



# RiskEstim: A Software Package to Quantify COVID-19 Importation Risk

Mingda Xu<sup>1,2,3†</sup>, Zhanwei Du<sup>1,2†</sup>, Songwei Shan<sup>1,2</sup>, Xiaoke Xu<sup>3</sup>, Yuan Bai<sup>1,2</sup>, Peng Wu<sup>1,2</sup>, Eric H. Y. Lau<sup>1,2</sup> and Benjamin J. Cowling<sup>1,2\*</sup>

<sup>1</sup>WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, Hong Kong SAR, China, <sup>2</sup>Laboratory of Data Discovery for Health, Hong Kong Science and Technology Park, Hong Kong, Hong Kong SAR, China, <sup>3</sup>College of Information and Communication Engineering, Dalian Minzu University, Dalian, China

## OPEN ACCESS

### Edited by:

Zhen Wang,  
Northwestern Polytechnical  
University, China

### Reviewed by:

Wenzhuo Song,  
Jilin University, China  
Xi Wang,  
The Chinese University of Hong Kong,  
Hong Kong SAR, China

Lu Zhong,  
Rensselaer Polytechnic Institute,  
United States

### \*Correspondence:

Benjamin J. Cowling  
bcowling@hku.hk

<sup>†</sup>These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Social Physics,  
a section of the journal  
Frontiers in Physics

Received: 15 December 2021

Accepted: 04 January 2022

Published: 24 January 2022

### Citation:

Xu M, Du Z, Shan S, Xu X, Bai Y, Wu P,  
Lau EHY and Cowling BJ (2022)  
RiskEstim: A Software Package to  
Quantify COVID-19 Importation Risk.  
Front. Phys. 10:835992.  
doi: 10.3389/fphy.2022.835992

We present an R package developed to quantify coronavirus disease 2019 (COVID-19) importation risk. Quantifying and visualizing the importation risk of COVID-19 from inbound travelers is urgent and imperative to trigger public health responses, especially in the early stages of the COVID-19 pandemic and emergence of new SARS-CoV-2 variants. We provide a general modeling framework to estimate COVID-19 importation risk using estimated pre-symptomatic prevalence of infection and air traffic data from the multi-origin places. We use Hong Kong as a case study to illustrate how our modeling framework can estimate the COVID-19 importation risk into Hong Kong from cities in Mainland China in real time. This R package can be used as a complementary component of the pandemic surveillance system to monitor spread in the next pandemic.

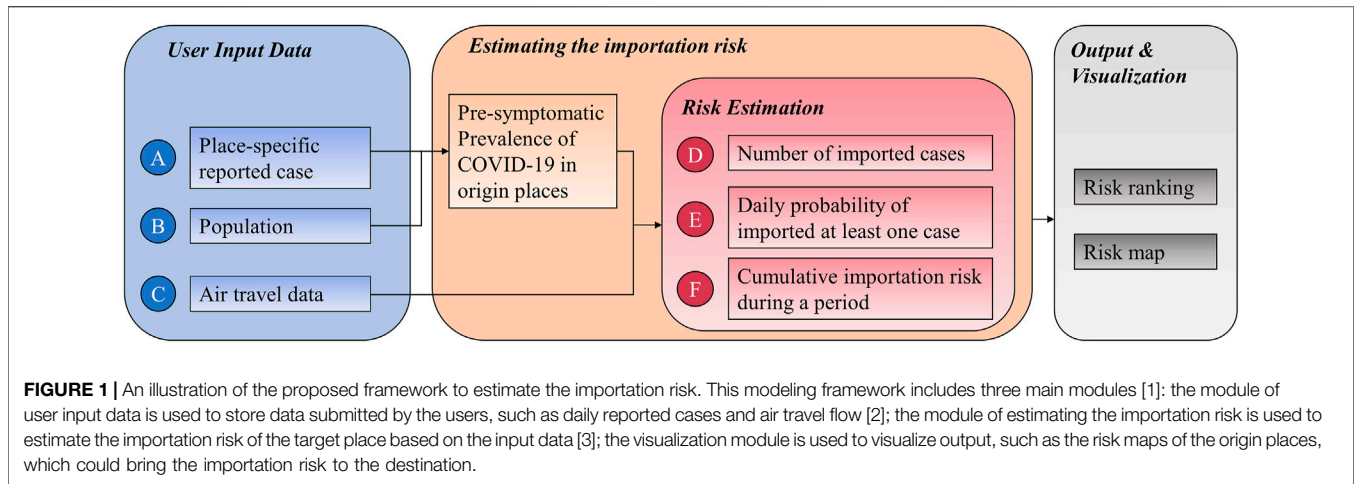
**Keywords:** COVID-19, SARS-CoV-2, importation risk, R package, introduction risk

## INTRODUCTION

The ongoing global pandemic of COVID-19 caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused incredible global disruption and challenges, in addition to the substantial health impact [1]. As of December 12, 2021, more than 269 million confirmed cases and 5.3 million deaths were reported worldwide [2].

Local outbreaks were often associated with the importation of infections. Quantifying and visualizing the importation risk of COVID-19 from inbound travelers is important for public health responses, especially in the early stage of an epidemic wave [3, 4]. For example, some studies showed that border control measures, such as flight restrictions and quarantine for inbound travelers from high-risk places (e.g., based on the number of new daily cases [5]), might have delayed epidemics in the destination countries [6–8]. In addition, assessment of the COVID-19 importation risk is needed for places where a high level of population immunity to COVID-19 has not been achieved in the target populations [9] or the government is considering relaxing border control measures [10–12]. Here, we present the R package *RiskEstim*, the latest codebase version developed to quantify COVID-19 importation risk. First, we outline the general modeling framework of the R package to estimate COVID-19 importation risk using daily pre-symptomatic prevalence data from multi-origin locations and air traffic data.

Hong Kong started the alert against COVID-19 and screening of the travellers from mainland China at the very early beginning of the pandemic [13]. Due to the sound public health infrastructure and the in-time response, the information for the reported cases in Hong Kong was highly reliable, and most of the imported cases at that time originated from mainland China [14]. From the above considerations, we used Hong Kong as a case study to illustrate how our modeling framework can estimate the multi-origin COVID-19 importation risk in real time.



**FIGURE 1** | An illustration of the proposed framework to estimate the importation risk. This modeling framework includes three main modules [1]: the module of user input data is used to store data submitted by the users, such as daily reported cases and air travel flow [2]; the module of estimating the importation risk is used to estimate the importation risk of the target place based on the input data [3]; the visualization module is used to visualize output, such as the risk maps of the origin places, which could bring the importation risk to the destination.

## METHODS

### The Modeling Framework of Estimating Importation Risk in the RiskEstim

To quantify the importation risk of COVID-19 from the place of origin to the destination, we first estimated the daily pre-symptomatic prevalence of the COVID-19 in each origin place. Then we calculated the number of potential imported cases using the estimated daily pre-symptomatic prevalence of origin places and the daily origin-destination air traffic data. Next, we estimated the probability of importing at least one case as the indicator of importation risk to rank the origin places and visualize the risk maps. The modeling framework is shown in **Figure 1**.

#### User Input Data in the RiskEstim

Using Hong Kong as an example, we applied the R package to estimate the importation risk from 15 high-risk cities in Mainland China into Hong Kong in early 2020 [15]. Daily confirmed COVID-19 cases reported by the Chinese Center for Disease Control and Prevention (China CDC) from January 1, 2020, to February 29, 2020 were obtained for the analysis, [16–18]. Because Hubei Province changed the definition of cases on February 12, 2020, which yielded a dramatic increase in the number of cases on February 12, 2020 and February 13, 2020 (14840 on February 12, 2020 and 4823 on February 13, 2020) [19]. To reduce the reporting bias due to different case definitions for COVID-19 during the study period [20], we assumed the number of reported cases in Wuhan on February 12 was the same as those on February 11, and that for February 13 were the same as that on February 14.

#### Estimating Pre-Symptomatic Prevalence of COVID-19 in Origin Places

The daily prevalence of pre-symptomatic infections could be estimated with the Package based on the daily reported cases in the origin place(s) input by the user. Let  $\omega_t^o$  be the number

of reported cases in the origin place  $O$  on day  $t$ . Then on average the cases reported on day  $t$  developed symptoms on day  $t - T_{rep}$  and were infected on day  $t - T_{rep} - T_{inc}$ , where  $T_{rep}$  and  $T_{inc}$  are the mean reporting delay and the median incubation period in days. Using this forward method, we estimated the daily numbers of infected individuals in the origin places.

In our case study of Hong Kong, we calculated the daily prevalence of pre-symptomatic COVID-19 in multiple origin places including cities from Hubei province and other provinces, and the estimates were consistent with the imported cases from these places during the early stage of the epidemic in Hong Kong [14]. Let  $y_t^o$  denote the place-specific pre-symptomatic prevalence of the place  $O$  on the day  $t$ , and  $H$  denote the cities in Hubei province. We used the median incubation period to denote the period where transmission would occur from infected cases. The place-specific pre-symptomatic prevalence is given by:

$$y_t^o = \begin{cases} \mu \sum_{d=t-T_{inc}+1}^t I_d^o, & o \in H \\ \sum_{d=t-T_{inc}+1}^t I_d^o, & otherwise \end{cases}$$

where  $\mu$  is the ascertainment rate ratio, representing the ascertainment rate of symptomatic cases in all non-Hubei provinces relative to Hubei province, which reflects the probability ratio of non-Hubei Provinces reporting a symptomatic case to Hubei Province [18].  $I_d^o$  denotes the incidence of SARS-CoV-2 infection in an origin place on day  $d$ . The parameters are summarized in **Table 1**.

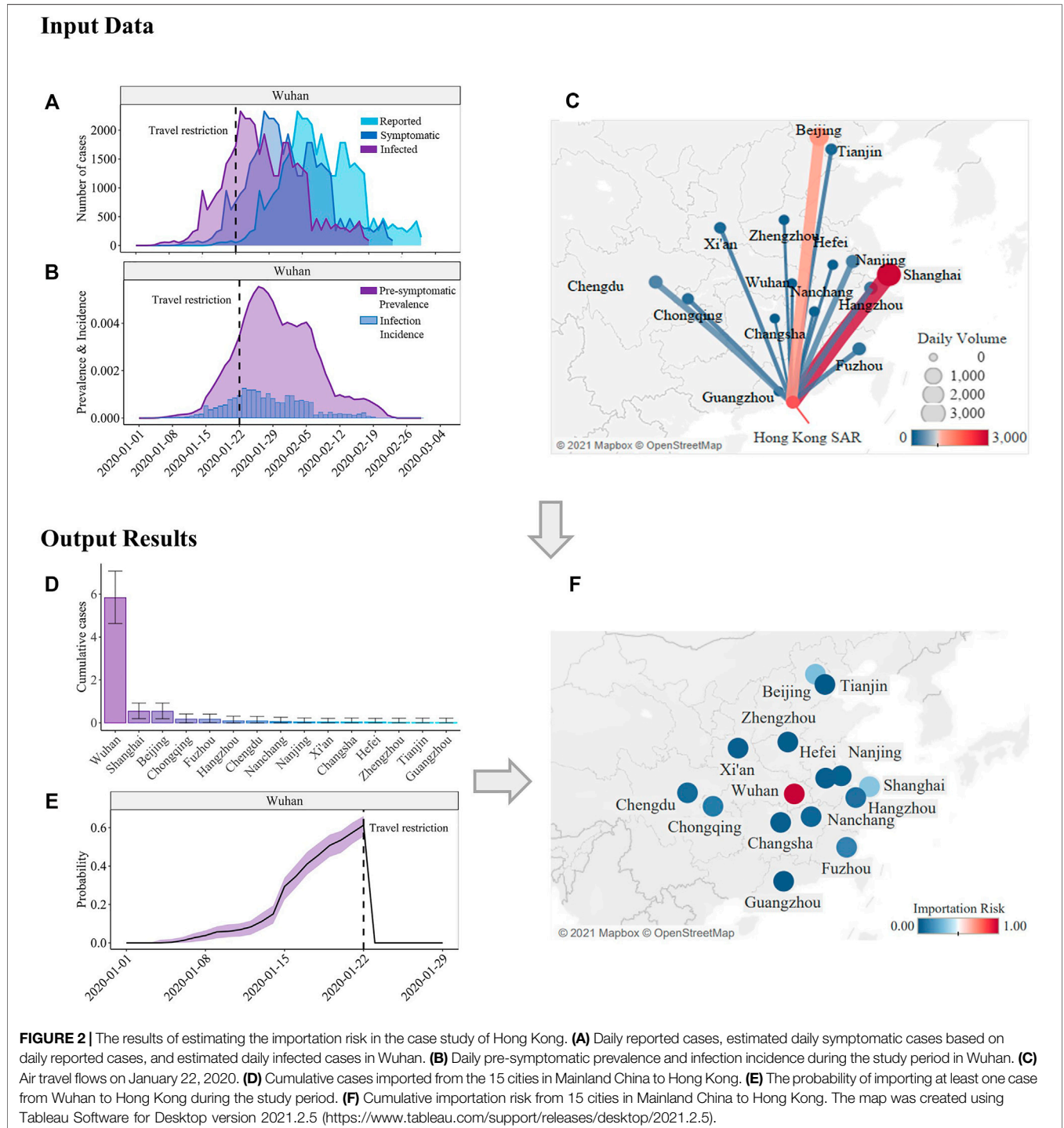
#### Estimating the Importation Risk

The place-specific importation risk of the destination was estimated based on [1]: daily pre-symptomatic prevalence of COVID-19 in origin places [2]; data on air passenger movements by place of origin and destination. Let  $\Gamma_t^{o,d}$  be the imported cases from the origin place  $O$  to the destination  $d$  on day  $t$ :

$$\Gamma_t^{o,d} = \alpha * y_t^o * M_t^{o,d}$$

**TABLE 1 |** Model parameters in the modeling framework.

Parameters	Description	Values	References
$T_{rep}$	Mean reporting delay	7 days	[21]
$T_{inc}$	Median incubation period	5 days	[22]
$\mu$	Ascertainment rate ratio of symptomatic cases in all non-Hubei provinces relative to that of Hubei province	5.14	[18]
$\alpha$	Scaling factor that relates the importation to scaled reported cases in high-surveillance places	5.05	[18]



where  $M_t^{o,d}$  represents the number of air passengers from origin place  $O$  to destination  $d$  on the day  $t$ , and  $\alpha$  is the scaling factor adjusting for the impact on the force of importation from varied surveillance efficiency on COVID-19 in different places [18]. With the assumption that the number of imported cases per day followed the Poisson distribution, we evaluated the 95% confidence interval (CI) of the imported cases based on 100 simulations. Following the study of estimating the probability of cases imported [23, 24], we estimated the cumulative importation risk  $\Phi_T^{o,d}$ , which denotes the cumulative probability of importing at least one case from the origin place  $O$  to the destination  $d$  during the period  $T$  between  $t_a$  and  $t_b$ , given by:

$$\Phi_T^{o,d} = 1 - \exp\left(-\int_{t=t_a}^{t_b} \Gamma_t^{o,d} dt\right)$$

## RESULTS

In our case study, we used daily reported cases of COVID-19 from 15 Mainland China cities, which were previously identified by Lai et al. [15] as high-risk cities COVID-19 imports during January 2020, to estimate the daily pre-symptomatic prevalence of these cities (Figures 2A,B). Based on the daily pre-symptomatic prevalence of these cities and the data on air travel flows between these 15 higher-risk Mainland China cities and Hong Kong (Figure 2C), we estimated the importation risk of Hong Kong (Figures 2D–F). The estimated number of imported cases from our model was 7.6 (95% CI: 5.0–12.1) from 15 higher-risk Mainland China cities into Hong Kong which was consistent with the reported 7 cases originating from Mainland China in Hong Kong before the Wuhan travel ban (January 23, 2020) [14, 25]. The estimated probability of importation of at least one case indicated that Wuhan exported the highest number of cases (5.8, 95% CI: 4.6–7.1) into Hong Kong, followed by Shanghai (0.5, 95% CI: 0.2–0.9) and Beijing (0.5, 95% CI: 0.2–0.9), during the study period.

## DISCUSSION AND CONCLUSION

This study aims to provide a general modeling framework to estimate COVID-19 importation risk. We illustrate the feasibility and reliability of the proposed framework with a case study which estimates the importation risk of COVID-19 to Hong Kong from multi-origin places using pre-symptomatic prevalence of infection and air traffic data. Notably, the method accommodates origin places where multiple variants circulate by estimating the importation risk of each variant separately then aggregating them in the destination places, given the availability of prevalence data and

human movement data. The method implemented in this study is from a previous study [18] and the reliability of it is demonstrated in the case study of Hong Kong, while proposing a technically innovative method with competitive accuracy is not our major focus. At the current time, only a main method is supported in our modeling framework, while it can be extended by other well-designed and fine-calibrated methods in the future, such as [24, 26–29]. These analyses of the correlation between importation risk and population movement data, preparedness, and vulnerability at the destination, will be further complemented.

This R package *RiskEstim* provides a general modeling framework to estimate the importation risk of infectious disease based on epidemiological and human movement data during an epidemic. The R package can be used as a complementary approach to the pandemic surveillance system to improve response to emerging SARS-CoV-2 variants and the next pandemic. In addition, the R package provides a modifiable codebase that can be extended to estimate the importation risk of other respiratory infectious diseases, such as influenza.

## DATA AVAILABILITY STATEMENT

Our software package is developed in R, called RiskEstim. All code to perform the analyses and generate the figures in this study are available from the corresponding author upon reasonable request. Publicly available datasets were analyzed in this study. This data can be found here: <https://doi.org/10.5281/zenodo.4266642>.

## AUTHOR CONTRIBUTIONS

BC, EL, XX, and ZD were involved in the conceptualization and design of the study. MX and ZD designed the statistical methods, conducted analyses, wrote the manuscript, and MX developed the R package. EL, ZD, SS, YB, and PW reviewed and edited the draft.

## FUNDING

This work was supported by AIR@InnoHK administered by the Innovation and Technology Commission. We acknowledge the financial support from an initiative funded by the COVID-19 Therapeutics Accelerator and convened by Health Data Research UK, Seed Fund for Basic Research for New Staff of the University of Hong Kong (grant no. 202009185062), and National Natural Science Foundation of China (grant nos. 72104208, 61773091 and 62173065).

## REFERENCES

- Chakraborty I, Maity P. COVID-19 Outbreak: Migration, Effects on Society, Global Environment and Prevention. *Sci Total Environ* (2020) 728:138882. doi:10.1016/j.scitotenv.2020.138882
- Weekly Epidemiological Update on COVID-19 - 14 December 2021 (2021). Available from: <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19-14-december-2021> (Accessed December 14, 2021).
- 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. *The Lancet* (2020) 395:689–97. doi:10.1016/s0140-6736(20)30260-9
- Du Z, Wang L, Cauchemez S, Xu X, Wang X, Cowling BJ, et al. Risk for Transportation of Coronavirus Disease from Wuhan to Other Cities in China. *Emerg Infect Dis* (2020) 26:1049–52. doi:10.3201/eid2605.200146
- CDC. How CDC Determines the Level for COVID-19 Travel Health Notices (2021). Available from: <https://www.cdc.gov/coronavirus/2019-ncov/travelers/how-level-is-determined.html> (Accessed December 06, 2021).
- Nakamura H, Managi S. Airport Risk of Importation and Exportation of the COVID-19 Pandemic. *Transport Policy* (2020) 96:40–7. doi:10.1016/j.tranpol.2020.06.018
- The Global Health Security Index (2019). Available from: <https://www.ghsindex.org/> (Accessed December 14, 2021).
- Grépin KA. Evidence of the Effectiveness of Travel-Related Measures during the Early Phase of the COVID-19 Pandemic: a Rapid Systematic Review. *BMJ Glob Health* (2021) 6, e004537. doi:10.1136/bmjgh-2020-004537
- Leung K, Wu JT, Leung GM. Effects of Adjusting Public Health, Travel, and Social Measures during the Roll-Out of COVID-19 Vaccination: a Modelling Study. *The Lancet Public Health* (2021) 6:e674–e682. doi:10.1016/s2468-2667(21)00167-5
- Devi S. COVID-19 Resurgence in Iran. *The Lancet* (2020) 395:1896. doi:10.1016/s0140-6736(20)31407-0
- Lai S. Assessing the Effect of Global Travel and Contact Reductions to Mitigate the COVID-19 Pandemic and Resurgence. *medRxiv* (2020) 7:914–923. doi:10.1101/2020.06.17.20133843
- Ruktanonchai NW, Floyd JR, Lai S, Ruktanonchai CW, Sadilek A, Rente-Lourenco P, et al. Assessing the Impact of Coordinated COVID-19 Exit Strategies across Europe. *Science* (2020) 369:1465–70. doi:10.1126/science.abc5096
- Government Launches Preparedness and Response Plan for Novel Infectious Disease of Public Health Significance. Available from: <https://www.info.gov.hk/gia/general/202001/04/P2020010400179.htm> (Accessed December 14, 2021).
- Lai CKC, Ng RWY, Wong MCS, Chong KC, Yeoh YK, Chen Z, et al. Epidemiological Characteristics of the First 100 Cases of Coronavirus Disease 2019 (COVID-19) in Hong Kong Special Administrative Region, China, a City with a Stringent Containment Policy. *Int J Epidemiol* (2020) 49:1096–105. doi:10.1093/ije/dyaa106
- Lai S, Bogoch II, Ruktanonchai NW, Watts A, Lu X, Yang W, et al. Assessing Spread Risk of Wuhan Novel Coronavirus within and beyond China, January–April 2020: a Travel Network-Based Modelling Study. *medRxiv* (2020). doi:10.1101/2020.02.04.20020479
- Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the Severity of Coronavirus Disease 2019: a Model-Based Analysis. *Lancet Infect Dis* (2020) 20:669–77. doi:10.1016/s1473-3099(20)30243-7
- Emery JC, Russell TW, Liu Y, Hellewell J, Pearson CA. The Contribution of Asymptomatic SARS-CoV-2 Infections to Transmission on the Diamond Princess Cruise Ship. *Elife* (2020) 9:e58699. doi:10.7554/eLife.58699
- Menkir TF, Chin T, Hay JA, Surface ED, De Salazar PM, Buckee CO, et al. Estimating Internationally Imported Cases during the Early COVID-19 Pandemic. *Nat Commun* (2021) 12:311. doi:10.1038/s41467-020-20219-8
- Tsang TK, Wu P, Lin Y, Lau EHY, Leung GM, Cowling BJ. Effect of Changing Case Definitions for COVID-19 on the Epidemic Curve and Transmission Parameters in mainland China: a Modelling Study. *The Lancet Public Health* (2020) 5:e289–e296. doi:10.1016/s2468-2667(20)30089-x
- Update on COVID-19 as of 24:00 on 12 February. Available from: <http://www.nhc.gov.cn/xcs/yqtb/202002/26fb16805f024382bff1de80c918368f.shtml> (Accessed December 14, 2021) (2021).
- Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. Evolving Epidemiology and Transmission Dynamics of Coronavirus Disease 2019 outside Hubei Province, China: a Descriptive and Modelling Study. *Lancet Infect Dis* (2020) 20:793–802. doi:10.1016/s1473-3099(20)30230-9
- Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, 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
- Du Z, Wang L, Yang B, Ali ST, Tsang TK, Shan S, et al. Risk for International Importations of Variant SARS-CoV-2 Originating in the United Kingdom. *Emerg Infect Dis* (2021) 27:1527–9. doi:10.3201/eid2705.210050
- Yang B, Tsang TK, Wong JY, He Y, Gao H, Ho F, et al. The Differential Importation Risks of COVID-19 from Inbound Travellers and the Feasibility of Targeted Travel Controls: A Case Study in Hong Kong. *Lancet Reg Health - West Pac* (2021) 13:100184. doi:10.1016/j.lanwpc.2021.100184
- Latest Situation of Novel Coronavirus Infection in Hong Kong. 2022. Available at: <https://chp-dashboard.geodata.gov.hk/covid-19/en.html> (Accessed December 14, 2021).
- Jia JS, Lu X, Yuan Y, Xu G, Jia J, Christakis NA. Population Flow Drives Spatio-Temporal Distribution of COVID-19 in China. *Nature* (2020) 582:389–94. doi:10.1038/s41586-020-2284-y
- Gilbert M, Pullano G, Pinotti F, Valdano E, Poletto C, Boëlle P-Y, et al. Preparedness and Vulnerability of African Countries against Importations of COVID-19: a Modelling Study. *The Lancet* (2020) 395:871–7. doi:10.1016/s0140-6736(20)30411-6
- Li X, Chen M, Nie F, Wang Q. Locality Adaptive Discriminant Analysis. in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)* (2017):2021–2027. doi:10.24963/ijcai.2017/306
- Li X, Chen M, Nie F, Wang Q. A Multiview-Based Parameter Free Framework for Group Detection. In: *Thirty-First AAAI Conference on Artificial Intelligence* (2017).

**Conflict of Interest:** BC reports honoraria from AstraZeneca, GSK, Moderna, Pfizer, Roche, and Sanofi Pasteur.

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

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

Copyright © 2022 Xu, Du, Shan, Xu, Bai, Wu, Lau and Cowling. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.