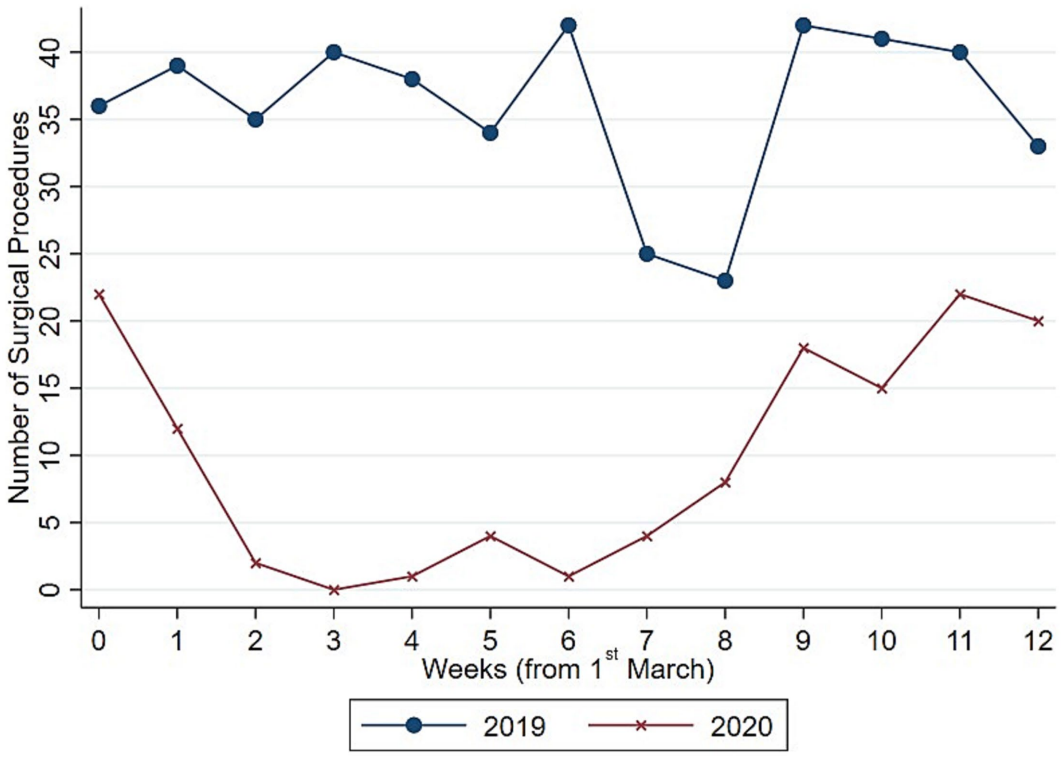

FIGURE 1
Evolution number of surgeries (total). Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure shows the weekly evolution in the Total Number of surgical procedures.

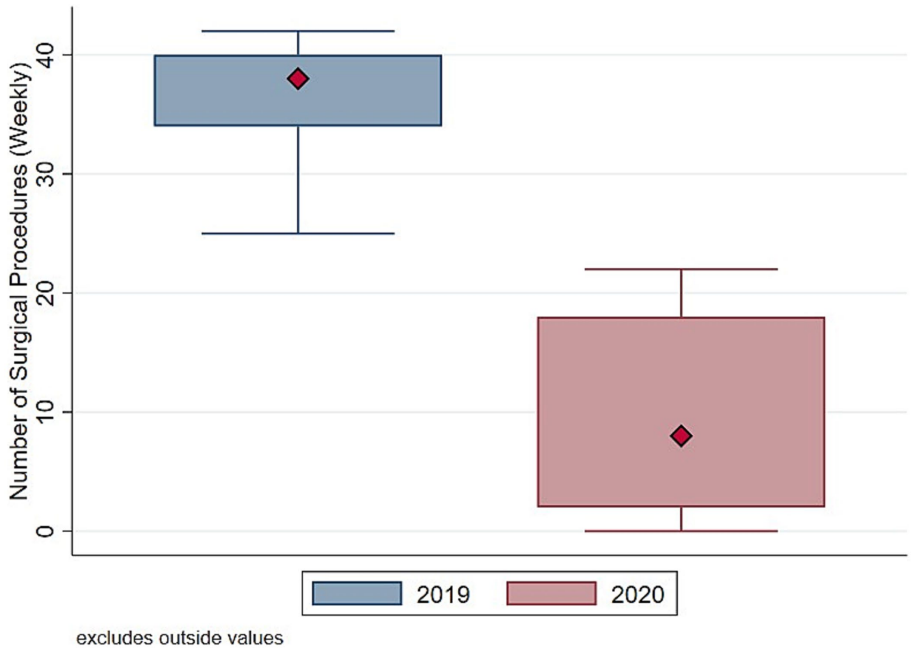

FIGURE 2
Distribution number of surgeries (total). Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure shows the boxplot of the distribution of weekly Total Number of surgical procedures.

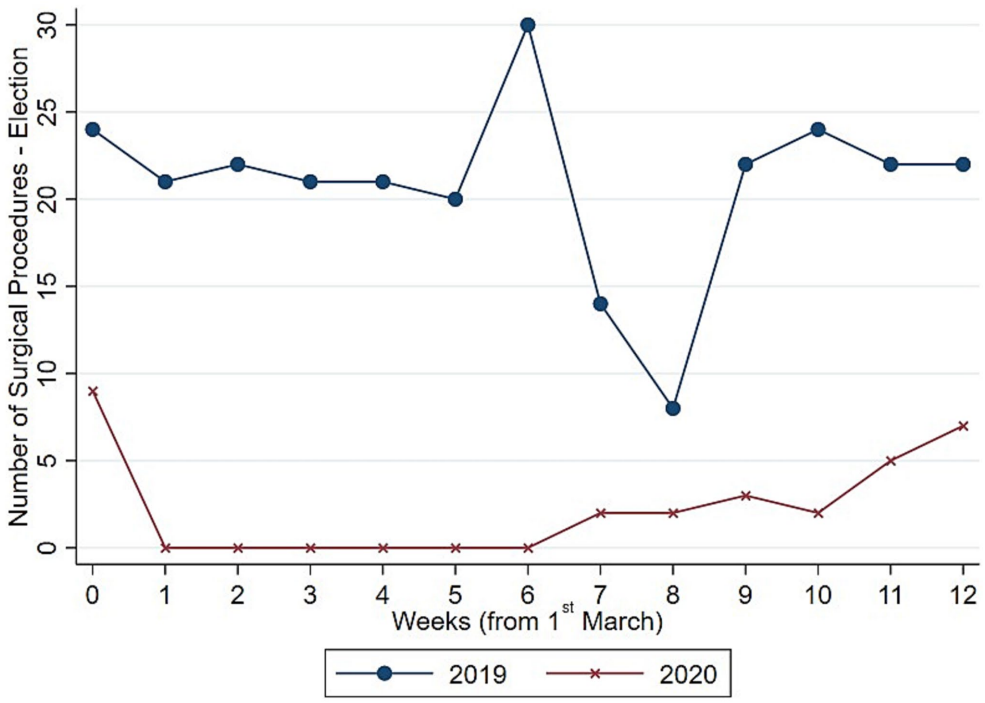

FIGURE 3
Evolution number of surgeries (elective). Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure shows the weekly evolution in the Number of Elective surgical procedures.

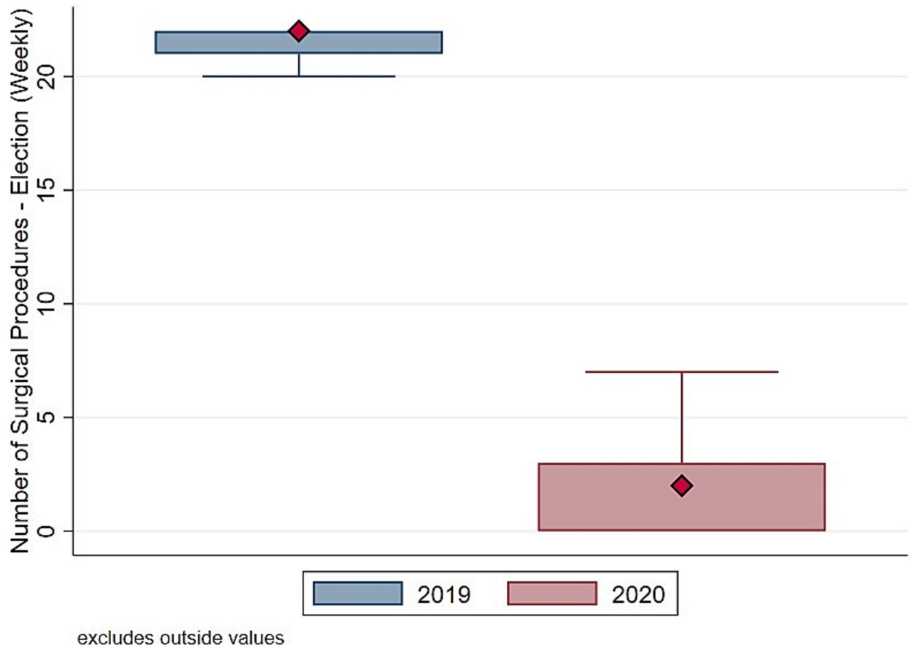

FIGURE 4
Distribution number of surgeries (elective). Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure shows the boxplot of the distribution of weekly Number of Elective surgical procedures.

and fibula fractures (8.1%), ORIF for ulna and radius fractures (8.1%), and ORIF for carpal-metacarpal fractures (5.0%). In 2019, after femur fracture surgeries, the next most common trauma procedures were ORIF for ulna and radius fractures (10.6%), ORIF for tibia and fibula fractures (9.6%), and hemiarthroplasty (7.6%).

Overall, these stylized facts provide three key pieces of evidence: (i) the total number of surgical procedures was significantly lower in

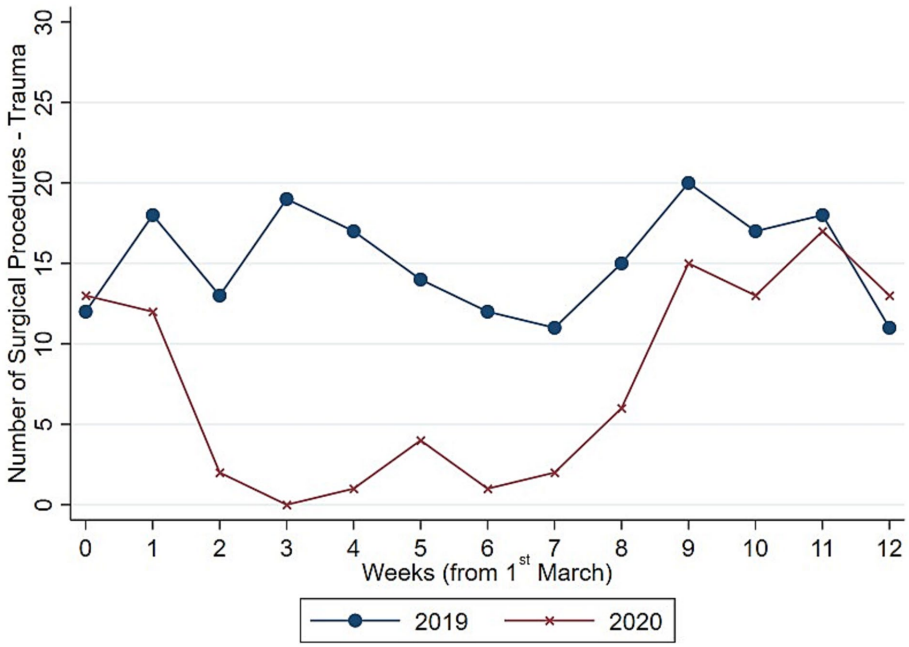

FIGURE 5
Evolution number of surgeries (trauma). Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure shows the weekly evolution in the Number of Trauma surgical procedures.

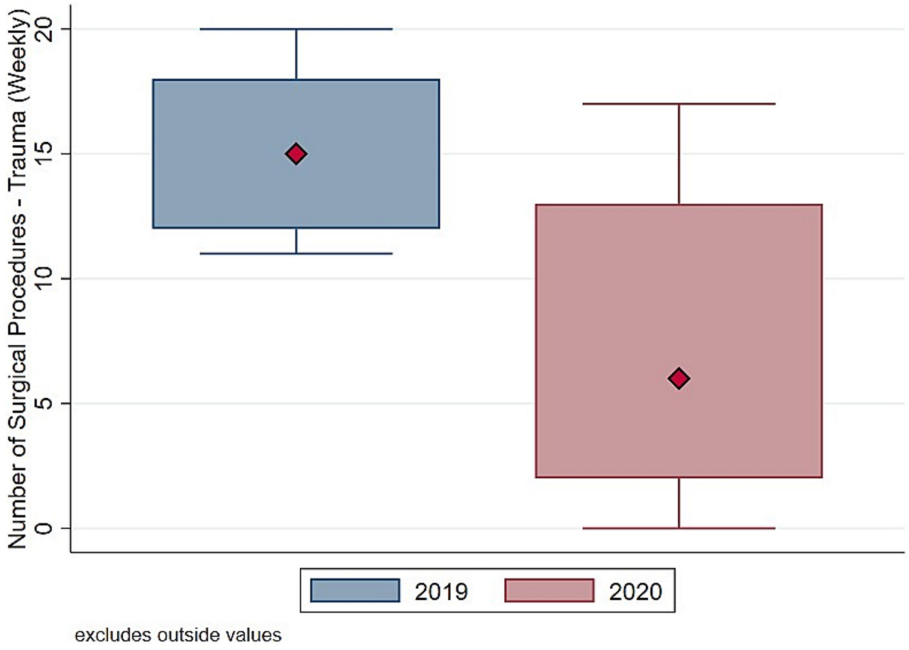

FIGURE 6
Distribution number of surgeries (trauma). Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure shows the boxplot of the distribution of weekly Number of Trauma surgical procedures.

2020 compared to 2019, (ii) the reduction was not uniform across elective and trauma procedures, and (iii) the types of elective procedures changed between 2019 and 2020, while this was less evident for trauma procedures. In the following section, we present our empirical strategy to validate these findings, controlling for confounding factors and the distribution of the outcome variables.

2.1 Statistical analysis

Our statistical analysis, conducted using STATA-16 for Windows, aims to determine whether the COVID-19 lockdown influenced (i) the volume of orthopedic surgical procedures and (ii) the types of procedures performed. After aggregating our individual data by week,

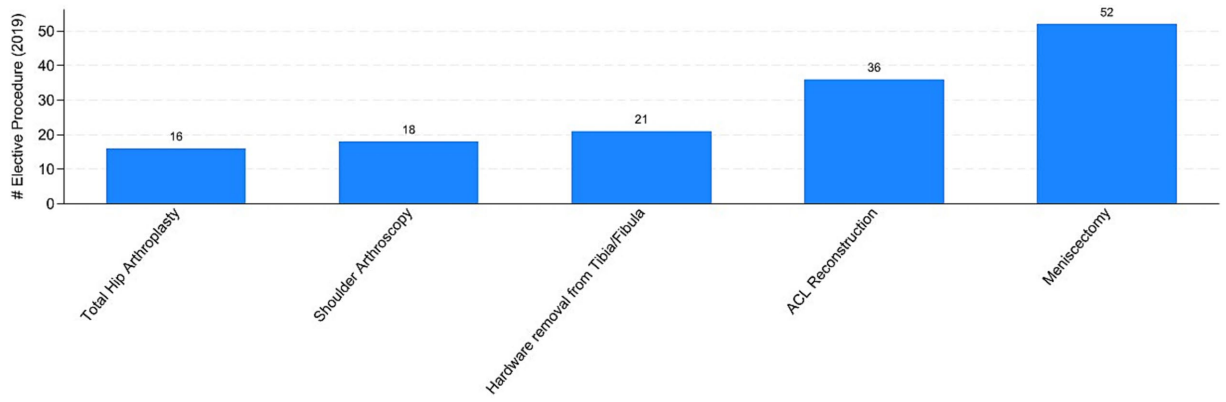

FIGURE 7
Top-5 elective procedures in 2019. Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure reports the top 5 elective surgical procedure in 2019. ACL stands for "Anterior Cruciate Ligament."

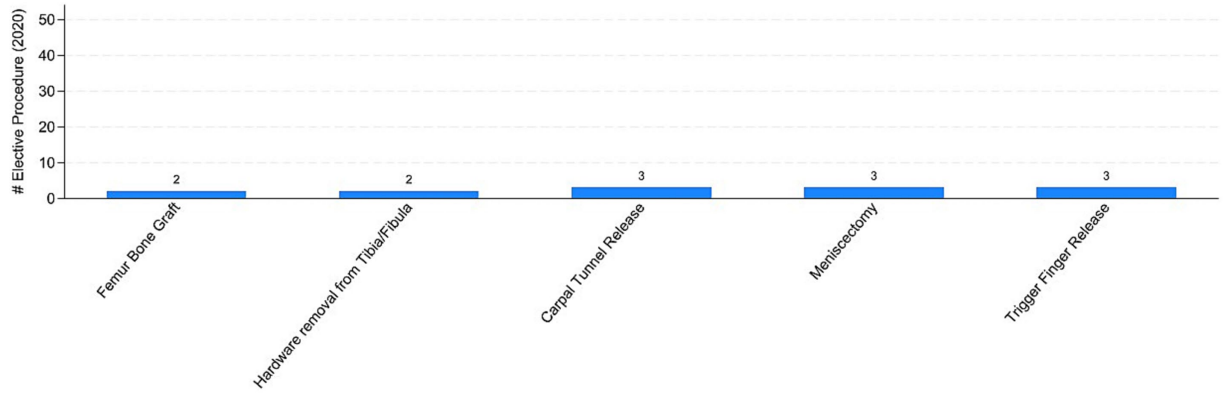

FIGURE 8
Top-5 elective procedures in 2020. Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure reports the top 5 elective surgical procedure in 2020.

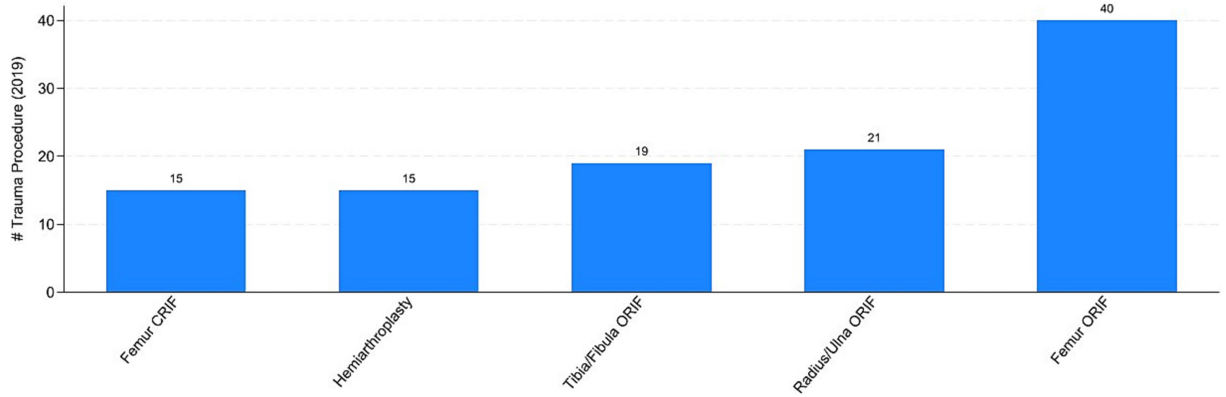

FIGURE 9
Top-5 trauma procedures in 2019. Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure reports the top 5 trauma surgical procedure in 2019. ORIF stands for "Open Reduction and Internal Fixation" and CRIF stands for "Closed Reduction and Internal Fixation."

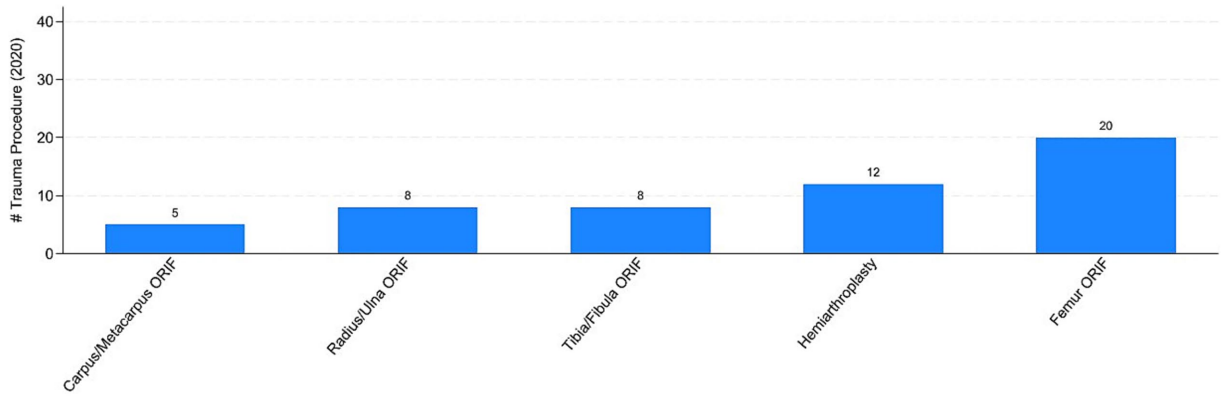

FIGURE 10
Top-5 trauma procedures in 2020. Authors' calculation overt Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The Figure reports the top 5 trauma surgical procedure in 2020. ORIF stands for "Open Reduction and Internal Fixation."

TABLE 1 Surgical procedures.

Year	(1) Total No.	(2) Trauma	(3) Trauma (%)	(4) Day(s) surg.	(5) Age	(6) Women (%)
2019	468	197	42.1	2.6	51.7	45.5
2020	129	99	76.7	2.6	55.8	48.1

Authors' calculations on data from Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). Columns (1)–(2) shows the total number of surgical procedures and trauma-specific, respectively. Column (3) presents the percentage of trauma-specific over the total. Column (4) shows the average number of days between the hospitalization and the surgery. Column (5) presents the average age of the patients, while column (6) shows the percentage of female patients.

we first compared the average number of surgical procedures between 2019 and 2020. To assess whether the differences in the number of procedures between these 2 years are statistically significant, we conducted a t-test and reported the results. Additionally, we examined whether average patient characteristics (age and gender) and the time from hospitalization to surgery differed between the lockdown cohort and the control cohort, ensuring that variations in the patient population did not account for differences in surgical activity.

To simultaneously control for confounding factors, we estimated the following linear model:

$$Y_w = \alpha + \beta Covid^{2020} + \Gamma X_w + \epsilon_w \quad (1)$$

where our outcome variable Y_w is the number of surgical procedure (total, elective, trauma) or the number of surgical procedures by type of procedure (e.g., Orhopaediatrics, Arthroplasty, etc.) for each week w $\in$ {0, 1, …, 12}. The variable of interest is $Covid^{2020}$, which is a dummy variable that takes value of one in the year 2020. The vector X_w includes a set of controls that can influence the volume and type of surgical procedure. The vector includes the average patient age, gender composition, and the average number of days between hospitalization to surgery. Finally, our model accounts for potential heteroskedasticity of the error term ϵ_w.

Estimating Equation 1 with a standard ordinary least square (OLS) estimator provides the estimates of the partial correlation between the lockdown period and surgical procedures $\left(\hat{\beta}\right)$ For instance, if the dependent variable is the total number of surgical procedures, a negative $\hat{\beta}$ suggests that in the lockdown period the weekly number of surgical procedures is smaller than the control period, after controlling for patient characteristics. Due to the substantial number of zeros in the dependent variable, OLS estimator can fit less precisely our data (6). For this reason, we follow (7) and we also estimate our baseline using the Poisson pseudo maximum likelihood estimator (hereafter PPML). Such an estimator is well suited to address large mass of zeros in the dependent variable and heteroskedasticity patterns in the error terms.

3 Results

3.1 Surgical volumes

From the early days of the outbreak, Fondazione IRCCS San Gerardo dei Tintori Hospital began reorganizing all departmental activities to address the emergency. Following the implementation of the lockdown on March 8th, 2020, more restrictive measures were adopted, including the suspension of all non-urgent surgical and outpatient services. Table 1 presents the total number of orthopedic and trauma procedures performed in 2019 and 2020. During the 61-day lockdown period (from March 1st to April 30th, 2020), a total of 129 cases were treated, compared to 468 cases during the same period in 2019. This represents a decrease of 339 cases, equating to a drop of approximately 72.4% (from an average of 36 ± 6.1 cases per week to 10.7 ± 8.4 cases per week; $p < 0.01$).

10.3389/fmed.2024.1454863

The reduction in orthopedic surgical volume is not proportionate across operative categories. Table 2 shows the differences across total, elective and trauma surgeries. Elective surgery has been almost completely suspended and it is the most affected with an 88.9% reduction (from 20.8 ± 5.2 to 4.3 ± 2.8 per week; $p < 0.01$). Emergency Trauma surgery suffers "only" a 49.7% reduction (from 15.1 ± 3.2 to 8.2 ± 6.1 per week; $p < 0.01$). Most of orthopedic surgical activities in the lockdown period consisted of emergency Trauma procedures (76.7%), with a Trauma-to-Elective (T:E) ratio of 3.3:1. Before the pandemic our control cohort registered a T:E ratio of 0.7:1, with Trauma surgery representing only 42.1% of total orthopedic OR activity. Notably, there were no significant differences in the demographic characteristics (sex and age) of the two patient cohorts, as shown in Table 2, indicating that variations in patient characteristics do not fully account for these differences. However, the number of days between hospitalization and surgery decreased in the emergency setting, dropping from an average of 5.2 days in 2019 to 2.1 days in 2020 ($p < 0.05$).

Finally, Table 3 presents the estimates from our benchmark Equation 1. The outcome variable includes the weekly total number of procedures [columns (1)–(4)], elective procedures [columns (2)–(5)], and trauma procedures [columns (3)–(6)]. Focusing on the OLS estimates [columns (1) to (3)], the results indicate that the lockdown significantly reduced the total number of orthopedic cases: on average, there were 25 fewer cases per week compared to the same period in 2019, including 18 elective cases and nearly 7 trauma cases. These coefficients are precisely estimated at the 1% significance level. The results are confirmed using a non-linear estimator (PPML) in columns (4) to (6). By interpreting the coefficients as semi-elasticities and controlling for confounding factors, we find that the pandemic decreased the weekly total number of orthopedic procedures by 71.1%. As expected, the results vary considerably between elective and trauma procedures: elective procedures were reduced by 88.4%, while trauma procedures experienced a milder reduction of 47.6% due to the pandemic.

3.2 Surgical procedures

Regarding the secondary outcome of the study, we assessed whether the pandemic influenced not only the total number of procedures but also affected the types of procedures conducted differently. Figures 7, 8 indicate that the most commonly performed elective procedures changed significantly between 2019 and 2020, while trauma procedures showed less variation. In this section, we test this observation by presenting estimates based on our empirical model outlined in Equation 1, after aggregating elective and trauma procedures into broad categories.

Concerning elective procedures, Table 4 presents the results on the number of procedures, aggregated into six broad categories: arthroplasty (col. 1), foot and ankle (col. 2), hardware removal (col. 3), non-unions and infections (col. 4), orthopediatrics (col. 5), and sports medicine (col. 6). Linear estimates are shown in the top panel, while the bottom panel presents non-linear estimates.

Focusing on the OLS results, we observe some degree of heterogeneity across procedures. During the pandemic, there were approximately 4.4 fewer procedures per week related to sports

TABLE 2 Surgical procedures—weekly statistics.

Type of Surgical Proc.	Total			Elective			Trauma		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	2019	2020	Difference	2019	2020	Difference	2019	2020	Difference
Number	36 (6.096)	10.75 (8.400)	-25.25*** (-8.65)	20.85 (5.178)	4.286 (2.812)	-16.56*** (-7.80)	15.15 (3.184)	8.250 (6.107)	-6.904*** (-3.59)
Patients' Age	51.59 (6.072)	58.10 (12.64)	6.504 (1.66)	45.91 (10.08)	47.14 (14.54)	1.231 (0.22)	59.39 (5.463)	60.27 (15.29)	0.872 (0.19)
Women (share)	0.454 (0.0677)	0.575 (0.253)	0.121 (1.66)	0.438 (0.154)	0.482 (0.339)	0.0440 (0.40)	0.495 (0.126)	0.593 (0.282)	0.0986 (1.14)
Day(s) Surg.	2.668 (1.474)	3.529 (5.544)	0.861 (0.54)	0.763 (0.900)	6.304 (14.66)	5.541 (1.39)	5.189 (3.040)	2.142 (1.477)	-3.047** (-3.14)
Observations	468	129	597	271	30	301	197	99	296

Authors' calculation on data from Orthopedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). The table shows the weekly average number of surgical procedures, patients' age, share of women and days between hospitalization and surgical operation over the years 2019 and 2020, and the differences between the 2 years over the total surgical procedures [col. (1)–(3)], over the elective surgical procedures [col. (4)–(6)] and over the trauma surgical procedures [col. (7)–(9)]. Standard errors are reported in parenthesis in columns (1), (2), (4), (5), (7), and (8) while p-values are reported in parenthesis in columns (3), (6), and (9). Level of significance: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

TABLE 3 Surgical procedures—regression results.

Estimation	OLS			PPML		
Surgical procedures	Total	Elective	Trauma	Total	Elective	Trauma
	(1)	(2)	(3)	(4)	(5)	(6)
COVID^{2020}	−25.121*** (3.171)	−18.301*** (1.952)	−6.821*** (1.929)	−1.243*** (0.225)	−2.158*** (0.361)	−0.646*** (0.209)
Age	0.254 (0.157)	0.046 (0.085)	0.209** (0.100)	0.017 (0.012)	0.006 (0.014)	0.024* (0.013)
Women	−18.194 (11.264)	−4.215 (5.964)	−13.980* (7.496)	−1.218 (0.885)	−0.834 (1.252)	−1.545* (0.892)
Day(s) Surg.	−0.125 (0.217)	−0.012 (0.109)	−0.113 (0.171)	−0.015 (0.026)	−0.005 (0.032)	−0.019 (0.030)
Observations	26	26	26	26	26	26
R^2	0.80	0.84	0.51	0.76	0.82	0.46

Authors' calculation on data from Orthopaedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). Robust standard errors are reported in parenthesis. Level of significance: *$p<0.1$, **$p<0.05$, ***$p<0.01$. The results report the coefficients from: OLS estimator [col. (1)–(3)], and PPML estimator [col. (4)–(6)]. The R2 associated to the PPML estimations is the standard Pseudo R^2.

medicine ($p<0.1$), 2.3 fewer for orthopediatrics ($p<0.05$), and 1.4 fewer associated with arthroplasty ($p<0.01$). Results for foot and ankle interventions, hardware removal, and non-unions/infections were not statistically different from zero. However, the lack of precision in some estimates may stem from the considerable number of zeros in the dependent variable, which can create estimation issues for a linear model. Indeed, the coefficients related to the pandemic period estimated from the non-linear model presented in panel B are all statistically significant, except for those associated with foot and ankle procedures.

Interpreting the estimates as semi-elasticities, elective procedures dropped by 68.9% for arthroplasty, 94.1% for hardware removal, 60.9% for non-unions/infections, 85.4% for orthopediatrics, and 93.4% for sports medicine during the lockdown. Overall, these results indicate that the lockdown influenced the various types of elective procedures differently.

Table 5 presents the linear and non-linear estimates associated with trauma surgical procedures, aggregated into the following seven groups: elbow, wrist, and hand fractures (col. 1); hip and femur fractures (col. 2); knee, ankle, and foot fractures (col. 3); orthogeriatric procedures (col. 4); pediatric procedures (col. 5); polytrauma (col. 6); and shoulder and upper arm fractures (col. 7). Interestingly, the estimated coefficients are not statistically significant or are only marginally significant ($p<0.1$) in a few cases. These results appear to confirm the evidence presented in Figures 9, 10. While there was a decrease in the total number of trauma cases, the types of procedures remained relatively consistent between 2019 and 2020.

4 Discussion

Our study demonstrates the significant impact of COVID-19 on elective and emergency orthopedic surgery. Given the unpredictable nature of the SARS-CoV-2 epidemic, particularly during the first wave, the lockdown policy implemented by the Italian government resulted in an overall reduction of approximately 70% in surgical volumes, dropping from 468 cases in 2019 to 129 cases in 2020. The reduction was more pronounced for elective surgeries compared to trauma emergency services (88.9% vs. 45.7%). During the initial weeks of the pandemic, international guidelines for orthopedic surgeons recommended minimizing in-hospital stays for all orthopedic patients and postponing any surgical procedures that could be deferred (8). Only non-deferrable conditions in the emergency setting, such as fractures and bone infections requiring urgent management, were admitted and treated accordingly. Notably, our data indicate a reduction in the number of days between hospitalization and surgery, decreasing from an average of 5.2 days in 2019 to 2.1 days in 2020 ($p<0.05$). However, these results can be partially explained by the redirection of residual surgical capacity toward managing these specific conditions and the reduction in in-hospital length of stay.

During the initial weeks of the pandemic, international guidelines for orthopedic surgeons recommended reducing the in-hospital length of stay for all orthopedic patients and postponing any surgical procedures that could be deferred (9, 10). Additionally, the decline in trauma-related admissions during the COVID-19 outbreak may be attributed to a reduction in public activities, fewer traffic-related accidents, and a decrease in sports injuries due to restrictive social distancing measures, such as the closure of gyms and sports clubs during the national lockdown (11, 12). Furthermore, the fear of contracting the virus may have deterred patients from presenting to emergency departments with minor traumas, contributing to the overall reduction in hospital admissions (4).

An epidemiological study conducted in Italy reported a significant drop in the incidence of proximal femoral fractures nationwide between 2019 and 2020. As illustrated in Figures 9, 10 our study corroborates this decline, with femoral fractures remaining the most frequently performed procedures (13).

The delay in treating non-urgent orthopedic surgeries represented a burden that the Italian health system is partially still facing (14). These restrictions also affected the diagnostic process, resulting in delays in both the diagnosis and subsequent treatment of several orthopedic conditions (15). Additionally, patients with confirmed diagnoses were unable to access treatment promptly. Consequently, the implementation of specific treatment protocols, which consider timeliness in the decision-making process, could not be realized (16, 17). The experiences of various medical centers have demonstrated that telemedicine could serve as a valuable tool to facilitate the diagnostic process. Further investigation is needed to equip patients and physicians with new, effective instruments for improving care (18), and to better explore the long-lasting consequences of such health shock.

Notably, Italian government regulations during this period also affected other emergency surgical services (19). Furthermore, an intriguing retrospective observational study revealed that the COVID-19 pandemic significantly impacted the volume of surgical

TABLE 4 Elective surgical procedures—regression results.

Type of procedure	Arthroplasty	Foot and ankle	Hardware removal	Non-unions infections	Orthopediatrics	Sports medicine
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS results						
COVID^{2020}	−1.414*** (0.495)	−0.536 (0.421)	−0.640 (0.426)	−0.585 (0.366)	−2.310** (0.915)	−4.421* (2.158)
Age	0.027** (0.010)	0.045** (0.019)	0.047* (0.025)	0.027** (0.009)	0.100* (0.054)	0.094* (0.048)
Women	−0.411 (0.785)	−1.016 (1.246)	−1.263 (1.487)	1.325 (0.860)	−0.125 (2.445)	1.473 (2.007)
Day(s) Surg.	−0.001 (0.006)	0.017 (0.023)	−0.481 (0.506)	−0.025 (0.049)	3.502 (2.807)	−1.724 (2.340)
Observations	26	26	26	26	26	26
R^2	0.81	0.64	0.57	0.59	0.60	0.89
Panel B: PPML results						
COVID^{2020}	−1.166*** (0.273)	−1.336 (1.057)	−2.824** (1.175)	−0.940** (0.435)	−1.924*** (0.313)	−2.713*** (0.572)
Age	0.046*** (0.007)	0.037*** (0.014)	0.027** (0.012)	0.037*** (0.007)	0.111*** (0.034)	0.022*** (0.008)
Women	−0.093 (0.292)	−0.605 (0.746)	−0.509 (0.583)	1.589*** (0.568)	0.138 (0.762)	1.755** (0.728)
Day(s) Surg.	0.009*** (0.003)	0.003 (0.010)	0.642 (0.403)	−0.014 (0.031)	0.624 (0.784)	1.951*** (0.750)
Observations	26	26	26	26	26	26
Pseudo R^2	0.79	0.47	0.46	0.53	0.56	0.90

Authors' calculation on data from Orthopaedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). Robust standard errors are reported in parenthesis. Level of significance: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Panel A reports the OLS estimates, while Panel B reports PPML estimates.

10.3389/fmed.2024.1454863

TABLE 5 Trauma Surgical procedures—regression results.

Type of procedure	Elbow wrist and hand fractures	Hip and femur fractures	Knee ankle and foot fractures	Orthogeriatric	Pediatric	Polytrauma	Shoulder and upper arm fractures
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS results							
COVID^{2020}	1.423 (0.937)	−0.146 (0.465)	−0.331 (0.509)	−1.222 (0.783)	−0.134 (0.236)	−0.457 (0.291)	0.008 (0.353)
Age	0.045*** (0.010)	0.044*** (0.009)	0.046*** (0.012)	0.023* (0.012)	0.152*** (0.026)	0.027*** (0.008)	0.053** (0.020)
Women	0.980 (0.811)	−1.439** (0.612)	−1.140 (0.819)	1.148 (1.012)	−0.186 (0.349)	−0.291 (0.656)	−1.679 (1.300)
Day(s) Surg.	0.102 (0.071)	−0.028 (0.054)	0.048 (0.050)	0.081 (0.159)	−0.500*** (0.165)	0.046** (0.018)	−0.092 (0.127)
Observations	26	26	26	26	26	26	26
R^2	0.58	0.54	0.54	0.41	0.72	0.67	0.50
Panel B: PPML results							
COVID^{2020}	0.761* (0.430)	−0.194 (0.316)	−0.071 (0.346)	−0.370* (0.217)	0.007 (0.391)	−0.778* (0.418)	0.314 (0.460)
Age	0.039*** (0.009)	0.035*** (0.008)	0.039*** (0.009)	0.038*** (0.011)	0.196*** (0.038)	0.036*** (0.008)	0.075*** (0.015)
Women	0.178 (0.320)	−0.721** (0.288)	−1.159** (0.464)	0.371 (0.313)	−0.058 (0.429)	−0.365 (0.444)	−1.817** (0.748)
Day(s) Surg.	0.075* (0.044)	−0.018 (0.027)	0.051* (0.027)	0.052 (0.037)	−0.538** (0.223)	0.024*** (0.008)	−0.051 (0.065)
Observations	26	26	26	26	26	26	26
Pseudo R^2	0.47	0.42	0.39	0.37	0.58	0.59	0.48

Authors' calculation on data from Orthopaedic and Trauma Department, Fondazione IRCCS San Gerardo dei Tintori Hospital, Monza (Italy). Robust standard errors are reported in parenthesis. Level of significance: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Panel A reports the OLS estimates, while Panel B reports PPML estimates.

10.3389/fmed.2024.1454863

procedures performed in a rural hospital. However, despite a shift towards non-operative management for conditions such as appendicitis and acute cholecystitis, there was no long-term increase in severe cases requiring surgery (20). The scientific community's research agenda should focus on investigating the extent to which changes in patient care plans during the COVID-19 pandemic affected health outcomes. The experiences of worldwide healthcare service reductions may provide valuable insights into minimizing unnecessary care (21).

The first wave of the pandemic exposed the world's unpreparedness to handle such an emergency. Since then, several improvements have been made. Flexibility and adaptability, along with scientific and technological advancements, have emerged as crucial factors in addressing this situation. Additionally, postoperative management has undergone various changes; the use of telemedicine has become a vital tool, enabling the monitoring of patients during rehabilitation without requiring them to visit a specialist in the hospital (22). Moreover, the COVID-19 pandemic highlighted the importance of primary care in ensuring continuity of care following hospital discharge.

In conclusion, the COVID-19 outbreak severely impacted the entire Italian national health system. Significant efforts were needed in terms of restrictions and the mobilization of healthcare workers, which hindered the ability to maintain standard care for various diseases. Our Orthopedic and Trauma department faced substantial challenges, leading to a significant reduction in all services, including emergency department consultations, outpatient consultations, and surgeries. Our experience may offer valuable insights for the future; however, addressing a future pandemic will require multidisciplinary coordination.

Data availability statement

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

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

Author contributions

MT: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. SG: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. LR: Conceptualization, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. GZ: Conceptualization, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. DM: Conceptualization, Data curation, Investigation, Writing – original draft, Writing – review & editing. MC: Conceptualization, Data curation, Methodology, Writing – original draft, Writing – review & editing. EB: Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. DP: Conceptualization, Data curation, Methodology, Writing – original draft, Writing – review & editing. MB: Conceptualization, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. RT: Conceptualization, Data curation, Formal analysis, 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 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

References

1. Zehender G, Lai A, Bergna A, Meroni L, Riva A, Balotta C, et al. Genomic characterization and phylogenetic analysis of SARS-COV-2 in Italy. *J Med Virol*. (2020) 92:1637–40. doi: 10.1002/jmv.25794

2. Alfano V, Ercolano S. The efficacy of lockdown against COVID-19: a cross-country panel analysis. *Appl Health Econ Health Policy*. (2020) 18:509–17. doi: 10.1007/s40258-020-00596-3

3. Piccinini D, Giunchi C, Olivieri M, Frattini F, Di Giovanni M, Prodi G, et al. COVID-19 lockdown and its latency in northern Italy: seismic evidence and socio-economic interpretation. *Sci Rep*. (2020) 10:16487. doi: 10.1038/s41598-020-73102-3

4. Wong JSH, Cheung KMC. Impact of COVID-19 on orthopaedic and trauma service. *J Bone Joint Surg*. (2020) 102:e80. doi: 10.2106/JBJS.20.00775

5. American College of Surgeons. Resources for optimal care of the injured patient. Chicago: Committee on Trauma (2014).

6. Angrist J, Pischke JS. (2009). Mostly harmless econometrics: an Empiricist's companion. In: Mostly Harmless Econometrics: An Empiricist's Companion.

7. Santos Silva JMC, Tenreyro S. Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Econ Lett*. (2011) 112:220–2. doi: 10.1016/j.econlet.2011.05.008

8. Massey PA, McClary K, Zhang AS, Savoie FH, Barton RS. Orthopaedic surgical selection and inpatient paradigms during the coronavirus (COVID-19) pandemic. *J Am Acad Orthop Surg*. (2020) 28:436–50. doi: 10.5435/JAAOS-D-20-00360

9. Wong KC, Leung KS. Transmission and prevention of occupational infections in orthopaedic surgeons. *J Bone Joint Surg*. (2004) 86:1065–76. doi: 10.2106/00004623-200405000-00029

10. Hirschmann MT, Hart A, Henckel J, Sadoghi P, Seil R, Mouton C. COVID-19 coronavirus: recommended personal protective equipment for the orthopaedic and trauma surgeon. *Knee Surg Sports Traumatol Arthrosc*. (2020) 28:1690–8. doi: 10.1007/s00167-020-06022-4

11. Qasim Z, Sjoholm LO, Volgraf J, Sailes S, Nance ML, Perks DH, et al. Trauma center activity and surge response during the early phase of the COVID-19 pandemic—the Philadelphia story. *J Trauma Acute Care Surg*. (2020) 89:821–8. doi: 10.1097/TA.0000000000002859

12. Piatti M, Turati M, Bigoni M, Gaddi D. Volleyball and COVID-19 emergency: experience of a high-level Italian club team. *Sport Sci Health*. (2021) 17:253–5. doi: 10.1007/s11332-020-00718-3

13. Ciatti C, Maniscalco P, Quattrini F, Gattoni S, Magro A, Capelli P, et al. The epidemiology of proximal femur fractures during COVID-19 emergency in Italy: a multicentric study. *Acta Biomed*. (2021) 92:e2021398. doi: 10.23750/abm.v92i5.11925

14. Bigoni M, Anghilieri FM, Piatti M, Ni Z, Turati M. From high-quality assistance to emergency orthopaedics during COVID-19 pandemic: a northern Italy scenario for sports medicine. *Sport Sci Health*. (2021) 17:489–91. doi: 10.1007/s11332-020-00721-8

15. Best MJ, Aziz KT, McFarland EG, Anderson GF, Srikumaran U. Economic implications of decreased elective orthopaedic and musculoskeletal surgery volume during the coronavirus disease 2019 pandemic. *Int Orthop*. (2020) 44:2221–8. doi: 10.1007/s00264-020-04713-8

16. Accadbled F, Turati M, Kocher MS. Osteochondritis dissecans of the knee: imaging, instability concept, and criteria. *J Child Orthop*. (2023) 17:47–53. doi: 10.1177/18632521221149054

17. Turati M, Rigamonti L, Giulivi A, Gaddi D, Accadbled F, Zanchi N, et al. Management of anterior cruciate ligament tears in Tanner stage 1 and 2 children: a narrative review and treatment algorithm guided by ACL tear location. *J Sports Med Phys Fitness*. (2023) 63:1218–26. doi: 10.23736/S0022-4707.21.12783-5

18. Moisan P, Barimani B, Antoniou J. Orthopedic surgery and telemedicine in times of COVID-19 and beyond: a review. *Curr Rev Musculoskelet Med*. (2021) 14:155–9. doi: 10.1007/s12178-021-09693-9

19. D'Urbano F, Fabbri N, Koleva Radica M, Rossin E, Carcoforo P. Emergency surgery in COVID-19 outbreak: has anything changed? Single center experience. *World J Clin Cases*. (2020) 8:3691–6. doi: 10.12998/wjcc.v8.i17.3691

20. Fabbri N, Pesce A, Uccellatori L, Greco S, Urgo MS, Oppici D, et al. Long term implications in surgical re-assisting (L.I.S.A. study) during the Covid-19 outbreak. A retrospective observational cohort study on a rural population. *Ann Ital Chir*. (2023) 94:195–202.

21. Moynihan R, Sanders S, Michaleff ZA, Scott AM, Clark J, To EJ, et al. Impact of COVID-19 pandemic on utilisation of healthcare services: a systematic review. *BMJ Open*. (2021) 11:e045343. doi: 10.1136/bmjopen-2020-045343

22. Basile G, Accetta R, Marinelli S, D'Ambrosi R, Petrucci QA, Giorgetti A, et al. Traumatology: adoption of the Sm@rtEven application for the remote evaluation of patients and possible medico-legal implications. *J Clin Med*. (2022) 11:3644. doi: 10.3390/jcm11133644

frontiers | Frontiers in Public Health

TYPE Perspective
PUBLISHED 18 March 2025
DOI 10.3389/fpubh.2025.1542344

OPEN ACCESS

EDITED BY
Vasiliki Epameinondas Georgakopoulou,
Laiko General Hospital of Athens, Greece

REVIEWED BY
Eric Shattuck,
Florida State University, United States

*CORRESPONDENCE
Claudio Russo
✉ crussoit@outlook.com

RECEIVED 17 December 2024
ACCEPTED 04 March 2025
PUBLISHED 18 March 2025

CITATION
Marsico G and Russo C (2025) Prolonged physical isolation, agonistic behaviour, and human resilience in pandemic times. *Front. Public Health* 13:1542344. doi: 10.3389/fpubh.2025.1542344

Prolonged physical isolation, agonistic behaviour, and human resilience in pandemic times

Giuseppina Marsico[1,2] and Claudio Russo[3]*

[1]Department of Human, Philosophical and Educational Sciences, University of Salerno, Fisciano, Italy, [2]Federal University of Bahia, Salvador, Brazil, [3]Stress Management and Therapy Clinic, Naples, Italy

With the purpose of enhancing a comprehensive approach to healthcare, public health initiatives have moved from managing the pandemic response towards an increased understanding of the sequelae, including but not limited to mental health issues triggered by societal limitations and precautionary measures. The long-term effects of the COVID-19 pandemic lie in the health system's capacity to promote a renewed sense of healthy communities, strengthen individual resilience, and mitigate environmental stressors in the future. Under these terms, the pandemic breakdown has been discussed in relation to the public health crisis and physical isolation resulting from SARS-CoV-2 disease.

KEYWORDS

physical isolation, pandemic bias, agonistic behaviour, health, resilience

Introduction

Social stressors are well-known to interfere with individual thoughts, triggering negative emotions and affecting human behaviour (1). Traditionally, the relationship between aggressive behaviour and social deprivation showed a response variability in laboratory studies of non-human animals and humans (2). Specifically, different types and models of aggression were proposed in the study of neural circuits behind the expression of aggressive behaviour, including environmental influences and the occurrence of social cues, emotions (e.g., fear or anxiety), motivational systems, and pleasure (3).

Since the fight against the SARS-CoV-2 pandemic has led us to 2 years of liminal feelings for the unknown consequences and cycle of the disease, human survival response to the infection transmission has resulted in a long-term impact on mental health (4). Nevertheless, expressed emotional states, one's lived experience, the healthcare system crisis, transnational policy interventions, and individual responses have exacerbated pre-existing health inequity and increased social disparities, which may affect human resilience (5).

Agonistic behaviour through the lens of contemporary science

In psychological terms, human agonistic behaviour may occur when external and/or internal stimuli elicit emotional processes, cognitive interpretation of events, or fight-or-flight responses. Previous studies in animal ecology have found long-lasting evidence that agonistic behaviour reduces reproduction and fertility and, conversely, it increases mortality and facilitates social dispersion (6). However, fear may drive aggressive behaviour, either in terms of primary or combined emotions expressed by human beings (7).

To develop a new framework on human agonistic behaviour, reinforced research efforts should move forward with the traditional model of cause-and-effect relationship (8), whereby

a study's focus should consider the health policy implications, the private and public dimensions, and the spatial distribution across dispersed geographical areas. Hence, the global effects of the pandemic are likely to represent "collective entities," namely distinct groups of individuals who are culturally responsive and whose actions are individually based on the "disease," the "policy," and the "economy" of the pandemic.

According to the proposed definition, multi-omics or social entities are meant to exist separately from other things. It follows that collective entities should be addressed as something having an independent existence, namely all the integrative entities, either individual or collective, which belong to the process of agonistic behaviours. Hence, our understanding of entities should be more focused on processes rather than their own existence.

Unveiling the complex ontology of individual biology on human society would identify discrete entities in past or future pandemics, including those collective entities that are either cohabitating or disengaging. In terms of pandemic outcomes, agonistic behaviour is hereby introduced at either the individual or population level.

Introducing the SARS-CoV-2 model of socio-behaviour analysis

First, the pandemic has resulted in emotional dysregulation, evidenced by manifestations of fear, anxiety, irritability, and frustration. These emotional responses are automatic and well-established reactions to deprivation. Second, experiencing the pandemic has likely resulted in episodic and semantic imprinting. This phenomenon involves space-oriented and one-time exposure to event memories, which are conveyed through a more general understanding of one's lived experiences. Third, self-isolation, quarantine, and limited physical interactions have exacerbated social deprivation through the enforcement of lockdown policies and social distance measures. Fourth, the immediate and delayed effects of the pandemic have led to unprecedented consequences on both the private and public spectrums. All four nodes of this model have produced a pattern of recovery (the "policy"), either in terms of individuals healing from the illness (the "disease") or societies recovering from economic constraints (the "economy"). A four-node representation of what we call the psyche of SARS-CoV-2 is shown in Figure 1.

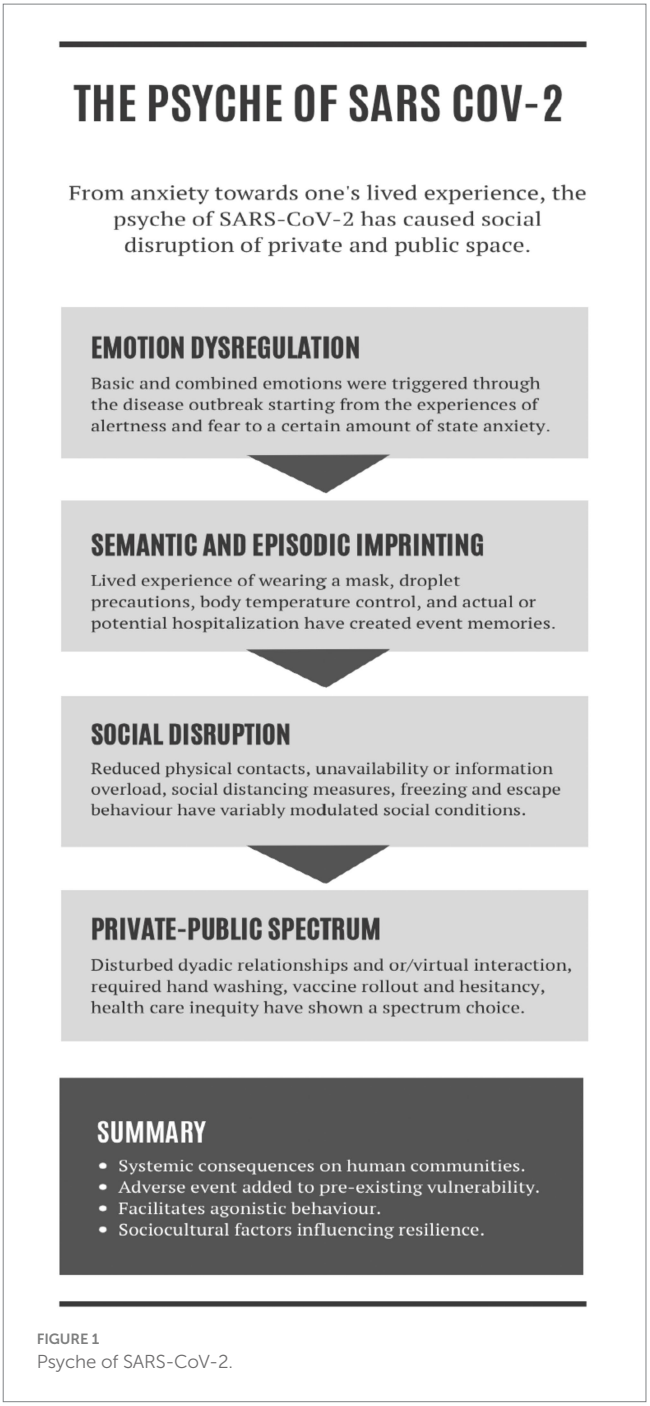

FIGURE 1
Psyche of SARS-CoV-2.

Understanding the social isolation of human ecosystems

The ongoing changes in how we interpret the environment have involved the replacement of natural spaces in response to ever-evolving human needs and new modalities of adaptation. Technological and economic infrastructures have added more complexity to cyclical patterns whilst combining previous health risks from exposure to new environments (9).

Following the SARS-CoV-2 pandemic, our research priorities should include the prolonged period of physical isolation and consider what the unforeseeable outcomes could have been in terms of agonistic behaviours. The intermittence of physical isolation may feasibly have a negative impact over prolonged periods or during human developmental time. Furthermore, the prolonged uncertainty over the disease and its transmission appeared to reinforce the effects of physical isolation on an individual basis ("disease") whilst influencing the real-time experiences amongst those collective entities ("economy" and "policy").

Nevertheless, unwarranted generalisations may raise controversies and debates about whether the scientific evidence gathered on an individual basis (micro level of analysis) is used to explain societal events of collective entities (macro level of analysis). In particular, the latter consequences refer to the physical isolation exacerbated by the "disease" and the "policy" on communities, whilst agonistic behaviour relies on the social dispersal of the "economy" resulting from the SARS-CoV-2 pandemic. Indeed, research efforts to unveil the relationship between social isolation and perceived

loneliness have shown the presence of complex proteomic networks, associations with morbidity and mortality profiles, and heterogeneity in health outcomes (10).

Discussion

By considering the reiteration of the events throughout human history that have inspired either societal change or defeat, we argue it is time for immediate action on public health policy. We have paid particular attention to one of the possible health outcomes of the COVID-19 pandemic, namely aggressive or defensive behaviour at individual, community-based, or global levels. In the post-pandemic world, we might suggest revising a dominant view when promoting individual and community health against the unified global threats (e.g., climate change, SARS-CoV-2, and war) and the divided global market or competing commercial interests involved.

However, unveiling individual and global phenomena in this era requires psychological science to provide its own traditional methods and novel strategies. Three levels of analysis were presented to argue how the "disease," the "policy," and the "economy" of the pandemic have shaped what we call the psyche of SARS-CoV-2. Our aim for proposing a new analysis model was to reflect upon the aggressive view of human behaviour and to interpret the complex societal patterns of human resilience (11).

Along with individual's readiness for positive change, the pandemic has triggered emotional dysregulation, created episodic and semantic imprinting, and generated social disruption over the private and public spectrum. Those collective entities, which also constitute the more socially disadvantaged ecosystems compared to the others, might jeopardise their own agonistic behaviour and be less likely to show collective resilience over time. As a result, agonistic behaviour might unintentionally increase systemic biases in medical research and policy.

Beyond the factors affecting an individual's resilience, we question what impact the pandemic has on global health systems and the social significance of human-induced actions, including the expression of agonistic behaviours worldwide.

In 1986, the Seville Statement on Violence concluded that the biological foundations of individual aggressive behaviour do not cause the war itself, whilst a historical attempt was made to prevent the confusion and misuse of either individual attitudes or political warfare (12).

By referring to collective entities as multi-omics or social entities, are the pandemic sequelae related to agonistic behaviour or showing an increase in the number of human casualties? For this purpose, new research is recommended as a crucial step to address a falsifiable and scientific integration of health, education, and culture (13).

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

GM: Conceptualization, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. CR: Conceptualization, Investigation, Writing – original draft, Writing – review & editing, Validation, Visualization.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Giuseppina Marsico acknowledges the Federal University of Bahia, Brazil (Edital Professor Visitante/PVB/UFBA and CAPES-PRINT for the financial support).

Acknowledgments

A preprint version of this article is available at: https://doi.org/10.31234/osf.io/hknd4 (14).

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 authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

References

1. Cunliffe VT. The epigenetic impacts of social stress: how does social adversity become biologically embedded? *Epigenomics*. (2016) 8:1653–69. doi: 10.2217/epi-2016-0075

2. Tóth M, Halász J, Mikics É, Barsy B, Haller J. Early social deprivation induces disturbed social communication and violent aggression in adulthood. *Behav Neurosci*. (2008) 122:849–54. doi: 10.1037/0735-7044.122.4.849

3. Nelson RJ, Trainor BC. Neural mechanisms of aggression. *Nat Rev Neurosci*. (2007) 8:536–46. doi: 10.1038/nrn2174

4. Manchia M, Gathier AW, Yapici-Eser H, Schmidt MV, De Quervain D, Van Amelsvoort T, et al. The impact of the prolonged COVID-19 pandemic on stress resilience and mental health: a critical review across waves. *Eur Neuropsychopharmacol*. (2022) 55:22–83. doi: 10.1016/j.euroneuro.2021.10.864

5. Russo C, Marsico G. Is the expression of agonistic behaviour precipitated by the pandemic events? *One Health*. (2023) 1:e14. doi: 10.1017/one.2023.6

6. King JA. The ecology of aggressive behavior. *Annu Rev Ecol Syst*. (1973) 4:117–38. doi: 10.1146/annurev.es.04.110173.001001

7. Anderson DJ, Adolphs R. A framework for studying emotions across species. *Cell*. (2014) 157:187–200. doi: 10.1016/j.cell.2014.03.003

8. Diener E, Northcott R, Zyphur MJ, West SG. Beyond experiments. *Perspect Psychol Sci*. (2022) 17:1101–19. doi: 10.1177/17456916211037670

9. Myers SS, Gaffikin L, Golden CD, Ostfeld RS, Redford KH, Ricketts TH, et al. Human health impacts of ecosystem alteration. *Proc Natl Acad Sci*. (2013) 110:18753–60. doi: 10.1073/pnas.1218656110

10. Shen C, Zhang R, Yu J, Sahakian BJ, Cheng W, Feng J. Plasma proteomic signatures of social isolation and loneliness associated with morbidity and mortality. *Nat Hum Behav*. (2025) 6:78. doi: 10.1038/s41562-024-02078-1

11. Luthar SS, Cicchetti D, Becker B. The construct of resilience: a critical evaluation and guidelines for future work. *Child Dev*. (2000) 71:543–62. doi: 10.1111/1467-8624.00164

12. Adams D. The Seville statement on violence: a progress report. *J Peace Res*. (1989) 26:113–21. doi: 10.1177/0022343389026002001

13. Napier AD, Depledge M, Knipper M, Lovell R, Ponarin E, Sanabria E, et al. Culture matters: using a cultural contexts of health approach to enhance policy making. Geneva: World Health Organization (2017).

14. Marsico G, Russo C. (2023). Prolonged physical isolation, agonistic behaviour, and human resilience. Available at: (https://osf.io/preprints/psyarxiv/hknd4_v1).

Frontiers in Medicine

Translating medical research and innovation into improved patient care

A multidisciplinary journal which advances our medical knowledge. It supports the translation of scientific advances into new therapies and diagnostic tools that will improve patient care.

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