Uncovering two phases of early intercontinental COVID-19 transmission dynamics

Running Title: Intercontinental COVID-19 transmission dynamics

Jing Yang, Ph.D[1,7], Juan Li, Ph.D[2,7], Shengjie Lai, MD, MPH, Ph.D[3,7*], Corrine W. Ruktanonchai, Ph.D[3,4], Weijia Xing, Ph.D[2], Alessandra Carioli, Ph.D[3], Peihan Wang, MS[2], Nick W. Ruktanonchai, Ph.D[3,4], Ruiyun Li, Ph.D[5], Jessica R. Floyd, Ph.D[3], Liang Wang, Ph.D[1], Yuhai Bi, Ph.D, DVM[1,6*], Weifeng Shi, Ph.D[2*], Andrew J. Tatem, Ph.D[3]

[1]CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Center for Influenza Research and Early-warning (CASCIRE), CAS-TWAS Center of Excellence for Emerging Infectious Diseases (CEEID), Chinese Academy of Sciences

[2]Shandong First Medical University & Shandong Academy of Medical Sciences

[3]WorldPop, School of Geography and Environmental Science, University of Southampton

[4]Population Health Sciences, Virginia Tech, Blacksburg

[5]MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, School of Public Health, Faculty of Medicine, Imperial College London

[6]University of Chinese Academy of Sciences

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

© International Society of Travel Medicine 2020. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com

[7]These authors contributed equally.

Correspondence to:

Dr Shengjie Lai, WorldPop, School of Geography and Environmental Science, University of Southampton, Southampton SO17 1BJ, UK, shengjie.lai@soton.ac.uk;

Prof Yuhai Bi, CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Center for Influenza Research and Early-warning (CASCIRE), CAS-TWAS Center of Excellence for Emerging Infectious Diseases (CEEID), Chinese Academy of Sciences, Beijing 100101, China, beeyh@im.ac.cn;

Prof Weifeng Shi, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian 271000, China, shiwf@ioz.ac.cn.

Funding:

This work was supported by the National Key Research and Development Programme of China, Chinese Academy of Sciences, Key Research and Development Project of Shandong Province, National Natural Science Foundation of China, Shandong First Medical University, Bill & Melinda Gates Foundation, European Union Horizon 2020, and Taishan Scholars Programme of Shandong Province.

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

Abstract

Background The COVID-19 pandemic has posed an ongoing global crisis, but how the virus spread across the world remains poorly understood. This is of vital importance for informing current and future pandemic response strategies.

Methods We performed two independent analyses, travel network-based epidemiological modelling and Bayesian phylogeographic inference, to investigate the intercontinental spread of COVID-19.

Results Both approaches revealed two distinct phases of COVID-19 spread by the end of March 2020. In the first phase, COVID-19 largely circulated in China during mid- to late January, 2020, and was interrupted by containment measures in China. In the second and predominant phase extending from late February to mid-March, unrestricted movements between countries outside of China facilitated intercontinental spread, with Europe as a major source. Phylogenetic analyses also revealed that the dominant strains circulating in the United States of America were introduced from Europe. However, stringent restrictions on international travel across the world since late March have substantially reduced intercontinental transmission.

Conclusions Our analyses highlight that heterogeneities in international travel have shaped the spatiotemporal characteristics of the pandemic. Unrestricted travel caused a large number of COVID-19 exportations from Europe to other continents between late February and mid-March, which facilitated the COVID-19 pandemic. Targeted restrictions on international travel from countries with widespread community transmission, together with improved capacity in testing, genetic sequencing and contact tracing, can inform timely strategies for mitigating and containing COVID-19 outbreaks post-lockdown.

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

Introduction

The coronavirus disease 2019 (COVID-19), caused by the highly contagious causative agent severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was first identified in Wuhan, Hubei province, in late 2019.[1,2] The World Health Organization declared COVID-19 a pandemic on March 11, 2020. As of October 1, 2020, there have been more than 33 million confirmed COVID-19 cases with 1.01 million fatalities worldwide, affecting more than 200 countries, territories or areas.[3] Non-pharmaceutical interventions are currently the only viable strategies available to contain and suppress COVID-19 spread, [4-6] especially for the travel-associated spread between and within countries.[7-13] For example, Wuhan's lockdown on January 23, 2020 delayed the growth and limited the size of the COVID-19 outbreak in China.[7-9] These travel restrictions also had an effect in reducing international transmission of SARS-CoV-2, with a 77% reduction in cases exported from mainland China to other countries.[10,11] Accordingly, epidemics triggered by international importations in multiple countries have been suppressed by substantial travel and social distancing interventions.[12,13] However, widespread community transmission of SARS-CoV-2 was still reported in some regions after easing these measures, with an increase in confirmed cases or subnational localized resurgences in the United States of America (USA), India, Japan, and some European countries since late June.[14,15]

Studies have attempted to quantify the spread risk of COVID-19 using epidemiological models within a country and at regional scales,[5,10,16-18] and accumulated divergences among genomic data of SARS-CoV-2 have also been used to trace the viral spread. For instance, it was found that most of the SARS-CoV-2 strains circulating in the New York area and Northern California were introduced from Europe via multiple independent importation events.[19,20] Similarly, the great majority of the SARS-CoV-2 strains circulating in UK were derived from

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

continental Europe in March via inbound travellers.[21] The effects of international travel and border control measures on the global spread of SARS-CoV-2 were also preliminarily explored.[10,22] However, questions remain as to the spatiotemporal characterizations and transmission routes of the COVID-19 pandemic, and which factors contributed to the seeding of the virus and the emergence of outbreaks. Answers to these questions are of vital importance in formulating effective measures to contain the ongoing COVID-19 pandemic and for future infectious disease outbreak planning.

Using a comprehensive dataset integrating epidemiological, travel, intervention, and genetic data, here we conduct epidemiological modelling and Bayesian phylogeographic inference to: 1) understand the changing patterns of international movements under COVID-19 interventions across the world, 2) measure the transmission dynamics of intercontinental COVID-19, 3) distinguish the main source of international transmission that facilitated the COVID-19 pandemic, and further 4) reveal the source of the dominant strains circulating in the USA. The findings of our study can be used to inform public health response efforts for ongoing and future waves across the world.

Methods

Simulating the COVID-19 spread using epidemiological model

Using domestic and international population mobility data and a susceptible-exposed-infectious-removed (SEIR) framework,[6] we built a global travel network-based stochastic metapopulation transmission model to simulate COVID-19 transmission across 221 countries, territories or areas globally from December 1, 2019 to March

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

31, 2020 (the R code of the SEIR model can be found at: https://github.com/wpgp/BEARmod). To initially parameterize the model, country-specific reproduction numbers (median 2.4, interquartile range [IQR] 2.0–2.8), before the implementation of physical distancing interventions, were estimated from daily case counts reported by each country, adjusted for reporting delays.[23] A publicly-available dataset of government COVID-19 countermeasures was obtained to define the timings of various interventions implemented in each country.[24] We also used initial epidemiological information (i.e. the incubation period and delays from symptom onset to report) estimated from case data during the early stages of the outbreak in Wuhan.[25]

Additionally, two mobile phone-derived publicly available anonymized population mobility datasets covering 133 countries, territories or areas in 2020 [26,27] were used in our simulations to account for the impact of travel and physical distancing interventions implemented to mitigate the COVID-19 pandemic across space and time. To simulate the international spread of COVID-19 through population movement, we parameterized the model using global air traffic data from December 1, 2019 to March 31, 2020, obtained from the Official Aviation Guide (https://www.oag.com/). As we simulated the COVID-19 spread by day, we used the proportion of daily volume of seats on scheduled flights over the total number of flight seats in each month and route to disaggregate air passenger counts from monthly to daily level between countries. Details of the model, parameters, and data sources and their collation are provided in the Supplementary Materials and Methods.

Bayesian phylogeographic inference of SARS-CoV-2

All available complete or near-complete genomic nucleotide sequences of SARS-CoV-2 were collected from GISAID (https://www.gisaid.org), NCBI GenBank (https://www.ncbi.nlm.nih.gov), NMDC (http://nmdc.cn/coronavirus), and NGDC

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

(https://bigd.big.ac.cn/ncov/) database on April 1, 2020. The latest collection date of the sequences was March 26, 2020. We excluded duplicated sequences, genomes with incomplete collection date, and low-quality sequences. We kept one representative of any known epidemiologically-linked case clusters (http://virological.org/t/phylodynamic-analysis-176-genomes-6-mar-2020/356). Further, we removed sequences isolated from Africa and South America due to a small number of available sequences as of April 1, 2020. We pooled the remaining genomes into five geographic regions, including China, the Asian countries minus China, North America, Europe, and Oceania. 14 jackknife resampled datasets were generated for phylogeographic analyses by randomly sampling at most 10 isolates per location per day. Each down-sampled dataset comprised a total of 1030 sequences, including China (n=233), the Asian countries without China (n=118), North America (n=311), Europe (n=296), and Oceania (n=72). Virus sequences were aligned with default parameters in MAFFT v7.[28]

To investigate the global phylodynamics of SARS-CoV-2, phylogeographic inferences were performed using the asymmetric discrete trait analysis in BEAST v1.10.4 (Supplementary Materials and Methods).[29] We employed the general time reversible (GTR) nucleotide substitution model and a model to account for rate heterogeneity across sites by combining the discrete gamma distribution with four rate categories and invariable-sites. A strict molecular clock was utilized to model sequence evolution. Two tree priors, exponential growth and Bayesian skyline, were employed to describe virus population dynamics. By evaluating the marginal likelihood and evolutionary rate estimates, the exponential growth model was selected in subsequent analyses. Also, we set up Markov jump count parameters to obtain the frequency of transitions between location traits along phylogenetic branches over time

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

(https://beast.community/markov_jumps_rewards). Analyses with the same parameter configuration were performed on 14 subsampled datasets to assess the robustness of phylogeographic inferences. We performed each analysis with 100 million iterations, later removing an appropriate burn-in (10%–40% of samples) to get an adequate effective sample size (basically ESS ≥100). The maximum clade credibility (MCC) tree was calculated from the posterior tree space in TreeAnnotator v1.10.4.

Phylogenetic and bioinformatics analyses of worldwide SARS-CoV-2

To investigate which lineage of SARS-CoV-2 contributed to the epidemic in different locations, we performed a maximum likelihood analysis of the global SARS-CoV-2 strains (n = 5482 and updated on April 14, 2020 in GISAID) using RAxML v8.2.9[30] with 100 bootstrap replicates and the GTR model. We classified global SARS-CoV-2 into two lineages based on two highly linked single nucleotide polymorphisms (SNPs): type S (8782T and 28144C) and type L (8782C and 28144T).[31,32]

Based on the maximum likelihood analysis and the specific nucleotide mutations highlighted on the Nextstrain website (https://nextstrain.org/), we proposed a simple classification frame that was able to capture the major clades of global SARS-CoV-2. In detail, the lineage L was further classified into L1(241C, 3037C, 14408C, 23403A) and L2 (241T, 3037T, 14408T, 23403G) (Supplementary Materials and Methods). The lineage L2 was further classified into L2.1 (28881–28883AAC), and L2.2 (25563T), which has a special sub-lineage L2.2.1 (1059T). Analogously, type S was further divided into S1 (18060C, 17858, and 17747) and S2 (18060T, 17858G, and 17747T). The numbering of nucleotide sites is referred to the NCBI reference sequence of SARS-CoV-2, *NC_045512*. Further, SARS-CoV-2 strains isolated from the USA by

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

May 20, 2020 were classified into different lineages. The distributions of different virus lineage within each state of the USA were summarized.

Results

Changing patterns of international travel

Compared with the levels of air travel between December 1, 2019 and January 22, 2020, international outbound travel from China following the lockdown of the country rapidly declined to a median of 16% (IQR 9%–55%). However, international air travel from countries outside of China did not significantly change in February, and reductions for European and Middle Eastern countries began to appear in early March. During the week of March 25–31, 2020, international air travel across the world reduced to a median of 37% (IQR 18%–63%) of the levels seen before the pandemic (Supplementary Figures 1, 2).

Two distinct phases of intercontinental spread

Two transmission phases of international exportations of COVID-19 were identified by both the epidemiological model and Bayesian phylogeographic inference (Figure 1). During the first phase in mid- to late January, 519 cases (95% credible interval [CI]: 369–682) were estimated to have been exported from China to other countries, with most of them (84%) occurring before February 1, 2020. However, three weeks later, more intercontinental exportations of COVID-19 occurred in the second phase, from late February to mid-March (Figure 1A and Supplementary Figure 3). We estimated that there were 5563 cases (95%CI: 2714–8627) exported from Asian countries, except China, to other continents by March 31. Europe was estimated to have the highest number of exported cases (17087, 95%CI: 11342–23136) to other continents, with >90%

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

of exported cases occurring in March. North America and Oceania were estimated to have exported 3462 cases (95%CI: 2169–4817) and 1428 cases (95%CI: 662–2231) to other continents, respectively. The decline of this transmission phase corresponded with the international travel reductions implemented in mid- to late March (Supplementary Figures 1, 2). Additionally, our estimated epicurves were consistent with the patterns seen for reported curves as of March 31, 2020 ($p<0.001$, $R^2=0.96$), and high correlations were also found between the reported data and the estimated incidence by country ($p<0.001$, $R^2=0.70$) and estimated imported cases from China ($p<0.001$, $R^2=0.65$), respectively (Supplementary Figures 4, 5).

We summarized the Markov jumps between defined locations in the phylogenies of the posterior tree space to describe the intercontinental transmissions of SARS-CoV-2, and the inferred phylodynamics of SARS-CoV-2 from 14 sub-sampling datasets are consistent and robust. Consistent with results from the SEIR modelling (Figure 1A), two peaks of global SARS-CoV-2 transmission were also identified by Bayesian phylogeographic inference (Figure 1B). The first peak of virus dissemination occurred during mid- to late January 2020, but shortly after the Wuhan shutdown, viral transmission from China reduced dramatically. The second peak, which was higher than the first, occurred around early to mid-March and decreased rapidly after March 15. The Bayesian inference also highlighted that China was the major contributor to the first spread peak, while multiple regions contributed to the second spread peak, with Europe being the dominant source.

Major drivers of the second phase facilitating the COVID-19 pandemic

We further studied the viral importation events for major continental regions. The network-based SEIR modelling revealed two distinct peaks of COVID-19 importation events in most continents

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

by March 31, 2020 (Figure 2). We estimated that there were 450 cases (95%CI: 338–568) exported from China into other Asian countries in the first phase and 11015 cases (95%CI: 7008–15235) from other continents into Asia in the second phase. Europe had 28 estimated cases (95%CI: 9–50) imported from China in the first phase and 5664 cases (95%CI: 2759–8832) from other continents in the second phase. Similarly, most of the imported cases in North America were from other continents (5559, 95%CI: 3447–7834) in the second phase, with only 23 cases (95%CI: 14–31) imported from China. Meanwhile, Africa, South America, and Oceania were estimated to have had 3316 (95%CI: 1846–4948), 2077 (95%CI: 1299–2909), and 1114 (95%CI: 677–1584) cases, respectively, imported from other continents in the second phase, with few cases exported from China into Africa (5 cases, 95%CI: 1–12), South America (1 case, 95%CI: 0–2), and Oceania (13 cases, 95%CI: 7–20). However, travel restrictions implemented in mid- to late March have significantly curbed the international spread of COVID-19 since the end of March.

The Bayesian inference analyses also found that viral exportations from China peaked during mid- to late January, and dropped sharply after the shutdown of Wuhan (Figure 3A). In contrast, exportations from Europe rapidly increased from mid-February and peaked in early March. During this period, the virus migration frequency from Europe to North America and Oceania was estimated to have increased until the WHO declared COVID-19 a pandemic on March 11 (Figure 3B). Additionally, virus diffusions from Europe to North America were likely to have already occurred before travel restrictions on European citizens were implemented by the USA on March 13, 2020. However, since March 26, 2020, shortly after the second transmission peak, the USA has been the country with the highest reported number of COVID-19 cases.

Origins of the dominant strains circulating in the USA

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

To further understand the impact of international travel on the introduction of SARS-CoV-2 in the USA, we further simulated importation events into the USA. Our epidemiological SEIR model revealed that most of the cases imported into the USA (as of March 31, 2020) likely originated from Europe (2673 cases, 95%CI: 1951–3401), other North American countries (1324 cases, 95%CI: 924–1730), and Oceania (192 cases, 95%CI: 89–297) (Figure 4A). However, only a limited number of cases (18, 95%CI: 12–23) were imported from China, with 685 cases (95%CI: 336–1040) coming from other Asian countries (Figure 4A).

Clearly, several SARS-CoV-2 variants of both S and L genotypes have been co-circulating in the USA (Figure 4C). The clade L2.2.1, characterized by 241T, 3037T, 14408T, 23404G, 25563T and 1059T, was dominant (approximately 51.07% of the USA strains available in GISAID as of May 20, 2020) in most of the States of the USA, including at least seven out of the top ten states that have reported the most COVID-19 cases as of May 20, particularly in New York (78.84%), New Jersey (84.00%), and Michigan (77.37%) (Figure 4C and Supplementary Table 1). In 13 out of the 14 MCC trees, L2.2.1 was predicted to have originated from Europe in late February (Supplementary Figure 6). Similarly, recent research also revealed that most of the SARS-CoV-2 strains circulating in the New York area and Northern California were introduced from Europe.[20,33] Apart from L2.2.1, there were also independent introductions from Europe to North America in other clades (Supplementary Figure 6). Although an S genotype variant, S2 (characterized by 8782T, 28144C and 18060T), was dominant (54.94%) in Washington state and predicted to have originated from China, it accounted for a minority of the confirmed SARS-CoV-2 cases in the USA (Supplementary Figure 7, 8 and Supplementary Table 1).[34] The dominant SARS-CoV-2 strains contributing the community transmission in the USA originated from Europe and were introduced before the implementation of travel restrictions to European

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

citizens, followed by extensive domestic transmission.

Discussion

Using comprehensive and integrated datasets on population mobility, interventions, case reporting, and virus genomes, we have conducted both epidemiological modelling and Bayesian phylogeographic inferences to uncover the intercontinental dynamics of COVID-19 transmission and virus evolution. The findings from the two independent methodologies were consistent, and both revealed two phases of COVID-19 transmission across the globe. In the first phase, the virus mainly circulated and propagated in China, and our modelling only found a small number of COVID-19 cases transmitted from China to other localities. This first phase was interrupted by strict containment measures implemented proactively across China and reduced outbound travel.[5,25]

However, in the second intercontinental dispersal phase, the seeding of viruses was exaggerated by population movements between countries outside of China from late February through mid-March, which finally led to the global pandemic. In particular, our study identified Europe as an epicenter of this phase through phylogenetic analyses, which is consistent with other studies that revealed the dominant strains circulating in the USA and the UK before travel restrictions were imported from continental Europe where community transmission was widespread.[19,21,35,36] The USA, as the most-affected country, had substantial airline network connections to Europe but imposed relatively late travel restrictions on European travellers. Our phylogenetic analyses revealed the dominant strains in seven out of ten states with the highest case numbers were imported from Europe in the second phase, rather than from China in the first phase, which was in accordance with the results from epidemiological modelling and further highlighted the role of international travel from Europe in SARS-CoV-2 spread into the USA and

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

across the globe.

More importantly, apart from Europe, several other regions outside of China, including North America and some Asian countries, were also estimated to have contributed to the international spread of SARS-CoV-2. However, the reductions in outbound travel from China since late January, as well as strict restrictions on international travel across the world since late March, have effectively reduced international spread of the virus.[6-7,10,22,25] However, many countries where COVID-19 had been contained or eliminated are now facing increasing risks of international importations after relaxing travel and social distancing interventions. For example, more than 2,000 internationally imported COVID-19 cases have been documented in China as of July 25, 2020.[37] Although strict testing, quarantine and contact tracing measures have been implemented, some cases have caused new outbreaks in different regions of China, including Harbin, Shulan, Beijing, Dalian, and Urumqi in May – July, 2020, with resurgences also reported in the USA, India, Japan, and several European countries.[3,37-40] The relaxation of travel restrictions and social distancing particularly in the summer holiday period and with the reopening of schools, might have contributed to the resurgence and ongoing outbreaks of COVID-19 in Europe. The potential impact of restoring domestic and international mobility on the resurgence and subsequent international spread of COVID-19 warrants further study. Further, the coordination between countries and regions in relaxing interventions and lifting international travel restrictions can greatly improve the likelihood of containing and reducing COVID-19 spread post-lockdown.[4]

Our findings should be considered in the context of several assumptions and data limitations. First, we only estimated the spread risk of COVID-19 via air travel here due to data availability. With the integration of land and sea transport models, the international travel networks are more

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

complex and our results may under-estimate the intercontinental risk of COVID-19 spread via travellers. Second, the accuracy of our modelling relies on accurate estimates of epidemiological parameters partially derived from reported case data, the quality of which might be constrained by case definitions, diagnosis and surveillance capacity, and other factors across countries/regions.[41-43] Third, mobile phone-derived data for parameterizing travel and physical distancing interventions in our model may not be representative of the population in each country due to variations in coverage and the spatiotemporal distribution of users in populations. Fourth, other factors and interventions, such as hand washing and wearing facemasks, may also contribute to COVID-19 spread or mitigation across space and time,[44-47] but our simulations did not specify their contributions to international transmission. Finally, our Bayesian phylogeographic inferences only included SARS-CoV-2 genomes available in GISAID by March 26, 2020, and the potential sampling bias among different continents may result in under-estimates of the peak of the second transmission phase and cause biases in the estimates.

However, our analyses highlight how heterogeneities in intercontinental travel have facilitated the intercontinental seeding of SARS-CoV-2 and shaped the spatiotemporal characteristics of the ongoing COVID-19 pandemic. Unrestricted movements between continents/countries after easing stringent lockdown measures would likely trigger a new wave of COVID-19 spread across the world. Our findings improve our understanding of early transmission dynamics of COVID-19 across continents and can help to tailor public health response strategies accordingly. Specifically targeted international travel interventions should be adapted to different phases of the pandemic for countries and their corresponding travellers,[48] together with timely testing, genetic sequencing, and contact tracing[49] for COVID-19 infections in travellers are also needed to monitor the ongoing pandemic and mitigate resurgences

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

post-lockdown.

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

Author contributions

Conceptualization: WS, YB, and SL; Data curation: JY, JL, and SL; Formal analysis: JY, JL, SL, and PW; Visualization: JY, JL, and SL; Writing: JY, SL, YB, WS, CWR, AC, NWR, AJT, JRF, and WX; Original draft: JY, SL, YB, WS, and AJT. All authors read and approved the manuscript.

Funding

This work was supported by Key research and development project of Shandong province [2020SFXGFY01 and 2020SFXGFY08]; the National Key Research and Development Programme of China [2016YFE0205800]; Strategic Priority Research Programme of the Chinese Academy of Sciences [XDB29010102 and XDA19090118]; National Natural Science Foundation of China (NSFC) [32041010 and 81773498]; National Science and Technology Major Project of China [2016ZX10004222-009]; the Academic Promotion Programme of Shandong First Medical University [2019QL006 and 2019PT008]; Bill & Melinda Gates Foundation [OPP1134076]; European Union Horizon 2020 [MOOD 874850]; National Major Project for Control and Prevention of Infectious Disease in China [2017ZX10104001 and 2018ZX10101004]; the NSFC Outstanding Young Scholars [31822055 to Y.B.]; Youth Innovation Promotion Association of CAS [2017122 to Y.B.]; Bill & Melinda Gates Foundation [OPP1170969 to N.R.; OPP1106427, OPP1032350, OPP1134076, and OPP1094793 to A.J.T.); the Clinton Health Access Initiative, the UK Department for International Development (DFID) and the Wellcome Trust [106866/Z/15/Z; 204613/Z/16/Z to A.J.T.]; Young Taishan Scholars Program of Shandong Province of China [tsqn20161046 to W.X.]; the Academic Promotion Programme of Shandong First Medical University [2019RC010 to W.X.] and Taishan Scholars Programme of Shandong Province

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

[ts201511056 to W.S.].

Acknowledgements

We thank the important work of SARS-CoV-2 genome data producers globally contributing sequence data to the GISAID, NCBI GenBank, NMDC, and NGDC databases. We also acknowledge Google and Baidu for sharing publicly available population mobility data.

Conflict of interest/disclosure

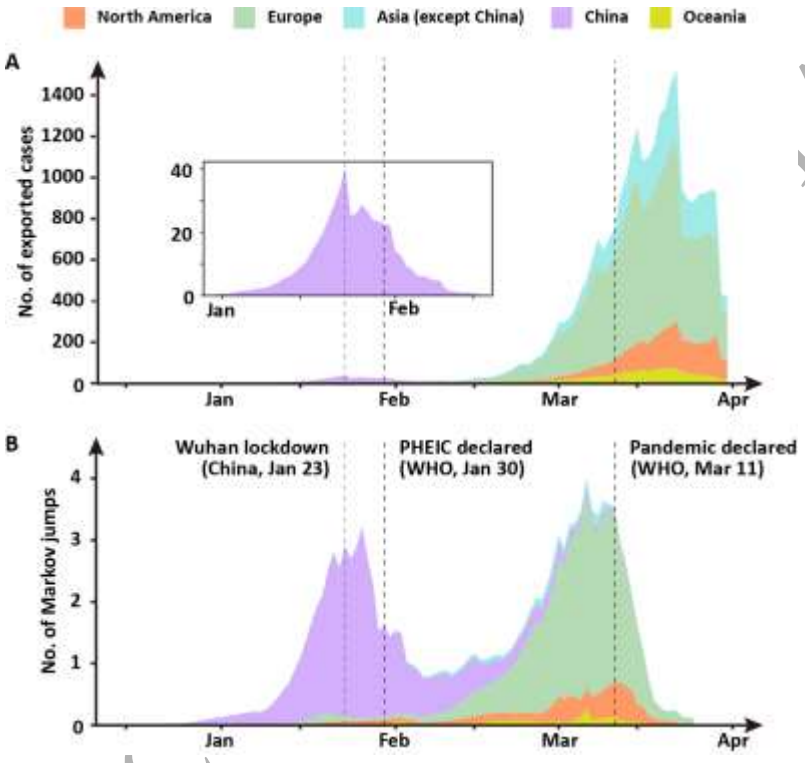
The authors have declared no conflicts of interest.

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

Figure Legends

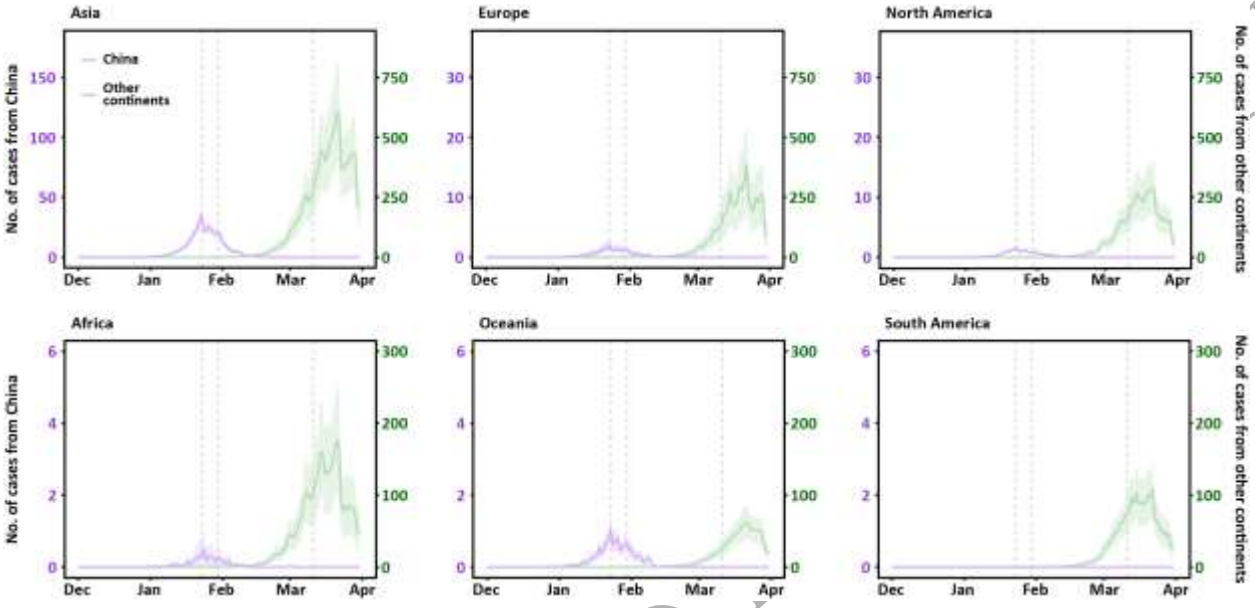
Figure 1. Dispersal dynamics of global COVID-19 cases and SARS-CoV-2 virus across regions. (**A**) The estimated numbers of COVID-19 cases exported from different regions by March 31, 2020. (**B**) Migration dynamics of SARS-CoV-2 virus from the different regions by March 26, 2020. The grey vertical dotted lines from left to right indicate the date of Wuhan shutdown in China, the date of COVID-19 declared as a Public Health Emergency of International Concern (PHEIC), and the date of COVID-19 declared as a pandemic by WHO, respectively. Colours in the plot represent the different regions defined in the present study.



UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

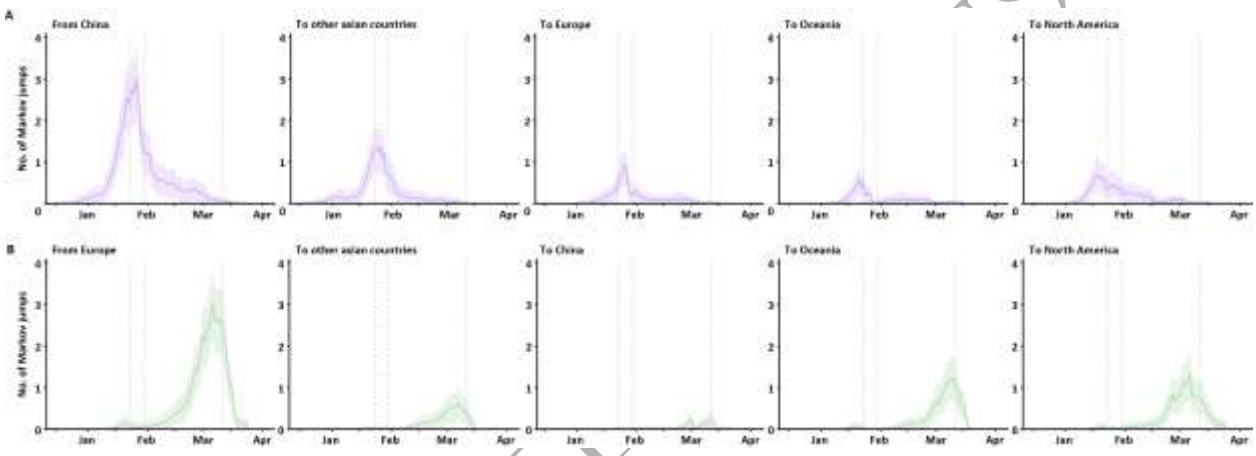
Figure 2. The estimated numbers of COVID-19 cases imported into different continents. This estimation ends at March 31, 2020. The median and interquartile range of estimates are provided. The three grey dotted vertical lines are the same as those in Figure 1.



UNCORRECTED

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

Figure 3. Migration events of SARS-CoV-2 from China and Europe to other continents. (**A**) Estimated Markov jumps of SARS-CoV-2 from China to other regions by March 26, 2020. (**B**) Estimated Markov jumps of SARS-CoV-2 from Europe to other regions by March 26, 2020. The mean and 95% high density intervals of estimates are provided. A Markov jump represents the frequency of transitions between locations along phylogenetic branches over time and can be regarded as one migration event of SARS-CoV-2. The three grey dotted vertical lines are the same as those in Figure 1.



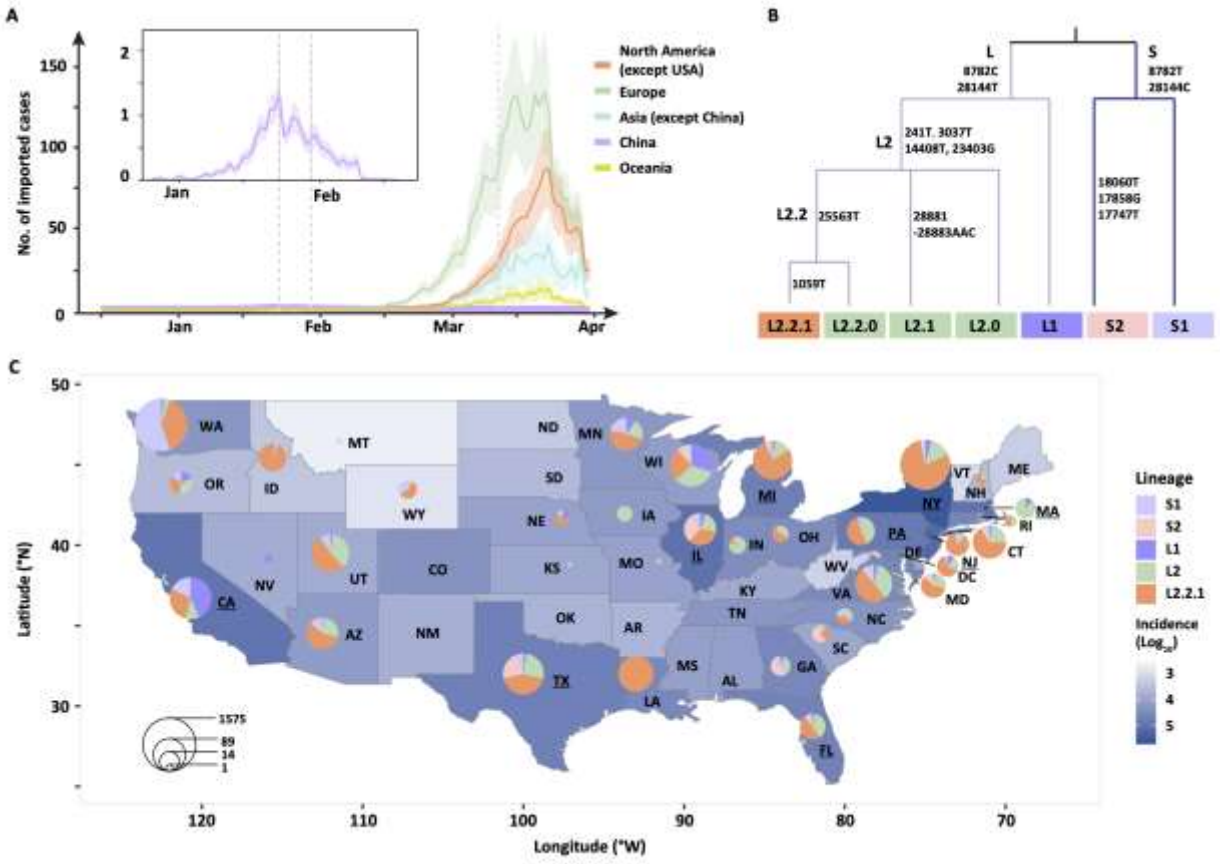
UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

Figure 4. Temporal dynamics of the imported COVID-19 cases and the circulating SARS-CoV-2 lineages in the USA. (**A**) Estimated COVID-19 case numbers imported into the USA from different sources (see legend) by March 31, 2020. Cases imported from China are highlighted in the inset box. The three grey dotted vertical lines are the same as those in Figure 1. (**B**) A simple classification framework for SARS-CoV-2 based on the maximum likelihood analysis and specific nucleotide substitutions along the SARS-CoV-2 genome. (**C**) Spatial distribution of the reported COVID-19 cases and the SARS-CoV-2 lineages within mainland USA. The shading from light to dark blue indicates the number of cumulative COVID-19 cases reported in each state by May 20, 2020. Names of the top ten states with the highest cumulative COVID-19 cases are underlined. The size of each circle represents the number of SARS-CoV-2 genome sequences in each state available from GISAID by May 20, 2020. Colours in the circles represent proportions of different virus lineages.

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020



UNCORRECTED

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

References:

1. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 2020; 579: 270–273.

2. Lu RJ, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 2020; 395: 565–574.

3. World Health Organization, "WHO Coronavirus Disease (COVID-19) Dashboard"; https://covid19.who.int.

4. Ruktanonchai NW, Floyd JR, Lai S, et al. Assessing the impact of coordinated COVID-19 exit strategies across Europe. Science 2020; doi: 10.1126/science.abc5096.

5. Lai S, Bogoch I, Ruktanonchai N, et al. Assessing spread risk of COVID-19 within and beyond China, January-April 2020: a travel network-based modelling study. medRxiv 2020. doi:10.1101/2020.02.04.20020479.

6. Lai S, Ruktanonchai NW, Zhou L, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature 2020. doi:10.1038/s41586-020-2293-x.

7. Tian H, Liu YH, Li YD, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in china. Science 2020; 368: 638–642.

8. Zhong P, Guo S, Chen T. Correlation between travellers departing from Wuhan before the Spring Festival and subsequent spread of COVID-19 to all provinces in China. J Travel Med 2020. doi: 10.1093/jtm/taaa036.

9. Zhao S, Zhuang Z, Cao P, et al. Quantifying the association between domestic travel and the exportation of novel coronavirus (2019-nCoV) cases from Wuhan, China in 2020: a

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

correlational analysis. J Travel Med 2020. doi: 10.1093/jtm/taaa022.

10. Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science 2020; 368: 395–400.

11. Daon Y, Thompson RN, Obolski U. Estimating COVID-19 outbreak risk through air travel. J Travel Med 2020. doi: 10.1093/jtm/taaa093.

12. Azad S, Devi S. Tracking the spread of COVID-19 in India via social networks in the early phase of the pandemic. J Travel Med 2020. doi: 10.1093/jtm/taaa130.

13. Candido DDS, Watts A, Abade L, et al. Routes for COVID-19 importation in Brazil. J Travel Med 2020. doi: 10.1093/jtm/taaa042.

14. European Centre for Disease Prevention and Control. Resurgence of reported cases of COVID-19 in the EU/EEA, the UK and EU candidate and potential candidate countries; https://www.ecdc.europa.eu/en/publications-data/rapid-risk-assessment-resurgence-reported-cases-covid-19.

15. Watts A, Au NH, Thomas-Bachli A, et al. Potential for inter-state spread of Covid-19 from Arizona, USA: Analysis of mobile device-location and commercial flight data. J Travel Med 2020. doi: 10.1093/jtm/taaa136.

16. 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: 689–697. doi:10.1016/S0140-6736(20): 30260-9.

17. Niehus R, De Salazar PM, Taylor AR, Lipsitch M. Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. Lancet Infect. Dis 2020. doi:10.1016/S1473-3099(20)30229-2.

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

18. Pullano G, Francesco P, Eugenio V, et al. Novel coronavirus (2019-nCoV) early-stage importation risk to Europe, January 2020. Euro surveillance: bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin 25, pii=2000057 (2020). doi:10.2807/1560-7917.ES.2020.25.4.2000057.

19. Centers for Disease Control and Prevention. Detection and Genetic Characterization of Community-Based SARS-CoV-2 Infections – New York City, March 2020; https://www.cdc.gov/mmwr/volumes/69/wr/mm6928a5.htm.

20. Deng XD, Gu W, Federman S, et al. Genomic surveillance reveals multiple introductions of SARS-CoV-2 into Northern California. Science 2020. doi: 10.1126/science.abb9263.

21. Pybus O, Rambaut A, Plessis LD, et al. "Preliminary analysis of SARS-CoV-2 importation & establishment of UK transmission lineages" 2020; https://virological.org/t/preliminary-analysis-of-sars-cov-2-importation-establishment-of-uk-transmission-lineages/507.

22. Wells CR, Sah P, Moghadas SM, et al. Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. PNAS 2020;117: 7504-7509.

23. Obadia T, Haneef R, Boelle PY. The R0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. BMC Medical Inform. Decis. Mak. 2012; 12: 147. doi:10.1186/1472-6947-12-147.

24. United Nations Office for the Coordination of Humanitarian Affairs, ACAPS COVID-19: Government Measures Dataset 2020; https://data.humdata.org/dataset/acaps-covid19-government-measures-dataset.

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

25. Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N. Engl. J. Med 2020. doi:10.1056/NEJMoa2001316.

26. Baidu Migration 2020; https://qianxi.baidu.com/.

27. Aktay A, Bavadekar S, Cossoul G, et al. Google COVID-19 community mobility reports: anonymization process description (version 1.0) 2020; https://arxiv.org/abs/2004.04145v2.

28. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 2013; 30: 772–780.

29. Suchard MA, Lemey P, Baele G, et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol 2018; 4: vey016.

30. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 2014; 30: 1312–1313.

31. Tang XL, Wu CC, Li X, et al. On the origin and continuing evolution of SARS-CoV-2. Natl. Sci. Rev. 2020. doi: 10.1093/nsr/nwaa036.

32. Rambaut A, Holmes EC, Hill V, et al. A dynamic nomenclature proposal for SARS-CoV-2 to assist genomic epidemiology. bioRxiv 2020. doi: 10.1101/2020.04.17.046086.

33. Gonzalez-Reiche AS, Hernandez MM, Sullivan MJ, et al. Introductions and early spread of SARS-CoV-2 in the New York city area. Science 2020. doi: 10.1126/science.abc1917.

34. Bedford T, Greninger AL, Roychoudhury P, et al. Cryptic transmission of sars-cov-2 in Washington state. medRxiv 2020. doi: 10.1101/2020.04.02.20051417.

35. Nadeau SA, Vaughan TG, Sciré J, et al. The origin and early spread of SARS-CoV-2 in Europe. medRxiv 2020. 2020.06.10.20127738; doi:

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

https://doi.org/10.1101/2020.06.10.20127738.

36. Centers for Disease Control and Prevention COVID-19 Response Team. Evidence for limited early spread of COVID-19 within the United States, January–February 2020 (2020). https://www.cdc.gov/mmwr/volumes/69/wr/mm6922e1.htm?s_cid=mm6922e1_w.

37. National Health Commission of the People's Republic of China. Situation report of COVID-19 in China. http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml.

38. Nuzzo J. Resurgence of COVID-19 in Japan. https://www.outbreakobservatory.org/outbreakthursday-1/7/16/2020/resurgence-of-covid-19-in-japan.

39. World Health Organization. India Situation Report. https://www.who.int/india/emergencies/coronavirus-disease-(covid-19)/india-situation-report.

40 Yang J, Niu P, Chen L, et al. Genetic tracing of HCoV-19 for the re-emerging outbreak of COVID-19 in Beijing, China. Protein Cell 2020. doi: 10.1007/s13238-020-00772-0.

41. Richardson S, Hirsch JS, Narasimhan M, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York city area. JAMA 2020. doi:10.1001/jama.2020.6775.

42. Grasselli G, Zangrillo A, Zanella A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy region, Italy. JAMA 2020; 323: 1574-1581.

43. Tsang TK, Wu P, Lin Y, et al. Effect of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China: a modelling study. Lancet Public Health 2020. doi:10.1016/S2468-2667(20)30089-X.

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020

44. Chu DK, Akl EA, Duda S, 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. doi:10.1016/S0140-6736(20)31142-9.

45. Walker PGT, Whittaker C, Watson OJ, et al. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. Science 2020. doi:10.1126/science.abc0035.

46. Hsiang S, Allen D, Annan-Phan A, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. Nature 2020. doi:10.1038/s41586-020-2404-8.

47. Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 2020. doi:10.1038/s41586-020-2405-7.

48. Costantino V, Heslop DJ, MacIntyre CR. The effectiveness of full and partial travel bans against COVID-19 spread in Australia for travellers from China during and after the epidemic peak in China. J Travel Med 2020. doi: 10.1093/jtm/taaa081.

49. Clifford S, Pearson CAB, Klepac P, et al. Effectiveness of interventions targeting air travellers for delaying local outbreaks of SARS-CoV-2. J Travel Med 2020. doi: 10.1093/jtm/taaa068.

UNCORRECTED MANUSCRIPT

Downloaded from https://academic.oup.com/jtm/advance-article/doi/10.1093/jtm/taaa200/5935386 by guest on 23 October 2020