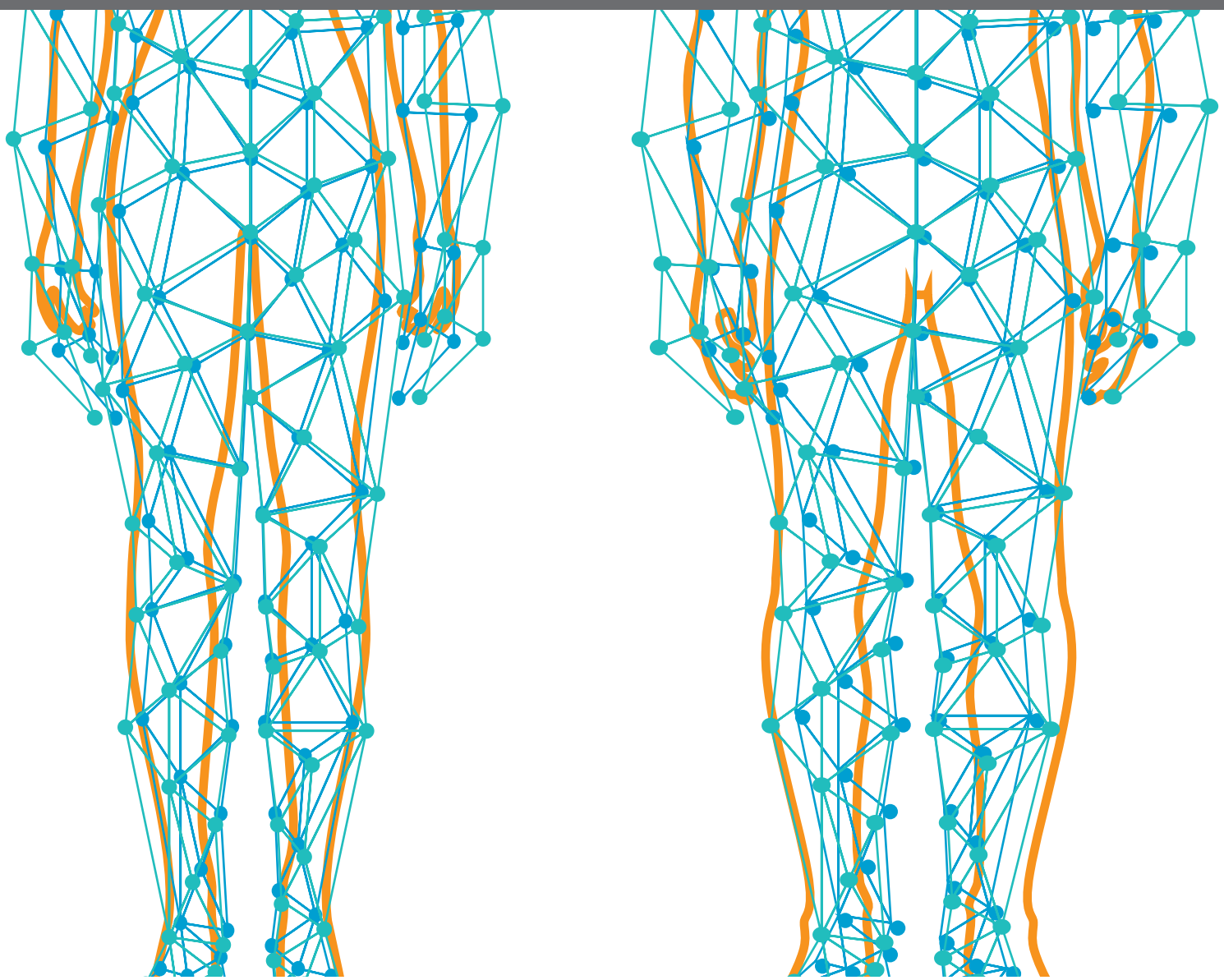
COVID ECOLOGY AND EVOLUTION: SYSTEMIC BIOSOCIAL DYNAMICS, 2nd Edition

EDITED BY: Matteo Convertino and Salvatore Flavio Pileggi
PUBLISHED IN: Frontiers in Medicine, Frontiers in Ecology and Evolution and Frontiers in Public Health



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-8325-2594-4
DOI 10.3389/978-2-8325-2594-4

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

COVID ECOLOGY AND EVOLUTION: SYSTEMIC BIOSOCIAL DYNAMICS, 2nd Edition

Topic Editors:
Matteo Convertino, Tsinghua University, China
Salvatore Flavio Pileggi, University of Technology Sydney, Australia

Due to the exceptional nature of the COVID-19 situation, Frontiers is waiving all article publishing charges for COVID-19 related research in this Research Topic.

Publisher's note: This is a 2nd edition due to an article retraction.

Citation: Convertino, M., Pileggi, S. F., eds. (2023). COVID Ecology and Evolution: Systemic Biosocial Dynamics, 2nd Edition.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2594-4

Table of Contents

frontiers in Ecology and Evolution

EDITORIAL
published: 03 September 2021
doi: 10.3389/fevo.2021.740213

Editorial: COVID Ecology and Evolution: Systemic Biosocial Dynamics

Matteo Convertino[1]* *and Salvatore Flavio Pileggi*[2]

[1] bluEco Lab, Tsinghua Shenzhen International Graduate School, Institute of Environment and Ecology, Tsinghua University, Shenzhen, China, [2] Centre on Persuasive Systems for Wise Adaptive Living (PERSWADE), Faculty of Engineering & IT, University of Technology Sydney, Sydney, NSW, Australia

Keywords: biocomplexity, environmental epidemiology, phylogenetics, healthcare, ecosystems, environmental change, pollution, complex networks

Editorial on the Research Topic

COVID Ecology and Evolution: Systemic Biosocial Dynamics

In a vastly impacted world by COVID-19, it is widely observable how ecosystem heterogeneity—i.e., biological and environmental more importantly, where the latter include socio-economical aspects—plays a fundamental role in shaping COVID-19 space-time impact and response (**Figure 1**). The lack to address ecosystem heterogeneity, that is partially reflected by administrative boundaries, within the decision making process can cause cascading extreme risks with multiple effects beyond morbidity and mortality in the human population. This is observed in many areas implementing disorganized control strategies or assuming non science-based "one size fits all" approaches. An ecosystem approach is needed where the collective ecology of COVID-19 is taken into account, from the portfolio of its determinants to controls. During the COVID-19 pandemic, due to a reduced human activity (e.g., in industrial production and mobility), surprising positive environmental outcomes also occurred, such as the decrease of air and water pollution. Meanwhile, many other still open research questions arose, such as the spread of SARS-CoV-2 through the hydrologic cycle and pathways.

In light of the above considerations this Special Issue aimed to embrace a systemic analysis of COVID-19 considering the multifaceted reality of the problem, including basic research and practical implications. Thirteen papers were published accordingly, spanning *biology* (Cueno and Imai; Cueno et al.), *environmental epidemiology* (Pansini and Fornacca; Rahalkar and Bahulikar; Ran et al.; Salom et al.), *public health and healthcare* (Marei et al.; Silva Campos et al.; Zhu et al.), and *health policy and social dynamics* (Chaudhury and Banerjee; Stöllberger and Winkler-Dworak; Wagner et al.; Yang et al.) of COVID-19. While certain papers aimed to investigate deeply the mechanisms underlying COVID-19 dynamics, others focused on determining socio-environmental associations useful for predictions and policy discussion. The differentiation between causality and predictability is a rather important point to emphasize due to the unfortunate generalized confusion between the two in "big data" and model-driven investigations that has deep policy implications.

Considering COVID-19 biology, very interestingly, Cueno and Imai found preliminary evidence suggesting that the structural conformation of SARS-CoV-2 (SARS2) spike protein is distinct from other known human-infecting CoVs; this may have triggered the viral pathogenesis (for humans) that mainly relies on the spike glycoprotein (with a furin-like cleavage site (FLC) as a structural feature) located on the virus surface. As a follow up Cueno et al. analyzed the COVID-19 genomic epidemiology network over time and detected nine SARS2 FLC patterns that potentially correspond to nine country clusters associated to the rapid evolution of the SARS2 genome and its infectivity. This evidently shed light into the environmental determinacy of viral biology and infectivity as well as SARS-CoV-2 uniqueness that lead to the pandemic.

OPEN ACCESS

Edited and reviewed by:
George L. W. Perry,
The University of Auckland,
New Zealand

***Correspondence:**
Matteo Convertino
matteo@ist.hokudai.ac.jp

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *12 July 2021*
Accepted: *16 August 2021*
Published: *03 September 2021*

Citation:
Convertino M and Pileggi SF (2021) Editorial: COVID Ecology and Evolution: Systemic Biosocial Dynamics. Front. Ecol. Evol. 9:740213. doi: 10.3389/fevo.2021.740213

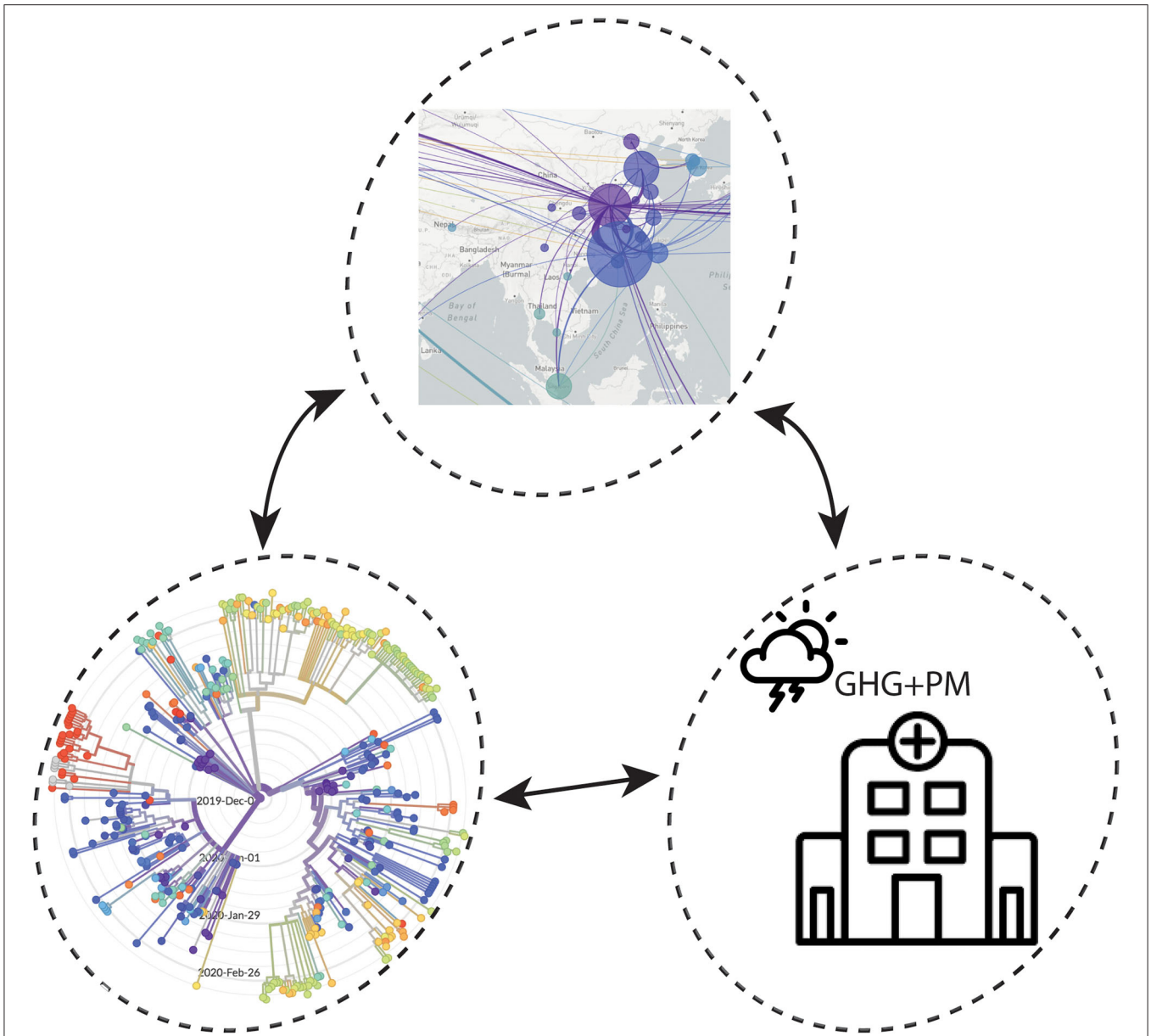


FIGURE 1 | Biocomplexity of COVID-19. The spread of COVID-19 worldwide manifested several ecological and evolutionary trajectories considering phylogenetic and socio-environmental determinants over space and time (including for instance climatological, air pollution, and healthcare factors underpinning vulnerability, exposure and control). Critical research and ecosystemic management of COVID-19 should consider the interdependencies of these factors underpinning novel and multidisciplinary areas such as eco-environmental phylogenetics and ecosystem health science. Phylogenetic tree and geographical spread map was made in https://nextstrain.org/sars-cov-2/; other icons are used under license from Shutterstock.com.

As for COVID-19 environmental vulnerability, Pansini and Fornacca, based on multiscale data from China, found a positive correlation between infections and mortality and Carbon Monoxide (CO), Formaldehyde, PM 2.5, and Nitrogen Dioxide (NO2) (retrieved by Sentinel-5 data) but lack of correlation with population density. These results emphasized how air pollutants are likely population vulnerability factors as suggested by other mechanistic studies highlighting the facilitation of virus spread (including SARS-CoV-2) via pollutant particles. This was also suggested by Salom et al. via detected positive correlations between the inferred reproduction number and pollution levels, as well as with temperature and humidity beyond other health and biological factors (such as cholesterol and blood Rh as individual vulnerabilities).

Contrarily to these findings, Ran et al. found that associations (adjusted by temperature and relative humidity) between three gaseous air pollutants (NO2, SO2, and CO) and COVID-19 basic reproductive number (assumed to represent transmissibility) are not statistically correlated. The multivariable linear regression model was used on different pollution estimates from China kriged via meteorological data.

Overall the discrepancy of findings between these studies emphasizes the caveat of predictions that are always based on input data, their space-time scale, model type and methodology (i.e., how a model is run) as predominant elements. Additionally, a warning is about the lack of consideration of non-linearity in these models that adopt statistical linear correlation for inferring "causality." It is rather important to stress that lack of correlation quite often is a signature of non-linear causation (as demonstrated by evidence on complex systems), and in general absence of correlation does not imply lack of causation between variables; this is easily provable by adopting a probabilistic rather than a categorical deterministic approach.

Yet, based on empirical evidence it is rather clear that persistent air pollution exposure is an ecological multi-scale public health issue where airborne pollutant particulates are likely acting as vector of COVID-19.

As for emphasizing the role of environmental disturbance into the ecology of emerging viruses, Rahalkar and Bahulikar found (based on previously published thesis) that a SARS-like CoV originating from Chinese horseshoe bats (Rhinolophus), was the predicted causative agent of Mojiang mineshaft miners' severe pneumonia-like illness in (2013). This provides very interestingly important clues into the potential ecosystem ontogeny of SARS-CoV-2.

As for public health more oriented toward population determinants, with Brazil data Silva Campos et al. tried to identify, via a MCDA model, the underpinning socio-economic and healthcare vulnerability of populations affected by COVID-19. Marei et al. emphasized the need of detailed identification of traveler's risk for minimizing COVID-19 spread. Zhu et al. highlighted how China prevention countermeasures for imported cases played an indispensable role in curbing COVID-19 spread, and suggested to prolong the 2-week quarantine period for monitoring asymptomatic patients.

In terms of health policy, Yang et al. discussed how COVID-19 represented a huge opportunity for clinical research. For this objective Stöllberger and Winkler-Dworak encouraged the sharing of clinical data for defining prognostic indicators and understanding the long-term pulmonary, cardiac, neurologic, and psychiatric consequences of COVID-19.

Lastly, Chaudhury and Banerjee suggested ecotherapy as a form of recovery from psychological effects of COVID-19.

In a broader perspective Wagner et al. reviewed economic and behavioral influencing factors of vaccine and antibiotics use and/or refusal globally. This has implications for the acceptance and coverage of COVID-19 vaccine: a topic that is extremely current due to the large vaccination hesitancy in many countries and subpopulations worldwide (also largely associated to misinformation generating risk aversion) that is impacting the curb of COVID-19 spread.

In conclusion, also in light of the diversity of this Special Issue papers, a firm belief should be that research and intervention for this pandemic and complex planetary issues alike, must overcome disciplinary boundaries as well as draw analogies and conclusions from previous events, including small unsuspicious local epidemics. This can raise early warning signals of disastrous systemic outbreaks whose pathogen biology may be largely unknown and extremely transmissible independently of many other factors, such the case of SARS-CoV-2 that "globalized" the world.

We underline the huge role of the environment for the ecology (and biology) of infectious diseases, considering both their emergence and spillover in humans due to environmental transformations, and disease control mediated by inhibitory or enhancing socio-environmental factors such as hygiene and air pollution. Environmental change, due to aggressive local development and ingrained climate change, is dramatically the core of problems like SARS2 emergence where organized and stable ecosystem connections (environment-biota including humans) are compromised. Therefore, in order to study and prevent future pandemics, the hope is to see more nature-oriented protection and monitoring before any public health control when the problem already occurred. Additionally, the hope is for science-based solutions in the decision-making process where accurate and salient information is gathered considering objectives, space-time scales, environmental and non-linear dynamics, especially when disease mechanisms are relevant and policy decision are made rather than for mere prediction exercises.

Open transdisciplinary collaborations are certainly creating the optimal platform for this modus operandi, targeting ecosystem or planetary health at the global scale.

AUTHOR CONTRIBUTIONS

MC developed the original framework of the Special Issue, acted as Chief Editor, and wrote the manuscript and realized the figure. SFP acted as Associate Editor for the Special Issue and revised the writing. All authors contributed to the article and approved the submitted version.

FUNDING

MC gratefully acknowledges the FY2020 SOUSEI Support Program and Award for Young Researchers awarded by the Executive Office for Research Strategy at Hokkaido University to the Top 20% scientists in terms of productivity and citations. MC acknowledges the High Talent Scheme ("A" Talents) of the Shenzhen Government as well as the Talents Hiring Program of Tsinghua University, Shenzhen International Graduate School.

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.

frontiers in Medicine

OPINION
published: 11 September 2020
doi: 10.3389/fmed.2020.00523

The Impact of COVID-19 Pandemic on Clinical Research in China: Challenges and Progress

Yuan Yang[1,2†], *Wen Li*[1†], *Ling Zhang*[3†], *Qinge Zhang*[3†], *Teris Cheung*[4], *Brian J. Hall*[5] *and Yu-Tao Xiang*[1,2*]

[1] Faculty of Health Sciences, Institute of Translational Medicine, University of Macau, Macao, China, [2] Southern Medical University Nanfang Hospital, Guangdong, China, [3] The National Clinical Research Center for Mental Disorders & Beijing Key Laboratory of Mental Disorders, Beijing Anding Hospital & The Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China, [4] School of Nursing, Hong Kong Polytechnic University, Hong Kong, China, [5] New York University (Shanghai), Shanghai, China

Keywords: COVID-19, China, clinical, research, progress

OPEN ACCESS

Edited by:
Salvatore Flavio Pileggi, University of Technology Sydney, Australia

Reviewed by:
Pedro Xavier-Elsas, Federal University of Rio de Janeiro, Brazil

***Correspondence:**
Yu-Tao Xiang
xyutly@gmail.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Infectious Diseases Surveillance, Prevention and Treatment, a section of the journal Frontiers in Medicine

Received: *07 July 2020*
Accepted: *27 July 2020*
Published: *11 September 2020*

Citation:
Yang Y, Li W, Zhang L, Zhang Q, Cheung T, Hall BJ and Xiang Y-T (2020) The Impact of COVID-19 Pandemic on Clinical Research in China: Challenges and Progress. Front. Med. 7:523. doi: 10.3389/fmed.2020.00523

INTRODUCTION

The Novel Coronavirus Disease 2019 (COVID-19 hereafter) was first reported in Wuhan, Hubei province, China at the end of 2019, and then was found in more than 200 countries worldwide. As of April 2020, the COVID-19 outbreak has been well-contained in China. It is beyond doubt that the COVID-19 outbreak has greatly transformed clinical research in China, and this transformation is of interest to the international community. Hence, we believe that a balanced and comprehensive overview describing the impact of COVID-19 on clinical research, the challenges researchers are facing, and an update on the progress of clinical research in China is warranted.

Research of Traditional Chinese Medicine During the COVID-19 Pandemic

Unlike other countries affected by COVID-19, traditional Chinese medicine (TCM) in China was widely prescribed in the prevention and treatment of COVID-19 among patients in clinical settings. A recent report released by the Information Office of the State Council, entitled "China Action Against the Novel Coronavirus Disease 2019," systematically introduced the functional role of TCM in the prevention and treatment of COVID-19 infections (1). In the past few months, around 92% of COVID-19 patients have been treated with TCM, with response rates reaching 90% in Hubei province, China (1). Some studies purported that certain TCM, such as Lianhuaqingwen (LH), a repurposed marketed product composed of 13 herbs, could effectively ameliorate COVID-19–related symptoms, such as fever, cough, and fatigue, and shorten the course of the disease (2). According to the 7th edition of the "Diagnosis and Treatment Protocol for Coronavirus Pneumonia" issued by the National Health Commission and the State Administration of Traditional Chinese Medicine, LH capsules were endorsed for the treatment of COVID-19 (3).

Ever since the outbreak of COVID-19, several preliminary studies have examined certain TCM with potential antiviral effects. For example, an open-label randomized controlled trial (RCT) found that LH capsules showed therapeutic effects on COVID-19 by improving the recovery rate of certain symptoms (especially mild symptoms), shortening the course of the illness, and improving chest radiologic abnormalities (2). An *in vitro* study found that LH significantly inhibits Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-COV-2) replication, affects virus morphology,

and exerts anti-inflammatory activity (4). These findings suggest that TCM provides antivirus effects along with symptomatic relief, and should be adopted as a novel supportive strategy in treating COVID-19. However, these findings are tentative due to methodological deficiencies, such as the lack of a double-blind design.

It is noteworthy that the recommendations on use of TCM for COVID-19 were mainly based on expert consensus, and solid evidence based on stringent RCTs demonstrating the safety and efficacy of TCM for COVID-19 is still insufficient. TCM emphasizes individualized treatment and syndrome differentiation even for patients with the same disease and severity (5). This individualization may hinder large-scale RCTs that examine the efficacy and safety of many TCM treatment approaches to COVID-19. In addition, the findings that 90% of COVID-19 patients improved from TCM treatments were mainly based on observational studies, but not RCTs. Therefore, considering that evidence of safety and effectiveness is indispensable, large-scale and double-blind RCTs on TCMs for COVID-19 are still warranted to meet the requirements of evidence-based medicine.

Clinical Research in China During the COVID-19 Pandemic

A recent review found that until March 26, 2020, a total of 681 COVID-19–related clinical trials have been registered on ClinicalTrials.gov and Chinese Clinical Trial Registry (ChiCTR), of which 481 had been conducted in China (6). Most of these registered studies focused on TCM, antiviral therapy, stem cell therapy, and plasma treatment; among these studies, 190 were RCTs (6, 7). On the one hand, the large number of registered trials reflect a heightened clinical research awareness and improved academic ability in China. On the other hand, however, many registered trials have obvious methodological limitations in terms of experimental design, biostatistics, data management, and preliminary data collection. For example, more than half of the registered clinical trials in China had a sample size of <100, whereas more than 30% of the clinical trials conducted in the USA, Italy, and France have a sample size of more than 500 (6). More importantly, a significant fraction of these registered clinical trials in China have failed to provide key information, such as treatment plan, drug dosage, study duration, or inclusion/exclusion criteria (7).

Because of the large number of registered clinical trials, alongside the implementation of quarantine measures in many areas, participant recruitment and follow-up assessments became immensely difficult in China (7). To ensure the quality of clinical trials and avoid wastage of resources, the Ministry of Science and Technology of China has thus enacted a strict supervision regulation, including a compulsory registration system and ethics examination procedure (7).

Since the COVID-19 outbreak, a huge volume of relevant publications drafted by researchers in China have been published, particularly in international peer-reviewed journals. For example, in a search for Severe Acute Respiratory Syndrome (SARS) and COVID-19–related publications in both PubMed and the China National Knowledge Infrastructure (CNKI) databases using the following search terms: "coronavir*," "severe acute respiratory syndrome," "SARS," "novel coronavir*," "COVID," "COVID-19," "China," and "Chinese," a total of 1,215 SARS-relevant English-language and 19,834 Chinese-language publications were retrieved from 2003 to 2004 (2 years). In contrast, during the period of January 1, 2020 to June 15, 2020 (only 5.5 months), 22,730 English and 17,098 Chinese COVID-19-related articles were retrieved. Specifically, only 58 SARS-related studies conducted in China were published on top international journals including the *Lancet*, *JAMA*, *British Medical Journal* (BMJ), and *New England Journal of Medicine* (NEJM) during the SARS outbreak. Nevertheless, during the period of January 1, 2020 to June 15, 2020, the corresponding figure of COVID-19–related publication has reached 417. This remarkable increase in the number of research outputs in China may be partly due to the growing attention to clinical research and the increasing academic communication with international community in the past two decades.

DISCUSSION

Despite the increasing number of publications during the COVID-19 outbreak in China, there are still some challenges to be overcome. First, standardized, sophisticated, and well-designed RCTs with large sample size are rare, which leads to insufficient empirical evidence affirming the safety and efficacy of TCM treatments for COVID-19. Second, most studies are cross-sectional and therefore, casual inferences are limited. Third, many of the clinical studies conducted in China are published in English journals, which compromises the dissemination of the study findings to frontline health professionals in China due to language barrier. Although some journals provide both English and Chinese versions of COVID-19–related articles, most other journals did not follow suit. In view of this, the Ministry of Science and Technology of China strongly encourage local researchers to publish their studies in Chinese journals instead of English journals since the end of January 2020. The Ministry of Education has also implemented new regulations that English articles are no longer listed as a pre-requisite requirement for professional/academic promotion (8). This state-sponsored push toward publications in Chinese may be a two-edge sword. On the one hand, this policy may make high-quality publications become more accessible to frontline health professionals who have difficulty reading English in China, but on the other hand, it will reduce the likelihood that they engage with a wider scientific community, thereby potentially limiting the rigor

and relevance of these studies relative to those reported in international journals.

In conclusion, the Chinese government and health sector have worked hard to make significant progress in improving clinical research. Consequently, clinical research has substantially advanced and flourished in China especially during the COVID-19 pandemic, although certain challenges are yet to be overcome. Lessons learned through the development of clinical research in China may be useful to address similar challenges in future pandemics.

AUTHOR CONTRIBUTIONS

YY, WL, LZ, and QZ: article draft. TC, BH, and Y-TX: article revision. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Science and Technology Major Project for investigational new drug (IND: 2018ZX09201-014) and the University of Macau (MYRG2019-00066-FHS).

REFERENCES

1. National Health Commission. *China Action Against the Novel Coronavirus Disease 2019* (in Chinese) (2020). Avaialble online at: http://m.news.cctv.com/2020/06/07/ARTIGx7s0Fu3qMsRNZ0kbwyx200607.shtml (accessed June 15, 2020).
2. Hu K, Guan WJ, Bi Y, Zhang W, Li L, Zhang B. Efficacy and safety of Lianhuaqingwen Capsules, a repurposed Chinese Herb, in Patients with Coronavirus disease 2019: a multicenter, prospective, randomized controlled trial. *Phytomedicine*. (2020) 16:153242. doi: 10.1016/j.phymed.2020.153242
3. National Health Commission. *Diagnosis and Treatment Protocol for Coronavirus Pneumonia*. 7th ed. (in Chinese) (2020). Available online at: http://www.nhc.gov.cn/yzygj/s7653p/202003/46c9294a7dfe4cef80dc7f5912eb1989.shtml (accessed June 15, 2020).
4. Runfeng L, Yunlong H, Jicheng H, Weiqi P, Qinhai M, Yongxia S. Lianhuaqingwen exerts anti-viral and anti-inflammatory activity against novel coronavirus (SARS-CoV-2). *Pharmacol Res*. (2020) 156:104761. doi: 10.1016/j.phrs.2020.104761
5. He YT, Ou AH, Yang XB, Chen W, Fu LY, Lu AP. Traditional Chinese medicine versus western medicine as used in China in the management of rheumatoid arthritis: a randomized, single-blind, 24-week study. *Rheumatol Int*. (2014) 34:1647–55. doi: 10.1007/s00296-014-3010-6
6. Li Q, Guo S, Kong Y, Zhang S. Analysis and consideration on clinical trials registration items during the epidemic of COVID-19 (in Chinese). *J Clin Exp Med*. (2020) 19:1029–33.
7. Huang Z, Yang G. Development and rethinking of the clinical trials during and after COVID-19 outbreak (in Chinese). *Chin J Clin Pharmacol Ther*. (2020) 25:591–4.
8. State Council. *Regulating the Use of Relevant Indexes of SCI Papers in Colleges* (in Chinese) (2020). Available online at: http://www.gov.cn/zhengce/zhengceku/2020-03/03/content_5486229.htm (accessed June 16, 2020).

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.

frontiers in Medicine

BRIEF RESEARCH REPORT
published: 24 September 2020
doi: 10.3389/fmed.2020.575839

Initial COVID-19 Transmissibility and Three Gaseous Air Pollutants (NO_2, SO_2, and CO): A Nationwide Ecological Study in China

Jinjun Ran[1,2†], Shi Zhao[3,4†], Lefei Han[5], Zhihang Peng[6], Maggie H. Wang[3,4], Yulan Qiu[2] and Daihai He[7†]*

[1] Li Ka Shing Faculty of Medicine, School of Public Health, The University of Hong Kong, Hong Kong, China, [2] School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [3] Jockey Club School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China, [4] Shenzhen Research Institute, Chinese University of Hong Kong, Shenzhen, China, [5] School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China, [6] School of Public Health, Nanjing Medical University, Nanjing, China, [7] Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China

In this study, we conducted an ecological study to examine their effects in the early phase of the pandemic (from December 2019 to February 2020) in China. We found that the associations between the average concentrations of NO_2, SO_2, and CO and the COVID-19 transmissibility are not statistically clear.

Keywords: gaseous air pollutant, COVID-19, transmissibility, ecological study, China

OPEN ACCESS

Edited by:
Matteo Convertino, Hokkaido University, Japan

Reviewed by:
Xiangjun Du, Sun Yat-sen University, China
Yu Zhao, Ningxia Medical University, China

****Correspondence:***
Daihai He
daihai.he@polyu.edu.hk

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

Specialty section:
This article was submitted to Infectious Diseases - Surveillance, Prevention and Treatment, a section of the journal Frontiers in Medicine

Received: *24 June 2020*
Accepted: *13 August 2020*
Published: *24 September 2020*

Citation:
Ran J, Zhao S, Han L, Peng Z, Wang MH, Qiu Y and He D (2020) Initial COVID-19 Transmissibility and Three Gaseous Air Pollutants (NO_2, SO_2, and CO): A Nationwide Ecological Study in China. Front. Med. 7:575839. doi: 10.3389/fmed.2020.575839

INTRODUCTION

The ongoing Coronavirus 2019–20 pandemic has caused a huge impact on global public health. Many studies showed that a high level of air pollutant concentrations increases the risk of pneumonia infection and deaths. The pollutant concentrations were significant for severe acute respiratory infection (1). Nitrogen dioxide (NO_2) exposure increases the occurrence, severity, hospitalization, and 30-day mortality, especially in cold months (2). Long-term exposure to higher levels of NO_2 increases the increased hospitalization for community-acquired pneumonia (CAP) in older adults (3). Short-term exposure to higher levels of NO_2 and carbon monoxide (CO) increases pneumonia-related hospitalization, emergency department visits, and outpatient visits in adults (4, 5). The concentrations of NO_2, sulfur dioxide (SO_2), and CO were positively associated with the upper respiratory tract infection and CAP (6). Becker and Soukop reported a non-linear dose-response effect between respiratory syncytial virus (RSV) internalization by airway epithelial cells and the level of NO_2 (positive association at a low-level 0.5 ppm of NO_2 and a negative association at a high-level 1.5 ppm) (7). Both the release of infectious virus 48-h post-exposure and virus-induced cytokine production were reduced at a high level of NO_2.

It is of significance to study the impact of gaseous air pollutants, especially NO_2, SO_2, and CO on the transmission of coronavirus diseases 2019 (COVID-19) in China. Many studies have been done on the impact of meteorological factors and air pollutants on the spread and COVID-19 and the

TABLE 1 | Descriptive statistics of the basic reproductive numbers, three gaseous pollutants, temperature, and relative humidity across 154 Chinese cities.

	Mean	SD	Min	25th	Median	75th	Max	IQR
COVID-19 transmissibility								
R_0	1.4	0.3	1.0	1.1	1.3	1.5	2.5	0.4
Gaseous pollutants								
NO_2 (μg/m^3)	31.0	7.8	13.1	25.7	31.1	36.1	49.3	10.4
SO_2 (μg/m^3)	11.9	6.0	5.3	7.8	9.7	14.3	42.2	6.5
CO (mg/m^3)	1.0	0.2	0.5	0.8	0.9	1.1	1.7	0.3
Weather conditions								
Temperature (°C)	4.3	8.3	−22.5	0.4	6.2	9.3	19.1	8.9
RH (%)	73.9	9.2	42.2	67.9	75.6	81.0	89.0	13.1

R_0, basic reproductive number; NO_2, nitrogen dioxide; SO_2, sulfur dioxide; CO, carbon monoxide; RH, relative humidity; SD, standard deviation; IQR, interquartile range.

patient outcome, for example, focusing in China (8–11), as well as other studies focusing in Italy, Spain, and the United States. However, some of these studies adopted a chronic disease approach, using either the linear regression model or time series model. However, COVID-19 is a highly transmissible infectious disease. We cannot merely adopt a statistical model for chronic disease to gain insights on the effects on the transmissibility (12). For instance, the serial correlation in daily new cases should be removed in a time series model (9). Riccò et al. (13) pointed out several other issues. For example, the lockdown of cities not only reduced the transmission of COVID-19 but also reduced the level of air pollution. Thus, the correlation between the level of certain pollutants and the transmission rate over time does not mean causation. The correlation between the transmission and meteorological conditions and air pollution could be non-linear (11).

METHODS

We collected daily confirmed COVID-19 cases from 303 cities in China from the Chinese provincial health agencies and China National Health Commission. We calculated the basic reproductive number, R_0, which is a unit-free measure of infectivity of a virus and is commonly used in infectious disease epidemiology. For the estimation of R_0, we first estimated the exponential (or intrinsic) growth rate, denoted by r, of the epidemic curve over the a 16-day period starting from the confirmation of the first case in each city (14–16). The number of cases at the t-th day, C_t, is modeled as $C_t = C_0\exp(rt)$, where C_0 denoted the number of seed cases at the start of the outbreak. Using the formula $R_0 = 1/M(-r)$ that is derived from the Lotka-Euler equation (17), we substitute r into the moment generation function, $M(\cdot)$, of the probability distribution of the COVID-19 serial interval. The distribution of serial interval is approximated by a Gamma distribution with mean at 5.5 days and standard deviation at 3.3 days (17–20). This analytical approach is also used in previous COVID-19 studies as well as in other infectious diseases (20–23).

We obtained air pollutant data in 1,642 observation stations from the China National Environmental Center and obtained meteorological data from the National Meteorological Data Center. We computed a raster for each pollutant by the kriging interpolation based on the averaged values across the period December 10, 2019, to February 29, 2020. Then we interpolated the level of each pollutant for each city. We adopted generalized (i) univariable, and (ii) multivariable linear regression, using

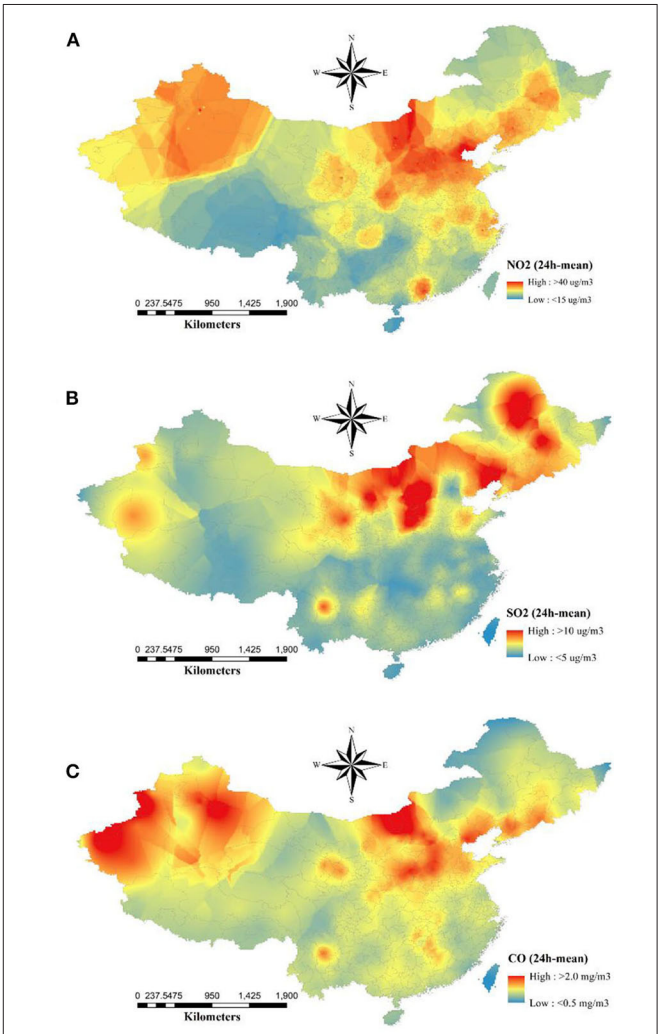


FIGURE 1 | Spatial distribution of average concentrations of ambient **(A)** NO_2, **(B)** SO_2, and **(C)** CO from December 10, 2019 to February 29, 2020. The raster maps were converted by the kriging interpolation after averaging their concentrations.

R_0 as the response and air pollutant as the factor, adjusted by temperature and relative humidity. In addition, we employ the spline regression for sensitivity analysis.

RESULTS

The descriptive statistics of the basic productive numbers, three pollutants, temperature, and relative humidity for 154 cities are given in **Table 1**. The maximum $R_0 = 2.5$ was detected in Wuhan City. The average concentrations of NO_2, SO_2, and CO were 31.0, 11.9 $\mu g/m^3$, and 1.0 mg/m^3, respectively. The spatial distributions of NO_2, SO_2, and CO are shown in **Figure 1**. Pearson and Spearman's ranked correlation coefficients between the level of pollutants and the basic reproductive number are given in **Table 2**. We observed no significant correlation.

In **Figure 2**, using either univariable or multivariable linear regression, we found no clear effects of NO_2, SO_2, and CO on the initial transmissibility of COVID-19 across Chinese cities. We performed sensitive tests by using spline regression (three degrees of freedom) and restricting the cities outside Hubei province only and found no significant association.

TABLE 2 | Correlation coefficients of COVID-19 R_0s with the three gaseous pollutants across 154 Chinese cities.

	Pearson's correlation		Spearman's ranked correlation	
	Estimate	***p*-value**	**Estimate**	***p*-value**
NO_2	0.01	0.896	0.02	0.866
SO_2	−0.09	0.264	−0.12	0.143
CO	−0.01	0.855	0.04	0.661

NO_2, nitrogen dioxide; SO_2, sulfur dioxide; CO, carbon monoxide.

DISCUSSION

The merit of this work is that we focus on the transmissibility in the initial phase of the pandemics in each city, which may reflect the intrinsic feature of the local outbreak to some extent. However, several limitations should be noted. Population density is a potential confounder since it may be positively related to gaseous pollutant concentrations and may catalyze the spread of COVID-19 (24, 25). Besides observed associations are not of statistical significance between gaseous air pollutants and the COVID-19 transmission, we may still overestimate the true effects because the unadjusted confounder, population density, would positively bias the associations (26). Population flow from Wuhan City may be another potential confounding factor since it may result in disentangled imported or local cases for other Chinese cities. An emerging study found that correlations between gaseous pollutants and R_0 varied in provinces with different population flow (27). However, the effect is likely uniform across cities, and the lockdown of Wuhan mitigated the COVID-19 exportation to other cities effectively. Previously works found the importation was reasonable uniform except for very few cities. This approach has been used previously by other teams and our team (28–30). We avoid using a time series model over a more extended period since it may suffer more issues discussed and demonstrated in a recent study (13). We argue that if there were indeed an "evident" statistical association, our approach should pick it up. At least, we avoid picking up a spurious association, for example, the entangled

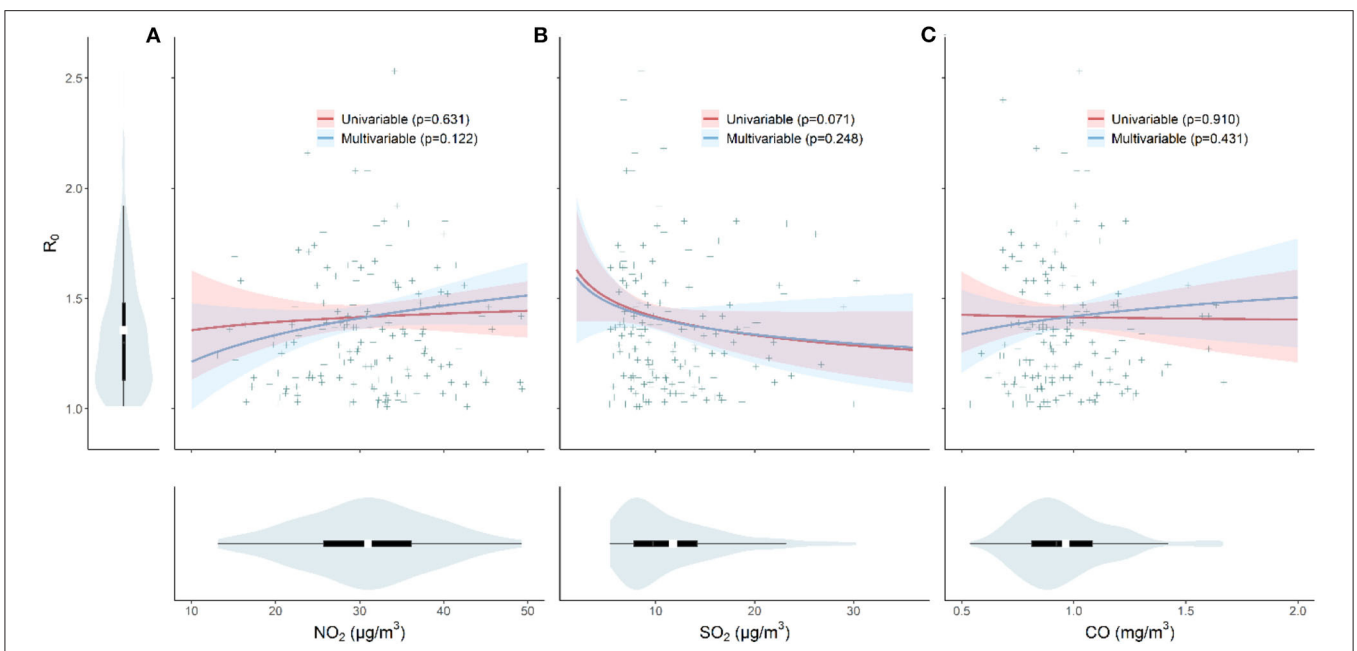


FIGURE 2 | The effects of pollutant **(A)** NO_2, **(B)** SO_2, and **(C)** CO concentrations and the initial transmissibility of COVID-19. Violin plots shows the distributions of R_0, NO_2, SO_2, and CO, respectively. The red line indicates the result from the univariable model, and the blue line represents result from the multivariable model (adjusted for temperature and relative humidity). The *p*-values are given in the legend.

effects between pollution reduction (if any) and city lockdown. Although the results are not of statistical significance, our method is straightforward on one hand, and we nevertheless consider the possibility of the non-linearity of association on the other hand.

In this large-scale ecological study, we find no significant association between the three gaseous air pollutants (NO_2, SO_2, and CO) and the initial transmissibility of COVID-19 in Chinese cities. Since we focus on the initial 16 days after the confirmation of the first case, the lockdown of cities should not have reduced the pollutant level in each city, as normally it will take time for pollutants to drop. Our averaged concentration should be a fair indicator of the level of air quality in each city.

Other factors for the insignificant results include that the outbreak period covered the spring festival when factories were closed. And the closure of factories was extended by the government. Many cities adopted different levels of lockdown, and emission of vehicles were reduced. Thus, the concentration of air pollutants may be low in general. Strict and effective mitigation measures stopped the transmission timely.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JR and SZ conceived the study and carried out the analysis. LH collected the data. JR, SZ, and DH drafted the letter and discussed the results. All authors critically read and revised the letter and gave final approval for publication.

FUNDING

DH was supported by the General Research Fund (Grant Number 15205119) of the Research Grants Council (RGC) of Hong Kong, China, and an Alibaba (China)-Hong Kong Polytechnic University Collaborative Research project. The funding agencies 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; or decision to submit the manuscript for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2020.575839/full#supplementary-material

REFERENCES

1. Silva DR, Viana VP, Müller AM, Livi FP, Dalcin PDTR. Respiratory viral infections and effects of meteorological parameters and air pollution in adults with respiratory symptoms admitted to the emergency room. *Influenza Other Respir Viruses*. (2014) 8:42–52. doi: 10.1111/irv.12158
2. Pirozzi CS, Jones BE, Van Derslice JA, Zhang Y, Paine R, Dean NC. Short-term air pollution and incident pneumonia a case-crossover study. *Ann Am Thorac Soc*. (2018) 15:449–59. doi: 10.1513/AnnalsATS.201706-495OC
3. Neupane B, Jerrett M, Burnett RT, Marrie T, Arain A, Loeb M. Long-term exposure to ambient air pollution and risk of hospitalization with community-acquired pneumonia in older adults. *Am J Respir Crit Care Med*. (2010) 181:47–53. doi: 10.1164/rccm.200901-0160OC
4. Fusco D, Forastiere F, Michelozzi P, Spadea T, Ostro B, Arcà M, et al. Air pollution and hospital admissions for respiratory conditions in Rome, Italy. *Eur Respir J*. (2001) 17:1143–50. doi: 10.1183/09031936.01.00005501
5. Cheng MF, Tsai SS, Wu TN, Chen PS, Yang CY. Air pollution and hospital admissions for pneumonia in a tropical city: Kaohsiung, Taiwan. *J Toxicol Environ Health A*. (2007) 70:2021–6. doi: 10.1080/15287390701601020
6. Li R, Jiang N, liu Q, Huang J, Guo X, Liu F, et al. Impact of air pollutants on outpatient visits for acute respiratory outcomes. *Int J Environ Res Public Health*. (2017) 14:47. doi: 10.3390/ijerph14010047
7. Becker S, Soukup JM. Effect of nitrogen dioxide on respiratory viral infection in airway epithelial cells. *Environ Res*. (1999) 81:159–66. doi: 10.1006/enrs.1999.3963
8. Yao Y, Pan J, Wang W, Liu Z, Kan H, Qiu Y, et al. Association ofparticulate matter pollution and case fatality rate of COVID-19 in 49 Chinese cities. *Sci Total Environ*. (2020) 741:140396. doi: 10.1016/j.scitotenv.2020.140396
9. Zhu Y, Xie J, Huang F, Cao L. Association between short-term exposure to air pollution and COVID-19 infection: evidence from China. *Sci Total Environ*. (2020) 727:138704. doi: 10.1016/j.scitotenv.2020.138704
10. Xu H, Yan C, Fu Q, Xiao K, Yu Y, Han D, et al. Possible environmental effects on the spread of COVID-19 in China. *Sci Total Environ*. (2020) 731:139211. doi: 10.1016/j.scitotenv.2020.139211
11. Zhang Z, Xue T, Jin X. Effects of meteorological conditions and air pollution on COVID-19 transmission: evidence from 219 Chinese cities. *Sci Total Environ*. (2020) 741:140244. doi: 10.1016/j.scitotenv.2020.140244
12. Zhao S. To avoid the noncausal association between environmental factor and COVID-19 when using aggregated data: simulation-based counterexamples for demonstration. *Sci Total Environ*. (2020) 748:141590. doi: 10.1016/j.scitotenv.2020.141590
13. Riccò M, Ranzieri S, Balzarini F, Bragazzi NL, Corradi M. SARS-CoV-2 infection and air pollutants: correlation or causation? *Sci Total Environ*. (2020) 734:139489. doi: 10.1016/j.scitotenv.2020.139489
14. de Silva UC, Warachit J, Waicharoen S, Chittaganpitch M. A preliminary analysis of the epidemiology of influenza A (H1N1) v virus infection in Thailand from early outbreak data, June-July 2009. *Eurosurveillance*. (2009) 14:19292. doi: 10.2807/ese.14.31.19292-en
15. Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. *Int J Infect Dis*. (2020) 92:214–7. doi: 10.1016/j.ijid.2020.01.050
16. Chowell G, Viboud C, Simonsen L, Moghadas SM. Characterizing the reproduction number of epidemics with early subexponential growth dynamics. *J R Soc Interface*. (2016) 13:20160659. doi: 10.1098/rsif.2016.0659
17. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc R Soc B Biol Sci*. (2007) 274:599–604. doi: 10.1098/rspb.2006.3754
18. 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
19. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*. (2020) 368:eabb6936. doi: 10.1101/2020.03.08.20032946
20. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med*. (2020) 382:1199–207. doi: 10.1056/NEJMoa2001316
21. Zhao S, Musa SS, Lin Q, Ran J, Yang G, Wang W, et al. Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the first half of January 2020: a data-driven modelling analysis of the early outbreak. *J Clin Med*. (2020) 9:388. doi: 10.3390/jcm9020388

22. Cowling BJ, Fang VJ, Riley S, Malik Peiris JS, Leung GM. Estimation of the serial interval of influenza. *Epidemiology*. (2009) 20:344–7. doi: 10.1097/EDE.0b013e31819d1092
23. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science*. (2009) 324:1557–61. doi: 10.1126/science.1176062
24. Rocklöv J, Sjödin H. High population densities catalyse the spread of COVID-19. *J Travel Med*. (2020) 27:1–2. doi: 10.1093/jtm/taaa038
25. Lamsal LN, Martin RV, Parrish DD, Krotkov NA. Scaling relationship for NO_2 pollution and urban population size: a satellite perspective. *Environ Sci Technol*. (2013) 47:7855–61. doi: 10.1021/es400744g
26. Mehio-Sibai A, Feinleib M, Sibai TA, Armenian HK. A positive or a negative confounding variable? A simple teaching aid for clinicians and students. *Ann Epidemiol*. (2005) 15:421–3. doi: 10.1016/j.annepidem.2004.10.004
27. Lin S, Wei D, Sun Y, Chen K, Yang L, Liu B, et al. Region-specific air pollutants and meteorological parameters influence COVID-19: a study from mainland China. *Ecotoxicol Environ Saf*. (2020) 204:111035. doi: 10.1016/j.ecoenv.2020.111035
28. Ran J, Zhao S, Han L, Liao G, Wang K, Wang MH, et al. A re-analysis in exploring the association between temperature and COVID-19 transmissibility: an ecological study with 154 Chinese cities. *Eur Respir J*. (2020) 56:2001253. doi: 10.1183/13993003.01253-2020
29. Ran J, Zhao S, Han L, Chen D, Yang Z, Yang L, et al. The ambient ozone and COVID-19 transmissibility in China: a data-driven ecological study of 154 cities. *J Infect*. (2020) 81:11–3. doi: 10.1016/j.jinf.2020.07.011
30. Yao Y, Pan J, Liu Z, Meng X, Wang W, Kan H, et al. No association of COVID-19 transmission with temperature or UV radiation in Chinese cities. *Eur Respir J*. (2020) 55:2000517. doi: 10.1183/13993003.00517-2020

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.

frontiers in Public Health

PERSPECTIVE
published: 20 October 2020
doi: 10.3389/fpubh.2020.581569

Lethal Pneumonia Cases in Mojiang Miners (2012) and the Mineshaft Could Provide Important Clues to the Origin of SARS-CoV-2

Monali C. Rahalkar[1]* *and Rahul A. Bahulikar*[2]

[1] C2, Bioenergy Group, MACS Agharkar Research Institute, Pune, India, [2] BAIF Development Research Foundation, Central Research Station, Pune, India

With the COVID-19 pandemic reaching its worst heights, people are interested in the origin of SARS-CoV-2. This study started with two important questions: first, were there any similar atypical pneumonia outbreaks, even on a smaller level, reported between SARS in 2004 and COVID-19 in 2019/20 in China. Second, examining the beta-coronavirus most closely related to date with SARS-CoV-2 at the genome sequence level, strain RaTG13 (CoV4991), which was sampled from a horseshoe bat in Yunnan province, we asked where exactly did it come from. It was found that RaTG13/CoV4991 was collected from Tongguan mineshaft in Mojiang, Yunnan, China, in 2013. Surprisingly, the same mineshaft was also associated with a severe pneumonia-like illness in miners in 2012 killing three of the six miners. A Master's thesis (in the Chinese language) was found on the cnki.net website which described in detail the severe illness in miners. The thesis concluded that a SARS-like CoV originating from Chinese horseshoe bats (*Rhinolophus*) was the predicted causative agent. The cases were remotely monitored by a prominent pulmonologist in China. Retrospective analysis of the pneumonia cases shows striking similarities with COVID-19. Bilateral pneumonia, vascular complications like pulmonary thromboembolism, and secondary infections are the main similarities. The treatment regimes were similar to the current treatments for COVID-19. We propose that the Mojiang mineshaft miners' illness could provide important clues to the origin of SARS-CoV-2. These cases should be studied by various academicians, researchers, and medical professionals as many important questions are raised in this context.

Keywords: RaTG13, SARS-CoV-2, pneumonia, mineshaft, Mojiang, origin, COVID-19

OPEN ACCESS

Edited by:
Paolo Vineis,
Imperial College London,
United Kingdom

Reviewed by:
Andre Goffinet,
Catholic University of Louvain, Belgium
Jonathan Couey,
University of Pittsburgh, United States

***Correspondence:**
Monali C. Rahalkar
monalirahalkar@aripune.org

Specialty section:
This article was submitted to Infectious Diseases - Surveillance, Prevention and Treatment, a section of the journal Frontiers in Public Health

Received: *09 July 2020*
Accepted: *17 September 2020*
Published: *20 October 2020*

Citation:
Rahalkar MC and Bahulikar RA (2020) Lethal Pneumonia Cases in Mojiang Miners (2012) and the Mineshaft Could Provide Important Clues to the Origin of SARS-CoV-2. Front. Public Health 8:581569. doi: 10.3389/fpubh.2020.581569

ONE LINE SUMMARY

Lethal pneumonia in Mojiang miners and the mine could provide an important link to the research investigating the origin of SARS-CoV-2.

INTRODUCTION

The global COVID-19 pandemic has now affected more than 26 million people with a death toll of 0.8 million affecting 188 countries and territories. Horseshoe bats (*Rhinolphus* sp.) are considered

to be the natural hosts or reservoirs of the Severe Acute Respiratory Syndrome (SARS)-CoV and SARS-CoV-2, the causative agent of the current pandemic COVID-19 (1). The horseshoe bats usually inhabit the Yunnan Province in China's southern sub-tropical zone (2) and Yunnan is also the likely region for the origin of SARS-CoV-2 (3).

Mojiang Mineshaft Associated With Lethal Pneumonia Cases in Miners (2012)

This study scientifically investigated any reports of atypical pneumonia cases covering the period 2004 to 2019, between the SARS (1) epidemic and COVID-19. We identified a case described in two scientific magazines. The first mention was in an interview with Dr. Zhengli Shi, a principal scientist of WIV, in the Scientific American journal (2). In the interview, Shi talked about a mineshaft in Mojiang where a lethal pneumonia-like disease occurred in six miners in 2012 (2). The discussion outlined that a diverse group of coronaviruses was discovered in the mine following the outbreak. These lethal pneumonia cases were also covered in news, including a more detailed article in Science magazine in 2014 (4). In April 2012, a pneumonia-like illness occurred in six miners who were cleaning bat feces from a copper mineshaft in Mojiang, Yunnan, in 2012, killing three of them (4). The Science article describes that a paramyxovirus, MojV was isolated from a rat sample in the same mine (4). Two further papers reported that no direct relationship between human infection and MojV could be established (4–6). In the Scientific American interview, Dr. Zhengli Shi outlines that fungus was responsible for pneumonia in the miners (2). However, no detailed information was elucidated in literature and the cause of the miners' illness remained a mystery.

Master's Thesis by Li Xu on the "Mojiang Miners Pneumonia" Illness

In 2013, Li Xu published a Master's thesis (7) that described in detail the symptoms suffered by six pneumonia patients. This thesis was found by a twitter user (@TheSeeker268), who emailed us the link after reading our pre-print (8) (on May 20, 2020). The thesis was found on the cnki.net website which is the official website for Master's and Ph.D. thesis in China and therefore considered to be a valid source. The original thesis is in Chinese (7) and we translated it using google translation. Currently, a professional translation of the thesis has been made available online by a research agency (https://biosciencesource.org/) that is currently examining the Master's thesis (https://www.documentcloud.org/documents/6981198-Analysis-of-Six-Patients-With-Unknown-Viruses.html).

According to the Master's thesis, in April 2012, six miners were given a job of clearing bat waste and bat feces from a copper mineshaft in Tongguan, Mojiang, Yunnan. After working for ∼14 days in the case of four miners, and 4–5 days in the case of the last two miners, they started facing breathing problems, cough, and fever which required immediate admission to the Kunming hospital in late April and early May (7). Three of the miners died in the course of ∼100 days and three survived (**Table 1A**). The thesis featured medical reports, radiological images such as CT scans, and detailed information regarding the diagnosis and treatment of the miners (7). (https://www.documentcloud.org/documents/6981198-Analysis-of-Six-Patients-With-Unknown-Viruses.html).

Severe Pneumonia and Illness in the Mojiang Miners Related to Horseshoe Bats in the Mojiang Mine (7)

The main clinical symptoms in the six patients from the Mojiang mine were cough and fever, and the main accompanying symptoms were dyspnoea, aching limbs, sputum/bloody sputum, and headache. The details of the course of illness and diagnosis for individual patients are summarized in **Supplementary Information A**. Radiography showed interstitial pneumonia, ground-glass opacities, and severe acute respiratory distress syndrome (ARDS) in the first four patients who also required a mechanical ventilator (patients 2–4). Some patients (1, 2, and 4) showed clotting complications such as pulmonary thromboembolism or thrombosis and elevated D-dimer values. Dr. Zhong Nanshan, a doctor for respiratory diseases and a national advisor for the SARS and COVID-19 epidemic, had provided remote consultation for patients 3 and 4, the most serious patients. Patients 3 and 4 remained in the hospital for more than 100 days. Four patients (1–4) a very low oxygenation index and classified as ARDS (Berlin criteria, 2012). Dr. Nanshan's diagnosis for patients 3 and 4 were interstitial pneumonia (primarily of viral origin), with a possibility of secondary infection (invasive pulmonary aspergillosis). He requested swab testing and SARS antibody testing (to be carried in WIV). He also asked the hospital staff to confirm with the Kunming Institute of Zoology for the type of bat. The

TABLE 1A | Summary of the six pneumonia patients in 2012 [as per (7)].

Number of the patient*	Age	Admitted to the hospital on	Worked in the mine for	Days in the hospital	Outcome/date of discharge/death
1.	63	26.04.2012	14 days	12	**Death** 07.05.2012
2.	42	25.04.2012	14 days	48	**Death** 12.06.2012
3.	45	27.04.2012	14 days	109	**Death** 13.08.2012
4.	46	26.04.2012	14 days	107 (actual days 137)	Improved and discharged on 10.09.2012
5.	30	02.05.2012	5 days	26	Alive, discharged on 28.05.2012
6.	32	26.04.2012	4 days	32	Alive, discharged on 28.05.2012

**Names not given.*

radiological findings were diffuse ground-glass opacities and areas of peripheral consolidation. The thesis concluded that the pneumonia cases were due to viral pneumonia, primarily from SARS-like coronaviruses originating from horseshoe bats. The percentage of lymphocytes, T, B, and NK cells decreased significantly after the admission of the patients, which indicated that the immune system of the patients was seriously damaged by a viral infection. Later, after the consultation of Dr. Zhong Nanshan, (∼after June 19, 2012), blood samples were sent to WIV for antibody testing. A chapter in a Ph.D. thesis by Canping Huang (supervised by Dr. George Gao, present Director China CDCP) also highlights these cases (9) (a translation of Chapter 3 is provided as **Supplementary Material**). According to the translation of the Ph.D. thesis (Lines 283–285, page 9), the "blood test results of four cases showed that: four people carried SARS virus IgG antibodies, of which two were discharged with higher antibody levels (patients 5 and 6) and two which were hospitalized had lower antibody levels (patients 3 and 4) (Wuhan, Chinese Academy of Sciences) Virology Institute)". Xu's Master's thesis, Huang's Ph.D. thesis, and Ge et al. (10), all report the dominance of Chinese horseshoe bats (*Rhinolophus sinicus* and *Rhinolophus affinis*) in the mine. The Kunming Institute of Zoology also confirmed that the six patients were exposed to Chinese horseshoe bats (*Rhinolophus* species). *Rhinolophus* species harbor SARS-like coronaviruses (11). The blood biochemical analysis from the pneumonia patients indicated elevated markers such as Serum Amyloid A (SAA) with a normal range of PCT (procalcitonin), which suggested that the patients had a viral infection. The treatment given to the pneumonia patients included antivirals (ganciclovir, acyclovir injections), steroids (methylprednisolone), antibiotics (meropenem, vancomycin, etc.), antifungals (caspofungin, fluconazole), and anti-thrombotic medicines (warfarin, low molecular weight heparin). The thesis concludes that severe pneumonia in miners was due to SARS-like CoV from horseshoe bats. Dr. Nanshan's conclusion that the Mojiang miners pneumonia appeared to be primarily viral and that it was most probably due to bat-related coronaviruses, is noteworthy.

Mojiang Mine and RaTG13

After the outbreak, WIV conducted longitudinal surveillance of the bat coronaviruses in the Mojiang mine (10). The mineshaft had six bat types of which the highest number of *Rhinolophus* sp. (horseshoe bats) were sampled. Sample collections were done four times between August 2012 and July 2013. A total of 150 alphacoronaviruses and only two betacoronaviruses, of which only one was SARS-like betacoronavirus (CoV/4991), were detected (10). The same virus 4991 was renamed as RaTG13, which is the next genetic relative of SARS-CoV-2 (12).

DISCUSSION

The retrospective analysis of the illness in the miners greatly resembles COVID-19 in the following aspects (**Table 1B**).

TABLE 1B | Common features observed in the six pneumonia patients and COVID-19.

Features	**COVID-19** (13, 14)	**Six pneumonia patients** (7) (master thesis 2013)
Major symptoms		
Fever	✓	✓
Dyspnoea/Fatigue	✓	✓
Cough	✓	✓
Minor symptoms		
Sputum/bloody sputum	/in some cases	✓
headache	(in some)	(in some)
ARDS	✓	✓
Laboratory results		
lymphocytes	decrease	decrease
Serum amyloid A protein, mg/L	High values	High values
D-dimer, mg/L	High value	High value
Radiology		
Chest C. T. scan prominent picture	Ground glass opacities, bilateral pneumonia, peripheral consolidation	Ground glass opacities, bilateral pneumonia, peripheral consolidation
Complications		
Pulmonary thromboembolism	✓	✓
Vascular complications	✓	✓
Hypoxia	✓	✓
Secondary infections (bacterial, fungal)	✓	✓
Role of age	✓	✓
Co-morbidities	✓	✓
Male sex	✓	All were males
Reason of death	Cardiac arrest, ARDS, pulmonary failure	Cardiac arrest, ARDS, pulmonary failure

1. **The radiological picture** seen in the CT scans of COVID-19 patients (13) and miners cases (7) is very similar, which includes ground-glass opacities, peripheral consolidation, and clear indications of bilateral pneumonia (characteristic in COVID-19). This is highly evident on pages 25, 26 and 35, 37 in the translation (https://www.documentcloud.org/documents/6981198-Analysis-of-Six-Patients-With-Unknown-Viruses.html).

2. **Elevated D-dimer values and pulmonary thromboembolism**, a complication seen in COVID-19 were also found in three of the six miners in 2012 (7). The use of heparin, warfarin, and anticlotting drugs was successful in treating the respiratory condition in the fourth miner. Similarly, in COVID-19, pulmonary thromboembolism and blood clotting have been a serious complication.

3. **Lymphocytopenia,** that is, low lymphocyte counts are another common feature characteristic of a viral disease, and common in both the miners' pneumonia cases and COVID-19.

4. **The similarity in treatments:** Treatment given to the miners were antivirals, steroids, mechanical ventilation, antibiotics (for treating the secondary bacterial infections), and antifungals (for treating the secondary fungal infections). Anti-thrombotic agents like warfarin, heparin were also given in the case of patient 4 who successfully recovered. Very similar

treatments are given for treating COVID-19 where an array of antivirals, steroids, blood thinners, antibiotics, and antifungals are given (in conjunction with the secondary infections).

5. Elevated Serum Amyloid A protein: is an inflammatory marker that shows characteristic high values in cases of viral infection. A high SAA value or an increasing trend is an indicator of bad prognosis in the case of COVID-19 (15). In the miners' pneumonia, this marker showed high initial values in the first four serious patients and later showed peaks of up to 1,000–1,200 mg/L in some cases.

(details of the similarities are given as **Supplementary Informations B** and **C**).

Based on the detailed evidence presented in the Master's thesis (7) and the Ph.D. thesis (9) and the discussion presented here, we do not think that fungus was the primary reason for the illness. Dr. Nanshan predicted the miners' illness to be a primary interstitial viral pneumonia (high probability) with invasive aspergillosis as a secondary infection (a condition commonly observed in COVID-19) (16). We think that if it was a fungal disease, only antifungals could have cured the illness. Vascular complications such as elevated D-dimer and thromboembolism are not common in fungal disease and have been observed in the miners' illness and COVID-19 (14). Elevated SAA (serum amyloid A) and declined lymphocytes are indicative of the fact that it was primary viral pneumonia **(Supplementary Information C).**

QUESTIONS

As has been stated, the miners' samples were sent to WIV for SARS Ab testing (7, 9), the same institute that also conducted surveillance of the bat coronaviruses in the Mojiang mineshaft (10). The link between the SARS-like CoV (4991/RaTG13) from mine where lethal pneumonia cases occurred, has not yet been discussed in scientific papers by the WIV laboratory before February 2020. We are curious to know what kind of samples the WIV received from the Mojiang miners, along with other questions, such as whether the samples are still stored in WIV, and whether they are available for study by other researchers. It would also be of particular value to know whether any viruses were isolated and if there is any DNA/RNA available from these samples. It would also be useful to know if PCR was performed on the miners' samples and available sequences. According to Huang's Ph.D. thesis, four miners tested positive in an Ab test against SARS-like CoV (**Supplementary Material**). However, further questions remain as to which antigen was used for the Ab detection in the pneumonia patients and what was the exact protocol used. Why is this information not available in any of the seroprevalence studies by WIV? Why were the severe pneumonia cases in 2012 not mentioned in any of the WIV publications before 2020? Were any SARS-like CoV isolated from the bat fecal samples collected in 2012–13? Why were the Mojiang miners pneumonia cases in 2012 not reported to any public health agency like the WHO? Why did programs like PREDICT not mention the lethal pneumonia cases as a mini-outbreak? Was the mineshaft in Mojiang closed, when? According to the literature, three research groups went to the Mojiang mine to collect samples between 2012 and October 2014 (5, 9, 10). The mine was promptly closed as per the (2). Why was the Mojiang mine being visited by researchers until October 2014? Questions also remain as to why Dr. Shi attributed the outbreak in Mojiang to a fungus in the interview with Scientific American. Was the mine open for researchers and were any samples brought after 2014? Did any of the researchers who visited the Mojiang mineshaft get infected by any coronavirus between 2012 and 2019? Are there any whole genome sequences available for SARS-like CoV originating from this mine? Why is the pathogen database (http://www.viruses.nsdc.cn/chinavpi/) associated with the project (2013FY113500) (10) not accessible anymore?

CONCLUSIONS

The striking similarities between the Mojiang pneumonia cases and COVID-19 are noteworthy, as is the fact that RaTG13/CoV4991, the next genomic relative of SARS-CoV-2 was found in the same mineshaft. The Master's thesis by Li Xu concludes that the pneumonia-illness in the miners was due to a SARS-like CoV from horseshoe bats. The remote consultation and diagnosis by a prominent pulmonologist in China, Dr. Nanshan, adds credibility to the diagnosis of the pneumonia cases in 2012. Although we cannot say that RaTG13 or SARS-CoV-2 infected the miners, there is a high chance that it could be a virus quite similar in genetic composition to these two. The coincidence between the 2012 illness in Mojiang miners, the subsequent samplings, and finding the nearest SARS-CoV-2 relative from this single mine warrants further inquiry, and the data along with the full history of this incident would be invaluable in the context of the current pandemic.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

MR: conceptualization. MR and RB: Writing of the paper. Both authors reviewed and approved the final version.

ACKNOWLEDGMENTS

We thank Dr. Luigi Warren, CA, USA, for important scientific discussions (@luigi_warren). We would also like to thank @TheSeeker268 for finding and sharing links for the Master's thesis by Li Xu (2013) and the Ph.D. thesis by Canping Huang (2016), which were both in Chinese. We are also very thankful to the Twitter group DRASTIC (Decentralized Radical Autonomous Search Team Investigating COVID-19) for invaluable discussions, which contributed to writing this paper, in particular, @BillyBostickson, @franciscodeasis @AntGDuarte @Real_Adam_B, @DrAntoniSerraT1,

@KevinMcH3, @internetperson1, and @Ersa Flavikins. We thank Dr. Rossanna Segreto for important suggestions and Dr. Jonathan Latham of bioscienceresource.org for arranging the translation of the Master's thesis and sharing the link to the professional translation. We also thank *@franciscodeasis* who translated relevant chapters in the Ph.D. thesis. In particular, would like to thank Professor Anand Rahalkar, MD radiology, Bharti Vidyapeeth Medical College, Pune, India, for reviewing the X-rays and CT scans from the Master's thesis by Li Xu (2013). We would like to specifically thank the reviewers and the editors for their valuable comments. An earlier version of the paper has been published as a pre-print on May 24, 2020 at [https://www.preprints.org/manuscript/202005.0322/v2], Rahalkar and Bahulikar (8).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2020.581569/full#supplementary-material

REFERENCES

1. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* (2020) 579:270–3. doi: 10.1038/s41586-020-2012-7
2. Qiu J. How China's "bat woman" hunted down viruses from sars to the new coronavirus. *Sci Am.* (2020) 322:24–32. doi: 10.1038/scientificamerican0620-24
3. Latinne A, Hu B, Olival KJ, Zhu G, Zhang L, Li H, et al. Origin and cross-species transmission of bat coronaviruses in China. *Nat Commun.* (2020) 11:4235. doi: 10.1038/s41467-020-17687-3
4. Stone R. A new killer virus in China? *Science.* (2014). Available online at: https://www.sciencemag.org/news/2014/03/new-killer-virus-china
5. Wu Z, Yang L, Yang F, Ren X, Jiang J, Dong J, et al. Novel henipa-like virus, mojiang paramyxovirus, in rats, China, 2012. *Emerging Infect Dis.* (2014) 20:1064–6. doi: 10.3201/eid2006.131022
6. Rissanen I, Ahmed AA, Azarm K, Beaty S, Hong P, Nambulli S, et al. Idiosyncratic Mojiang virus attachment glycoprotein directs a host-cell entry pathway distinct from genetically related henipaviruses. *Nat Commun.* (2017) 8:16060. doi: 10.1038/ncomms16060
7. Xu L. *The analysis of 6 patients with severe pneumonia caused by unknown viruses* (Master's Thesis). The First Clinical Medical College of Kunming Medical University, Kunming Medical University, Kunming, China (2013). http://eng.oversea.cnki.net/Kcms/detail/detail.aspx?filename=1013327523.nh&dbcode=CMFD&dbname=CMFD2014
8. Rahalkar MC, Bahulikar RA. Understanding the origin of 'BatCoVRaTG13', a virus closest to SARS-CoV-2. *Preprints.* (2020) 2020050322. doi: 10.20944/preprints202005.0322.v1
9. Huang C. *Novel virus discovery in bat and the exploration of receptor of bat coronavirus HKU9* (PhD Thesis). National Institute for Viral Disease Control and Prevention, Beijing: China (2016). Available online at: http://eng.oversea.cnki.net/kcms/detail/detail.aspx?dbcode=CDFD&QueryID=11&CurRec=1&dbname=CDFDLAST2018&filename=1017118517.nh&UID=WEEvREcwSlJHSldTTEYzWEpEZktmRXB3Sm9JeHRKZExVOG5ySkJjK0xHMD0%3d%249A4hF_YAuvQ5obgVAqNKPCYcEjKensW4IQMovwHtwkF4VYPoHbKxJw!!&autoLogin=0
10. Ge XY, Wang N, Zhang W, Hu B, Li B, Zhang YZ, et al. Coexistence of multiple coronaviruses in several bat colonies in an abandoned mineshaft. *Virol Sin.* (2016) 31:31–40. doi: 10.1007/s12250-016-3713-9
11. Yua P, Hua B, Zhengli S, Cuia J. Geographical structure of bat SARS-related coronaviruses. *Infect Genet Evol.* (2019) 2019:224–9. doi: 10.1016/j.meegid.2019.02.001
12. Cohen J. Wuhan coronavirus hunter Shi Zhengli speaks out. *Science.* (2020) 369:487–8. doi: 10.1126/science.369.6503.487
13. Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *Lancet.* (2020) 20:425–34. doi: 10.1016/S1473-3099(20)30086-4
14. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5
15. Li H, Xiang X, Ren H, Xu L, Zhao L, Chen X, et al. Serum amyloid A is a biomarker of severe coronavirus disease and poor prognosis. *J Infect.* (2020) 80:646–55. doi: 10.1016/j.jinf.2020.03.035
16. Johns S. *COVID-19 Patients Could be at Greater Risk of Fungal Infections, Researchers Say*. London: Imperical College London (2020).

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.

frontiers in Medicine

PERSPECTIVE
published: 08 December 2020
doi: 10.3389/fmed.2020.594623

Some Questions to Our Chinese Colleagues Pioneering Research Into Coronavirus Disease 2019 (COVID-19)

Claudia Stöllberger[1]* *and Maria Winkler-Dworak*[2]

[1] Medizinische Abteilung, Klinik Landstrasse, Vienna, Austria, [2] Wittgenstein Centre for Demography and Global Human Capital, Vienna Institute of Demography of the Austrian Academy of Sciences, Vienna, Austria

A pandemic has developed, so physicians worldwide are particularly interested in the experiences of their Chinese Colleagues which are frequently cited. To assess the long-term pulmonary, cardiac, neurologic, and psychiatric consequences after COVID-19, the outcome of patients included in the early publications and the association with baseline findings is of particular interest. Thus, we review the methods of early Coronavirus disease 2019 (COVID-19) publications. Reports published before March 19th 2020, comprising >40 patients were included, considering especially cardiologic aspects. It remains unclear whether patients were described several times, or they were different patients. Only patients with confirmed COVID-19 were described, and no differences in findings of patients with initially suspected and later confirmed, or excluded infection. It remains unclear in how many cases information was missing, since missing values were not reported. Medication before hospital admission, level of education and occupation, household size and composition, weight or body mass index are lacking. No details about electrocardiographic findings are given. Patients still in follow-up, constituting the major part of observations, were excluded. The data should be re-analyzed. A comparison between confirmed and excluded cases could be carried out. By now, in November 2020, the reported patients will most probably have recovered. Thus, it would be possible to differentiate prognostic indicators more precisely. Laboratory tests and electrocardiograms could be analyzed in more detail to shed light on the spectrum of this disease and to solve some of the unanswered questions related with COVID-19.

Keywords: statistics, infection, electrocardiogram, mortality, cardiology

OPEN ACCESS

Edited by:
Matteo Convertino, Hokkaido University, Japan

Reviewed by:
Ying Luo, Tongji Hospital, China
Kuldeep Dhama, Indian Veterinary Research Institute (IVRI), India

***Correspondence:**
Claudia Stöllberger
claudia.stoellberger@chello.at

Specialty section:
This article was submitted to Infectious Diseases – Surveillance, Prevention and Treatment, a section of the journal Frontiers in Medicine

Received: *13 August 2020*
Accepted: *19 November 2020*
Published: *08 December 2020*

Citation:
Stöllberger C and Winkler-Dworak M (2020) Some Questions to Our Chinese Colleagues Pioneering Research Into Coronavirus Disease 2019 (COVID-19). Front. Med. 7:594623. doi: 10.3389/fmed.2020.594623

BACKGROUND

At the beginning of 2020, symptoms, laboratory findings, complications, and outcome of patients with Coronavirus disease 2019 (COVID-19) were reported from China. Since then, a pandemic has developed, and accordingly, a great number of articles have been published. On November 18th 2020, the search term "COVID-19" yielded 75,182 results in PubMed. Since the first COVID-19 infections were reported from Wuhan in China, physicians worldwide are particularly interested in the experiences of their colleagues and the early publications are frequently cited (**Table 1**) (1–6).

TABLE 1 | Studies from China about COVID-19 published until March 19th, 2020.

Authors	Hospital/Place	Sample size	Design	Statistical analysis	Methodological concerns	Missing information	PubMed citations on November 18th 2020
(1)	Jin Yin-tan Hospital (Wuhan, China)	41	Group comparisons between ICU ($n = 13$) and non ICU patients ($n = 28$) admitted to hospital between December 16, 2019 to Jan 2, 2020.	Mann Whitney *U*-test, chi2 test, Fisher's exact	Definition of comorbidities are missing, no reference values for laboratory tests.	Weight, BMI, ECG, CK levels, BNP levels, medication at admission, urine tests. N/A information is missing.	7,023
(2)	Jin Yin-tan and Tongji hospital (Wuhan, China)	150	Group comparison between discharged ($n = 82$) and deceased ($n = 68$) patients.	*t*-test, Mann Whitney–Wilcoxon test, chi2 test, Fisher's exact test	Definition of comorbidities are lacking. Recruitment of patients and time period of enrollment is unclear. No survival analysis. The term "myocarditis" is used without definition.	Weight, BMI, ECG, BNP levels, medication at admission, urine tests. N/A information is missing.	850
(3)	Jin Yin-tan hospital (Wuhan, China)	52	Group comparisons of deaths ($n = 32$) and survivors ($n = 20$) of critically ill adult patients with pneumonia admitted from December 24, 2019 to January 12, 2020; final date of follow-up February 9, 2020.	*t*-test, Wilcoxon rank sum test, chi2 test, Fisher's exact test, Kaplan–Meier curves	Definition of comorbidities are missing, no reference values for laboratory tests. Cardiac injury is only defined by troponin elevation.	Weight, BMI, ECG, BNP levels, medication at admission, urine tests. N/A information is missing.	1,850
(4)	Zhongnan (Wuhan, China)	138	Group comparisons of ICU ($n = 36$) and non ICU ($n = 132$) patients with pneumonia admitted from January 1 to January 28, 2020; final date of follow-up February 3, 2020.	*t*-test, Mann Whitney–Wilcoxon test, generalized linear mixed model, chi2 test, Fisher's exact test	Definition of comorbidities are missing. "Arrhythmia" is not defined.	Weight, BMI, ECG, BNP levels, medication at admission, urine tests. N/A information is missing.	4,179
(5)	552 hospitals including 132 patients from Jin Yin-tan Hospital (Wuhan, China)	1,099	Group comparison of non-severe ($n = 926$) and severe patients ($n = 173$) and of those experiencing a primary endpoint event ($n = 67$) and not ($n = 1,032$) out of 7,735 COVID-19 patients admitted from December 11, 2019 to January 29, 2020. Composite end points: ICU, mechanical ventilation or death. Patients censored with no outcome by January 31, 2020.	Descriptive, median and interquartile range, count, and percentage	Definition of comorbidities are missing, no reference values for laboratory tests.	Weight, BMI, ECG, medication at admission, urine tests, cardiac problems. N/A information is partially missing.	4,282
(6)	Jin Yin-tan Hospital and Wuhan Pulmonary Hospital (Wuhan, China)	191	Group comparisons between survivors ($n = 137$) and non-survivors ($n = 54$) out of 813 adult COVID-19 patients, admitted from December 29, 2019 to January 31, 2020, who died or were discharged by January 31, 2020.	Mann–Whitney *U*-test, χ^2 test, or Fisher's exact, logistic regression	Definition of comorbidities are missing. Censored patients not included.	Weight, BMI, ECG, medication at admission, urine tests, cardiac problems. N/A information is missing.	4,112

ICU, Intensive care unit; BMI, Body mass index; ECG, electrocardiogram; BNP, brain natriuretic pepetide; N/A, not available.

Follow-up data about the consequences of COVID-19 infection on the health of affected patients are not yet available. There are indications for long-term pulmonary, cardiac, neurologic, and psychiatric illnesses and disabilities after COVID-19 (7). In that respect, again, the colleagues in Wuhan could be pioneers in reporting the long-term outcome of patients included in their early publications and relate them with clinical and instrumental findings at baseline. These baseline-data, however, were obtained under pressure of an evolving epidemic, which is also reflected by a high rate of published errata (1–3, 6). There are concerns about the methodological quality of the data (8).

Thus, the present article endeavors to review methods and results of early COVID-19 publications. We have included articles from Wuhan, China, published before March 19th 2020, which provided clinical details about >40 patients in English language (1–6). Not considered were case reports, epidemiologic, virologic, or radiologic studies without clinical data, publications with <40 patients and publications from which it was evident that they are duplicate publications. Since one of the authors is a cardiologist, the cardiological aspects were especially considered.

MAIN BODY

Several Publications From the Same Hospital

Reviewing the articles, it became obvious that four of them originate from the same institution (1–3, 6). It remains unclear whether patients were described several times, or whether they were different patients.

Suspected Covid-19 Cases

The articles described only patients with confirmed COVID-19. In clinical routine, however, it is important to differentiate between COVID-19 and other diseases. It would be very useful to know differences in symptoms and laboratory findings in patients with initially suspected and later confirmed, or excluded, COVID-19 infection.

This issue is mentioned in one article: "The absence of fever in COVID-19 is more frequent than in Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Middle East Respiratory Syndrome Coronavirus (MERS-CoV) infection, so afebrile patients may be missed if the surveillance case definition focuses on fever detection" (5). In a further article, the consequences of atypical presentations are described: "A total of 14 patients (10%) initially presented with diarrhea and nausea 1 to 2 days prior to development of fever and dyspnea.… One patient in the current study presented with abdominal symptoms and was admitted to the surgical department. More than 10 health care workers in this department were presumed to have been infected by this patient…" (4).

Missing Values

The data in all included studies were collected retrospectively. It remains unclear, however, in how many cases information was missing, since just one publication reported missing values in some, not all, tables (5).

Other Symptoms

Ageusia or anosmia were not reported in the included articles, but only in more recent ones from China (9) and Italy (10, 11). Why were these not reported in the included articles? One answer could be that symptoms were assessed in accordance with the International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) case record form, as indicated in the methods section of the articles. In this form, there is no space for "other symptoms" (12).

Missing Baseline Information

Data about medication before hospital admission are lacking, thus no information can be obtained whether specific drugs render the patients more prone to a severe course of the disease. This issue, relating angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, and statins has been extensively discussed in the meantime (13, 14).

Data about level of education and occupation, household size and composition are not reported. Neither weight nor body mass index of the patients are reported. The reason might be the above-mentioned ISARIC case record form, as it does not require such information. Among the comorbidities, the terms "obesity" and "malnutrition" "as defined by the clinical staff" are registered in the case record form, but not the body mass index as an objective measure of obesity. Whether obesity influences the prognosis in COVID-19 is a still controversially assessed issue, and it is uncertain whether ethnic and socioeconomic differences might play a role for the discrepant findings (15–17).

We miss urine analysis data, which might help to characterize the type and pathomechanism of renal failure in COVID-19 (18). Is the frequently encountered hypalbuminemia due to renal loss of protein? Again, this lack of data may be related to the ISARIC case record form, where no such data are requested.

Lack of Information About the Heart

"Cardiac injury" is frequently reported in COVID-19 patients. The definition for "cardiac injury" was adopted from a publication in which cardiac biomarkers, electrocardiography (ECG), or echocardiography among hospitalized patients infected with influenza A (H7N9) virus had been reported (19). "Myocarditis" is reported in one publication, however, it remains unclear how it was diagnosed (2). In another publication, arrhythmia is mentioned, but no description is given (4). In yet another publication, "cardiac injury" is only defined by elevations of the troponin (3).

No details about ECG findings are given in any of the included articles. Whereas, echocardiography should only be performed if it is expected to provide clinical benefit (and not as screening investigation), because of close proximity of patient and echocardiographer; this is not the case with ECG (20). An ECG is easy to perform and yields useful information about the cardiac condition, especially in patients with preexisting cardiovascular diseases. It would be of great interest to know how often and which types of arrhythmias were registered, the prevalence and dynamics of ST-elevations or depressions, and their association with cardiac biomarkers like troponin or natriuretic peptides. Up to now, our knowledge about ECG

findings in COVID-19 patients from China is limited to small series comprising 63 (21), 102 (22), and 112 patients (23).

Statistical Issues

The negligence of relevant variables can have important consequences. The variables may be confounders and the omission of these confounders will produce biased results. Moreover, the consequences might be enduring in future analyses, as model-building is often based (and even advised in some textbooks) on results of prior studies (24, 25). The selection of patients is unclear or biased (4, 5). In addition, patients still in follow-up, constituting the major part of observations, were excluded from the sample in two studies (3, 6). This phenomenon, that patients are still hospitalized when the article is written, is not only encountered in publications from China but also from Europe and the U.S.A (17, 26).

CONCLUSIONS

It remains uncertain if such missing clinical information on COVID-19 patients was due to a rush to publish as rapid due to competition or due to underreporting. The retrospectively collected data, obtained under pressure of an evolving epidemic, will not have the quality of prospectively collected information. Nevertheless, it would be worthwhile to re-analyse and re-collect data of the patients included in these early studies. By doing this, a comparison of symptoms and findings between initially suspected and eventually confirmed and excluded COVID-19 cases could possibly be carried out. This would be helpful for physicians in emergency rooms worldwide. Moreover, by now, in November 2020, the reported patients will most probably have experienced an outcome, i.e., either recovered and are discharged from hospital or died from COVID-19. Thus, it would be possible differentiate indicators for survival or death more precisely with the complete observations in the full samples. Furthermore, laboratory tests, ECG recordings and—if available–results of imaging studies could be analyzed in more detail to shed light on the spectrum of this disease. Furthermore, it is of high interest to know the long-term consequences of patients who have suffered from COVID-19. The Chinese collegues could be pioneers in assessing these questions by carrying out follow-up investigations regarding pulmonary, cardiac, endocrinologic, neurologic, psychiatric, and immunologic impairments of post-COVID-19 patients (27).

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 author/s.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5
2. Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Med.* (2020) 46:846–8. doi: 10.1007/s00134-020-05991-x
3. Yang X, Yu Y, Xu J, Shu H, Xia J, Liu H, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med.* (2020) 8:475–81. doi: 10.1016/S2213-2600(20)30079-5
4. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA*. (2020) 323:1061–9. doi: 10.1001/jama.2020.1585
5. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. (2020) 382:1708–20. doi: 10.1101/2020.02.06.20020974
6. Zhou F, Yu T, Du R, Fan G, Liu Y, Lu 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
7. Marshall M. The lasting misery of coronavirus long-haulers. *Nature.* (2020) 585:339–41. doi: 10.1038/d41586-020-02598-6
8. London AJ, Kimmelman J. Against pandemic research exceptionalism. *Science.* (2020) 368:476–7. doi: 10.1126/science.abc1731
9. Mao L, Jin H, Wang M, Hu Y, Chen S, He Q, et al. Neurologic manifestations of hospitalized patients with coronavirus disease 2019 in Wuhan, China. *JAMA Neurol.* (2020) 77:683–90. doi: 10.1001/jamaneurol.2020.1127
10. Giacomelli A, Pezzati L, Conti F, Bernacchia D, Siano M, Oreni L, et al. Self-reported olfactory and taste disorders in patients with severe acute respiratory coronavirus infection: a cross-sectional study. *Clin Infect Dis.* (2020) 71:889–90. doi: 10.1093/cid/ciaa330
11. Spinato G, Fabbris C, Polesel J, Cazzador D, Borsetto D, Hopkings C, et al. Alterations in smell or taste in mildly symptomatic outpatients with SARS-CoV-2 infection. *JAMA.* (2020) 323:2089–90. doi: 10.1001/jama.2020.6771
12. ISARIC. *International Severe Acute Respiratory and Emerging Infection Consortium.* (2020). Available online at: https://isaric.tghn.org/COVID-19-CRF/ (accesed April 26, 2020).
13. Dambha-Miller H, Albasri A, Hodgson S, Wilcox CR, Khan S, Islam N, et al. Currently prescribed drugs in the UK that could upregulate or downregulate ACE2 in COVID-19 disease: a systematic review. *BMJ Open.* (2020) 10:e040644. doi: 10.1136/bmjopen-2020-040644
14. Daniels LB, Sitapati AM, Zhang J, Zou J, Bui QM, Ren J, et al. Relation of statin use prior to admission to severity and recovery among COVID-19 inpatients. *Am J Cardiol.* (2020) 136:149–55. doi: 10.1016/j.amjcard.2020.09.012
15. Rottoli M, Bernante P, Belvedere A, Balsamo F, Garelli S, Giannella M, et al. How important is obesity as a risk factor for respiratory failure, intensive care admission and death in hospitalised COVID-19 patients? Results from a single Italian centre. *Eur J Endocrinol.* (2020) 183:389–97. doi: 10.1530/EJE-20-0541
16. Arbel Y, Fialkoff C, Kerner A, Kerner M. Can reduction in infection and mortality rates from coronavirus be explained by an obesity survival paradox? An analysis at the US statewide level. (2020) *Int J Obes*. (2020) 15:1–4. doi: 10.1038/s41366-020-00680-7
17. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York city: prospective cohort study. *BMJ.* (2020) 369:m1966. doi: 10.1136/bmj.m1966

18. Zheng X, Yang H, Li X, Li H, Xu L, Yu Q, et al. Prevalence of kidney injury and associations with critical illness and death in patients with COVID-19. *Clin J Am Soc Nephrol.* (2020) 15:1549–53. doi: 10.2215/CJN.04780420
19. Gao C, Wang Y, Gu X, Shen X, Zhou D, Zhou S, et al. Community-acquired pneumonia–china network, association between cardiac injury and mortality in hospitalized patients infected with avian influenza A (H7N9) virus. *Crit Care Med.* (2020) 48:451–8. doi: 10.1097/CCM.0000000000004207
20. Kirkpatrick JN, Mitchell C, Taub C, Kort S, Hung J, Swaminathan M. ASE statement on protection of patients and echocardiography service providers during the 2019 novel coronavirus outbreak. *J Am Coll Cardiol.* (2020) 75:3078–84. doi: 10.1016/j.jacc.2020.04.002
21. Chen L, Feng Y, Tang J, Hu W, Zhao P, Guo X, et al. Surface electrocardiographic characteristics in coronavirus disease 2019: repolarization abnormalities associated with cardiac involvement. *ESC Heart Fail.* (2020) doi: 10.1002/ehf2.12991
22. Xu H, Hou K, Xu R, Li Z, Fu H, Wen L, et al. Clinical characteristics and risk factors of cardiac involvement in COVID-19. *J Am Heart Assoc.* (2020) 9:e016807. doi: 10.1161/JAHA.120.016807
23. Deng Q, Hu B, Zhang Y, Wang H, Zhou X, Hu W, et al. Suspected myocardial injury in patients with COVID-19: evidence from front-line clinical observation in Wuhan, China. *Int J Cardiol.* (2020) 311:116–21. doi: 10.1016/j.ijcard.2020.03.087
24. Kleinbaum DG, Klein D. *Logistic Regression. A Self-Learning Text*. 2nd Edn. New York, NY: Springer (2002).
25. Harrel FE Jr. *Regression Modeling Strategies. With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. 2nd ed. Heidelberg: Springer (2015).
26. Haase N, Plovsing R, Christensen S, Poulsen LM, Brøchner AC, Rasmussen BS, et al. Characteristics, interventions and longer-term outcomes of COVID-19 ICU patients in Denmark - a nationwide, observational study. *Acta Anaesthesiol Scand.* (2020). doi: 10.1111/aas.13701
27. Hayden MR. An immediate and long-term complication of COVID-19 may be type 2 diabetes mellitus: the central role of β-cell dysfunction, apoptosis and exploration of possible mechanisms. *Cells.* (2020) 9:E2475. doi: 10.3390/cells9112475

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.

frontiers in Public Health

REVIEW
published: 21 December 2020
doi: 10.3389/fpubh.2020.614113

Economic and Behavioral Influencers of Vaccination and Antimicrobial Use

Caroline E. Wagner[1†], Joseph A. Prentice[2†], Chadi M. Saad-Roy[2,3†], Luojun Yang[2†], Bryan T. Grenfell[2,4,5], Simon A. Levin[2,6] and Ramanan Laxminarayan[6,7]*

[1] Department of Bioengineering, McGill University, Montreal, QC, Canada, [2] Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, United States, [3] Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, United States, [4] Princeton School of Public and International Affairs, Princeton University, Princeton, NJ, United States, [5] Fogarty International Center, National Institutes of Health, Bethesda, MD, United States, [6] Princeton Environmental Institute, Princeton University, Princeton, NJ, United States, [7] Center for Disease Dynamics, Economics & Policy, Washington, DC, United States

Despite vast improvements in global vaccination coverage during the last decade, there is a growing trend in vaccine hesitancy and/or refusal globally. This has implications for the acceptance and coverage of a potential vaccine against COVID-19. In the United States, the number of children exempt from vaccination for "philosophical belief-based" non-medical reasons increased in 12 of the 18 states that allowed this policy from 2009 to 2017 (1). Meanwhile, the overuse and misuse of antibiotics, especially in young children, have led to increasing rates of drug resistance that threaten our ability to treat infectious diseases. Vaccine hesitancy and antibiotic overuse exist side-by-side in the same population of young children, and it is unclear why one modality (antibiotics) is universally seen as safe and effective, while the other (vaccines) is seen as potentially hazardous by some. In this review, we consider the drivers shaping the use of vaccines and antibiotics in the context of three factors: individual incentives, risk perceptions, and social norms and group dynamics. We illustrate how these factors contribute to the societal and individual costs of vaccine underuse and antimicrobial overuse. Ultimately, we seek to understand these factors that are at the nexus of infectious disease epidemiology and social science to inform policy-making.

Keywords: COVID-19, vaccination, antimicrobial, behavior, hesitancy

OPEN ACCESS

Edited by:
Matteo Convertino, Hokkaido University, Japan

Reviewed by:
Marisa Silvia Castro, Institute of Studies on Humoral Immunity (IDEHU), Argentina
Feng Fu, Dartmouth College, United States

***Correspondence:**
Ramanan Laxminarayan
ramanan@cddep.org

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Infectious Diseases - Surveillance, Prevention and Treatment, a section of the journal Frontiers in Public Health

Received: *15 October 2020*
Accepted: *01 December 2020*
Published: *21 December 2020*

Citation:
Wagner CE, Prentice JA, Saad-Roy CM, Yang L, Grenfell BT, Levin SA and Laxminarayan R (2020) Economic and Behavioral Influencers of Vaccination and Antimicrobial Use. Front. Public Health 8:614113. doi: 10.3389/fpubh.2020.614113

INTRODUCTION

Vaccines are among the most cost-effective health technologies of all time. They have been responsible for the two instances, smallpox and rinderpest, in which an infectious disease has been eradicated (2). By choosing to be vaccinated, an individual protects themself but also protects their community by preventing disease transmission. Although immunizing enough individuals in a community above a critical proportion can help prevent outbreaks, actual vaccination levels tend to fall short of epidemiological goals due to vaccine hesitancy and refusal. Vaccine hesitancy is as old as vaccines themselves, but has gained momentum in recent years due to a growing distrust in science and institutions. One recent impetus for this was the subsequently retracted and discredited 1998 study by Wakefield and coauthors which falsely claimed a link between the measles-mumps-rubella (MMR) vaccine and autism in children (3). Vaccine hesitancy runs the range from doubts about a specific vaccine to a complete rejection of all forms of immunization. It is relevant not just to childhood immunizations but also to adult vaccines including those

being developed against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes coronavirus disease 2019 (COVID-19).

According to the United States Centers for Disease Control and Prevention (CDC), national vaccination coverage has remained constant in recent years: 91.9–91.5% for MMR, 91.2–91.0% for varicella, 94.1–94.0% for diphtheria, tetanus, and pertussis (DTaP; ≥3 doses), and 72.6–73.2% for rotavirus from 2013 to 2017 (4), but vaccine coverage in certain states and communities has declined. For example, among children aged 5 months, up-to-date status for recommended vaccines declined from 67.9% in 2019 to 49.7% in 2020 in Michigan (5).

Another important method of disease control is through antimicrobial treatment. However, the effectiveness of antimicrobials has decreased in recent years due to the emergence of antimicrobial-resistant strains. Indeed, resistance genes, such as Enterobacteriaceae-producing extended-spectrum β-lactamase (ESBL), NDM-1, and *Klebsiella pneumoniae* carbapenemase (KPC), are widespread and represent major burdens to public health (6). The emergence of antimicrobial-resistant strains has been driven largely by overconsumption and misuse of antimicrobials, which are additionally associated with altered microbiome communities (7), obesity, and irritable bowel syndrome (8).

Despite the contextual complexity of health behaviors, a recent meta-analysis (9) based on conditioned risk questions found that people are more likely to accept vaccination when they perceive a high risk of contracting the disease when unvaccinated (12 studies, effect size 0.26), greater personal vulnerability to the disease (five studies, effect size 0.24), and greater severity of the disease (31 studies, effect size 0.16). Vaccine refusal arises from underestimated risk of disease or overestimated risk of vaccine-induced adverse effects. Risk (mis)perceptions also contribute to overuse of antibiotics. While a lack of awareness about antimicrobial resistance is associated with high rates of antibiotic use among self-medicated individuals (10), clinicians' misperceptions of antibiotic harmlessness are also associated with higher antibiotic prescribing rates in Emergency Departments (11). Studies such as these raise important questions that we must understand to better tackle both vaccine underuse and antibiotic overuse.

What makes people comfortable with the idea of using antibiotics, while being concerned about vaccination, even when the target population tends to be small children? How are these decisions influenced by perceptions of the benefits of antibiotic treatment or immunization, and perceptions of side-effects associated with these interventions? Why do some people perceive vaccines to be unsafe but think that antibiotics are safe? Are individuals likely to take into consideration the benefits (of vaccination) or costs (of antibiotic resistance) that they create for others as a consequence of their actions? And how influenced are they by social norms or peer groups in their behavior? These concerns have increased relevance in the context of COVID-19, where a potential vaccine or set of vaccines are likely to form part of the long-term strategy to keep the disease in check (12).

In this review, we examine factors that shape vaccine hesitancy and antimicrobial overconsumption and characterize the risk and cost they exert upon individuals and societies. First, in Section 2: Risk Perception, we examine the role of risk perception. In Section 3: Free-Riding and Individual Incentives, we look at the issue of individual incentives and external consequences. In Section 4: Social Norms and Group Dynamics, we examine the types of norms and community histories that govern vaccine- and antibiotic-related health behaviors. In Section 5: Actual Risks, Costs, and Benefits of Vaccine Underuse and Antibiotic Overuse, we estimate the risks and costs to individuals and societies associated with vaccination, vaccine hesitancy, antimicrobial use and antimicrobial resistance. In Section 6: Policy Interventions, we propose a series of policy interventions in an effort to curb vaccine hesitancy and antimicrobial overuse and conclude with Section 7: Conclusion.

SECTION 2: RISK PERCEPTION

One reason why people underuse vaccines and overuse antibiotics is that their perception of risk differs from what the evidence may suggest. In the United States, the perceived risk of vaccine use in the general population is several orders of magnitude greater than the actual probability of adverse vaccine-associated events (13). Likewise, across several countries, even when antibiotic use is unlikely to have a significant benefit for infection prognosis, a large proportion of patients have been shown to desire antibiotic prescription (14, 15).

Most health-related decisions are made under uncertainty. Game-theoretic models of vaccination behavior predict that individuals will free-ride on the herd protection afforded by others' vaccination status, particularly when the risk associated with vaccination outweighs the risk of infection (16). In fact, vaccine refusal can emerge even when rational assessment favors vaccination, due to bounded rationality of individual decision-makers (17, 18). With imperfect information and limited processing capacity, individuals' perceived risks may deviate from the actual risks in positive or negative directions. This idea also applies to the overuse of antibiotics for viral infections or other situations in which their prescription is inappropriate due to the general misperception that they are harmless. According to prospect theory (19), such misperceptions could arise from a complex combination of cognitive biases such as framing effects, loss aversion, and diminishing sensitivity. Misperception of risks may also explain why individual decisions often deviate from predictions of expected utility theory, or the classic risk-benefit model of health behavior.

A paradox of vaccination is that while disease transmission is eliminated, so is collective memory of the disease, which results in an underestimation of the harm caused by the disease. On the other hand, the immediate risk of adverse effects such as fever, anaphylaxis, and vaccine-mediated infection may cause concerns among parents, particularly considering the underdeveloped and vulnerable nature of the immune systems of newborns. Further, because of omission bias, individuals tend to be more concerned with consequences arising from their actions rather than inactions, and therefore may overestimate the risk of rare adverse effects of vaccines (20, 21).

With antibiotics, there is a perception that they are completely safe. Even when the risk of emerging antibiotic resistance due to antibiotic overuse is acknowledged, individuals tend to assess their personal immediate risk of resistance-related effects to be low (22). The majority of antibiotics prescribed in primary care are done so unnecessarily for viral conditions such as acute cough and diarrhea (23, 24) but based on the perception that they could prevent secondary bacterial infections. The misuse of broad-spectrum antibiotics can lead to the disruption of patients' gut microbiomes (25, 26), obesity, and irritable bowel syndrome (8), as well as the emergence of antibiotic-resistant bacteria which may cause severe infections and spread within a community (27, 28). However, the consideration of such long-term effects is often superseded by the urgency of relieving symptoms, which is associated with antibiotic treatment (29). The problem of antibiotic overuse and emerging antibiotic resistance is multifaceted (30). Risk perceptions of physicians, patients, pharmacists, and livestock owners all have a propensity to drive antibiotic consumption. Policy interventions to encourage judicious use of antibiotics should also consider the multi-scale information flow and interactions among these players.

Cognitive Biases in Risk Perception

How do people form risk perceptions? In general, this is done using heuristics rather than reflective thinking. One example is the availability heuristic that most people use to assess the likelihood of a catastrophic event based on how readily examples come to mind (31). As mass vaccination successfully reduces the population-level prevalence of an infectious disease, knowledge of the disease also declines over time, leading to underestimations of its severity. While antibiotic resistance is a growing threat to global public health, it is also a relatively recent, rarely reported one. Consequently, a lack of experience with the threat and the relatively low number of reported cases of antibiotic-resistant infections to date may increase the perceived safety of antibiotic use. Another aspect shaping the perceived risk of antibiotic resistance is that it is a "slowly emerging" problem similar to climate change, and the uncertainty surrounding its ultimate severity results in it being assigned a lower priority status compared to more immediate threats (32). Such present bias or discounting also influence medical decisions through physicians feeling pressured to prescribe antibiotics in order to satisfy patients' expectations (33). Fortunately, such heuristics also suggest the possibility of using nudges to influence decision making to improve societal outcomes, as will be discussed in Section 6: Policy Interventions.

Individual health decisions are made based on perceived individual risk rather than societal risks. Parents may choose not to vaccinate their children, despite statistics favoring vaccination, if they perceive their children to be more vulnerable to side effects of vaccination or more resistant to the disease. While epidemiological statistics based on large datasets offer "one-fit-all" recommendations for health decisions, individuals may think such conclusions do not apply to their personal cases (20). For example, older individuals may perceive themselves to be more vulnerable than average to influenza, and thus may vaccinate more accordingly (34). However, in the case of antibiotic use, people are often overconfident that they personally would not be affected by antibiotic resistance despite their awareness of the threat at the community level (22, 32).

The framing of choices also affects how individuals assess decisions and outcomes. For example, Emergency Department clinicians tend to prescribe more antibiotics when they view their possible outcomes as either improving a patients' health or having them remain ill, compared to those who frame this decision as one of balancing the potential harm of therapy vs. a patient's continued illness (11). One potential explanation for the drastically different decision patterns in vaccine underuse and antibiotic overuse is that when deciding whether to vaccinate, people are balancing the risks of vaccine adverse effects vs. disease contraction, similar to the second type of clinicians; while when deciding whether to take antibiotics, patients are more likely to frame the decision as the first type of clinicians do, focusing on the potential positive consequences of antibiotic use while ignore the possible negative outcomes. Combined with the cognitive bias of loss aversion, individuals will refrain from accepting the risk of contracting vaccine-related side effects if the benefit of vaccination is only to keep them healthy with no additional gains to their status quo. In keeping with this notion, omission bias, or the tendency for individuals to feel more responsible for a negative outcome when it is due to their action rather than inaction, further increases the general tendency to avoid risks associated with even very rare events (21). Therefore, when the possible adverse side effects of a vaccine are known, even if the chances of them occurring are very low, individuals tend to be more cautious about actively getting vaccinated compared to the potentially riskier inaction of doing nothing. One reason that omission bias does not seem to affect antibiotic use may once again be due to the misperception that antibiotics are completely safe.

The association between perceived risk and health behavior is not definitive. Other individual or social factors such as emotion and trust could also influence individual decisions (35). Anticipated regret plays an important role in health behavior as people try to minimize regret they expect to experience (36). The stronger anticipation of regret for taking the action of getting vaccinated as well as the weaker anticipation of regret associated with inaction will both encourage vaccine refusal (36). Implementation of certain policies to promote responsible health behavior may actually be destructive for arousing negative emotions about the enforced action, or may reduce individual trust in governments. In an experimental vaccination game, Betsch and Bohm (37) showed that compulsory vaccination increased anger and negative attitudes about vaccination among participants, and decreased vaccination uptake in later voluntary vaccination among the vaccine hesitant individuals.

The social context is crucial in shaping individual trust in authoritative recommendations on vaccine and antibiotic use. Organized resistance to vaccination has a long history dating back to the nineteenth century following the enforcement of the smallpox vaccine in England (38). One interesting characteristic of vaccine critical groups is that they consider trust in others (especially the government) to be passive, and instead associate responsibility and empowerment with

the act of individually assessing parental decisions, including those related to vaccination (38). In the case of antibiotic use, the consumer-provider relationship between patients and physicians plays an important role in antibiotic prescription. Physicians often feel pressured to prescribe antibiotics to satisfy perceived patient expectations, which are often overestimated (39). This overestimation can lead to a vicious cycle of escalated antibiotic prescribing practices by leading patients to believe that antibiotics are actually necessary for some self-limiting illnesses (39).

SECTION 3: FREE-RIDING AND INDIVIDUAL INCENTIVES

The effects of an individual's use of vaccines and antibiotics extend beyond the first order prophylactic and/or treatment benefits they incur. In the case of vaccination, when uptake is sufficiently high, susceptible portions of a population are protected from infection by the presence of immune individuals, a concept known as herd immunity (40). Antibiotic use results in both positive and negative societal effects, or externalities, that are generally not accounted for by the individual when choosing whether or not to take a drug. On the one hand, when an antibiotic succeeds in curing an individual, society benefits from the positive externality of a reduced chance of transmission of that pathogen (41). On the other hand, treatment with a specific antibiotic results in the selection of pathogen strains that are resistant to that treatment, thereby reducing the expected future usefulness of the drug (41).

In economic terms, the non-excludable and non-rivalrous nature of herd immunity makes it a public good, and as a consequence it is vulnerable to free-riding, or use by individuals who do not contribute to maintaining it (42). More concretely, individuals will choose to vaccinate at a rate that is lower than optimal for society since herd immunity protects them from disease even in the absence of vaccination. Similarly, economic theory predicts that the negative externality of resistance associated with antibiotic use results in individual levels of consumption exceeding those that are societally optimal, since the additional cost of decreased drug effectiveness is not borne by the individual user. These concepts are further complicated by the global nature of pathogen spread and antibiotic resistance, which may alter the incentives for policy implementation in a single geographical region.

Considerations by Individuals

When the perceived risk of a disease is higher, or in other words the perceived benefits of undergoing a prophylactic intervention against it increase, individual decision makers are more likely to take preventative action (9, 43), a phenomenon known as prevalence-dependent behavior. In some cases, this pattern of behavior may result in counterintuitive health outcomes such as observed increases in cases of HIV and other sexually transmitted infections (STIs) following the widespread use of antiretroviral therapies (ARTs) due to increased risky behavior when the perceived risk of disease is reduced (43). In this sense then, the incredible success of vaccines as a public health initiative may also be related to the decline in their voluntary uptake due to the near elimination of previously common diseases from recent memory.

In the simplest terms, balancing the perceived risk of acquiring a disease is the risk of receiving its medical intervention. In the case of vaccination, although there are undoubtedly real associated risks such as the possibility of the oral polio vaccine (OPV) reverting to a pathogenic form of the virus *in vivo* (44), the incidence of such events is extremely low. On the other hand, despite evidence of lasting effects of antibiotics on human health, elevated levels of consumption reflect a general societal sentiment of safety toward antibiotics. The low perceived risks and high negative externalities associated with antibiotic use are evidenced in a number of theoretical economic studies on individual antibiotic uptake in the presence and absence of a social planner that conclude that antibiotic allocation in an uncontrolled market economy will differ from that of the social optimum (45, 46) and even potentially the Nash equilibrium (47). As an example, despite the absence of rigorous testing regarding the efficacy of using growth promoting antibiotics (GPAs) in broiler chicken production, and even evidence that removal of GPAs may *increase* the net value of the flocks (48), the practice was not banned in the USA until 2017.

The costs of these medical interventions, including not only direct medical costs but also indirect costs such as lost time, also influence individual decision making. In one study, 26.1% of respondents in a state with personal-belief exemptions for vaccination stated that they submitted such an exemption for convenience purposes to enroll their children in school, as doing so was less costly than fulfilling the vaccination requirements (49). Therefore, the cost of vaccination must be maintained at low levels, at least relative to the cost of opting out, to encourage uptake in order to compensate for the associated positive externality of herd immunity. In the case of antibiotics, the choice of drug treatment is often made on the basis of cost-effectiveness (41). This, along with institutional control of drug procurement and the fact that clinical treatment guidelines are typically issued by national public health bodies, results in frequently uniform antibiotic choices for given conditions (41). However, the use of a single drug increases the likelihood that a resistant strain will evolve (41), which consequently decreases its future effectiveness. As a result, in order to preserve cost effectiveness while also minimizing the emergence of resistance, economic models for the extraction of non-renewable resources have been used to study the timing of antibiotic use (50, 51). Another solution may be to simultaneously prescribe a variety of drugs randomized over patients in order to mitigate excessive selection pressure toward a single drug or drug class (41). Overall, since individual users do not bear the cost of the negative externality of resistance associated with antibiotic use, there is a need for policy interventions to adjust the price of antibiotics accordingly.

In addition to the perceived cost and benefits of intervention, there is evidence that individual values and sentiments of social responsibility may shape medical decisions. Previous studies found altruism to be a strong motivator in the decision to obtain a vaccination (52), yet in one study where 69% of participants

qualified as pro-social as opposed to pro-self as quantified by a social value orientation score, 89% of participants switched behaviors at least once from vaccination to non-vaccination depending on the conditions of the game and their perceived individual infection risks and vaccination costs (53). Importantly, individual values generally cannot be considered in isolation, and must be evaluated within the context of the social norms relevant to the groups they belong to. For instance, one study found that a stronger motivator than either altruism or free-riding for getting vaccinated was the behavior of bandwagoning, or making a decision in line with that of others (52). This type of behavior is consistent with the notion of an availability cascade, or the self-reinforcing process of collective belief formation within groups to avoid individual reputational harm (49). In this way, it has been suggested that one method of increasing vaccination rates may be to preferentially target individuals who form "hubs" of social networks as a result of the social influence they exert over others (54).

Considerations Across Hospital and National Boundaries

The transboundary nature of disease spread and emergence of antibiotic resistance have prompted a large number of studies into the conditions for cooperation between decision making bodies and how this affects their individual behavior. At the level of hospitals, the control of hospital-acquired infections (HAIs) would likely result in real economic benefits since hospital stays are typically longer for patients infected with resistant bacterial strains (47). In urban settings where patients are exchanged between numerous facilities, game theoretical studies have concluded that the amount a single hospital will invest in hospital infection control (HIC) is dependent on the proportion of patients potentially carrying resistant bacterial strains as well as the strain transmissibility (55). The same study also found that in the absence of coordination, the number of hospitals who will act selfishly and free-ride on HIC investments of other facilities is expected to grow as the number of hospitals in the network increases (55). Therefore, regional coordination and planning between hospitals is likely essential for controlling HAIs (55). Indeed, targeted HIC interventions such as government subsidies and universal decolonization have shown promise in both theoretical (56) and empirical studies (57), respectively.

This same logic also applies at the global scale, and suggests that a global coordinated response may be necessary for the control of antibiotic resistance (58). In the absence of coordination, countries have the incentive to free-ride off of the vaccination efforts of their neighbors without ensuring that their own coverage levels are at the social optimum (59). This notion has received a lot of attention particularly in the context of disease eradication due to the massive potential gains arising from the eliminated need to maintain vaccination. For instance, it is estimated that the annual global benefit of smallpox eradication is about $1.35 billion (using 1967 as a base year), while the total cost of its elimination from endemic countries between 1967 and 1979 was about $300 million (60). While there are a number of pathogen-specific biological, sociology, and epidemiological reasons that complicate the eradication of a particular disease, global cooperation and the incentives of individual nations are also very important considerations (44, 60–64).

SECTION 4: SOCIAL NORMS AND GROUP DYNAMICS

Vaccine hesitancy and refusal are prominent in geographical and socioeconomic or religious clusters (65). This suggests that an important feature of vaccine-related behaviors is their propagation at the community level. Indeed, a number of studies have supported the idea that vaccine hesitancy and refusal are social norms. Social norms can be broken down into two categories: descriptive and injunctive. Descriptive norms include behaviors that are performed by community members (i.e., what *is* done), while injunctive norms describe behaviors that receive approval or disapproval from the community (i.e., what *ought* to be done) (66). Social norms principally spread via contagion, and this effect is amplified by homophily (67, 68). Studies from numerous sociocultural contexts have illustrated the influence of vaccine-related norms on individual behavior. These outcomes are explained using the Theory of Planned Behavior, which posits that the performance of a behavior is principally the result of its antecedent intention (69). The intention, in turn, is informed by a social norm.

In a number of cases, the norm of vaccine acceptance has been shown to predict individual behaviors. Agarwal et al. (70) investigated a combination of descriptive and injunctive pro-vaccination norms in the context of college students' vaccination behaviors. They found that four out of the six norms tested showed statistically significant correlation with actual vaccination behaviors (70). A study on Nigerian mothers' acceptance of the Bacille Calmette-Guérin (BCG) vaccine showed that living in a community with pro-immunization activism predicted a more than twofold increase in the odds of BCG coverage (71). A similar study on Tongolese mothers found that the communication of a pro-immunization message by the chief, as well as vaccination by the chief himself (a descriptive norm), led to increases in vaccination by mothers for themselves and their children (72).

However, social norms can have negative effects on vaccine uptake if they are anti-vaccination or more generally anti-medicine and anti-establishment in nature. For instance, Brunson found that within an individual's social network, increases in the number of people recommending non-vaccination were associated with an increased likelihood of that behavior (73). Similarly, a study on homeschooling parents' vaccination behaviors found that more parents disagreed with the statement "friends think I should vaccinate my children" than agreed with it. Further, more parents agreed with the statement "friends think the risks of vaccination outweigh the benefits" than disagreed with it. These statements refer primarily to injunctive norms and their effects were reflected in vaccine coverage: only 38% of parents stated that their children had received all the recommended vaccines (74). In the case of the HPV vaccine, gender norms have led to substandard coverage for men. Clinical

trial data for the HPV vaccine had originally come from female subjects, so FDA approval was female-specific. Over time, the HPV vaccination recommendations have maintained a gender discrepancy, creating the sense that HPV disproportionately affects women, when in reality, it affects all genders. The feminization of the HPV vaccine has had undesirable effects on the distribution of uptake between the genders: in 2014, 40% of women and 22% of men completed the vaccine series (75). In addition, studies have documented the effects of paternal beliefs surrounding vaccines on the likelihood that mothers will vaccinate their children (74, 76).

Given the weight of evidence suggesting a link between vaccination norms and behaviors, it is important to understand the motivations underlying norm conformity. These motivations can be different for individuals entering a community (social newcomers) and long-standing community members. It is therefore necessary to consider them separately.

Anti-vaccination as Capital

Case Study: Waldorf (Steiner) School Parents

The vaccination tendencies of newcomers to communities with pre-existing anti-vaccination norms can be understood from studies on Waldorf school parents. Waldorf schools offer alternative forms of education and are thought to account for a significant number of children whose parents file personal belief exemption forms. These forms would allow their children to remain un- or under-vaccinated on the basis of non-religious and non-medical explanations and still attend school (77). Sobo found that Waldorf school parents appealed to the traditions of the school culture, which emphasize "[looking] away from biomedicine." One of the study's participants was quoted as saying, "the school philosophy actually embraces illness because they believe that when your body has a strong illness, particularly a fever, it precedes a developmental leap in the child." Parents received significant pressure to follow these traditions (77). Based on the analysis of Waldorf school parents, it appears that social newcomers are motivated by a desire to belong, and they adopt community traditions to do so.

Attwell et al. (78) provide a theoretical framework for understanding these types of motivations. They use Bourdieu's theories of "capital" and "habitus" to analyze vaccination decisions. For instance, they suggest that the induction of parents into non-vaccinating communities, can be seen as a drive to acquire cultural capital in a new social context (78). Reich (79) uses a similar line of reasoning for communities of mothers who see themselves as autonomous actors in relation to their children's' health. They mobilize this social capital to gain validation for their rejection of vaccines (79). Attwell et al. also read the tendency to reject vaccines as the acquisition of symbolic capital, which encompasses behaviors that are seen as a "positive sign" and distinguish the group from another, similar to injunctive norms (78). This adds to previous research on "cultural cognition" —the tendency for individuals to match their ideas to those of the broader community as a way of avoiding cognitive dissonance and building solidarity (80–82). This tendency is accentuated when the idea in question is a distinguishing feature of the group (83). A significant finding from Atwell et al. was that individuals moving from one community to another felt a sense of instability, which was then resolved by conformity to the community's accepted "habitus." "Habitus" refers to the largely unconscious dispositions that a community shares—in this case, the tendency to reject or accept vaccines (78). Therefore, the vaccination decision can be seen as a source of cognitive resolution among social newcomers when the new community holds acceptance or rejection of vaccination as a social norm.

Two cognitive mechanisms have been implicated in norm-related vaccine hesitancy: omission bias (as previously described) and the credibility heuristic (84). The credibility heuristic refers to the tendency of individuals to evaluate the merit of an argument based on the perceived credibility of the source. Importantly, individuals tend to confer credibility to sources with whom they share an in-group connection (84, 85). These mechanisms are consistent with the framework of social identity theory, which posits that the social context for vaccination decisions puts them outside the realm of individual rationality (86, 87). One particularly pertinent example of a long-standing anti-vaccination group is those believing in complementary and alternative medicine (CAM). This group encompasses an ensemble of communities that believe in an entirely DIY approach to medical care. A study on these communities found that they selectively rely on information about the failures of Western medicine to justify their tendencies, while ignoring information about its utility (88). These biases also highlight the possibility of in-group attachment being channeled into outgroup hate. A study of the ways in which non-vaccinating parents portray the vaccinating mainstream found that they constructed a narrative of lifestyle, health, and decision-making inferiority for vaccinating parents. This narrative helped to further cement their beliefs (88).

Other Group Dynamics

There are certain groups for which the social norm of non-vaccination has distinct historical origins, and it is worthwhile to consider them separately. Several studies have found that African-Americans refuse vaccines more than other racial groups (89–91). Historical tensions between the medical community and African-Americans have been proposed to account for this tendency to reject vaccines. Mistrust of the medical establishment among African-Americans stems back to the era of slavery, when slaves were used as subjects for medical experimentation without their consent or personal gain. Out of this phase of selective experimentation grew the sense among African-Americans that medical technologies were weapons designed to be used against them. This type of stigma has persisted (92).

Orthodox protestant communities have also been studied for their vaccination decisions. Historically, religious arguments both in favor of and against vaccination have been circulated in orthodox protestant groups. Ruijs et al. (93) distinguish

between appeals to tradition and "deliberate" choice in religion-based decision making. In traditions-oriented families, they found little evidence for choice consideration at all. A Bourdieusian analysis of this approach to the decision would theorize that non-vaccination was "habitus" for these families. Among the deliberate choice group, most participants cited personal religious experiences (for instance, praying to God for help in the decision) rather than consultation with religious leaders as the predominant factor in their choice (93). This finding suggests that orthodox protestant communities are structured to prioritize personal experience over the influence of social leaders.

Norm Effects on Antibiotic Use and Prescription

Within various groups, including the aforementioned CAM community, there exists a belief that Western medical practices are "unnatural" (94). In these groups, the reasoning for rejection of vaccines and antibiotics tends to overlap. Looking at public beliefs about antibiotics, Norris et al. (95) found that for many of their study's participants, aversion to antibiotic use reflected a more general reluctance to take any sort of medications (95). Previous research had elucidated a relevant psychological mechanism: namely, the effect of illness perceptions on help-seeking behaviors (96) (e.g., inquiring about the possibility of taking a course of antibiotics). Beliefs in either 'holistic' medical practices or the body's innate power to fight off infections was associated with a decrease in such help-seeking behaviors (97–100), and therefore, likely a decrease in willingness to use antibiotics. However, rejection of antibiotics in general does not appear to be held as an injunctive norm in the same way that vaccine hesitancy can be. More often, norms around antibiotic use lead to overuse, and this trend is pertinent to the discussion of antibiotic resistance. Norms play a role in two ways: by affecting patient expectations and by affecting prescriber approaches.

The tendency for individuals to use antibiotics is related both to their perceptions of antibiotic efficacy and their perceptions of whether antibiotics are needed for their particular ailments. As mentioned in the context of vaccine decisions, the basis on which individuals form these sorts of perceptions is in large part the behaviors of the individuals in their social networks. This is true of other medical interventions as well. For instance, Zikmund-Fisher et al. (101) found that when presented with descriptive norms relating to cancer treatment, study participants took behavioral cues from those norms (101). There is some evidence to suggest that patients take similar social cues in relation to their antibiotic use, but it is limited (102, 103). More research must be done to investigate this connection. However, a significant body of literature has shown a related social connection to antibiotic use: namely, that cultural values affect patient demands for prescription antibiotics. This connection has been studied in Australia (104, 105), England (106), France and Germany (107), Europe in general (99, 100, 108, 109), Egypt (110), and Tanzania (111). In all of these cases, ideas about the physician's role in the physician-patient relationship, the need for antibiotics, the efficacy of antibiotics, and the dangers of antibiotics affected patient demand for prescription antibiotics. Cultural values can also have an impact on the use of non-prescription antibiotics. For example, Widayati et al. (112) found that overuse of non-prescription antibiotics in Yogyakarta City, Indonesia, was attributable in part to the prominent belief that medical consultation is a waste of time (112). Interestingly, notions of self-efficacy were associated with high rates of non-prescription antibiotic use in Lithuania, even when the majority of knowledge about antibiotics was coming from medical professionals (113). In Palestine, similar notions of self-efficacy, combined with a positive attitude toward medications and a lack of public education about antibiotics was associated with high rates of non-prescription antibiotic use (114).

In addition, prescription norms and physicians' own ideas about their roles in relation to their patients can influence trends in antibiotic use. For example, Chan et al. (115), studying a hospital in Singapore, found that junior physicians deferred to the practices of senior physicians, thus setting up a norms-based prescribing pattern in their hospital. This pattern was problematic, because physicians tended to focus on their subjective clinical judgements, as opposed to the hospital's guidelines on antibiotic prescription (115). A similar hierarchical, norms-based prescription structure was found by Papoutsi et al. (116) in their study of doctors-in-training (116). In a comparison of antibiotic use in France and Germany, Harbarth et al. found that one of the primary factors accounting for the greater antibiotic use in France was the difference in prescription norms. In Germany, for suspected cases of respiratory tract infection, diagnostic tests were performed, whereas in France, in the face of diagnostic ambiguity, prescription of antibiotics was the default practice (107). Further, physicians' ideas about the benefits of prescription, which are culturally influenced by the degree of community emphasis on guideline adherence and patient satisfaction have been shown to affect their likelihood of prescribing antibiotics, following the Theory of Planned Behavior (117).

Physicians' ideas about their roles in the physician-patient relationship can also have an impact on rates of antibiotic use. For instance, Butler et al. (118) interviewed physicians on their prescribing behaviors, and one participant who admitted to ignoring the guidelines on antibiotic prescription said, "I'm quite well aware of the lack of firm evidence that antibiotics treat [upper respiratory tract infections] and that in terms of evidence-based medicine we overprescribe antibiotics, but my own view is that I don't really care… your goals at the end of the conversation is for both you and the mother and the baby to be satisfied." (118) In this case, the physician sees his primary responsibility as his patient's satisfaction. Similarly, Kandeel and colleagues' study on antibiotic use practices in Minya, Egypt, found that antibiotic prescription was significantly associated with the patients' preferences for such treatment (110). Thus, the over-prescription of antibiotics appears to be influenced by a combination of hospital-specific descriptive norms and broader cultural ideas about physician responsibilities.

SECTION 5: ACTUAL RISKS, COSTS, AND BENEFITS OF VACCINE UNDERUSE AND ANTIBIOTIC OVERUSE

Vaccine hesitancy is a growing issue (119) which poses risks and costs to societies and individuals alike, including increased infection rates, economic costs, and decreased herd immunity. Vaccine hesitancy is facilitated by a number of factors, including the option to obtain non-medical exemptions in several states (120) which increases the likelihood of disease outbreaks. An overwhelming driver of vaccine hesitancy is the belief of adverse reactions that an individual can have to vaccines. To address these concerns, the Vaccine Adverse Event Reporting System (VAERS) (121) was implemented to factually track adverse reactions. Furthermore, the Vaccine Safety Datalink (VSD) actively studies adverse effects post-vaccination, in addition to generally ensuring safety (122).

How does vaccine hesitancy compare to antimicrobial use? Antimicrobial resistance is a worldwide problem that has plagued society since the introduction of antibiotics in the 1940s (123), and has been exacerbated in the present day by antimicrobial overuse. The sustained emergence of resistant pathogen strains results in the need for continued development of more powerful antimicrobials, and current drug discovery efforts are unlikely to be sustainable. Furthermore, it is believed that significant improvements to current antibiotics will remain elusive (124). Motivated by this, in 2015 the WHO enacted the "Global action plan on antimicrobial resistance" (125) in a large-scale effort to curb antimicrobial resistance and develop strategies to address this issue.

In this section, we first examine general societal risks and costs associated with vaccine refusal, along with societal benefits of vaccination, in addition to risks and costs of antimicrobial resistance. Then, we briefly summarize risks, costs, and benefits at the individual-level. Subsequently, to highlight and contextualize these general ideas, we present two contrasting cases studies for vaccination and conclude with a specific example of antimicrobial resistance and its consequences.

Societal Aspects

Vaccination

Perhaps the most obvious societal cost of vaccine underuse is the cost of treating vaccine preventable illnesses (126). This cost takes two forms: actual hospitalization, treatment and mortality, and downstream effects. Oftentimes, the cost of even a single hospitalization far outweighs the cost of immunization. For example, a "successful" infection by Diphteria and Tetanus invariably lead to hospitalization. Whitney et al. (126) estimate that the costs of treating these diseases (i.e. hospitalizations) are about $17,000 and $100,000, respectively, although in 2017 the total costs of treating a single unvaccinated child for Tetanus could exceed $800,000 (127).

It is also important to consider both direct and indirect consequences of vaccine refusal. On a broader scale for the USA, Ozawa et al. (128) estimated that adults lacking immunizations represented an economic cost of about $7 billion annually. This number was reduced to $2 billion for infected vaccinated individuals. Moreover, for vaccination against varicella, Zhou et al. (129) showed that the benefit to cost ratio in the USA is about 4 for a single dose of the vaccine, and about 3 for two doses. Furthermore, Omer et al. (130) found that increased non-medical exemptions in the USA resulted in an increased disease burden of pertussis.

Antimicrobial Resistance

In a landmark study, Michaelidis et al. (131) computed the "hidden" societal cost of a single course of antibiotics. These authors focused on costs attributed to antibiotic resistance, and concluded that each course of antibiotics imposes a societal burden equivalent to $13 on average. This cost is non-negligible considering that the actual antibiotic cost to the individual may vary from a few dollars to a few hundred dollars.

It is currently estimated that antibiotic resistance poses a significant societal burden through infection: In the US, infections are on the order of millions ($\sim$2 million), leading to several thousand deaths ($\geq$23,000) (132). The societal implications of these infections include not only the costs associated with these high annual infection and death rates, but also the increasing probability of these resistant bacteria infecting susceptible hosts as the number of infected individuals rises. For instance, in a study examining *Salmonella* outbreaks, Varma et al. (133) found that infections with resistant pathogens led to 14% more hospitalizations than infections with non-resistant strains. Despite this, most current societal cost estimates only consider the direct consequences of infection with a resistant bacterial strain, and consequently underestimate their true burden (e.g., surgeries would lead to substantially more infections) (134).

Individual Aspects

Vaccination

Individual risks and costs of vaccination lie almost entirely along two axes. First, as previously discussed, there are quantifiable risks of adverse effects associated with every vaccine, as discussed by Stratton et al. (135). In a study spanning multiple vaccines, Bohlke et al. (136) found a rate of 0.65 cases of anaphylaxis per million doses of vaccines administered. Out of five such cases identified, anaphylaxis did not lead to mortality. The second dominant risk is that of an individual's infection probability (137), although as previously mentioned in the context of herd immunity, this probability, and thus its associated cost are functions of the vaccination decisions of others. If vaccination achieves a high enough threshold such that herd immunity is maintained, then the individual that refuses vaccination will be effectively protected with a very low probability of infection. On the other hand, if herd immunity is not achieved, the probability of infection for an unvaccinated individual depends upon the value of R_0 for the specific disease. For high values of R_0, Susceptible-Infected-Recovered (SIR) disease models (138, 139) predict that nearly all initially susceptible individuals will be infected. Thus, refusing vaccination before such an epidemic will very likely result in infection and any complications that may ensue.

Antimicrobial Resistance

Antimicrobial resistance gives rise to many individual risks and costs for the individual. First and foremost, incorrect administration of antimicrobials can enhance selection for resistant pathogens and aggravate infection (140). There are also adverse effects associated with using more potent antibiotics to combat resistant pathogens, such as perturbations to the human gut microbiome (141), which may be correlated with severe downstream consequences ranging from weight-gain to increased susceptibility to other infections and even to cancer (142). Thus, conditional upon successful treatment, infection with pathogens that are resistant to milder antibiotics may have effects that last significantly longer than the actual infection.

Case Study: Influenza Viruses

Influenza viruses are single-stranded, negative sense, segmented RNA viruses that exert significant yearly seasonal burdens on human populations (143) largely due to rapid evolution of the immunodominant hemagglutinin (HA) surface protein (144). Current vaccination strategies elicit immune responses to exposed HA regions that are known targets of antibodies (145), but only have moderate efficacy (146). To further increase immunity in human populations and reduce the number of necessary vaccines to maintain herd immunity, there are currently significant efforts aimed at developing and understanding the impact of Universal Influenza Vaccines (UIVs) that would provide broad protection across multiple strains for multiple years (147–149).

Influenza vaccine refusal is often shaped by preconceived notions of low effectiveness (150). A meta-analysis of studies pertaining to influenza vaccination and health care workers revealed similar notions, in addition to beliefs of low personal risk associated with actual influenza infection (151). Thus, individuals believe that the risk of adverse reactions due to influenza vaccines far outweighs the protection they provide (152). Yet, recent work indicates that vaccination dramatically decreases actual individual risk, both in children [e.g., (153, 154)] and adults [e.g., (155)].

For individuals that are at low risk of influenza complications, the single most compelling reason for annual vaccination is to increase herd immunity and thus lower the probability of transmission to individuals that are at risk of complications. In general, those at elevated risk include children younger than 1 year of age, adults older than 65, pregnant women, and those with chronic illnesses (156). Despite the current "imperfect" seasonal vaccines that require yearly updating and that have mixed efficacy, their societal impacts have been important. Arinaminpathy et al. (157) estimated both direct protection, i.e., a vaccinated host successfully resisting influenza infection, and indirect protection, i.e., a potential host averting infection due to reduced transmission from vaccinated hosts. These authors concluded that vaccination reduced seasonal influenza burden in the United States by between roughly 10 and 37 million cases. Overall, advantages of UIV would include decreased yearly cases and transmission, as well as a change in the evolutionary dynamics of the influenza virus, which may potentially reduce its future burden (158).

Case Study: Measles

Measles morbilliviruses are singled-stranded, negative-sense, non-segmented RNA viruses that principally infect school-aged children (159). Measles infections can lead to severe complications or damage to the central nervous system years after infection. Due to long-lasting immunity following immunization, vaccination against measles has been highly successful. For example, following the successful use of the MMR vaccine in the USA, endemic measles has been eliminated (160). Following the publication of a retracted paper on the association of the MMR vaccine with autism, vaccine safety has been the subject of extensive studies that reveal no association with autism (161). Yet, vaccine hesitancy introduces pockets of susceptibility in populations. In conjunction with high transmissibility, this results in a loss of herd immunity and possible measles epidemics. For example, in Washington state, there was recently a measles outbreak (162). These outbreaks impose a significant burden on public health infrastructure, in addition to exerting significant direct and indirect costs on individuals.

The economic consequences of measles vaccine hesitancy can be significant [e.g., see predictions of (163)]. Furthermore, a retrospective modeling analysis following the 2012–2013 outbreak in Merseyside, UK established that this outbreak could have been averted with 11,793 vaccines (182,909 pounds) instead of costing 4.4 million pounds due to infections (164). In addition, there are serious public health risks associated with vaccine hesitancy, including substantially more infections (163) and infection risk for those that refuse vaccines [e.g., for measles risk pertaining to exemptions in children, see (165)]. But the individual risk tied to MMR vaccine refusal extends beyond immediate infection into long-term immunological consequences for children that contract measles. Indeed, immunomodulation by measles infection in children leads to immunosuppression (166). Subsequently, other opportunistic pathogens can infect these immunocompromised children, resulting in increased childhood mortality (166).

Case Study: Methicillin-Resistant Staphylococcus Aureus (MRSA)

Staphylococcus aureus is a bacterial human pathogen that causes skin and blood infections. Perhaps the archetype of antibiotic resistance, *Staphylococcus aureus* is also responsible for significant nosocomial infections. How did methicillin-resistant *Staphylococcus Aureus* (MRSA) emerge? Treatment for this pathogen began through the introduction of penicillin in the 1940s. Yet, the emergence of resistant strains to penicillin led to the use of methicillin to successfully combat this pathogen (167). Eventually, however, resistance to methicillin developed, leading to current MRSA. This "specialization" of resistant bacteria acquiring further resistance illustrates that antimicrobial resistance is not novel. Furthermore, genetic analyses have revealed that these MRSA pathogens are further specializing and becoming vancomycin resistant (167).

At the individual level, the odds of dying following surgery when infection with MRSA is acquired compared to without infection are 11.4:1. In contrast, these same odds compared

to infection with a non-resistant *Staphylococcus Aureus* strain are 3.4:1 (168). The individual cost of an MRSA infection extends beyond the direct health implications of such infections. Described as the "twenty-first century lepers" (169), MRSA patients also face tremendous stigma due to fear of contagion. Therefore, the social isolation that results from this stigma may lead to significant psychological harm, substantially increasing the individual burden of a MRSA infection.

What are the societal costs of MRSA outbreaks? Stigmatization can lead to decreased cooperativity and productivity, affecting proper societal functioning. Furthermore, MRSA outbreaks can exert pressure on societal resources. For example, a recent MRSA outbreak in a Finnish hospital facility in 2003–2004 lasted 14 months and had tremendous direct and indirect costs (170).

SECTION 6: POLICY INTERVENTIONS

Direct Manipulation of Cost or Supply of Vaccines and Antibiotics

Several policy proposals to target the overuse of antibiotics and the underuse of vaccines suggest direct manipulation of drug supply or pricing. Among these, extended durations for antibiotic patents have been suggested, which incentivize patent owners to curtail their usage (171). Another proposed solution has been to promote a single buyer for antibiotics, as this actor would have the incentive to take future resistance and drug effectiveness into consideration during their purchasing (171). Pigouvian taxes on antibiotics to absorb the externality of resistance development have also been proposed (47, 58), although a clear disadvantage of such a policy is that it disproportionately affects poor users while doing little to limit the consumption of more affluent ones. In this way then, Hollis and Maybarduk suggest that such a tax may be better applied to industrial and agricultural uses where sensitivity to taxation is more equal (58). Nevertheless, even if reductions in antibiotic use are successful, novel drug classes will need to be developed for sustained treatment options in the future (6). To this end proposals for reimagining the business models for drug development have been proposed (6), as well as incentivizing drug discovery through financial rewards that are delinked from drug prices or volumes (58).

Similarly, it has been suggested that one approach to incentivize vaccination may be through the implementation of fines or rewards (42). However, whether or not such policies would succeed in increasing vaccine uptake are debatable. Indeed, when disease prevalence is high, numerous studies suggest that public uptake of vaccines will also rise, making subsidies irrelevant if they are not properly timed (43). Further, there is evidence of utilization of fines as a means to "buy out" of the action they are designed to incentivize in the first place (42).

Changes to Medical and Prescriber Practices

Another proposed approach to altering vaccine and antibiotic use is through changes in medical practices and physician prescribing behavior. In terms of antibiotic prescribing, studies have found that interactive, computer-based antibiotic guidelines have been successful at reducing the number of inappropriate drug prescriptions (172). Further, the prescription of an antibiotic has sociological ramifications for both the physician and the patient, such as the conclusion that a diagnosis has been made and that the visit is terminated (172). Therefore, for a patient to feel satisfied without an antibiotic prescription, it is recommended that additional explanations detailing the true medical usefulness of these drugs be provided by the physician (172). Indeed, simple interventions to encourage physician-patient communication about the use and risks of antibiotics were shown to effectively reduce antibiotic prescription by 60% (39). Necessarily, however, such recommendations will need to be adapted to the local country and culture within which they are implemented. For instance, antibiotic use without prescription is common in many low-and-middle-income countries (6), and it will be necessary to ensure that those who need the drugs most are still able to obtain them following any policy changes.

Along a similar vein, the way that medical information is presented can also play an important role in treatment uptake. For example, the erroneous over-identification of human papilloma virus (HPV) as a female-specific disease has disadvantaged males from receiving its vaccine, despite growing evidence of a causal role for this virus in penile cancer, anal cancer, and other conditions men are susceptible to (75). This insufficient vaccine uptake among men is exacerbated by an inadequate uptake of the vaccine among women, and hence the inability to generate herd immunity (75). Furthermore, there is evidence to suggest that the way that medical information is presented plays an important role in the magnitude of the risk perceived by the patient (49). In one of the most basic forms of such an intervention, pneumococcal vaccination rates among high-risk patients were found to increase when simple educational information was provided and patient-physician communication about the vaccine was encouraged (173). Further, several studies have found that individuals perceive a risk as greater when it is presented in terms of a frequency (i.e., a 1 in 10 chance) as opposed to a probability (i.e., a 10% chance) (49). Consequently, one strategy for increased vaccination uptake may be for physicians to also present the risk of non-vaccination in frequency terms (49). Similarly, betrayal aversion, or the emotional reaction associated with an object of trust betraying its implicit promise of protection, has been found to occur less frequently when visual aids are employed to communicate risk, suggesting that this may also be a useful strategy for healthcare workers in discussing vaccination (49).

Nudging and Influences to Social Behavior

The range of biases shaping the perceived benefits of vaccination have resulted in individuals with a spectrum of opinions, ranging from those entirely opposed to vaccination to those who are vaccine hesitant, i.e., susceptible to the anti-vaccination message but whose preferences are not hard-and-fast (49). Drawing on concepts from behavioral law and economics, a number of recommendations, or nudges, targeted at vaccine hesitant individuals in order to correct predictable errors

in risk assessment of vaccination caused by cognitive biases without restricting individual choices have been suggested (49). For instance, similarly to the notion of bandwagoning previously introduced, there is evidence that simply informing individuals that others are choosing to vaccinate can increase their uptake (42). Additionally, receiving concrete information through face-to-face interactions may be more effective at initiating action than abstract information, from say a pamphlet (174). Finally, although the ethics of this form of nudging may be more questionable, one study found that when combined with a specific plan for action regarding how and when to get vaccinated, fear was a successful initiator of increased tetanus vaccine uptake among seniors at Yale University (175).

Modifications to laws and government policies regarding vaccination have also been proposed as nudging tools. Herd immunity could be made an excludable good by restricting the community and social activities unvaccinated children can participate in (42), as was done in Rockland County NY in March 2019 following a large measles outbreak. Alternatively, and perhaps more of a shove than a nudge, it has been suggested that removing all philosophical exemptions from vaccination may be an effective way to reduce free-riding and increase uptake (49). Indeed, Mississippi is the state with the highest vaccination rates in the USA, and is one of two states that do not permit religious or philosophical exemptions (49). However, there are concerns that such a hardline approach may not successfully increase uptake given the likelihood of it facing legal challenges and political pushback, and more subtle options in which vaccination is made the default option have been proposed (49). Further, Chapman et al. (34) found that uptake of flu vaccinations increased when the appointments to receive them were pre-scheduled, and allocation of patients to doctors based on geographical proximity has been proposed as a mechanism to alleviate patient attrition in response to reduced antibiotic prescription (176).

Educational Campaigns and Media Coverage

Based on the association between risk perception and health behavior, campaigns to increase the perceived risk of non-vaccination or reduce the perceived risk of vaccination may be effective in promoting vaccine uptake. Fear arousal alone through presentation of dramatic narratives or pictures about the danger of vaccine-preventable diseases has been shown to be ineffective at increasing the intention or action to vaccinate (175, 177). However, it has been shown that individuals overestimate the risk or frequency of occurrence of events that are highly publicized (49), and as a result one important way to rectify such a cognitive bias may be through the regulation of media coverage of outlier cases of negative vaccine consequences, perhaps by enforcing that equal coverage of cases of the diseases that they are preventing be broadcast.

In general, social norms-based campaigns have been found to be most effective when they stress a positive injunctive message (178). That is, messages have the greatest positive impact when they express what the individual *should* do: "take your vaccine" instead of "don't refuse your vaccine." In addition, effective messages convey injunctive rather than descriptive norms: "get your child vaccinated for his or her well-being" instead of "get your child vaccinated because that's what others have done."

Finally, educational campaigns that emphasize the societal consequences of vaccine refusal and antibiotic overuse and set the social expectations for responsible health behavior could also be effective strategies for encouraging the incorporation of societal risk as part in individual risk calculations. For instance, it is commonly observed that individuals misunderstand their personal role in spreading antimicrobial resistance through their use of antibiotic drugs, and consequently believe that the responsibility for control measures lies uniquely with health organizations (22). In this case, better information regarding the causes of antibiotic resistance and specific instructions on how individuals could contribute to controlling antibiotic resistance would likely prove effective.

SECTION 7: CONCLUSION

In this paper, we reviewed the personal, societal, and economic factors affecting vaccine hesitancy and antimicrobial overuse. These insights are helpful to understand the uptake of a potential vaccine against COVID-19. A variety of misperceptions about risk contributes, in part, to the imbalance of vaccine uptake and antimicrobial use relative to their socially optimal levels of consumption. For instance, individuals may underestimate the risk of a disease because of herd immunity and engage in free riding. They may further overestimate the risk of adverse events from vaccination and underestimate the risk of antimicrobial overuse. This may be particularly relevant in the context of COVID-19, and the importance of rigourous vaccine testing to maintain public trust has been emphasized, despite the simultaneous need for unprecedentedly fast vaccine development (179).

From a policy perspective, complications arise from the multifactorial nature of information flow, involving prescribing physicians, patients, pharmacists, the government, etc…, and the cognitive biases that reinforce misperceptions. Social norms of non-vaccination also appear to push individuals toward vaccine hesitancy and refusal. In communities that hold such norms, social newcomers are highly incentivized to conform as a way of building solidarity and securing their positions. For long-standing community members, on the other hand, non-vaccination may become an unconscious behavior or a deeply-rooted belief. In the latter case, individuals are susceptible to a variety of cognitive biases. Finally, specific racial and ethnic communities may have unique relationships to vaccines for a variety of historical reasons, which will be important for policy makers to consider.

The prevalence of vaccine hesitancy and antimicrobial overuse warrants consideration from policy makers because of the individual, societal, and economic costs that they entail. We investigated several policy interventions aimed at encouraging

a shift toward the socially optimal levels of vaccine and antibiotic consumption. These included direct manipulations of the costs of vaccines and antibiotics, changes to prescriber practices, and nudging through modifications of the choice structure or direct regulations on vaccine exemption. NGOs can also engage in nudging through directed campaigns, but the efficacy of these initiatives depends largely on characteristics of the messages that they portray. Based on the successes and failures of previous campaigns, it appears that these messages should be informed by social norms theory. It is critical that any interventions be coordinated between regional actors to match the global nature of pathogenic spread and antimicrobial resistance.

AUTHOR CONTRIBUTIONS

The review was drafted by CW, JP, CS-R, and LY with input and supervision from BG, SL, and RL. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the James S. McDonnell Foundation 21st Century Science Initiative Collaborative Award in Understanding Dynamic and Multiscale Systems, the US Centers for Disease Control and Prevention, the RAPIDD program of the Science and Technology Directorate Department of Homeland Security and the Fogarty International Center, NIH, and the National Science Foundation (grant CNS-2027908), the National Science Foundation Expeditions (grant CCF1917819), and the C3.ai Digital Transformation Institute (grant AWD1006615), and Gift from Google, LLC. 21st Century Science Initiative Collaborative Award in Understanding Dynamic and Multiscale Systems, the US Centers for Disease Control and Prevention, the RAPIDD program of the Science and Technology Directorate Department of Homeland Security and the Fogarty International Center, NIH, and the National Science Foundation (grant CNS-2027908), the National Science Foundation Expeditions (grant CCF1917819), and the C3.ai Digital Transformation Institute (grant AWD1006615), and Gift from Google, LLC, CS-R acknowledges support from the Natural Sciences and Engineering Research Council of Canada through a Postgraduate-Doctoral Scholarship.

ACKNOWLEDGMENTS

We acknowledge useful discussions with all members of the May 2019 *Extending the Cure* meeting.

REFERENCES

1. Olive JK, Hotez PJ, Damania A, Nolan MS. The state of the antivaccine movement in the United States: a focused examination of nonmedical exemptions in states and counties. *PLoS Med*. (2018) 15:e1002578. doi: 10.1371/journal.pmed.1002578
2. Breman JG, Arita I. The confirmation maintenance of smallpox eradication. *N Engl J Med*. (1980) 303:1263–73. doi: 10.1056/NEJM198011273032204
3. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger JA. Vaccine hesitancy: an overview. *Hum Vaccin Immunother*. (2013) 9:1763–73. doi: 10.4161/hv.24657
4. Hill HA, Elam-Evans LD, Yankey D, Singleton JA, Kang, Morbidity Y, et al. Vaccination coverage among children aged 19–35 months — United States, 2017. *MMWR*. (2018) 67:1123–8. doi: 10.15585/mmwr.mm6740a4
5. Bramer CA, Kimmins LM, Swanson R, Kuo J, Vranesich P, Jacques-Carroll LA, et al. Decline in child vaccination coverage during the COVID-19 pandemic — michigan care improvement registry, May 2016-May 2020. *MMWR*. (2020) 69:630–1. doi: 10.15585/mmwr.mm6920e1
6. Laxminarayan R, Duse A, Wattal C, Zaidi AKM, Wertheim HFL, Sumpradit N, et al. Antibiotic resistance-the need for global solutions. *Lancet Infect Dis*. (2013) 13:1057–98. doi: 10.1016/S1473-3099(13)70318-9
7. Blaser MJ. Antibiotic use and its consequences for the normal microbiome. *Science*. (2016) 352:544–5. doi: 10.1126/science.aad9358
8. Gilbert JA, Blaser MJ, Caporaso JG, Jansson JK, Lynch SV, Knight R. Current understanding of the human microbiome. *Nat Med*. (2018) 24:392–400. doi: 10.1038/nm.4517
9. Brewer NT, Chapman GB, Gibbons FX, Gerrard M, McCaul KD, Weinstein ND. Meta-analysis of the relationship between risk perception and health behavior: the example of vaccination. *Heal Psychol*. (2007) 26:136–45. doi: 10.1037/0278-6133.26.2.136
10. Topor G, Grosu I-A, Ghiciuc CM, Strat AL, Lupuşoru CE. Awareness about antibiotic resistance in a self-medication user group from Eastern Romania: a pilot study. *PeerJ*. (2017) 5:e3803. doi: 10.7717/peerj.3803
11. Klein EY, Martinez EM, May L, Saheed M, Reyna V, Broniatowski DA, et al. Categorical risk perception drives variability in antibiotic prescribing in the emergency department: a mixed methods observational study. *J Gen Intern Med*. (2017) 32:1083–9. doi: 10.1007/s11606-017-4099-6
12. Saad-Roy CM, Wagner CE, Baker RE, Morris SE, Farrar J, Graham AL, et al. Immune life-history, vaccination, and the dynamics of SARS-CoV-2 over the next five years. *Science*. (2020) 370:811–8. doi: 10.1126/science.abd7343
13. Song G. Understanding public perceptions of benefits and risks of childhood vaccinations in the United States. *Risk Anal*. (2014) 34:541–55. doi: 10.1111/risa.12114
14. Zanichelli V, Tebano G, Gyssens IC, Vlahović-Palčevski V, Monnier AA, Stanic Benic M, et al. Patient-related determinants of antibiotic use: a systematic review. *Clin Microbiol Infect*. (2019) 25:48–53. doi: 10.1016/j.cmi.2018.04.031
15. Coenen S, Francis N, Kelly M, Hood K, Nuttall J, Little P, et al. Are patient views about antibiotics related to clinician perceptions, management and outcome? a multi-country study in outpatients with acute cough. *PLoS ONE*. (2013) 8:e76691. doi: 10.1371/journal.pone.0076691
16. Bauch CT, Earn DJ. Vaccination, and the theory of games. *Proc Natl Acad Sci USA*. (2004) 101:13391–94. doi: 10.1073/pnas.0403823101
17. Oraby T, Bauch CT. Bounded rationality alters the dynamics of paediatric immunization acceptance. *Sci Rep*. (2015) 5:10724. doi: 10.1038/srep10724
18. Voinson M, Billiard S, Alvergne A. Beyond rational decision-making: modelling the influence of cognitive biases on the dynamics of vaccination coverage. *PLoS ONE*. (2015) 10:e0142990. doi: 10.1371/journal.pone.0142990
19. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science*. (1981) 211:453–8. doi: 10.1126/science.7455683
20. Meszaros JR, Asch DA, Baron J, Hershey JC, Kunreuther H, Schwartz-Buzaglo J, et al. Cognitive processes and the decisions of some parents to forego pertussis vaccination for their children. *J Clin Epidemiol*. (1996) 49:697–703. doi: 10.1016/0895-4356(96)00007-8
21. Asch DA, Baron J, Hershey JC, Kunreuther H, Meszaros J, Ritov I, et al. Omission bias and pertussis vaccination. *Med Decis Mak*. (1994) 14:118–23. doi: 10.1177/0272989X9401400204
22. Brooks L, Shaw A, Sharp D, Hay AD. Towards a better understanding of patients' perspectives of antibiotic resistance and MRSA: a qualitative study. *Fam Pract*. (2008) 25:341–8. doi: 10.1093/fampra/cmn037

23. Rosenstock IM. The health belief model and preventive health behavior. *Health Educ Monogr.* (1974) 2:354–86. doi: 10.1177/109019817400200405
24. Reynolds L, McKee M. Factors influencing antibiotic prescribing in China: an exploratory analysis. *Health Policy.* (2009) 90:32–6. doi: 10.1016/j.healthpol.2008.09.002
25. Dethlefsen L, Relman DA. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc Natl Acad Sci USA*. (2011) 108:4554–61. doi: 10.1073/pnas.1000087107
26. Modi SR, Collins JJ, Relman DA. Antibiotics the gut microbiota. *J Clin Invest.* (2014) 124:4212–8. doi: 10.1172/JCI72333
27. Masterton RG. Antibiotic de-escalation. *Crit Care Clin.* (2011) 27:149–62. doi: 10.1016/j.ccc.2010.09.009
28. Hay AD, Thomas M, Montgomery A, Wetherell M, Lovering A, McNulty C, et al. The relationship between primary care antibiotic prescribing and bacterial resistance in adults in the community: a controlled observational study using individual patient data. *J Antimicrob Chemother.* (2005) 56:146–53. doi: 10.1093/jac/dki181
29. Wood F, Simpson S, Butler CC. Socially responsible antibiotic choices in primary care: a qualitative study of GPs' decisions to prescribe broad-spectrum and fluroquinolone antibiotics. *Fam Pract.* (2007) 24:427–34. doi: 10.1093/fampra/cmm040
30. Lundborg CS, Tamhankar AJ. Understanding changing human behaviour-antibiotic mainstreaming as an approach to facilitate modification of provider consumer behaviour. *UPS J Med Sci.* (2014) 119:125–33. doi: 10.3109/03009734.2014.905664
31. Serpell L, Green J. Parental decision-making in childhood vaccination. *Vaccine.* (2006) 24:4041–6. doi: 10.1016/j.vaccine.2006.02.037
32. Ancillotti M, Eriksson S, Veldwijk J, Nihlén Fahlquist J, Andersson DI, Godskesen T. Public awareness and individual responsibility needed for judicious use of antibiotics: a qualitative study of public beliefs and perceptions. *BMC Public Health*. (2018) 18:1153. doi: 10.1186/s12889-018-6047-8
33. Mangione-Smith R, McGlynn EA, Elliott MN, McDonald L, Franz CE, Kravitz RL. Parent expectations for antibiotics, physician-parent communication, and satisfaction. *Arch Pediatr Adolesc Med.* (2001) 155:800–6. doi: 10.1001/archpedi.155.7.800
34. Chapman GB, Coups EJ. Predictors of influenza vaccine acceptance among healthy adults. *Prev Med*. (1999) 29:249–62. doi: 10.1006/pmed.1999.0535
35. Dubé E, Vivion M, Sauvageau C, Gagneur A, Gagnon R, Guay M. Nature does things well, why should we interfere?: vaccine hesitancy among mothers. *Qual Health Res*. (2016) 26:411–25. doi: 10.1177/1049732315573207
36. Brewer NT, DeFrank JT, Gilkey MB. Anticipated regret and health behavior: a meta-analysis. *Heal Psychol*. (2016) 35:1264–75. doi: 10.1037/hea0000294
37. Betsch C, Böhm R. Detrimental effects of introducing partial compulsory vaccination: experimental evidence. *Eur J Public Health.* (2015) 26:378–81. doi: 10.1093/eurpub/ckv154
38. Hobson-West P. Trusting blindly can be the biggest risk of all: Organised resistance to childhood vaccination in the UK. *Sociol Heal Illn.* (2007) 29:198–215. doi: 10.1111/j.1467-9566.2007.00544.x
39. Altiner A, Brockmann S, Sielk M, Wilm S, Wegscheider K, Abholz HH. Reducing antibiotic prescriptions for acute cough by motivating GPs to change their attitudes to communication and empowering patients: a cluster-randomized intervention study. *J Antimicrob Chemother.* (2007) 60:638–44. doi: 10.1093/jac/dkm254
40. Fine PEM. Herd immunity: history, theory, practice. *Epidemiol Rev.* (1993) 15:265–302. doi: 10.1093/oxfordjournals.epirev.a036121
41. Laxminarayan R, Weitzman ML. On the implications of endogenous resistance to medications. *J Health Econ*. (2002) 21:709–18. doi: 10.1016/S0167-6296(02)00034-6
42. Buttenheim AM, Asch DA. Making vaccine refusal less of a free ride. *Hum Vaccin Immunother.* (2013) 9:2674–5. doi: 10.4161/hv.26676
43. Klein E, Laxminarayan R, Smith DL, Gilligan CA. Economic incentives and mathematical models of disease. *Environ Dev Econ*. (2007) 12:707–32. doi: 10.1017/S1355770X0700383X
44. Barrett S. Economic considerations for the eradication endgame. *Philos Trans R Soc B Biol Sci.* (2013) 368:20120149. doi: 10.1098/rstb.2012.0149
45. Gersovitz M, Hammer JS. The economical control of infectious diseases. *Econ J.* (2004) 114:1–27. doi: 10.1046/j.0013-0133.2003.0174.x
46. Brown G, Layton DF. Resistance economics: social cost and the evolution of antibiotic resistance. *Environ Dev Econ.* (1996) 1:349–55. doi: 10.1017/S1355770X0000067X
47. Rudholm N. Economic implications of antibiotic resistance in a global economy. *J Health Econ.* (2002) 21:1071–83. doi: 10.1016/S0167-6296(02)00053-X
48. Graham JP, Boland JJ, Silbergeld E. Growth promoting antibiotics in food animal production: an economic analysis. *Public Health Rep.* (2007) 122:79–87. doi: 10.1177/003335490712200111
49. Laskowski M. Nudging towards vaccination: a behavioral law and economics approach to childhood immunization policy. *Tex Law Rev.* (2016) 94:601–28.
50. Laxminarayan R, Brown GM. Economics of antibiotic resistance: a theory of optimal use. *J Environ Econ Manage.* (2001) 42:183–206. doi: 10.1006/jeem.2000.1156
51. Laxminarayan R. Act now or later? economics of malaria resistance. *Am J Trop Med Hyg.* (2004) 71:187–95. doi: 10.4269/ajtmh.2004.71.187
52. Hershey JC, Asch DA, Thumasathit T, Meszaros J, Waters VV. The roles of altruism, free riding, and bandwagoning in vaccination decisions. *Organ Behav Hum Decis Process*. (1994) 59:177–87. doi: 10.1006/obhd.1994.1055
53. Böhm R, Betsch C, Korn L. Selfish-rational non-vaccination: experimental evidence from an interactive vaccination game. *J Econ Behav Organ.* (2016) 131:183–95. doi: 10.1016/j.jebo.2015.11.008
54. Fu F, Rosenbloom DI, Wang L, Nowak MA. Imitation dynamics of vaccination behaviour on social networks. *Proc R Soc B Biol Sci.* (2011) 278:42–9. doi: 10.1098/rspb.2010.1107
55. Smith DL, Levin SA, Laxminarayan R. Strategic interactions in multi-institutional epidemics of antibiotic resistance. *Proc Natl Acad Sci USA*. (2005) 102:3153–8. doi: 10.1073/pnas.0409523102
56. Drohan SE, Levin SA, Grenfell BT, Laxminarayan R. Incentivizing hospital infection control. *Proc Natl Acad Sci USA*. (2019) 116:6221–5. doi: 10.1073/pnas.1812231116
57. Huang SS, Septimus E, Kleinman K, Moody J, Hickok J, Avery TR, et al. Targeted versus universal decolonization to prevent ICU infection. *N Engl J Med.* (2013) 368:2255–65. doi: 10.1056/NEJMoa1207290
58. Hollis A, Maybarduk P. Antibiotic resistance is a tragedy of the commons that necessitates global cooperation. *Law J Med Ethics.* (2015) 43:33–7. doi: 10.1111/jlme.12272
59. Klepac P, Megiddo I, Grenfell BT, Laxminarayan R. Self-enforcing regional vaccination agreements. *J R Soc Interface.* (2016) 13:20150907. doi: 10.1098/rsif.2015.0907
60. Barrett S, Hoel M. Optimal disease eradication. *Environ Dev Econ.* (2007) 12:627–52. doi: 10.1017/S1355770X07003816
61. Barrett S. Global disease eradication. *J Eur Econ Assoc.* (2003) 1:591–600. doi: 10.1162/154247603322391224
62. Barrett S. Stop! the polio vaccination cessation game. *World Bank Econ Rev*. (2010) 24:361–85. doi: 10.1093/wber/lhq018
63. Barrett S. Polio eradication: strengthening the weakest links. *Health Aff.* (2009) 28:1079–90. doi: 10.1377/hlthaff.28.4.1079
64. Barrett S. The smallpox eradication game. *Public Choice.* (2006) 130:179–207. doi: 10.1007/s11127-006-9079-z
65. Omer SB, Enger KS, Moulton LH, Halsey NA, Stokley S, Salmon DA. Geographic clustering of nonmedical exemptions to school immunization requirements associations with geographic clustering of pertussis. *Am J Epidemiol*. (2008) 168:1389–96. doi: 10.1093/aje/kwn263
66. Cialdini RB, Kallgren CA, Reno RR. A focus theory of normative conduct: a theoretical refinement and reevaluation of the role of norms in human behavior. *Adv Exp Soc Psychol.* (1991) 24:201–34. doi: 10.1016/S0065-2601(08)60330-5
67. Centola D. The spread of behavior in an online social network experiment. *Science.* (2010) 329:1194–7. doi: 10.1126/science.1185231
68. Centola D. An experimental study of homophily in the adoption of health behavior. *Science.* (2011) 334:1269–72. doi: 10.1126/science.1207055
69. Conner M, Norman P. Health behaviour: current issues and challenges. *Psychol Heal.* (2017) 32:895–906. doi: 10.1080/08870446.2017.1336240

70. Agarwal V. A/H1N1 vaccine intentions in college students: an application of the theory of planned behavior. *J Am Coll Heal.* (2014) 62:416–24. doi: 10.1080/07448481.2014.917650
71. Babalola S, Lawan U. Factors predicting BCG immunization status in northern Nigeria: a behavioral-ecological perspective. *J Child Heal Care.* (2009) 13:46–62. doi: 10.1177/1367493508098380
72. Eng E, Naimoli J, Naimoli G, Parker KA, Lowenthal N. The acceptability of childhood immunization to Togolese mothers: a socio-behavioral perspective. *Health Educ Q.* (1991) 18:97–110. doi: 10.1177/109019819101800110
73. Brunson EK. The impact of social networks on parents' vaccination decisions. *Pediatrics.* (2013) 131:e1397–404. doi: 10.1542/peds.2012-2452
74. Thorpe EL, Zimmerman RK, Steinhart JD, Lewis KN, Michaels MG. Homeschooling parents' practices and beliefs about childhood immunizations. *Vaccine.* (2012) 30:1149–53. doi: 10.1016/j.vaccine.2011.12.019
75. Daley EM, Vamos CA, Thompson EL, Zimet GD, Rosberger Z, Merrell L, et al. The feminization of HPV: how science, politics, economics and gender norms shaped US HPV vaccine implementation. *Papillomavirus Res.* (2017) 3:142–8. doi: 10.1016/j.pvr.2017.04.004
76. Frew PM, Saint-Victor DS, Owens LE, Omer SB. Socioecological and message framing factors influencing maternal influenza immunization among minority women. *Vaccine.* (2014) 32:1736–44. doi: 10.1016/j.vaccine.2014.01.030
77. Sobo EJ. Social cultivation of vaccine refusal delay among Waldorf (Steiner) school parents. *Med Anthropol Q.* (2015) 29:381–99. doi: 10.1111/maq.12214
78. Attwell K, Meyer SB, Ward PR. The social basis of vaccine questioning refusal: a qualitative study employing bourdieu's concepts of 'capitals' 'habitus.' *Int J Environ Res Public Health.* (2018) 15:1044. doi: 10.3390/ijerph15051044
79. Reich JA. "We are fierce, independent thinkers and intelligent": social capital and stigma management among mothers who refuse vaccines. *Soc Sci Med.* (2018) 257:112015. doi: 10.1016/j.socscimed.2018.10.027
80. Kahan DM, Jenkins-Smith H, Braman D. Cultural cognition of scientific consensus. *J Risk Res.* (2011) 14:147–74. doi: 10.1080/13669877.2010.511246
81. Kahan DM, Peters E, Wittlin M, Slovic P, Ouellette LL, Braman D, et al. The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nat Clim Chang.* (2012) 2:732–5. doi: 10.1038/nclimate1547
82. Mesch GS, Schwirian KP. Confidence in government vaccination willingness in the USA. *Health Promot Int.* (2014) 30:213–21. doi: 10.1093/heapro/dau094
83. Kahan DM. Ideology, motivated reasoning, cognitive reflection. *Judgement Decis Mak.* (2013) 8:407–24. doi: 10.2139/ssrn.2182588
84. Kahan DM, Braman D, Cohen GL, Gastil J, Slovic P. Who fears the HPV vaccine, who doesn't, and why? an experimental study of the mechanisms of cultural cognition. *Law Hum Behav.* (2010) 34:501–6. doi: 10.1007/s10979-009-9201-0
85. Brewer MB. Perpetrators of prejudice; the psychology of prejudice: ingroup love outgroup hate? *J Soc Stud.* (1999) 55:429–44. doi: 10.1111/0022-4537.00126
86. Attwell K, Smith DT. Parenting as politics: social identity theory vaccine hesitant communities. *Int J Heal Gov.* (2017) 22:183–98. doi: 10.1108/IJHG-03-2017-0008
87. Huddy L. From social to political identity : a critical examination of social identity theory. *Polit Psychol.* (2001) 22:127–56. doi: 10.1111/0162-895X.00230
88. Attwell K, Ward PR, Meyer SB, Rokkas PJ, Leask J. "Do-it-yourself": vaccine rejection and complementary and alternative medicine (CAM). *Soc Sci Med.* (2018) 196:106–14. *doi: 10.1016/j.socscimed.2017.11.022*
89. Quinn SC, Hilyard KM, Jamison AM, An J, Hancock GR, Musa D, et al. The influence of social norms on flu vaccination among African American and white adults. *Health Educ Res.* (2017) 32:473–86. doi: 10.1093/her/cyx070
90. Jamison AM, Quinn SC, Freimuth VS. "You don't trust a government vaccine": narratives of institutional trust and influenza vaccination among African American and white adults. *Soc Sci Med.* (2019) 221:87–94. doi: 10.1016/j.socscimed.2018.12.020
91. Hirth JM, Fuchs EL, Chang M, Fernandez ME, Berenson AB. Variations in reason for intention not to vaccinate across time, region, by race/ethnicity. NIS-Teen. (2008–2016). *Vaccine.* (2019) 37:595–601. doi: 10.1016/j.vaccine.2018.12.017
92. Sadr W, Capps L. The challenge of minority recruitment in clinical trials for AIDS. *JAMA J Am Med Assoc.* (1992) 267:954–7. doi: 10.1001/jama.267.7.954
93. Ruijs WL M, Hautvast JL A, van IJzendoorn G., van Ansem WJ C, van der Velden K., Hulscher MEJ, et al. How orthodox protestant parents decide on the vaccination of their children: a qualitative study. *BMC Public Health.* (2012) 12:408. doi: 10.1186/1471-2458-12-408
94. McKee J. Holistic health and the critique of western medicine. *Soc Sci Med.* (1988) 26:775–84. doi: 10.1016/0277-9536(88)90171-2
95. Norris P, Chamberlain K, Dew K, Gabe J, Hodgetts D, Madden, et al. Public beliefs about antibiotics, infection and resistance: a qualitative study. *Antibiotics.* (2013) 2:465–76. doi: 10.3390/antibiotics2040465
96. de Melker RA, Touw-Otten FWM M, Kuyvenhoven MM. Transcultural differences in illness behaviour and clinical outcome: an underestimated aspect of general practice? *Fam Pract.* (1997) 14:472–7. doi: 10.1093/fampra/14.6.472
97. van Duijn H, Kuyenhoven M, Tudor Jones R, Butler C, Coenen S, Van Royen P. Patient's views on respiratory tract symptoms and antibiotics. *Br J Gen Pract.* (2003) 53:491–2.
98. Rosman S, Le Vaillant M, Schellevis F, Clerc P, Verheij R, Pelletier-Fleury N. Prescribing patterns for upper respiratory tract infections in general practice in France and in the Netherlands. *Eur J Public Health.* (2008) 18:312–6. doi: 10.1093/eurpub/ckm118
99. Grigoryan L, Burgerhof JG M, Degener JE, Deschepper R, Stålsby Lundborg C, Monnet DL, et al. Attitudes, beliefs and knowledge concerning antibiotic use and self-medication: a comparative European study. *Pharmacoepidemiol Drug Saf.* (2017) 16:1234–43. doi: 10.1002/pds.1479
100. Deschepper R, Vander Stichele RH, Haaijer-Ruskamp FM. Cross-cultural differences in lay attitudes and utilisation of antibiotics in a Belgian and a Dutch city. *Patient Educ Couns.* (2002) 48:161–9. doi: 10.1016/S0738-3991(02)00017-4
101. Zikmund-Fisher BJ, Windschitl PD, Exe N, Ubel PA. I'll do what they did: social norm information and cancer treatment decisions. *Patient Educ Couns.* (2011) 85:225–9. doi: 10.1016/j.pec.2011.01.031
102. Gaygisiz Ü, Lajunen T, Gaygisiz E. Socio-economic factors cultural values, national personality antibiotics use: a cross-cultural study among European countries. *J Infect Public Health.* (2017) 10:755–60. doi: 10.1016/j.jiph.2016.11.011
103. Broom A, Broom J, Kirby E, Adams J. The social dynamics of antibiotic use in an Australian hospital. *J Sociol.* (2016) 52:824–39. doi: 10.1177/1440783315594486
104. Degeling C, Johnson J, Iredell J, Nguyen AE, Norris JM, Turnidge JD, et al. Assessing the public acceptability of proposed policy interventions to reduce the misuse of antibiotics in Australia: a report on two community juries. *Heal Expect.* (2018) 21:90–9. doi: 10.1111/hex.12589
105. Gaarslev C, Yee M, Chan G, Fletcher-Lartey S, Khan R. A mixed methods study to understand patient expectations for antibiotics for an upper respiratory tract infection. *Antimicrob Resist Infect Control.* (2016) 5:39. doi: 10.1186/s13756-016-0134-3
106. Hawking MK D, Lecky DM, Touboul Lundgren P, Aldigs E, Abdulmajed H, Ioannidou E, et al. Attitudes and behaviours of adolescents towards antibiotics and self-care for respiratory tract infections: a qualitative study. *BMJ Open.* (2017) 7:e015308. doi: 10.1136/bmjopen-2016-015308
107. Harbarth S, Albrich W, Brun-Buisson C. Outpatient antibiotic use and prevalence of antibiotic-resistant pneumococci in France and Germany: a sociocultural perspective. *Emerg Infect Dis.* (2002) 8:1460–7. doi: 10.3201/eid0812.010533
108. Deschepper R, Grigoryan L, Stålsby Lundborg C, Hofstede G, Cohen J, van der Kelen G, et al. Are cultural dimensions relevant for explaining cross-national differences in antibiotic use in Europe? *BMC Health Serv. Res.* (2008) 8:123. doi: 10.1186/1472-6963-8-123
109. Kamekis A, Bertsias A, Moschandreas J, Petelos E, Papadakaki M, Tsiantou V, et al. Patients' intention to consume prescribed and non-prescribed medicines: a study based on the theory of planned behaviour in selected European countries. *J Clin Pharm Ther.* (2018) 43:26–35. doi: 10.1111/jcpt.12601

110. Kandeel A, El-Shoubary W, Hicks L, Fattah M, Dooling K, Lohiniva A, et al. Patient attitudes and beliefs and provider practices regarding antibiotic use for acute respiratory tract infections in Minya, Egypt. *Antibiotics.* (2014) 3:632–44. doi: 10.3390/antibiotics3040632
111. Kamat VR. Cultural interpretations of the efficacy and side effects of antimalarials in Tanzania. *Anthropol Med.* (2009) 16:293–305. doi: 10.1080/13648470903246854
112. Widayati A, Suryawati S, de Crespigny C, Hiller JE. Beliefs about the use of nonprescribed antibiotics among people in Yogyakarta City, Indonesia. *Asia Pacific J Public Heal.* (2012) 27:NP402–13. doi: 10.1177/1010539512445052
113. Pavyde E, Veikutis V, Mačiuliene A, Mačiulis V, Petrikonis K, Stankevičius E. Public knowledge, beliefs and behavior on antibiotic use and self-medication in Lithuania. *Int J Environ Res Public Health.* (2015) 12:7002–16. doi: 10.3390/ijerph120607002
114. Sawalha AF. Self-medication with antibiotics: a study in Palestine. *Int J Risk Saf Med.* (2008) 20:213–22. doi: 10.3233/JRS-2008-0445
115. Chan YY, Bin Ibrahim MA, Wong CM, Ooi CK, Chow A. Determinants of antibiotic prescribing for upper respiratory tract infections in an emergency department with good primary care access: a qualitative analysis. *Epidemiol Infect.* (2019) 147:e111. doi: 10.1017/S095026881800331X
116. Papoutsi C, Mattick K, Pearson M, Brennan N, Briscoe S, Wong G. Social and professional influences on antimicrobial prescribing for doctors-in-training: a realist review. *J Antimicrob Chemother.* (2017) 72:2418–30. doi: 10.1093/jac/dkx194
117. Walker AE, Grimshaw JM, Armstrong EM. Salient beliefs intentions to prescribe antibiotics for patients with a sore throat. *Br J Health Psychol.* (2001) 6:347–60. doi: 10.1348/135910701169250
118. Butler CC, Rollnick S, Pill R, Maggs-Rapport F, Stott N. Understanding the culture of prescribing: qualitative study of general practitioners' patients' perceptions of antibiotics for sore throats. *BMJ.* (1998) 317:637–42. doi: 10.1136/bmj.317.7159.637
119. Phadke VK, Bednarczyk RA, Salmon DA, Omer SB. Association between vaccine refusal vaccine-preventable diseases in the United States: a review of measles pertussis. *JAMA.* (2016) 315:1149–58. doi: 10.1001/jama.2016.1353
120. McDonald P, Limaye RJ, Omer SB, Buttenheim AM, Mohanty S, Klein NP, et al. Exploring California's new law eliminating personal belief exemptions to childhood vaccines and vaccine decision-making among homeschooling mothers in California. *Vaccine.* (2019) 37:742–50. doi: 10.1016/j.vaccine.2018.12.018
121. Chen RT, Rastogi SC, Mullen JR, Hayes SW, Cochi SL, Donlon JA, et al. The Vaccine Adverse Event Reporting System (VAERS). *Vaccine.* (1994) 12:542–50. doi: 10.1016/0264-410X(94)90315-8
122. Chen RT, Glasser JW, Rhodes PH, Davis RL, Barlow WE, Thompson RS, et al. Vaccine safety datalink project : a new tool for improving vaccine safety monitoring in the United States. *Pediatrics.* (1997) 99:765–73. doi: 10.1542/peds.99.6.765
123. Ventola CL. The antibiotic resistance crisis: part 1: causes and threats. *Pharm Ther.* (2015) 40:277–83.
124. Cormican M, Vellinga A. Existing classes of antibiotics are probably the best we will ever have. *BMJ.* (2012) 344:e3369. doi: 10.1136/bmj.e3369
125. World Health Organisation. *Global Action Plan on Antimicrobial Resistance*. Geneva: World Health Organisation (2015).
126. Whitney CG, Zhou F, Singleton J, Schuchat A. Benefits from immunization during the vaccines for children program era — United States, 1994–2013. *MMWR Morb MortalWkly Rep.* (2014) 63:352–5.
127. Guzman-Cottrill JA, Lancioni C, Eriksson C, Cho J.-Y, Liko J. Notes from the field: Tetanus in an unvaccinated child — Oregon, 2017. *MMWR Morb Mortal Wkly Rep.* (2019) 68:231–2. doi: 10.15585/mmwr.mm6809a3
128. Ozawa S, Portnoy A, Getaneh H, Clark S, Knoll M, Bishai D, et al. Modeling the economic burden of adult vaccine-preventable diseases in the United States. *Health Aff.* (2016) 35:2124–32. doi: 10.1377/hlthaff.2016.0462
129. Zhou F, Ortega-Sanchez IR, Guris D, Shefer A, Lieu T, Seward JF. An economic analysis of the universal varicella vaccination program in the United States. *J Infect Dis.* (2008) 197:S156–64. doi: 10.1086/522135
130. Omer SB, Pan WK Y, Halsey NA, Stokley S, Moulton LH, Navar AM, et al. Nonmedical exemptions to school immunization requirements: secular trends association of state policis with pertussis incidence. *JAMA.* (2006) 296:1757–63. doi: 10.1001/jama.296.14.1757
131. Michaelidis CI, Fine MJ, Lin CJ, Linder JA, Nowalk MP, Shields RK, et al. The hidden societal cost of antibiotic resistance per antibiotic prescribed in the United States: an exploratory analysis. *BMC Infect Dis.* (2016) 16:655. doi: 10.1186/s12879-016-1990-4
132. US Department of Health and Human Sciences, Centres for Disease Control and Prevention. *Antibiotic Resistance Threats in the United States* (2013). Available online at: https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf
133. Varma JK, Greene KD, Ovitt J, Barrett TJ, Medalla F, Angulo FJ. Hospitalization and antimicrobial resistance in salmonella outbreaks, 1984-2002. *Emerg Infect Dis.* (2005) 11:943–6. doi: 10.3201/eid1106.041231
134. Smith R, Coast J. The true cost of antimicrobial resistance. *BMJ.* (2013) 346:f1493. doi: 10.1136/bmj.f1493
135. Committee to Review Adverse Effects of Vaccines, Institute of Medicine. *Adverse Effects of Vaccines: Evidence and Causality*. Washington, DC: The National Academies Press (2012).
136. Bohlke K, Davis RL, Marcy SM, Braun MM, Destefano F, Black SB, et al. Risk of anaphylaxis after vaccination of children and adolescents. *Pediatrics.* (2003) 112:815–20. doi: 10.1542/peds.112.4.815
137. Siddiqui M, Salmon DA, Omer SB. Epidemiology of vaccine hesitancy in the United States. *Hum Vaccin Immunother.* (2013) 9:2643–8. doi: 10.4161/hv.27243
138. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. *Proc R Soc Lond A Contain Pap Math Phys Character.* (1927) 115:700–21. doi: 10.1098/rspa.1927.0118
139. Ma J, Earn DJ. Generality of the final size formula for an epidemic of a newly invading infectious disease. *Bull Math Biol.* (2006) 68:679–702. doi: 10.1007/s11538-005-9047-7
140. Holmberg SD, Solomon SL, Blake PA. Health and economic impacts of antimicrobial resistance. *Rev Infect Dis.* (1987) 9:1065–78. doi: 10.1093/clinids/9.6.1065
141. Cosgrove SE, Carmeli Y. The impact of antimicrobial resistance on health and economic outcomes. *Clin Infect Dis.* (2003) 36:1433–7. doi: 10.1086/375081
142. Bhatt AP, Redinbo MR, Bultman SJ. The role of the microbiome in cancer development therapy. *CA Cancer J Clin.* (2017) 67:326–44. doi: 10.3322/caac.21398
143. Rolfes MA, Foppa IM, Garg S, Flannery B, Brammer L, Singleton JA, et al. Annual estimates of the burden of seasonal influenza in the United States: a tool for strengthening influenza surveillance and preparedness. *Influenza Other Resp Viruses.* (2018) 12:132–7. doi: 10.1111/irv.12486
144. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of Influenza A viruses. *Microbiol Rev.* (1992) 56:152–79. doi: 10.1128/MMBR.56.1.152-179.1992
145. Carrat F, Flahault A. Influenza vaccine: the challenge of antigenic drift. *Vaccine.* (2007) 25:6852–62. doi: 10.1016/j.vaccine.2007.07.027
146. Omer SB, Yildirim I. Influenza vaccine effectiveness: a glass both half full and half empty. *Clin Infect Dis.* (2016) 63:1574–6. doi: 10.1093/cid/ciw637
147. Saad-Roy CM, McDermott AB, Grenfell BT. Dynamic perspectives on the search for a universal influenza vaccine. *J Infect Dis.* (2019) 219:S46–56. doi: 10.1093/infdis/jiz044
148. Viboud C, Gostic K, Nelson MI, Price GE, Perofsky A, Sun K, et al. Beyond clinical trials: evolutionary and epidemiological considerations for development of a universal influenza vaccine. *PLoS Pathog.* (2020) 16:e1008583. doi: 10.1371/journal.ppat.1008583
149. Arinaminpathy N, Riley S, Barclay WS, Saad-Roy C, Grenfell B. Population implications of the deployment of novel universal vaccines against epidemic and pandemic influenza. *J R Soc Interface.* (2020) 17:20190879. doi: 10.1098/rsif.2019.0879
150. Fiebach NH, Viscoli CM. Patient acceptance of influenza vaccination. *Am J Med.* (1991) 91:393–400. doi: 10.1016/0002-9343(91)90157-S
151. Hofmann F, Ferracin C, Marsh G, Dumas R. Influenza vaccination of healthcare workers: a literature review of attitudes and beliefs. *Infection.* (2006) 34:142–7. doi: 10.1007/s15010-006-5109-5
152. Poland GA, Jacobsen SJ. Influenza vaccine Guillain–Barré syndrome, chasing zero. *Vaccine.* (2012) 40:5801–3. doi: 10.1016/j.vaccine.2012.06.093

153. Ferdinands JM, Olsho LE W, Agan AA, Bhat N, Sullivan RM, Hall M, et al. Effectiveness of influenza vaccine against life-threatening RT-PCR-confirmed influenza illness in US children, 2010 – 2012. *J Infect Dis.* (2014) 210:674–83. doi: 10.1093/infdis/jiu185
154. Flannery B, Reynolds SB, Blanton L, Santibanez TA, O'Halloran AP, Lu J, et al. Influenza vaccine effectiveness against pediatric deaths: 2010–2014. *Pediatrics*. (2017) 139:e20164244. doi: 10.1542/peds.2016-4244
155. Thompson MG, Pierse N, Huang QS, Prasad N, Duque J, Newbern EC, et al. Influenza vaccine effectiveness in preventing influenza-associated intensive care admissions and attenuating severe disease among adults in New Zealand 2012 – 2015. *Vaccine.* (2018) 36:5916–25. doi: 10.1016/j.vaccine.2018.07.028
156. Rothberg MB, Haessler SD, Brown RB. Complications of viral influenza. *Am J Med*. (2008) 121:258–64. doi: 10.1016/j.amjmed.2007.10.040
157. Arinaminpathy N, Kim IK, Gargiullo P, Haber M, Foppa IM, Gambhir M, et al. Estimating direct and indirect protective effect of influenza vaccination in the United States. *Am J Epidemiol.* (2017) 186:92–100. doi: 10.1093/aje/kwx037
158. Arinaminpathy N, Ratmann O, Koelle K, Epstein SL, Price GE, Viboud C, et al. Impact of cross-protective vaccines on epidemiological and evolutionary dynamics of influenza. *Proc Natl Acad Sci USA*. (2012) 109:3173–7. doi: 10.1073/pnas.1113342109
159. de Vries RD, Duprex WP, de Swart RL. Morbillivirus infections: an introduction. *Viruses*. (2015) 7:699–706. doi: 10.3390/v7020699
160. Orenstein WA, Papania MJ, Wharton ME. Measles elimination in the United States. *J Infect Dis*. (2004) 189:S1–3. doi: 10.1086/377693
161. Omer SB, Yildrim I. Further evidence of MMR vaccine safety: scientific and communications considerations. *Ann Intern Med.* (2019) 170:567–8. doi: 10.1093/med/9780199343560.003.0038
162. Sundaram ME, Guterman LB, Omer SB. The true cost of measles outbreaks during the postelimination era. *JAMA*. (2019) 321:1155–6. doi: 10.1001/jama.2019.1506
163. Lo NC, Hotez PJ. Public health and economic consequences of vaccine hesitancy for measles in the United States. *JAMA Pediatr*. (2017) 171:887–92. doi: 10.1001/jamapediatrics.2017.1695
164. Ghebrehewet S, Thorrington D, Farmer S, Kearney J, Blissett D, McLeod H, et al. The economic cost of measles: healthcare, public health and societal costs of the 2012-13 outbreak in Merseyside, UK. *Vaccine.* (2016) 34:1823–31. doi: 10.1016/j.vaccine.2016.02.029
165. Salmon DA, Haber M, Gangarosa EJ, Phillips L, Smith NJ, Chen RT. Health consequences of religious philosophical exemptions from immunization laws. *JAMA*. (1999) 281:47–54. doi: 10.1001/jama.282.1.47
166. Mina MJ, Metcalf CJE, de Swart RL, Osterhaus ADM, Grenfell BT. Long-term measles-induced immunomodulation increases overall childhood infectious disease mortality. *Science.* (2015) 348:694–700. doi: 10.1126/science.aaa3662
167. Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG. The evolutionary history of methicillin-resistant Staphylococcus aureus (MRSA). *Proc Natl Acad Sci USA*. (2002) 99:7687–92. doi: 10.1073/pnas.122108599
168. Cosgrove SE. The relationship between antimicrobial resistance and patient outcomes: mortality, length of hospital stay, and health care costs. *Clin Infect Dis*. (2006) 42:S82–9. doi: 10.1086/499406
169. Mozzillo KL, Ortiz N, Miller LG. Patients with meticillin-resistant Staphylococcus aureus infection: twenty-first century lepers. *J Hosp Infect*. (2010) 75:132–4. doi: 10.1016/j.jhin.2009.10.031
170. Kanerva M, Blom M, Tuominen U, Kolho EV, Anttila J, Vaara M, et al. Costs of an outbreak of meticillin-resistant *Staphylococcus aureus*. *J Hosp Infect*. (2007) 66:22–8. doi: 10.1016/j.jhin.2007.02.014
171. Horowitz JB, Moehring HB. How property rights and patents affect antibiotic resistance. *Health Econ.* (2004) 13:575–83. doi: 10.1002/hec.851
172. Avorn J, Solomon DH. Cultural and economic factors that (mis) shape antibiotic use: The nonpharmacologic basis of therapeutics. *Ann Intern Med*. (2000) 133:128–35. doi: 10.7326/0003-4819-133-2-200007180-00012
173. Jacobson TA, Thomas DM, Morton FJ, Offutt G, Shevlin J, Ray S. Use of a low-literary patient education tool to enhance pneumococcal vaccination rates. *J Am Med Assoc.* (1999) 282:646–50. doi: 10.1001/jama.282.7.646
174. Borgida E, Nisbett RE. The differential impact of abstract vs concrete information on decisions. *J Appl Soc Psychol.* (1977) 7:258–71. doi: 10.1111/j.1559-1816.1977.tb00750.x
175. Leventhal H, Singer R, Jones S. Effects of fear and specificity of recommendation upon attitudes and behavior. *J Pers Soc Psychol.* (1965) 2:20–9. doi: 10.1037/h0022089
176. Chandy SJ, Mathai E, Thomas K, Faruqui AR, Holloway K, Lundborg, et al. Antibiotic use and resistance: perceptions and ethical challenges among doctors, pharmacists and the public in Vellore, South India. *Indian J Med Ethics*. (2013) 10:20–7. doi: 10.20529/IJME.2013.005
177. Nyhan B, Reifler J, Richey S, Freed GL. Effective messages in vaccine promotion: a randomised trial. *Pediatrics*. (2014) 133:e835–42. doi: 10.1542/peds.2013-2365
178. Cialdini RB, Demaine LJ, Sagarin BJ, Barrett DW, Rhoads K, Winter PL. Managing social norms for persuasive impact. *Soc Influ.* (2006) 1:3–15. doi: 10.1080/15534510500181459
179. Farrar J. *Let's Get Real. No Vaccine Will Work as if by Magic, Returning Us to 'Normal.' Guard.* (2020). Available online at: https://www.theguardian.com/commentisfree/2020/sep/06/lets-get-real-no-vaccine-will-work-as-if-by-magic-returning-us-to-normal.

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.

frontiers in Public Health

ORIGINAL RESEARCH
published: 23 December 2020
doi: 10.3389/fpubh.2020.573925

Demographic and Clinical Characteristics of Early Travel-Associated COVID-19 Cases

Reham M. Marei[1], Mohamed M. Emara[2], Omar M. Elsaied[2], Gheyath K. Nasrallah[1,3], Tawanda Chivese[2], Hamad E. Al-Romaihi[4], Mohamed H. Althani[4], Asmaa A. Al Thani[1,3], Elmoubasher A. Farag[4] and Hadi M. Yassine[1,4]**

[1] Biomedical Research Center, Qatar University, Doha, Qatar, [2] College of Medicine-QU Health, Qatar University, Doha, Qatar, [3] College of Health Sciences-QU Health, Qatar University, Doha, Qatar, [4] Public Health Department, Ministry of Public Health, Doha, Qatar

Background: SARS-CoV-2 continues to claim hundreds of thousands of people's lives. It mostly affects the elderly and those with chronic illness but can also be fatal in younger age groups. This article is the first comprehensive analysis of the epidemiological and clinical outcomes of the travel-associated SARS-CoV-2 cases until April 19, 2020.

Methods: Demographic and clinical data of travel-associated SARS-CoV-2 cases were collected for the period between January 16, 2020 and April 19, 2020. More than one hundred and eighty databases were searched, including the World Health Organization (WHO) database, countries' ministries websites, and official media sites. Demographic and clinical data were extracted and analyzed.

Results: A total of 1,186 cases from 144 countries meeting the inclusion criteria were reported and included in the analysis. The mean age of the cases was 44 years, with a male to female ratio of 1.6:1. Travel-associated cases originated from more than 40 countries, with China, Italy, and Iran reporting the highest numbers at 208, 225, and 155, respectively. Clinical symptoms varied between patients, with some reporting symptoms during the flights (117 cases; 9.87%). A total of 312 (26.31%) cases were hospitalized, of which 50 cases (4.22%) were fatal.

Conclusion: Major gaps exist in the epidemiology and clinical spectrum of the COVID-19 travel-associated cases due to a lack of reporting and sharing data of many counties. The identification and implementation of methodologies for measuring traveler's risk to coronavirus would help in minimizing the spread of the virus, especially in the next waves.

Keywords: coronavirus, traveler's risk to infection, clinical outcome, epidemiology, COVID-19

OPEN ACCESS

Edited by:
Matteo Convertino, Hokkaido University, Japan

Reviewed by:
José Eduardo Levi, University of São Paulo, Brazil
William Davis, Centers for Disease Control and Prevention (CDC), United States

***Correspondence:**
Elmoubasher A. Farag
eabdfarag@moph.gov.qa
Hadi M. Yassine
hyassine@qu.edu.qa

Specialty section:
This article was submitted to Infectious Diseases - Surveillance, Prevention and Treatment, a section of the journal Frontiers in Public Health

Received: *18 June 2020*
Accepted: *30 November 2020*
Published: *23 December 2020*

Citation:
Marei RM, Emara MM, Elsaied OM, Nasrallah GK, Chivese T, Al-Romaihi HE, Althani MH, Al Thani AA, Farag EA and Yassine HM (2020) Demographic and Clinical Characteristics of Early Travel-Associated COVID-19 Cases. Front. Public Health 8:573925. doi: 10.3389/fpubh.2020.573925

INTRODUCTION

Out of seven coronaviruses known to infect humans, four are seasonal (229E, NL63, OC43, and HKU1) and cause common respiratory infections similar to influenza (1). Additionally, three novel coronaviruses of animal origin (zoonotic) have emerged in the last two decades and caused unprecedented outbreaks in human: the Severe Acute Respiratory Syndrome (SARS-CoV) that emerged in 2003 (2–4), the Middle East Respiratory Syndrome (MERS-CoV) that emerged in 2012

(5, 6), and the SARS-CoV-2 that emerged at the end of 2019 (7–10). All of the zoonotic viruses were associated with server respiratory illnesses.

SARS-CoV-2, which is the causative agent of Coronavirus disease 2019 (COVID-19), was first identified in Wuhan, China, in December 2019 (10). The virus then rapidly spread to other countries around the globe. The first ten countries which reported the virus after China were Thailand, Japan, South Korea, the USA, Singapore, Vietnam, Australia, France, and Malaysia (11, 12). The rapid spread of the virus to the different continents evoked WHO to declare the pandemic stage on March 11, 2020, making it the first zoonotic-origin coronavirus to reach this stage (13).

The clinical spectrum of COVID-19 ranges from asymptomatic/mild upper respiratory tract illness in about 80% of the patients to severe viral pneumonia with respiratory failure and even death. The major risk factors associated with disease severity include chronic illness, older age, and genetic dispositions. The incubation period of COVID ranges from 1 to 14 days (median 5 days), during which infected patients can transmit the virus even before experiencing any symptoms (7, 10, 14, 15).

The risk of acquiring an infectious agent during travel is relatively high. As per the 2015 records of the World Tourism Organization (WTO), 1.2 billion international tourist arrivals were documented (16). These figures are projected to increase further and reach 1.4 billion by 2020 and 1.8 billion by 2030, signifying the crucial public health impact that travel will have on global health security (17). Consequently, travelers represent a cohort of epidemiological importance because of the associated dynamicity, exposure to infectious diseases, and risk of circulating infections among populations. This concern has been further fueled by the rapid emergence and spread of various agents such as Ebola, MERS-CoV, Zika, and most recently, the SARS-CoV-2, although each of them is characterized by a different mode of transmission (18, 19).

Although MERS-CoV was identified in 2012, only 2,494 cases have been reported so far (20–25). Further, SARS-CoV-1 was contained in <2 years after affecting more than 8,000 cases, including 800 deaths (26, 27). As of October 15, 2020, SARS-CoV-2 had affected more than 40 millions worldwide, including at least one million reported deaths (fatality rate of about 2.5%), making it the most transmissible coronaviruses (zoonotic) identified so far (28). Although the significance of SARS-CoV-2 spread and transmission among populations is well recognized, there is a noticeable lack of the information about travel-associated cases. In this study, we report on SARS-CoV-2 travel-associated cases demography, clinical outcome, and risk factors associated with virus transmission.

METHODOLOGY

The aim of this study was to analyze the epidemiological and clinical characteristics of the early SARS-CoV-2 cases or so-called travel-associated cases. Demographic and clinical data of travel-associated SARS-CoV-2 cases were collected for the period between January 16 and April 19, 2020. One hundred eighty-one cases databases were searched, including the World Health Organization (WHO) database, countries' ministries websites, and official media sites (**Supplementary Table 1**). A comprehensive data collection sheet of the individual study parameters was prepared using Microsoft Excel. The following information was extracted and summarized: age, gender, the origin of cases, occupation, destination, comorbidities, hospitalization, and clinical manifestation, and clinical outcome, symptoms appearance during travel, and the timeline for imported cases. Data were presented as medians, means, standard deviation (SD), percentage, and cumulative numbers. The data set was not complete for all reported cases, and hence, rates were calculated according to the available data. It is worth noting that we had collected the data to the best of our knowledge and capacity, relying on the resources listed in **Supplemental Table 1**.

RESULTS

Demographic Characteristics

The demographic characteristics of the imported cases up to April 19, 2020 are described in **Table 1**. A total of 1,186 cases meeting the inclusion criteria were reported by that date. More than half of the data were missing for the age group (65.85%). People of all ages were affected, with a mean age of the cases being 44 (SD 18) years. The majority of cases were reported for those older than 20 years, while those <20 years of age represented only 3.20% of the documented cases. Higher cases were reported in males (27.82%) compared to females (17.37%), with a female to male ratio of (1:1.6).

Cases Occupation

The cases had varying occupations as described in **Table 2**, with 1,143 cases occupation (96.37%) being unidentified. For the available data, students represented the highest percentage of (1.01%), followed by Doctors (0.84%) and Business staff (0.76%).

TABLE 1 | Demographic characteristics of travel-associated cases as of April 19 2020 (n = 1,186).

Demographics	Criteria	Number (%)
Age, years	Mean (SD)	44 (SD = 18)
Age, years	Median (IQR)	42
Age group, *n* (%)	0–20	38 (3.20)
	20–40	126 (10.62)
	40–60	143 (12.06)
	60–80	78 (6.58)
	>80	20 (1.69)
	Unspecified	781 (65.85)
Gender, *n* (%)	Female	206 (17.37)
	Male	330 (27.82)
	Not specified	650 (54.67)

Timeline of SARS-CoV-2 Spread Across the Globe

The timeline of SARS-CoV-2 spread to different countries between January 13 and April 11, 2020, is shown in **Figure 1**. The first ten countries to report the virus after China were: Thailand, Japan, South Korea, the USA, Singapore, Vietnam, Australia, France, and Malasia. Until February 27, 2020, China was the main hub for exporting the virus to 23 countries. After the mid-February, Italy and Iran became the epicenters for Europe and the Middle East, respectively. On March 17, China reported the first travel-associated case from abroad, reaching 264 cases by April 19, 2020. Imported cases per region and country are shown in **Table 3**.

TABLE 2 | Cases occupation ($N = 1186$).

Occupation	*N* (%)
Student	12 (1.01)
Nurse	4 (0.34)
Business staff	9 (0.76)
Professor	2 (0.17)
Doctor	10 (0.84)
Music producer	1 (0.08)
Military staff	2 (0.17)
Office worker	3 (0.25)
Unspecified	1,143 (96.37)

Origin of Travel-Associated Cases

A critical part of our analysis was to identify the origin of travel-associated cases. The majority of travel-associated cases originated from Italy 255 (21.50%), followed by China 208 (17.53%), Iran 155(13.07%), and Diamond Princess Cruise ship 141 (11.89%). Few other cases were reported from other countries, including the USA 72 (6.07%), UK, and France 46 (3.88%). Twenty-three subjects (1.94%) had missing data of the place of origin (**Table 4**).

Clinical Presentation and Symptoms

Most of the travel-associated cases presented with fever (45.87%), followed by cough (38.69%), flu-like symptoms (6.22%), sore throat (1.60%), and pneumonia (1.01%). Several cases ($n = 117$; 9.87%) reported symptoms during the flight from China. For the non-symptomatic cases on the flights (48; 4.05 %), symptoms were developed between 1 and 22 days after travel (mean = 7.4; Median = 7, SD = 4.63) (**Table 5**).

Comorbidities, Hospitalization, Severity, and Outcome of Travel-Associated Cases

One thousand ninety (91.91%) of the imported cases had unspecified data about former health conditions. About 32 (2.70 %) of the cases were apparently healthy. Comorbidities were reported in 96 (8.09 %) cases only. Comorbidities included hypertension ($n = 19$), cardiovascular diseases ($n = 12$), Diabetes mellitus ($n = 12$), cancer ($n = 10$), pulmonary diseases ($n = 9$) and kidney diseases ($n = 2$) (**Table 6**).

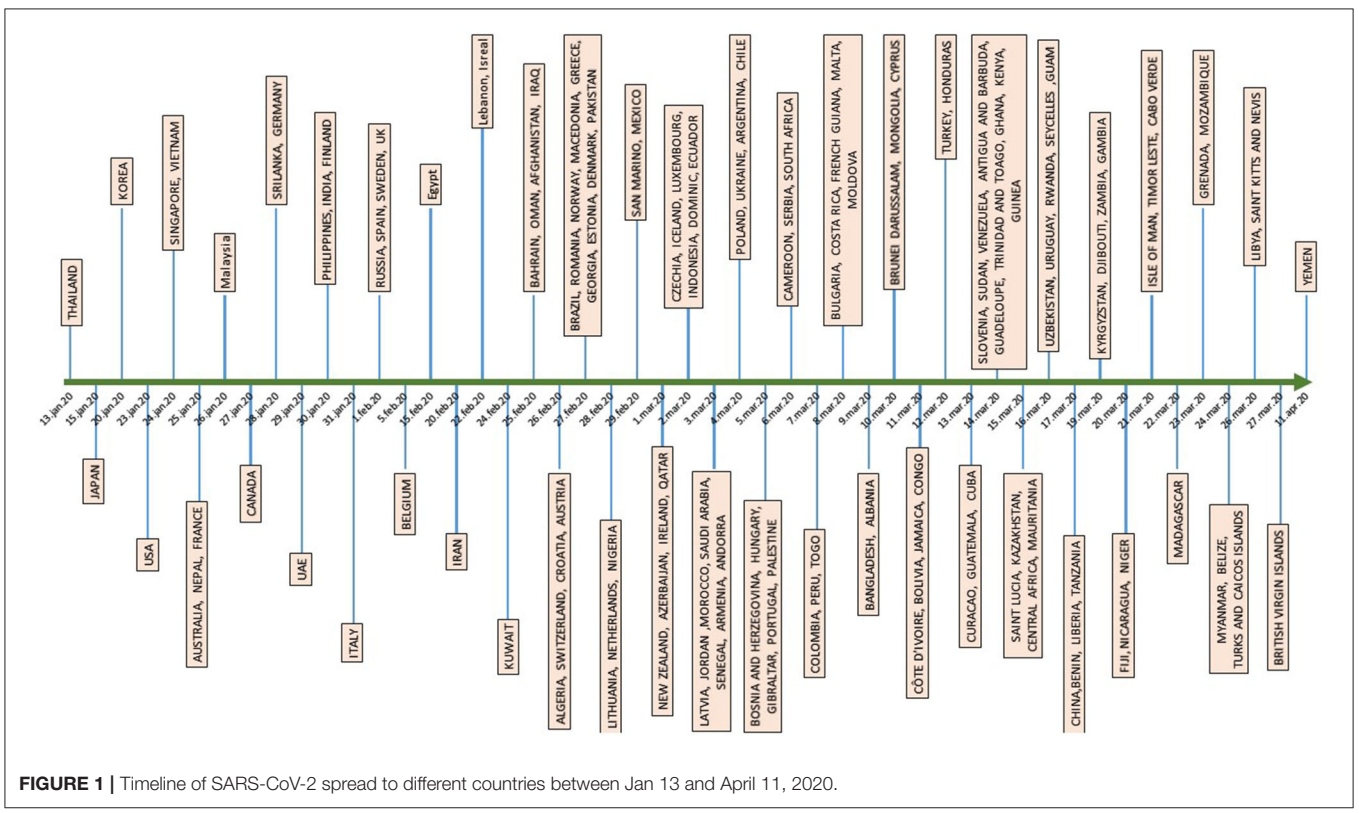


FIGURE 1 | Timeline of SARS-CoV-2 spread to different countries between Jan 13 and April 11, 2020.

TABLE 3 | Travel-associated cases per region, country, and territory.

Region, country, territory	*N* (%)
Western Pacific Region	421
Malaysia	11 (0.93)
New Zealand	3 (0.25)
China	264 (22.26)
Australia	33 (2.78)
Korea	14 (1.18)
Philippines	9 (0.76)
Singapore	32 (2.61)
Vietnam	26 (2.19)
Japan	24 (2.02)
Mongolia	2 (0.17)
Fiji	1 (0.08)
Brunei Darussalam	2 (0.17)
Territories	2
Guam	2 (0.17)
European Region	294
Spain	17 (1.43)
Italy	6 (0.50)
Germany	2 (0.17)
The United Kingdom	27 (2.28)
France	3 (0.25)
Turkey	2 (0.17)
Belgium	15 (1.26)
Netherlands	11 (0.93)
Switzerland	6 (0.50)
Portugal	6 (0.50)
Ireland	3 (0.25)
Austria	4 (0.34)
Sweden	13 (1.01)
Israel	10 (0.84)
Romania	19 (1.60)
Denmark	7 (0.59)
Norway	7 (0.59)
Serbia	1 (0.08)
Ukraine	1 (0.08)
Poland	4 (0.34)
Luxembourg	2 (0.17)
Finland	5 (0.42)
Hungary	1 (0.08)
Greece	5 (0.42)
Moldova	5 (0.42)
Croatia	5 (0.42)
Estonia	2 (0.17)
Azerbaijan	11 (0.93)
Slovenia	5 (0.42)
Lithuania	6 (0.50)
Uzbekistan	6 (0.50)
Bosnia& Herzegovina	6 (0.50)
Armenia	5 (0.42)
Kazakhstan	8 (0.67)
Andorra	1 (0.08)
Bulgaria	1 (0.08)
Kyrgyzstan	9 (0.76)
Albania	3 (0.25)
Macedonia	3 (0.25)
Iceland	13 (1.01)
Russia	5 (0.42)
Czech Republic	4 (0.34)
Cyprus	3 (0.25)
Latvia	4 (0.34)
Malta	8 (0.67)
Georgia	4 (0.34)
Territories	3
Isle of Man	1 (0.08)
Gibraltar	2 (0.17)
South-East Asia Region	79
Thailand	38 (3.20)
India	10 (0.84)
Indonesia	2 (0.17)
Bangladesh	3 (0.25)
Myanmar	2 (0.17)
Nepal	4 (0.34)
Timor-Leste	1 (0.08)
Taiwan	7 (0.59)
sir lanka	12 (1.01)
Eastern Mediterranean Region	151
UAE	10 (0.84)
Iran	4 (0.34)
Kuwait	42 (3.54)
Bahrain	10 (0.84)
Afghanistan	2 (0.17)
Oman	18 (1.52)
Iraq	18 (1.52)
Yemen	1 (0.08)
Lebanon	10 (0.84)
Qatar	7 (0.59)
Saudi Arabia	9 (0.76)
Tunisia	4 (0.34)
Jordan	1 (0.08)
Libya	1 (0.08)
Pakistan	8 (0.67)
Morocco	2 (0.17)
Sudan	1 (0.08)
Egypt	2 (0.17)
Djibouti	1 (0.08)
Territories	2
State of Palestine	2 (0.17)
Region of the Americas	153
USA	68 (5.73)
Canada	21 (1.77)
Brazil	4 (0.34)
Peru	6 (0.50)

(Continued)

TABLE 3 | Continued

Region, country, territory	*N* (%)
Argentina	2 (0.17)
Chile	1 (0.08)
Colombia	3 (0.25)
Mexico	8 (0.67)
Bolivia	4 (0.34)
Costa Rica	1 (0.08)
Cuba	1 (0.08)
Dominican Republic	2 (0.17)
Ecuador	1 (0.08)
Uruguay	4 (0.34)
Guatemala	2 (0.17)
Venezuela	6 (0.50)
Jamaica	2 (0.17)
Trinidad and Tobago	3 (0.25)
Honduras	5 (0.42)
Antigua & Barbuda	3 (0.25)
Saint Kitts and Nevis	2 (0.17)
Grenada	1 (0.08)
Belize	1 (0.08)
Nicaragua	1 (0.08)
saint Lucia	1 (0.08)
Territories	8
French Guiana	3 (0.25)
Curacao	1 (0.08)
British Virgin Islands	2 (0.17)
Turks and Caicos Islands	1 (0.08)
Guadeloupe	1 (0.08)
African Region	72
Congo	5 (0.42)
Cameroon	1 (0.08)
Nigeria	4 (0.34)
Madagascar	8 (0.67)
South Africa	20 (1.69)
guinea	1 (0.08)
Ghana	4 (0.34)
Niger	1 (0.08)
Kenya	2 (0.17)
Algeria	2 (0.17)
Tanzania	1 (0.08)
Senegal	1 (0.08)
Togo	1 (0.08)
Liberia	1 (0.08)
Cabo Verde	1 (0.08)
Zambia	3 (0.25)
Rwanda	1 (0.08)
Benin	1 (0.08)
Central African Republic	1 (0.08)
Mozambique	2 (0.17)
Seychelles	2 (0.17)
Mauritania	8 (0.67)
Gambia	1 (0.08)
Côte d'Ivoire	1 (0.08)

TABLE 4 | Origin of travel-associated cases (*N* = 1,186).

Region, Country, Territory	*N* (%)
Western Pacific Region	231
China	208 (17.53)
Korea	4 (0.34)
Japan	8 (0.67)
Australia	4 (0.34)
Philippines	2 (0.17)
Singapore	2 (0.17)
Malaysia	3 (0.25)
European Region	507
Spain	44 (3.71)
Italy	255 (21.50)
Germany	23 (1.94)
France	46 (3.88)
Turkey	11 (0.93)
Belgium	22 (1.85)
Netherlands	2 (0.17)
Switzerland	18 (1.52)
Portugal	2 (0.17)
Ireland	3 (0.25)
Austria	19 (1.60)
Israel	2 (0.17)
UK	46 (3.88)
Russia	2 (0.17)
Czech Republic	1 (0.08)
Poland	1 (0.08)
Norway	2 (0.17)
Denmark	1 (0.08)
Greece	4 (0.34)
Hungary	2 (0.17)
Azerbaijan	1 (0.08)
South-East Asia Region	201
India	7 (0.59)
Indonesia	1 (0.08)
Thailand	1 (0.08)
Iran	155 (13.07)
Saudi Arabia	2 (0.17)
Pakistan	2 (0.17)
UAE	17 (1.43)
Qatar	1 (0.08)
Egypt	12 (1.01)
Morocco	1 (0.08)
Iraq	2 (0.17)
Oman	2 (0.17)
Region of the Americas	75
USA	72 (6.07)
Canada	1 (0.08)
Panama	1 (0.08)
Colombia	1 (0.08)
African Regions	8
South Africa	4 (0.34)
Burkina Faso	1 (0.08)
Mauritius	1 (0.08)
Senegal	4 (0.34)
Diamond Princess cruise ship	141 (11.89)
Unspecified	23 (1.94)

TABLE 5 | Symptoms at presentation ($N = 1,186$).

Symptoms	Number (%)
Cough	460 (38.69)
Fever	544 (45.87)
Flu-like symptoms	74 (6.22)
pneumonia	12 (1.01)
Gastrointestinal tract	1 (0.08)
Sore throat	19 (1.60)
Unspecified	76 (6.40)
Symptomatic during flight	117 (9.87)
Non-symptomatic during flight	48 (4.05)
Nonspecific	1,021 (86.09)
Average days to develop symptoms after arrival	1–22 day
For days to develop symptoms	Mean = 7.4
For days to develop symptoms	Median = 7
For days to develop symptoms	SD = 4.63

TABLE 6 | Comorbidities in travel-associated cases ($N = 1,186$).

Comorbidity diseases	*N* (%)
Hypertension	19 (1.60)
Cardiovascular disease	12 (1.01)
Diabetes mellitus	12 (1.01)
Pulmonary disease	9 (0.76)
Cancer	10 (0.84)
Kidney disease	2 (0.17)
Healthy	32 (2.70)
Unspecified	1,090 (91.91)

TABLE 7 | Reasons for hospitalization of travel-associated cases ($N = 1,186$).

Reason	*N* (%)
Pneumonia	20 (1.69)
Tested positive (isolation)	256 (21.59)
Symptomatic	35 (2.95)
Already admitted due to a different condition	1 (0.08)
Unspecified	874 (73.69)

Hospitalization data were reported for 312 (26.31 %) cases. The majority were hospitalized because they "tested positive" for COVID19 infection (21.59%). Symptoms of COVID-19 was the second leading cause of hospitalization (2.95%), while some were hospitalized for developing pneumonia (1.69%). Other reasons for hospitalization are shown in **Table 7**.

Out of 1,186 analyzed cases, 1,049 (88.45%) had a mild illness, and 135 (11.38%) were sever or critical cases. The fatality was reported in 50 (4.22%) cases (**Table 8**).

TABLE 8 | Cases severity and outcome ($N = 1,186$).

Case severity	*N* (%)
Mild	1,049 (88.45)
Sever	135 (11.38)
CASE OUTCOME	
Recovery	400 (33.73)
Death	50 (4.22)
Unspecified	736 (62.06)

DISCUSSION

Emerging infections continue to be a major threat to the human population. Several major outbreaks and pandemics have been reported in the past two decades due to emerging or re-emerging infections, including but not limited to the SARS outbreak in 2002 (Affecting >30 countries and ∼9,000 people), pandemic H1N1 in 2009 (Worldwide; millions), MERS in 2012 (27 countries; ∼2,500), Ebola in 2013 (10 countries; 30,000), and Zika in 2015 (87 countries; >100,000) (29–32).

The intrusion and spread of novel pathogens into the human population is attributed to several factors, majorly humans' behavior, social and environmental changes. In the past 20 years, six new coronaviruses emerged in the human population: SARS (2003), NL63 (2004), HKU1 (2005), MERS (2012), and SARS-CoV-2 (2019) (33). While SARS-CoV-1 was contained in <2 years, little is known about how the NL63 and HKU1 had emerged and spread in the human population. Both viruses (NL63 and HKU1) continue to circulate in humans, and they constitute up to 30% of the respiratory infections along with OC43 and 229E coronaviruses. However, we lack knowledge about the clinical manifestations in the primary patients who acquired these viruses. Unlike the other highly pathogenic coronaviruses, SARS-CoV-1, and MERS-CoV, only the SARS-CoV-2 reached the pandemic stage in a relatively short period (11). Within <11 months of its emergence, the virus affected more than 40 million, including more than one million confirmed deaths (12, 34).

Studies have shown that infected individuals with SARS-CoV-2 could spread the virus several days before clinical symptoms appear, and many of which do not even present any clinical manifestations (35). Accordingly, the risk of spread and transmission from these two groups is considered high. A recent modeling study suggests that asymptomatic persons might be the major drivers for the growth of the virus nationally and globally (36). This partially explains the late detection of the virus in China and, possibly, the reason for the widespread of the virus globally. Besides, a high viral load close to the onset of the symptoms suggests that the virus can be easily transmissible at an early stage of the infection (36).

The SARA-CoV-2 emerged in Wuhan, China, in December 2019. By February 1, 2020, the virus was reported in Thailand, Japan, South Korea, USA, Vietnam, Singapore, Australia, France, Nepal, Malaysia, Canada, Cambodia, Srilanka, Germany, United Arab Emirates, Philippines, India, Finland, and Italy

(chronological order) (11, 12). Italy and Iran reported the first cases on January 31 and February 20, becoming the epicenter for Europe and the Middle East, respectively. By February 26, 2020, the virus massively spread to many countries, resulting in declare of the pandemic on March 11, 2020.

Although the reproduction number (R_0) for SARS-CoV-1 and SARS-CoV-2 is relatively similar (24), the containment measures were only efficient against the earlier virus. Many factors could have contributed to the rapid spread of SARS-CoV-2, including host, viral and environmental factors. Although travel, specifically air travel, is a significant player in infectious diseases spread across countries, limited studies have been conducted in this domain. Here, we summarize the main epidemiological and clinical features of travel-associated COVID-19 up to April 19, 2020.

Travel-related introduction and tourism-related spread contributed substantially to the transmission of the virus across and within countries during the early phase of the COVID-19 pandemic. After an extensive search of national and international databases (**Supplementary Table 1**), we reported data from 145 countries and territories. Unfortunately, not all clinical and epidemiological data were available for all cases, which highlights the importance and the need for a better recording system, supervised by WHO and other international organizations, to preserve and share data.

In the absence of other studies on travel-associated CoV illness, our analysis is limited but remains essential for understanding the early phase of the pandemic. The mean age of the COVID-confirmed cases in our study was 44 (Range: 1–88), compared to 39.9 and 56 for MERS and SARS, respectively. On the other hand, the male to female ratio was calculated at 1.6:1 compared to 3.3:1 and 1:1.3 for MERS and SARS, respectively (37, 38). Accordingly, our data suggest that at the early stages of the outbreak, COVID-19 affected more males than females, similar to MERS, but younger age groups like SARS. According to several studies, age is considered a major risk factor for severe diseases (9, 14). Hence, the relatively young population (Mean of 44) of travelers could have been asymptomatic carriers who spread the virus across the globe. It is worth noting that most of the SARS and MERS cases were reported in healthcare or closed settings (21, 26, 38). The occupation information was missing for most of the travel-associated cases in our study. Healthcare workers (nurses and doctors) and students reported 1.18 and 1% of the occupations according to the available data. All other occupations represented proportions <1%.

The major obstacle of our analysis was the difficulty of finding reliable data from several countries. This limitation was noticeable in both developed and developing countries, indicating a lack of proper response to the crisis since the early stages. For example, the United States and Italy were among the earliest countries to report cases. Still, data about the travel-associated illness was very scarce. This partially explains the tremendous increase in cases at later stages in both countries. On the other hand, most of the organized and informative data were collected from Thailand, Malaysia, and South Korea, which had better control of the outbreak. This again emphasizes the importance of sharing data and making it available for public health officials and experts in the field to assess the problem and undertake the proper measures.

Initially, most of the travel-associated cases originated from China; however, Italy and Iran became the two main epicenters for Europe and The Middle East, respectively. Interestingly, between February 24 and March 13, China reported 264 imported cases from 17 countries. Further, the origin of travel-associated cases was identified in more than 40 countries during the period of the study, indicating the easiness of virus spread regardless of all implemented containment measures.

Identifying the incubation period and onset of symptoms in patients is very critical while preparing guidelines and plans to battle the pandemic. Unfortunately, this essential information was missing for most of the cases in our study. In general, patients presented a variety of symptoms, similar to what has been reported in other studies (8, 10, 39). Interestingly, 117 of the travel-associated cases presented symptoms during the flight from China. These were traveling to Singapore, Malaysia, Thailand, South Korea, Philippines, Germany, Vietnam, and Iran. On the other hand, data was available for 48 cases who developed symptoms up to 22 after travel. Accordingly, the implemented quarantine period of 14 days might be a little short and perhaps requires further consideration.

On a related aspect, it is essential to study viral survival on airplanes furniture, filters, and other parts, and its suspension as an aerosol during the flight. During the SARS-CoV-1 outbreak, studies had investigated the spread of the virus among passengers who traveled on 40 flights carrying symptomatic patients (40). Transmission seems to have occurred on board of only five out of the 40 flights, with only one reporting major spread. In one flight, a 72-year-old man, infected with SARS-CoV-1, transmitted the virus to 22 other passengers, including two aircraft crew (40, 41). Infection in the passengers was related to the physical proximity to the index patient. On the contrary, in another flight that carried four symptomatic patients, only one patient got infected. Few studies had reported on probable aircraft transmission of SARS-CoV-2 (42), and further studies on this aspect are needed. Most of the airlines had suspended their international flights, which affected the aviation economy globally (43, 44). This situation might last for an extended period until a herd immunity is achieved via national infection or vaccination. Until then, the airline companies are preparing a new set of guidelines that enable them to operate while maintaining a safe environment for passengers, including social distancing and other measures in the airports and onboard.

In about 8 months (November 2002–July 2003), the SARS-CoV-1 reached 31 counties, most of which (23; 75%) reported <10 cases (32). In more than 8 years, MERS-CoV was reported in 27 countries (23), with major outbreaks happened in Saudi Arabia and South Korea in hospital settings. Interestingly, MERS-CoV reported minimal spreading during crowd seasons in Saudi Arabia during Haj and Umra (38). SARS-CoV-2, on the other hand, reached all over the globe and spread rapidly in communities in a very short period. Accordingly, the three zoonotic CoV seems to follow a different trend of human-to-human transmission and spread.

Travel has led to the introduction and spread of SARS-CoV-2 in several ways, mainly through the mobility of individuals during trips (airports, airplanes, ships, and others), or afterward through family and social gatherings. Standard non-pharmaceutical measures are the most critical approaches for controlling the spread of COVID-19 in all settings, including during travel. Such measures include physical distancing, hand hygiene, respiratory etiquette, as well as other infection prevention and control measures. Information about the risk and symptoms of COVID-19, and advice to avoid travel while experiencing any of the COVI-19 related symptoms is essential (36, 45).

The SARS-CoV-2 emerged during the wintertime of temperate regions, which could have eased the spread of the virus. Interestingly, the virus spread did not slow down during the summertime, and it continues to surge in multiple countries. More flights are now in motion, easing the re-introduction of the virus to countries that might have controlled the spread on infection. This becomes more worrisome if the virus mutates to become more pathogenic and transmissible. Accordingly, several measures have to be implemented to reduce virus transmission with travelers, including airport screening, quarantine of contacts, and isolation of confirmed cases after travel. More importantly, WHO and other international organizations should urge countries to report in detail all the epidemiological and clinical characteristics of identified travel-associated cases, along with the status of transmission on all the flights. Proper decontamination of aircraft and social distancing in the flight might reduce the transmission, and hence, the control of the virus spread internationally.

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/s.

AUTHOR CONTRIBUTIONS

ME, TC, EF, and HY designed the concept. RM and OE collected and categorized data. RM analyzed data and generated tables and figures. AA provided funding. HY wrote the first draft. All authors revised the draft and agreed on the final version of the paper.

FUNDING

This work was supported by Qatar University internal grant # QUCG-BRC-20/21-1.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2020.573925/full#supplementary-material

REFERENCES

1. Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol.* (2019) 17:181–92. doi: 10.1038/s41579-018-0118-9
2. Drosten C, Günther S, Preiser W, van der Werf S, Brodt H-R, Becker S, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med.* (2003) 348:1967–76. doi: 10.1056/NEJMoa030747
3. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med.* (2003) 348:1953–66. doi: 10.1056/NEJMoa030781
4. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, Lim W, et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet.* (2003) 361:1319–25. doi: 10.1016/S0140-6736(03)13077-2
5. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med.* (2012) 367:1814–20. doi: 10.1056/NEJMoa1211721
6. Zumla A, Hui DS, Perlman S. Middle east respiratory syndrome. *Lancet.* (2015) 386:995–1007. doi: 10.1016/S0140-6736(15)60454-8
7. Chang D, Lin M, Wei L, Xie L, Zhu G, Dela Cruz CS, et al. Epidemiologic and clinical characteristics of novel coronavirus infections involving 13 patients outside Wuhan, China. *JAMA.* (2020) 323:1092–3. doi: 10.1001/jama.2020.1623
8. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5
9. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* (2020) 579:270–3. doi: 10.1038/s41586-020-2951-z
10. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med.* (2020) 382:727–33. doi: 10.1056/NEJMoa2001017
11. *Timeline-COVID-19 [WHO]. WHO Timeline - COVID-19*. WHO (2020). Available online at: https://www.who.int/news/item/27-04-2020-who-timeline---covid-19
12. John Hopkins University. *Animated Maps USA*. John Hopkins University (2020). Available online at: https://coronavirus.jhu.edu/data/animated-world-map
13. WHO-Virtual Conference. *Virtual Press Conference on COVID-19 – March 11 2020* (2020).
14. Guan W-J, Zhong N-S. Clinical characteristics of Covid-19 in China. *N Engl J Med.* (2020) 382:1859–62. doi: 10.1056/NEJMc2005203
15. Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *N Engl J Med.* (2020) 382:970–1. doi: 10.1056/NEJMc2001468
16. UNWTO. *Annual Report*. UNWTO (2015).
17. UNWTO. *Tourism Towards 2030 / Global Overview*. UNWTO (2011).
18. Parrish C, Holmes E, Morens D, Park E-C, Burke D, Calisher C, et al. Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol Mol Biol Rev.* (2008) 72:457–70. doi: 10.1128/MMBR.00004-08
19. Petersen E, Wilson ME, Touch S, McCloskey B, Mwaba P, Bates M, et al. Rapid spread of Zika virus in The Americas–implications for public health preparedness for mass gatherings at the 2016 Brazil Olympic Games. *Int J Infect Dis.* (2016) 44:11–5. doi: 10.1016/j.ijid.2016.02.001
20. Al Kahlout RA, Nasrallah GK, Farag EA, Wang L, Lattwein E, Müller MA, et al. Comparative serological study for the prevalence of anti-MERS coronavirus antibodies in high- and low-risk groups in Qatar. *J Immunol Res.* (2019) 2019:1386740. doi: 10.1155/2019/1386740
21. Azhar EI, Hui DSC, Memish ZA, Drosten C, Zumla A. The middle east respiratory syndrome (MERS). *Infect Dis Clin North Am.* (2019) 33:891–905. doi: 10.1016/j.idc.2019.08.001
22. Choi WS, Kang C-I, Kim Y, Choi J-P, Joh JS, Shin H-S, et al. Clinical presentation and outcomes of middle east respiratory

syndrome in the Republic of Korea. *Infect Chemother.* (2016) 48:118–26. doi: 10.3947/ic.2016.48.2.118
23. WHO-MERS. *Middle East Respiratory Syndrome Coronavirus (MERS-CoV)*. WHO (2019). Available online at: https://www.who.int/csr/don/26-december-2019-mers-qatar/en/
24. Petrosillo N, Viceconte G, Ergonul O, Ippolito G, Petersen E. COVID-19, SARS and MERS: are they closely related? *Clin Microbiol Infect.* (2020) 26:729–34. doi: 10.1016/j.cmi.2020.03.026
25. Saad M, Omrani AS, Baig K, Bahloul A, Elzein F, Matin MA, et al. Clinical aspects and outcomes of 70 patients with Middle East respiratory syndrome coronavirus infection: a single-center experience in Saudi Arabia. *Int J Infect Dis.* (2014) 29:301–6. doi: 10.1016/j.ijid.2014.09.003
26. Lee N, Hui D, Wu A, Chan P, Cameron P, Joynt GM, et al. A major outbreak of severe acute respiratory syndrome in Hong Kong. *N Engl J Med.* (2003) 348:1986–94. doi: 10.1056/NEJMoa030685
27. Yu IT, Li Y, Wong TW, Tam W, Chan AT, Lee JH, et al. Evidence of airborne transmission of the severe acute respiratory syndrome virus. *N Engl J Med.* (2004) 350:1731–9. doi: 10.1056/NEJMoa032867
28. WHO-Dashboard. *WHO Coronavirus Disease (COVID-19) Dashboard*. WHO (2020). Available online at: https://covid19.who.int/
29. WHO-Ebola. *Ebola Virus Disease News*. WHO (2020). Available online at: https://www.who.int/news-room/fact-sheets/detail/ebola-virus-disease
30. WHO-Zika. *Zika Epidemiology Update*. WHO (2019). Available online at: https://www.who.int/emergencies/diseases/zika/epidemiology-update/en/
31. Reperant LA, Osterhaus A. AIDS, Avian flu, SARS, MERS, Ebola, Zika… what next? *Vaccine.* (2017) 35 (35 Pt A):4470–4. doi: 10.1016/j.vaccine.2017.04.082
32. WHO-SARS. *Summary of Probable SARS Cases With Onset of Illness From November 1 2002 to July 31 2003*. WHO (2004).
33. Kawana A. SARS, MERS and coronavirus infections. *Nihon Rinsho.* (2016) 74:1967–72.
34. Google-COVID19. *Coronavirus (COVID-19)*. Google (2020). Available online at: https://news.google.com/covid19/map?hl=en-US&gl=US&ceid=US:en
35. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City Area. *JAMA.* (2020) 323:2052–9. doi: 10.1001/jama.2020.6775
36. ECDC. *Considerations for Travel-Related Measures to Reduce Spread of COVID-19 in the EU/EEA*. European Centre for Disease Prevention and Control (2020). Available online at: https://www.ecdc.europa.eu/en/publications-data/considerations-travel-related-measures-reduce-spread-covid-19-eueea
37. Aggarwal S, Garcia-Telles N, Aggarwal G, Lavie C, Lippi G, Henry BM. Clinical features, laboratory characteristics, and outcomes of patients hospitalized with coronavirus disease 2019 (COVID-19): early report from the United States. *Diagnosis.* (2020) 7:91–6. doi: 10.1515/dx-2020-0046
38. Al-Tawfiq JA, Zumla A, Memish ZA. Travel implications of emerging coronaviruses: SARS and MERS-CoV. *Travel Med Infect Dis.* (2014) 12:422–8. doi: 10.1016/j.tmaid.2014.06.007
39. Assiri A, Al-Tawfiq JA, Al-Rabeeah AA, Al-Rabiah FA, Al-Hajjar S, Al-Barrak A, et al. Epidemiological, demographic, and clinical characteristics of 47 cases of Middle East respiratory syndrome coronavirus disease from Saudi Arabia: a descriptive study. *Lancet Infect Dis.* (2013) 13:752–61. doi: 10.1016/S1473-3099(13)70204-4
40. Olsen SJ, Chang HL, Cheung TY, Tang AF, Fisk TL, Ooi SP, et al. transmission of the severe acute respiratory syndrome on aircraft. *N Engl J Med.* (2003) 349:2416–22. doi: 10.1056/NEJMoa031349
41. Pavia AT. Germs on a plane: aircraft, international travel, and the global spread of disease. *J Infect Dis.* (2007) 195:621–2. doi: 10.1086/511439
42. Eldin C, Lagier JC, Mailhe M, Gautret P. Probable aircraft transmission of Covid-19 in-flight from the Central African Republic to France. *Travel Med Infect Dis.* (2020) 35:101643. doi: 10.1016/j.tmaid.2020.101643
43. Willis Towers Watson. *COVID-19 and the Aviation Industry.* Willis Towers Watson (2020). Available online at: https://www.willistowerswatson.com/en-GB/Insights/2020/03/covid-19-and-the-aviation-industry (accessed April, 2020).
44. *International Airport Review. Coronavirus Roundtable: How is the Aviation Industry Responding to the COVID-19 Pandemic?* (2020). Available online at: https://www.internationalairportreview.com/article/114585/aviation-industry-covid-19-pandemic/ (accessed April 2020).
45. WHO-IRIS. *Non-pharmaceutical Public Health Measures for Mitigating the Risk and Impact of Epidemic and Pandemic Influenza: Annex: Report of Systematic Literature Reviews*. World Health Organization (2019). Available online at: https://apps.who.int/iris/handle/10665/329439

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.

frontiers in Medicine

ORIGINAL RESEARCH
published: 14 January 2021
doi: 10.3389/fmed.2020.594439

Structural Comparison of the SARS CoV 2 Spike Protein Relative to Other Human-Infecting Coronaviruses

Marni E. Cueno and Kenichi Imai**

Department of Microbiology, Nihon University School of Dentistry, Tokyo, Japan

Coronaviruses (CoV) are enveloped positive-stranded RNA viruses and, historically, there are seven known human-infecting CoVs with varying degrees of virulence. CoV attachment to the host is the first step of viral pathogenesis and mainly relies on the spike glycoprotein located on the viral surface. Among the human-infecting CoVs, only the infection of SARS CoV 2 (SARS2) among humans resulted to a pandemic which would suggest that the protein structural conformation of SARS2 spike protein is distinct as compared to other human-infecting CoVs. Surprisingly, the possible differences and similarities in the protein structural conformation between the various human-infecting CoV spike proteins have not been fully elucidated. In this study, we utilized a computational approach to generate models and analyze the seven human-infecting CoV spike proteins, namely: HCoV 229E, HCoV OC43, HCoV NL63, HCoV HKU1, SARS CoV, MERS CoV, and SARS2. Model quality assessment of all CoV models generated, structural superimposition of the whole protein model and selected S1 domains (S1-CTD and S1-NTD), and structural comparison based on RMSD values, Tm scores, and contact mapping were all performed. We found that the structural orientation of S1-CTD is a potential structural feature associated to both the CoV phylogenetic cluster and lineage. Moreover, we observed that spike models in the same phylogenetic cluster or lineage could potentially have similar protein structure. Additionally, we established that there are potentially three distinct S1-CTD orientation (Pattern I, Pattern II, Pattern III) among the human-infecting CoVs. Furthermore, we postulate that human-infecting CoVs in the same phylogenetic cluster may have similar S1-CTD and S1-NTD structural orientation. Taken together, we propose that the SARS2 spike S1-CTD follows a Pattern III orientation which has a higher degree of similarity with SARS1 and some degree of similarity with both OC43 and HKU1 which coincidentally are in the same phylogenetic cluster and lineage, whereas, the SARS2 spike S1-NTD has some degree of similarity among human-infecting CoVs that are either in the same phylogenetic cluster or lineage.

Keywords: C-terminal domain (CTD), modeling, N-terminal domain (NTD), SARS coronavirus 2 (SARS2), spike

OPEN ACCESS

Edited by:
Matteo Convertino, Hokkaido University, Japan

Reviewed by:
Soham Saha, Institut Pasteur, France
Takahiro Watanabe, Nagoya University, Japan

****Correspondence:***
Kenichi Imai
imai.kenichi@nihon-u.ac.jp
Marni E. Cueno
marni.cueno@nihon-u.ac.jp

Specialty section:
This article was submitted to Infectious Diseases - Surveillance, Prevention and Treatment, a section of the journal Frontiers in Medicine

Received: *13 August 2020*
Accepted: *14 December 2020*
Published: *14 January 2021*

Citation:
Cueno ME and Imai K (2021) Structural Comparison of the SARS CoV 2 Spike Protein Relative to Other Human-Infecting Coronaviruses. Front. Med. 7:594439. doi: 10.3389/fmed.2020.594439

INTRODUCTION

Coronaviruses (CoV) are enveloped positive-stranded RNA viruses that belong to the family *Coronaviridae* and order *Nidovirales* with the subfamily *Othocoronavirinae* composed of four genera, namely: *alphacoronavirus*, *betacoronavirus*, *gammacoronavirus*, and *deltacoronavirus* (1). Historically, there are seven known CoVs capable of infecting humans, namely: human CoV (HCoV)-229E (1962), HCoV-OC43 (1967), severe acute respiratory syndrome (SARS)-CoV 1 (SARS1) (2002), HCoV-NL63 (2004), HCoV-HKU1 (2005), Middle East respiratory syndrome (MERS)-CoV (2012), and SARS-CoV 2 (SARS2) (2019) (2–8). In general, CoVs cause serious health problems to both human and animal hosts and, in particular, CoV infections primarily affect the respiratory and gastrointestinal tracts (9). Moreover, CoVs have the largest genome among all known RNA viruses which in-turn is packed in a helical capsid comprised of a nucleocapsid protein (N) and surrounded by a viral envelope which in-turn is associated with structural proteins, namely: membrane, envelope, and spike (10). Among the structural proteins, the spike protein has been involved in mediating viral entry, determinant of host tropism, inducing viral pathogenesis, and major inducer of host immune responses (9–12). This would highlight the significance of the CoV spike protein in terms of viral pathogenesis.

Spike protein (a class I viral fusion protein) follows a metastable prefusion conformation upon translation and, likewise, forms trimers that resemble club-shaped spikes along the CoV membrane surface (13). Additionally, the spike protein is comprised of three segments, namely: the large ectodomain, single-pass transmembrane anchor, and short intracellular tail (10). With regards to the ectodomain, it is further divided into the S1 receptor-binding subunit that mainly functions in viral attachment and S2 membrane-fusion subunit that facilitates virus-cell fusion (9, 10). In a CoV infection scenario, S1 would bind to a suitable receptor on the host cell surface enabling viral attachment and, subsequently, S2 fuses both the host and viral membranes, thereby, allowing viral genomes to enter host cells (9, 10). This shows that receptor binding and membrane fusion are important initial and key steps in CoV pathogenesis. Interestingly, the receptor-binding domain (RBD) and host receptor differ among the known human-infecting CoVs. In particular, known host receptors include: aminopeptidase N (APN) for 229E; angiotensin-converting enzyme 2 (ACE2) for NL63, SARS1, and SARS2; *O*-acetylated sialic acid (*O*-ac Sia) for OC43 and HKU1; and dipeptidyl peptidase-4 (DPP4) for MERS (9, 14). Among the human-infecting CoVs, infection in the upper respiratory tract has been associated with 229E, OC43, NL63, and HKU1 (15), whereas, infection in the lower respiratory tract has been associated with SARS1, MERS, and SARS2 (7, 16, 17). Interestingly, among the human-infecting CoVs, only SARS2 infection resulted to a pandemic which would insinuate that the protein structure of the SARS2 spike protein has a structural conformation that is distinct as compared to other human-infecting CoVs (18, 19). However, to our knowledge, the structural comparison between SARS2 and the other human-infecting CoVs has not been fully elucidated. A better understanding of the possible differences and similarities in the protein structural conformation of the SARS2 spike protein compared to the spike proteins of the other human-infecting CoV may shed a light on how this particular virus more effectively cause infection and, more importantly, establish the potential cross-reactivity of SARS2 with other human-infecting CoV which in-turn may lead to novel therapeutic strategies.

MATERIALS AND METHODS

CoV Spike Modeling

Representative CoV spike amino acid sequences from 229E, OC43, NL63, HUK1, SARS1, MERS, and SARS2 were collected from the National Center for Biological Information (NCBI) and UniProt Web sites. To obtain the most accurate monomeric spike model that could serve as the representative prefusion model for each CoV strain, a minimum of five generated sequence models were first analyzed and spike models with similar Root Mean Square Deviation (RMSD) values and Template Modeling scores (Tm-scores) based on superimposition using the default setting of Tm align (20) were used for further downstream analyses. The following representative amino acid sequences were utilized for spike modeling with Genebank accession number indicated: 229E (ABB90513), OC43 (AXX83297), NL63 (QED88040), HKU1 (ARB07617), SARS1 (AAR07625), MERS (AHX00731), and SARS2 (YP_009724390). In addition, representative SARS2 spike S1 C-terminal domain (S1-CTD) and N-terminal domain (S1-NTD) models were generated based on UniProt reference number P0DTC2. In the whole study, all generated models were made using the default settings of Phyre2 web server (21) and Jmol applet (22) was used for protein visualization.

Model Quality Assessment

All generated spike models were assessed for quality prior to further analyses. Both protein model:crystal structure superimposition and contact mapping were utilized for model quality estimation. Representative crystal structure used for superimposition was the 1998 strain (PDB ID: 6VXX). Additionally, a monomeric 6VXX model (crystal model) was generated using Phyre 2 and, subsequently, superimposed for comparison to the 6VXX crystal structure to further serve as model quality check. Representative CoV spike models and crystal structure were superimposed using Tm align (20). For the purpose of this study, we considered spike models adequate for further analyses if RMSD values between superimposed sequence model:crystal and crystal model:crystal are close. Moreover, CMView applet (Contact type: Cα; Distance cut-off: 8.0; Needleman-Wunsch alignment) was used to establish protein contact map of both the model and crystal in order to determine common contact (23) and, consequently, higher common contact would mean more structural similarities between the model and crystal (24) which in-turn would further indicate whether the model is suitable for further analyses.

Comparison Among CoV Spike Models

Three different sets of structural comparisons were performed. In one analysis conducted, all generated CoV spike models

were compared (Visually, RMSD value, and Tm score) to the SARS2 spike model through superimposition. Subsequently, superimposition and comparison (both RMSD value and Tm score) between the various CoV spike models were likewise made. In another separate analysis, SARS2 spike S1-CTD and S1-NTD models were similarly superimposed and compared (Tm score only) to the other generated CoV spike models. For mutant spike model comparisons, superimpositions were done with the following: (1) original SARS2 spike model; and (2) original SARS2 spike S1-CTD and S1-NTD models. Visual observation (simply looking at the structure), RMSD value, Tm score, and protein common contact were established using Jmol, Tm align, and CMView, respectively.

RESULTS

Generated CoV Spike Models Are Reliable

Model quality assessment prior to further downstream analyses on either experimental (i.e., crystallized) or theoretical (i.e., computer-based) protein structures generated has long been recommended (25). In line with this, to elucidate the accuracy and reliability of all CoV spike models generated throughout this study, both structural and protein contact map superimpositions were performed. Three representative structures [SARS2 crystal structure (**Figure 1A**), SARS2 crystal model (**Figure 1B**), SARS2 sequence model (**Figure 1C**)] were used for superimposition. For model-crystal superimpositions, only spike monomers were considered. We found that RMSD values between crystal model:crystal [RMSD 1.78] (**Figure 1D**) and sequence model:crystal model [RMSD 1.77] (**Figure 1E**) were relatively close, which would imply that both generated models are structurally similar. In addition, the sequence model:crystal superimposition [RMSD 1.19] (**Figure 1F**) was below 1.5 Å which in-turn was considered adequate for further analyses (26). Furthermore, we observed that protein contact map superimposition between crystal model:crystal (**Figure 1G**), sequence model:crystal model (**Figure 1H**), and sequence model:crystal (**Figure 1I**) were above common contact 70%, which demonstrates the high contact similarity between the superimposed structures. Taken together, we believe that these results indicate that the generated models can be used for further downstream analyses.

CoV Spike Models Differ Based on Phylogenetic Cluster and Lineage

Currently, there are seven known human-infecting CoV (1) and the spike protein for each CoV has been thoroughly studied (9, 10). CoV spike proteins are divided into two functionally distinct subunits (S1 and S2 subunits), wherein, the S1 subunit is further distinguished by four distinct domains (NTD comprised of domain A; CTD comprised of domains B, C, D) serving as receptor binding domains highlighting the importance of these domains in viral pathogenicity (9, 14, 27). To visualize and compare the CoV spike proteins, each human-infecting CoV spike model was generated and both visual observation and structural comparison mainly focused on both S1-CTD and S1-NTD. As seen in **Figures 2A–G**, through visual observation, a prominent structural difference between the spike models is the S1-CTD orientation (indicated in red dashed lines), whereas, S1-NTD orientation generally looked the same (indicated in blue dashed lines). More specifically, we were able to identify three possible patterns of spike S1-CTD orientation: (1) 229E and NL63; (2) OC43, HKU1, SARS1, and SARS2; and (3) MERS. Among the human-infecting CoVs, two strains (229E and NL63) belong to the alpha-CoV phylogenetic cluster, whereas, the remaining five strains belong to the beta-CoV phylogenetic cluster which can be further divided into the A (OC43 and HKU1), B (SARS1 and SARS2), and C (MERS) lineages (2–4, 6, 8, 28, 29). In this regard, we hypothesize that the similarities in S1-CTD orientation among the spike models is a possible structural feature associated to both the CoV phylogenetic cluster and lineage.

To further compare the CoV spike models, both RMSD value and Tm scores were determined. RMSD values measure similarity between two superimposed atomic coordinates, whereas, Tm scores measure the similarity between protein structures without relying on protein size (20, 30). Both measurements are used to establish structural similarities between two superimposed proteins (30). In this regard, we observed that superimposed CoV spike models that have RMSD < 1.0 (**Figure 2H**) and Tm score > 0.95 (**Figure 2I**) either belong to the same phylogenetic cluster (229E and NL63) or lineage (OC43 and HKU1, SARS1 and SARS2, MERS). This would further imply that spike models in the same phylogenetic cluster or lineage generally may have similar protein structure as well. Noticeably, NL63 model normalized to the 229E model measured Tm score 0.84366 which is lower compared to other Tm scores measured from other CoVs spike models within the same phylogenetic cluster and lineage. This may suggest that there is some structural difference between these two spike models, which we suspect is related to viral evolution of NL63 from 229E (31).

It is worth mentioning that although all human-infecting CoVs are in the same protein structural fold (Tm score > 0.50) (32), among the superimposed spike models, we found certain beta-CoV strains that belong to separate lineages (OC43 and SARS1; OC43 and SARS2; HKU1 and SARS1; HKU1 and SARS2) have Tm score > 0.70 which (asides from being consistent to belonging to the same CoV genera) may likewise insinuate some degree of structural similarity albeit to a lesser extent compared to those in the same lineage (Tm score > 0.95). Additionally, in possible future works, it would be interesting to determine specific conformational features, establish known conformations and structural domains of S1-CTD, and elucidate domain classification among the seven human-infecting CoV spike protein.

SARS1 and SARS2 Spike Models Are Structurally Similar

Among the seven human-infecting CoVs, only SARS2 resulted to a pandemic (18, 19) which may suggest that the overall SARS2 spike protein differs from the other human-infecting

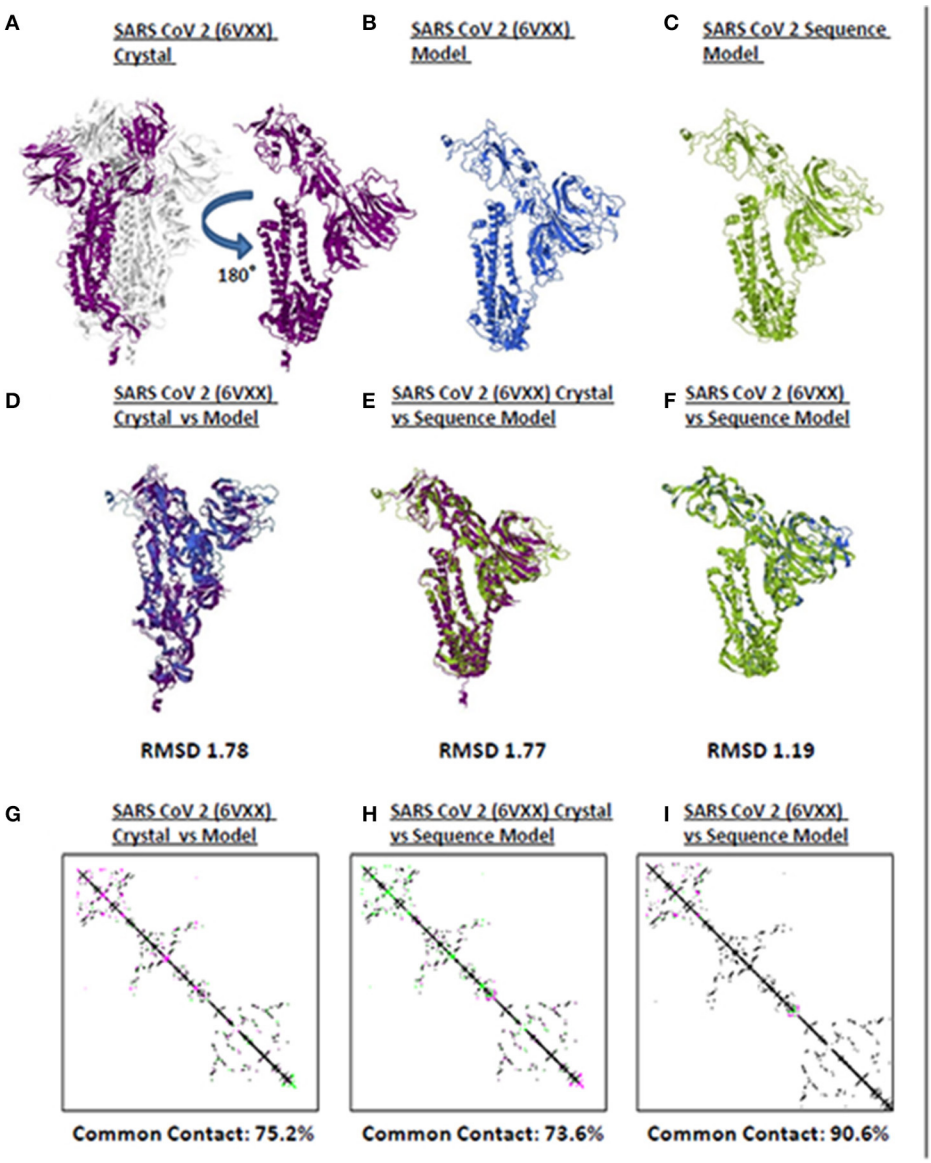
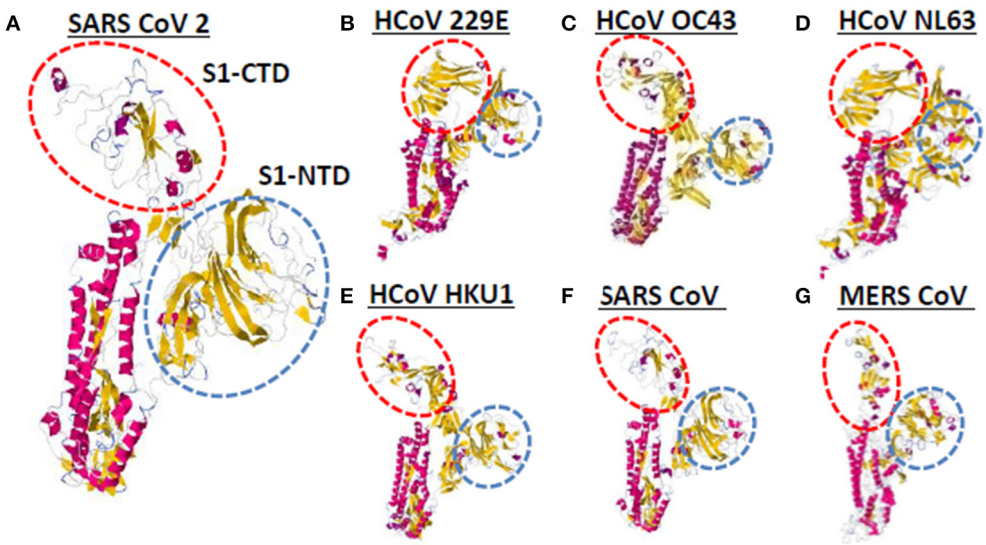


FIGURE 1 | Quality assessment of monomeric human-infecting coronavirus spike protein models generated. Representative SARS CoV 2 **(A)** 6VXX crystal, **(B)** 6VXX model, and **(C)** sequence model of monomeric spike proteins are shown. Sumperimposition between **(D)** 6VXX crystal and model, **(E)** 6VXX crystal and sequence model, and **(F)** 6VXX and sequence models are presented. RMSD scores of the superimposed protein structures are indicated below. Contact maps of **(G)** 6VXX crystal and model, **(H)** 6VXX crystal and sequence model, and **(I)** 6VXX and sequence models are shown. Contacts present in both protein structures (black) and present in one of the protein structures [either pink (first protein structure uploaded) or green (second protein structure uploaded)] are indicated. Common contact of the protein structures being compared are labeled below. SARS CoV 2 6VXX crystal (violet), 6VXX model (royal blue), and sequence model (yellow green) are indicated.

CoVs. To structurally differentiate SARS2 and other human-infecting CoV spike models, model superimposition was performed. As a follow-up from our earlier results (**Figure 1**), we utilized representative spike models (NL63, MERS, SARS1) for superimposition against SARS2 since these models putatively share different S1-CTD orientation (based on visual observation) and have both RMSD < 1.0 and Tm score > 0.95 among spike models within the same phylogenetic cluster and lineage. For purposes of this study, we classified distinct S1-CTD orientations as patterns and, likewise, established which among the spike protein models share the same S1-CTD orientation, whereby, spike protein models with the same S1-CTD orientation would be classified into one pattern. In this regard, we observed three distinct S1-CTD orientations which we classified into three patterns among the superimposed spike models: (1) Pattern I (NL63 and SARS2 superimposition; **Figure 3A**); (2) Pattern II (MERS and SARS2 superimposition; **Figure 3B**), and (3) Pattern III (SARS1 and SARS2 superimposition; **Figure 3C**). This is consistent with our earlier observations (**Figures 2A–G**) which would further suggest that spike models within the same



H

Strain	HCoV 229E	HCoV OC43	HCoV NL63	HCoV HKU1	SARS CoV	MERS CoV	SARS CoV 2
HCoV 229E		5.94	0.84	5.93	5.74	6.25	5.78
HCoV OC43			6.19	0.99	4.35	4.73	4.39
HCoV NL63				6.05	5.72	6.38	5.81
HCoV HKU1					4.47	4.70	4.45
SARS CoV						4.36	0.70
MERS CoV							4.38
SARS CoV 2							

I

Strain	(HCoV 229E)	(HCoV OC43)	(HCoV NL63)	(HCoV HKU1)	(SARS CoV)	(MERS CoV)	(SARS CoV 2)
HCoV 229E		0.68060 (0.58837)	0.98914 (0.84366)	0.67916 (0.59781)	0.65198 (0.62908)	0.62417 (0.54765)	0.64977 (0.62958)
HCoV OC43			0.58774 (0.58992)	0.95898 (0.97838)	0.75170 (0.84187)	0.66455 (0.67328)	0.74868 (0.84213)
HCoV NL63				0.58798 (0.59642)	0.56526 (0.62885)	0.54389 (0.54870)	0.56322 (0.62914)
HCoV HKU1					0.76049 (0.83595)	0.67150 (0.66786)	0.75740 (0.83626)
SARS CoV						0.71186 (0.64377)	0.98095 (0.98547)
MERS CoV							0.64606 (0.71751)
SARS CoV 2							

FIGURE 2 | Whole protein structural comparison of the various monomeric human-infecting coronavirus spike protein models. Representative spike protein models of **(A)** SARS CoV 2, **(B)** HCoV 229E, **(C)** HCoV OC43, **(D)** HCoV NL63, **(E)** HCoV HKU1, **(F)** SARS CoV 1, and **(G)** MERS CoV are shown. S1-CTD (red dashed circle) and S1-NTD (blue dashed circle) are indicated. **(H)** RMSD and **(I)** Tm scores of superimposed spike models are tabulated. Tm scores normalized to a spike model is distinguished by having or not having a parenthesis.

phylogenetic cluster and lineage share the same spike S1-CTD model orientation. In this regard, based on **Figures 2A–G**, we think that 229E follows a Pattern I orientation while both OC43 and HKU1 follows a Pattern III orientation.

To further differentiate SARS2 and other human-infecting CoV spike models, contact map overlap (CMO) analyses were done. Contact maps provide information with regards to the pairwise spatial and functional relationship of residues in a given protein and, likewise, unifies certain aspects of protein folding and structure prediction which in-turn allows protein reconstruction (33, 34). We found that among the CoV spike models compared to SARS2 model (**Figures 3D–I**), only SARS1

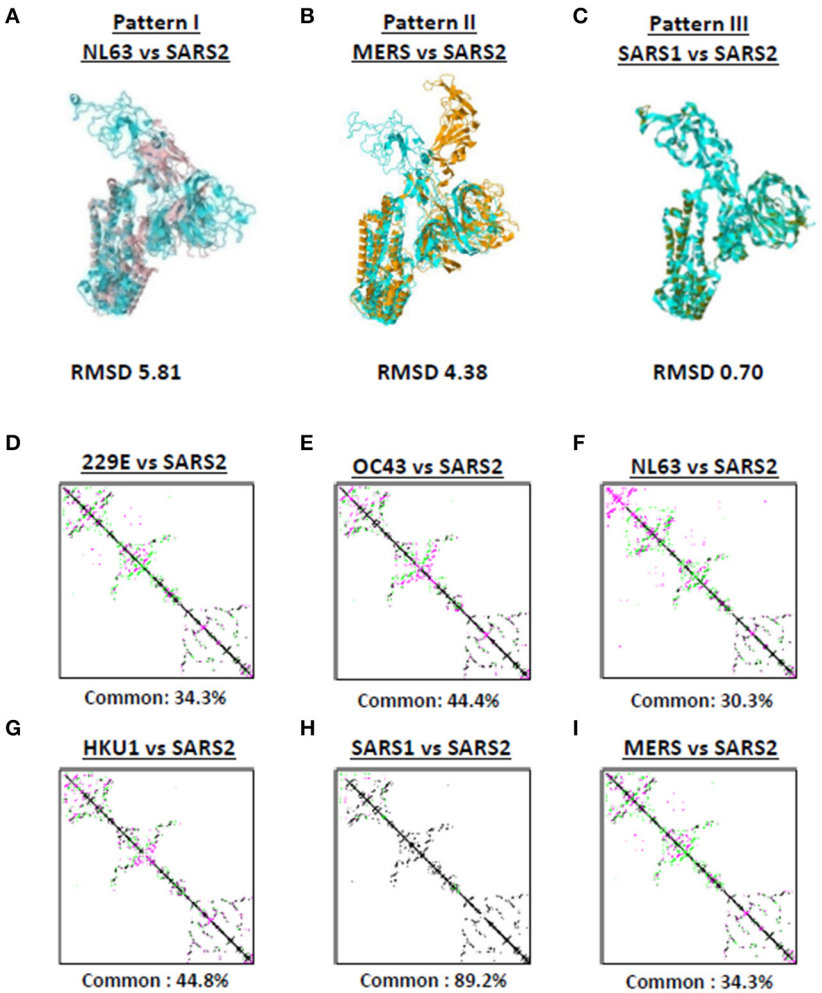


FIGURE 3 | Putative structural patterns among the monomeric human-infecting conronavirus spike protein models. Representative **(A)** Pattern I based on superimposition of NL63 and SARS2 models, **(B)** Pattern II based on superimposition of MERS and SARS2 models, and **(C)** Pattern III based on superimposition of SARS1 and SARS2 models are shown. RMSD scores of the superimposed protein structures are indicated below. Contact maps of **(D)** 229E and SARS2, **(E)** OC43 and SARS2, **(F)** NL63 and SARS2, **(G)** HKU1 and SARS2, **(H)** SARS1 and SARS2, and **(I)** MERS and SARS2 models are shown. Contacts present in both protein structures (black) and present in one of the protein structures [either pink (first protein structure uploaded) or green (second protein structure uploaded: SARS2)] are indicated. Common contact of the protein structures being compared are labeled below.

has more common contact (89.2%) with SARS2 (**Figure 3H**) which in-turn would indicate that SARS1 and SARS2 spike models have high structural similarity compared to SARS2 and other human-infecting CoV spike models (common contact < 50%). SARS1 and SARS2 viral genomes have ∼80% nucleotide identity (35, 36), whereas, SARS1 and SARS2 spike proteins have a 75–81% nucleotide similarity (37). In this regard, asides from belonging to the same lineage (29), we correlated the high common contact between SARS1 and SARS2 spike models to high nucleotide similarity.

Considering the results at this point, it is worth mentioning that RMSD values, Tm score, and CMO analyses were all based on superimposition of full-length CoV spike protein models. However, since spike S1-CTD model orientation varied while spike S1-NTD model orientation seem to be the same (**Figures 2A–G**, **3A–C**), structural comparison focusing only on both S1-CTD and S1-NTD is merited.

SARS2 Spike S1-CTD and S1-NTD Models Have Some Degree of Similarity Among Beta-CoV Spike Models

Both S1-CTD and S1-NTD are major domains located in the globular S1 subunit of CoV spike proteins that have been associated to receptor recognition (9, 10). Considering the spike S1-CTD orientation varied while the spike S1-NTD orientation were similar among the human-infecting CoVs, we compared the SARS2 spike S1-CTD and S1-NTD models from selected human-infecting CoV spike models through model superimposition. Subsequently, visual observation of the superimposed structure

was performed and, for confirmation, Tm score normalized to either the SARS2 spike S1-CTD or S1-NTD model were likewise measured. For S1-CTD model superimposition, only human-infecting CoV spike models following Pattern III orientation were superimposed to the SARS2 spike S1-CTD model. For S1-NTD model superimposition, all spike models were used since all S1-NTD orientation seemed to be the same (**Figures 2A–G**). Based on visual observation, we observed that both SARS2 spike S1-CTD (**Figures 4A–C**, upper panels) and S1-NTD (**Figures 4D–I**, lower panels) model superimpositions showed few structural overlaps compared to other human-infecting CoV spike models, whereas, SARS2 spike S1-CTD model seems to suggest higher structural overlap with SARS1 spike model (**Figure 4C**, upper panel). Similarly, Tm score measurements of either SARS2 spike S1-CTD (**Figures 4A–C**, lower panels) or S1-NTD (**Figures 4D–I**, lower panels) model superimposition showed a Tm score > 0.70 except for 229E (**Figure 4D**) and NL63 (**Figure 4F**).

Taken together, we postulate that these results would insinuate that: (1) both SARS1 spike S1-CTD and S1-NTD have higher structural similarity (Tm > 0.90) with SARS2 spike S1-CTD and S1-NTD, respectively; (2) OC43 and HKU1 spike S1-CTD and S1-NTD have some degree of similarity (Tm > 0.70) with SARS2 spike S1-CTD and S1-NTD, respectively; (3) MERS spike S1-NTD may likewise have some degree of similarity (Tm > 0.70) with SARS2 spike S1-NTD; and (4) 229E and NL63 spike S1-NTD are not structurally similar (Tm < 0.50) with SARS2 spike S1-NTD. We likewise suspect that this is correlated to whether the human-infecting CoV spike model belongs to the same phylogenetic cluster and lineage consistent with our earlier results.

DISCUSSION

SARS2 is the causative agent of coronavirus disease 2019 (COVID-19) pandemic (19). Interestingly, pre-existing SARS2 immunity has been observed among certain unexposed individuals in the general population (38, 39). Moreover, it was speculated that SARS2-specific T cells among unexposed individuals is attributable to memory T cells exposed to common cold CoVs (39) which in-turn would suggest the possible occurrence of immune cross-reactivity. By definition, cross-reactivity refers to immune responses that have non-specific targeting against a particular antigen ascribable to the flexible interaction between both B- and T-cell receptors and antigens (40). Considering all CoV infections start with spike protein binding, thereby, making it the first antigen recognized by the immune response (9, 10, 41), structural similarities between the spike proteins of human-infecting CoVs may play an important role in stimulating pre-existing SARS2 immunity. Throughout this study, we attempted to establish the putative structural differences and similarities among the seven known human-infecting CoVs spike protein conformations.

Epitopes serve as antigenic determinants that are normally found along the regions of an antigen that are recognized by B- and T-cells and can be classified as either sequential or conformational, whereby, sequential epitopes (continuous or linear amino acid stretch) do not rely on protein conformation while conformational epitopes (discontinuous amino acid stretch) rely on protein folding and conformation (42–44). In addition, conformational epitopes make up ∼90% of total antigen:antibody complexes (45) which would emphasize the importance of conformational epitopes. On the other hand, complimentary determinants (paratopes) within the antibody variable region recognize and interact with epitopes, wherein, this particular interaction goes beyond amino acid sequence recognition but rather is at the level of epitope:paratope steric complementarity and ionic charge (40). Considering every antibody paratope could interact with multiple antigen epitopes, this could lead to polyclonal immune response which is a fundamental principle of cross-reactivity (40, 46). Thus, this would mean that antibody binding to conformational epitopes could potentially lead to cross-reactivity. Considering the CoV S glycoprotein is the primary target for neutralizing antibody (47), we assume that any possible structural similarities between CoV spike proteins would likewise mean putatively comparable conformational epitopes. Earlier works have shown that immune cross-reactivity (and some cases of cross-neutralization) among human-infecting CoVs has been observed in the following human-infecting CoVs: between SARS1 and SARS2 (48–50); between SARS1 and NL63 (51); between SARS1 and 229E (51, 52); between SARS1 and OC43 (51–53); between SARS1 and MERS (54); and between NL63 and 229E (55). In this regard and considering our results, we hypothesize that some degree of structural similarity (Tm > 0.70) between SARS2 and other human-infecting CoVs spike S1-CTD and S1-NTD may suggest the possibility of cross-reactivity, whereby, potential neutralizing antibodies that recognize conformational epitopes along the S1-CTD and S1-NTD of human-infecting CoVs spike protein could likewise recognize conformational epitopes along the SARS2 spike S1-CTD and S1-NTD. In fact, consistent with our proposed hypothesis, earlier works have shown that pre-existing T cells recognizing SARS2 can be detected in a significant portion of the global human population (38, 39, 56) possibly attributable to humans being exposed to at least one form of human-infecting CoV (57). Thus, we believe that this would further support the possibility of having pre-existing immunity against SARS2 via cross-reactive immune response from other human-infecting CoVs with at least some degree of structural similarity (particularly in S1-CTD and S1-NTD).

It is worth mentioning that levels of neutralizing antibody response between human-infecting CoVs may likewise vary as previously observed (52), wherein, SARS1 and OC43 cross-reactive immune response was found to be higher compared to SARS1 and 229E cross-reactive immune response. Considering the results we obtained in this study, we speculate that varying immune cross-reactivity among human-infecting CoV spike protein might be ascribable to whether one or both S1-CTD and S1-NTD have higher structural similarity which in-turn may be influenced by both CoV phylogenetic cluster and lineage. Moreover, since immune responses (both humoral and cellular) to CoV diminishes at a certain time which in-turn allow for future re-infection (58–60), we likewise suspect that this may impact immune cross-reactivity of SARS2 and other human-infecting CoVs which consequently may affect the severity of SARS2

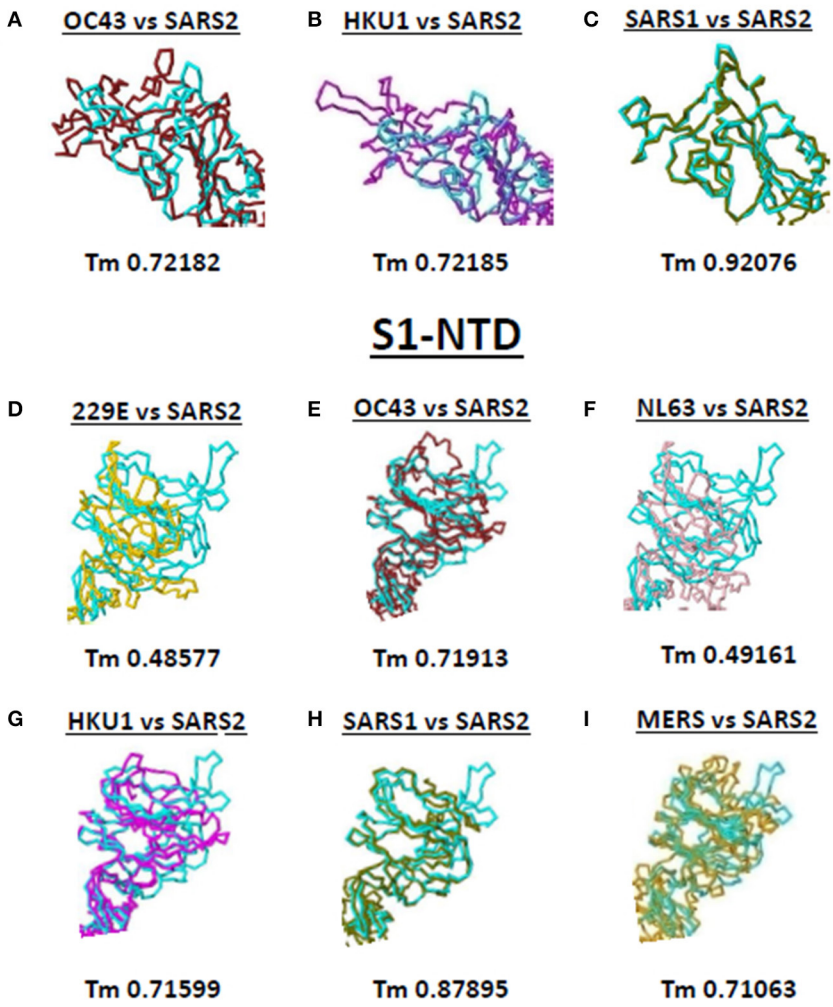


FIGURE 4 | Structural comparison of S1-CTD and S1-NTD between the various human-infecting coronaviruses and SARS CoV 2 spike models. Spike S1-CTD comparison between superimposed **(A)** OC43 and SARS2, **(B)** HKU1 and SARS2, and **(C)** SARS1 and SARS2 models are shown. Spike S1-NTD comparison between superimposed **(D)** 229E and SARS2, **(E)** OC43 and SARS2, **(F)** NL63 and SARS2, **(G)** HKU1 and SARS2, **(H)** SARS1 and SARS2, and **(I)** MERS and SARS2 models are presented. Tm scores normalized to the SARS2 model are labeled below all superimposed structures. 229E (gold), OC43 (brown), NL63 (pink), HKU1 (magenta), SARS1 (olive), MERS (orange), and SARS2 (cyan) are indicated.

infection. To speculate on the impact, patients with a relatively recent CoV infection (not SARS2) may develop a less severe form of COVID-19 while patients infected by another human-infecting CoV more than a year ago may result into a more severe form of COVID-19 with both scenarios being affected by the presence or absence of cross-reactive immune response from a prior human-infecting CoV contagion with some degree of structural similarity to one or both SARS2 S1-CTD and S1-NTD. We emphasize that these are speculations and would ultimately require both laboratory and clinical experimentation to prove. Similarly, we would like to highlight that the entire study is performed with predicted monomeric protein conformations, however, in cells a trimer of spike protein usually attaches to the host receptor (depending on the human-infecting CoV strain). In this regard, our results and interpretation to these results may differ in a CoV infection scenario.

In summary, we putatively established the differences and similarities in the structural conformation of the spike protein among human-infecting CoVs. In particular, we postulate on the following: (1) structural orientation of S1-CTD is a possible structural feature associated to both the CoV phylogenetic cluster and lineage; (2) spike models in the same phylogenetic cluster or lineage could potentially have similar protein structure; (3) there are potentially three distinct S1-CTD orientation among the human-infecting CoVs; and (4) human-infecting CoVs in the same phylogenetic cluster possibly have similar S1-CTD and S1-NTD. Overall, we propose that the SARS2 spike S1-CTD follows a Pattern III orientation which has a higher degree

of similarity with SARS1 and some degree of similarity with both OC43 and HKU1 which coincidentally are in the same phylogenetic cluster and lineage, whereas, the SARS2 spike S1-NTD has some degree of similarity among human-infecting CoVs that are either in the same phylogenetic cluster or lineage.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

Both authors were involved in formulating the idea, performing the structural analyses, and writing the manuscript.

FUNDING

This work was supported by JSPS KAKENHI Grant Numbers 19K10078 and 19K10097 and Uemura Fund, Dental Research Center, Nihon University School of Dentistry.

REFERENCES

1. King AMQ, Lefkowitz EJ, Mushegian AR, Adams MJ, Dutilh BE, Gorbalenya AE, et al. Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2018). *Arch Virol*. (2018) 163:2601–31. doi: 10.1007/s00705-018-3847-1
2. Hamre D, Procknow JJ. A new virus isolated from the human respiratory tract. *Proc Soc Exp Biol Med*. (1966) 121:190–3. doi: 10.3181/00379727-121-30734
3. Kapikian AZ, James HD Jr, Kelly SJ, Dees JH, Turner HC, Mcintosh K, et al. Isolation from man of "avian infectious bronchitis virus-like" viruses (coronaviruses) similar to 229E virus, with some epidemiological observations. *J Infect Dis*. (1969) 119:282–90. doi: 10.1093/infdis/119.3.282
4. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med*. (2003) 348:1953–66. doi: 10.1056/NEJMoa030781
5. Fouchier RA, Hartwig NG, Bestebroer TM, Niemeyer B, De Jong JC, Simon JH, et al. A previously undescribed coronavirus associated with respiratory disease in humans. *Proc Natl Acad Sci USA*. (2004) 101:6212–6. doi: 10.1073/pnas.0400762101
6. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J Virol*. (2005) 79:884–95. doi: 10.1128/JVI.79.2.884-895.2005
7. Zaki AM, Van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med*. (2012) 367:1814–20. doi: 10.1056/NEJMoa1211721
8. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*. (2020) 382:727–33. doi: 10.1056/NEJMoa2001017
9. Hulswit RJ, De Haan CA, Bosch BJ. Coronavirus spike protein and tropism changes. *Adv Virus Res*. (2016) 96:29–57. doi: 10.1016/bs.aivir.2016.08.004
10. Li F. Structure, function, and evolution of coronavirus spike proteins. *Annu Rev Virol*. (2016) 3:237–61. doi: 10.1146/annurev-virology-110615-042301
11. Lu G, Wang Q, Gao GF. Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends Microbiol*. (2015) 23:468–78. doi: 10.1016/j.tim.2015.06.003
12. Millet JK, Whittaker GR. Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. *Virus Res*. (2015) 202:120–34. doi: 10.1016/j.virusres.2014.11.021
13. Bosch BJ, Van Der Zee R, De Haan CA, Rottier PJ. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol*. (2003) 77:8801–11. doi: 10.1128/JVI.77.16.8801-8811.2003
14. Wang N, Shang J, Jiang S, Du L. Subunit vaccines against emerging pathogenic human coronaviruses. *Front Microbiol*. (2020) 11:298. doi: 10.3389/fmicb.2020.00298
15. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol*. (2015) 1282:1–23. doi: 10.1007/978-1-4939-2438-7_1
16. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, Lim W, et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet*. (2003) 361:1319–25. doi: 10.1016/S0140-6736(03)13077-2
17. Liu SL, Saif L. Emerging viruses without borders: the Wuhan coronavirus. *Viruses*. (2020) 12:130. doi: 10.3390/v12020130
18. Moore JB, June CH. Cytokine release syndrome in severe COVID-19. *Science*. (2020) 368:473–4. doi: 10.1126/science.abb8925
19. Tay MZ, Poh CM, Renia L, Macary PA, Ng LFP. The trinity of COVID-19: immunity, inflammation and intervention. *Nat Rev Immunol*. (2020) 20:363–74. doi: 10.1038/s41577-020-0311-8
20. Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res*. (2005) 33:2302–9. doi: 10.1093/nar/gki524
21. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc*. (2009) 4:363–71. doi: 10.1038/nprot.2009.2
22. Herraez A. Biomolecules in the computer: Jmol to the rescue. *Biochem Mol Biol Educ*. (2006) 34:255–61. doi: 10.1002/bmb.2006.494034042644
23. Vehlow C, Stehr H, Winkelmann M, Duarte JM, Petzold L, Dinse J, et al. CMView: interactive contact map visualization and analysis. *Bioinformatics*. (2011) 27:1573–4. doi: 10.1093/bioinformatics/btr163
24. Holm L, Sander C. Mapping the protein universe. *Science*. (1996) 273:595–603. doi: 10.1126/science.273.5275.595
25. Berman HM, Burley SK, Chiu W, Sali A, Adzhubei A, Bourne PE, et al. Outcome of a workshop on archiving structural models of biological macromolecules. *Structure*. (2006) 14:1211–7. doi: 10.1016/j.str.2006.06.005
26. Hevener KE, Zhao W, Ball DM, Babaoglu K, Qi J, White SW, et al. Validation of molecular docking programs for virtual screening against dihydropteroate synthase. *J Chem Inf Model*. (2009) 49:444–60. doi: 10.1021/ci800293n
27. Li F. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. *J Virol*. (2015) 89:1954–64. doi: 10.1128/JVI.02615-14
28. Chiu SS, Chan KH, Chu KW, Kwan SW, Guan Y, Poon LL, et al. Human coronavirus NL63 infection and other coronavirus infections in children hospitalized with acute respiratory disease in Hong Kong, China. *Clin Infect Dis*. (2005) 40:1721–9. doi: 10.1086/430301
29. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol*. (2020) 5:562–9. doi: 10.1038/s41564-020-0688-y
30. Kufareva I, Abagyan R. Methods of protein structure comparison. *Methods Mol Biol*. (2012) 857:231–57. doi: 10.1007/978-1-61779-588-6_10
31. Pyrc K, Dijkman R, Deng L, Jebbink MF, Ross HA, Berkhout B, et al. Mosaic structure of human coronavirus NL63, one thousand years of evolution. *J Mol Biol*. (2006) 364:964–73. doi: 10.1016/j.jmb.2006.09.074
32. Yang J, Zhang W, He B, Walker SE, Zhang H, Govindarajoo B, et al. Template-based protein structure prediction in CASP11 and retrospect of I-TASSER in the last decade. *Proteins*. (2015) 84 (Suppl. 1):233–46. doi: 10.1002/prot.24918
33. Wang Z, Xu J. Predicting protein contact map using evolutionary and physical constraints by integer programming. *Bioinformatics*. (2013) 29:i266–73. doi: 10.1093/bioinformatics/btt211
34. Bittrich S, Schroeder M, Labudde D. StructureDistiller: structural relevance scoring identifies the most informative entries of a contact map. *Sci Rep*. (2019) 9:18517. doi: 10.1038/s41598-019-55047-4
35. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for

virus origins and receptor binding. *Lancet*. (2020) 395:565–74. doi: 10.1016/S0140-6736(20)30251-8
36. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. (2020) 579:270–3. doi: 10.1038/s41586-020-2012-7
37. Robson B. Computers and viral diseases. Preliminary bioinformatics studies on the design of a synthetic vaccine and a preventative peptidomimetic antagonist against the SARS-CoV-2 (2019-nCoV, COVID-19) coronavirus. *Comput Biol Med*. (2020) 119:103670. doi: 10.1016/j.compbiomed.2020.103670
38. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell*. (2020) 181:1489–501 e1415. doi: 10.1016/j.cell.2020.05.015
39. Sette A, Crotty S. Pre-existing immunity to SARS-CoV-2: the knowns and unknowns. *Nat Rev Immunol*. (2020) 20:457–8. doi: 10.1038/s41577-020-0389-z
40. Vojtek I, Buchy P, Doherty TM, Hoet B. Would immunization be the same without cross-reactivity? *Vaccine*. (2019) 37:539–49. doi: 10.1016/j.vaccine.2018.12.005
41. Salvatori G, Luberto L, Maffei M, Aurisicchio L, Roscilli G, Palombo F, et al. SARS-CoV-2 SPIKE PROTEIN: an optimal immunological target for vaccines. *J Transl Med*. (2020) 18:222. doi: 10.1186/s12967-020-02392-y
42. Jerne NK. Immunological speculations. *Annu Rev Microbiol*. (1960) 14:341–58. doi: 10.1146/annurev.mi.14.100160.002013
43. Benjamin DC, Berzofsky JA, East IJ, Gurd FR, Hannum C, Leach SJ, et al. The antigenic structure of proteins: a reappraisal. *Annu Rev Immunol*. (1984) 2:67–101. doi: 10.1146/annurev.iy.02.040184.000435
44. Gershoni JM, Roitburd-Berman A, Siman-Tov DD, Tarnovitski Freund N, Weiss Y. Epitope mapping: the first step in developing epitope-based vaccines. *BioDrugs*. (2007) 21:145–56. doi: 10.2165/00063030-200721030-00002
45. Haste Andersen P, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci*. (2006) 15:2558–67. doi: 10.1110/ps.062405906
46. Sewell AK. Why must T cells be cross-reactive? *Nat Rev Immunol*. (2012) 12:669–77. doi: 10.1038/nri3279
47. Fung TS, Liu DX. Human coronavirus: host-pathogen interaction. *Annu Rev Microbiol*. (2019) 73:529–57. doi: 10.1146/annurev-micro-020518-115759
48. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. (2020) 181:271–80 e278. doi: 10.1016/j.cell.2020.02.052
49. Lv H, Wu NC, Tsang OT, Yuan M, Perera R, Leung WS, et al. Cross-reactive antibody response between SARS-CoV-2 and SARS-CoV Infections. *Cell Rep*. (2020) 31:107725. doi: 10.1016/j.celrep.2020.107725
50. Pinto D, Park YJ, Beltramello M, Walls AC, Tortorici MA, Bianchi S, et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature*. (2020) 583:290–5. doi: 10.1038/s41586-020-2349-y
51. Chan KH, Cheng VC, Woo PC, Lau SK, Poon LL, Guan Y, et al. Serological responses in patients with severe acute respiratory syndrome coronavirus infection and cross-reactivity with human coronaviruses 229E, OC43, and NL63. *Clin Diagn Lab Immunol*. (2005) 12:1317–21. doi: 10.1128/CDLI.12.11.1317-1321.2005
52. Che XY, Qiu LW, Liao ZY, Wang YD, Wen K, Pan YX, et al. Antigenic cross-reactivity between severe acute respiratory syndrome-associated coronavirus and human coronaviruses 229E and OC43. *J Infect Dis*. (2005) 191:2033–7. doi: 10.1086/430355
53. Patrick DM, Petric M, Skowronski DM, Guasparini R, Booth TF, Krajden M, et al. An outbreak of human coronavirus OC43 infection and serological cross-reactivity with SARS coronavirus. *Can J Infect Dis Med Microbiol*. (2006) 17:330–6. doi: 10.1155/2006/152612
54. Chan KH, Chan JF, Tse H, Chen H, Lau CC, Cai JP, et al. Cross-reactive antibodies in convalescent SARS patients' sera against the emerging novel human coronavirus EMC (2012) by both immunofluorescent and neutralizing antibody tests. *J Infect*. (2013) 67:130–40. doi: 10.1016/j.jinf.2013.03.015
55. Dijkman R, Jebbink MF, El Idrissi NB, Pyrc K, Muller MA, Kuijpers TW, et al. Human coronavirus NL63 and 229E seroconversion in children. *J Clin Microbiol*. (2008) 46:2368–73. doi: 10.1128/JCM.00533-08
56. Weiskopf D, Schmitz KS, Raadsen MP, Grifoni A, Okba NMA, Endeman H, et al. Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci Immunol*. (2020) 5:eabd2071. doi: 10.1126/sciimmunol.abd2071
57. Gorse GJ, Patel GB, Vitale JN, O'connor TZ. Prevalence of antibodies to four human coronaviruses is lower in nasal secretions than in serum. *Clin Vaccine Immunol*. (2010) 17:1875–80. doi: 10.1128/CVI.00278-10
58. Callow KA, Parry HF, Sergeant M, Tyrrell DA. The time course of the immune response to experimental coronavirus infection of man. *Epidemiol Infect*. (1990) 105:435–46. doi: 10.1017/S0950268800048019
59. Channappanavar R, Zhao J, Perlman S. T cell-mediated immune response to respiratory coronaviruses. *Immunol Res*. (2014) 59:118–28. doi: 10.1007/s12026-014-8534-z
60. Sariol A, Perlman S. Lessons for COVID-19 immunity from other coronavirus infections. *Immunity*. (2020) 53:248–63. doi: 10.1016/j.immuni.2020.07.005

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.

frontiers in Public Health

BRIEF RESEARCH REPORT
published: 22 January 2021
doi: 10.3389/fpubh.2020.597753

COVID-19 Higher Mortality in Chinese Regions With Chronic Exposure to Lower Air Quality

Riccardo Pansini[1,2*†] *and Davide Fornacca*[1,3*†]

[1] *Institute of Eastern-Himalaya Biodiversity Research, Dali University, Dali, China,* [2] *Behavioral and Experimental Economics Research Center, Statistic and Mathematics College, Yunnan University of Finance and Economics, Kunming, China,* [3] *Institute for Environmental Sciences, University of Geneva, Geneva, Switzerland*

We investigated the geographical character of the COVID-19 infection in China and correlated it with satellite- and ground-based measurements of air quality. Controlling for population density, we found more viral infections in those prefectures (U.S. county equivalent) afflicted by high Carbon Monoxide, Formaldehyde, PM 2.5, and Nitrogen Dioxide values. Higher mortality was also correlated with relatively poor air quality. When summarizing the results at a greater administrative level, we found that the 10 provinces (U.S. state equivalent) with the highest rate of mortality by COVID-19, were often the most polluted but not the most densely populated. Air pollution appears to be a risk factor for the incidence of this disease, despite the conventionally apprehended influence of human mobility on disease dynamics from the site of first appearance, Wuhan. The raw correlations reported here should be interpreted in a broader context, accounting for the growing evidence reported by several other studies. These findings warn communities and policymakers on the implications of long-term air pollution exposure as an ecological, multi-scale public health issue.

Keywords: air pollution, SARS-CoV-2, risk factors, virulence, climate change

OPEN ACCESS

Edited by:
Matteo Convertino, Hokkaido University, Japan

Reviewed by:
Sneha Gautam, Karunya Institute of Technology and Sciences, India
Elza Bontempi, University of Brescia, Italy

***Correspondence:**
Riccardo Pansini
r.pansini@gmail.com
Davide Fornacca
fornacca@eastern-himalaya.cn

[†] *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Environmental Health, a section of the journal Frontiers in Public Health

Received: *21 August 2020*
Accepted: *07 December 2020*
Published: *22 January 2021*

Citation:
Pansini R and Fornacca D (2021) COVID-19 Higher Mortality in Chinese Regions With Chronic Exposure to Lower Air Quality. Front. Public Health 8:597753. doi: 10.3389/fpubh.2020.597753

HIGHLIGHTS

- There is a significant correlation between air pollution and COVID-19 spread and mortality in China.
- The correlation stands at a second-order administration level for several air pollutants, after controlling for varying population densities and removing Wuhan and Hubei from the dataset.
- Living in an area with low air quality is a risk factor for becoming infected and dying from this new form of coronavirus.

INTRODUCTION

COVID-19, initially detected in China and rapidly spread to the rest of the world, has ignited a pandemic causing exorbitant human and economic cost (1). Within a few months since its discovery in December 2019, eastern and western doctors, biologists, and sociologists alike have turned their attention to disentangling the etiology of this airborne disease, a highly contagious respiratory illness caused by a novel coronavirus (2). Various risk factors have been implicated with the fast spread of the virus, assuming different characters, whether considered within or

between countries. Even if free health care was dispensed to everyone in the exceptional case of this epidemics, the Chinese health system, like those of many other countries, is not adequate without proper identification and evaluation of the multiple epidemiological risk factors (3). On an individual level, an older age, the male gender and smoking status have all been shown to increase the coronavirus, SARS-CoV-2 (4). In particular, the angiotensin-converting enzyme 2 (ACE2) receptors in our respiratory system, hit by smoking and this new coronavirus, can bind with air pollutants (5).

From the standpoint of the natural sciences and geography, we can take a broader perspective to appreciate how a coronavirus, transmitted once more from an animal species to us, may show certain patterns in the way it affects and spreads among people, which go beyond virological and medical mechanisms, spatial proximity or apparently chaotic patterns. For example, elements including human and livestock overpopulation, biodiversity loss and climate change played a critical role in making the ground suitable for yet again a new epidemics to flourish (6). A multidisciplinary approach to study cultural and socioeconomic factors may be included when studying the likeliness of the populations to show stronger morbidity to this disease (7). Pertaining to climate change, air pollution is notoriously known to cause health problems and, in particular, viral respiratory infections and pneumonia to individuals chronically exposed to air pollutants (8, 9).

We therefore hypothesize a numerical and geographical association between chronic exposure to air pollution and the spread of SARS-CoV-2 (10). We investigated this possible correlation taking China as a unique case study (11), and have updated and expanded these findings here. A positive correlation had been found between chronic high levels of air pollution perceived as particulate matter found in 9 large Asian cities (three of those being Chinese) and higher lethality related to COVID-19 (12). Despite the strong containment measures adopted over there, if pollution still plays a role, it should be considered as an element of high concern in relation to this disease.

DATA AND METHODS

We collected COVID-19 infection and fatality figures for every prefecture of the People's Republic of China (2nd administrative divisions, equivalent to U.S. counties) from the Chinese government health commission (**Table 1**). We normalized these epidemiological values per 10,000 inhabitants of each prefecture. The data included COVID-19 cases and deaths.

The dataset of COVID-19 cases and deaths analyzed in this study captured the first and unique wave of SARS-CoV-2 infection for this country (19 December 2019–23 May 2020). It includes the 17 April update, when an increase of 1,290 casualties was reported, following a revised WHO guideline, showing a drastic rise of about 50% from the prior figure. This update included an increase of 325 infections for the city of Wuhan only.

While infections and fatalities inform on the extent of the pandemic, mortality rates (fatalities/infections * 100) provide additional information on the severity of the virus in each prefecture. It is important to note that China did not see a systematic COVID-19 testing at a national scale. Tests were mostly performed for people presenting symptoms and registered in hospitals. In some later cases, large scale testing was performed only to prevent localized outbreaks. As a result, asymptomatic cases are not included in the data, and mortality rates may appear inflated compared to other regions of the world.

Population densities were defined using the population totals of each prefecture divided by its surface area. Air pollution measurements from localized ground stations (monthly averages 2014–2016) as well as continuous tropospheric vertical column density measurements (whole year 2019) of several air pollutants were aggregated as the average values at the prefecture level. Time series information of atmospheric air pollutants was retrieved by the Sentinel-5, a satellite mission launched in October 2017 as part of the Copernicus program of the European Space Agency (13). The Google Earth Engine platform (14) was employed to compute the 2019 averages of each air pollutant measurements derived from satellite, namely the UV Aerosol Index, Carbon Monoxide (CO), Formaldehyde (HCHO), Nitrogen Dioxide (NO_2), Ozone (O_3), and Sulfur Dioxide (SO_2). Air pollutants collected from ground stations were PM 2.5, PM 10, O_3, NO_2, SO_2, and CO. Data types and their sources are shown in **Table 1** and the fully compiled dataset, including an aggregated version at the provincial level, are available on a dedicated GitHub repository (https://github.com/DavideFornacca/COVID19/tree/master/China).

Correlation and significance analyses between air pollution, population, and the three COVID-19 variables (infections/100,000 inhabitants, fatalities/100,000 inhabitants, mortality rate) were performed for the prefecture-level dataset using non-parametric Kendall rank correlation coefficient because of the distributions of COVID-19 and population variables being mostly skewed. To assess the potential influence of outliers, we repeated the same tests by firstly removing the prefecture of Wuhan and then the whole Hubei province from the dataset. The significance threshold was set to <0.05. Using the aggregated version of the dataset (mean values at the provincial level), we identified the first 10 Chinese provinces showing the highest values of each variable separately, and we used this for comparative analysis. Furthermore, thematic maps comparing COVID-19 and air pollution distributions in China were produced for visual assessment.

Data processing and mapping was done with QGIS. Statistical analysis was performed in Python programming environment.

RESULTS

A descriptive statistics' summatory table for the different satellite- and ground-based air quality measurements can be found in **Supplementary Table 1**.

Higher amounts of viral infections per 100,000 inhabitants, fatalities per 100,000 inhabitants, and mortality rates (fatalities/infections * 100) were found in those Chinese prefectures afflicted by several pollutants of the air: CO, HCHO, PM 2.5, PM 10, and NO_2, as shown by the significant positive correlation coefficients in **Table 2**. In particular, stronger associations for infections and fatalities were found with

TABLE 1 | The analyzed datasets and their sources.

Data	Measuring unit	Time period	Format	Source
COVID-19	No. of infections, No. of deaths	Updated on 23 May 2020	Tabular prefecture level	DXY - DX Doctor: http://ncov.dxy.cn/ncovh5/view/en_pneumonia from Chinese government health commission
Population	No. of residents	Estimates 2017	Tabular prefecture level	https://www.citypopulation.de/ Data from provincial governments
AIR QUALITY, GROUND STATIONS				
PM2.5, PM10, O_3, NO_2, SO_2, CO	Air Quality Index (AQI)	2014-2016	Tabular GPS points	University of Harvard Dataverse: https://dataverse.harvard.edu Data from http://aqicn.org
AIR QUALITY, SATELLITE				
UV Aerosol Index	Qualitative Index	2019	Continuous grid (0.01 arc deg.)	Sentinel-5 Atmospheric variables https://developers.google.com/earth-engine/datasets/tags/air-quality
CO, HCHO, NO_2, O_3, SO_2	mol/m^2			

TABLE 2 | Correlation between satellite- and ground-based air quality variables with (i) cumulated COVID-19 infections per 100,000 inhabitants, (ii) fatalities per 100,000 inhabitants, and (iii) mortality rate (fatalities / infections) in China at a prefectural level, until 23 May 2020.

	Infections (/100k pop)			Fatalities (/100k pop)			Mortality (fatalities/infections)		
	df (n-2)	tau	*p*-value	df (n-2)	tau	*p*-value	df (n-2)	tau	*p*-value
CO sat	337	0.28	**<0.001**	337	0.19	**<0.001**	313	0.16	**<0.001**
NO_2 sat	337	0.23	**<0.001**	337	0.14	**0.001**	313	0.12	**0.006**
O_3 sat	337	−0.08	**0.030**	337	0.00	0.967	313	0.02	0.635
SO_2 sat	337	−0.10	**0.005**	337	−0.02	0.634	313	0.00	0.964
Aerosol sat	337	−0.12	**0.001**	337	−0.03	0.488	313	0.00	0.950
HCHO sat	337	0.34	**<0.001**	337	0.20	**<0.001**	313	0.17	**<0.001**
PM 2.5 ground	302	0.15	**<0.001**	302	0.18	**<0.001**	285	0.18	**<0.001**
PM 10 ground	302	0.04	0.330	302	0.12	**0.006**	285	0.13	**0.005**
CO ground	302	−0.01	0.840	302	0.11	**0.012**	285	0.12	**0.007**
NO_2 ground	302	0.12	**0.002**	302	0.12	**0.005**	285	0.12	**0.007**
O_3 ground	302	−0.03	0.477	302	−0.02	0.585	285	−0.03	0.482
SO_2 ground	302	−0.01	0.843	302	0.04	0.409	285	0.06	0.178
population	337	0.23	**<0.001**	337	0.17	**<0.001**	313	0.14	**<0.001**
pop density	337	0.32	**<0.001**	337	0.16	**<0.001**	313	0.12	**0.004**

Note that the degrees of freedom (df) are different for the ground station results because of the limited data availability. $p < 0.05$ are marked in bold characters.

tropospheric column values of Formaldehyde ($r_\tau = 0.34$, $p < 0.001$ and $r_\tau = 0.20$, $p < 0.001$) and Carbon Monoxide values ($r_\tau = 0.28$, $p < 0.001$ and $r_\tau = 0.20$, $p < 0.001$), while for mortality rates, PM 2.5 was the most incident pollutant ($r_\tau = 0.18$, $p < 0.001$). This trend holds also after removing in succession (i) Wuhan city and (ii) the whole Hubei province from the dataset (see **Supplementary Tables 3**, **4**). Levels of particulate matter measured by ground stations, especially the finer PM 2.5, were associated with a greater number of fatalities and higher mortality rates. Conversely, aerosol data from the satellite, which potentially include PM 2.5 and PM 10, were not associated with fatalities or mortality rates. They negatively correlated with infections weakly. This is not surprising, given that the measurement is related to UV-absorbing particles, which are in general non-pollutant, being inert particulates such as dust, sand, and sea salt, but they also include smoke from volcano ash and biomass burning. To note that these sources of dust, however, are often found far from highly-polluted development areas. Higher levels of O_3 and SO_2 from both satellite and ground data were not associated with more COVID-19 deaths and mortality rates. This goes against the trend shown by the other pollutants, an aspect that will require further investigation. Levels of CO, HCHO, and PM 2.5 showed stronger correlation than the population variables when analyzing fatalities and mortality rates. As expected, several air pollutants correlated with population density except for Sulfur Dioxide. Conversely, O_3 and Aerosol index showed weak negative correlations, suggesting their presence in low populated areas (**Supplementary Table 2**).

TABLE 3 | Summary table of mean values of COVID-19 (orange), air pollution (blue), and population density (green) variables at the provincial level.

Province (prefectures *n*)	Pop density (pop/m^2)	Total infections	Infections /100k pop	Total fatalities	Fatalities /100k pop	Mortality (%)	CO sat (μmol/m^2)	NO_2 sat (μmol/m^2)	O_3 sat (μmol/m^2)	SO_2 sat (μmol/m^2)	Aerosol sat (index)	HCHO sat (μmol/m^2)	PM25 gr (AQI)	PM10 gr (AQI)	CO gr (AQI)	NO_2 gr (AQI)	O_3 gr (AQI)	SO_2 gr (AQI)
Hubei (15)	378.92	68,135	115.44 ± 106.39	4,512	7.6449 ± 8.3312	6.62 ± 1.63	46,393 ± 4845	51.60 ± 21.80	123,670 ± 1,342	20.17 ± 13.46	−1.04 ± 0.03	176.50 ± 27.82	133 ± 13	71 ± 12	10 ± 3	13 ± 4	24 ± 5	11 ± 3
Hainan (3)	541.16	169	1.83 ± 3.25	6	0.0648 ± 0.0660	3.55 ± 3.44	40,725 ± 1,175	21.00 ± 4.42	116,021 ± 101	−19.02 ± 8.73	−1.11 ± 0.09	146.33 ± 17.63	61 ± 5	29 ± 4	6 ± 0	6 ± 1	20 ± 2	2 ± 1
Heilongjiang (13)	93.29	559	1.47 ± 0.98	13	0.0341 ± 0.0580	2.33 ± 2.76	38,689 ± 1,588	21.34 ± 8.40	165,885 ± 2,284	93.47 ± 37.60	−0.93 ± 0.08	94.29 ± 16.13	91 ± 20	49 ± 11	6 ± 2	10 ± 4	24 ± 5	9 ± 4
Gansu (14)	124.46	91	0.35 ± 0.34	2	0.0076 ± 0.0142	2.20 ± 1.68	28,325 ± 3,199	25.23 ± 13.34	132,895 ± 4,450	72.38 ± 29.55	−0.79 ± 0.20	91.62 ± 17.26	98 ± 13	69 ± 14	10 ± 3	14 ± 5	30 ± 9	13 ± 5
Jilin (9)	165.48	154	0.56 ± 0.20	3	0.0109 ± 0.0098	1.95 ± 2.36	40,300 ± 1,983	36.23 ± 14.48	158,741 ± 2,422	105.60 ± 11.59	−0.93 ± 0.15	104.95 ± 12.74	113 ± 19	64 ± 12	9 ± 2	13 ± 3	24 ± 2	12 ± 3
Hebei (11)	514.53	318	0.42 ± 0.25	6	0.0080 ± 0.0144	1.89 ± 3.40	51,139 ± 7,776	129.64 ± 46.39	144,916 ± 4,197	190.11 ± 46.96	−0.92 ± 0.08	175.63 ± 41.13	157 ± 30	108 ± 28	12 ± 3	21 ± 5	24 ± 3	23 ± 6
Xinjiang (16)	145.23	56	0.23 ± 0.23	1	0.0041 ± 0.0490	1.79 ± 9.45	28,671 ± 4,747	23.41 ± 20.33	142,304 ± 4,832	68.02 ± 74.28	−0.53 ± 0.32	78.33 ± 17.42	108 ± 32	87 ± 43	11 ± 3	13 ± 5	24 ± 5	7 ± 2
Henan (17)	675.57	1,273	1.33 ± 0.87	22	0.0229 ± 0.0189	1.73 ± 3.40	49,781 ± 3,351	96.52 ± 30.76	131,876 ± 2,818	93.28 ± 38.83	−1.01 ± 0.03	190.87 ± 14.88	145 ± 8	91 ± 9	15 ± 4	19 ± 4	23 ± 6	21 ± 7
Taiwan (1)	648.32	441	1.87	7	0.0297	1.59	34,858	39.1	115,054	−21.2	−1.09	121.27	–	–	–	–	–	–
Liaoning (14)	319.55	128	0.29 ± 0.20	2	0.0046 ± 0.0133	1.56 ± 4.82	45,853 ± 2,591	72.03 ± 21.47	153,539 ± 2,061	154.65 ± 17.93	−0.97 ± 0.08	117.78 ± 8.57	120 ± 10	70 ± 8	13 ± 3	17 ± 3	25 ± 4	22 ± 5
Tianjin (1)	1330.81	192	1.23	3	0.0193	1.56	54,124	175.04	146,921	193.42	−0.87	198.66	135	84	14	21	21	17
Beijing (1)	1313.10	593	2.75	9	0.0418	1.52	45,974	123.19	148,350	134.26	−0.94	173.28	145	83	11	24	27	11
Guizhou (9)	244.30	146	0.41 ± 0.19	2	0.0056 ± 0.0125	1.37 ± 3.35	38,014 ± 1,975	30.87 ± 7.21	117,525 ± 910	21.63 ± 12.63	−0.96 ± 0.07	136.68 ± 5.70	106 ± 7	54 ± 7	7 ± 0	13 ± 1	20 ± 1	10 ± 1
Inner Mongolia (12)	73.21	77	0.30 ± 0.33	1	0.0039 ± 0.0171	1.30 ± 3.77	33,581 ± 2,486	33.47 ± 32.68	150,964 ± 8,343	92.09 ± 48.43	−0.78 ± 0.15	81.38 ± 13.90	90 ± 18	58 ± 18	9 ± 6	11 ± 6	27 ± 7	14 ± 10
Shaanxi (10)	280.14	243	0.63 ± 0.41	3	0.0078 ± 0.0106	1.23 ± 0.79	37,972 ± 3,157	51.47 ± 26.39	130,735 ± 4,029	69.58 ± 22.96	−0.97 ± 0.04	140.57 ± 20.20	121 ± 20	72 ± 16	15 ± 5	16 ± 4	20 ± 3	16 ± 11
Yunnan (18)	127.40	174	0.36 ± 0.33	2	0.0041 ± 0.0229	1.15 ± 2.51	31,788 ± 4,068	18.43 ± 6.32	115,465 ± 593	18.41 ± 15.22	−1.23 ± 0.08	118.73 ± 22.77	74 ± 12	37 ± 8	7 ± 2	7 ± 2	26 ± 6	8 ± 4

(Continued)

TABLE 3 | Continued

Province (prefectures *n*)	Pop density (pop/m^2)	Total infections	Infections /100k pop	Total fatalities	Fatalities /100k pop	Mortality (%)	CO sat (μmol/m^2)	NO_2 sat (μmol/m^2)	O_3 sat (μmol/m^2)	SO_2 sat (μmol/m^2)	Aerosol sat (index)	HCHO sat (μmol/m^2)	PM25 gr (AQI)	PM10 gr (AQI)	CO gr (AQI)	NO_2 gr (AQI)	O_3 gr (AQI)	SO_2 gr (AQI)
Shanghai (1)	3513.05	667	2.76	7	0.0289	1.05	45,641	178.15	124,858	54.4	−0.94	156.15	116	57	7	21	35	9
Chongqing (1)	376.25	579	1.87	6	0.0193	1.04	41,471	35.8	121,772	20	−0.93	150.42	124	62	9	19	18	10
Shandong (15)	641.28	763	0.76 ± 0.69	7	0.0070 ± 0.0140	0.92 ± 2.06	51,789 ± 2,492	125.14 ± 22.32	138,473 ± 2,630	152.21 ± 23.38	−1.00 ± 0.03	178.85 ± 21.53	146 ± 22	94 ± 18	11 ± 4	21 ± 4	31 ± 6	25 ± 8
Guangxi (14)	247.65	252	0.51 ± 0.76	2	0.0041 ± 0.0171	0.79 ± 1.13	43,533 ± 1,523	29.83 ± 4.08	116,588 ± 312	−16.53 ± 8.68	−0.94 ± 0.04	160.19 ± 8.65	97 ± 13	47 ± 7	9 ± 2	9 ± 3	26 ± 7	9 ± 3
Anhui (15)	673.01	991	1.57 ± 1.26	6	0.0095 ± 0.0359	0.61 ± 0.78	47,941 ± 1,974	80.50 ± 22.67	126,466 ± 2,866	56.81 ± 13.74	−1.08 ± 0.03	181.35 ± 7.91	117 ± 16	59 ± 10	8 ± 2	13 ± 3	25 ± 5	11 ± 4
Sichuan (19)	385.74	563	0.68 ± 1.36	3	0.0036 ± 0.0041	0.53 ± 0.40	38,603 ± 7,695	37.53 ± 16.58	121,486 ± 2,521	36.45 ± 23.37	−0.97 ± 0.12	144.81 ± 35.70	109 ± 27	57 ± 14	8 ± 2	13 ± 4	25 ± 6	9 ± 4
Guangdong (19)	1270.33	1,590	1.40 ± 1.39	8	0.0071 ± 0.0143	0.50 ± 2.42	42,819 ± 1,462	66.15 ± 43.52	116,397 ± 212	−32.14 ± 13.12	−1.05 ± 0.07	168.06 ± 23.56	93 ± 7	44 ± 4	9 ± 1	12 ± 4	28 ± 5	7 ± 2
Hunan (14)	356.49	1,018	1.48 ± 0.82	4	0.0058 ± 0.0082	0.39 ± 0.35	45,067 ± 2,212	37.94 ± 10.56	119,555 ± 1,383	1.25 ± 9.58	−0.99 ± 0.03	173.23 ± 17.23	114 ± 15	59 ± 8	9 ± 2	11 ± 4	27 ± 6	12 ± 3
Hong Kong SAR (1)	6598.01	1,065	14.29	4	0.0537	0.38	42,420	139.17	116,276	−48.4	−1.02	176.42	92	46	10	17	28	4
Fujian (9)	652.50	296	0.75 ± 0.49	1	0.0025 ± 0.0043	0.34 ± 0.46	39,331 ± 848	36.99 ± 12.96	117,021 ± 627	−10.92 ± 11.61	−1.09 ± 0.05	142.37 ± 10.48	76 ± 10	37 ± 7	7 ± 1	10 ± 3	24 ± 7	5 ± 2
Jiangxi (11)	340.40	934	2.02 ± 0.23	1	0.0022 ± 0.0035	0.11 ± 0.40	45,074 ± 1,956	43.39 ± 10.75	119,799 ± 1,251	18.69 ± 14.93	−1.07 ± 0.03	173.25 ± 11.72	100 ± 7	51 ± 7	7 ± 2	10 ± 3	21 ± 4	12 ± 4
Zhejiang (11)	641.36	1,182	2.09 ± 1.45	1	0.0018 ± 0.0033	0.08 ± 0.06	42,655 ± 2,191	71.76 ± 34.48	121,312 ± 1,686	20.05 ± 14.83	−1.05 ± 0.05	163.04 ± 19.40	114 ± 10	55 ± 6	8 ± 1	17 ± 4	30 ± 4	9 ± 3
Jiangsu (13)	856.03	631	0.78 ± 2.99	0	0	0	48,727 ± 1,420	125.84 ± 30.98	128,531 ± 2,714	75.05 ± 24.14	−1.05 ± 0.03	175.42 ± 13.30	120 ± 8	66 ± 8	5 ± 1	16 ± 4	29 ± 3	12 ± 3
Shanxi (11)	270.16	134	0.36 ± 0.27	0	0	0	39,921 ± 4,031	89.14 ± 17.85	138,704 ± 3,975	150.68 ± 27.36	−1.05 ± 0.05	136.37 ± 25.60	125 ± 11	77 ± 14	17 ± 4	17 ± 3	24 ± 5	28 ± 4
Ningxia (5)	155.56	74	1.09 ± 0.85	0	0	0	32,917 ± 3,598	49.00 ± 28.86	136,458 ± 3,274	121.81 ± 43.19	−0.82 ± 0.06	103.16 ± 9.05	112 ± 0	81 ± 1	9 ± 1	15 ± 3	24 ± 1	26 ± 5
Macao SAR (1)	19136.39	45	6.89	0	0	0	43,309	125.72	116,336	−42.5	−0.98	170.45	–	–	–	–	–	–
Qinghai (8)	60.22	18	0.30 ± 0.41	0	0	0	20,107 ± 3,157	13.54 ± 10.12	129,362 ± 3,269	36.08 ± 36.84	−0.89 ± 0.23	61.29 ± 12.75	109 ± 18	71 ± 5	10 ± 3	11 ± 6	31 ± 8	12 ± 3
Xizang (7)	5.70	1	0.03 ± 0.07	0	0	0	17,503 ± 1,374	8.91 ± 1.60	118,934 ± 3,000	6.94 ± 6.13	−0.86 ± 0.22	59.39 ± 5.16	62 ± 25	37 ± 23	8 ± 5	6 ± 2	26 ± 5	6 ± 5

Colored cells represent the 12 highest values for each column. The table is in descending order according to mortality rates. Standard deviations of COVID-19 and pollution variables are also displayed. Sat, satellite measures; gr, ground stations; AQI, Air Quality Index.

A comprehensive statistical output of these data is reported in **Table 2**.

The values shown in **Table 3** include the prefectural detail aggregated at a coarser provincial level, and they are sorted to highlight the 12 provinces with the highest rates of mortality. These provinces are often reported among the 12 most polluted ones, except for Taiwan and Hainan. However, these provinces are not the most densely populated, with the exception of Henan, Taiwan (a densely populated island), Tianjin and Beijing (two relatively small but highly populated provincial-level municipalities).

The maps (**Figure 1**) offer a synoptic view of the correlations. Those ones referring to different pollutants display continuous values at 0.01 arc-degree (about 1 km) resolutions, which can potentially highlight within-prefecture differences. Prefectural mean values of each air pollutant can be found in the dataset available in the dedicated repository. A clear longitudinal pattern ranging from the northeast region down to Hong Kong is visible for both COVID-19 variables and air pollution. In addition to Wuhan and Hubei appearing heavily affected compared to the rest of the country, the map shows mortality occurring in several other hotspots, also in less populated yet industrial provinces of central China.

DISCUSSION AND CONCLUSION

The present study suggests a strong association between the incidence of COVID-19 and chronic exposure to air pollution in China. Comparative analyses made in this study indicate the role of air pollution as a critical risk cofactor for COVID-19 in China, with a stronger influence of Formaldehyde and Carbon Monoxide levels.

Our finding is in line with other studies observing a similar influence of air pollutants. (i) Firstly, testing the more proximal hypothesis that COVID-19 outbreaks could follow with a temporal delay from days with high NO_2 presence in the air, colleagues in Shanghai have published detailed time series data pointing at a lag of 12 days before hospitalizations for the Hubei province (16). This suggests the role of air pollutants as airborne vectors for this virus, also highlighted by another study (18) conducted in three cities in Hubei province and a further one illustrating a potential role of PMs in other Asian cities (12). (ii) Secondly, in the United States, the correlation of respiratory illness with chronic exposure to PM 2.5 was observed stronger than 11 other demographic covariables, including population density, patients' age, socioeconomic status, ethnicity, education, obesity, smoking status, number of hospital beds per unit population, average daily temperature and relative humidity, and lockdown state (15) This same finding for the U.S. was replicated by others, e.g., (17). (iii) Thirdly, in Italy, a similar positive correlation between COVID-19 occurrence and chronic exposure to NO_2, O_3 and PMs was also reported (20) controlling in addition for the extra five demographic variables of mobility, temperature, housing density, health care density, and age of the population (21), with the additional and novel evidence that fragments of the RNA from this virus were found in the particulate matter of the harshly hit northern Italian city of Bergamo (19), laying in the most polluted European area of the Po valley, severely affected by the virus (22). (iv) Finally, our couple of multinational investigations employing different satellite-based datasets has also observed trends similar to the current paper in at least five countries other than China, including Italy, United States, Iran, France, and U.K. (23, 24). Having run and posted these significant correlations at different moments throughout the expansion of the pandemic is an element which further signals that the correlations stand throughout time.

Our study combined data from two different sources: localized ground-stations, directly measuring ambient air pollution, and continuous, grid-based satellite observations. These two sources present several technical and methodological differences that could explain some degree of discordance in the resulting correlation analyses, such as the one found between the satellite-based UV Aerosol Index and the ground values of particulate matter. We maintained the assessment of pollutants from the traditional ground stations because they are more representative of pollution levels to which populations living in their proximity are exposed. Notwithstanding that, they are unevenly scattered and their spatial coverage is very limited. On the other hand, satellite observations offer the unique advantage of global coverage, highlighting pollution differences at a regular and finer spatial unit. We believe that the combined use of these two data sources is instrumental in interpreting any detrimental role of specific air pollutants.

A significant obstacle to the interpretability of these findings in China is the availability of only limited information about the associated covariables, such as high-resolution data regarding health services and infrastructures, epidemiological traits, and population movement, to ascertain the relative importance of air pollution among many socio-environmental driving factors of COVID-19 infections. For example, the spatial effect of the strict lockdown adopted in China can be deduced by the huge difference in the number of infections and fatalities between Wuhan, its surrounding prefectures, and the rest of China (see maps in **Figure 1**). We have considered population density as chief cofactor and found that its correlation with COVID-19 is similar to air pollution, although slightly weaker. Moreover, highly populated areas are often more polluted (**Supplementary Table 2**). These results preclude us from understanding specific health consequences of air pollutants and call for a pressing need to further investigate this matter. In the past, the correlation between air pollution and human illness was notified and attributable to PMs and NO_2 acting as vectors for the spread and extended survival of bioaerosols (25–30) in relation to pathogenic microbes including the avian influenza, measles and the syncytial virus (31–35). To generalize the results of the present study, higher mortality rates were found in provinces with the worst air pollution problems and only some of them were among the most densely populated ones. Regardless, the probability of dying from the virus once it infects is higher where air pollution is heavier. We have also observed a similar pattern for populations at risk of chronic exposure of PM 2.5 and NO_2 in the two less densely-populated countries of Italy and Iran (23).

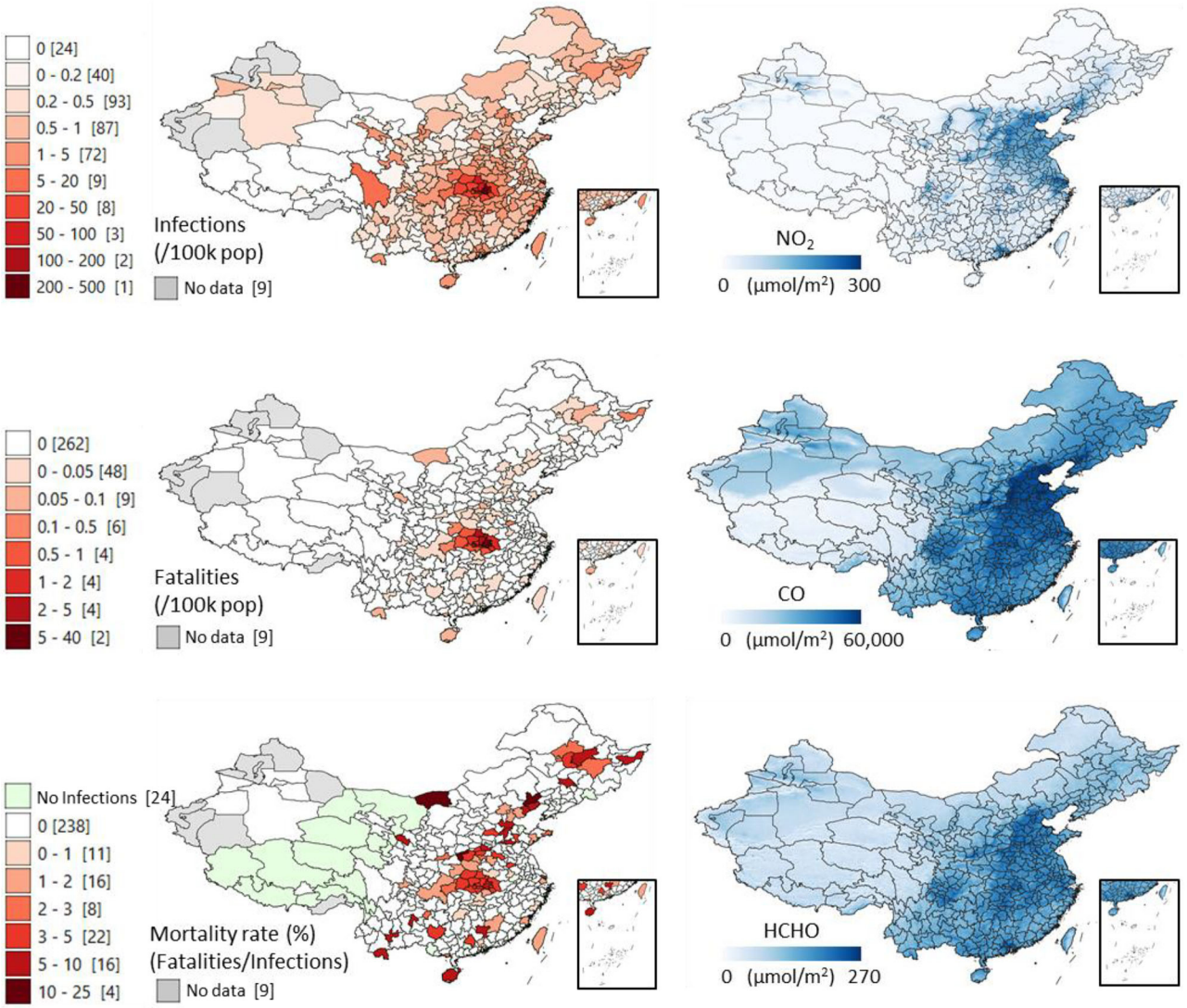


FIGURE 1 | Distribution of COVID-19 infections, fatalities and mortality rates (fatalities/infections * 100) across the prefectures of China (updated on 23 May 2020), and the distribution of the tropospheric column amounts of three representative air pollutants derived from the 2019 averaged satellite measures of: Nitrogen Dioxide (NO_2), Carbon Monoxide (CO), and Formaldehyde (HCHO). The values in the square brackets show the COVID-19 cases' counts of administrative units.

To conclude, despite the fact the SARS-CoV-2 was first detected in Wuhan and that the first location of the pathogen assumes a key role in the geographical spread of the infection, the detrimental effect of air pollution on patients infected by the virus remains evident. To overcome the limitations of our study, longitudinal screenings performed on patients from retrospective cohorts will help clarify the role of air pollution as a cofactor for these types of airborne transmittable diseases (36). In this century, in China as elsewhere, health policymaking is not adequate unless following human and environmental "one health" approaches. As a clear and immediate action to prevent the trajectory of this and future epidemics, curbing climate change (37) must be endorsed way more seriously. Will the smallest of the parasites be able to awaken us, this time, so that we convincingly start caring about the health of the environment as much as we have clumsily been caring about our public health?

DATA AVAILABILITY STATEMENT

The data are available at https://github.com/DavideFornacca/COVID19/tree/master/China.

AUTHOR CONTRIBUTIONS

RP: concept. RP and DF: design, interpretation, drafting of the manuscript, and critical revision of the manuscript. DF: data

acquisition and statistical analysis. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

Livia Ottisova and three reviewers improved the manuscript. Chun Chen commented on statistics. RP did this work while in isolation in the Italian Po valley, due to the ongoing pandemic.

A simpler version of this manuscript was released as a pre-print at medRxiv (11).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2020.597753/full#supplementary-material

REFERENCES

1. Bashir MF, Ma B, Shahzad L. A brief review of socio-economic and environmental impact of Covid-19. *Air Qual Atmos Health.* (2020) 13:1403–9. doi: 10.1007/s11869-020-00894-8
2. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet.* (2020) 395:470–3. doi: 10.1016/S0140-6736(20)30185-9
3. Zhai T, Goss J. Health system reform in China: the challenges of multimorbidity. *Lancet Global Health.* (2020) 8:e750–1. doi: 10.1016/S2214-109X(20)30225-4
4. 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
5. Lukassen S, Chua RL, Trefzer T, Kahn NC, Schneider MA, Muley T, et al. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *EMBO J.* (2020) 39:e105114. doi: 10.15252/embj.2020105114
6. Bedford J, Farrar J, Ihekweazu C, Kang G, Koopmans M, Nkengasong J. A new twenty-first century science for effective epidemic response. *Nature.* (2019) 575:130–6. doi: 10.1038/s41586-019-1717-y
7. Bontempi E, Vergalli S, Squazzoni F. Understanding COVID-19 diffusion requires an interdisciplinary, multi-dimensional approach. *Environ Res.* (2020) 188:109814. doi: 10.1016/j.envres.2020.109814
8. Lelieveld J, Evans JS, Fnais M, Giannadaki D, Pozzer A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature.* (2015) 525:367–71. doi: 10.1038/nature15371
9. Krewski D. Evaluating the effects of ambient air pollution on life expectancy. *N Engl J Med.* (2009) 360:413–5. doi: 10.1056/NEJMe0809178
10. Comunian S, Dongo D, Milani C, Palestini P. Air pollution and Covid-19: the role of particulate matter in the spread and increase of Covid-19's morbidity and mortality. *Int J Environ Res Public Health.* (2020) 17:4487. doi: 10.3390/ijerph17124487
11. Pansini R, Fornacca D. COVID-19 higher morbidity and mortality in Chinese regions with lower air quality. *medRxiv.* (2020). doi: 10.1101/2020.05.28.20115832
12. Gupta A, Bherwani H, Gautam S, Anjum S, Musugu K, Kumar N, et al. Air pollution aggravating COVID-19 lethality? Exploration in Asian cities using statistical models. *Environ Dev Sustain.* (2020) 1–10. doi: 10.1007/s10668-020-00878-9
13. ESA. *European Space Agency - Sentinel-5.* (2018). Available online at: https://developers.google.com/earth-engine/datasets/tags/air-quality (accessed March 25, 2020).
14. Google. *Google Earth Engine.* (2020). Available online at: https://developers.google.com/earth-engine/datasets/catalog/sentinel-5p (accessed March 30, 2020).
15. Wu X, Nethery RC, Sabath MB, Braun D, Dominici F. Air pollution and COVID-19 mortality in the United States: strengths and limitations of an ecological regression analysis. *Sci Adv.* (2020) 6:eabd4049. doi: 10.1126/sciadv.abd4049
16. Yao Y, Pan J, Liu Z, Meng X, Wang W, Kan H, et al. Ambient nitrogen dioxide pollution and spread ability of COVID-19 in Chinese cities. *Ecotoxicol Environ Safety.* (2020) 208:111421. doi: 10.1101/2020.03.31.20048595
17. Liang D, Shi L, Zhao J, Liu P, Sarnat JA, Gao S, et al. Urban air pollution may enhance COVID-19 case-fatality and mortality rates in the United States. *Innovation.* (2020) 1:100047. doi: 10.1016/j.xinn.2020.100047
18. Jiang Y, Wu X-J, Guan Y-J. Effect of ambient air pollutants and meteorological variables on COVID-19 incidence. *Infect Control Hospital Epidemiol.* (2020) 41:1–11. doi: 10.1017/ice.2020.222
19. Setti L, Passarini F, De Gennaro G, Barbieri P, Perrone MG, Borelli M, et al. SARS-Cov-2 RNA found on particulate matter of bergamo in Northern Italy: first evidence. *Environ Res.* (2020) 188:109754. doi: 10.1016/j.envres.2020.109754
20. Fattorini D, Regoli F. Role of the chronic air pollution levels in the Covid-19 outbreak risk in Italy. *Environ Pollut.* (2020) 264:114732. doi: 10.1016/j.envpol.2020.114732
21. Pluchino A, Inturri G, Rapisarda A, Biondo AE, Le Moli R, Zappala' C, et al. A novel methodology for epidemic risk assessment: the case of COVID-19 outbreak in Italy. *arXiv.* 2004.02739 (2020).
22. Setti L, Passarini F, De Gennaro G, Barbieri P, Licen S, Perrone MG, et al. Potential role of particulate matter in the spreading of COVID-19 in Northern Italy: first observational study based on initial epidemic diffusion. *BMJ Open.* (2020) 10:e039338. doi: 10.1136/bmjopen-2020-039338
23. Pansini R, Fornacca D. Higher virulence of COVID-19 in the air-polluted regions of eight severely affected countries. *medRxiv.* (2020). doi: 10.1101/2020.04.30.20086496
24. Pansini R, Fornacca D. Initial evidence of higher morbidity and mortality due to SARS-CoV-2 in regions with lower air quality. *medRxiv.* (2020). doi: 10.1101/2020.04.04.20053595
25. Burge HA. Biological airborne pollutants. In: Foster WM, Costa DL, editors. *Lung Biology in Health Disease*. Boca Raton, FL: Taylor & Francis Group (2005).
26. Wong G, Ko F, Lau T, Li S, Hui D, Pang S, et al. Temporal relationship between air pollution and hospital admissions for asthmatic children in Hong Kong. *Clin Exp Allergy.* (2001) 31:565–9. doi: 10.1046/j.1365-2222.2001.01063.x
27. Smets W, Moretti S, Denys S, Lebeer S. Airborne bacteria in the atmosphere: presence, purpose, and potential. *Atmos Environ.* (2016) 139:214–21. doi: 10.1016/j.atmosenv.2016.05.038
28. Dong L, Qi J, Shao C, Zhong X, Gao D, Cao W, et al. Concentration and size distribution of total airborne microbes in hazy and foggy weather. *Sci Total Environ.* (2016) 541:1011–8. doi: 10.1016/j.scitotenv.2015.10.001
29. Wei K, Zou Z, Zheng Y, Li J, Shen F, Wu C-y, et al. Ambient bioaerosol particle dynamics observed during haze and sunny days in Beijing. *Sci Total Environ.* (2016) 550:751–9. doi: 10.1016/j.scitotenv.2016.01.137
30. Li Y, Fu H, Wang W, Liu J, Meng Q, Wang W. Characteristics of bacterial and fungal aerosols during the autumn haze days in Xi'an, China. *Atmos Environ.* (2015) 122:439–47. doi: 10.1016/j.atmosenv.2015.09.070
31. Cao C, Jiang W, Wang B, Fang J, Lang J, Tian G, et al. Inhalable microorganisms in Beijing's PM2.5 and PM10 pollutants during a severe smog event. *Environ Sci Technol.* (2014) 48:1499–507. doi: 10.1021/es4048472
32. Ye Q, Fu J-f, Mao J-h, Shang S-q. Haze is a risk factor contributing to the rapid spread of respiratory syncytial virus in children. *Environ Sci Pollut Res.* (2016) 23:20178–85. doi: 10.1007/s11356-016-7228-6
33. Chen P-S, Tsai FT, Lin CK, Yang C-Y, Chan C-C, Young C-Y, et al. Ambient influenza and avian influenza virus during dust storm days and background days. *Environ Health Perspect.* (2010) 118:1211–6. doi: 10.1289/ehp.0901782
34. Chen G, Zhang W, Li S, Williams G, Liu C, Morgan GG, et al. Is short-term exposure to ambient fine particles associated with measles incidence in China? A multi-city study. *Environ Res.* (2017) 156:306–11. doi: 10.1016/j.envres.2017.03.046

35. Peng L, Zhao X, Tao Y, Mi S, Huang J, Zhang Q. The effects of air pollution and meteorological factors on measles cases in Lanzhou, China. *Environ Sci Pollut Res.* (2020) 27:13524–33. doi: 10.1007/s11356-020-07903-4
36. Villeneuve Paul J, Goldberg Mark S. Methodological considerations for epidemiological studies of air pollution and the SARS and COVID-19 coronavirus outbreaks. *Environ Health Perspect.* (2020) 128:095001. doi: 10.1289/EHP7411
37. Haines A, Ebi K. The imperative for climate action to protect health. *N Engl J Med.* (2019) 380:263–73. doi: 10.1056/NEJMra1807873

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.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 25 January 2021
doi: 10.3389/fevo.2020.617841

Effects of Demographic and Weather Parameters on COVID-19 Basic Reproduction Number

Igor Salom[1], *Andjela Rodic*[2], *Ognjen Milicevic*[3], *Dusan Zigic*[1], *Magdalena Djordjevic*[1] *and Marko Djordjevic*[2]*

[1] *Institute of Physics Belgrade, National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia,* [2] *Quantitative Biology Group, Faculty of Biology, University of Belgrade, Belgrade, Serbia,* [3] *Department for Medical Statistics and Informatics, School of Medicine, University of Belgrade, Belgrade, Serbia*

It is hard to overstate the importance of a timely prediction of the COVID-19 pandemic progression. Yet, this is not possible without a comprehensive understanding of environmental factors that may affect the infection transmissibility. Studies addressing parameters that may influence COVID-19 progression relied on either the total numbers of detected cases and similar proxies (which are highly sensitive to the testing capacity, levels of introduced social distancing measures, etc.), and/or a small number of analyzed factors, including analysis of regions that display a narrow range of these parameters. We here apply a novel approach, exploiting widespread growth regimes in COVID-19 detected case counts. By applying nonlinear dynamics methods to the exponential regime, we extract basic reproductive number R_0 (i.e., the measure of COVID-19 inherent biological transmissibility), applying to the completely naïve population in the absence of social distancing, for 118 different countries. We then use bioinformatics methods to systematically collect data on a large number of potentially interesting demographics and weather parameters for these countries (where data was available), and seek their correlations with the rate of COVID-19 spread. While some of the already reported or assumed tendencies (e.g., negative correlation of transmissibility with temperature and humidity, significant correlation with UV, generally positive correlation with pollution levels) are also confirmed by our analysis, we report a number of both novel results and those that help settle existing disputes: the absence of dependence on wind speed and air pressure, negative correlation with precipitation; significant positive correlation with society development level (human development index) irrespective of testing policies, and percent of the urban population, but absence of correlation with population density *per se*. We find a strong positive correlation of transmissibility on alcohol consumption, and the absence of correlation on refugee numbers, contrary to some widespread beliefs. Significant tendencies with health-related factors are reported, including a detailed analysis of the blood type group showing consistent tendencies on Rh factor, and a strong positive correlation of transmissibility with cholesterol levels. Detailed comparisons of obtained results with previous findings, and limitations of our approach, are also provided.

Keywords: COVID-19 transmissibility, environmental factors, basic reproduction number, COVID-19 demographic dependence, COVID-19 weather dependence

OPEN ACCESS

Edited by:
Matteo Convertino,
Hokkaido University, Japan

Reviewed by:
Essam Rashed,
Nagoya Institute of Technology, Japan
Subhas Khajanchi,
Presidency University, India
Kankan Sarkar,
Malda College, India

***Correspondence:**
Marko Djordjevic
dmarko@bio.bg.ac.rs

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *15 October 2020*
Accepted: *28 December 2020*
Published: *25 January 2021*

Citation:
Salom I, Rodic A, Milicevic O, Zigic D, Djordjevic M and Djordjevic M (2021) Effects of Demographic and Weather Parameters on COVID-19 Basic Reproduction Number. Front. Ecol. Evol. 8:617841. doi: 10.3389/fevo.2020.617841

INTRODUCTION

The ancient wisdom teaches us that "knowing your adversary" is essential in every battle—and this equally applies to the current global struggle against the COVID-19 pandemic. Understanding the parameters that influence the course of the pandemic is of paramount importance in the ongoing worldwide attempts to minimize the devastating effects of the virus which, to the present moment, has already taken a toll of more than a million lives (Dong et al., 2020), and resulted in double-digit recession among some of the major world economies (World Bank, 2020a). Of all such factors, the ecological ones (both abiotic such as meteorological factors and biotic such as demographic and health-related population properties) likely play a prominent role in determining the dynamics of disease progression (Qu et al., 2020).

However, making good estimates of the effects that general demographic, health-related, and weather conditions, have on the spread of COVID-19 infection is beset by many difficulties. First of all, these dependencies are subtle and easily overshadowed by larger-scale effects. Furthermore, as the effective rate of disease spread is an interplay of numerous biological, medical, social, and physical factors, a particular challenge is to differentiate the dominating effects of local COVID-19-related policies, which are both highly heterogeneous and time-varying, often in an inconsistent manner. And this is precisely where, in our view, much of the previous research on this subject falls short.

There are not many directly observable variables that can be used to trace the progression of the epidemics on a global scale (i.e., for a large number of diverse countries). The most obvious one—the number of detected cases—is heavily influenced both by the excessiveness of the testing (which, in turn, depends on non-uniform medical guidelines, variable availability of testing kits, etc.) and by the introduced infection suppression measures (where the latter are not only non-homogeneous but are also erratically observed (Cohen and Kupferschmidt, 2020). Nevertheless, the majority of the research aimed to establish connections of the weather and/or demographic parameters with the spread of COVID-19 seeks correlations exactly with the raw number of detected cases (Adhikari and Yin, 2020; Correa-Araneda et al., 2020; Fareed et al., 2020; Gupta et al., 2020; Iqbal et al., 2020; Li et al., 2020; Pourghasemi et al., 2020; Rashed et al., 2020; Singh and Agarwal, 2020). For the aforementioned reasons, the conclusions reached in this way are questionable. Other variables that can be directly measured, such as the number of hospitalized patients or the number of COVID-19 induced deaths (Pranata et al., 2020; Tosepu et al., 2020; Ward, 2020), again depend on many additional parameters that are difficult to take into account: level of medical care and current hospital capacity, advancements, and changing practices in treating COVID-19 patients, the prevalence of risk groups, and even on the diverging definitions of when hospitalization or death should be attributed to the COVID-19 infection. As such, these variables are certainly not suitable as proxies of the SARS-CoV-2 transmissibility *per se*.

On the other hand, as we here empirically find [and as theoretically expected (Anderson and May, 1992; Keeling and Rohani, 2011)] the initial stage of the COVID-19 epidemic (in a given country or area) is marked by a period of a nearly perfect exponential growth for a wide range of countries, which typically lasts for about 2 weeks (based on our analysis of the available data). One can observe widespread dynamical growth patterns for many countries, with a sharp transition between exponential, superlinear (growth faster than linear), and sublinear (growth slower than linear) regimes (see **Figure 1**)—the last two representing a subexponential growth. We here concentrate on the initial exponential growth of the detected-case data (marked in red in **Figure 1**), characterizing the period *before* the control measures took effect, and with a negligible fraction of the population resistant to infection. Note that dates which correspond to the exponential growth regime (included in **Supplementary Table 1**) are different for each country, corresponding to the different start of COVID-19 epidemic in those countries.

We use the exponential growth regime to deduce the basic reproduction number R_0 (Martcheva, 2015), following a simple and robust mathematical (dynamical) model presented here. R_0 is a straightforward and important epidemiological parameter characterizing the inherent biological transmissibility of the virus, in a completely naïve population, and the absence of social distancing measures (Bar-On et al., 2020; Eubank et al., 2020). To emphasize the absence of social distancing in the definition (and inference) of R_0 used here, the term $R_{0,free}$ is also used, — for simplicity, we further denote $R_0 \equiv R_{0,free}$. R_0 is largely independent of the implemented COVID-19 policies and thus truly reflects the characteristics of the disease itself, as it starts to spread unhampered through the given (social and meteorological) settings. Namely, the exponential period ends precisely when the effect of control measures kick in, which happens with a delay of ∼10 days after their introduction

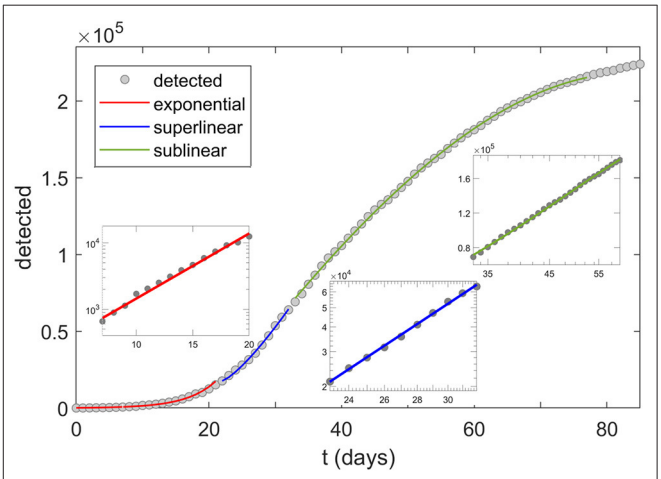


FIGURE 1 | COVID-19 growth regimes. Transitions of the growth patterns (here shown for Italy) from exponential (red), to superlinear (blue) and sublinear (green) regime. The three insets correspond to the log-linear scale (exponential), log-log scale (superlinear), and linear-log scale (sublinear). Dots correspond to detected infections, starting from 20.02.2020. In this study, $R_{0,free}$ is extracted from the slope of the first (exponential, i.e., log-linear) inset, corresponding to dates 29.02–13.03 in the case of Italy.

(The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, 2020), corresponding to the disease latent period, and to the time between the symptom onset and the disease confirmation. Not only that very few governments had enacted any social measures before the occurrence of a substantial number of cases (Cohen and Kupferschmidt, 2020), but also the length of the incubation period makes it likely that the infection had been already circulating for some time through the community even before the first detected case (and that the effects of the measures are inescapably delayed in general). Also, the transition from the exponential to the subsequent subexponential phase of the epidemics is readily visible in the COVID data (see **Figure 1**). Furthermore, R_0 is invariant to the particular testing guidelines, as long as these do not significantly vary over the (here relatively short) studied period. Note that in **Figure 1** cumulative number of positive cases (also known as cumulative infection incidence) is shown, which has to monotonically increase—though with a decreasing rate, once the infection starts to slow down, i.e., once the subexponential growth (sublinear and superlinear regimes) is reached.

In the analysis presented here, we consider 42 different weather, demographic, and health-related population factors, whose analyzed ranges correspond to their variations exhibited in 118 world countries (not all of the parameters were available for all of the countries, as discussed in Section "Demographic and Weather Data Acquisition"). While some authors prefer more coherent data samples to avoid confusing effects of too many different factors (Adhikari and Yin, 2020; Correa-Araneda et al., 2020; Fareed et al., 2020; Rashed et al., 2020; Singh and Agarwal, 2020; Tosepu et al., 2020), this consideration is outweighed by the fact that large ranges of the analyzed parameters serve to amplify the effects we are seeking to recognize and to more reliably determine the underlying correlations. For example, while the value of the Human Development Index (HDI, a composite index of life expectancy, education, and per capita income indicators) varies from 0.36 to 0.96 over the set of analyzed countries, this range would drop by an order of magnitude (Global Data Lab, 2020) if the states of the US were chosen as the scope of the study (other demographic parameters exhibit similar behavior). The input parameters must take values in some substantial ranges to have measurable effects on R_0 (i.e., small variations may lead to effects that are easily lost in statistical fluctuations).

The number of considered parameters is also significant, especially when compared to other similar studies (Adhikari and Yin, 2020; Copat et al., 2020; Fareed et al., 2020; Iqbal et al., 2020; Rashed et al., 2020; Rychter et al., 2020; Singh and Agarwal, 2020; Thangriyal et al., 2020; Tosepu et al., 2020). In a model where a large number of factors are analyzed under the same framework, consistency of the obtained results, in terms of agreement with other studies, common-sense expectations, and their self-consistency, becomes an important check of applied methodology and analysis. Furthermore, a comprehensive and robust analysis is expected to generate new findings and lead to novel hypotheses on how environmental factors influence COVID-19 spread. Overall, we expect that the understanding achieved here will contribute to the ability to understand the behavior of the pandemics in the future and, by the same token, to timely and properly take measures in an attempt to ameliorate the disease effects.

MODEL AND PARAMETER EXTRACTION

Modified SEIR Model and Relevant Approximations

There are various theoretical models and tools used to investigate and predict the progress of an epidemic (Keeling and Rohani, 2011; Martcheva, 2015). We here opted for the SEIR compartmental model, up to now used to predict or explain different features of COVID-19 infection dynamics (Maier and Brockmann, 2020; Maslov and Goldenfeld, 2020; Perkins and España, 2020; Tian et al., 2020; Weitz et al., 2020). The model is sufficiently simple to be applied to a wide range of countries while capturing all the features of COVID-19 progression relevant for extracting the R_0 values. The model assumes dividing the entire population into four (mutually exclusive) compartments with labels: (S)usceptible, (E)xposed, (I)fected, and (R)ecovered.

The dynamics of the model (which considers gradual transitions of the population from one compartment to the other) directly reflects the disease progression. Initially, a healthy individual has no developed SARS-CoV-2 virus immunity and is considered as "susceptible." Through contact with another infected individual, this person may become "exposed"—denoting that the transmission of the virus has occurred, but the newly infected person at this point has neither symptoms nor can yet transmit the disease. An exposed person becomes "infected"—in the sense of becoming contagious—on average after the so-called "latent" period which is, in the case of COVID-19, approximately 3 days. After a certain period of the disease, this person ceases to be contagious and is then considered as "recovered" (from the mathematical perspective of the model, "recovered" are all individuals who are no longer contagious, which therefore also includes deceased persons). In the present model, the recovered individuals are taken to be no longer susceptible to new infections (irrespectively of whether the COVID-19 immunity is permanent or not, it is certainly sufficiently long in the context of our analysis).

Accordingly, almost the entire population initially belongs to the susceptible class. Subsequently, parts of the population become exposed, then infected, and finally recovered. SARS-CoV-2 epidemic is characterized by a large proportion of asymptomatic cases (or cases with very mild symptoms) (Day, 2020), which leads to a large number of cases that remain undiagnosed. For this reason, only a portion of the infected will be identified (diagnosed) in the population, and we classify them as "detected." This number is important since it is the only direct observable in our model, i.e., the only number that can be directly related to the actual COVID-19 data.

This dynamic is schematically represented in **Figure 2**, and is governed by the following set of differential equations:

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \tag{1}$$

$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E \tag{2}$$

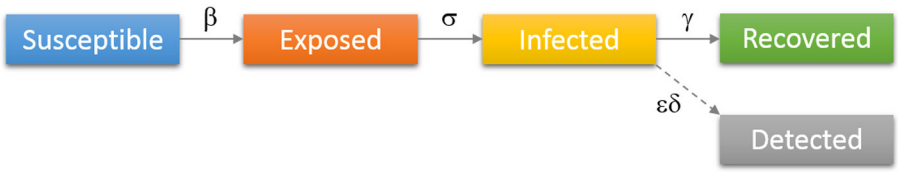


FIGURE 2 | Diagrammatic representation of the SEIR model with the added class of "Detected" patients. Individuals move (denoted by solid arrows) from Susceptible to Exposed to Infected to Recovered, with the rates indicated above arrows in the figure. Some of the infected are detected (diagnosed/confirmed), indicated by the dashed arrow.

$$\frac{dI}{dt} = \sigma E - \gamma I \tag{3}$$

$$\frac{dR}{dt} = \gamma I \tag{4}$$

$$\frac{dD}{dt} = \varepsilon \delta I \tag{5}$$

In the above equations, S, E, I, and R denote numbers of individuals belonging to, respectively, susceptible, exposed, infected, and recovered compartments, D is the cumulative number of detected cases, while N is the total population. Parameter β denotes the transmission rate, which is proportional to the probability of disease transmission in contact between a susceptible and an infectious subject. Incubation rate σ determines the rate at which exposed individuals become infected and corresponds to the inverse of the average incubation period. Recovery rate γ determines the transition rate between infected and recovered parts of the population, (i.e., $1/\gamma$ is the average period during which an individual is infectious). Finally, ε and δ are detection efficiency and the detection rate. All these rate parameters are considered constant during the analyzed (brief) period. Also, note that the constants in our model do not correspond to transition probabilities *per se*, but rather to transition rates (with units 1/time), so that e.g. γ and $\varepsilon\delta$ do not add to one. While rates in the model can be rescaled and normalized to directly correspond to transition probabilities, our formulation (with rates rather than probabilities) is rather common (see e.g., Keeling and Rohani, 2011), and also has a direct intuitive interpretation, where the transition rates correspond to the inverse of the period that individuals spend in a given compartment (see e.g., the explanation for γ above).

In the first stage of the epidemic, when essentially the entire population is susceptible (i.e., $S/N \approx 1$) and no distancing measures are enforced, the average number of secondary infections, caused directly by primary-infected individuals, corresponds to the basic reproduction number R_0. The infectious disease can spread through the population only when $R_0 > 1$ (Khajanchi et al., 2020a), and in these cases, the initial growth of the infected cases is exponential. Though R_0 is a characteristic of the pathogen, it also depends on environmental abiotic (e.g., local weather conditions), as well as biotic factors (e.g., prevalence of health conditions, and population mobility tightly related to the social development level).

Note that, as we seek to extract the basic reproduction number R_0 from the model for a wide range of countries, the social distancing effects are not included in the model presented above. That is, the introduced model serves only to explain the exponential growth phase—note that this growth regime characterizes part of the infection progression where the social distancing interventions still did not take effect, and where the fraction of resistant (non-susceptible) population is still negligible. It is only this phase which is relevant for extracting R_0 that is used in the subsequent analysis. R_0 should not be confused with the effective reproduction number R_e, which takes into account also the effects of social distancing interventions and the decrease in the number of susceptibles due to acquired infection resistance. R_e is not considered in this work, as we are concerned with the factors that affect the inherent biological transmissibility of the virus, independently from the applied measures. That is, by considering R_0 rather than R_e, we disentangle the influence of meteorological and demographic factors on transmissibility (the goal of this study), from the effects of social distancing interventions (not analyzed here). The model can, however, be straightforwardly extended to include social distancing measures, as we did in (Djordjevic et al., 2020)—social distancing measures were also included through other frameworks (Khajanchi and Sarkar, 2020; Maier and Brockmann, 2020; Maslov and Goldenfeld, 2020; Perkins and España, 2020; Samui et al., 2020; Sarkar et al., 2020; Tian et al., 2020; Weitz et al., 2020). Such extensions are needed to explain the subexponential growth that emerges due to intervention measures (i.e., superlinear and sublinear growth regimes that are illustrated in **Figure 1** for Italy but are common for other countries as well).

COVID-19 Growth Regimes

If we observe the number of total COVID-19 cases (e.g., in a given country) as a function of time, there is a regular pattern that we observe: the growth of the detected COVID cases is initially exponential but slows down after some time—when we say it enters the subexponential regime. The subexponential regime can be further divided into the superlinear (growing asymptotically faster than a linear function) and sublinear regime (the growth is asymptotically slower than a linear function). This typical behavior is illustrated, in the case of Italy, in **Figure 1** above. The transition to the subexponential regime occurs relatively soon, much before a significant portion of the population gains immunity, and is a consequence of the introduction of the infection suppression measures.

Inference of the Basic Reproduction Number R_0

In the initial exponential regime, a linear approximation to the model can be applied. Namely, in this stage, almost the entire population is susceptible to the virus, i.e., S/N≈1, which simplifies the Equation (2) to:

$$\frac{dE}{dt} = -\sigma E + \beta I. \tag{6}$$

By combining expressions (3) and (6) one obtains:

$$\frac{d}{dt}\begin{pmatrix} E \\ I \end{pmatrix} = \begin{pmatrix} -\sigma & \beta \\ \sigma & -\gamma \end{pmatrix}\begin{pmatrix} E \\ I \end{pmatrix} = A\begin{pmatrix} E \\ I \end{pmatrix}, \tag{7}$$

where we have introduced a two-by-two matrix:

$$\mathrm{A} = \begin{pmatrix} -\sigma & \beta \\ \sigma & -\gamma \end{pmatrix} \tag{8}$$

The solution for the number of infected individuals can now be written:

$$I(t) = C_1 \cdot e^{\lambda_+ t} + C_2 \cdot e^{\lambda_- t}, \tag{9}$$

where λ_+ and λ_- denote eigenvalues of the matrix A, i.e., the solutions of the equation:

$$\det(A - \lambda I) = 0. \tag{10}$$

The eigenvalues must satisfy:

$$\begin{vmatrix} -\sigma - \lambda & \beta \\ \sigma & -\gamma - \lambda \end{vmatrix} = 0,$$

leading to:

$$(\lambda + \sigma) \cdot (\lambda + \gamma) - \beta \cdot \sigma = 0. \tag{11}$$

The solutions of (11) are:

$$\lambda_\pm = \frac{-(\gamma + \sigma) \pm \sqrt{(\gamma - \sigma)^2 + 4\beta\sigma}}{2}. \tag{12}$$

Since $\lambda_- < 0$, the second term in (9) can be neglected for sufficiently large *t*. More precisely, numerical analysis shows that this approximation is valid already after the second day, while, for the extraction of R_0 value we will anyhow ignore all data before the fifth day (for the analyzed countries, numbers of cases before the fifth day were generally too low, hence this early data is dominated by stochastic effects/fluctuations). Hence, $I(t)$ is proportional to $exp(\lambda_+ t)$, i.e.:

$$I(\mathrm{t}) = I(0) \cdot e^{\lambda_+ t}. \tag{13}$$

By using β from (12) and $R_0 = \frac{\beta}{\gamma}$ (Keeling and Rohani, 2011; Martcheva, 2015), we obtain:

$$R_0 = 1 + \frac{\lambda_+ \cdot (\gamma + \sigma) + {\lambda_+}^2}{\gamma \cdot \sigma}. \tag{14}$$

From (13) and (5) we compute:

$$\mathrm{D}(\mathrm{t}) = \varepsilon \cdot \delta \cdot I(0) \cdot \frac{(e^{\lambda_+ \mathrm{t}} - 1)}{\lambda_+}. \tag{15}$$

By taking the logarithm, the above expression leads to:

$$\log(\mathrm{D}(\mathrm{t})) = \log(\varepsilon\delta I(0)/\lambda_+) + \lambda_+ \cdot \mathrm{t}, \tag{16}$$

from which λ_+ can be obtained as the slope of the log (D (t)) function. From Equation (14), we thus obtain the R_0 value as a function of the slope of log (D (t)), where the latter can be efficiently inferred from the plot of the number of detected COVID-19 cases for a large set of countries.

The SEIR model and the above derivation of R_0 assume that the population belonging to different compartments is uniformly mixed. Possible heterogeneities may tend to increase R_0 values (Keeling and Rohani, 2011). However, this would not influence the results obtained below, as our R_0 values are consistently inferred for all analyzed countries by using the same model, methodology, and parameter set. Moreover, our R_0 values are in agreement with the prevailing estimates in the literature (Najafimehr et al., 2020).

Demographic and Weather Data Acquisition

For the countries for which R_0 was determined through the procedure above, we also collect a broad spectrum of meteorological and demographic parameter values. Overall, 118 countries were selected for our analysis, based on the relevance of the COVID-19 epidemiological data. Namely, a country was considered as relevant for the analysis if the number of detected cases on June 15th was higher than a threshold value of 1,000. A few countries were then discarded from this initial set, where the case count growth was too irregular to extract any results, possibly due to inconsistent or irregular testing policies. As a source for detected cases, we used (World Bank, 2020b; Worldometer, 2020).

In the search for factors correlated with COVID-19 transmissibility, we have analyzed overall 42 parameters, 11 of which are related to weather conditions, 30 to demographics or health-related population characteristics, and one parameter quantifying a delay in the epidemic's onset (data provided in **Supplementary Tables 2–5**). Not all of these parameters were available for all of the considered countries. In particular, data on the prevalence of blood types (**Supplementary Table 4** in the Supplement) was possible to find for 83 of the 118 countries, while, primarily due to scarce data on pollutant concentrations during the epidemics, almost 30% of entries in **Supplementary Table 5** in Supplement had to be left blank for this category. Nevertheless, we opted to include these parameters in our report: despite the lower number of values, some of these parameters exhibited strong and highly statistically significant correlations with R_0, warranting their inclusion.

Our main source of weather data was project POWER (Prediction of Worldwide Energy Resources) of the NASA agency (NASA Langley Research Center, 2020). A dedicated

Python script was written and used to acquire weather data via the provided API (Application Programming Interface). NASA project API allows a large set of weather parameters to be obtained for any given location (specified by latitude and longitude) and given date (these data are provided in the **Supplementary Table 7**). From this source, we gathered data on temperature (estimated at 2 m above ground), specific humidity (estimated at 2 m above ground), wind speed (estimated at 2 m above ground), and precipitation (defined as the total column of precipitable water). Data on air pressure (at ground level) and UV index (international standard measurement of the strength of sunburn-producing ultraviolet radiation) were collected via similar API from World Weather Online source (World Weather Online, 2020), using the same averaging methodology. Since we needed to assign a single value to each country (for each analyzed parameter), the following method was used for averaging meteorological data. In each country, a number of largest cities[1] were selected and weather data was taken for the corresponding locations. These data was then averaged, weighted by the population of each city, followed by averaging over the period used for R_0 estimation (more precisely, to account for the time between disease transmission and the case confirmation, we shifted this period 12 days into the past). The applied averaging method used here can be of limited adequacy in countries spreading over multiple climate zones, but is still expected to provide reasonable single-value estimates of the weather parameters, particularly since the averaging procedure was formulated to reflect the most likely COVID-19 hotspots in a given country.

Demographic data was collected from several sources. Percentage of the urban population, refugees, net migration, social and medical insurance coverage, infant mortality, and disease (CVD, cancer, diabetes, and CRD) risk was taken from the World Bank organization (World Bank, 2020b). The HDI was taken from the Our World in Data source (Our World in Data, 2020), while median age information was obtained from the CIA website. The source of most of the considered medical parameters: cholesterol, raised blood pressure, obesity, inactivity, BSG vaccination as well as data on alcohol consumption and smoking prevalence was World Health Organization (World Health Organization, 2020). Data for blood types were taken from the Wikidata web site. BUCAP parameter, representing population density in the built-up area, was taken from GHS Urban Center Database 2015 (European Commission Global Human Settlement, 2020). The onset parameter, determining the delay (in days) of the epidemic's start, was inferred from COVID-19 counts data. We used the most recent available data for all the parameters.

RESULTS

The $\log(D(t))$ function, for a subset of selected countries, is shown in **Figure 3**. The obvious linear dependence confirms that the progression of the epidemic in this stage is almost perfectly exponential. Note that our model exactly reproduces this early exponential growth (see Equation 13), happening under the assumption of a small fraction of the population being resistant, and the absence of the effect of social distancing interventions. From **Figure 3**, we see that this behavior, predicted by the model for the early stage of the epidemic, is also directly supported by the data, i.e., the exponential growth in the cumulative number of confirmed cases is indeed observed for a wide range of countries. For each country, the parameter λ_+ is directly obtained as the slope of the corresponding linear fit of the $\log(D(t))$, and the basic reproduction number R_0 is then calculated from Equation (14). Here, we used the following values for the incubation rate, $\sigma = 1/3\ \text{day}^{-1}$, and for the recovery rate $\gamma = 1/4\ \text{day}^{-1}$, per the commonly accepted values in the literature (Bar-On et al., 2020). Note that possible variations in these two values would not significantly affect any conclusions about R_0 correlations, due to the mathematical properties of the relation (14): it is a strictly monotonous function of λ_+ and the linear term $\lambda_+ \cdot (\gamma + \sigma)/\gamma \cdot \sigma$ dominantly determines the value of R_0.

Supplementary tables contain the values for 42 variables, for all countries. Correlations of each of the variables with R_0 are given in **Supplementary Table 6**. Values for the Pearson correlation coefficient are further shown below, though consistent conclusions are also obtained by Kendall and Spearman correlation coefficients (which do not assume a linear relationship between variables). Correlation coefficients were calculated in the usual manner: as the correlation of the vector of parameter values with the vector of R_0 values, by taking into account all available data (for parameters that were available across all of the countries, both of the vectors were 118 dimensional; if values were missing for certain countries, these countries were simply ignored and lower-dimensional vectors were compared).

The first set of results that corresponds to, roughly speaking, general demographic data, is presented in **Figure 4**. The plot in panel A shows the distribution of R_0 vs. HDI values for all countries, where a higher HDI score indicates the more prosperous country concerning life expectancy, education, and per capita income (Sagar and Najam, 1998). This parameter was included in the study due to a reasonable expectation that a higher level of social development also implies a higher level of population interconnectedness and mixing (stronger business and social activity, more travelers, more frequent contacts, etc.), and hence that HDI could be related to the SARS-CoV-2 transmissibility. Indeed, we note a strong, statistically highly significant correlation between the HDI and the R_0 value, with $R = 0.37$, and $p = 4 \cdot 10^{-5}$, demonstrating that the initial expansion of COVID-19 was faster in more developed societies.

The social security and health insurance coverage (INS) "shows the percentage of population participating in programs that provide old age contributory pensions (including survivors and disability) and social security and health insurance benefits (including occupational injury benefits, paid sick leave, maternity, and other social insurance)" (World Bank, 2020b). Reflecting the percentage of the population covered by medical insurance and likely feeling more protected from the financial

[1]This number was determined for each country by the following condition: the total population of the cities taken into consideration had to surpass 10 percent of the overall population of the country.

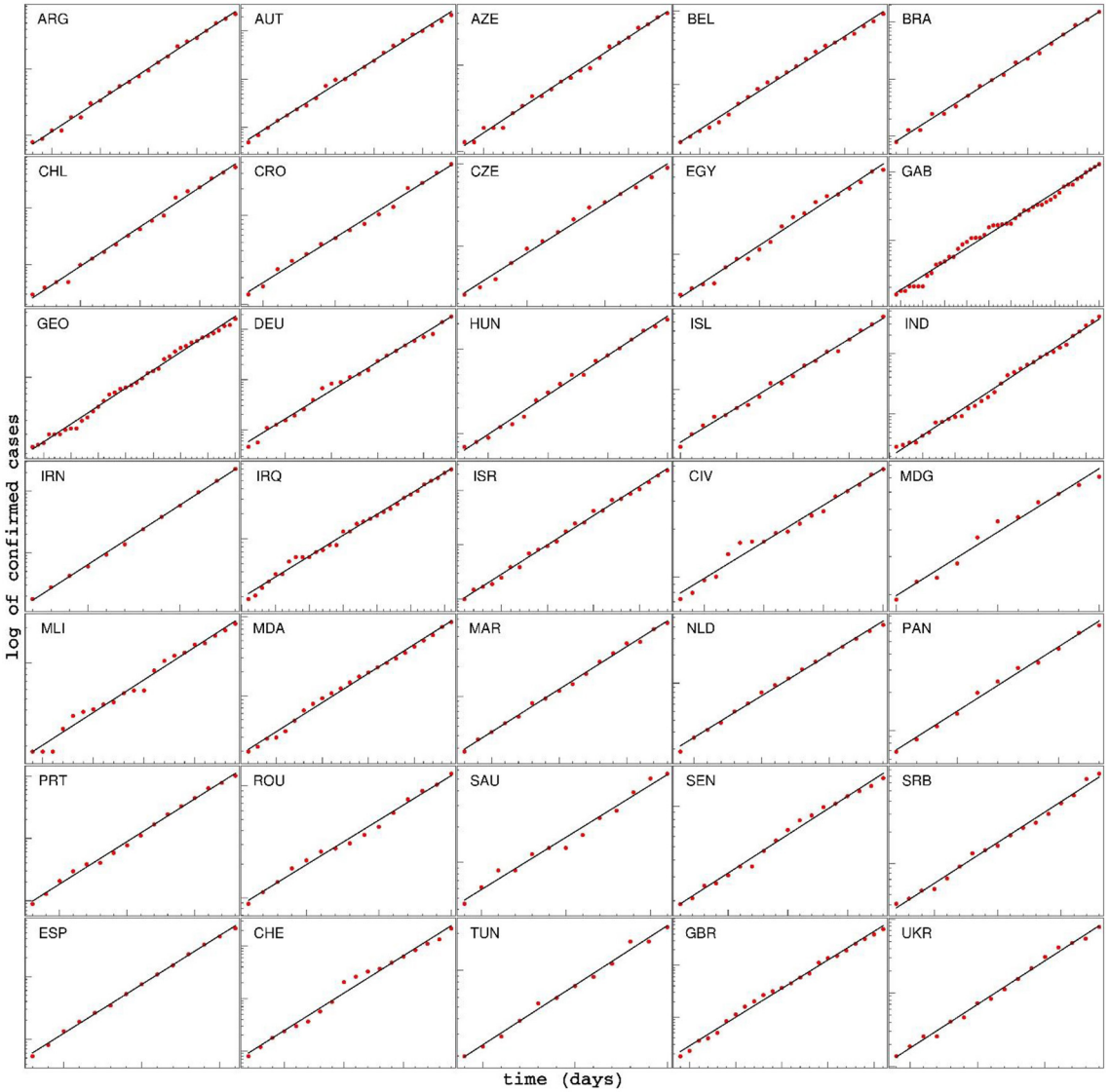


FIGURE 3 | Time dependence of the detected cases for various countries, during the initial period of the epidemic, shown on a log-linear scale. The linear fit of log(D) shows that the spread of COVID-19 in this phase is very well approximated by exponential growth. Note that the values on axes are chosen differently for each country, in order to emphasize the exponential character of the growth. For each country, the start and end dates of the exponential regime, together with the extracted slope λ_+, are provided in the **Supplementary Table 1**. ARG, Argentine; AUT, Austria; AZE, Azerbaijan; BEL, Belgium; BRA, Brazil; CHL, Chile; CRO, Croatia; CZE, Czech Republic; EGY,Egypt; GAB, Gabon; GEO, Georgia; DEU, Germany; HUN, Hungary; ISL, Iceland; IND, India; IRN, Iran; IRQ, Iraq; ISR, Israel; CIV, Cote d'Ivoire; MDG, Madagascar; MLI, Mali; MDA, Moldova; MAR, Morocco; NLD, Netherlands; PAN, Panama; PRT, Portugal; ROU, Romania; SAU, Saudi Arabia; SEN, Senegal; SRB, Serbia; ESP, Spain; CHE, Switzerland; TUN, Tunis; GBR, Great Britain; UKR, Ukraine.

effects of the epidemics, this indicator shows a strong ($R = 0.4$) and highly significant ($p = 4{\cdot}10^{-4}$) positive correlation with R_0. The percentage of urban population (UP) and BUCAP density (BAP) are both included as measures of how concentrated is the population of the country. While the UP value simply shows what percentage of the population lives in cities, the BUCAP parameter denotes the amount of built-up area per person. Of the two, the former shows a highly significant positive correlation with the COVID-19 basic reproduction number, whereas the latter shows no correlation. Median age (MA) should be of obvious potential relevance in COVID-19 studies since it is well known that the disease more severely affects the older population (Jordan et al., 2020). Thus, we wanted to investigate also if there is any connection of age with the virus transmissibility. Our results are suggestive of such a connection, since we obtained a strong positive correlation of age with R_0, with very high statistical confidence. Infant mortality (IM) is defined as the number of infants dying before reaching 1 year of age, per 1,000 live births. Lower IM rates can serve as another indicator of the prosperity of a society, and it turns out that this measure is also strongly correlated, but negatively, with $R = -0.36$ and $p = 8{\cdot}10^{-5}$ (showing again that more developed countries, i.e.,

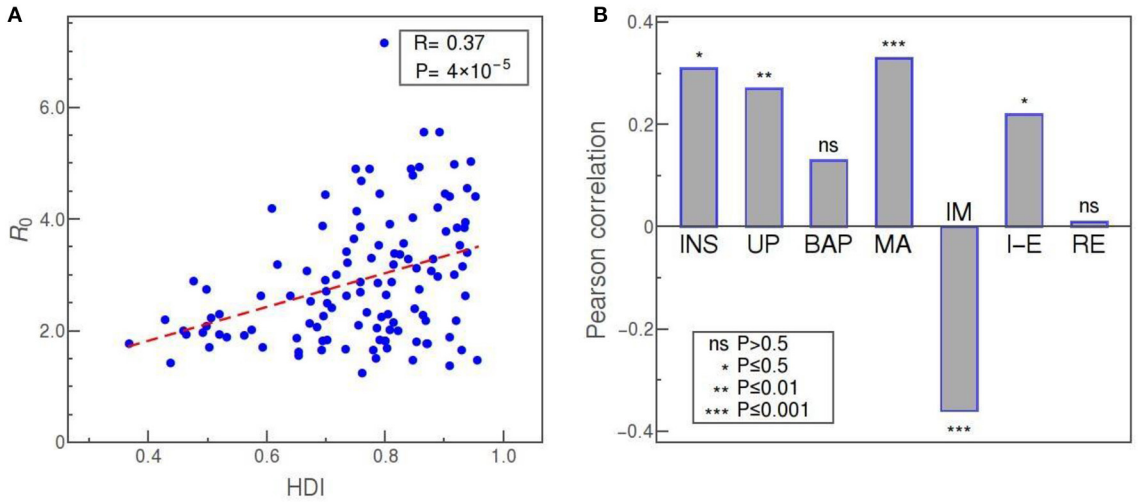


FIGURE 4 | (A) R_0 vs. HDI as an example of transmissibility dependence on demographic data (Pearson correlation denoted as R). **(B)** Pearson correlations of R_0 with (from left to right): social security and health insurance coverage (INS), percentage of urban population (UP), BUCAP measure of population density (BAP), median age (MA), infant mortality (IM), net migration (I-E), and percentage of refugee population by country or territory of asylum (RE). The statistical significance of each correlation is indicated in the legend, while "ns" stands for "no significance."

those with lower IM rates, have experienced more rapid spread of the virus infection). Net migration (I-E) represents the 5-year estimates of the total number of immigrants less the annual number of emigrants, including both citizens and non-citizens. This number, related to the net influx of foreigners, turns out to be positively correlated, in a statistically significant way, with R_0. However, according to our data, the percentage of refugees, defined as the percentage of the people in the country who are legally recognized as refugees and were granted asylum in that country, is not correlated with R_0 at all.

Another set of parameters corresponds to medically-related demographic parameters and is shown in the upper part of **Figure 5**. The plot in panel A represents the average blood cholesterol level (in mmol/L) in the population of various countries, plotted against the value of R_0. The two parameters are strongly correlated, with $R = 0.4$, and $p = 6{\cdot}10^{-6}$. Another demographic parameter with clear medical relevance, that has a comparatively strong and significant positive correlation with R_0, is the alcohol consumption per capita (ALC), as shown in panel B of **Figure 5**. Our data shows that R_0 is also positively correlated, with high statistical significance, with the prevalence of obesity and to a somewhat smaller extent with the percentage of smokers. Here, obesity is defined as having a body-mass index over 30. A medical parameter that is strongly, but negatively, correlated with R_0, is a measure of prevalence and severity of COVID-19 relevant chronic diseases in the population (CD). This parameter is defined as "the percent of 30-year-old-people who would die before their 70th birthday from any of cardiovascular disease, cancer, diabetes, or chronic respiratory disease, assuming that s/he would experience current mortality rates at every age and s/he would not die from any other cause of death" (World Bank, 2020b). The percentage of people with raised blood pressure (RBP) is also negatively correlated with R_0, though this correlation is not as strong and as statistically significant as in the case of the CD parameter. Here, raised blood pressure is defined as systolic blood pressure over 140 or diastolic blood pressure over 90, in the population older than 18. The percentage of smokers exhibits statistically significant (though not large) positive correlation. Two medical-demographic parameters that show no correlation with R_0 in our data are the prevalence of insufficient physical activity among adults aged over 18 (IN) and BCG immunization coverage among 1-year-olds (BCG).

In **Figure 5C** we see that blood types are, in general, strongly correlated with R_0. The highest positive correlation is exhibited by A^- and O^- types, with a Pearson correlation of 0.4 and 0.39, and a very high statistical significance of $p = 10^{-4}$ and $p = 2{\cdot}10^{-4}$, respectively. Taken as a whole, group A is still strongly and positively correlated with R_0, albeit with a bit lower statistical significance (A^+ type correlation has p-value two orders of magnitude higher than A^-). This is not so for group O that, overall, does not seem to be correlated to R_0 (O^+ even shows a certain negative correlation but without statistical significance). Our data reveals a highly significant positive correlation also for AB^- subtype ($R = 0.31$, $p = 0.003$), while neither the AB^+ subtype nor overall AB group is significantly correlated with the basic reproduction number. Clear negative correlation is exhibited only by B blood group ($R = -0.31$, $p = 0.004$), mostly due to the negative correlation of its B^+ subtype ($R = -0.34$, $p = 0.001$), whereas B^- subtype is not significantly correlated with R_0 in our data. If we consider the rhesus factor alone, we again observe very strong correlations with R_0 and with very high statistical significance: Rh^- and Rh^+ correlate positively ($R = 0.4$) and negatively ($R = -0.4$), respectively, with very high statistical significance ($p = 2{\cdot}10^{-4}$). The tendency of Rh^- and Rh^+ to, respectively, increase and decrease the transmissibility,

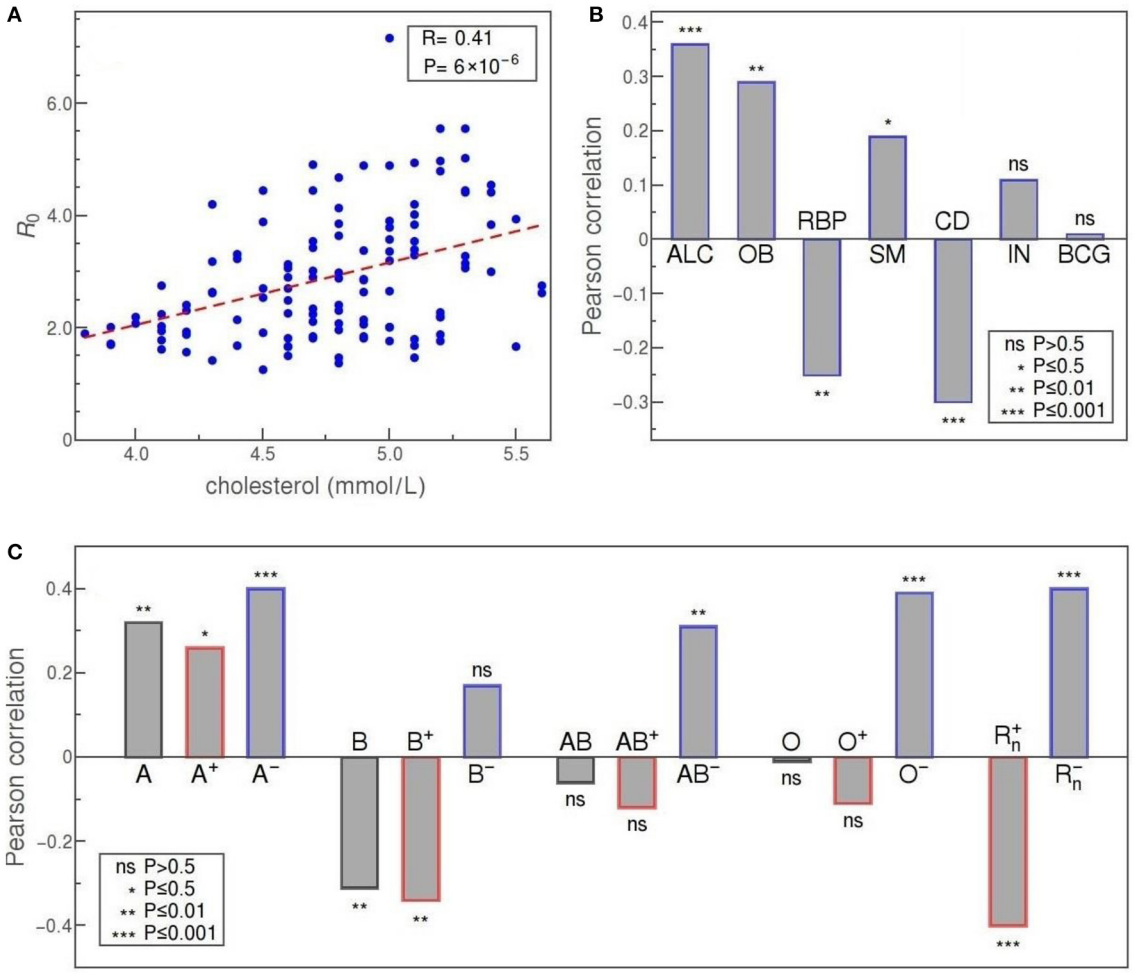


FIGURE 5 | (A) R_0 vs cholesterol level, as an example of a health-related parameter dependence. **(B)** Pearson correlation of R_0 with (from left to right): alcohol consumption per capita (ALC); the prevalence of obesity (OB); severity of COVID-19 relevant chronic diseases in the population (CD); a percentage of people with raised blood pressure (RBP); a percentage of smokers (SM); the prevalence of insufficient physical activity among adults (IN); BCG immunization coverage among 1-year-olds (BCG) **(C)** Correlation of blood types with R_0 in order: A, B, AB, and O (from left to right); overall value for that group, correlation only for Rh^+ subtype of the group, and correlation for Rh^- subtype is shown. The two rightmost bars correspond to the overall correlation of Rh^+ and the overall correlation of the Rh^- blood type with R_0. The convention for representing the statistical significance of each correlation is the same as in **Figure 4**.

is therefore consistent with the results obtained for all four individual blood-groups.

In **Figure 6**, the onset represents the delay of the exponential phase and is defined, for each country, as the number of days from February 15 to the start of the exponential growth of detected cases. The motivation was to check for a possible correlation between the delay in the onset of the epidemic and the rate at which it spreads. Indeed, our data shows that such correlation exists and that it is strong and statistically significant: $R = -0.48$ and $p = 4{\cdot}10^{-8}$. In other words, the later the epidemic started, the lower (on average) is the basic reproduction number.

Panel B of **Figure 6** shows the correlation of R_0 with some of the commonly considered air pollutants. Our data reveal a statistically significant positive correlation of R_0 with NO_2 and SO_2 concentrations. Other pollutants—CO, PM2.5 (fine inhalable particles, with diameters that are generally 2.5 μm and smaller), and PM10 (inhalable particles, up to 10 nm in diameter)—show no statistically significant correlation with R_0.

Next, we consider weather factors. Panels C and D of **Figure 6** show correlations of precipitation, temperature, specific humidity, UV index, air pressure, and wind speed with the reproduction number R_0. Of these, precipitation, temperature, specific humidity, and UV index show a strong negative correlation, at a high level of statistical significance. Of the other two parameters, both air pressure and wind speed are not correlated at all with R_0 in our data.

DISCUSSION

The present paper aimed to establish relations between the COVID-19 transmissibility and a large number of

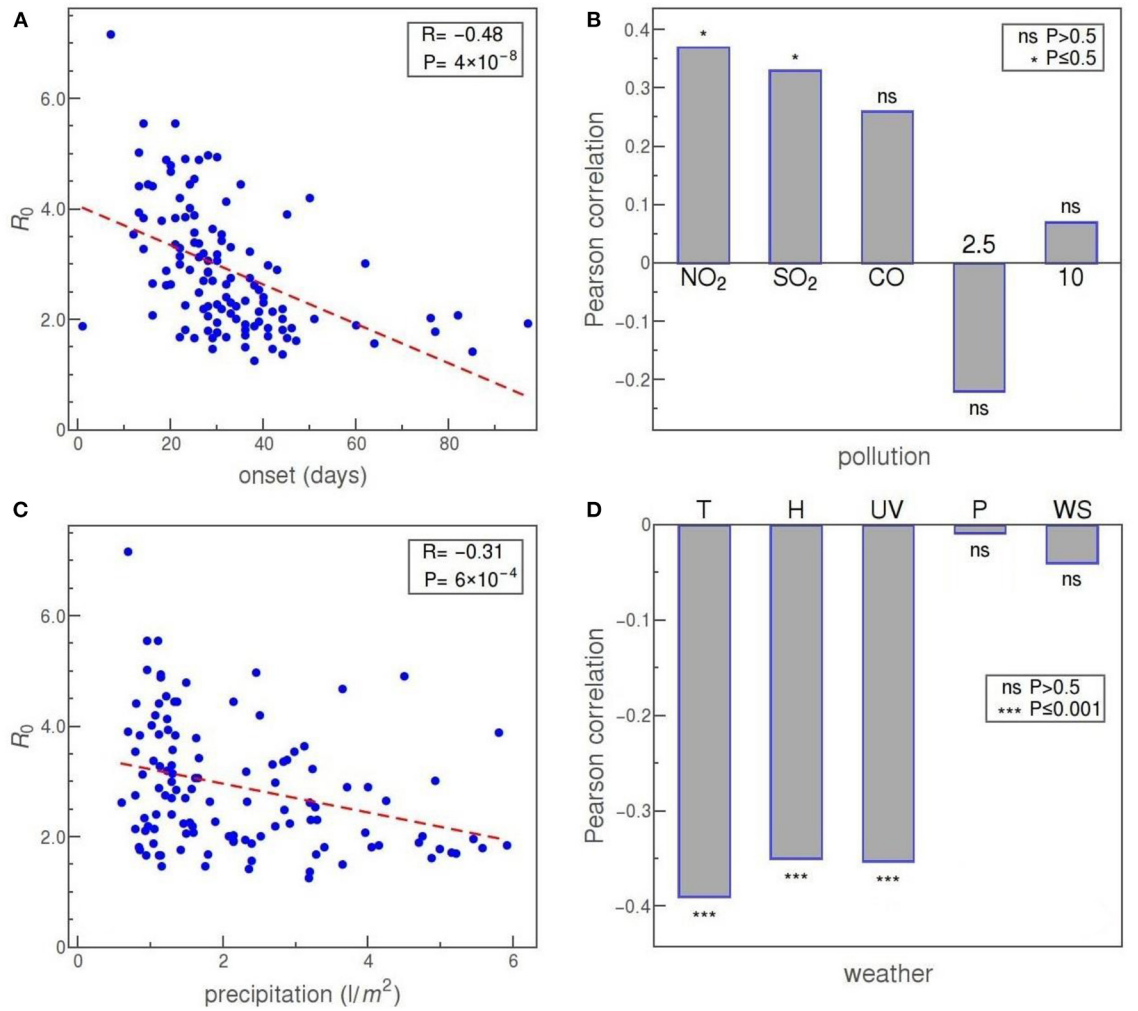


FIGURE 6 | (A) R_0 vs. the time delay of epidemic onset. **(B)** Pearson correlation of R_0 with pollutants (from left to right): NO_2, SO_2, CO, PM2.5, and PM10 (inhalable particles with 2.5 and 10 μm, respectively). **(C)** R_0 vs. precipitation. **(D)** Pearson correlation of R_0 with (from left to right): temperature, specific humidity, UV index, air pressure, and wind speed. The convention for representing the statistical significance of each correlation is the same as before.

demographic and weather parameters. As a measure of COVID-19 transmissibility, we have chosen the basic reproduction number R_0—a quantity that is essentially independent of the variations in both the testing policies and the introduced social measures (as discussed in the Introduction), in distinction to many studies on transmissibility that relied on the total number of detected case counts [see e.g., (Adhikari and Yin, 2020; Correa-Araneda et al., 2020; Fareed et al., 2020; Gupta et al., 2020; Iqbal et al., 2020; Li et al., 2020; Pourghasemi et al., 2020; Rashed et al., 2020; Singh and Agarwal, 2020)]. We have covered a substantial number of demographic and weather parameters, and included in our analysis all world countries that were significantly affected by the COVID-19 pandemic (and had a reasonable consistency in tracking the early phase of infection progression). While a number of manuscripts have been devoted to factors that may influence COVID-19 progression, only a few used an estimate of R_0 or some of its proxies (Coccia, 2020; Contini and Costabile, 2020; Copiello and Grillenzoni, 2020)—these studies were however limited to China, and included a small set of meteorological variables, with conflicting results obtained for their influence on R_0. Therefore, a combination of (i) using a reliable and robust measure of COVID-19 transmissibility, and (ii) considering a large number of factors that may influence this transmissibility within the same study/framework, distinguishes our study over prior work. We, however, must be cautious when it comes to further interpretation of the obtained data. As always, we must keep in mind that "correlation does not imply causation" and that further research is necessary to identify possible confounding factors and establish which of these parameters truly affect the COVID-19 transmissibility. Due to the sheer number of studied variables, an even larger number of parameters that might be relevant but are inaccessible to study (or even impossible to quantify), as well as due to possible intricate mutual relations of

the factors that may influence COVID-19 transmission, this is a highly non-trivial task. While we postpone any further analysis in this direction to future studies, we will, nevertheless, consider here the possible interpretations of the obtained correlations, assuming that they also probably indicate the existence of at least some causation. We below provide a detailed comparison between our results and previous findings. While a detailed discussion is presented, despite our best effort, we may have missed some of the relevant references due to an extremely rapidly developing field. Nevertheless, we point out a clear distinction of our work with previous studies, as outlined in this paragraph.

We will first consider the demographic variables presented in **Figure 4**. The obtained correlation of the HDI with the basic reproduction number is both strong and hardly surprising. The level of prosperity and overall development of a society is necessarily tied with the degree of population mobility and mixing, traffic intensity (in particular air traffic), business and social activity, higher local concentrations of people, and other factors that directly or indirectly increase the frequency and range of personal contacts (Gangemi et al., 2020), rendering the entire society more vulnerable to the spread of viruses. In this light, it is reasonably safe to assume that the obtained strong and highly statistically significant correlation of HDI with R_0 reflects a truly causal connection. However, some authors offer also a different explanation: that higher virus transmission in more developed countries is a consequence of more efficient detection of COVID-19 cases due to the better-organized health system (Gangemi et al., 2020)—but since our R_0 measure does not depend on detection efficiency, presented results can be taken as evidence against such hypothesis.

The interpretation is less clear for other demographic parameters, for example, the percentage of the population covered by medical and social insurance programs (INS). While there seem to be no previous studies discussing this parameter, one possibility is to attribute its strong positive correlation with R_0 to a hypothetical tendency of population to more easily indulge in the epidemiologically-risky behavior if they feel well-protected, both medically and financially, from the risks posed by the virus; conversely, that the population that cannot rely on professional medical care in the case of illness is likely to be more cautious not to contract the virus. The other is, of course, to see this correlation as an indirect consequence of the strong correlation of this parameter with HDI—which is also, almost certainly, the underlying explanation of the infant mortality (IM) correlation, where low mortality ratios point to a better medical system, which goes hand in hand with the overall prosperity and development of the country (Ruiz et al., 2015) (thus the negative correlation with R_0).

Similarly, the strong positive correlation of median age (MA) with R_0 might be a mere consequence of its clear relation with the overall level of development of the country (Gangemi et al., 2020), but it can be also considered in the light of the fact that clinical and epidemiological studies have unanimously shown that the elderly are at higher risk of developing a more severe clinical picture, and our result may indicate that the virus also spreads more efficiently in the elderly population. Possible explanations may include: drugs frequently prescribed to this population that increase levels of ACE2 receptors (Shahid et al., 2020), a general weakening of the immune system with age leading to a greater susceptibility to viral infections (Pawelec and Larbi, 2008), and a large number of elderly people grouped in nursing homes, where the virus can expand very quickly (Kimball et al., 2020).

The correlation of population density with R_0, or the lack of thereof, is more challenging to explain. Naively, one could expect that COVID-19 spreads much more rapidly in areas with a large concentration of people, but, if exists, this effect is not that easily numerically captured. As the standard population density did not show any correlation with the reproduction number R_0 (not shown), we explored some more subtle variants. Namely, the simplest reason why the data shows no correlation of R_0 with population density would be that the density, calculated in the usual way, is too averaged out: the most densely-populated country on our list, Monaco, has roughly 10,000 times more people per square kilometer than the least densely-populated Australia. However, Melbourne downtown has a similar population density as Monaco and far more people, so one would expect no a priori reason that its infection progression would be slower (and the R_0 rate for Australia as a whole will be dominantly determined by the fastest exponential expansion occurring anywhere on its territory). For this reason, we included the BUCAP parameter into the analysis, which takes into account only population density in built-up areas. Surprisingly, even this parameter did not exhibit any statistically significant correlation. Actually, several studies may serve as examples showing that the correlation of population density with the rate of COVID-19 expansion can be expected only under certain conditions since the frequency of contacts between people is to a large extent modulated by additional geographical, economic, and sociological factors (Berg et al., 2020; Carozzi, 2020; Pourghasemi et al., 2020; Rashed et al., 2020). Our observed absence of a correlation could be therefore expected and possibly indicates that such a correlation should be sought at the level of smaller populated areas—for example, individual cities (Yu et al., 2020). This conclusion is somewhat supported by the obtained highly significant and strong positive correlation of R_0 with the percentage of the population living in cities (UP) and which probably reflects the higher number of encounters between people in a more densely populated, urban environment (Li et al., 2020). It is also possible that virus spread might have a highly non-linear dependence on the population density—namely, that an outbreak in a susceptible population requires a certain threshold value of its density, while below that value population density ceases to be a significant factor influencing virus (Scheffer, 2009; Carozzi, 2020; Coro, 2020).

Another demographic parameter that exhibits a significant correlation with R_0 in our data is the net migration (I-E), denoting the number of immigrants less the number of emigrants. Unlike this number, which shows a positive correlation, the number of refugees (RE) seems not to be correlated at all. By definition, migrants deliberately choose to move to improve their prospects, while refugees have to move to save their lives or preserve their freedom. Migrants (e.g., in

economic or academic migration), arguably tend to stay in closer contact with the country of their origin and have more financial means for that, which likely contributes to more frequent border crossings and more intensive passenger traffic (Fan et al., 2020), thereby promoting the infection spread. On the other hand, refugees are mostly stationed in refugee camps, there is less possibility of spreading the virus outside through contacts with residents, but there is a high possibility of escalation of the epidemic within camps with a high concentration of people (Hargreaves et al., 2020). We did not find any other attempt in the literature to examine this issue. In any case, our results demonstrate that refugees are certainly not a primary cause of concern in the pandemics, contrary to fears expressed in some media.

Of the medical factors, the strongest correlation of R_0 is established with elevated cholesterol levels, as shown in **Figure 5**. Cholesterol may be associated with a viral infection and further disease development through a complex network of direct and indirect effects. *In vitro* studies of the role of cholesterol in virus penetration into the host body, done on several coronaviruses, indicate that its presence in the lipid rafts of the cell membrane is essential for the interaction of the virus with the ACE2 receptor, and also for the latter endocytosis of the virus (Radenkovic et al., 2020). Obesity prevalence (OB) also exhibits a highly significant, though somewhat weaker correlation with R_0, which might be a consequence of the common connection between obesity and cholesterol: in principle, obesity might be a relevant factor in the COVID-19 epidemic exactly due to the effects of cholesterol on SARS-CoV-2 susceptibility. Of course, other effects might be at play, e.g., the fact that the adipose tissue of obese people excessively produces pro-inflammatory cytokines (Sattar et al., 2020). In the case of obesity, a simple explanation via relation to HDI is not available, since obesity does not show a simple correlation with the society development (Haidar and Cosman, 2011). Overall, while the correlation of obesity with a more severe prognosis in COVID-19 is well established in the literature, its relation to COVID-19 transmissibility is only mentioned in Li et al. (2020) and hitherto unexplained.

Often related to obesity is also raised blood pressure (RBP), and we have discovered that this factor is also correlated, at high statistical significance, with R_0. While this seems to be the first study correlating high blood pressure with the SARS-CoV-2 transmission rate, it is known that, based on clinical studies, RBP appears to be a risk factor for hospitalization and death due to COVID-19 (Ran et al., 2020a; Schiffrin et al., 2020). In this light, it might be surprising that the correlation between RBP and R_0 turns out to be negative. On the other hand, this result supports the existing hypothesis about the beneficial effect of ACE inhibitors and ARBs (Ran et al., 2020a; Schiffrin et al., 2020) (standardly used in the treatment of hypertension). Similarly unintuitive correlation we report in the case of chronic diseases that are known to be relevant for the COVID-19 outcome. Namely, our data show, at very high statistical significance, a strong negative correlation of R_0 with the risk of death from a batch of chronic diseases (cardiovascular disease, cancer, diabetes, and chronic respiratory disease), agreeing in this regard with some recent research (Chiang et al., 2020; Li et al., 2020). These diseases are identified as relevant comorbidities in the context of COVID-19, leading to a huge increase in the severity of the infection and poorer prognosis (An et al., 2020; Zheng et al., 2020) and, therefore, the discovered negative correlation comes as a surprise—particularly when contrasted to the positive correlation of obesity (where both are recognized risk factors in COVID-19 illness). One possible explanation is that the correlation may be due to potentially lower mobility of people with chronic diseases compared to the general mobility of the population. Additionally, it is possible that these people, being aware to belong to a high-risk group, behaved more cautiously even before the official introduction of social distancing measures.

According to our analysis, the prevalence of certain health-hazard habits is also significantly correlated to COVID-19 transmissibility. Chronic excessive alcohol consumption has, in general, a detrimental effect on immunity to viral and bacterial infections, which, judging by the strong positive correlation we obtained, most likely applies also to SARS-CoV-2 virus infection. This correlation contradicts the belief that alcohol can be used as a protective nostrum against COVID-19, which has spread in some countries and even led to cases of alcohol poisoning (Chick, 2020).

Regarding the impact of smoking on SARS-CoV-2 virus infection—the results are controversial (Chatkin and Godoy, 2020). The positive correlation of smoking with COVID-19 transmissibility that we obtained seems to support the reasoning that, since the SARS-CoV-2 virus enters cells by binding to angiotensin-converting enzyme 2 (ACE2) receptors and that the number of these receptors is significantly higher in the lungs of smokers, the smokers will be more affected and easily infected (Brake et al., 2020; Hoffmann et al., 2020). Accordingly, our result contradicts the hypothesis that a weakened immune response of smokers to virus infection may prove beneficial in the context of inflammation caused by intense cytokine release (Garufi et al., 2020).

Another result that addresses the association of unhealthy lifestyle with greater susceptibility to SARS-CoV-2 infection is the slight positive correlation we obtained for the prevalence of insufficient physical activity (IN) in adults, which is however not statistically significant. In this sense, in the case of COVID-19, we could not fully confirm the findings from (Jurak et al., 2020), who found that physical activity significantly reduces the risk of viral infections.

Despite the recent media interest (Gallagher, 2020), our findings neither could confirm that BCG immunization has any beneficial effect in the case of COVID-19, at least as far as reducing the risk of contracting and transmitting the disease is concerned. While it is known that the BCG vaccine provides some protection against various infectious agents, unfortunately, there is no clear evidence for such an effect against SARS-CoV-2 (O'Neill and Netea, 2020). Our analysis suggests that BCG immunization simply does not correlate with SARS-CoV-2 virus transmission.

SARS-CoV-2 target cells are typically capable of synthesizing ABH antigens and certain arguments exist, both theoretical and experimental, for a potential relation of blood groups

with COVID-19 progression and transmission (Guillon et al., 2008; Dai, 2020; Gérard et al., 2020). While the results of epidemiological studies on COVID-19 patients mostly support the proposed effect of blood groups on the development of COVID-19 disease, the relationship between virus transmission and blood group prevalence and Rh phenotype has been significantly less studied. Our analysis showed strong positive correlations of virus transmission with the presence of A blood group and Rh^- phenotype, as well as strong negative correlations for B blood group and Rh^+ phenotype, while for AB and O blood group no significant correlations were obtained (**Figure 5C**). This result coincides significantly with the correlations obtained in a study conducted for 86 countries (Ansari-Lari and Saadat, 2020). However, another study focused on hospitalized patients in Turkey reported that the Rh^+ phenotype represents a predisposition to infection (Arac et al., 2020), contradicting our findings. Similar results regarding the Rh factor were obtained in a study (Latz et al., 2020) on hospitalized patients in the US (this study further reported no correlation of blood types with the severity of the disease). One way to reconcile these results with ours would be to speculate that the virus is more efficiently transmitted in a population with a higher proportion of Rh^- phenotype because these people show a milder clinical picture compared to Rh^+, so their movement is not equally limited, which is why they have more ability to pass on the infection.

Our data (**Figure 6A**) shows a strong negative correlation with the date of the epidemic onset. Curiously, it seems that the later the epidemic started in a given country, it is more likely that the disease expansion will be slower. Instead of interpreting this result as an indication that the virus has mutated and changed its properties over such a short period, we offer the following simpler explanation: pandemic reached first those countries that are most interconnected with the rest of the world (at the same time, those are the countries characterized by great mobility of people overall), so it is expected that also the progression of the local epidemics in these countries is more rapid. Another contributing factor could be the effect of media, which had more time to raise awareness about the risks of COVID-19 in the countries that were hit later (Khajanchi et al., 2020b).

Another segment of our interest were air pollutants, shown in **Figure 6B**. Air pollution can have a detrimental effect on the human immune system and lead to the development (or to worsening) of respiratory diseases, including those caused by respiratory viral infections (Becker and Soukup, 1999; Copat et al., 2020). Several papers have already investigated air pollution in the context of COVID-19 and reported a positive correlation between the death rate due to COVID-19 and the concentration of PM2.5 in the environment (Wu et al., 2020; Yao et al., 2020c). Positive correlations were also found for the spread of the SARS-CoV-2 virus, but mainly by considering daily numbers of newly discovered cases—a method that, as we have already argued, may strongly depend on testing policies, as well as on state measures to combat the epidemic (Copat et al., 2020). It has been suggested that virus RNA can be adsorbed to airborne particles facilitating thus its spread over greater distances (Coccia, 2020; Setti et al., 2020), but these arguments were contested by examination of air samples in Wuhan (Contini and Costabile, 2020; Liu et al., 2020). The latter conclusions concur with the results of a study in which no correlation was obtained between the basic reproductive number of SARS-CoV-2 infection for 154 Chinese cities and the concentration of PM2.5 and PM10 particles, while the correlation of these factors with the death rate (CFR) was shown (Ran et al., 2020b). The statistically insignificant and relatively weak correlations we obtained for PM2.5 and PM10 pollutants also do not support the hypothesis of a potentially significant role of these particles in the transmission of this virus. In contrast, significant positive correlations were shown by our analysis for concentrations of NO_2, SO_2, and CO in the air (although the correlation for CO is not statistically significant), which is generally supported by the results of other studies. For example, a positive correlation of NO_2 levels with the basic reproductive number of infection was obtained from data for 63 Chinese cities (Yao et al., 2020a). Also, it has been shown that the number of detected cases of COVID-19 in China is strongly positively correlated with the level of CO, while in Italy and the USA such correlation exists with NO_2 (Pansini and Fornacca, 2020). The mentioned study failed to establish a clear correlation with the level of SO_2. Possible mechanisms of interaction were also proposed (Daraei et al., 2020). Also, it is important to emphasize that the atmospheric concentration of NO_2 strongly depends on the levels of local exhaust emissions, so its correlation with virus transmission can be interpreted by the connection with the urban environment, characterized by more intensive traffic (Goldberg et al., 2020).

Finally, we have also obtained some interesting correlations of the meteorological parameters with R_0, shown in **Figures 6C,D**. The statistically very highly significant negative correlation of the basic reproductive number of SARS-CoV-2 virus infection with both the mean temperature and humidity obtained in our research (**Figure 6D**) is consistent with the results of other relevant papers, e.g., (Mecenas et al., 2020). For example, a similar correlation was obtained in a study that analyzed COVID-19 outbreak in the cities of Chile—a country that covers several climate zones, but where it is still safe to assume that social patterns of behavior and introduced epidemic control measures do not drastically differ throughout the country (Correa-Araneda et al., 2020). Effectively the same conclusion—that fewer COVID-19 cases were reported in countries with higher temperatures and humidities—was reached in a study covering over 200 countries in the world (Iqbal et al., 2020). While an established correlation between virus transmission and a certain factor is not, in general, a telltale sign of a direct causal relationship between them, in the case of temperature and humidity such connection is firmly indicated also by results of experimental research (Lowen et al., 2007; Casanova et al., 2010; Chan et al., 2011; van Doremalen et al., 2020). Nevertheless, some studies yielded different conclusions, most likely due to the method of calculating R_0 or due to choosing a small/uninformative sample of populations in which the number of infected cases was monitored (Guo et al., 2020; Lin et al., 2020; Yao et al., 2020b). For example, a study focused on the suburbs of New York, Queens, obtained a positive correlation between virus transmission and temperature, which seems unexpected given the prevailing observations of other studies (Adhikari and Yin,

2020). This result is most likely a consequence of analyzing data for a small area (Queens only) where the temperature varies in a relatively narrow range of values, as well as correlating the number of detected cases, which may be sensitive to variations in the testing procedure.

Another environmental agent that can destroy or inactivate viruses is UV radiation from sunlight, and the properties of a particular virus determine how long it can remain infectious when exposed to radiation. For example, epidemics of influenza have a seasonal character precisely due to the susceptibility of influenza viruses to UV radiation (Sagripanti and Lytle, 2007). Our analysis found, at very high statistical significance, a strong negative correlation between the transmission of the SARS-CoV-2 virus and the intensity of UV radiation, which is consistent with the results of other studies obtained for the cities of Brazil and the provinces of Iran (Ahmadi et al., 2020; Mendonça et al., 2020). It is worth mentioning that lower temperatures, humidity, and sunlight levels usually occur in combination and directly affect not only the virus but also the human behavior, so the observed higher transmission of the virus in such conditions can alternatively be interpreted by indirect effects of other factors that act together in cold weather, such as more time spent indoors where the virus spreads more easily, or weakening of the immune system that increases susceptibility to infections (Abdullahi et al., 2020).

While the results related to COVID-19 correlations with temperature, humidity, and UV radiation are fairly frequent in the literature, this is less so for the results on the precipitation levels. Very few other studies have examined the association of precipitation with SARS-CoV-2 transmission, with either no correlation found (Pourghasemi et al., 2020), or looking at precipitation as a surrogate for humidity and generally receiving a negative correlation with infection rate (Araujo and Naimi, 2020; Coro, 2020). Our results, however, shown in **Figure 6D**, confirm natural expectations: just like humidity, the precipitation exhibits a strong negative correlation with R_0, only slightly lower than in the case of T, H, and UV, at a very high level of statistical significance. Such results also concur with some general conclusions about the behavior of similar viruses (Agrawal et al., 2009; Pica and Bouvier, 2012).

Our analysis did not reveal any statistically significant correlation either between the wind speed or between air pressure and SARS-CoV-2 transmissibility. In the case of wind speed, this result agrees with the findings in some other papers (Gupta et al., 2020; Oliveiros et al., 2020). A positive correlation of wind speed with COVID-19 transmissibility was obtained in a study in Chilean cities, but, as the authors themselves note, the interpretation of the effect of this factor is complicated by its observed significant interaction with temperature (Correa-Araneda et al., 2020). The role of wind in transmitting the virus to neighboring buildings is predicted by the SARS virus spread model within the Amoy Gardens residential complex in Hong Kong, but such an effect may relate to local air currents and virus transmission over relatively short distances and does not imply a correlation of mean wind speeds in the area with virus transmission (McKinney et al., 2006; Pica and Bouvier, 2012). As for the air pressure, the potential connection is hardly at all investigated in the literature. An exception is a study (Cambaza et al., 2020) reporting a positive correlation of air pressure with the number of COVID-19 cases in parts of Mozambique, but our results do not confirm such a conclusion.

CONCLUSION

While there is by now a significant amount of research on a crucial problem of how environmental factors affect COVID19 spread, several features set this analysis apart from the existing research. First is the applied methodology: instead of basing analysis directly on the number of detected COVID-19 cases (or some of its simple derivatives), we employ an adapted SEIR model to extract the basic reproduction number R_0 from the initial stage of the epidemic. By taking into account only data in the exponential growth regime, i.e., before the social measures took effect (as explained in the "Methods" section), we ensured that the correlations we have later identified were not confounded with the effects of local COVID-19 policies. Even more importantly, our method is also invariant to variations in COVID-19 testing practices, which, as is well known, used to vary in quite an unpredictable manner between different countries. Another important factor is the large geographical scope of our research: we collected data from 118 countries worldwide, more precisely, from all the countries that were above a certain threshold for the number of confirmed COVID-19 cases (except for several countries with clearly irregular early growth data). The third factor was the number of analyzed parameters: we calculated correlations for the selected 42 different variables (of more than a hundred that we initially considered overall) and looked for viable interpretations of the obtained results.

These results should also help in resolving some of the existing disputes in the literature. For example, our findings indicate that correlation of HDI with R_0 is not a consequence of the COVID-19 testing bias, as was occasionally argued. Of the opposing opinions, our data seem to support assertions that blood types are indeed related to COVID-19 transmissibility, as well as arguments that the higher prevalence of smoking does increase the virus transmissibility (though weakly). On the other hand, in the dispute about the effects of the pollution, our correlations give an edge to claims that there is no correlation between PM2.5 and PM10 particles and transmissibility (whereas we agree with the prevailing conclusions about the positive correlation of other considered pollutants). In the case of the effects of the wind, based on the obtained results we tend to side with those denying any connection. In certain cases our findings contradict popular narratives: there are no clear indications that either number of refugees or physical inactivity intensifies the spread of COVID-19. Unfortunately, our data also suggest that BCG immunization may not help in subduing the epidemic. Additionally, the obtained correlations hint to possible new alleys of research, e.g., those that would help us understand the connection between cholesterol levels and SARS-CoV-2 transmissibility.

Overall, we believe that the presented results can be a useful contribution to the ongoing attempts to better understand the first pandemic of the twenty-first century—and the better we understand it, the sooner we may hope to overcome it.

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/s.

AUTHOR CONTRIBUTIONS

MarD and MagD conceived the research. The work was supervised by IS, MagD, and MarD. Data acquisition by OM, IS, DZ, and AR. Data analysis by DZ, OM, and IS, with the help of MagD. Figures and tables made by DZ and AR with the help of MagD. A literature search by AR. Manuscript written by IS, AR, and MarD. All authors contributed to the article and approved the submitted version.

FUNDING

This work was partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

ACKNOWLEDGMENTS

The authors thank Anica Brzakovic for reproducing a subset of the results presented in this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.617841/full#supplementary-material

REFERENCES

Abdullahi, I. N., Emeribe, A. U., Mustapha, J. O., Fasogbon, S. A., Ofor, I. B., Opeyemi, I. S., et al. (2020). Exploring the genetics, ecology of SARS-COV-2 and climatic factors as possible control strategies against COVID-19. *Infez. Med.* 28, 166–173.

Adhikari, A., and Yin, J. (2020). Short-term effects of ambient ozone, PM2.5, and meteorological factors on COVID-19 confirmed cases and deaths in queens, New York. *Int. J. Environ. Res. Public Health* 17:4047. doi: 10.3390/ijerph17114047

Agrawal, A. S., Sarkar, M., Chakrabarti, S., Rajendran, K., Kaur, H., Mishra, A. C., et al. (2009). Comparative evaluation of real-time PCR and conventional RT-PCR during a 2 year surveillance for influenza and respiratory syncytial virus among children with acute respiratory infections in Kolkata, India, reveals a distinct seasonality of infection. *J. Med. Microbiol.* 58, 1616–1622. doi: 10.1099/jmm.0.011304-0

Ahmadi, M., Sharifi, A., Dorosti, S., Ghoushchi, S. J., and Ghanbari, N. (2020). Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *Sci. Total Environ.* 729:138705. doi: 10.1016/j.scitotenv.2020.138705

An, C., Lim, H., Kim, D. W., Chang, J. H., Choi, Y. J., and Kim, S. W. (2020). Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. *Sci. Rep.* 10:18716. doi: 10.1038/s41598-020-75767-2

Anderson, R. M., and May, R. M. (1992). *Infectious Diseases of Humans: Dynamics and Control.* New York, NY:: Oxford University Press.

Ansari-Lari, M., and Saadat, M. (2020). The morbidity and mortality of COVID-19 are associated with ABO and Rh blood groups. *Eur. J. Prev. Cardiol.* doi: 10.1177/2047487320939216. [Epub ahead of print].

Arac, E., Solmaz, I., Akkoc, H., Donmezdil, S., Karahan, Z., Kaya, S., et al. (2020). Association between the Rh blood group and the Covid-19 susceptibility. *UHOD-Uluslar. Hematol.* 30, 81–86. doi: 10.4999/uhod.204247

Araujo, M. B., and Naimi, B. (2020). Spread of SARS-CoV-2 coronavirus likely constrained by climate. medRxiv [Preprint]. Available online at https://europepmc.org/article/ppr/ppr117563 (accessed December 16, 2020).

Bar-On, Y. M., Flamholz, A., Phillips, R., and Milo, R. (2020). SARS-CoV-2 (COVID-19) by the numbers. *Elife* 9:e57309. doi: 10.7554/eLife.57309.sa2

Becker, S., and Soukup, J. M. (1999). Exposure to urban air particulates alters the macrophage-mediated inflammatory response to respiratory viral infection. *J. Toxicol. Environ. Health Part A* 57, 445–457. doi: 10.1080/009841099157539

Berg, M. K., Yu, Q., Salvador, C. E., Melani, I., and Kitayama, S. (2020). Mandated Bacillus Calmette-Guérin (BCG) vaccination predicts flattened curves for the spread of COVID-19. *Sci. Adv.* 6:eabc1463. doi: 10.1126/sciadv.abc1463

Brake, S. J., Barnsley, K., Lu, W., McAlinden, K. D., Eapen, M. S., and Sohal, S. S. (2020). Smoking upregulates angiotensin-converting enzyme-2 receptor: a potential adhesion site for novel coronavirus SARS-CoV-2 (Covid-19). *J. Clin. Med.* 9:841. doi: 10.3390/jcm9030841

Cambaza, E. M., Viegas, G. C., and Cambaza, C. M. (2020). Potential impact of temperature and atmospheric pressure on the number of cases of COVID-19 in Mozambique, Southern Africa. *J. Public Health Epidemiol.* 12, 246–260. doi: 10.5897/JPHE2020.1258

Carozzi, F. (2020). Urban density and COVID-19. *Institute for the Study of Labor (IZA)* [Online], 13440. Available online at https://ssrn.com/abstract=3643204 (accessed December 16, 2020).

Casanova, L. M., Jeon, S., Rutala, W. A., Weber, D. J., and Sobsey, M. D. (2010). Effects of air temperature and relative humidity on coronavirus survival on surfaces. *Appl. Environ. Microbiol.* 76, 2712–2717. doi: 10.1128/AEM.02291-09

Chan, K.-H., Peiris, J. M., Lam, S., Poon, L., Yuen, K., and Seto, W. H. (2011). The effects of temperature and relative humidity on the viability of the SARS coronavirus. *Adv. Virol.* 2011:734690. doi: 10.1155/2011/734690

Chatkin, J. M., and Godoy, I. (2020). Are smoking, environmental pollution, and weather conditions risk factors for COVID-19? *J. Bras. Pneumol.* 46:e20200183. doi: 10.36416/1806-3756/e20200183

Chiang, W.-H., Liu, X., and Mohler, G. (2020). Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates. *medRxiv [Preprint]*. Available online at: https://europepmc.org/article/ppr/ppr172763 (accessed December 16, 2020).

Chick, J. (2020). Alcohol and COVID-19. *Alcohol Alcohol.* 55, 341–342. doi: 10.1093/alcalc/agaa039

Coccia, M. (2020). The effects of atmospheric stability with low wind speed and of air pollution on the accelerated transmission dynamics of COVID-19. *Int. J. Environ. Stud.* doi: 10.1080/00207233.2020.1802937. [Epub ahead of print].

Cohen, J., and Kupferschmidt, K. (2020). Countries test tactics in 'war' against COVID-19. *Science* 367, 1287–1288. doi: 10.1126/science.367.6484.1287

Contini, D., and Costabile, F. (2020). Does air pollution influence COVID-19 outbreaks? *Atmosphere* 11:377. doi: 10.3390/atmos11040377

Copat, C., Cristaldi, A., Fiore, M., Grasso, A., Zuccarello, P., Signorelli, S. S., et al. (2020). The role of air pollution (PM and NO_2) in COVID-19 spread and lethality: a systematic review. *Environ. Res.* 191:110129. doi: 10.1016/j.envres.2020.110129

Copiello, S., and Grillenzoni, C. (2020). The spread of 2019-nCoV in China was primarily driven by population density. Comment on "Association between short-term exposure to air pollution and COVID-19 infection: evidence from China" by Zhu et al. *Sci. Total Environ.* 744:141028. doi: 10.1016/j.scitotenv.2020.141028

Coro, G. (2020). A global-scale ecological niche model to predict SARS-CoV-2 coronavirus infection rate. *Ecol. Modell.* 431:109187. doi: 10.1016/j.ecolmodel.2020.109187

Correa-Araneda, F., Ulloa-Yañez, A., Núñez, D., Boyero, L., Tonin, A., Cornejo, A., et al. (2020). Environmental determinants of COVID-19 transmission across a

wide climatic gradient in Chile. Research Square [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr166276 (accessed December 16, 2020).

Dai, X. (2020). ABO blood group predisposes to COVID-19 severity and cardiovascular diseases. *Eur. J. Prev. Cardiol.* 27, 1436–1437. doi: 10.1177/2047487320922370

Daraei, H., Toolabian, K., Kazempour, M., and Javanbakht, M. (2020). The role of the environment and its pollution in the prevalence of COVID-19. *J. Infect.* 81, e168–e169. doi: 10.1016/j.jinf.2020.06.019

Day, M. (2020). Covid-19: four fifths of cases are asymptomatic, China figures indicate. *BMJ* 369:m1375. doi: 10.1136/bmj.m1375

Djordjevic, M., Djordjevic, M., Salom, I., Rodic, A., Zigic, D., Milicevic, O., et al. (2020). COVID-19 puzzle in China: a serendipitous interplay between transmissibility and social distancing measures. arXiv [Preprint]. Available online at: https://arxiv.org/abs/2005.09630 (accessed December 16, 2020).

Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *Lancet. Infect. Dis.* 20, 533–534. doi: 10.1016/S1473-3099(20)30120-1

Eubank, S., Eckstrand, I., Lewis, B., Venkatramanan, S., Marathe, M., and Barrett, C. (2020). Commentary on Ferguson, et al., "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand". *Bull. Math. Biol.* 82:52. doi: 10.1007/s11538-020-00726-x

European Commission Global Human Settlement (2020). *Urban Centre Database UCDB R2019A* [Online]. Available online at: https://ghsl.jrc.ec.europa.eu/ghs_stat_ucdb2015mt_r2019a.php (accessed September 26, 2020).

Fan, C., Cai, T., Gai, Z., and Wu, Y. (2020). The relationship between the migrant population's migration network and the risk of COVID-19 transmission in China—empirical analysis and prediction in prefecture-level cities. *Int. J. Environ. Res. Public Health* 17:2630. doi: 10.3390/ijerph17082630

Fareed, Z., Iqbal, N., Shahzad, F., Shah, S. G. M., Zulfiqar, B., Shahzad, K., et al. (2020). Co-variance nexus between COVID-19 mortality, humidity, and air quality index in Wuhan, China: new insights from partial and multiple wavelet coherence. *Air Qual. Atmos. Health* 8, 1–10. doi: 10.1007/s11869-020-00847-1

Gallagher, J. (2020, October 11). BCG: can a vaccine from 1921 save lives from Covid-19? *BBC News*.

Gangemi, S., Billeci, L., and Tonacci, A. (2020). Rich at risk: socio-economic drivers of COVID-19 pandemic spread. *Clin. Mol. Allergy* 18:12. doi: 10.1186/s12948-020-00127-4

Garufi, G., Carbognin, L., Orlandi, A., Tortora, G., and Bria, E. (2020). Smoking habit and hospitalization for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)-related pneumonia: the unsolved paradox behind the evidence. *Eur. J. Intern. Med.* 77, 121–122. doi: 10.1016/j.ejim.2020.04.042

Gérard, C., Maggipinto, G., and Minon, J.-M. (2020). COVID-19 and ABO blood group: another viewpoint. *Brit. J. Haematol.* 190, e57–e94. doi: 10.1111/bjh.16884

Global Data Lab (2020). *Subnational Human Development Index (4.0)* [Online]. Available online at://globaldatalab.org/shdi/(accessed September 27, 2020).

Goldberg, D. L., Anenberg, S. C., Griffin, D., McLinden, C. A., Lu, Z., and Streets, D. G. (2020). Disentangling the impact of the COVID-19 lockdowns on urban NO2 from natural variability. *Geophys. Res. Lett.* 47:e2020GL089269. doi: 10.1002/essoar.10503396.1

Guillon, P., Clément, M., Sébille, V., Rivain, J.-G., Chou, C.-F., Ruvoën-Clouet, N., et al. (2008). Inhibition of the interaction between the SARS-CoV spike protein and its cellular receptor by anti-histo-blood group antibodies. *Glycobiology* 18, 1085–1093. doi: 10.1093/glycob/cwn093

Guo, X.-J., Zhang, H., and Zeng, Y.-P. (2020). Transmissibility of COVID-19 in 11 major cities in China and its association with temperature and humidity in Beijing, Shanghai, Guangzhou, and Chengdu. *Infect. Dis. Poverty* 9:87. doi: 10.1186/s40249-020-00708-0

Gupta, A., Banerjee, S., and Das, S. (2020). Significance of geographical factors to the COVID-19 outbreak in India. *Model. Earth Syst. Environ.* 6, 2645–2653. doi: 10.1007/s40808-020-00838-2

Haidar, Y. M., and Cosman, B. C. (2011). Obesity epidemiology. *Clin. Colon Rectal Surg.* 24, 205–210. doi: 10.1055/s-0031-1295684

Hargreaves, S., Kumar, B. N., McKee, M., Jones, L., and Veizis, A. (2020). Europe's migrant containment policies threaten the response to covid-19. *BMJ* 368:m1213. doi: 10.1136/bmj.m1213

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181, 271–280. doi: 10.1016/j.cell.2020.02.052

Iqbal, M. M., Abid, I., Hussain, S., Shahzad, N., Waqas, M. S., and Iqbal, M. J. (2020). The effects of regional climatic condition on the spread of COVID-19 at global scale. *Sci. Total Environ.* 739:140101. doi: 10.1016/j.scitotenv.2020.140101

Jordan, R. E., Adab, P., and Cheng, K. K. (2020). Covid-19: risk factors for severe disease and death. *BMJ* 368:m1198. doi: 10.1136/bmj.m1198

Jurak, G., Morrison, S. A., Leskošek, B., Kovač, M., HadŽić, V., Vodičar, J., et al. (2020). Physical activity recommendations during the coronavirus disease-2019 virus outbreak. *J. Sport Health Sci.* 9, 325–327. doi: 10.1016/j.jshs.2020.05.003

Keeling, M. J., and Rohani, P. (2011). *Modeling Infectious Diseases in Humans and Animals*. Princeton, NJ: Princeton University Press. doi: 10.2307/j.ctvcm4gk0

Khajanchi, S., Bera, S., and Roy, T. K. (2020a). Mathematical analysis of the global dynamics of a HTLV-I infection model, considering the role of cytotoxic T-lymphocytes. *Math. Comput. Simulat.* 180, 354–378. doi: 10.1016/j.matcom.2020.09.009

Khajanchi, S., and Sarkar, K. (2020). Forecasting the daily and cumulative number of cases for the COVID-19 pandemic in India. *Chaos* 30:071101. doi: 10.1063/5.0016240

Khajanchi, S., Sarkar, K., Mondal, J., and Perc, M. (2020b). Dynamics of the COVID-19 pandemic in India. arXiv [Preprint]. Available at: https://arxiv.org/abs/2005.06286 (accessed December 16, 2020). doi: 10.21203/rs.3.rs-27112/v1

Kimball, A., Hatfield, K. M., Arons, M., James, A., Taylor, J., Spicer, K., et al. (2020). Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility—King County, Washington, March 2020. *MMWR Morb. Mortal. Wkly Rep.* 69, 377–381. doi: 10.15585/mmwr.mm6913e1

Latz, C. A., DeCarlo, C., Boitano, L., Png, C. M., Patell, R., Conrad, M. F., et al. (2020). Blood type and outcomes in patients with COVID-19. *Ann. Hematol.* 99, 2113–2118. doi: 10.1007/s00277-020-04169-1

Li, M., Zhang, Z., Cao, W., Liu, Y., Du, B., Chen, C., et al. (2020). Identifying novel factors associated with COVID-19 transmission and fatality using the machine learning approach. *Sci. Total Environ.* 142810. doi: 10.1016/j.scitotenv.2020.142810. [Epub ahead of print].

Lin, S., Wei, D., Sun, Y., Chen, K., Yang, L., Liu, B., et al. (2020). Region-specific air pollutants and meteorological parameters influence COVID-19: a study from mainland China. *Ecotoxicol. Environ. Saf.* 204:111035. doi: 10.1016/j.ecoenv.2020.111035

Liu, Y., Ning, Z., Chen, Y., Guo, M., Gali, N. K., Sun, L., et al. (2020). Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature* 582, 557–560. doi: 10.1038/s41586-020-2271-3

Lowen, A. C., Mubareka, S., Steel, J., and Palese, P. (2007). Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog.* 3, 1470–1476. doi: 10.1371/journal.ppat.0030151

Maier, B. F., and Brockmann, D. (2020). Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science* 368, 742–746. doi: 10.1126/science.abb4557

Martcheva, M. (2015). *An Introduction to Mathematical Epidemiology*. Boston, MA: Springer. doi: 10.1007/978-1-4899-7612-3

Maslov, S., and Goldenfeld, N. (2020). Window of opportunity for mitigation to prevent overflow of ICU capacity in Chicago by COVID-19. medRxiv [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr118721 (accessed December 16, 2020).

McKinney, K. R., Yu, Y. G., and Lewis, T. G. (2006). Environmental transmission of SARS at amoy gardens. *J. Environ. Health* 68, 26–30.

Mecenas, P., Bastos, R., Vallinoto, A. C. R., and Normando, D. (2020). Effects of temperature and humidity on the spread of COVID-19: a systematic review. *PLoS ONE* 15:e0238339. doi: 10.1371/journal.pone.0238339

Mendonça, F., Anjos, M., Collischonn, E., Murara, P., DE, F., Limberger, L., et al. (2020). Climate and Covid-19–upgrade and solar radiation influences based on Brazil cases. Research Square [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr174953 (accessed December 16, 2020).

Najafimehr, H., Mohamed Ali, K., Safari, S., Yousefifard, M., and Hosseini, M. (2020). Estimation of basic reproduction number for COVID-19 and the reasons for its differences. *Int. J. Clin. Pract.* 74:e13518. doi: 10.1111/ijcp.13518

NASA Langley Research Center (2020). *The Prediction of Worldwide Energy Resources (POWER) Project [Online]*. Available online at: https://power.larc.nasa.gov/ (accessed October 6, 2020).

Oliveiros, B., Caramelo, L., Ferreira, N. C., and Caramelo, F. (2020). Role of temperature and humidity in the modulation of the doubling time of COVID-19 cases. medRxiv [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr116498 (accessed December 16, 2020).

O'Neill, L. A., and Netea, M. G. (2020). BCG-induced trained immunity: can it offer protection against COVID-19? *Nat. Rev. Immunol.* 20, 335–337. doi: 10.1038/s41577-020-0337-y

Our World in Data (2020). *Research and Data to Make Progress Against the World's Largest Problems* [Online]. Available online at: https://ourworldindata.org/ (accessed May, 2020).

Pansini, R., and Fornacca, D. (2020). COVID-19 higher induced mortality in Chinese regions with lower air quality. medRxiv [Preprint]. Available online at //www.medrxiv.org/content/10.1101/2020.05.28.20115832v2 (accessed December 16, 2020).

Pawelec, G., and Larbi, A. (2008). Immunity and ageing in man: annual review 2006/2007. *Exp. Gerontol.* 43, 34–38. doi: 10.1016/j.exger.2007.09.009

Perkins, T. A., and España, G. (2020). Optimal control of the COVID-19 pandemic with non-pharmaceutical interventions. *Bull. Math. Biol.* 82:118. doi: 10.1007/s11538-020-00795-y

Pica, N., and Bouvier, N. M. (2012). Environmental factors affecting the transmission of respiratory viruses. *Curr. Opin. Virol.* 2, 90–95. doi: 10.1016/j.coviro.2011.12.003

Pourghasemi, H. R., Pouyan, S., Heidari, B., Farajzadeh, Z., Fallah Shamsi, S. R., Babaei, S., et al. (2020). Spatial modeling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between February 19 and June 14, 2020). *Int. J. Infect. Dis.* 98, 90–108. doi: 10.1016/j.ijid.2020.06.058

Pranata, R., Lim, M. A., Huang, I., Raharjo, S. B., and Lukito, A. A. (2020). Hypertension is associated with increased mortality and severity of disease in COVID-19 pneumonia: a systematic review, meta-analysis and meta-regression. *J. Renin Angiotensin Aldosterone Syst.* 21:1470320320926899. doi: 10.1177/1470320320926899

Qu, G., Li, X., Hu, L., and Jiang, G. (2020). An imperative need for research on the role of environmental factors in transmission of novel coronavirus (COVID-19). *Environ. Sci. Technol.* 54, 3730–3732. doi: 10.1021/acs.est.0c01102

Radenkovic, D., Chawla, S., Pirro, M., Sahebkar, A., and Banach, M. (2020). Cholesterol in relation to COVID-19: should we care about it? *J. Clin. Med.* 9:1909. doi: 10.3390/jcm9061909

Ran, J., Song, Y., Zhuang, Z., Han, L., Zhao, S., Cao, P., et al. (2020a). Blood pressure control and adverse outcomes of COVID-19 infection in patients with concomitant hypertension in Wuhan, China. *Hypertens. Res.* 43, 1267–1276. doi: 10.1038/s41440-020-00541-w

Ran, J., Zhao, S., Han, L., Qiu, Y., Cao, P., Yang, Z., et al. (2020b). Effects of particulate matter exposure on the transmissibility and case fatality rate of COVID-19: a Nationwide Ecological Study in China. *J. Travel Med.* 27:taaa133. doi: 10.1093/jtm/taaa133

Rashed, E. A., Kodera, S., Gomez-Tames, J., and Hirata, A. (2020). Influence of absolute humidity, temperature and population density on COVID-19 spread and decay durations: multi-prefecture study in Japan. *Int. J. Environ. Res. Public Health* 17:5354. doi: 10.3390/ijerph17155354

Ruiz, J. I., Nuhu, K., McDaniel, J. T., Popoff, F., Izcovich, A., and Criniti, J. M. (2015). Inequality as a powerful predictor of infant and maternal mortality around the world. *PLoS ONE* 10:e0140796. doi: 10.1371/journal.pone.0140796

Rychter, A. M., Zawada, A., Ratajczak, A. E., Dobrowolska, A., and Krela-Kazmierczak, I. (2020). Should patients with obesity be more afraid of COVID-19? *Obes. Rev.* 21:e13083. doi: 10.1111/obr.13083

Sagar, A. D., and Najam, A. (1998). The human development index: a critical review. *Ecol. Econ.* 25, 249–264. doi: 10.1016/S0921-8009(97)00168-7

Sagripanti, J.-L., and Lytle, C. D. (2007). Inactivation of influenza virus by solar radiation. *Photochem. Photobiol.* 83, 1278–1282. doi: 10.1111/j.1751-1097.2007.00177.x

Samui, P., Mondal, J., and Khajanchi, S. (2020). A mathematical model for COVID-19 transmission dynamics with a case study of India. *Chaos Solitons Fractals* 140:110173. doi: 10.1016/j.chaos.2020.110173

Sarkar, K., Khajanchi, S., and Nieto, J. J. (2020). Modeling and forecasting the COVID-19 pandemic in India. *Chaos Solitons Fractals* 139:110049. doi: 10.1016/j.chaos.2020.110049

Sattar, N., McInnes, I. B., and McMurray, J. J. (2020). Obesity a risk factor for severe COVID-19 infection: multiple potential mechanisms. *Circulation* 142, 4–6. doi: 10.1161/CIRCULATIONAHA.120.047659

Scheffer, M. (2009). *Critical Transitions in Nature and Society*. Princeton, NJ: Princeton University Press. doi: 10.1515/9781400833276

Schiffrin, E. L., Flack, J. M., Ito, S., Muntner, P., and Webb, R. C. (2020). Hypertension and COVID-19. *Am. J. Hypertens.* 33, 373–374. doi: 10.1093/ajh/hpaa057

Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M. G., Borelli, M., et al. (2020). SARS-Cov-2RNA found on particulate matter of Bergamo in Northern Italy: first evidence. *Environ. Res.* 188:109754. doi: 10.1016/j.envres.2020.109754

Shahid, Z., Kalayanamitra, R., McClafferty, B., Kepko, D., Ramgobin, D., Patel, R., et al. (2020). COVID-19 and older adults: what we know. *J. Am. Geriatr. Soc.* 68, 926–929. doi: 10.1111/jgs.16472

Singh, K., and Agarwal, A. (2020). Impact of weather indicators on the COVID-19 outbreak: a multi-state study in India. medRxiv [Preprint]. Available online at https://europepmc.org/article/ppr/ppr176027 (accessed December 16, 2020).

Thangriyal, S., Rastogi, A., Tomar, A., and Baweja, S. (2020). Impact of temperature and sunshine duration on daily new cases and death due to COVID-19. medRxiv [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr176097 (accessed December 16, 2020).

The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team (2020). The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19)—China, 2020. *China CDC Weekly* [Online], 2. Available online at: https://slma.lk/wp-content/uploads/2020/02/TheEpidemiologicalCharacteristicsofanOutbreakof2019NovelCoronavirusDiseases28COVID-1929E28094China2C20201.pdf. doi: 10.46234/ccdcw2020.032

Tian, H., Liu, Y., Li, Y., Wu, C. H., Chen, B., Kraemer, M. U. G., et al. (2020). An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 368, 638–642. doi: 10.1126/science.abb6105

Tosepu, R., Gunawan, J., Effendy, D. S., Ahmad, O. A. I., Lestari, H., Bahar, H., et al. (2020). Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia. *Sci. Total Environ.* 725:138436. doi: 10.1016/j.scitotenv.2020.138436

van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., et al. (2020). Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N. Engl. J. Med.* 382, 1564–1567. doi: 10.1056/NEJMc2004973

Ward, D. (2020). *Actions Speak Louder Than Age: Explaining Wide Variations in COVID-19 Deaths* [Online]. Available online at: https://www.researchgate.net/publication/341599695_Actions_Speak_Louder_than_Age_Explaining_Wide_Variations_in_COVID-19_Deaths (accessed December 16, 2020).

Weitz, J. S., Beckett, S. J., Coenen, A. R., Demory, D., Dominguez-Mirazo, M., Dushoff, J., et al. (2020). Modeling shield immunity to reduce COVID-19 epidemic spread. *Nat. Med.* 26, 849–854. doi: 10.1038/s41591-020-0895-3

World Bank (2020a). *Global Economic Prospects, June 2020*. Washington, DC: World Bank.

World Bank (2020b). *World Bank Open Data* [Online]. Available online at: https://www.worldbank.org/ (accessed May, 2020).

World Health Organization (2020). *The Global Health Observatory: Explore a World of Health Data* [Online]. Available online at: https://www.who.int/data/gho (accessed October 6, 2020).

World Weather Online (2020). *Local Weather History API* [Online]. Available online at: https://www.worldweatheronline.com/developer/api/historical-weather-api.aspx (accessed December 9, 2020).

Worldometer (2020). *COVID-19 Coronavirus Pandemic* [Online]. Available online at: https://www.worldometers.info/coronavirus/ (accessed June 14, 2020).

Wu, X., Nethery, R. C., Sabath, M. B., Braun, D., and Dominici, F. (2020). Air pollution and COVID-19 mortality in the United States: strengths and limitations of an ecological regression analysis. *Sci. Adv.* 6:eabd4049. doi: 10.1126/sciadv.abd4049

Yao, Y., Pan, J., Liu, Z., Meng, X., Wang, W., Kan, H., et al. (2020a). Ambient nitrogen dioxide pollution and spread ability of COVID-19 in Chinese cities. *Ecotoxicol. Environ. Saf.* 208:111421. doi: 10.1016/j.ecoenv.2020.111421

Yao, Y., Pan, J., Liu, Z., Meng, X., Wang, W., Kan, H., et al. (2020b). No association of COVID-19 transmission with temperature or UV radiation in Chinese cities. *Eur. Respir. J.* 55:2000517. doi: 10.1183/13993003.00517-2020

Yao, Y., Pan, J., Wang, W., Liu, Z., Kan, H., Qiu, Y., et al. (2020c). Association of particulate matter pollution and case fatality rate of COVID-19 in 49 Chinese cities. *Sci. Total Environ.* 741:140396. doi: 10.1016/j.scitotenv.2020.140396

Yu, Q., Salvador, C., Melani, I., Berg, M., Neblett, E., and Kitayama, S. (2020). Racial residential segregation and economic disparity jointly exacerbate the COVID-19 fatality in large American cities. *PsyArXiv [Preprint]*. Available online at: https://psyarxiv.com/xgbpy/ (accessed December 16, 2020).

Zheng, Z., Peng, F., Xu, B., Zhao, J., Liu, H., Peng, J., et al. (2020). Risk factors of critical and mortal COVID-19 cases: a systematic literature review and meta-analysis. *J. Infect.* 81:e16–e25. doi: 10.1016/j.jinf.2020.04.021

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.

frontiers in Medicine

BRIEF RESEARCH REPORT
published: 10 March 2021
doi: 10.3389/fmed.2021.613412

Insights on the Structural Variations of the Furin-Like Cleavage Site Found Among the December 2019–July 2020 SARS-CoV-2 Spike Glycoprotein: A Computational Study Linking Viral Evolution and Infection

Marni E. Cueno[1,2,3], Miu Ueno[3], Rinako Iguchi[3], Tsubasa Harada[3], Yoshifumi Miki[2], Kanae Yasumaru[2], Natsumi Kiso[2], Kanta Wada[2], Koki Baba[2] and Kenichi Imai[1]*

[1] *Department of Microbiology, Nihon University School of Dentistry, Tokyo, Japan,* [2] *Immersion Biology Class, Department of Science, Tokyo Gakugei University International Secondary School, Tokyo, Japan,* [3] *Immersion Physics Class, Department of Science, Tokyo Gakugei University International Secondary School, Tokyo, Japan*

OPEN ACCESS

Edited by:
Matteo Convertino, Hokkaido University, Japan

Reviewed by:
Takahiro Watanabe, Nagoya University, Japan
Vladimir N. Uversky, University of South Florida, United States

***Correspondence:**
Marni E. Cueno
marni.cueno@nihon-u.ac.jp

Specialty section:
This article was submitted to Infectious Diseases Surveillance, Prevention and Treatment, a section of the journal Frontiers in Medicine

Received: *02 October 2020*
Accepted: *16 February 2021*
Published: *10 March 2021*

Citation:
Cueno ME, Ueno M, Iguchi R, Harada T, Miki Y, Yasumaru K, Kiso N, Wada K, Baba K and Imai K (2021) Insights on the Structural Variations of the Furin-Like Cleavage Site Found Among the December 2019–July 2020 SARS-CoV-2 Spike Glycoprotein: A Computational Study Linking Viral Evolution and Infection. Front. Med. 8:613412. doi: 10.3389/fmed.2021.613412

The SARS-CoV-2 (SARS2) is the cause of the coronavirus disease 2019 (COVID-19) pandemic. One unique structural feature of the SARS2 spike protein is the presence of a furin-like cleavage site (FLC) which is associated with both viral pathogenesis and host tropism. Specifically, SARS2 spike protein binds to the host ACE-2 receptor which in-turn is cleaved by furin proteases at the FLC site, suggesting that SARS2 FLC structural variations may have an impact on viral infectivity. However, this has not yet been fully elucidated. This study designed and analyzed a COVID-19 genomic epidemiology network for December 2019 to July 2020, and subsequently generated and analyzed representative SARS2 spike protein models from significant node clusters within the network. To distinguish possible structural variations, a model quality assessment was performed before further protein model analyses and superimposition of the protein models, particularly in both the receptor-binding domain (RBD) and FLC. Mutant spike models were generated with the unique ^{681}PRRA^{684} amino acid sequence found within the deleted FLC. We found 9 SARS2 FLC structural patterns that could potentially correspond to nine node clusters encompassing various countries found within the COVID-19 genomic epidemiology network. Similarly, we associated this with the rapid evolution of the SARS2 genome. Furthermore, we observed that either in the presence or absence of the unique ^{681}PRRA^{684} amino acid sequence no structural changes occurred within the SARS2 RBD, which we believe would mean that the SARS2 FLC has no structural influence on SARS2 RBD and may explain why host tropism was maintained.

Keywords: furin-like cleavage site, infection clusters, SARS-CoV-2 (SARS2), spike glycoprotein, structural variations

INTRODUCTION

Coronaviruses (CoV) are enveloped positive-stranded RNA viruses that have the largest genome among all known RNA viruses and, at present, there are seven known CoVs capable of infecting humans (1–7). Among CoV structural proteins, the spike protein is a class I viral fusion protein that is involved in viral entry, host tropism determination, viral pathogenesis, and host immune response induction (8–11). The spike protein is comprised of three segments (large ectodomain, single-pass transmembrane anchor, and short intracellular tail) (11), with the ectodomain further divided into the S1 receptor-binding subunit and S2 membrane-fusion subunit (10, 11). During a typical CoV infection, S1 binds to an ideal host receptor enabling viral attachment and, consequently, S2 would fuse the host and viral membranes, allowing viral genetic material to enter host cells (10, 11).

Interestingly, prior to SARS-CoV-2 (SARS2), there were six human pathogenic coronaviruses (10), with SARS2 resulting in a pandemic causing the coronavirus disease 2019 (COVID-19) (12, 13). With regards to the homotrimeric spike protein, the SARS2 spike protein follows the same mechanism of viral entry used by SARS-CoV-1, wherein, the SARS2 spike protein binds to a functional receptor human angiotensin-converting enzyme 2 (ACE2) via the 6-residue (L455, F486, Q493, S494, N501, Y505) receptor-binding domain (RBD) (10, 14). One notable structural feature of the SARS2 spike protein is the presence of a polybasic (furin-like) cleavage site (^{682}RRAR^{685}) which has been found to be disordered (15, 16) and, likewise, linked to effective furin cleavage that could help determine viral pathogenesis and host tropism (17–19). Moreover, the comparative analysis of the intrinsic disorder predisposition of spike protein from SARS2, SARS, and Bat CoV revealed that the furin-like cleavage site of SARS spike is incorporated within the longer disordered region ^{676}TQTNSPRRARSVAS^{691}, which is not present in spike proteins from SARS and Bat CoV (20). The presence of disorder in a region containing a polybasic (furin-like) cleavage site is an extremely important point, as an intrinsic disorder at the cleavage site is crucial for efficient protease action (20, 21). Furthermore, aside from the presence of the polybasic cleavage site (^{682}RRAR^{685}), SARS2 likewise has an inserted leading proline (P681), which is suggested to improve protease active site accessibility not only by furin proteases but other proteases as well (21). Thus, this would mean that the inserted sequence unique for SARS2 is the ^{681}PRRA^{684} sequence (18).

The structural orientation of either individual or a series of amino acids plays an important role in establishing both protein configuration and protein-protein complexes (22), which likewise may affect protein function (23). This would imply that any probable changes in structural orientation occurring in the SARS2 spike furin-like cleavage (including P681) site (FLC) may have an impact on viral infectivity (24). However, to our knowledge, this has never been fully elucidated. A better understanding of the potential effects of the structural orientation changes occurring within the SARS2 FLC site may shed light on the occurrence of varying SARS2 variants and, more importantly, its role in viral reinfection, potentially leading to novel drug design and therapeutic strategies.

MATERIALS AND METHODS

COVID-19 Genomic Epidemiology Network Design and Analyses Between December 2019 and July 2020

Network analyses were performed in order to gather a holistic understanding of the phylogeny of the COVID-19 genomic epidemiology (25). For this study, network design followed the phylogenetic tree of the COVID-19 genomic epidemiology, based on the GISAID website (www.gisaid.org) between December 2019 and July 2020. A total of 2,793 genomes were used for both network design and analyses. We used Cytoscape for both network design and analyses (26). For network design, nodes were made to represent the countries (indicated as a box) and phylogenetic branch points (indicated as dots) while the edges represent the phylogenetic lineage originating from either a country or branch point. For network analyses, the following centrality measurements were initially analyzed: (1) stress centrality (identifying important nodes); (2) eccentricity centrality (identifying accessible nodes); (3) closeness centrality (identifying relevant nodes); (4) betweenness centrality (identifying crucial nodes); and (5) edge betweenness centrality (identifying significant edges) (27). Briefly, nodes (**Supplementary Figure 1**) and edges (**Supplementary Figure 2**) above a computed threshold for each centrality were considered significant. A unified network was designed based on all centrality measurements used for this study (both nodal and edge centralities) and, more importantly, nodes that were linked to either nodes or edges that are above the threshold based on all five centrality measurements used were determined.

SARS2 Spike Protein Modeling

Representative SARS2 spike amino acid sequences ($n = 263$) deposited between December 2019 and July 2020 were collected from the National Center for Biological Information (NCBI). The selection of sequences was based on the results obtained from our previous COVID-19 genomic epidemiology network analyses. Moreover, representative monomeric SARS2 spike models were selected using Tm align (28). Briefly, a minimum of 10 generated sequence models were initially obtained. Further structural analyses used spike models with similar Root Mean Square Deviation (RMSD) values and Template Modeling scores (Tm-scores) based on superimposition. In particular, the SARS2 spike models used for further structural analyses were based on structural variations in SARS2 FLC and have the following Genebank accession numbers: MT019529, MN994468, MT020781, MT825091, MT467261, MT658503, MT499218, MT549887, and MT461625. The Phyre2 web server (29) was used to generate all protein models while the Jmol applet (30) was used for protein visualization.

Protein Model Quality Assessment

To confirm the accuracy and suitability of the generated SARS2 spike protein models for further analyses, both contact

mapping and protein model:crystal structure superimposition were performed for model quality assessment. A protein contact map was made using the CMView applet to determine the common contact between the model and crystal (31). Moreover, higher common contact (>90%) would mean more structural similarities (32), which would mean that the generated model is suitable for further analyses. Subsequently, representative SARS2 spike cryo-EM structure (PDB ID: 6XR8) (15) and a monomeric 6XR8 model (cryo-EM model) generated using Phyre 2 were used for superimposition (using Tm align) to serve as a model quality check. For this study, SARS2 spike models were considered suitable for further analyses if superimposed sequence model:crystal and crystal model:crystal have RMSD < 1.50.

Comparison of SARS2 Spike Models

All structural comparisons conducted focused on both the SARS2 FLC and RBD. Moreover, two sets of structural comparisons were made. The first set of structural comparisons focused on contrasting the SARS2 FLC and RBD among all representative SARS2 spike models through superimposition. One of the representative models (generated from MT019529) was used as the common model for superimposition. The second set of structural comparisons involved producing mutants from all representative SARS2 spike models without the ^{681}PRRA^{684} sequence unique in SARS2. A protein threading approach (via Phyre 2) was used to generate the mutant models. Similarly, focusing on SARS2 FLC and RBD, the original model (with ^{681}PRRA^{684}) was compared to the mutated model (without ^{681}PRRA^{684}) through superimposition using Tm align. Model superimposition (focusing on SARS2 FLC and RBD), RMSD values, and Tm scores were established using Jmol and Tm align, respectively.

RESULTS

Nine Node Clusters From the COVID-19 Genomic Epidemiology Network Were Established Between December 2019 and July 2020

The SARS2 genome is constantly evolving, and genome distribution varies in terms of geographic location (33, 34). To establish possible node clusters within the COVID-19 genomic epidemiology network established between December 2019 and July 2020, network analytics was performed to elucidate the holistic and simultaneous analyses of complementary data (27, 35). One of the key points of network analytics is centrality analysis, which involves collecting network components in order to distinguish important elements and, likewise, requires several centrality measurements to be considered fully efficient for analyzing networks (27, 36). Considering this and the five different centrality measurements used to identify node clusters, this would suggest that the results obtained are reliable. Interestingly, we were able to identify nine node clusters, encompassing various SARS2 genomic clades classified by the GISAID website (**Figure 1A**). We observed that some of the countries identified among the nine node clusters are likewise found in other node clusters (regardless of belonging to different SARS2 clades) (**Figure 1B**). These results could mean that the putative significant node clusters are not dependent on SARS2 clades, which coincidentally are based on viral genome mutations (34). This insinuates that there could be other similarities among the node clusters with regard to SARS2 pathogenesis. Considering that the SARS2 FLC is crucial for viral pathogenesis and host tropism (17–19), which we believe would imply that the SARS2 FLC is a conserved structural feature (18), we postulate that the SARS2 FLC could be a common structural feature among the node clusters. We wish to emphasize that our current study mainly focused on the SARS2 FLC structural feature. In possible future work, it would be interesting to recognize other possible spike protein structural features found among the node clusters identified.

SARS2 Spike Models Are Suitable for Structural Analyses

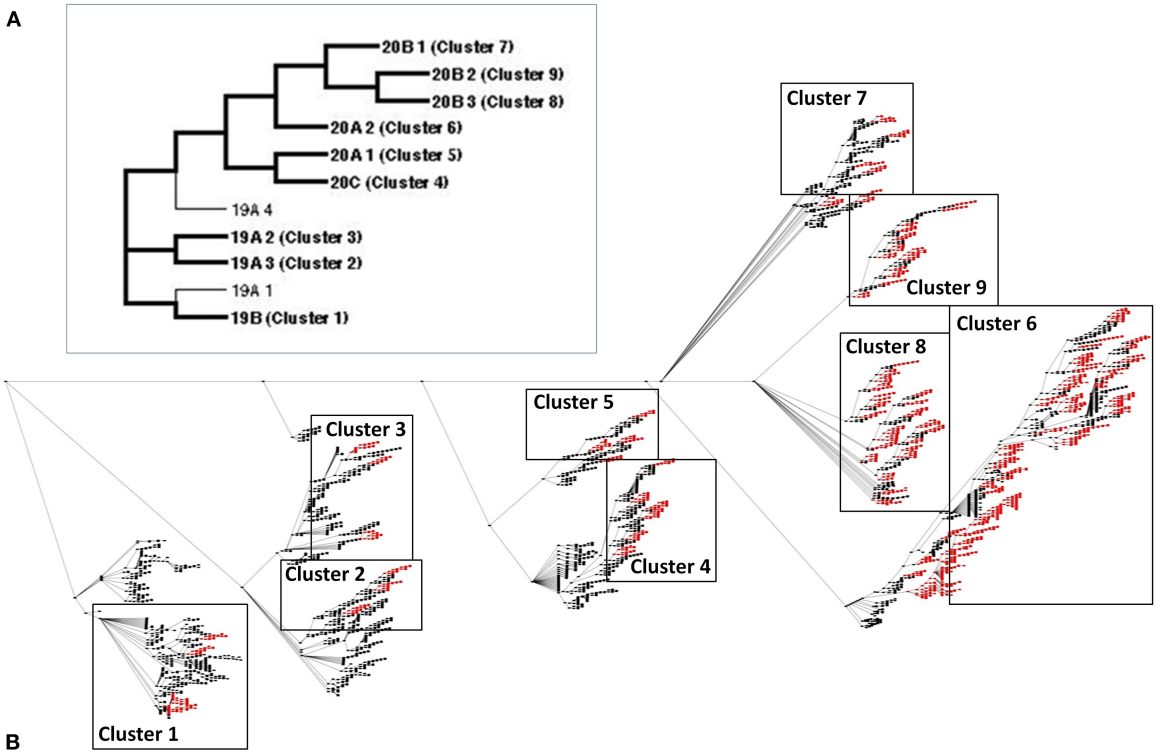
It has long been recommended that model quality assessment be performed prior to any downstream structural analyses using protein structures generated from either experimental (i.e., crystallized) or theoretical (i.e., computer-based) methods (37). To establish the reliability and suitability of all SARS2 spike models generated, both protein contact maps and structural superimpositions were performed. Representative SARS2 crystal structure (**Figure 2A**), SARS2 crystal model (**Figure 2B**), and SARS2 sequence model (**Figure 2C**) were used for all superimpositions conducted. We observed that protein contact map superimposition between crystal model:crystal structure (**Figure 2D**), sequence model:crystal structure (**Figure 2E**), and sequence model:crystal model (**Figure 2F**) have high common contact (>90%), which implies that there is high contact similarity between the superimposed structures. We only considered SARS2 spike monomers when examining structural superimpositions. We also observed that RMSD values between cryo-EM model:crystal structure [RMSD 0.75] (**Figure 2G**), sequence model:cryo-EM structure [RMSD 0.66] (**Figure 2H**), and sequence model:cryo-EM model [RMSD 1.07] (**Figure 2I**) were RMSD < 1.5 which in-turn were considered adequate for further analyses (38). These results (both protein contact map and structural superimpositions) would suggest that the generated SARS2 spike models are suitable for further structural analyses.

Nine SARS2 FLC Structural Patterns Were Identified Among the Nine Node Clusters

Protein structure and conformation dynamics have often been correlated to biological function, which emphasizes the importance of protein structural pattern variations (23). To elucidate the possible SARS2 FLC structural variations among the 9 node clusters, representative SARS2 models from each node cluster were superimposed with the SARS2 model generated from MT019529 (Wuhan, China) as a comparison. Since SARS2 FLC also affects host tropism, SARS2 RBD was similarly checked.

As seen in **Figure 3A**, both SARS2 RBD (box dash lines) and FLC (box solid lines) structural changes were the focus of

A



B

Cluster 1		Cluster 2		Cluster 3		Cluster 4			Cluster 5
.Australia	.Nepal	.Austria	.Singapore	.Australia	.Spain	.Austria	.Jordan	.Taiwan	.Australia
.Bangladesh	.New Zealand	.Belgium	.South Africa	.England	.Taiwan	.Bosnia and Herzegovina	.Luxembourg	.Thailand	.Bangladesh
.Benin	.Oman	.Bulgaria	.Taiwan	.Gambia		.Canada	.Morocco	.UAE	.Belgium
.Canada	.Pakistan	.Canada	.USA	.Georgia		.Chile	.New Zealand	.USA	.Chile
.Costa Rica	.Poland	.England	.China (Wuhan)	.Guam		.Colombia	.Nigeria		.Colombia
.Finland	.SaudiArabia	.Germany		.India		.Croatia	.Norway		.Egypt
.Germany	.Senegal	.Indonesia		.Ireland		.Denmark	.Poland		.Greece
.Ghana	.Singapore	.Luxembourg		.Israel		.England	.Romania		.India
.Iceland	.Sweden	.Malaysia		.Malaysia		.Finland	.Russia		.Kazakhstan
.India	.Taiwan	.Mexico		.Norway		.France	.Senegal		.Latvia
.Japan	.Tunisia	.Oman		.Oman		.Germany	.Singapore		.Morocco
.Jordan	.Uganda	.Poland		.Poland		.Israel	.Spain		.Saudi Arabia
.Malaysia	.Uruguay	.Senegal		.Singapore			.Sweden		.Turkey
.Mexico	.USA								.Venezuela
	.Vietnam								

Cluster 6			Cluster 7			Cluster 8			Cluster 9
.Argentina	.Georgia	.Portugal	.Argentina	.Italy	.United Arab Emirates	.Australia	.Israel	.Switzerland	.Argentina
.Australia	.Germany	.Romania	.Australia	.Japan	.Uruguay	.Bangladesh	.Italy	.Taiwan	.Australia
.Austria	.Greece	.Russia	.Austria	.Jordan	.Vietnam	.Beijing	.Japan	.Turkey	.Bangladesh
.Bangladesh	.Guangdong	.Scotland	.Bangladesh	.Kazakhstan		.Belgium	.Jordan	.Vietnam	.Belgium
.Belgium	.Hungary	.Senegal	.Beijing	.Kenya		.Bosnia and Herzegovina	.Kenya	.Wales	.Chile
.Bosnia and Herzegovina	.Iceland	.Serbia	.Belgium	.Kuwait		.Brazil	.Kuwait		.Colombia
.Brazili	.India	.Slovenia	.Brazil	.Latvia		.Bulgaria	.Latvia		.Cyprus
.Bulgaria	.Israel	.South Africa	.Canada	.Morocco		.Canada	.Luxembourg		.England
.Canada	.Italy	.Spain	.Chile	.Netherlands		.Chile	.Netherlands		.Finland
.Chile	.Japan	.Sri Lanka	.Costa Rica	.Oman		.Colombia	.New Zealand		.Germany
.Colombia	.Kenya	.Sweden	.Cyprus	.Peru		.Croatia	.Norway		.Ghana
.Costa Rica	.Kuwait	.Switzerland	.Czech Republic	.Poland		.Cyprus	.Oman		.HongKong
.Croatia	.Latvia	.Taiwan	.DRC	.Portugal		.Czech Republic	.Poland		.Italy
.Czech Republic	.Luxembourg	.Thailand	.Ecuador	.Romania		.Ecuador	.Portugal		.Jordan
.Denmark	.Mexico	.Turkey	.England	.Russia		.England	.Romania		.Morocco
.DRC	.Morocco	.UAE	.Gambia	.SaudiArabia		.Germany	.Russia		.Oman
.Ecuador	.Netherlands	.Uganda	.Germany	.Senegal		.Greece	.SaudiArabia		.Romania
.Egypt	.New Zealand	.Uruguay	.Greece	.South Africa		.HongKong	.Serbia		.Russia
.England	.Norway	.USA	.HongKong	.Sweden		.Hungary	.Singapore		.South Africa
.Finland	.Oman	.Vietnam	.Hungary	.Switzerland		.Iceland	.South Africa		.Sweden
	.Panama	.Wales	.India	.Taiwan		.India	.Spain		.Switzerland
	.Poland	.Zhejiang	.Israel	.Turkey			.Sweden		.Turkey
							.Switzerland		.Vietnam
									.Wales

FIGURE 1 | Nine significant node clusters within the COVID-19 genomic epidemiology network designed between December 2019 and July 2020. **(A)** COVID-19 genomic epidemiology network. (Upper panel) Simplified network, with the genomic clades and node clusters labeled. (Lower panel) Actual network, with the significant nodes (red) as determined by centrality analyses are shown. Nodes (dots) and edges (lines) are indicated. Node clusters are boxed and labeled. **(B)** List of countries identified by the significant nodes and classified according to node cluster.

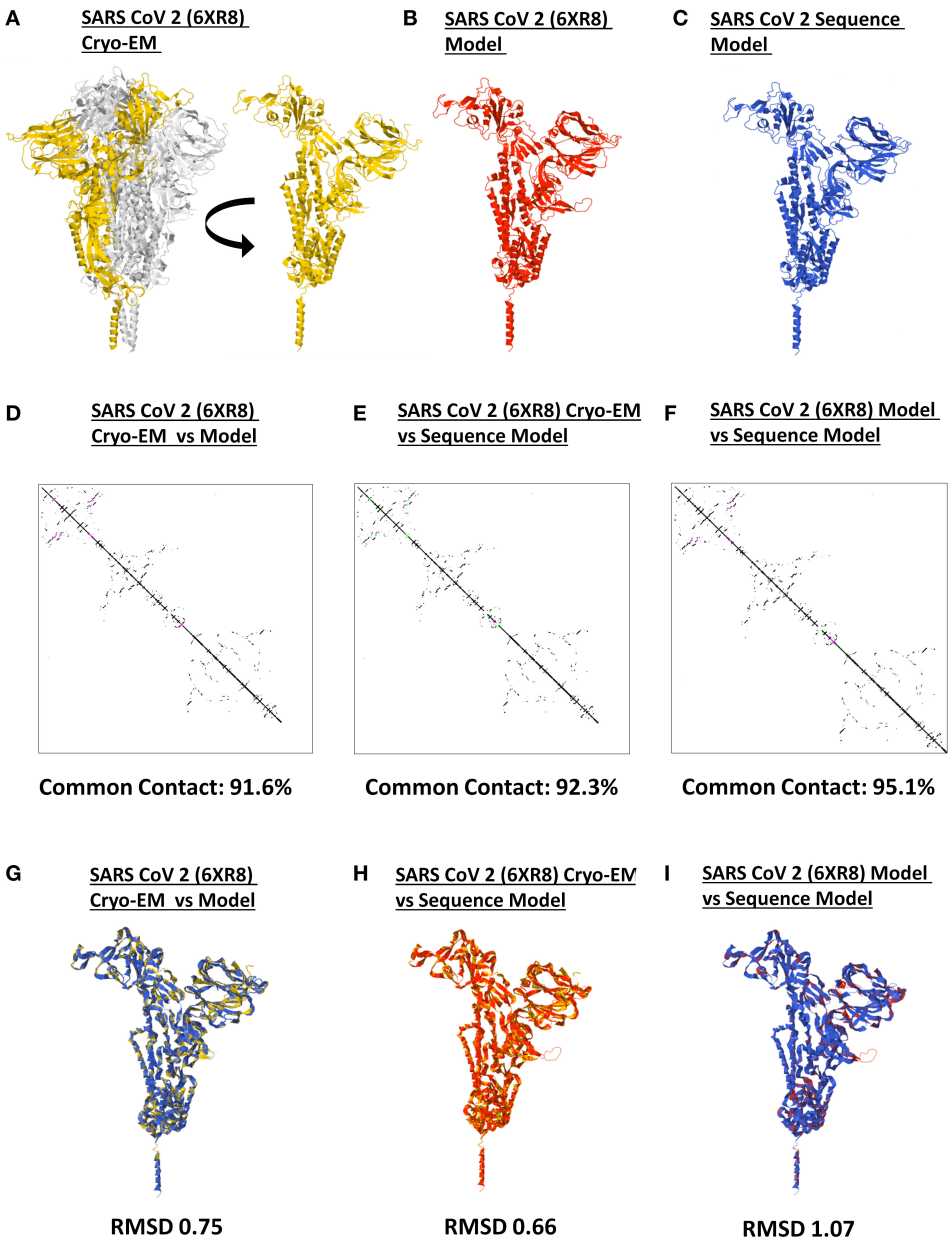


FIGURE 2 | Model quality assessment of a generated monomeric SARS-CoV-2 spike protein. Representative SARS-CoV-2 **(A)** 6XR8 cryo-EM, **(B)** 6XR8 model, and **(C)** sequence model of monomeric spike proteins are indicated. Contact maps of **(D)** 6XR8 cryo-EM and model, **(E)** 6XR8 cryo-EM and sequence model, and **(F)** 6XR8 model and sequence models are shown. The common contact of the protein structures being compared is labeled below. Superimposition between **(G)** 6XR8 cryo-EM and model, **(H)** 6XR8 cryo-EM and sequence model, and **(I)** 6XR8 model and sequence models are presented. RMSD scores of the superimposed protein structures are indicated below. SARS CoV 2 6XR8 cryo-EM (yellow), 6XR8 model (red), and sequence model (royal blue) are indicated.

the study. Interestingly, we found nine SARS2 FLC structural patterns (**Figures 3B–J**, *left panel*), which coincidentally match with the nine node clusters identified earlier (**Figure 1A**). This insinuates that the SARS2 FLC structural pattern identified in each node cluster is a unique structural feature for the node cluster. However, we emphasize that the SARS2 FLC might not be the only factor determining the nine node clusters. In this regard and as possible future works, additional experimental evidence is needed to further prove the presence of the nine SARS2 FLC structural patterns from the nine nodal clusters, and, equally important, it would be interesting to likewise determine other factors that may explain the presence of the nine node clusters. Subsequently, we observed that no structural changes occurred in the SARS2 RBD (**Figures 3B–J**, *right panel*). In all the superimpositions made, no significant structural changes (RMSD $<$ 1.0; Tm align $>$ 0.96) occurred between

superimposed SARS2 models (**Figures 3B–J**, *lower panel*), which is consistent with SARS2 maintaining its genomic integrity across propagation (34).

It was previously reported that the SARS2 FLC naturally undergoes polymorphisms, which in-turn affects viral transmissibility and tropism (39). In this regard, we suspect that the putative nine SARS2 FLC structural patterns are a product of natural polymorphism and, similarly, finding one of the SARS2 FLC structural patterns in one of the node clusters identified could suggest that certain countries (or continents) with overlapping node clusters may have varying levels of viral transmissibility and virulence (33, 34, 39). Since cleavage of the SARS2 FLC is a prerequisite for pathogenesis (17–19), we think that cleavage among the nine SARS2 FLC structural patterns may likewise vary (possibly depending on how exposed the FLC is), which in turn, could directly affect viral transmissibility. Additionally, with regards to host tropism, there seems to be no noticeable structural change in the SARS2 RBD, insinuating that host tropism is unchanged. This indicates that, regardless of any structural variations in SARS2 FLC, host tropism will not be consistently affected by genomic integrity (34). However, it is unclear whether the absence of SARS2 FLC (particularly ^{681}PRRA^{684}) would affect SARS2 RBD.

SARS2 RBD Residues Did Not Change in the Absence of the Unique ^{681}PRRA^{684} Sequence

SARS2 has been reported to infect multiple species as well as humans due to variations in ACE2 receptors across species (40), which emphasizes the potential significance of the SARS2 RBD with regards to host tropism. Similarly, SARS2 FLC was found to likewise affect host tropism (17–19). This may suggest that SARS2 FLC (particularly ^{681}PRRA^{684}) could affect SARS2 RBD. To establish the possible structural influence of the unique ^{681}PRRA^{684} amino acid sequence on SARS2 RBD structural orientation, we generated mutant SARS2 models with the unique ^{681}PRRA^{684} amino acid sequence deleted in all nine SARS2 FLC structural patterns and, subsequently, superimposed each mutant to the original model for comparison. This study undertook a side-by-side comparison of an original (*left panel*) and mutant (*right panel*) SARS2 model with a focus on SARS2 RBD (box dash lines) and FLC (box solid lines) structural changes (**Figure 4A**). As expected, in the absence of the ^{681}PRRA^{684} amino acid sequence we observed structural variations in the SARS2 FLC (**Figures 4B–J**, *left panel*). Nevertheless, no significant structural changes were observed (RMSD $<$ 1.0; Tm align $>$ 0.82) between superimposed original and mutated SARS2 models (**Figures 4B–J**, *lower panel*). Most surprisingly, no structural variations were observed in the SARS2 RBD (**Figures 4B–J**, *right panel*). This would suggest that SARS2 FLC (particularly ^{681}PRRA^{684}) has no structural influence on SARS2 RBD, which is consistent with earlier works (41) that showed that SARS2 FLC may not be as critical as previously thought for the high fusion capacity of SARS2. However, it is worth mentioning that regions with high levels of the disorder typically do not have stable structures, and thus, would not have much of an effect on the remaining structured parts of the protein (20) consistent with our observations. Taken together, the lack of a stable structure in the FLC site and its surroundings may explain why no structural changes occurred within the SARS2 RBD after the removal of a unique ^{681}PRRA^{684} region. Nevertheless, we presume that regardless of the absence of any structural variations within the SARS2 RBD, viral pathogenesis was unaffected since one important factor that determines virulence is high-affinity virus receptor interaction and, likewise, takes into account multiple host factors (40). This may explain why SARS2 infection in humans varies among COVID-19 infected patients. Additional experiments are needed to further prove this point.

DISCUSSION

SARS2 FLC is a conserved structural feature that is crucial for viral entry to host cells (39, 42) and, more importantly, can influence viral pathogenesis and host tropism (17–19, 40). In addition, the SARS2 FLC was found to have a naturally occurring polymorphism that can affect both transmissibility and host tropism (39). Throughout this study, we attempted to show that the SARS2 FLC has structural orientation variations putatively associated with the SARS2 genomic distribution particularly between December 2019 and July 2020.

SARS2 genome has continued to mutate since its emergence in December 2019 and SARS2 was found to have a $>$7.23 actual mutation rate with genetic changes occurring every other week (33, 34). These mutational changes are made possible through host-dependent RNA editing associated with the APOBEC mechanism (43). Cluster infections have also been associated with SARS2 incubation period infection and, likewise, play an important role in the rapid evolution of COVID-19 transmission (44, 45). This highlights how quickly the SARS2 genome is changing and, similarly, may explain how multiple variants of the virus can evolve easily and spread worldwide (33, 34). Several of the SARS2 nucleotide changes are nonsynonymous, thus, amino acid changes likewise occur (33) that may result in protein structural changes among SARS2 viral proteins. In particular, several structural changes have been reported with regards to the SARS2 spike protein (39, 42, 46, 47). Considering that we observed nine SARS2 FLC structural patterns from nine node clusters distributed worldwide, we postulate that this observation is putatively correlated to mutational changes that occurred within the SARS2 spike genome during the timeframe studied which in-turn affected the resulting amino acid sequence and, subsequently, lead to structural changes that may affect virulence and tropism.

It is worth mentioning that COVID-19 symptoms vary in the human population and, similarly, animal species (40). SARS2 infection in the human population often affects the lower respiratory tract (48) and follows a distinguishable order of symptom onset with varying levels of severity (49–51). COVID-19 reinfection has been clinically observed (52–56) and we suspect it is associated with varying SARS2 variants. In this regard, we hypothesize that COVID-19 reinfection could potentially be linked to SARS2 FLC structural variations

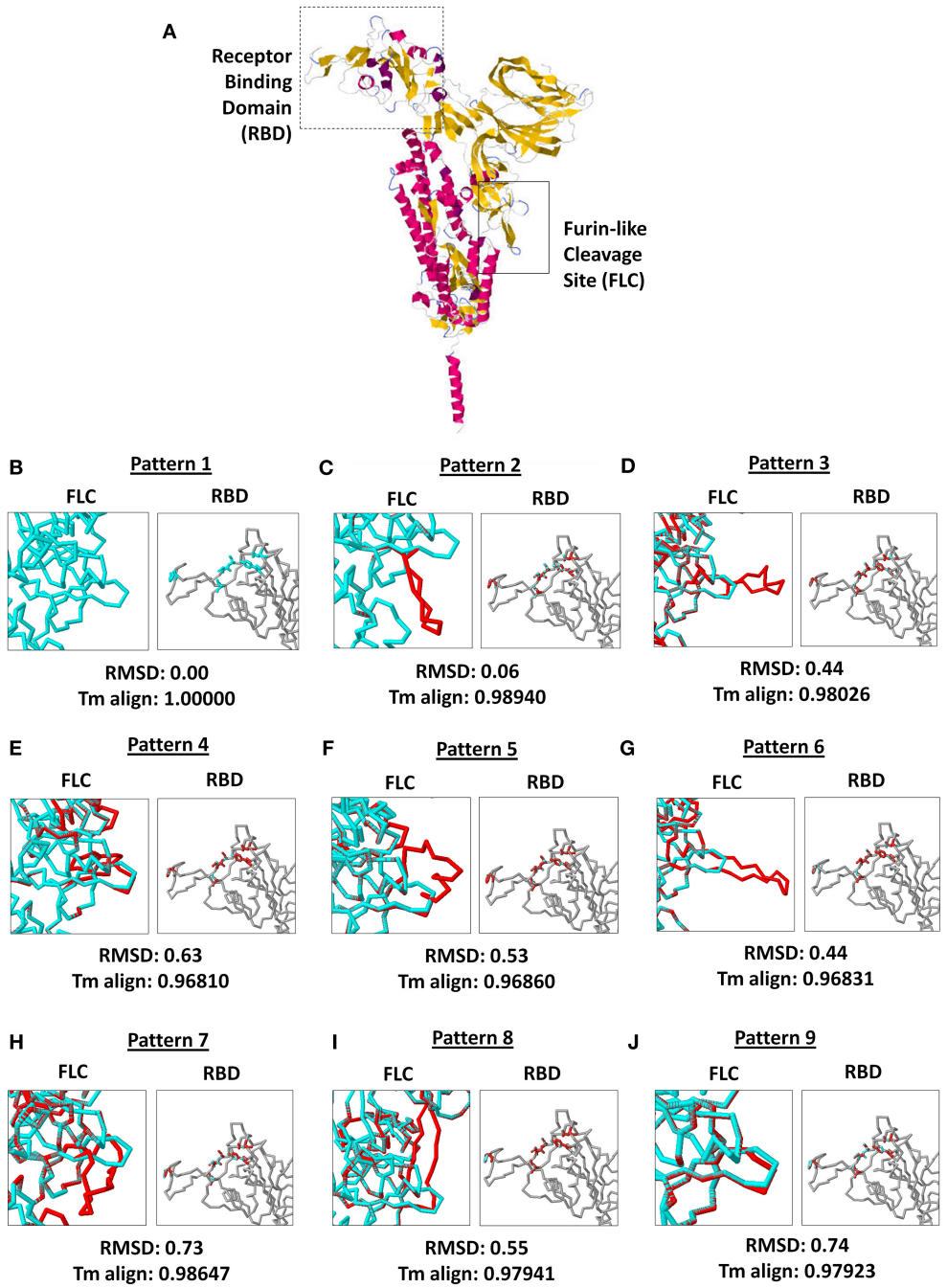


FIGURE 3 | Comparison of the 9 SARS-CoV-2 spike protein furin-like cleavage site structural patterns and corresponding receptor binding domains. **(A)** Representative monomeric SARS-CoV-2 spike protein model with the receptor binding domain (boxed dash lines) and furin-like cleavage site (boxed solid lines) indicated. **(B–J)** Superimposed spike protein models showing the nine structural patterns of the furin-like cleavage site (*left panel*) and receptor binding domain (*right panel*). Pattern 1 SARS-CoV-2 spike protein model (cyan) and the eight other structural patterns (red) are shown. RMSD scores and Tm align values normalized to Pattern 1 SARS-CoV-2 spike protein model are indicated below.

since SARS2 FLC affects viral pathogenesis, tropism, and transmissibility. Admittedly, additional experiments are needed to further prove this hypothesis.

In summary, we propose that between December 2019 and July 2020, nine SARS2 FLC structural patterns could putatively correspond to the nine node clusters found within the COVID-19 genomic epidemiology network. Similarly, we associated this with the rapid evolution of the SARS2 genome. We observed that either in the presence or absence of the unique ^{681}PRRA^{684} amino acid sequence no

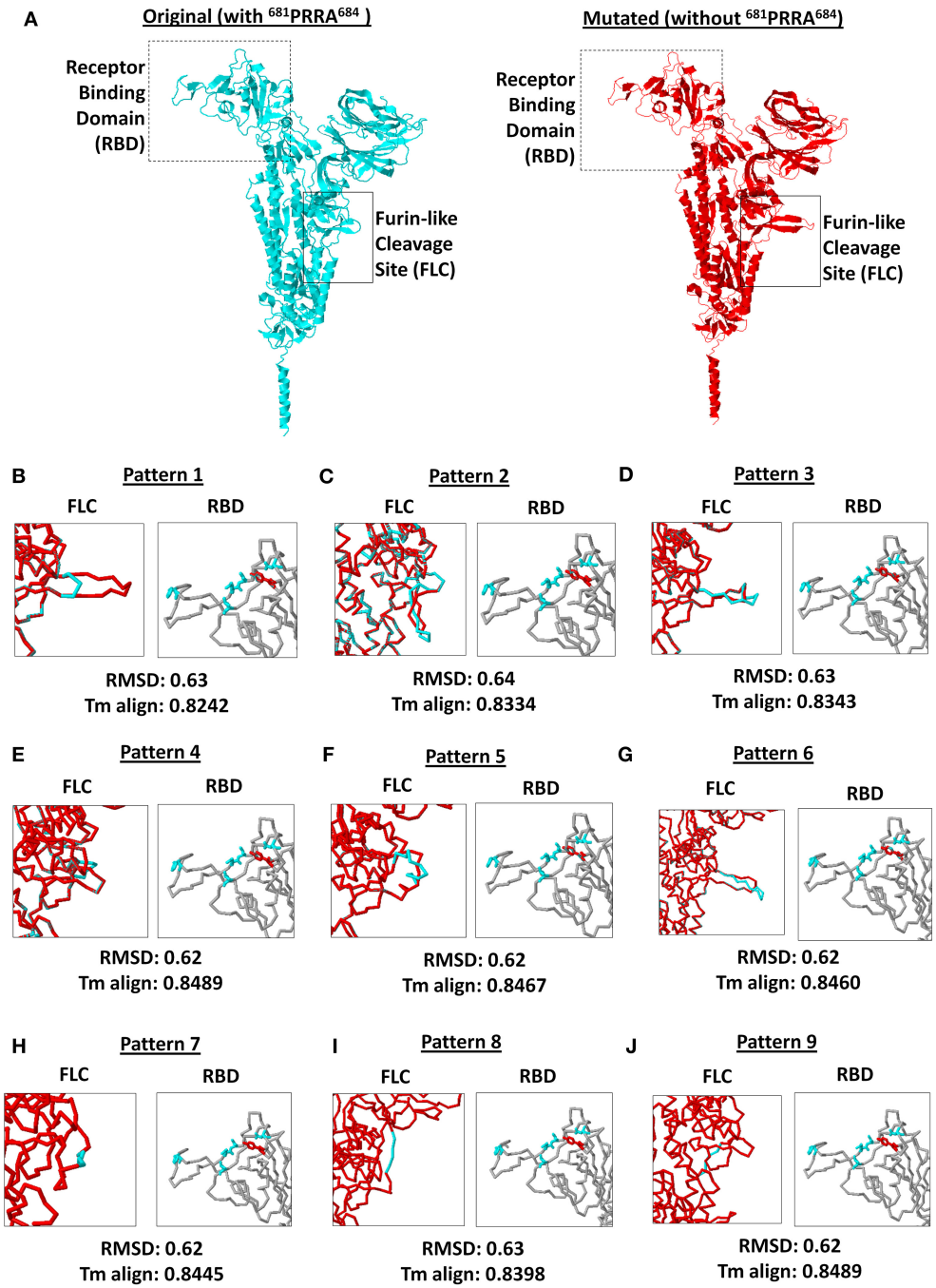


FIGURE 4 | Comparison between original (with ^{681}PRRA^{684}) and mutated (without ^{681}PRRA^{684}) forms of the 9 SARS-CoV-2 spike protein furin-like cleavage site structural patterns and corresponding receptor binding domains. **(A)** Original (cyan) and mutated (red) representative monomeric SARS-CoV-2 spike proteins are shown. Receptor binding domain (boxed dash lines) and furin-like cleavage site (boxed solid lines) indicated. **(B–J)** Superimposed spike protein models showing the 9 structural patterns of the furin-like cleavage site (*left panel*) and receptor binding domain (*right panel*). Original (cyan) and mutated (red) SARS-CoV-2 spike protein furin-like cleavage site structural patterns and corresponding receptor binding domains are shown. RMSD scores and Tm align values normalized to the original SARS-CoV-2 spike protein model are indicated below.

structural changes occurred within the SARS2 RBD, which we believe could mean that the SARS2 FLC has no structural influence on SARS2 RBD and may explain why host tropism was maintained.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

MC and KI conceptualized the idea, provided feedback, helped in both structural and network analyses, and wrote the paper. MU, RI, and TH generated the protein models and analyzed the structural changes. YM, KY, NK, KW, and KB designed and analyzed the network. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by JSPS KAKENHI Grant Numbers 19K10078 and 19K10097 and Uemura Fund, Dental Research Center, Nihon University School of Dentistry.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2021.613412/full#supplementary-material

REFERENCES

1. Hamre D, Procknow JJ. A new virus isolated from the human respiratory tract. *Proc Soc Exp Biol Med.* (1966) 121:190–3. doi: 10.3181/00379727-121-30734
2. Kapikian AZ, James HDJr, Kelly SJ, Dees JH, Turner HC, Mcintosh K, et al. Isolation from man of "avian infectious bronchitis virus-like" viruses (coronaviruses) similar to 229E virus, with some epidemiological observations. *J Infect Dis.* (1969) 119:282–90. doi: 10.1093/infdis/119.3.282
3. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med.* (2003) 348:1953–66. doi: 10.1056/NEJMoa030781
4. Fouchier RA, Hartwig NG, Bestebroer TM, Niemeyer B, De Jong JC, Simon JH, et al. A previously undescribed coronavirus associated with respiratory disease in humans. *Proc Natl Acad Sci USA.* (2004) 101:6212–6. doi: 10.1073/pnas.0400762101
5. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J Virol.* (2005) 79:884–95. doi: 10.1128/JVI.79.2.884-895.2005
6. Zaki AM, Van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med.* (2012) 367:1814–20. doi: 10.1056/NEJMoa1211721
7. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med.* (2020) 382:727–33. doi: 10.1056/NEJMoa2001017
8. Lu G, Wang Q, Gao GF. Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends Microbiol.* (2015) 23:468–78. doi: 10.1016/j.tim.2015.06.003
9. Millet JK, Whittaker GR. Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. *Virus Res.* (2015) 202:120–34. doi: 10.1016/j.virusres.2014.11.021
10. Hulswit RJ, De Haan CA, Bosch BJ. Coronavirus spike protein and tropism changes. *Adv Virus Res.* (2016) 96:29–57. doi: 10.1016/bs.aivir.2016.08.004
11. Li F. Structure, function, and evolution of coronavirus spike proteins. *Annu Rev Virol.* (2016) 3:237–61. doi: 10.1146/annurev-virology-110615-042301
12. Moore JB, June CH. Cytokine release syndrome in severe COVID-19. *Science.* (2020) 368:473–4. doi: 10.1126/science.abb8925
13. Tay MZ, Poh CM, Renia L, Macary PA, Ng LFP. The trinity of COVID-19: immunity, inflammation and intervention. *Nat Rev Immunol.* (2020) 20:363–74. doi: 10.1038/s41577-020-0311-8
14. Wang N, Shang J, Jiang S, Du L. Subunit vaccines against emerging pathogenic human coronaviruses. *Front Microbiol.* (2020) 11:298. doi: 10.3389/fmicb.2020.00298
15. Cai Y, Zhang J, Xiao T, Peng H, Sterling SM, Walsh RMJr, et al. Distinct conformational states of SARS-CoV-2 spike protein. *Science.* (2020) 369:1586–92. doi: 10.1126/science.abd4251
16. Ord M, Faustova I, Loog M. The sequence at Spike S1/S2 site enables cleavage by furin and phospho-regulation in SARS-CoV2 but not in SARS-CoV1 or MERS-CoV. *Sci Rep.* (2020) 10:16944. doi: 10.1038/s41598-020-74101-0
17. Nao N, Yamagishi J, Miyamoto H, Igarashi M, Manzoor R, Ohnuma A, et al. Genetic predisposition to acquire a polybasic cleavage site for highly pathogenic avian influenza virus hemagglutinin. *MBio.* (2017) 8:e02298–16. doi: 10.1128/mBio.02298-16
18. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. *Nat Med.* (2020) 26:450–2. doi: 10.1038/s41591-020-0820-9
19. Coutard B, Valle C, De Lamballerie X, Canard B, Seidah NG, Decroly E. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res.* (2020) 176:104742. doi: 10.1016/j.antiviral.2020.104742
20. Giri R, Bhardwaj T, Shegane M, Gehi BR, Kumar P, Gadhave K, et al. Understanding COVID-19 via comparative analysis of dark proteomes of SARS-CoV-2, human SARS and bat SARS-like coronaviruses. *Cell Mol Life Sci.* (2020) 78:1655–88. doi: 10.1007/s00018-020-03603-x
21. Jaimes J, Millet J, Whittaker G. Proteolytic cleavage of the SARS-CoV-2 spike protein and the role of the novel S1/S2 Site. *SSRN.* (2020) 3581359. doi: 10.2139/ssrn.3581359
22. Kortemme T, Morozov AV, Baker D. An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein-protein complexes. *J Mol Biol.* (2003) 326:1239–59. doi: 10.1016/S0022-2836(03)00021-4
23. Chen SC, Bahar I. Mining frequent patterns in protein structures: a study of protease families. *Bioinformatics.* (2004) 20 (Suppl. 1):i77–85. doi: 10.1093/bioinformatics/bth912
24. Seyran M, Takayama K, Uversky VN, Lundstrom K, Palu G, Sherchan SP, et al. The structural basis of accelerated host cell entry by SARS-CoV-2dagger. *FEBS J.* (2020). doi: 10.1111/febs.15651
25. Gilman A, Arkin AP. Genetic "code": representations and dynamical models of genetic components and networks. *Annu Rev Genomics Hum Genet.* (2002) 3:341–69. doi: 10.1146/annurev.genom.3.030502.111004
26. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* (2003) 13:2498–504. doi: 10.1101/gr.1239303
27. Koschutzki D, Schreiber F. Centrality analysis methods for biological networks and their application to gene regulatory networks. *Gene Regul Syst Bio.* (2008) 2:193–201. doi: 10.4137/GRSB.S702
28. Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* (2005) 33:2302–9. doi: 10.1093/nar/gki524
29. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc.* (2009) 4:363–71. doi: 10.1038/nprot.2009.2
30. Herraez A. Biomolecules in the computer: Jmol to the rescue. *Biochem Mol Biol Educ.* (2006) 34:255–61. doi: 10.1002/bmb.2006.494034042644
31. Vehlow C, Stehr H, Winkelmann M, Duarte JM, Petzold L, Dinse J, et al. CMView: interactive contact map visualization and analysis. *Bioinformatics.* (2011) 27:1573–4. doi: 10.1093/bioinformatics/btr163

32. Holm L, Sander C. Mapping the protein universe. *Science.* (1996) 273:595–603. doi: 10.1126/science.273.5275.595
33. Day T, Gandon S, Lion S, Otto SP. On the evolutionary epidemiology of SARS-CoV-2. *Curr Biol.* (2020) 30:R849–R857. doi: 10.1016/j.cub.2020.06.031
34. Mercatelli D, Giorgi FM. Geographic and genomic distribution of SARS-CoV-2 mutations. *Front Microbiol.* (2020) 11:1800. doi: 10.3389/fmicb.2020.01800
35. Przulj N, Malod-Dognin N. NETWORK ANALYSIS. Network analytics in the age of big data. *Science.* (2016) 353:123–4. doi: 10.1126/science.aah3449
36. Wuchty S, Stadler PF. Centers of complex networks. *J Theor Biol.* (2003) 223:45–53. doi: 10.1016/S0022-5193(03)00071-7
37. Berman HM, Burley SK, Chiu W, Sali A, Adzhubei A, Bourne PE, et al. Outcome of a workshop on archiving structural models of biological macromolecules. *Structure.* (2006) 14:1211–7. doi: 10.1016/j.str.2006.06.005
38. Hevener KE, Zhao W, Ball DM, Babaoglu K, Qi J, White SW, et al. Validation of molecular docking programs for virtual screening against dihydropteroate synthase. *J Chem Inf Model.* (2009) 49:444–60. doi: 10.1021/ci800293n
39. Xing Y, Li X, Gao X, Dong Q. Natural Polymorphisms Are Present in the Furin Cleavage Site of the SARS-CoV-2 Spike Glycoprotein. *Front Genet.* (2020) 11:783. doi: 10.3389/fgene.2020.00783
40. Sarkar J, Guha R. Infectivity, virulence, pathogenicity, host-pathogen interactions of SARS and SARS-CoV-2 in experimental animals: a systematic review. *Vet Res Commun.* (2020) 44:101–10. doi: 10.1007/s11259-020-09778-9
41. Xia S, Lan Q, Su S, Wang X, Xu W, Liu Z, et al. The role of furin cleavage site in SARS-CoV-2 spike protein-mediated membrane fusion in the presence or absence of trypsin. *Signal Transduct Target Ther.* (2020) 5:92. doi: 10.1038/s41392-020-0184-0
42. Walls AC, Park YJ, Tortorici MA, Wall A, Mcguire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell.* (2020) 181:281–92 e286. doi: 10.1016/j.cell.2020.02.058
43. Di Giorgio S, Martignano F, Torcia MG, Mattiuz G, Conticello SG. Evidence for host-dependent RNA editing in the transcriptome of SARS-CoV-2. *Sci Adv*. (2020) 6:eabb5813. doi: 10.1126/sciadv.abb5813
44. Gao Y, Shi C, Chen Y, Shi P, Liu J, Xiao Y, et al. A cluster of the Corona Virus Disease 2019 caused by incubation period transmission in Wuxi, China. *J Infect.* (2020) 80:666–70. doi: 10.1016/j.jinf.2020.03.042
45. Liu T, Gong D, Xiao J, Hu J, He G, Rong Z, et al. Cluster infections play important roles in the rapid evolution of COVID-19 transmission: a systematic review. *Int J Infect Dis.* (2020) 99:374–80. doi: 10.1016/j.ijid.2020.07.073
46. Davidson AD, Williamson MK, Lewis S, Shoemark D, Carroll MW, Heesom KJ, et al. Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Med.* (2020) 12:68. doi: 10.1186/s13073-020-00763-0
47. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking Changes in SARS-CoV-2 spike: evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell*. (2020) 182:812–27 e819. doi: 10.1016/j.cell.2020.06.043
48. Liu SL, Saif L. Emerging *Viruses* without Borders: The Wuhan coronavirus. *Viruses.* (2020) 12:130. doi: 10.3390/v12020130
49. Grant MC, Geoghegan L, Arbyn M, Mohammed Z, Mcguinness L, Clarke EL, et al. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): a systematic review and meta-analysis of 148 studies from 9 countries. *PLoS ONE.* (2020) 15:e0234765. doi: 10.1371/journal.pone.0234765
50. Koutsakos M, Kedzierska K. A race to determine what drives COVID-19 severity. *Nature.* (2020) 583:366–8. doi: 10.1038/d41586-020-01915-3
51. Larsen JR, Martin MR, Martin JD, Kuhn P, Hicks JB. Modeling the onset of symptoms of COVID-19. *Front Public Health.* (2020) 8:473. doi: 10.3389/fpubh.2020.00473
52. Bonifacio LP, Pereira APS, Araujo D, Balbao V, Fonseca B, Passos ADC, et al. Are SARS-CoV-2 reinfection and Covid-19 recurrence possible? A case report from Brazil. *Rev Soc Bras Med Trop.* (2020) 53:e20200619. doi: 10.1590/0037-8682-0619-2020
53. Gousseff M, Penot P, Gallay L, Batisse D, Benech N, Bouiller K, et al. Clinical recurrences of COVID-19 symptoms after recovery: Viral relapse, reinfection or inflammatory rebound? *J Infect.* (2020) 81:816–46. doi: 10.1016/j.jinf.2020.06.073
54. Madan M, Kunal S. COVID-19 reinfection or relapse: an intriguing dilemma. *Clin Rheumatol.* (2020) 39:3189. doi: 10.1007/s10067-020-05427-3
55. Parry J. Covid-19: Hong Kong scientists report first confirmed case of reinfection. *BMJ.* (2020) 370:m3340. doi: 10.1136/bmj.m3340
56. To KK, Hung IF, Chan KH, Yuan S, To WK, Tsang DN, et al. Serum antibody profile of a patient with COVID-19 reinfection. *Clin Infect Dis.* (2020) ciaa1368. doi: 10.1093/cid/ciaa1368

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.

frontiers in Public Health

ORIGINAL RESEARCH
published: 13 April 2021
doi: 10.3389/fpubh.2021.586670

A Vulnerability Analysis for the Management of and Response to the COVID-19 Epidemic in the Second Most Populous State in Brazil

Igor Silva Campos[1], *Vinícius Ferreira Aratani*[1], *Karina Baltor Cabral*[1], *Jean Ezequiel Limongi*[2] *and Stefan Vilges de Oliveira*[3]*

[1] Undergraduate Medical, Faculty of Medicine, Federal University of Uberlândia, Uberlândia, Brazil, [2] Undergraduate Course in Collective Health, Institute of Geography, Federal University of Uberlândia, Uberlândia, Brazil, [3] Department of Collective Health, Faculty of Medicine, Federal University of Uberlândia, Uberlândia, Brazil

The COVID-19 pandemic has the potential to affect all individuals, however in a heterogeneous way. In this sense, identifying specificities of each location is essential to minimize the damage caused by the disease. Therefore, the aim of this research was to assess the vulnerability of 853 municipalities in the second most populous state in Brazil, Minas Gerais (MG), in order to direct public policies. An epidemiological study was carried out based on Multi-Criteria Decision Analysis (MCDA) using indicators with some relation to the process of illness and death caused by COVID-19. The indicators were selected by a literature search and categorized into: demographic, social, economic, health infrastructure, population at risk and epidemiological. The variables were collected in Brazilian government databases at the municipal level and evaluated according to MCDA, through the Program to Support Decision Making based on Indicators (PRADIN). Based on this approach, the study performed simulations by category of indicators and a general simulation that allowed to divide the municipalities into groups of 1–5, with 1 being the least vulnerable and 5 being the most vulnerable. The groupings of municipalities were exposed in their respective mesoregions of MG in a thematic map, using the software Tabwin 32. The results revealed that the mesoregion of Norte de Minas stands out with more than 40% of its municipalities belonging to group 5, according to economic, social and health infrastructure indicators. Similarly, the Jequitinhonha mesoregion exhibited almost 60% of the municipalities in this group for economic and health infrastructure indicators. For demographic and epidemiological criteria, the Metropolitana de Belo Horizonte was the most vulnerable mesoregion, with 42.9 and 26.7% of the municipalities in group 5, respectively. Considering the presence of a population at risk, Zona da Mata reported 42.3% of the municipalities in the most vulnerable group. In the joint analysis of data, the Jequitinhonha, Vale do Mucuri and Vale do Rio Doce mesoregions were the most vulnerable in the state of MG. Thus, through the outlined profile, the present study proved how socioeconomic diversity affects the vulnerability of the municipalities to face COVID-19 outbreak, highlighting the need for interventions directed to each reality.

Keywords: COVID-19, social vulnerability, disease outbreaks, epidemics, policy formulation, health policy

OPEN ACCESS

Edited by:
Salvatore Flavio Pileggi, University of Technology Sydney, Australia

Reviewed by:
Lei Gong, Anhui Center of Disease Control and Prevention, China
Francesco Bogliacino, National University of Colombia, Colombia

***Correspondence:**
Stefan Vilges de Oliveira
stefanbio@yahoo.com.br

Specialty section:
This article was submitted to Infectious Diseases - Surveillance, Prevention and Treatment, a section of the journal Frontiers in Public Health

Received: *23 July 2020*
Accepted: *15 March 2021*
Published: *13 April 2021*

Citation:
Campos IS, Aratani VF, Cabral KB, Limongi JE and Oliveira SVd (2021) A Vulnerability Analysis for the Management of and Response to the COVID-19 Epidemic in the Second Most Populous State in Brazil. Front. Public Health 9:586670. doi: 10.3389/fpubh.2021.586670

INTRODUCTION

In late December 2019, hospitals in Wuhan, China, identified numerous patients with pneumonia of unknown cause (1). After investigating the possible etiologic agent involved, on January 7, 2020, Chinese scientists isolated a new type of coronavirus from an individual and, therefore, were able to sequence its genome (2).

The SARS-CoV-2 or 2019-nCoV virus is the causative agent of the clinical syndrome known as COVID-19 (Coronavirus 19 disease) (3). Although SARS-CoV-2 belongs to the same gender as the viruses responsible for Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS), the new coronavirus appears to be related to mild infections disorders but with a high rate of transmissibility (3–5). Considering the high levels of transmission, on March 11, the World Health Organization (WHO) characterized COVID-19 as a pandemic due to the rapid spread across countries, such as Italy, Spain and, later, United States, that currently has the highest number of cases of the novel coronavirus disease (6).

In Brazil, on February 26, 2020, the first case of COVID-19 was confirmed in the state of São Paulo and the first death on March 17, in the same state. During the months of April, May and June, the number of cases and deaths increased exponentially and, then, on June 20, 2020, Brazil was the second country in the world with the highest number of confirmed cases, more than 1 million, and also the second country with the most confirmed deaths, about 50 thousand (6).

In this context, the state of Minas Gerais, the second most populous state in the country, initially stood out for presenting an apparently controlled situation. While neighboring states in the southeastern region accumulated more than 300,000 cases and almost 22,000 deaths by COVID-19, Minas Gerais was one of the states with the least confirmed cases, approximately 27,000 (6). However, recent researches pointed to a possible underreporting scenario owing to the unprecedented increase in deaths from causes clinically similar to COVID-19, including SARS, respiratory failure and pneumonia, and due to the low number of tests performed by the state in comparison with the others, according to data obtained from the Minas Gerais Department of Health (7).

From this perspective, the high number of cases in neighboring states and underreporting in Minas Gerais rendered the region extremely susceptible to the increase in the number of cases of COVID-19, with the occurrence and diffusion of new cases. In this situation, delimiting and defining the main regions of vulnerability in the state is essential to guide the population, managers, public policies and government healthcare workforces. The concept of vulnerability, in this case, considers aspects that can measure whether the resources designated to the protection of people are available or in need. Thus, a complex set of indicators may determine a higher or lower vulnerability (8, 9).

In this work, the indicators, selected based on the literature to identify the vulnerability of the cities of Minas Gerais, describe physical, social and individual characteristics that enable to assess and qualify the regions with greater difficulty in managing the pandemic. These indicators, divided into demographic, social, economic, health infrastructure, population at risk and epidemiological, are directly related to the increase in illness and death due to the life-threatening condition.

Therefore, the aim of this study was to identify areas of higher vulnerability in the state of Minas Gerais and, from there, assess the region in a segmented and directed way (mesoregions of Minas Gerais) in order to contribute to the elaboration of public policies for prevention and combat of COVID-19.

METHODS

Study Design

This is an epidemiological study to assess vulnerability based on Multi-Criteria Decision Analysis (MCDA) (10). The indicators were selected to allow the assessment of the state of Minas Gerais according to the vulnerability to COVID-19 at the municipal level.

Study Area

The study assesses the state of Minas Gerais (MG), one of the 27 federative units in Brazil, located in the southeastern region of the country, and characterized as the second most populous state and the fourth with the largest territorial extension. It borders the states of São Paulo (south and southwest), Mato Grosso do Sul (west), Goiás (northwest), Bahia (north and northeast), Espírito Santo (east), and Rio de Janeiro (southeast). According to data provided by the Brazilian Institute of Geography and Statistics (IBGE), MG has an estimated population of 21,168,791, demographic density of 33.41 inhabitants per km² (11, 12) in a territory of 586,521,123 km², which corresponds, approximately, to the sum of the territorial areas of Spain and Portugal. The population is predominantly urban, 85.3% (12), composed of 22.3% of people from 0 to 15 years old, 69.3% from 15 to 64 years old and 8.1% above 65 years old. MG is the state with the third highest gross domestic product (GDP) in Brazil and has Human Development Index (HDI) of 0.731 (11, 12).

The state of MG, whose capital is located in the municipality of Belo Horizonte, has 853 municipalities distributed in 12 mesoregions: Noroeste de Minas, Norte de Minas, Jequitinhonha, Vale do Mucuri, Triângulo Mineiro and Alto Paranaíba, Central Mineira, Metropolitana de Belo Horizonte, Vale do Rio Doce, Oeste de Minas, Sul e Sudoeste de Minas, Campos das Vertentes, and Zona da Mata. Created in 1990 by IBGE, these mesoregions are guided by regional particularities related to social and administrative processes (13) (**Figure 1**).

Data Analysis

The indicators used in this analysis were categorized into: demographic, social, economic, health infrastructure, population at risk, and epidemiological. In the process of selection of the indicators, articles that describe or analyze factors that could interfere directly in the increase of damage caused by COVID-19 in the population were selected based on a literature search in Scielo and PubMed databases. Then, among 18 articles selected (14–31), 23 indicators were identified, categorized and used in this paper. Other indicators that could potentially have some relation to the disease and death caused by coronavirus infection

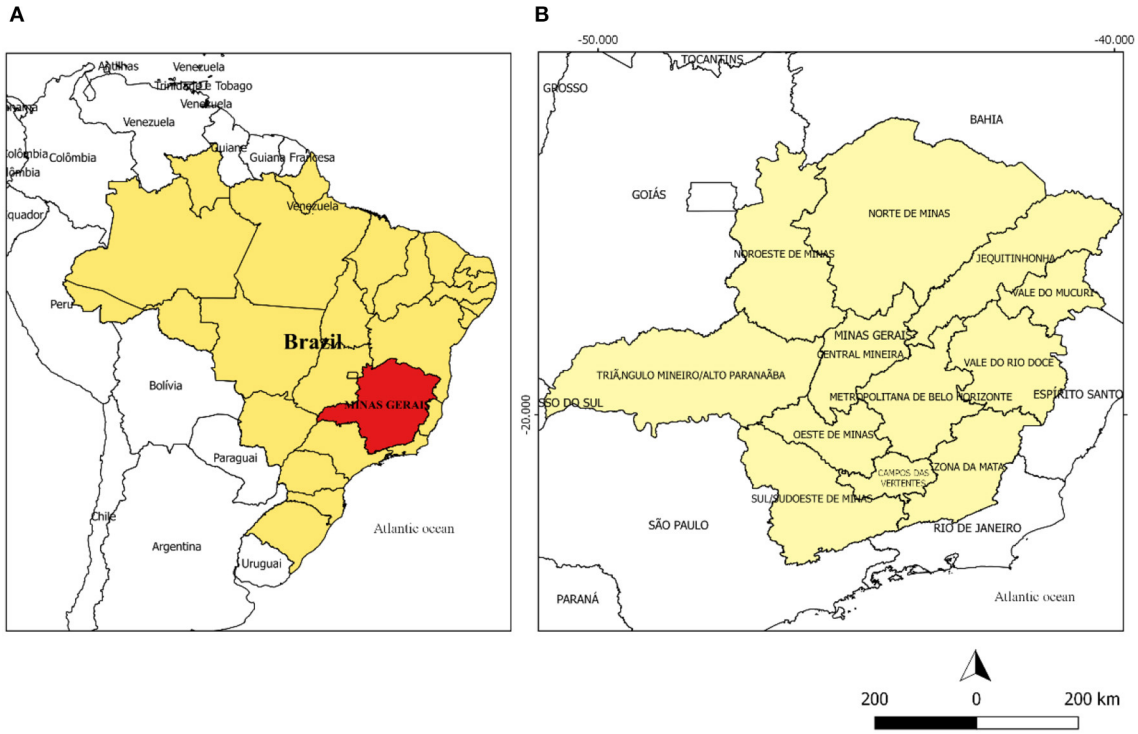


FIGURE 1 | (A) Map of Brazil highlighting the state of Minas Gerais in red. **(B)** Map of the state of Minas Gerais divided into its mesoregions.

were not used because the official data available were limited at the municipalities level.

After selecting the indicators, these variables were collected from the databases of the Brazilian Institute of Geography and Statistics ("IBGE") (https://www.ibge.gov.br/), Atlas of Human Development of Brazil 2013 (http://hdr.undp.org/en/content/new-atlas-human-development-brazil), from the National Supplementary Health Agency (NSHA) (http://www.ans.gov.br/), Mortality Information System (MIS) (http://www2.datasus.gov.br/DATASUS/index.php?area=060701), the National Registry of Health Establishments (NRHE) (http://www2.datasus.gov.br/DATASUS/index.php?area=0204&id=6906) and from the State Secretariat of Health of Minas Gerais (https://www.saude.mg.gov.br/). Then, 23 indicators remained, grouped according to categories (**Table 1**). The tabulated absolute data were adjusted to the respective values in percentage, incidence or mortality per 1,000 inhabitants, using information from the population projection of Minas Gerais for the year 2019 (11).

Statistical Analysis

The vulnerability of the municipalities of Minas Gerais to COVID-19 was evaluated by a complex set of factors, herein represented by the indicators (32). For the joint analysis of these factors, an instrument known as Multi-criteria Decision Analysis (MCDA) was applied through the Program to Support Decision Making Based on Indicators (PRADIN) software, in order to represent the vulnerability of these municipalities (33).

The use of this approach in Health Surveillance is recommended by the Ministry of Health of Brazil as an analysis methodology for the management process of the Brazilian Unified Health System (SUS) (34). Importantly, the method has already been used to classify areas of vulnerability for *Trypanosoma cruzi* (35) and to map the vulnerability to hantavirus infections in Brazil (36). Likewise, in the United States, the methodology was used to analyze the risk of Chagas disease in Texas (37).

In order to facilitate the understanding of the statistical analysis performed, we divided the session into: **Decision-making**, seeking to explain the tools used in the analysis and choice of indicators; **Simulations**, explaining what the simulations are and how they work; **Categorization**, showing how we organize the indicators in the study; **Weights**, explaining the distribution of the weights of the indicators; **Organization of results**, explaining how the data were grouped and placed on thematic maps.

Decision-Making

Multi-criteria Decision Analysis (MCDA) can be conceptualized as a set of strategies to assist in making decisions about a complex problem. MCDA uses and evaluates several criteria and perspectives aimed at identifying priorities, better health management and planning solutions (10). The technique allows the decision to occur based on what the decision-makers consider to be relevant to the situation (10). Thus, applying the concept of this methodology (MCDA) to the present study, different

TABLE 1 | Indicators by category (demographic, social, economic, health infrastructure, population at risk, epidemiological) used in the Multi-criteria Decision Analysis (MCDA) to assess the vulnerability of the municipalities of Minas Gerais to COVID-19.

Category	Indicators	Description	Application in the context of COVID-19	Source	References
Demographic	Percentage of the population living in an urban area	Percentage of the population residing in a situation of urban domicile in the municipalities of Minas Gerais.	There is a correlation between the increase in population density, proportion of built area, industrial concentration and other parameters associated with urbanization and increased morbidity by COVID-19. Additionally, air pollution, which is prevalent in locations with high rates of urbanization, contributes to the probability of infections.	2010 IBGE population census	(14, 15)
	Demographic density	Demographic density of the territorial unit (Inhabitant per square kilometer)	There is evidence that population density affects the number of COVID-19 daily cases.	2010 IBGE population census	(14, 16, 17)
Social	Percentage of inadequate sanitation	Households without basic sanitation condition, that is, they were not connected to the general water supply network, to sanitary sewage and had no access to garbage collection.	The virus that causes COVID-19 has already been detected in sewage samples in several countries and in the feces of infected patients, hence demonstrating the need for proper waste treatment.	2010 IBGE population census	(15, 18, 19)
	Illiteracy percentage	Rate of people aged 15 and over who cannot read or write	Important factor of social vulnerability, especially considering that one of the bases for combating the disease is information. Studies also indicate a higher prevalence of certain comorbidities in people with low levels of education.	2010 IBGE population census	(20, 21)
	Gini index	It measures the degree of inequality that exists in the distribution of individuals according to per capita household income.	Studies indicate that the Gini index can be extremely useful to measure the exposure-disease relationship.	Atlas of human development in Brazil 2010 (http://hdr.undp.org/en/content/new-atlas-human-development-brazil)	(22, 23)
	Municipal human development index	Geometric mean of the indexes related to income (per capita income indicator), Education (geometric average of the school attendance sub-index, with 2/3 weight, and of the schooling sub-index, with 1/3 weight) and Longevity (obtained through the life expectancy at birth), with equal weights.	The HDI can allow the assessment of social vulnerability by measuring the level of development of each region from three essential factors for quality of life.	Atlas of human development in Brazil 2010 (http://hdr.undp.org/en/content/new-atlas-humandevelopment-brazil)	(24)
Economics	Percentage of the population with per capita monthly income of up to 70 reais (BRL) (equivalent to US$ 13)	Population considered extremely poor.	Poverty and unemployment, characteristics of a population with such a low monthly income, are social determinants directly related to higher mortality caused by COVID-19.	2010 IBGE population census	(17)
	Percentage of the population with health insurance	Percentage of people by municipality who have access to health insurance.	Considering the need to treat more severe cases in Intensive Care Units (ICU) and the low availability of beds due to high demand, access to a private health network becomes an important indicator of less vulnerability.	National supplementary health agency TabNet DataSUS (March/2020)	(25, 26)
	Gross domestic product per capita	Proportion between the wealth produced by a municipality and its number of inhabitants.	The concentration of financial resources facilitates the promotion of measures to contain the pandemic, such as increasing the number of tests in the population.	2010 IBGE population census	(19)
Health infrastructure	Number of beds per 1,000 inhabitants	Proportion of the number of hospitalization beds by municipality.	Due to the pandemic, a great increase in the demand for health services is expected, thus it is essential to identify the most vulnerable regions and optimize the use of services and dimension resources that will be necessary to strengthen the response capacity of the health system regionally and locally.	TabNet DataSUS - Brazilian national registry of health facilities (NRHE) - Physical resources - 2019	(25, 27)

(Continued)

TABLE 1 | Continued

Category	Indicators	Description	Application in the context of COVID-19	Source	References
	Number of respirators per 1,000 inhabitants	Proportion of the number of respirators by municipality.			
	Number of doctors per 1,000 inhabitants	Proportion of the number of doctors by municipality.		TabNet DataSUS - Brazilian national registry of health establishments (NRHE) - Human resources - 2019	
	Number of nurses per 1,000 inhabitants	Proportion of the number of nurses by municipality.			
	Number of rapid tests per 1,000 inhabitants	Proportion of the number of rapid tests for COVID-19 performed by municipality.	Although total population testing is impractical, a well-designed program is essential to determine the prevalence of COVID-19 in the general society, in specific subgroups (including healthcare workers) and at-risk groups.	Data provided by the state health department of Minas Gerais on 06/22/2020	(28)
	Number of molecular tests (RT-PCR) per 1,000 inhabitants	Proportion of the number of molecular tests (RT-PCR) performed by municipality.			
Population at risk	Percentage of the population aged 60 or over	Percentage of the resident population in the municipalities of Minas Gerais aged 60 or over.	According to the World Health Organization, the mortality rate caused by COVID-19 increases with older age, with higher mortality among people over 80 years old.	2010 IBGE population census	(29)
	Mortality from diseases of the respiratory system per 1,000 inhabitants	Deaths per residence - Chapter ICD-10: X. Diseases of the respiratory system in 2018.	Cancer, hypertension, diabetes, Chronic Obstructive Pulmonary Disease (COPD), heart and cerebrovascular diseases are major risk factors for patients with COVID-19. Thus, municipalities with numerous cases of these life-threatening conditions become more vulnerable.	TabNet DataSUS mortality information system - MIS - 2018	(29–31)
	Mortality from diabetes per 1,000 inhabitants	Deaths per residence - Chapter ICD-10: IV. Diabetes (E10–E14) in 2018.			
	Mortality from neoplasms per 1,000 inhabitants	Deaths per residence - Chapter ICD-10: II. Neoplasms (tumors) in 2018.			
	Mortality from diseases of the circulatory system per 1,000 inhabitants	Deaths per residence - Chapter ICD-10: IX. Circulatory system diseases in 2018.			
Epidemiological	Incidence of COVID-19	Proportion between new cases of COVID-19 of a municipality and its population.	These occurrence measures are essential to compose an overview of COVID-19 in the municipalities of Minas Gerais, in addition to informing the evolution of the infectious illness in the state, a fact that would not be achieved only with the exposure of the absolute data of cases and deaths from the disease.	Epidemiological bulletin of the secretary of health of Minas Gerais on 06/22/2020	(31)
	Mortality of COVID-19	Number of COVID-19 deaths per 1,000 inhabitants.			
	Lethality of COVID-19	Proportion between the number of deaths caused by COVID-19 and the population affected.			

indicators selected by the authors were evaluated to identify the most vulnerable areas of Minas Gerais to COVID-19 and, therefore, priorities for public policies to combat the pandemic.

MCDA is a qualitative and quantitative process (**Figure 2**). The qualitative phase corresponds to the choice of factors related to the analyzed problem, the indicators (**Table 1**). In turn, the quantitative phase refers to the techniques that are employed in the search for a multi-criteria solution, which have already been described and analyzed by the literature (10, 38). The choice of the technique depends on how the problem presents itself and how it relates to the qualitatively selected factors (10, 38).

The present work used the Preference Ranking Organization Method for Enrichment of Evaluations, the PROMETHEE technique, specifically variant 2, because this is an over-classification method that allowed the municipalities to be structured hierarchically, based on indicators related to the complex problem, in order to determine priority levels (38). The technique allows simulations to be performed (**Table 2**) by comparing the municipalities for each defined indicator, according to functions capable of defining the overcoming of one city in relation to another.

The PROMETHEE II method was computationally implemented to the Program to Support Decision Making Based on Indicators (PRADIN) in order to facilitate its operationalization. Therefore, using the functions of the PROMETHEE algorithm, associated with PRADIN, simulations can be performed for the selected indicators. The values (0–1) received by each city in the simulations are known as multi-criteria indicators in the app, and from there, the hierarchy for vulnerability classification occurs.

Simulations

After collecting the data related to the chosen indicators, these data were inserted in the program, PRADIN, to be compared and hierarchized. In the present work, it was called simulation every time the program performs this process of comparison and hierarchization of the chosen indicators. Thus, these simulations indicate how the data was organized in the app.

Categorization

To facilitate the use of this analysis by health managers, the study divided the indicators into categories (demographic, social, economic, health infrastructure, population at risk, epidemiological) and carried out separate simulations with indicators belonging to each one. Thereby, the vulnerability can be analyzed in different perspectives, allowing segmented interventions based on the categories, which are usefull according to the interest of health managers.

Therefore, the simulations occurred in a segmented manner, according to groupings, evaluating the group of indicators by the pre-defined categories (**Table 2**). Posteriorly, an analysis was undertaken from a joint simulation of all multi-criteria indicators to establish the vulnerability scenario for Minas Gerais at the municipal level for COVID-19. This general simulation enabled a more complete view of the pandemic scenario at Minas Gerais. Six simulations were conducted, with the groupings of indicators and a general simulation, covering all 23 indicators, therefore totaling seven simulations (**Table 2**).

For the development of the MCDA, each city of Minas Gerais was included in a spreadsheet and received a set of data related to the indicators (**Table 1**), which were analyzed by the PRADIN program according to the total number of municipalities in order to classify them according to their vulnerability (worst indicators reflect greater vulnerability).

Weights

The weights given to the indicators (0–10) in the MCDA were established by the decision-makers. In this work, all indicators received the same weight in the hierarchy process, since no evidence was found on which factor influences COVID-19's illness and death processes to a higher or lower extent. For the authors of the present study, giving different weights to the indicators is inconsistent in view of the complexity of the pandemic scenario experienced in Brazil, a country full of structural problems, so that any type of assumption about the greater or lesser importance of the indicators without accurate legitimation by the scientific literature could invalidate the results obtained.

Results Organization

After the analysis of the indicators by MCDA, each simulation was classified into quintiles, according to the multi-criteria indicator (MCI) and divided into groups of vulnerability according to municipalities, mesoregions and population size, classified in an ascending order. Groups 1 and 2 were composed by the municipalities with the least vulnerability, group 3 with moderate vulnerability and groups 4 and 5 were those with the greatest vulnerability.

After the classification and definition of the groups, the codes of the municipalities and the program Tabwin 32 (http://www2.datasus.gov.br/DATASUS/index) were used to construct the vulnerability maps for COVID-2019. In thematic maps, the darker colors of the municipalities represent greater vulnerability, while the lightest colors indicate lower vulnerability. For the organization and evaluation of the collected data, the regionalization of the municipalities of Minas Gerais in Mesoregions was used (**Figure 1**).

To compose the data analysis, the concept of population size of each municipality was also used (39), categorizing in small municipalities those with up to 25 thousand inhabitants, in medium-sized municipalities those with population between 25 and 100 thousand inhabitants and in large municipalities those with more than 100 thousand inhabitants.

Due to the advantages of the technique, other studies used this tool to assist in decision making considering disease surveillance programs (35–37, 40–42). Interestingly, a recent study carried out in India performed an analysis very similar to this study, in which researchers mapped the vulnerability of India to COVID-19 (41). Given the importance of the method and its increasingly recurrent use, a systematic review was conducted to assess and synthesize articles that used multi-criteria analysis for decision making in health area (42). This review highlighted the methodological variety that can be used to construct MCDA,

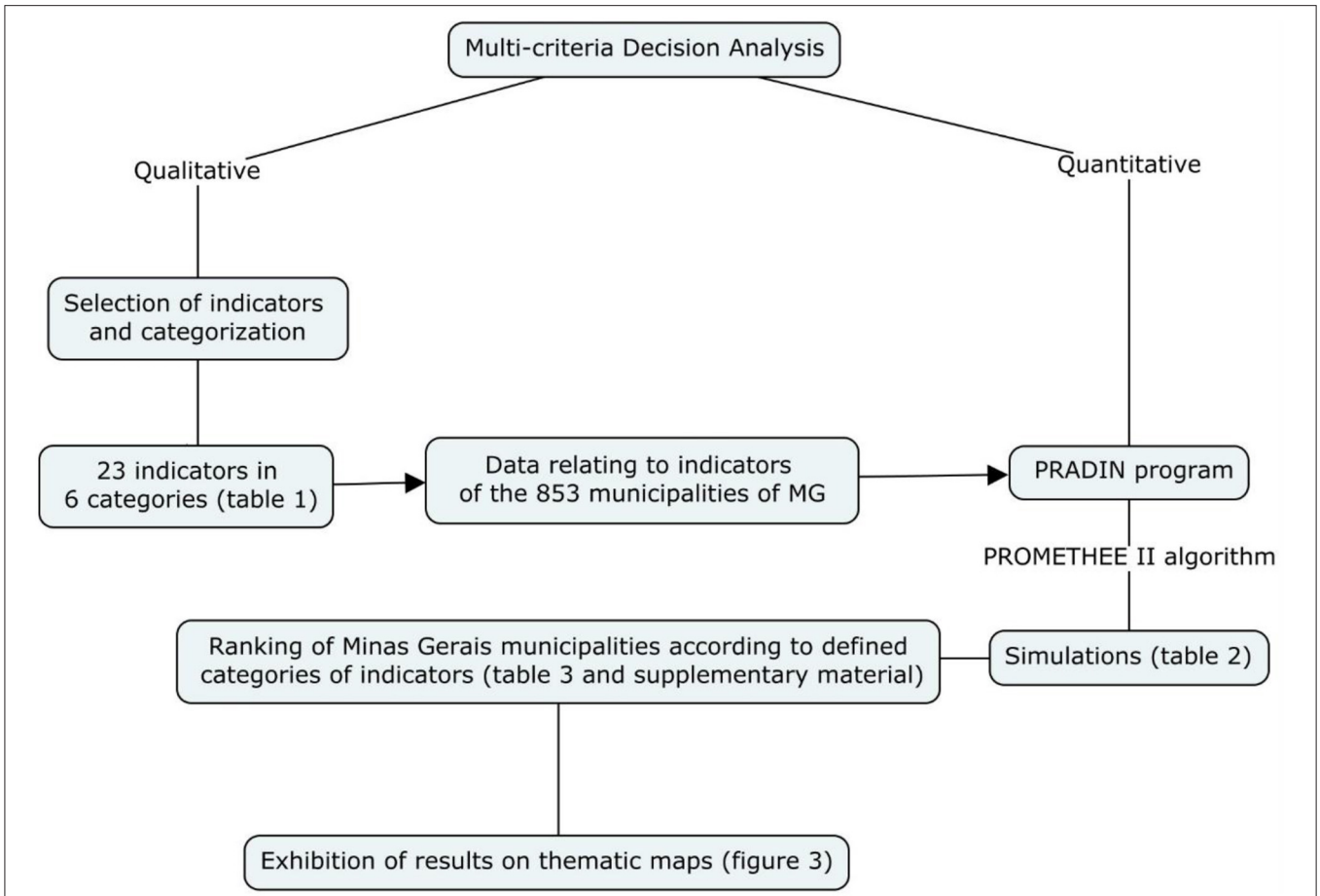


FIGURE 2 | Flowchart with step-by-step description of the methodology.

with the collaboration of both literature and decision makers and experts in the process of evaluating the best criteria and discussing the results for evaluation of the decision (42).

RESULTS

The results obtained by the multi-criteria analysis of the groupings of indicators are depicted in **Figure 3**, according to the analysis of the 853 municipalities in Minas Gerais (**Supplementary Table 1**). **Table 3** shows the mesoregions of the most and least vulnerable municipalities in Minas Gerais for COVID-19, in accordance with the grouping of indicators used in the multi-criteria analysis of decision.

From the analysis of demographic indicators, the results indicated that the Norte de Minas mesoregion is the least vulnerable (**Figure 3**-1). Accordingly, this mesoregion had 80.9% of the municipalities with the least vulnerability, groups 1 and 2, and 9 municipalities among the 10 least vulnerable in the state for these indicators. Contrarily, the Metropolitana de Belo Horizonte had 53.4% of its municipalities classified in groups 4 and 5, that is, with greater vulnerability (**Table 3**). Thus, 7 of the 10 most vulnerable municipalities in Minas Gerais were located in this mesoregion. The findings also revealed that 97.0% of the municipalities categorized as large ones were included in group 5 of vulnerability (**Table 4**).

From the analysis of social indicators, the results showed that the mesoregions Oeste de Minas and Metropolitana de Belo Horizonte occupy, together, the position of the least vulnerable. Both have 63.0% of their municipalities integrating groups 1 and 2. In contrast, Norte de Minas, Vale do Rio Doce and Vale do Mucuri stand out as the most vulnerable (**Figure 3**-2). While the first two mesoregions have, respectively, 85.3 and 81.4% of their municipalities distributed between groups 4 and 5, the Vale do Mucuri mesoregion exhibits 91.3% of its municipalities integrating the most vulnerable groups, where 4 of the 10 most vulnerable cities are located in this mesoregion (**Table 3**). Besides, 84.9% large and 45.2% medium-sized municipalities were among the least vulnerable in relation to social indicators (**Table 4**).

Regarding economic indicators, Triângulo Mineiro/Alto Paranaíba mesoregion was the least vulnerable region in the state, with 68.2% of the municipalities classified in groups 1 and 2 (**Figure 3**-3). The Jequitinhonha and Norte de Minas mesoregions had 56.9% and 47.2% of their municipalities categorized in group 5, respectively. Importantly, 6 most

TABLE 2 | Simulations performed by category of indicators (1–6) and the general simulation (7), gathering all the indicators simultaneously.

Indicators	Simulations						
	1st	2nd	3rd	4th	5th	6th	7th
Demographic							
Population percentage living in urban area	X						X
Demographic density	X						X
Social							
Percentage of inadequate sanitation		X					X
Human development index		X					X
Illiteracy percentage		X					X
Gini index		X					X
Economic							
Population percentage with monthly income higher than 70 reais (equivalent to US$ 13)			X				X
Population percentage with health insurance			X				X
Gross domestic product			X				X
Healthcare infrastructure							
Number of respirators by 1,000 inhabitants				X			X
Number of beds by 1,000 inhabitants				X			X
Number of nurses by 1,000 inhabitants				X			X
Number of doctors by 1,000 inhabitants				X			X
Number of rapid tests by 1,000 inhabitants				X			X
Number of molecular tests (RT-PCR) by 1,000 inhabitants				X			X
Population at risk							
Mortality from respiratory diseases by 1,000 inhabitants					X		X
Mortality from cardiovascular diseases by 1,000 inhabitants					X		X
Mortality from neoplasm by 1,000 inhabitants					X		X
Mortality from diabetes by 1,000 inhabitants					X		X
Population percentage with 60 years or more					X		X
Epidemiological							
COVID-19 incidence by 1,000 inhabitants						X	X
COVID-19 mortality by 1,000 inhabitants						X	X
COVID-19 lethality						X	X
All indicators (general)	X	X	X	X	X	X	X

vulnerable municipalities in the state belong to these regions (**Table 3**). Further, within this analysis category, 52.2 and 75.7% of medium-sized and large cities, respectively, are part of the groups 1 and 2 and, therefore, have lower vulnerability (**Table 4**).

When analyzing the indicators that assess health infrastructure, the results reported that Oeste de Minas and Central Mineira mesoregions are less vulnerable, with 31.8 and 30.0% of the municipalities in group 1, respectively (**Table 3**). Norte de Minas and Vale de Mucuri were among the most vulnerable areas, with 41.6 and 34.8% of the municipalities in group 5, respectively (**Figure 3**-4) (**Table 3**). In this regard, medium-sized and large municipalities with, respectively, 80.8 and 84.9% in groups 1 and 2, have lower vulnerability (**Table 4**).

Considering the estimation of population at risk, the findings showed that the mesoregion of Norte de Minas was the least vulnerable, with 61.8% of its municipalities classified in group 1. Notwithstanding that, Zona da Mata mesoregion proved to be the most vulnerable, with 42.3% of its municipalities being part of group 5 (**Table 3**) (**Figure 3**-5). Furthermore, 22.0% of the municipalities classified as small-sized were part of group 5 (**Table 4**).

Based on epidemiological indicators, the Norte de Minas and Jequitinhonha mesoregions stood out due to their lower vulnerability with, respectively, 59.6 and 56.9% of the municipalities classified in groups 1 and 2. In turn, Metropolitana de Belo Horizonte and Triângulo Mineiro/Alto Paranaíba revealed higher vulnerability, with 54.3% and 57.6% of the municipalities classified in groups 4 and 5, respectively (**Table 3**) (**Figure 3**-6). However, 7 of the 10 most vulnerable municipalities in the state are located in the mesoregion of Zona da Mata. Moreover, the findings reported that 45.2% of medium-sized municipalities and 97.0% of large municipalities were classified in group 5 (**Table 4**).

The analysis of all indicators jointly demonstrated that Triângulo Mineiro/Alto Paranaíba is the least vulnerable mesoregion (**Table 3**) (**Figure 3**-7). With 57.6% of its municipalities classified in group 1, this region contains 5 of the

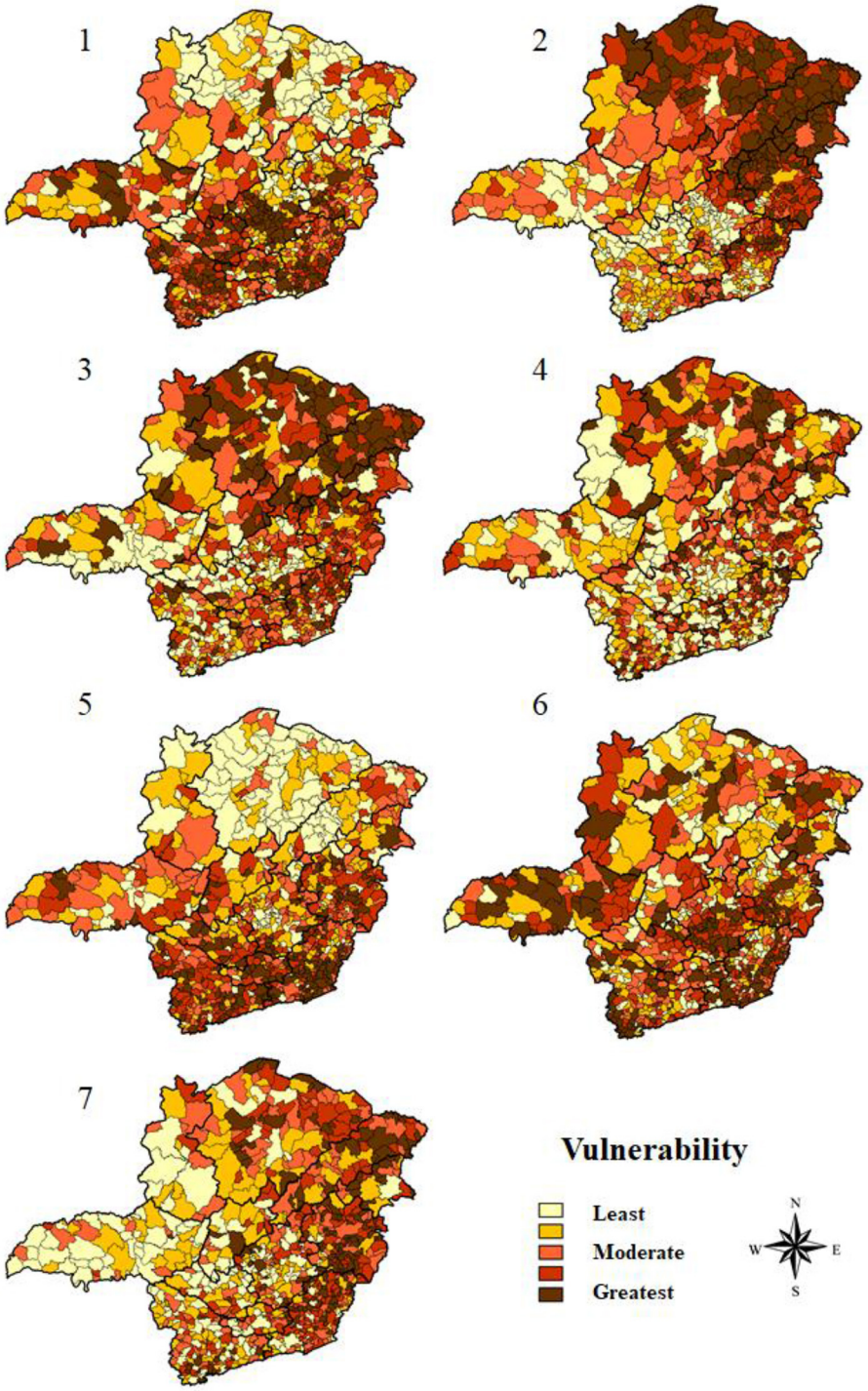


FIGURE 3 | Thematic maps of the vulnerability simulations of the municipalities of Minas Gerais for COVID-19, based on the multi-criteria decision analysis.

10 least vulnerable municipalities. It is noteworthy to highlight the positions of Vale do Mucuri and Vale do Rio Doce. While the first mesoregion has the highest rate of municipalities occupying group 5 (47.8%), Vale do Rio Doce comprises 5 of the 10 most vulnerable municipalities, with 44.1% of its municipalities integrating group 5 (**Table 3**) (**Figure 3**-7). Additionally, the

general analysis revealed that small municipalities are among the most vulnerable, with 43.7% of their representatives divided between groups 4 and 5 (**Table 4**).

DISCUSSION

The use of MCDA to generate these results was based on the perspective that the health-illness-care process depends on several factors determined by the individuals' living conditions. Thus, exploring the structure and spatial dynamics of the population is essential for the characterization of health situations to plan actions and allocate resources (15, 43). From this perspective, in order to draw a panorama of reality, the use of indicators becomes an important instrument to measure it in a succinct, objective, quick and efficient way, aiming to support an intervention (43).

In the context of the COVID-19 pandemic, it is emphasized that the novel coronavirus infection has the potential to affect everyone in society, however in a heterogeneous way (44), therefore requiring identification of areas of vulnerability. Although various strategies for mitigating the rate of disease transmission are recommended for the entire community, it is pivotal to examine areas based on their unique characteristics, including demographic variation, economic aspects, health conditions of population and characteristics of the health system, in order to produce improved and targeted interventions (45). Thus, in a state with a projection of more than 20 million inhabitants in 2019 and more than 580 thousand km² of area (IBGE), the peculiarities of each region become even more accentuated.

In this scenario, Minas Gerais occupied, in 2012, the 9th position in the national urbanization ranking (24), but its territorial vastness causes important internal inequalities regarding the urbanization rates of the municipalities, a relevant fact for vulnerability analysis of COVID-19 (46). Indeed, COVID-19 is closely related with high population density owing to the high degree of social interactions. In this sense, individuals living in urban areas are more likely to test positive for the disease when compared to individuals living in rural areas (47, 48). Thus, the lower vulnerability verified in Norte de Minas mesoregion is due to the much lower urbanization rates in comparison to the average state, in addition to the low demographic density and predominance of smaller cities (12). Furthermore, proving this strong relationship, it is important to mention the specificity of Montes Claros (**Figure 3**-1), a municipality in the Norte de Minas that differs in terms of higher rates of urbanization and, for this reason, exposed a similar profile to the municipalities located in the center-south portion of Minas Gerais, with high vulnerability, comparable to the Metropolitana de Belo Horizonte mesoregion (49).

According to the Atlas of Social Vulnerability in Brazilian Municipalities (50), by 2015, all thirty municipalities in the Southeast region classified as high social vulnerability were located in Minas Gerais. Further, considering the COVID-19 transmission scenario, locations with better education, sanitation and development indicators concentrate greater instructional and sanitary capacity to contain the spread of the virus (18). Thus, in the analysis of vulnerabilities, large and medium-sized municipalities, predominant in the central-southern portion of the state, demonstrated less risk, given that the best social indicators are verified in Oeste de Minas and Metropolitana de Belo Horizonte mesoregions. In contrast, the northern mesoregions, especially Vale do Mucuri, are the ones with the worst educational and housing conditions (50). The northern portion of the state herein proved to be more socially vulnerable to COVID-19, demanding public policies directed to improving these indicators to overcome the contagion.

In addition, from the analysis of economic indicators, the results indicated greater vulnerability in the northern region of Minas Gerais, mainly in Jequitinhonha and Norte de Minas, in contrast to Triângulo Mineiro/Alto Paranaíba and Metropolitana de Belo Horizonte. Remarkably, this information is relevant as, in the context of the pandemic, the politicians have focused on populations at risk considering mainly comorbidities and age (46). However, socioeconomic issues have been in the background, which may favor COVID-19 exposure and mortality. Additionally, economically disadvantaged people are more likely to live in accommodation with high number of people and less access to open areas, besides to having unstable occupations that do not allow remote office work (46). In this sense, the current prevention model based mainly on social isolation can be fragile and limited when applied to needy, isolated and low-educated populations (51). Therefore, poverty represents a hurdle to effective measures to contain the pandemic and must be taken into account in public policy decision making.

Regarding the health infrastructure indicators, a higher vulnerability was also found in the northern regions. This category is especially important when considering more severe cases of the disease, which require hospitalization in Intensive Care Units (ICU) (26). Inadequate health infrastructure directly influences the mortality rate caused by COVID-19. In this context, the Brazilian health regions with the highest mortality rates are located in places where the shortage of ICU beds and ventilators is more prevalent (27). Thus, the saturation of ICU and respirators resulting from the increasing demand becomes an aggravating factor for the COVID-19 pandemic and requires attention from managers (25). In this context, the construction of temporary hospitals, as has already been done in other parts of Brazil, may be an alternative.

Considering the number of tests for COVID-19, an important factor in determining the prevalence of infection/disease in the population (28), Brazil, as the vast majority of other developing countries, has a very modest number when compared to developed countries (52). Despite being the second country with the highest number of absolute deaths and the fifth with the highest number of deaths per million inhabitants, Brazil is only the 14th country testing patients, hence demonstrating a serious concern (53). The context of Minas Gerais is even more worrying, considering that the state has the third lowest number of tests per thousand inhabitants among the 26 states and 1 federative unit (6). This has failed to identify potential transmitters and directly influences the number of reported cases, which may be much lower than the actual number. Additionally, questions

TABLE 3 | Mesoregions of municipalities and groups of vulnerability (1 and 2 representing lower vulnerability, 3 moderate vulnerability and 4 and 5 greater vulnerability) for COVID-19 in the state of Minas Gerais, according to indicators used in the multi-criteria decision analysis.

Mesoregions and groups		Indicators (%)						
		Demographic	Social	Economic	Healthcare infrastructure	Population at risk	Epidemiological	General
Campo das vertentes	1	8.3	19.4	11.1	22.2	5.6	16.7	25.0
	2	25.0	33.3	27.8	25.0	16.7	13.9	22.2
	3	22.2	33.3	41.7	19.4	16.7	30.6	25.0
	4	19.4	13.9	11.1	13.9	27.8	19.4	19.4
	5	25.0	0.0	8.3	19.4	33.3	19.4	8.3
Central mineira	1	23.3	16.7	23.3	30.0	20.0	33.3	40.0
	2	26.7	30.0	20.0	20.0	20.0	13.3	23.3
	3	23.3	33.3	16.7	20.0	20.0	16.7	13.3
	4	20.0	16.7	26.7	20.0	26.7	23.3	16.7
	5	6.7	3.3	13.3	10.0	13.3	13.3	6.7
Jequitinhonha	1	37.3	0.0	3.9	9.8	45.1	23.5	0.0
	2	33.3	2.0	13.7	17.6	21.6	33.3	17.6
	3	19.6	7.8	11.8	21.6	21.6	17.6	23.5
	4	9.8	31.4	13.7	29.4	11.8	15.7	21.6
	5	0.0	58.8	56.9	21.6	0.0	9.8	37.3
Metropolitana de belo horizonte	1	17.1	42.9	25.7	28.6	28.6	15.2	27.6
	2	21.9	15.2	25.7	20.0	25.7	13.3	25.7
	3	7.6	16.2	20.0	26.7	19.0	17.1	21.9
	4	10.5	12.4	17.1	12.4	12.4	27.6	13.3
	5	42.9	13.3	11.4	12.4	14.3	26.7	11.4
Noroeste de minas	1	42.1	10.5	5.3	26.3	31.6	10.5	31.6
	2	26.3	15.8	31.6	15.8	42.1	36.8	26.3
	3	31.6	42.1	26.3	0.0	10.5	5.3	31.6
	4	0.0	15.8	15.8	26.3	15.8	36.8	10.5
	5	0.0	15.8	21.1	31.6	0.0	10.5	0.0
Norte de minas	1	57.3	1.1	11.2	7.9	61.8	30.3	5.6
	2	23.6	2.2	11.2	11.2	25.8	29.2	19.1
	3	7.9	11.2	11.2	15.7	10.1	20.2	25.8
	4	6.7	40.4	19.1	23.6	2.2	6.7	28.1
	5	4.5	44.9	47.2	41.6	0.0	13.5	21.3
Oeste de minas	1	15.9	50.0	40.9	31.8	6.8	13.6	36.4
	2	9.1	25.0	20.5	20.5	27.3	20.5	22.7
	3	13.6	25.0	22.7	25.0	20.5	38.6	22.7
	4	31.8	0.0	9.1	6.8	29.5	15.9	15.9
	5	29.5	0.0	6.8	15.9	15.9	11.4	2.3
Sul/sudoeste de minas	1	6.2	32.2	26.7	23.3	6.8	18.5	26.0
	2	11.0	43.8	26.7	24.7	11.6	23.3	28.1
	3	27.4	19.9	26.0	16.4	19.9	19.9	17.1
	4	28.8	3.4	12.3	16.4	34.2	16.4	17.1
	5	26.7	0.7	8.2	19.2	27.4	21.9	11.6
Triângulo mineiro/alto paranaíba	1	19.7	28.8	39.4	25.8	22.7	13.6	57.6
	2	19.7	30.3	28.8	31.8	24.2	16.7	24.2
	3	19.7	34.8	16.7	19.7	31.8	12.1	13.6
	4	24.2	6.1	7.6	19.7	13.6	30.3	3.0
	5	16.7	0.0	7.6	3.0	7.6	27.3	1.5
Vale do mucuri	1	43.5	0.0	17.4	4.3	13.0	26.1	4.3
	2	17.4	0.0	8.7	17.4	30.4	21.7	17.4

(Continued)

TABLE 3 | Continued

Mesoregions and groups		Indicators (%)						
		Demographic	Social	Economic	Healthcare infrastructure	Population at risk	Epidemiological	General
	3	17.4	8.7	13.0	17.4	39.1	30.4	21.7
	4	21.7	13.0	17.4	26.1	13.0	8.7	8.7
	5	0.0	78.3	43.5	34.8	4.3	13.0	47.8
Vale do rio doce	1	16.7	4.9	9.8	11.8	7.8	19.6	3.9
	2	26.5	5.9	19.6	12.7	20.6	22.5	6.9
	3	18.6	7.8	11.8	25.5	20.6	15.7	17.6
	4	24.5	42.2	36.3	28.4	26.5	24.5	27.5
	5	13.7	39.2	22.5	21.6	24.5	17.6	44.1
Zona da mata	1	6.3	12.7	14.8	19.0	7.0	21.1	9.2
	2	16.9	19.0	11.3	21.1	12.0	11.3	14.1
	3	30.3	26.1	24.6	19.0	19.7	22.5	19.0
	4	23.9	26.8	32.4	21.8	19.0	20.4	30.3
	5	22.5	15.5	16.9	19.0	42.3	24.6	27.5

TABLE 4 | Results of the multi-criteria decision analysis according to the population size of municipalities in the state of Minas Gerais and their vulnerability classification, in accordance with the grouping of indicators used in the multi-criteria decision analysis for COVID-19.

Cities by population size	Total of cities	Groups	Indicators (%)						
			Demographic	Social	Economic	Healthcare infrastructure	Population at risk	Epidemiological	General
Small size (until 25 thousand inhabitants)	705	1	23.3	12.9	16.2	11.1	18.6	24.3	17.0
		2	22.7	19.9	20.0	20.0	18.9	21.7	18.2
		3	22.8	22.3	21.7	21.3	20.6	21.8	21.1
		4	20.7	22.8	21.1	23.5	20.0	20.1	21.1
		5	10.5	22.1	21.0	24.1	22.0	12.1	22.6
Medium size (25–100 thousand inhabitants)	115	1	6.1	45.2	32.2	56.5	26.1	0.0	31.3
		2	9.6	23.5	20.0	24.3	26.1	15.7	27.8
		3	8.7	11.3	13.9	13.9	16.5	14.8	14.8
		4	20.9	8.7	17.4	4.3	20.9	24.3	18.3
		5	54.8	11.3	16.5	0.9	10.4	45.2	7.8
Large size (more than 100 thousand inhabitants)	33	1	0.0	84.8	54.5	78.8	30.3	0.0	45.5
		2	0.0	12.1	21.2	6.1	24.2	0.0	33.3
		3	0.0	3.0	6.1	15.2	21.2	0.0	15.2
		4	3.0	0.0	6.1	0.0	18.2	3.0	3.0
		5	97.0	0.0	12.1	0.0	6.1	97.0	3.0

have been raised about possible underreporting of cases, which further aggravates the state's situation, thus making government intervention urgent (7).

The factors of comorbidities and age, which compose the population at risk indicators, were raised during the pandemic in order to draw a well-defined profile of people more susceptible to the complications of COVID-19. Both in Wuhan, China (54) and in the Italian states (55), respiratory and cardiac diseases, as well as neoplasms, diabetes and advanced age are considered factors for complications of the clinical condition and of higher mortality. In this sense, access to health infrastructure and education acts as an aggravation of diseases, and the region that is able to provide more appropriately these resources to the population ensures better conditions to prolong life (56). Importantly, a greater longevity is also accompanied by an increase in the elderly population, which is more affected by chronic diseases, and with regard to COVID-19, these diseases act as complicating factors of the clinical condition. Thus, as reported by this study, the worst social and health infrastructure indicators in the Norte de Minas may be associated with lower longevity, hence leading the northern portion of the State to have fewer people in the risk group (57). On the other hand,

the Zona da Mata is better assisted, which leads its population to be longer-lived and, consequently, to have a higher number of people more susceptible to the health complications caused by COVID-19.

Another pivotal issue is that coping with the COVID-19 pandemic involves changes in the health system and also requires political decisions that affect the management of chronic non-communicable diseases, as well as patients' adherence to treatment, especially those from less favored social classes (58). Furthermore, the pandemic scenario increases patients' fear of seeking health services, which can increase mortality from events related to the chronic illness (59). Based on the identification of this regional vulnerability profile, it is possible to outline public policies that address the major diseases associated with the worsening of the clinical condition of patients with COVID-19, as well as specialized care for the elderly, markedly more affected by the condition.

With regard to epidemiological aspects, greater vulnerability was found in the metropolitan mesoregions of Metropolitana de Belo Horizonte, Triângulo Mineiro/Alto Paranaíba and Zona da Mata and, in a lesser extent, in the Norte de Minas and Vale do Mucuri. In this sense, the observance of the great impact of the mesoregions further south and southeast of the state becomes relevant when considering that these places are located on the border with the states of São Paulo and Rio de Janeiro, which concentrate the highest number of cases in Brazil. Besides that, the lower socioeconomic development in the north of Minas Gerais favors the scenario of underreporting of cases in the state, which can be associated with a large north-south discrepancy in the numbers found. Thus, social distance should be considered through reliable measures, including travel restrictions or even the institution of lockdown, which have proven effective in countries such as China, South Korea, Iran, Italy, France and the United States (60). In addition to these measures, others widely used worldwide must be promulgated with greater avidity in the most affected municipalities, including awareness about the use of personal protective equipment, social distance, closing schools and business buildings, quarantine, cleaning and disinfection and increase in the number of tests (19, 61–64).

Considering the joint analysis of all indicators, the lower vulnerability of Triângulo Mineiro and Alto Paranaíba was proven, with better social, economic, population at risk and health infrastructure indicators. Thus, the highest human development indexes, in addition to a diverse and historically integrated economy to the State of São Paulo, associated with a higher presence of young people and the concentration of hospital resources, integrate factors to reduce vulnerability to COVID- 19 (65, 66). Similarly, the findings also reported the greatest vulnerability in Vale do Mucuri, followed by Vale do Rio Doce. In these areas, the economy is fragile, basically composed of the primary sector. Education and sanitation indicators are remarkably low and there is a predominance of higher age groups (65). Additionally, the few existing hospital resources are concentrated especially in the municipalities of Teófilo Otoni and Governador Valadares, not reaching all the surrounding municipalities (66).

Besides to the evident differences revealed by the indicators between the mesoregions of Minas Gerais, data also exhibited differences between the municipalities grouped by population size, with emphasis on the numerical predominance of small municipalities. In this sense, the analysis performed in the present study is essential to better understand the possible particularities of these municipalities in the face of the pandemic, highlighting the need to formulate specific strategies and public policies according to the size of the population.

The high transmissibility of SARS-COV-2 in large urban centers with population agglomerations results in a rapid and exponential increase in the number of cases and deaths from the entry of the virus into the population (65). In fact, the rapid increase in the absolute number of cases requires the development of containment measures and, in some cases, when applied with due urgency, relative success is achieved. However, the pandemic is not restricted to large municipalities, but also reaches medium-sized and small areas, consolidating the internalization of the disease, which in Brazil reaches more than 90% of the municipalities (6, 67).

The smaller cities did not have large absolute number of cases and deaths compared to the others, but when investigating the proportion measures, including incidence, lethality and mortality, as performed by this work, it is clear that several small municipalities were in a serious situation. Thus, the exposed and disseminated absolute data of these cities do not cause as much impact on the population and public managers as they should, making containment measures, such as social distance, take time to be employed or adhered by the population and managers.

This scenario is particularly worrying for the small municipalities that were, for the most part, more vulnerable, especially in relation to health infrastructure and financial resources (68). A large portion of these ones reported few ICU beds, few or no respirators and a reduced number of health professionals, leading to a high lethality of the disease, since the basic care conditions of the most serious cases of COVID-19 are not guaranteed (45).

Based on the regionalization process defined by the Minas Gerais Health Regionalization Master Plan, these small municipalities should be assisted by medium and large-sized cities, enabling access to medium and high complexity services in severe cases of the disease (27). However, the reality of the state of Minas Gerais does not meet this proposal, as the assessment demonstrated several discontinuities and inequalities in all indicators of the state. Alarmingly, the small municipalities are isolated in the middle of the pandemic, without support from the medium-sized and large municipalities and without enough resources to improve their own health infrastructure (66, 68). Municipalities listed to offer this support face the overcrowding of beds and the lack of respirators (25).

Importantly, some indicators that could contribute to the mapping of the vulnerability of Minas Gerais to COVID-19 were not included due to the absence of data related to the municipalities, hence hindering the tabulation. The data included also had a difference in the dates on which they were made available, since some are only accessed by the 2010 demographic census conducted by IBGE. Another limitation

is also related to the database available for consultation, since various indicators may be out of date, especially in small municipalities where the registration process does not occur or is not done properly, showing the presence of under-notification of cases. Nevertheless, it is important to use these data to evaluate the vulnerability considering the risk of worsening disease, since this article intends to measure the capacity of the cities to contain COVID-19 and its complications.

In addition, in the data tabulation process for MCDA, one of the steps consists in defining weights for the different data included, herein determining which ones would have different intensities of influence. However, the present study chose to keep all data with the same weight owing to the lack of evidence, showing which factors would have a higher or lower influence in the pandemic and the difficulty of stipulating the proportion of this influence. Therefore, the maps and the findings of this study should be used only as an instrument of guidance for public policies with other existing tools, and not as the only resource. Also, a segmented analysis may be performed by category of indicators in order to avoid possible differences in the influence of indicators in the compilation of the final result.

From the vulnerability analysis performed, it is clear that the demands of the municipalities of Minas Gerais in the context of COVID-19 are different, varying according to the region in which they are located and their population size. Thus, a public policy planned for the state will have totally different applicability and effectiveness depending on the region or municipality in question. Therefore, a more segmented analysis of the state should be conducted, as proposed by this work, in order to identify the particularities of each municipality and mesoregion in the search for interventions that have an effect in a faster and more practical way, as the context requires. In this scenario, measures are needed to contain the spread of the disease in the state as a whole, not just in the most economically important regions. For this, the problem of capillarity of the state related to social, economic and health indicators must be solved, so that everyone has a similar capacity to fight the pandemic. As a result, the state will not only benefit from combating the COVID-19 pandemic, but also from combating all inequalities that have been consolidated in Minas Gerais and directly affect the quality of life of the population in the less-assisted regions.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

SO contributed to the conception of the study. SO and IC contributed to the acquisition, analysis and interpretation of data, contributed to the statistical analysis, interpretation of data, and creation of table and figures. VA, KC, and JL participated in revising it critically for important intellectual content for discussion topic. All the authors co-wrote the paper and give final approval to the version to be submitted.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2021.586670/full#supplementary-material

REFERENCES

1. Jiang F, Deng L, Zhang L, Cai Y, Cheung CW, Xia Z. Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). *J Gen Intern Med.* (2020) 35:1545–9. doi: 10.1007/s11606-020-05762-w
2. Hui DS, Azhar EI, Madani TA, Ntoumi F, Kock R, Dar O, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health: the latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis.* (2020) 91:264–6. doi: 10.1016/j.ijid.2020.01.009
3. Petrosillo N, Viceconte G, Ergonul O, Ippolito G, Petersen E. COVID-19, SARS and MERS: are they closely related? *Clin Microbiol Infect.* (2020) 26:729–34. doi: 10.1016/j.cmi.2020.03.026
4. Kolifarhood G, Aghaali M, Saadati HM, Taherpour N, Rahimi S, Izadi N, et al. Epidemiological and clinical aspects of Covid-19; a narrative review. *Arch Acad Emerg Med.* (2020) 8:1.
5. Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R. COVID-19 infection: origin, transmission, and characteristics of human coronaviruses. *J Adv Res.* (2020) 24:91–8. doi: 10.1016/j.jare.2020.03.005
6. Brasil. *Ministério da Saúde: Situação Epidemiológica da COVID-19.* (2021). Available online at: https://www.gov.br/saude/pt-br (accessed February 11, 2021).
7. Alves THE, de Souza TA, Silva AS, Ramos NA, de Oliveira SV. Underreporting of death by COVID-19 in Brazil's second most populous state. *medRxiv.* (2020) Available online at: https://www.medrxiv.org/content/10.1101/2020.05.20.20108415v1 (accessed July 3, 2020).
8. Ayres JRCM, França Júnior I, Calazans G, Salleti H. Vulnerabilidade e prevenção em tempos de aids. In: Barbosa R, Parker R, editors. *Sexualidade Pelo Avesso: Direitos, Identidades e Poder*. Rio de Janeiro: Relume Dumará (1999). p. 50–71.
9. Sánchez AIM, Bertolozzi MR. Pode o conceito de vulnerabilidade apoiar a construção do conhecimento em Saúde Coletiva? *Ciênc Saúde Colet.* (2007) 12:319. doi: 10.1590/S1413-81232007000200007
10. Gomes LFAM, Araya MCG, Carignano C. *Tomada de Decisão em Cenários Complexos: Introdução aos Métodos Discretos do Apoio Multicritério à Decisão*. São Paulo: Pioneira Thomson Learning. (2004). p. 126.
11. Brazilian Institute of Geography and Statistics. *Estimativas Populacionais dos Municípios em 2019.* (2019). Available online at: https://sidra.ibge.gov.br/tabela/6579 (accessed June 22, 2020).
12. Brazilian Institute of Geography and Statistics. *Censo Demográfico de 2010: Características da População e dos Domicílios - Resultados do Universo.* (2011). Available online at: https://biblioteca.ibge.gov.br/visualizacao/periodicos/93/cd_2010_caracteristicas_populacao_domicilios.pdf (accessed June 22, 2020).
13. Diniz AMA, Batella WB. O Estado de minas gerais e suas regiões: um resgate histórico das principais propostas oficiais de regionalização. *Soc Nat.* (2005) 17:59–77.
14. You H, Wu X, Guo X. Distribution of COVID-19 morbidity rate in association with social and economic factors in Wuhan, China: implications for urban development. *Int J Environ Res Public Health.* (2020) 17:3417. doi: 10.3390/ijerph17103417

15. Barcellos CC, Sabroza PC, Peiter P, Rojas LI. Spatial organization, health and quality of life: use of spatial analysis and indicators in health situation analysis. *Inf Epidemiol Sus*. (2002) 11:129–38. doi: 10.5123/S0104-16732002000300003
16. Pirouz B, Haghshenas SS, Pirouz B, Haghshenas SS, Piro P. Development of an assessment method for investigating the impact of climate and urban parameters in confirmed cases of covid-19: a new challenge in sustainable development. *Int J Environ Res Public Health*. (2020) 17:2801. doi: 10.3390/ijerph17082801
17. Ramírez IJ, Lee J. COVID-19 emergence and social and health determinants in Colorado: a rapid spatial analysis. *Int J Environ Res Public Health*. (2020) 17:3856. doi: 10.3390/ijerph17113856
18. Zhang Y, Chen C, Zhu S, Shu C, Wang D, Song J, et al. Isolation of 2019-nCoV from a stool specimen of a laboratory-confirmed case of the coronavirus disease 2019 (COVID-19). *China CDC Weekly*. (2020) 2:123–4. doi: 10.46234/ccdcw2020.033
19. Güner R, Hasanoglu I, Aktas F. COVID-19: Prevention and control measures in community. *Turk J Med Sci*. (2020) 50:571–7. doi: 10.3906/sag-2004-146
20. Malta, DC, Bernal RTI, de Souza MFM, Szwarcwald CL, Lima M, et al. Social inequalities in the prevalence of self-reported chronic non-communicable diseases in Brazil: national health survey 2013. *Int J Equity Health*. (2016) 15:153. doi: 10.1186/s12939-016-0427-4
21. Da Silva MHA, Procópio IM. A fragilidade do sistema de saúde brasileiro e a vulnerabilidade social diante da COVID-19. *Rev Bras Promoç Saúde*. (2020) 33:10724. doi: 10.5020/18061230.2020.10724
22. Llorca J, Delgado-Rodríguez M. Visualising exposure-disease association: the Lorenz curve and the Gini index. *Med Sci Monit*. (2002) 8:MT193–7.
23. Lee WC. Characterizing exposure–disease association in human populations using the Lorenz curve and Gini index. *Stat Med*. (1997) 16:729–39. doi: 10.1002/(SICI)1097-0258(19970415)16:7<729::AID-SIM491>3.0.CO;2-A
24. Brazilian Institute of Geography and Statistics. *Indicadores Sociais Municipais: Uma Análise dos Resultados do Universo do Censo Demográfico 2010 - Estudos & Pesquisas: Informação Demográfica e Socioeconômica*. (2011). Available online at: https://biblioteca.ibge.gov.br/visualizacao/livros/liv54598.pdf (accessed June 22, 2020).
25. Rache B, Rocha R, Nunes L, Spinola P, Malik AM, Massuda A. *Necessidades de Infraestrutura do SUS em Preparo à COVID-19: Leitos de UTI, Respiradores e Ocupação Hospitalar*. (2020). Available online at: https://ieps.org.br/wp-content/uploads/2020/04/IEPS-NT3.pdf (accessed July 3, 2020).
26. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. (2020) 395:507–13. doi: 10.1016/S0140-6736(20)30211-7
27. Moreira, RS. COVID-19: intensive care units, mechanical ventilators, and latent mortality profiles associated with case-fatality in Brazil. *Cad Saúde Púb*. (2020) 36:e00080020. doi: 10.1590/0102-311x00080020
28. Krammer F, Simon V. Serology assays to manage COVID-19. *Science*. (2020) 368:1060–1. doi: 10.1126/science.abc1227
29. World Health Organization. *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. (2020). Available online at: https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf (accessed June 15, 2020).
30. Wang B, Li R, Lu Z, Huang Y. Does comorbidity increase the risk of patients with COVID-19: evidence from meta-analysis. *Aging*. (2020) 12:6049–57. doi: 10.18632/aging.103000
31. Pizzichini MMM, Patino CM, Ferreira JC. Medidas de frequência: calculando prevalência e incidência na era do COVID-19. *J Bras Pneumol*. (2020) 46:20200243. doi: 10.36416/1806-3756/e20200243
32. Schumann LRMA, Moura LBA. Índices sintéticos de vulnerabilidade: uma revisão integrativa de literatura. *Ciência Saúde Coletiva*. (2015) 20:2105–20. doi: 10.1590/1413-81232015207.10742014
33. Brans J, Mareshcal B. Promethee methods. In: S. Greco, editor. *Multiple Criteria Decision Analysis: State of the Art Surveys*. Berlin: Spriger. (2005). p. 163–195. Available online at: https://www.springer.com/gp/book/9780387230818
34. Ministério da Saúde, Secretaria-Executiva, Área de Economia da Saúde e Desenvolvimento. *Avaliação de Tecnologias em Saúde: Ferramentas Para a Gestão do SUS*. Brasília: Ministério da Saúde (2009). p. 110.
35. Vinhaes MC, de Oliveira SV, Reis PO, de Lacerda Sousa AC, Silva RA, Obara MT, et al. Assessing the vulnerability of Brazilian municipalities to the vectorial transmission of *Trypanosoma cruzi* using multi-criteria decision analysis. *Acta Trop*. (2014) 137:105–10. doi: 10.1016/j.actatropica.2014.05.007
36. de Oliveira SV, Fonseca LX, de Araújo Vilges KM, Maniglia FV, Pereira SV, de Caldas EP, et al. Vulnerability of Brazilian municipalities to hantavirus infections based on multi-criteria decision analysis. *Emerg Themes Epidemiol*. (2015) 12:15. doi: 10.1186/s12982-015-0036-5
37. Sarkar S, Strutz SE, Frank DM, Rivaldi CL, Sissel B, Sánchez-Cordero V. Chagas disease risk in Texas. *PLoS Negl Trop Dis*. (2010) 4:e836. doi: 10.1371/journal.pntd.0000836
38. Jannuzzi PDM, Miranda WD, Silva DD. Análise multicritério e tomada de decisão em políticas públicas: aspectos metodológicos, aplicativo operacional e aplicações. *Inform Púb*. (2009) 11:69–87.
39. Calvo MCM, de Lacerda JT, Colussi CF, Schneider IJC, Rocha TAH. Estratificação de municípios brasileiros para avaliação de desempenho em saúde. *Epidemiol Serv Saúde*. (2016) 25:767–76. doi: 10.5123/S1679-49742016000400010
40. Hussey LK, Arku G. Prioritizing climate-sensitive infectious diseases under a changing climate in Ghana: a multicriteria evaluation analysis approach. *Reg Environ Change*. (2020) 20:1–14. doi: 10.1007/s10113-020-01582-0
41. Acharya R, Porwal A. A vulnerability index for the management of and response to the COVID-19 epidemic in India: an ecological study. *Lancet Glob Health*. (2020) 8:e1142–51. doi: 10.1016/S2214-109X(20)30300-4
42. Frazão T, Camilo D, Cabral E et al. Multicriteria decision analysis (MCDA) in health care: a systematic review of the main characteristics and methodological steps. *BMC Med Inform Decis Mak*. (2018) 18:90. doi: 10.1186/s12911-018-0663-1
43. Pitchon A, Girodo AM, Gomes CC, Gomes DHP, Pimenta Júnior FG, Freire F, et al. *Índice de Vulnerabilidade da Saúde 2012*. Belo Horizonte: Secretaria Municipal de Saúde (2013). p. 25.
44. Smith JA, Judd J. COVID-19: Vulnerability and the power of privilege in a pandemic. *Health Promot J Austr*. (2020) 31:158. doi: 10.1002/hpja.333
45. Requia WJ, Kondo EK, Adams MD, Gold DR, Struchiner CJ. Risk of the Brazilian health care system over 5572 municipalities to exceed health care capacity due to the 2019 novel coronavirus (COVID-19). *Sci Total Environ*. (2020) 730:139–44. doi: 10.1016/j.scitotenv.2020.139144
46. Patel JA, Nielsen FBH, Badiani AA, Assi S, Unadkat VA, Patel B, et al. Poverty, inequality and COVID-19: the forgotten vulnerable. *Public Health*. (2020) 183:110–11. doi: 10.1016/j.puhe.2020.05.006
47. Wang KW, Gao J, Wang H, Wu XL, Yuan QF, Guo FY, et al. Epidemiology of 2019 novel coronavirus in Jiangsu Province, China after wartime control measures: a population-level retrospective study. *Travel Med Infect Dis*. (2020) 35:101654. doi: 10.1016/j.tmaid.2020.101654
48. de Lusignan S, Dorward J, Correa A, Jones N, Akinyemi O, Amirthalingam G, et al. Risk factors for SARS-CoV-2 among patients in the Oxford royal college of general practitioners research and surveillance centre primary care network: a cross-sectional study. *Lancet Infect Dis*. (2020) 20:1034–42. doi: 10.1016/S1473-3099(20)30371-6
49. Costa MA, Marguti BO. *Atlas da Vulnerabilidade Social nos Municípios Brasileiros*. Brasília: Instituto de Pesquisa Econômica Aplicada (2015). p. 77.
50. Cardoso DF, Santana Ribeiro LC. Índice relativo de qualidade de vida para os municípios de minas gerais. *Planej Polít*. (2015) 45:347–75.
51. Chung RY, Dong D, Li MM. Socioeconomic gradient in health and the covid-19 outbreak. *BMJ*. (2020) 369:m1329. doi: 10.1136/bmj.m1329
52. COVID- 19 Coronavirus Pandemic (2020). Available online at: https://www.worldometers.info/coronavirus/ (accessed June 5, 2020).
53. Shams SA, Haleem A, Javaid M. Analyzing COVID-19 pandemic for unequal distribution of tests, identified cases, deaths, and fatality rates in the top 18 countries. *Diabetes Metab Syndr*. (2020) 14:953–61. doi: 10.1016/j.dsx.2020.06.051
54. Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Zhonghua Liu Xing Bing Xue Za Zhi*. (2020) 41:145–51. doi: 10.3760/cma.j.issn.0254-6450.2020.02.003
55. Pluchino A, Biondo AE, Giuffrida N, Inturri G, Latora V, Le Moli R, et al. A novel methodology for epidemic risk assessment: the case of COVID-19

outbreak in Italy. *arXiv*. (2020). Available online at: https://arxiv.org/abs/2004.02739 (accessed July 3, 2020).
56. Ishitani LH, Franco GC, Perpétuo IHO, França E. Desigualdade social e mortalidade precoce por doenças cardiovasculares no Brasil. *Rev Saúde Púb.* (2006) 40:684–91. doi: 10.1590/S0034-89102006000500019
57. Pereira CS, Hespanhol AN. Região e regionalização no estado de minas gerais e suas vinculações com as políticas públicas. *Formação.* (2015) 2:42–7. doi: 10.33081/formacao.v1i22.3510
58. Basu S. Non-communicable disease management in vulnerable patients during Covid-19. *Indian J Med Ethics.* (2020) 2:103–5. doi: 10.20529/IJME.2020.041
59. Hammad TA, Parikh M, Tashtish N, Lowry CM, Gorbey D, Forouzandeh F, et al. Impact of COVID-19 pandemic on ST-elevation myocardial infarction in a non-COVID-19. *Catheter Cardiovasc Interv.* (2020) 97:208–14. doi: 10.1002/ccd.28997
60. Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, et al. The effect of large-scale anti-contagion policies on the COVID-19. *Nature.* (2020) 584:262–7. doi: 10.1038/s41586-020-2691-0
61. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schunemann H, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet.* (2020) 395:1973–87. doi: 10.1016/S0140-6736(20)31142-9
62. Cook TM. Personal protective equipment during the coronavirus disease (COVID) 2019 pandemic - a narrative review. *Anaesthesia.* (2020) 75:920–7. doi: 10.1111/anae.15071
63. Davies NG, Kucharski AJ, Eggo RM, Gimma A, Edmund WJ. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *Lancet Public Health.* (2020) 5:e375–85. doi: 10.1016/S2468-2667(20)30133-X
64. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*. (2020) 584:257–61. doi: 10.1038/s41586-020-2405-7
65. Stier A, Berman M, Bettencourt L. COVID-19 attack rate increases with city size. *medRxiv.* (2020). Available online at: https://www.medrxiv.org/content/10.1101/2020.03.22.20041004v2 (accessed June 22, 2020).
66. Simões R, Guimarães C, Godoy N, Velloso T, Araújo T, Galinari R, et al. Rede Urbana da oferta de serviços de saúde: uma análise de clusters espaciais para minas gerais. In: *Anais do XIV Encontro Nacional de Estudos Populacionais.* Caxambu: Minas Gerais (2016). p. 1–27.
67. Fundação Oswaldo Cruz. *Nota Técnica: Interiorização e as Redes de Atendimento em Saúde*. (2020). Available online at: https://portal.fiocruz.br/documento/nota-tecnica-interiorizacao-do-covid-19-e-redes-de-atendimento-em-saude-04/05 (accessed June 25, 2020).
68. Pinafo E, Nunes EFPA, Carvalho BG, Mendonça FF, Domingos CM, Silva CR. Problemas e estratégias de gestão do SUS: a vulnerabilidade dos municípios de pequeno porte. *Ciênc Saúde Coletiva.* (2020) 25:1619–28. doi: 10.1590/1413-81232020255.34332019

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.

frontiers in Medicine

ORIGINAL RESEARCH
published: 20 May 2021
doi: 10.3389/fmed.2021.573726

Challenges Caused by Imported Cases Abroad for the Prevention and Control of COVID-19 in China

Jianfei Zhu[1,2†], Qingqing Zhang[3†], Chenghui Jia[1†], Shuonan Xu[4†], Jie Lei[1†], Jiakuan Chen[1], Yanmin Xia[1], Wenchen Wang[1], Xuejiao Wang[1], Miaomiao Wen[1], Hongtao Wang[2], Zhipei Zhang[1], Wuping Wang[1], Jinbo Zhao[1*] and Tao Jiang[1*]*

[1] Department of Thoracic Surgery, Tangdu Hospital, Air Force Military Medical University (Fourth Military Medical University), Xi'an, China, [2] Department of Thoracic Surgery, Shaanxi Provincial People's Hospital, Xi'an, China, [3] Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Xi'an Medical University, Xi'an, China, [4] Department of Cardiothoracic Surgery, The First Affiliated Hospital of Xi'an Medical University, Xi'an, China

Background: Overseas imported cases of COVID-19 continue to increase in China, so we conducted this study to review the epidemiological characteristics of these patients.

Methods: From February 26 to April 4, 2020, the imported cases from abroad were enrolled in this study. The effect of prevention countermeasures in curbing the spread of COVID-19 was assessed in this study. Moreover, we defined incubation period and confirmed time as from the date of leaving the epicenter to date of symptom onset and date of final diagnosed, respectively, and the interval of symptom onset to final diagnosed time was defined as diagnostic time. Categorical variables were summarized as numbers and percentages, and the difference among the variables were analyzed.

Results: For 670 cases imported from abroad, 555 were Chinese and 115 were foreigners. Apparently, confirmed cases had significantly decreased after China was compelled to temporarily suspend the entry of foreign passport holders with valid visas or residence permits; 6 days after implement of controlled measures, the daily new confirmed cases were reduced to 13 cases. Moreover, about 84.3% of patients (166/197) presented symptoms 1 week after leaving the epicenter, and notably seven patients (3.6%) had symptoms 2 weeks after leaving the epicenter. The median incubation period was 3.0 days (inter quartile range, 1.0 to 6.0), the 95th percentile was 11.6 days. Additionally, most of cases (92.9%) were detected positively of nucleic acid after symptom onset with 4 days, the median diagnostic time was 2.0 days (interquartile range, 1.0 to 3.0), and the 95th percentile of the distribution was 5.0 days. Finally, about 5.8% of patients were healthy carriers, and the median confirmed time of asymptomatic patients was 4.0 days (interquartile range, 2.0 to 9.0). The following variables might be associated with confirmed time: symptom type ($P = 0.005$), exported regions ($P < 0.001$), and symptom onset time ($P < 0.001$).

OPEN ACCESS

Edited by:
Matteo Convertino, Hokkaido University, Japan

Reviewed by:
Rongrong Ma, The Second Affiliated Hospital of Xi'an Jiaotong University, China
Ruchong Chen, Guangzhou Institute of Respiratory Health, China
Xiaohong Chang, Yanan University Affiliated Hospital, China

***Correspondence:**
Wuping Wang
wwp._8@163.com
Jinbo Zhao
zhaojinbo@aliyun.com
Tao Jiang
jiangtaotd@163.com

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

Specialty section:
This article was submitted to Infectious Diseases - Surveillance, Prevention and Treatment, a section of the journal Frontiers in Medicine

Received: *17 June 2020*
Accepted: *26 March 2021*
Published: *20 May 2021*

Citation:
Zhu J, Zhang Q, Jia C, Xu S, Lei J, Chen J, Xia Y, Wang W, Wang X, Wen M, Wang H, Zhang Z, Wang W, Zhao J and Jiang T (2021) Challenges Caused by Imported Cases Abroad for the Prevention and Control of COVID-19 in China. Front. Med. 8:573726. doi: 10.3389/fmed.2021.573726

Conclusions: The prevention countermeasures for imported cases implemented by the Chinese government played an indispensable role in curbing the spread of COVID-19; the time of departure from epicenter could provide an estimate of the incubation period; and a confirmed time, 2-week quarantine period might need to be prolonged, while asymptomatic patients should be closely monitored.

Keywords: COVID-19, global pandemic, imported case, epidemiological characteristics, prevention

INTRODUCTION

The coronavirus disease (COVID-19) (1) epidemic broken out in Wuhan City of China in December 2019, and it had spread rapidly around the world (2) due to economic globalization. Internationally, on March 11, the epidemic spread worldwide in 114 countries or regions, a total of 118,319 confirmed cases and 4,292 deaths worldwide (3) were reported by WHO. Under this situation, WHO announced COVID-19 as a global pandemic (4). In China mainland, the epidemic was controlled effectively, on March 27, 20 provinces or autonomous regions had reported no domestic cases for more than 4 weeks (5), and some regions of China reduced the public health emergency to Level Three Response (6, 7).

To our disappointment, on February 26, the first imported cases were reported in Ningxia Hui Autonomous Region of China, imported from Iran (8). Epidemic prevention and control measures to guard against imported cases were brought out by the authorities, such as all entering people received temperature test and home quarantine. With the increasing number of imported patients confirmed, on March 18, professor Zhong, a top Chinese epidemiologist, suggested that all entry population from epicenter should undergo nucleic acid testing (9), which resulted in a continuous increase in the number of imported cases, as on March 26, the new imported cases reached 54 cases (10). To more powerfully curb the spread of COVID-19, on March 28, China was compelled to temporarily suspend the entry of foreign passport holders with valid visas or residence permits (11).

Therefore, prevention and treatment measures for imported cases based on the epidemiological characteristics of imported cases can be formulated. However, for imported cases, most of them could not provide the definite exposure time, which was indispensable for infected disease. In real life study, it was important to utilize the available information to predict the outcome of the imported cases. So, we conducted this study to analyze the epidemiological characteristics of COVID-19 cases imported from abroad by using the available information.

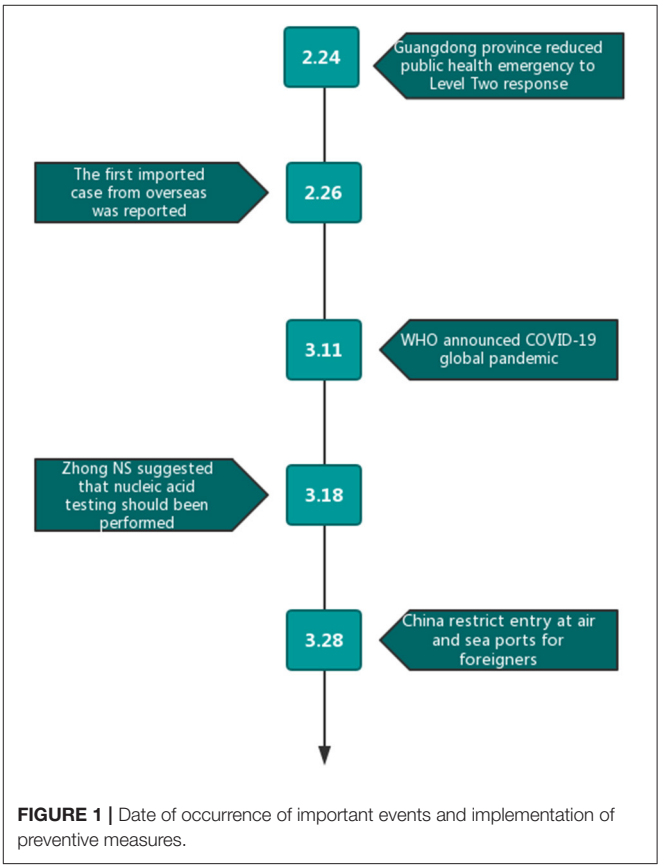


FIGURE 1 | Date of occurrence of important events and implementation of preventive measures.

MATERIALS AND METHODS

Study Design

After strict prevention and treatment of COVID-19, on February 29, 20 provinces of China mainland announced no domestic COVID-19 cases (5). However, the confirmed cases outside of China continued to increase, and by March 11, WHO announced COVID-19 as a global pandemic (4). As a result, China's epidemic has reversed, changing from the exporting country to the importing one. The first imported case from Iran was reported on February 26, and China was compelled to temporarily suspend the entry of foreign passport holders with valid visas or residence permits on March 11 to curb the spread of COVID-19. The implementation date of critical public health prevention initiatives is displayed in **Figure 1**. This study was designed to review the epidemiological characteristics and outcomes of imported cases from abroad (from February 26 to April 4). The study was approved by the Review Board of the Air Force Medical University. Case enrollment needs to meet the following conditions: (1) imported from abroad; (2) underwent quarantine at designated places; (3) diagnosed by the results of nucleic acid test; and (4) adequate clinical information and available follow-up data.

Data Collection

We collected the departure time and the country or region, the time of entry, symptoms, time of symptom onset, confirmed age and gender from the news reports and press releases reported by National Commission and Chinese Center for Diseases Control and Prevention. The degree of severity and diagnostic criteria refer to the 7th edition of the National New Coronavirus Pneumonia Diagnosis and Treatment Program. The data collection was performed by doctor Zhu, doctor Zhang, and doctor Jia, and any major disagreement among these three doctors was checked by the other two reviewers (doctor Xu and doctor Lei).

Statistical Analysis

Due to lack of definite exposure time, we defined incubation period as from the date of leaving epicenter to date of symptom onset, and the interval of symptom onset to final diagnosed time was defined as diagnostic time (12). We also defined confirmed time as from date of leaving epicenter to date of final diagnosis. The incubation period and diagnostic time were estimated by fitting a Weibull distribution on the date of leaving epicenter and symptom onset. Categorical variables were summarized as numbers and percentages. *T* test were performed to compare the differences of incubation period among groups, and when the cases were not normal data distribution, Mann-Whitney U test or Kruskal–Wallis *H* test was used. It was considered statistically significant that $P < 0.05$. SPSS 20.0 was used to finish the above analyses.

RESULTS

Clinical Characteristics

As of April 4, 2020, 1 week after China suspended the entry of foreign passport holders with valid visas or residence permits (February 28, 2020), cumulative imported cases from abroad increased to 913 cases, with 11 domestic cases related to imported cases. Finally, a total of 670 imported cases (Chinese: 555 cases and foreigners: 115 cases) were enrolled in this study. Europe was the main exported area (59.4%), followed by North America (20.9%) and South America (1.2%) (**Table 1**). The distribution of cases in exported countries and entry province, autonomous region or municipality was shown in **Figure 2**, and we show that United Kingdom and United States were main exported countries, and the main imported cities were metropolises, such as Shanghai and Guangzhou. For these cases, 269 cases were male, 209 were female, and 192 cases were unknown. Among 458 patients with clear age information, the oldest patient was 71 years old, the youngest was only 2 months old, and the median age of the patients was 30 years. **Figure 3** showed that patients aged between 20 and 49 years accounted for 69.2% (317/458); the main reason may be that a large proportion of the overseas imported cases were students studying abroad (32.8%). The detailed information of the patients is listed in **Table 1**.

TABLE 1 | Clinical characteristics of imported patients of COVID-19.

Characteristics	Cases	Proportion
Age, (median: 30 years, range from 2 month to 71 years)	*N* (670)	(%)
Gender	(null: 192)	28.7%
Male	269	40.2%
Female	209	31.2%
Age[a]	(null: 212)	31.6%
<30	228	34.0%
≥30	230	34.3%
Time of entry		
Before the global explosion	48	7.2%
Global explosion	7	1.0%
After the global explosion	615	91.8%
Nationality		
Foreigner	115	17.2%
Chinese	555	82.8%
Region of cases exported		
Asia	100	14.9%
Africa	24	3.6%
Europe	398	59.4%
South America	8	1.2%
North America	140	20.9%
Occupation		
International student	220	32.8%
Non-international students	450	67.2%
Symptoms		
Fever	107	16.0%
Respiratory symptoms	65	9.7%
Gastrointestinal symptoms	4	0.6%
Other symptom	26	3.9%
Without symptom	39	5.8%
Symptoms unknown	429	64.0%

[a] *Age: median age of all cases.*

Public Health Interventions

Since the first imported case from Iran was reported on February 26, China began to be concerned about the imported cases from abroad, with the reversal of domestic and foreign epidemics; on March 11, WHO announced COVID-19 as global pandemic. March 11 was an obvious watershed, the confirmed cases before and after were 28 cases and 642 cases, respectively. With the increase of confirmed cases, on March 18, Zhong Nanshan, a top Chinese epidemiologist, suggested that all entry population from epicenter should undergo nucleic acid testing, and 1 week after nucleic acid test was performed, the daily newly increased imported patients reached a peak with 50 cases. To more powerfully curb the spread of COVID-19, on March 28, China was compelled to temporarily suspend the entry of foreign passport holders with valid visas or residence permits. Apparently both arrived cases and confirmed cases had significant decreased, 6 days after implement of controlled measures, on April 4, and the daily new confirmed cases were reduced to 13 cases. The public health intervention events and related dynamic results are exhibited in **Figure 4**.

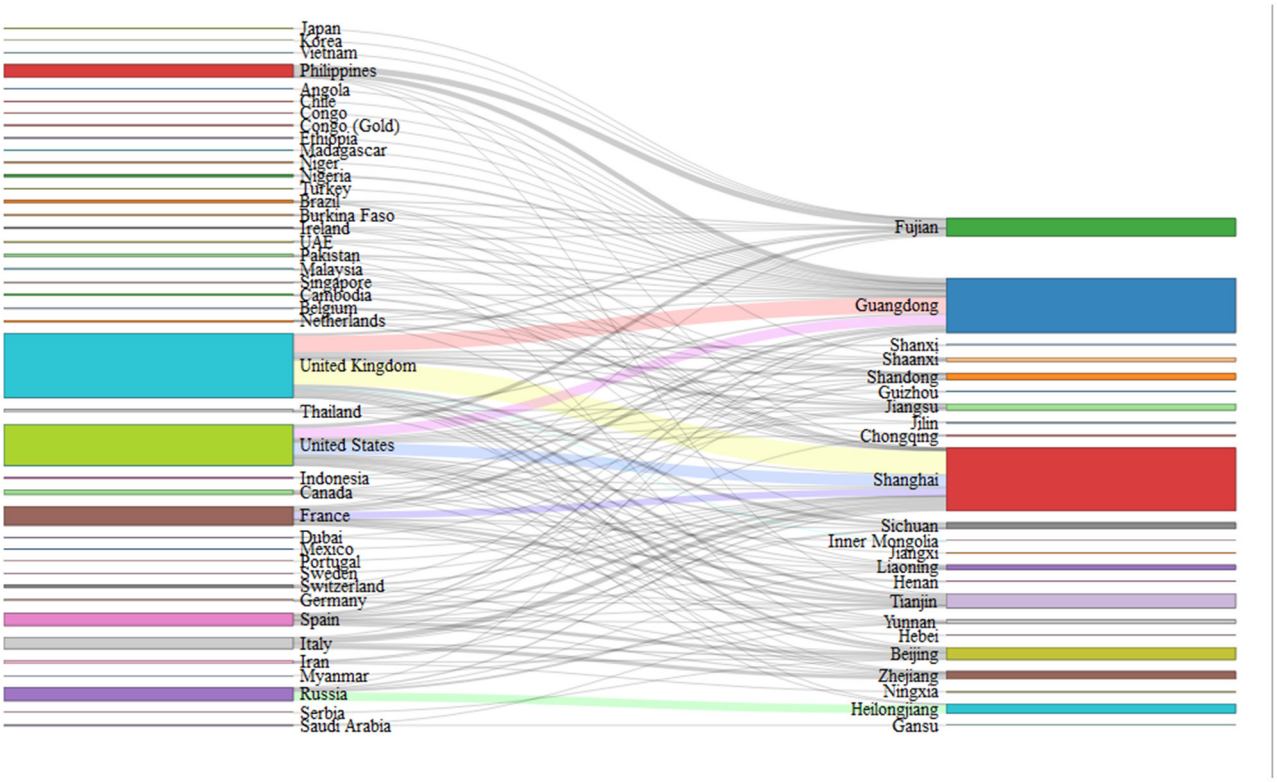


FIGURE 2 | Distribution of exported countries and distribution of imported provinces.

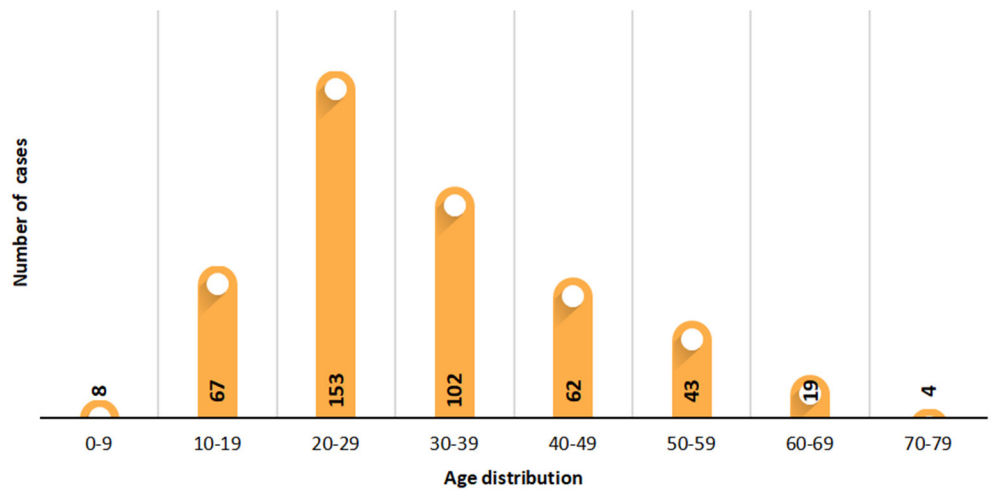


FIGURE 3 | Age distribution of imported cases from abroad.

Epidemiological Characteristics

For symptomatic patients, 37 cases had symptoms before departure, 17 patients had symptoms on the day of departure, and 180 patients had symptoms after leaving the epicenter. About 84.3% of patients (166/197) presented symptoms 1 week after leaving the epicenter, and 7 patients (3.6%) had symptoms 2 weeks after leaving the epicenter. Interesting, the peak of symptom onset emerged was on the first day (42 patients). Incubation period was estimated by fitting Weibull distribution, and patients who presented symptoms before departure had been excluded from the analysis. The median incubation period was 3.0 days (interquartile range, 1.0 to 6.0), and the 95th percentile

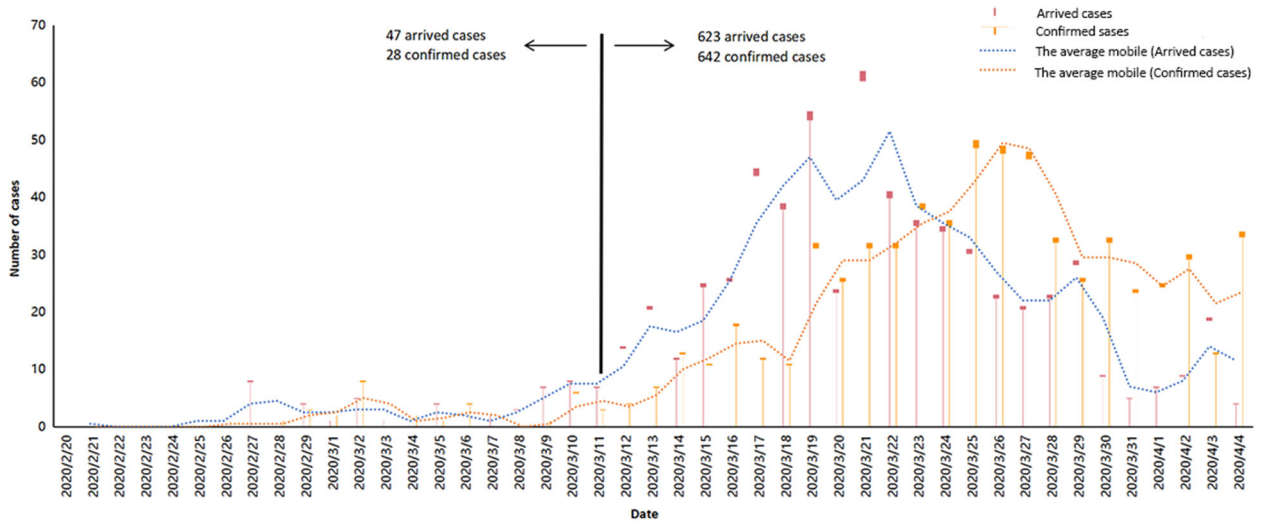


FIGURE 4 | Time distribution of confirmed cases from abroad.

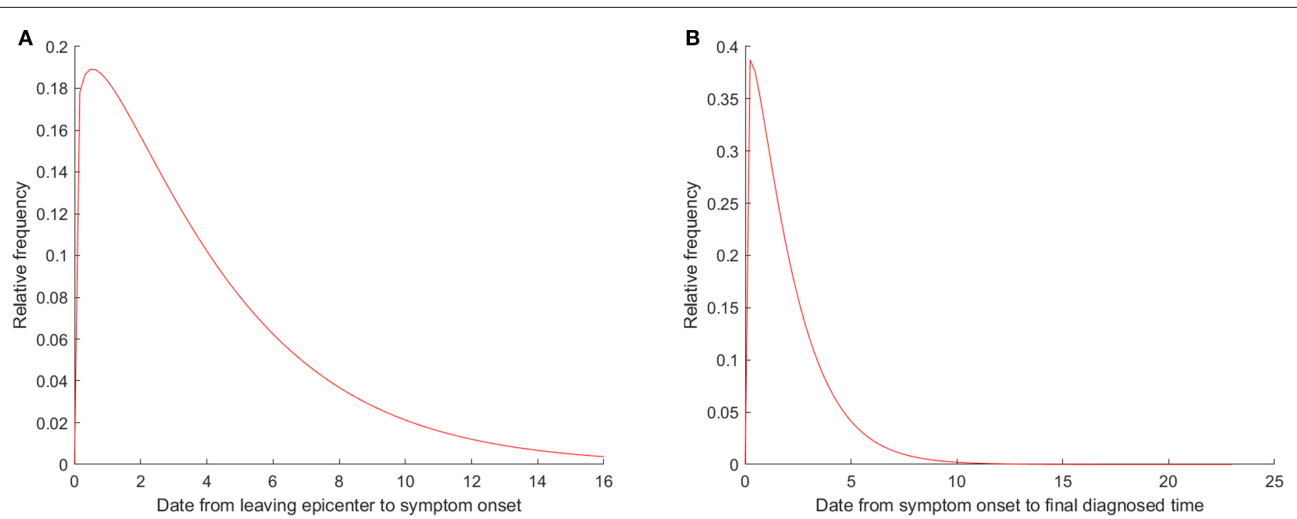


FIGURE 5 | Estimation of incubation period and diagnostic time by fitting Weibull distribution. **(A)** Incubation period estimated on the date of leaving epicenter; **(B)** Diagnostic time were estimated on the date of symptom onset.

was 11.6 days (**Figure 5A**). Moreover, cases exported from Africa presented symptoms later than patients exported from other regions ($P = 0.007$), which may due to the outbreak time varied in different regions (**Table 2**). Studies related to the incubation period are listed in **Table 3**.

Furthermore, we also researched the serial interval of symptom onset to final diagnosed (positive results of nucleic acid test). Two patients presented symptoms after final diagnosed; most cases (92.9%) were detected positively of nucleic acid after symptom onset with 4 days. For symptomatic patients, the median diagnostic time was 2.0 days (interquartile range, 1.0 to 3.0), the 95th percentile of the distribution was 5.0 days, it was showed in **Figure 5B**, furthermore, age and occupation were associated with diagnostic time ($P = 0.017$ and $P = 0.044$) (**Table 2**).

Finally, the median confirmed time of all patients was 4.0 days (interquartile range, 2.0 to 6.0). It was worth noting that about 5.8% patients were healthy carriers, and the median confirmed time of asymptomatic patients was 4.0 days (interquartile range, 2.0 to 9.0). **Figure 6** showed that the peak of confirmed time was on the third day (139 cases); there were still 37 patients (5.5%) who were final diagnosed 2 weeks after leaving the epicenter. The following variables might be associated with confirmed time: symptom type ($P = 0.005$), exported regions ($P < 0.001$), and symptom onset time ($P < 0.001$) (**Table 4**).

TABLE 2 | Epidemiological characteristics of symptomatic patients imported from abroad.

Variable	Incubation period (days)	IQI[b]	*P*	Diagnostic time (days)	IQI[b]	*P*
In total (197)	3.0	(1.0, 6.0)		2.0	(1.0, 2.0)	
Gender			0.116			0.708
Male (87)	4.0	(2.0, 7.0)		1.0	(1.0, 2.0)	
Female (80)	3.0	(1.0, 5.0)		2.0	(1.0, 2.0)	
Age[a]			0.749			0.017
<27 (77)	4.0	(2.0, 6.0)		1.0	(0.0, 2.0)	
≥27 (84)	4.0	(2.0, 6.0)		2.0	(1.0, 2.0)	
Occupation			0.570			0.044
International student (71)	4.0	(1.0, 6.0)		1.0	(1.0, 2.0)	
Non-international students (126)	3.0	(1.0, 6.0)		2.0	(1.0, 2.25)	
Symptoms			0.067			0.307
Fever (89)	4.0	(2.0, 6.0)		2.0	(1.0, 2.0)	
Respiratory symptoms (41)	2.0	(1.0, 6.0)		1.0	(1.0, 2.0)	
Gastrointestinal symptoms (3)	3.0	(2.5, 3.0)		1.0	(1.0, 1.5)	
Other symptom (17)	4.0	(1.0, 7.0)		1.0	(0.5, 2.0)	
Symptoms unknown (47)	1.0	(1.0, 5.0)		2.0	(1.0, 2.0)	
Nationality			0.805			0.815
Foreigner (27)	2.0	(1.0, 8.0)		1.0	(1.0, 2.0)	
Chinese (170)	3.0	(1.0, 6.0)		2.0	(1.0, 2.0)	
Region of cases exported			0.007			0.309
Asia (33)	2.0	(0.0,5.0)		2.0	(1.0, 3.0)	
Africa (7)	8.0	(3.0,14.0)		1.0	(1.0, 2.0)	
Europe (114)	3.0	(1.0,5.0)		1.5	(1.0, 2.0)	
South America (3)	5.0	(5.0,5.5)		1.0	(1.0, 2.0)	
North America (40)	4.0	(2.0,6.0)		2.0	(1.0, 3.0)	

[a]*Age, median age of symptomatic cases.*[b]*IQI, interquartile intervals.*

DISCUSSION

With the global outbreak of COVID-19, the confirmed cases and deaths continue to rise globally (26–29). In China, the domestic epidemic was controlled effectively but the challenge of cross-border COVID-19 transmissions emerged. To our knowledge, this study is the first designed to review the epidemiological characteristics of COVID-19 cases imported from abroad in China. Most of them were exported from Europe and North American, which might be related with the economic and academic exchanges between China and these regions.

Internationally, according to statistics, about 145 million tourists traveled to China from abroad in 2019 (30), and most of them came from Europe and North America. We also know that Europe and North America were the most popular areas for Chinese students (31). The distribution of exported regions of COVID-19 was associated with above data, and in this study, about 74.3% cases came from Europe and North America and nearly one-third of exported cases were Chinese overseas students. As exported country, previous studies have shown most cases from China had spread to Asian neighbors, followed by Europe (32). In our study, although there were few cases (4.8%) exported from Africa and South America, the imported cases in these two regions could not be ignored. On February 26, the first imported case from Iran was reported; after that, the confirmed increased steadily, as on March 11, COVID-19 was defined as a global pandemic, and the imported cases increased sharply.

Incubation period, from the earliest exposure time to the time of symptoms onset, was necessary for preventing and controlling the epidemic, so numerous studies were focused on this issue. The median incubation period was ranged from 3.9 days to 8.6 days due to the different definition of exposure time (13, 14, 16–19, 21–25, 33, 34). However, in a real world study, most patients could not provide the specific exposure time. It was significant to estimate the incubation period by using the available information. Jing et al. (12) estimated the median incubation period of COVID-19 was 8.13 days (95% CI: 7.37–8.91), and they defined it as from the time of departure from Wuhan City to the symptom onset by using the well-known renewal theory in probability. For the imported cases, preventive measures were formulated based on the incubation period; unfortunately, we could not obtain the definite exposure time, and we supposed that the time of leaving epicenter could predict the progression of COVID-19. In this study, we found that the median incubation of the imported COVID-19 cases was 3.0 days (interquartile range, 1.0 to 6.0) after they leaving the epicenter, shorter than the previous study. The following reasons can explain its short incubation period: a minority of them were infected on the way to China, or the mutations in the novel coronavirus (35). Notably, there were still 3.6% of patients who had symptom 2 weeks after leaving the epicenter, which indicated that the 2-week quarantine period might need to be prolonged. The outbreak of sporadic epidemics since mid-2020 were related to the onset of imported cases after the 2-week quarantine period; therefore, the Chinese government had formulated a 2-week plus 7-day quarantine policy for overseas cases.

We also explored the optimal nucleic acid detection time for imported cases in this study. For the symptomatic patients, the median diagnostic time was 2.0 days after they presented symptoms. A study from South Korea (20) showed that the median of symptom-onset to diagnosis was 5.2 days, the longest time of symptom-onset to diagnosis was 16.0 days reported by Xiao, and a total of 301 patients were analyzed (36). In this study, the positive results of nucleic acid test were detected earlier than the related studies, which might be due to that all inbound population from epicenter were suggested to underwent nucleic acid test when they arrived in China.

Moreover, our experience of fighting the COVID-19 indicated that public health intervention could reduce its transmission (37). The public health intervention, including strictly enforced segregation and travel bans, entry screening, suspension of public transportation in the city, closure of

TABLE 3 | Study list of incubation period of COVID-19 estimation.

References	*N*	Distribution	Mean, days (95% CI)	Median, days (95% CI)	Percentiles
Zhang et al. (13)	49	Log-normal	5.2 (1.8, 12.4)	-	95th: 10.5
Li et al. (14)	Null	Gamma	7.2 (6.8, 7.6)	-	-
Leung et al. (15)	175	Weibull	(Travelers to Hubei) 1.8 (1.0, 2.7) (Non-travelers) 7.2 (6.1, 8.4)	-	(Travelers to Hubei) 95th: 3.2 (Non-travelers) 95th: 14.6
Qian et al. (16)	262	-	6.7	-	-
Ping et al. (17)	758(93)	Log-normal	-	8.06 (6.9, 9.4).	95th: 21.9
Qian et al. (16)	91	-	-	6 (3.0, 8.0)	-
Wang et al. (18)	483	Logarithm normal	7.4	7.0	-
Lu et al. (19)	265	Weibull	6.4 (5.3, 7.6)	-	5th: 1.0 95th: 13.1
Ki et al. (20)	28	-	3.9 (0, 15.0)	3.0	-
Guan et al. (21)	1,099	-	-	4.0 (2.0, 7.0)	-
Li et al. (14)	10	Log-normal	5.2 (4.1, 7.0)	-	95th:12.5(9.2, 18)
Backer et al. (22)	88	Weibull Gamma Lognormal	6.4 (5.6, 7.7) 6.5 (5.6, 7.9) 6.8 (5.7, 8.8)	-	95th:10.3 (8.6, 14.1) 95th:11.3 (9.1, 15.7) 95th:13.3 (9.9, 20.5)
Lauer et al. (23)	181	Outside mainland China ($n = 108$).	-	5.5 (4.4, 7.0)	95th: (2.1, 14.7)
Lauer et al. (23)	181	Inside mainland China ($n = 73$)	-	4.8 (4.2, 5.6)	95th: (2.5, 9.2)
Lauer et al. (24)	181	Lognormal	-	5.1 (4.5,5.8)	97.5th: 11.5 (8.2,15.6)
Linton et al. (25)	Excluding Wuhan residents ($n = 52$)	Lognormal Weibull Gamma	5.0 (4.2, 6.0) 5.4 (4.3, 6.6) 5.3(4.3, 6.6)	4.3(3.5, 5.1) 4.7(3.6, 5.8) 4.7(3.8,5.7)	95th:10.6 (8.5, 14.1) 95th:12.0 (9.8, 15.6) 95th:11.3 (9.2, 14.5)
Linton et al. (25)	Including Wuhan residents ($n = 158$).	Lognormal Weibull Gamma	5.6 (5.0, 6.3) 5.8 (5.2, 6.5) 6.0 (5.3, 6.7)	5.0(4.4, 5.6) 5.3(4.7, 6.0) 5.6(4.9, 6.4)	95th:10.8 (9.3, 12.9) 95th:11.0 (9.6, 12.9) 95th:11.7 (10.3, 13.4)
Jing et al. (12)	1,211	Weibull	8.6 (8.0, 9.23)	8.1 (7.4, 8.9)	90th:14.7 (14.0, 15.3) 99th:20.6(19.5, 21.6)
Han et al. (24)	59	Monte Carlo simulation	5.8	5.0	-

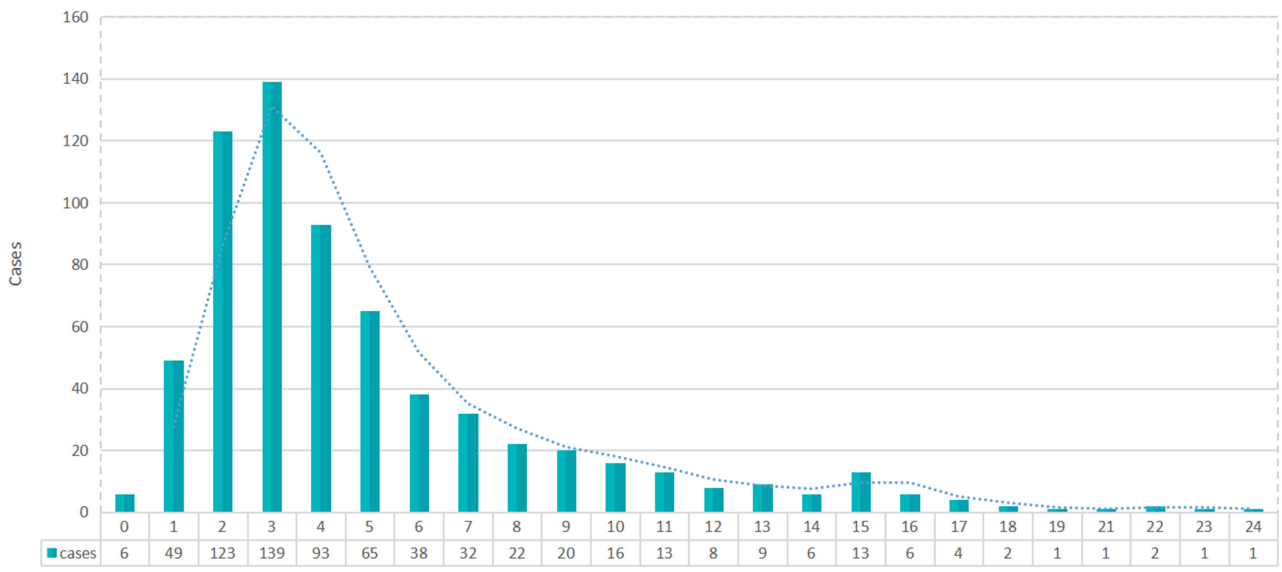


FIGURE 6 | Distribution of confirmed time after confirmed cases leave the epidemic area.

TABLE 4 | Confirmed time of imported COVID-19 patients.

Variable	Confirmed time (median, days)	IQI[b]	*P*
In total (670)	4.0	(2.0, 6.0)	
Gender			0.776
Male (269)	4.0	(2.0, 7.0)	
Female (209)	4.0	(2.0, 7.0)	
Age[a]			0.281
<30 (228)	3.0	(2.0, 6.0)	
≥30 (230)	4.0	(2.0, 7.0)	
Occupation			0.1
International student (220)	4.0	(3.0, 6.0)	
Non-international students (450)	4.0	(2.0, 6.25)	
Symptoms			0.005
Fever (107)	5.0	(3.0, 8.0)	
Respiratory symptoms (65)	3.0	(2.0, 6.0)	
Gastrointestinal symptoms (4)	4.0	(3.25, 4.75)	
Other symptom (26)	3.0	(2.0, 5.5)	
Without symptom (39)	4.0	(2.0, 9.0)	
Symptoms unknown (429)	4.0	(3.0, 6.0)	
Nationality			0.066
Foreigner (115)	4.0	(3.0, 8.0)	
Chinese (555)	4.0	(2.0, 6.0)	
Region of cases exported			<0.001
Asia (100)	4.0	(2.0, 7.0)	
Africa (24)	8.0	(4.25, 13.75)	
Europe (398)	3.0	(2.0, 5.0)	
South America (8)	8.5	(7.25, 9.75)	
North America (140)	4.0	(3.0, 7.0)	
Symptoms onset			<0.001
Symptom onset is clear (234)	4.0	(3.0, 7.0)	
Symptoms during isolation (79)	6.0	(4.0, 8.0)	
Asymptomatic (39)	4.0	(2.0, 9.0)	
Unknown time of symptoms (318)	3.0	(2.0, 5.0)	

[a]*Age, median age of all cases.* [b]*IQI, interquartile intervals.*

entertainment venues, and bans on public gatherings, played an important role in controlling the epidemic (38–40). One study proved that the Wuhan City shutdown delayed the spread of COVID-19 to other cities for 2.91days (95% CI: 2.54–3.29) (39). Therefore, on March 28 (11), China was compelled to temporarily suspend the entry of foreign passport holders with valid visas or residence permits. Apparently both arrived cases and confirmed cases had significantly decreased.

Finally, with asymptomatic patients, because the condition was hidden and might be contagious, we should be more vigilant. In our study, 39 asymptomatic patients were detected during their quarantine. A study about a health carrier from Wuhan caused widespread concern, she transmitted new coronavirus to her five family relatives, but she did not present any symptoms during her disease course (41). A series of subsequent studies confirmed that asymptomatic patients were infectious (42–45). The median of confirmed of asymptomatic patients was 4.0 days, basically consistent with symptomatic patients. Therefore, nucleic acid testing of all immigrants was essential for screening asymptomatic patients.

The shortcoming of this study was that the retrospective study did not include all people returning from overseas (243 cases were lost), so it could not represent the epidemiological characteristics of all imported confirmed cases; secondly, we estimated the incubation period for COVID-19; the departure time from the epidemic area was recognized as the exposure time, which may lead to a reduction in the real incubation period. Third, we could not quantify that interventions had multiple indirect effects on changes in human behavior. Finally, most patients were still in hospital, so the clinical outcomes had not been analyzed.

It is important to plan for the early stages of imported cases from overseas during a pandemic to avoid a national outbreak. In this study, we investigated the epidemiological characteristics of COVID-19 imported from overseas and observed interventions implemented by the government. We found that the departure time from epicenter could be used to predict the progression of the COVID-19. Promulgation of a ban on entry, a combination of quarantine, and quarantine methods was effective. For imported cases, the 2-week quarantine period might need to be prolonged and asymptomatic patients should be closely monitored.

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/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Tangdu Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TJ, JZha, and WuW participated in study design and study conception. JZhu, QZ, CJ, SX, JL, HW, and ZZ performed data analysis. JZhu, QZ, CJ, SX, JC, YX, WeW, XW, and MW recruited patients. JZhu, QZ, CJ, WuW, and JC drafted the manuscript. All authors provided critical review of the manuscript and approved the final draft for publication.

FUNDING

This research was supported by grants from the Wu Jieping Medical Foundation (320.6750.17527) and the Provincial Key R & D Program of Shaanxi (2017ZDCXL-SF-01-04-01).

REFERENCES

1. *Novel Coronavirus–China*. (2020). Available online at: https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/ (accessed January 12, 2020).
2. Hui DS, Azhar EI, Madani TA, Ntoumi F, Kock R, Dar O, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health-The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis.* (2020) 91:264–6. doi: 10.1016/j.ijid.2020.01.009
3. *Coronavirus Disease (COVID-2019) Situation*. Available online at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn$=$1ba62e57_10 (accessed March 11, 2020).
4. *Coronavirus Outbreak is a Pandemic, Says WHO*. Available online at: https://www.who.int/zh/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19-$-$11-march-2020 (accessed March 11, 2020).
5. *China Daily in March 28, 2020*. Available online at: https://www.chinadaily.com.cn/a/202003/28/WS5e7f2ba6a310128217282b8c.html (accessed March 28, 2020).
6. *Chongqing Center for Disease Control and Prevention*. Available online at: http://www.cqcdc.org/html/content/20/03/4617.shtml (accessed March 25, 2020).
7. *Shanxi Provincial Center for Disease Control and Prevention*. Avavailble online at: http://www.sxcdc.cn/Article/20196.html (accessed March 10, 2020).
8. *The Confirmed Cases of Imported New Coronary Pneumonia in Ningxia were Isolated After Arriving in Zhongwei City, and Have Been Intensively Treated on the 26th*. Available online at: https://wap.peopleapp.com/article/5208841/5110610 (accessed February 26, 2020).
9. *Press Conference of Guangzhou Municipal Government*. Available online at: http://www.gz.gov.cn/zt/gzsrmzfxwfbh/fbt/content/post_5744314.html (accessed March 18, 2020).
10. *National Health Commission of the People's Republic of China*. Available online at: http://www.nhc.gov.cn/xcs/yqtb/202003/c521093a01734df3b3fbc156064ba19f.shtml (accessed March 27, 2020).
11. *People's Republic of China Ministry of Foreign Affairs*. Available online at: https://www.mfa.gov.cn/web/zyxw/t1761858.shtml (accessed March 26, 2020).
12. Jing Q, You C, Lin QS, Hu TJ, Yu SC, Zhou XH, et al. Estimation of incubation period distribution of COVID-19 using disease onset forward time: a novel cross-sectional and forward follow-up study. *Sci Adv.* (2020). 6:c1202. doi: 10.1126/sciadv.abc1202
13. Zhang JJ, Litvinova M, Wang W, Wang Y, Deng XW, Chen XH, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *Lancet Infect Dis.* (2020) 20:93–802. doi: 10.1016/S1473-3099(20)30230-9
14. Li Q, Guan X, Wu P, Wang XY, Zhou L, Tong YQ, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med.* (2020) 382:1199–207. doi: 10.1056/NEJMoa2001316
15. Leung C. The difference in the incubation period of 2019 novel coronavirus (SARS-CoV-2) infection between travelers to Hubei and nontravelers: the need for a longer quarantine period. *Infect Control Hosp Epidemiol*. (2020) 41:594–6. doi: 10.1017/ice.2020.81
16. Qian GQ, Yang NB, Ding F, Ma AHY, Wang ZY, Shen YF, et al. Epidemiologic and clinical characteristics of 91 hospitalized patients with COVID-19 in Zhejiang, China: a retrospective, multi-centre case series. *QJM.* (2020) 113:474–81. doi: 10.1093/qjmed/hcaa089
17. Ping KK, Lei MY, Gou Y. Epidemiologic characteristics of COVID-19 in Guizhou, China. *J Infect Dev Ctries.* (2021). 15:389–97. doi: 10.3855/jidc.12818
18. Wang P, Lu J-a, Jin Y, Zhu M, Wang L, Chen S. Statistical and network analysis of 1212 COVID-19 patients in Henan, China. *Int J Infect Dis*. (2020) 95:391–8. doi: 10.1016/j.ijid.2020.04.051
19. Lu HZ, Ai JW, Shen YZ, Li Y, Li T, Zhou X, et al. A descriptive study of the impact of diseases control and prevention on the epidemics dynamics and clinical features of SARS-CoV-2 outbreak in Shanghai, lessons learned for metropolis epidemics prevention. *medRxiv [Preprint].* (2020). doi: 10.1101/2020.02.19.20025031
20. Ki M, Task Force for-nCo V. Epidemiologic characteristics of early cases with 2019 novel coronavirus(2019-nCoV) disease in Korea. *Epidemiol Health.* (2020) 42:e2020007. doi: 10.4178/epih.e2020007
21. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of coronavirus disease 2019 in China. *N Engl J Med.* (2020) 382:1708–20. doi: 10.1056/NEJMoa2002032
22. Backer JA, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus(2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Euro Surveill.* (2020) 25:2000062. doi: 10.2807/1560-7917.ES.2020.25.5.2000062
23. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng QL, Meredith HR, et al. The incubation period of coronavirus disease 2019(COVID-19)from publicly reported confirmed cases: estimation and application. *Ann Intern Med.* (2020) 172:577–82. doi: 10.7326/M20-0504
24. Han H. Exploring the association between compliance with measures to prevent the spread of COVID-19 and big five traits with Bayesian generalized linear model. *Pers Individ Dif.* (2021) 176:110787. doi: 10.1016/j.paid.2021.110787
25. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung S, et al. Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. *J Clin Med.* (2020) 9:538. doi: 10.3390/jcm9020538
26. Zhao S, Zhuang Z, Cao P, Cao PH, Ran JJ, Gao DZ, et al. Quantifying the association between domestic travel and the exportation of novel coronavirus(2019-nCoV)cases from Wuhan, China in 2020: a correlational analysis. *J Travel Med.* (2020) 27:taaa022. doi: 10.1093/jtm/taaa022
27. Bogoch, II, Watts A, Thomas-Bachli A, Huber C, Kraemer M, Khan K, et al. Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel. *J Travel Med.* (2020) 27:taaa008. doi: 10.1093/jtm/taaa008
28. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet.* (2020) 395:689–97. doi: 10.1016/S0140-6736(20)30260-9
29. Findlater A, Bogoch, II. Human mobility and the global spread of infectious diseases: a focus on air travel. *Trends Parasitol.* (2018) 34:772–83. doi: 10.1016/j.pt.2018.07.004
30. *Basic Situation of the Tourism Market in 2019*. Available online at: https://mp.weixin.qq.com/s/VQ4VVeV-rmiZoq8oAUWN9Q (accessed March 10, 2020).
31. *Overseas Chinese Students*. Available online at: http://zz.sll.cn/news/01083742.html (accessed October 31, 2019).
32. Haider N, Yavlinsky A, Simons D, Osman AY, Ntoumi F, Zumla A, et al. Passengers' destinations from China: low risk of Novel Coronavirus (2019-nCoV)transmission into Africa and South America. *Epidemiol Infect.* (2020) 148:e41. doi: 10.1017/S0950268820000424
33. Li ML, Chen P, Yuan QQ. Transmission characteristics of the COVID-19 outbreak in China: a study driven by data. *medRxiv.* (2020).
34. Tian S, Hu N, Lou J, Chen K, Kang XQ, Xiang ZJ, et al. Characteristics of COVID-19 infection in Beijing. *J Infect.* (2020) 80:401–6. doi: 10.1016/j.jinf.2020.02.018
35. Forster P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc Natl Acad Sci USA*. (2020) 117:9241–3. doi: 10.1073/pnas.2004999117
36. Xiao AT. Tong YX, Gao C. Dynamic profile of RT-PCR findings from 301 COVID-19 patients in Wuhan, China: a descriptive study. *SSRN.* (2020)127:104346. doi: 10.2139/ssrn.3548769
37. Iwata K, Miyakoshi C. A simulation on potential secondary spread of novel coronavirus in an exported country using a stochastic epidemic SEIR model. *J Clin Med.* (2020) 9:E944. doi: 10.3390/jcm9040944
38. Boldog P, Tekeli T, Vizi Z, Dénes A, Bartha FA, Röst G. Risk assessment of novel coronavirus COVID-19 outbreaks outside China. *J Clin Med.* (2020) 9:571.doi: 10.3390/jcm9020571
39. Tian H, Liu Y, Li Y, Wu C, Chen B, Kraemer M, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science.* (2020) 368:638–42. doi: 10.1126/science.abb6105
40. Zhou XK, Wu ZG, Yu RR, Cao SN, Fang W, Jiang Z, et al. Modelling-based evaluation of the effect of quarantine control by the Chinese government

in the coronavirus disease 2019 outbreak. *Sci China Life Sci.* (2020) 63:1–4. doi: 10.1007/s11427-020-1717-9
41. Bai Y, Yao L, Wei T, Tian F, Jin DY, Chen L, et al. Presumed asymptomatic carrier transmission of COVID-19. *JAMA*. (2020) 323:1406–7.doi: 10.1001/jama.2020.2565
42. Wolfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature.* (2020) 581:465–9. doi: 10.1038/s41586-020-2196-x
43. Lai CC, Liu YH, Wang CY, Wang Y-H, Hsueh S-C, Yen M-Y, et al. Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2(SARS-CoV-2): facts and myths. *J Microbiol Immunol Infect.* (2020) 53:404–12. doi: 10.1016/j.jmii.2020.02.012
44. Huang R, Xia J, Chen YX, Wang YH, Hsueh SC, Yen M, et al. A family cluster of SARS-CoV-2 infection involving 11 patients in Nanjing, China. *Lancet Infect Dis.* (2020) 20:534–5. doi: 10.1016/S1473-3099(20)30147-X
45. Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *N Engl J Med.* (2020) 382:970–1. doi: 10.1056/NEJMc2001468

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership