International Society of Travel Medicine
Established 1991
Promoting healthy travel worldwide

Research Letter

Quantifying the association between domestic travel and the exportation of novel coronavirus (2019-nCoV) cases from Wuhan, China in 2020: a correlational analysis

Shi Zhao[1,2], Zian Zhuang[3], Peihua Cao[4], Jinjun Ran[5], Daozhou Gao[6], Yijun Lou[3], Lin Yang[7], Yongli Cai[8], Weiming Wang[8], Daihai He[3,*], and Maggie H. Wang[1,2]

[1]JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China, [2]Shenzhen Research Institute of Chinese University of Hong Kong, Shenzhen, China, [3]Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China, [4]Clinical Research Centre, Zhujiang Hospital, Southern Medical University, Guangzhou, China, [5]School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China, [6]Mathematics and Science College, Shanghai Normal University, Shanghai, China, [7]School of Nursing, Hong Kong Polytechnic University, Hong Kong, China and [8]Department of Mathematics, Huaiyin Normal University, Huai'an, China

*To whom correspondence should be addressed. Email: daihai.he@polyu.edu.hk

Submitted 2 February 2020; Revised 10 February 2020; Editorial Decision 11 February 2020; Accepted 11 February 2020

In the end of 2019, a novel coronavirus (2019-nCoV) emerged in Wuhan, China and was causing a serious outbreak of acute respiratory illness.[1] Wuhan locates in the centre of mainland China with a population of 14 million and is very conveniently connected to other parts of China through airlines and high-speed rails.[2] As of 31 January 2020 (5:00 p.m., GMT + 8), there were 9809 confirmed 2019-nCoV cases in mainland China, including 213 deaths and 180 discharges.[3, 4] Cases infected in Wuhan were also detected in many foreign countries or regions including Thailand, Japan, Republic of Korea, the United States, Canada and some European countries.[4] The World Health Organization (WHO) has declared that the novel coronavirus outbreak is a public health emergency of international concern. Official reports on the newly confirmed cases are released very rapidly (several times a day) after January 16,[3–5] as the official diagnosis protocol was released by the WHO on January 17.[6] Recent studies indicated the likelihood of travel-related risks of 2019-nCoV spreading both domestically and internationally.[7, 8] Many major cities in mainland China reported the finding of imported cases, including Beijing, Shanghai, Guangzhou and Shenzhen. The outbreak is still on-going with an increasing trend in daily new cases.[3, 4] Before the Wuhan lockdown (official travel restriction) on January 23, virtually all cases found in other major cities were exported cases from Wuhan. Population flow data between major cities in mainland China are available online due to the rapid development of internet in recent

decades, see https://qianxi.baidu.com/ (in Chinese). In this work, we quantified the association between the domestic travel load and the number of cases exported from Wuhan to other city-clusters in mainland China. Our city-clusters are the pool of top five cities in the top 10 provinces (in number of cumulative cases). Thus, we included 10 city-clusters in the analysis, and the details of the selected city-clusters can be found in Supplementary Data S1.

We examined the association between the load of domestic passengers departed from Wuhan and the number of confirmed cases to the 10 city-clusters (including the three municipalities, Beijing, Shanghai and Chongqing). Data sets of the daily numbers of domestic passengers were obtained from the location-based services database of Baidu company from January 1 to 20. We selected the top 10 provincial regions (except Hubei) with the largest cumulative number of cases, which accounts for 68% of all cases reported outside Hubei, before the implementation of the city lockdown, on 23 January 2020. Cases from other provinces are scattered and showed no clear pattern. The daily numbers of passengers from Wuhan to the city-clusters of each province were adopted to measure the load of domestic passengers departed from Wuhan to that cluster. From now on, province means the city-clusters in that province. The daily number of cases time series are obtained via the online outbreak situation reports.[3, 4] Daily cases for each cluster are scaled from daily province total. The detail is given in Supplementary Data. The association was formulated

Table 1. Summary of the estimated percentage change rate (%) modelled in Equation (1) between daily passengers and number of exported 2019-nCoV cases

Delay term (τ)	Percent change, Δ (%, per 100 passengers per day)			LR test	Pseudo-R^2
	Per 1 new case	Per 10 new cases	Per 100 new cases		
3 days	0.23 (0.19, 0.26)	2.27 (1.96, 2.59)	25.23 (21.47, 29.10)	$p < 0.001$	0.56
4 days	0.19 (0.17, 0.21)	1.93 (1.74, 2.12)	21.02 (18.78, 23.30)	$p < 0.001$	0.61
5 days	0.15 (0.14, 0.16)	1.52 (1.40, 1.64)	16.25 (14.86, 17.66)	$p < 0.001$	0.62
6 days	0.11 (0.10, 0.11)	1.06 (0.99, 1.14)	11.15 (10.30, 12.00)	$p < 0.001$	0.52
7 days	0.08 (0.07, 0.08)	0.79 (0.74, 0.85)	8.22 (7.64, 8.81)	$p < 0.001$	0.43

The highlighted estimates were treated as the main results.

Note: the 'LR test' is the likelihood-ratio (LR) test of the model in Equation (1) against the model in Equation (2). The 'pseudo-R^2' is the McFadden's pseudo-R-squared.

as follows:

$$\log\left[\mathbf{E}\left(c_{i,t}\right)\right] = \alpha_i \cdot \text{province}_i + \beta \cdot \xi_{i,t-\tau} \cdot \varepsilon_{t-\tau}. \quad (1)$$

Here, $c_{i,t}$ represents the daily number of new cases in the ith provincial region on day t. The function $E(\cdot)$ is the expectation; the 'province$_i$' denotes the dummy variable for the ith provincial region in accounting for the heterogenicity among different provinces, and thus, α_i is a locality-varying interception term. Term ξ_i is the daily number of passengers from Wuhan to the ith province. The time index was denoted by t, and the term τ modelled the delay from exposure to being detected. The ε represented the daily number of new cases in Wuhan. Hence, the product term ($\xi \cdot \varepsilon$) was proportional to the rate of cases exported from Wuhan to other places, and this setting was consistent with the framework in Imai *et al.*,[9] which estimated the outbreak size through exported 2019-nCoV cases overseas, as well as in more complex frameworks.[10] The β is the regression coefficient to quantify the association between the load of passengers multiplied by the local infectivity in Wuhan and the number of cases reported outside Wuhan. Hence, the change rate, denoted by $\Delta = [\exp(\beta \times 100) - 1] \times 100\%$, may be interpreted as the expected percentage change in the daily number of cases offsite found associated with per 100 increase in the daily number of passengers departed from Wuhan where there was one new case daily. We estimated β for three means of transportation, and the *P*-value less than 0.05 was considered as statistical significance.

We also considered a baseline version of the model by replacing term '$\beta \cdot \xi_{i,t-\tau} \cdot \varepsilon_{t-\tau}$' in Equation (1) with '$\beta_0 \cdot \xi_{i,t-\tau}$', which ignored the variation in the force of infection in Wuhan, thus a constant, as in Equation (2)

$$\log\left[\mathbf{E}\left(c_{i,t}\right)\right] = \alpha^{(0)}_i \cdot \text{province}_i + \beta_0 \cdot \xi_{i,t-\tau}. \quad (2)$$

The likelihood ratio (LR) test was adopted to justify the model structure in Equation (1) against the baseline form in Equation (2). For the delay term (τ), it was expected to be equivalent to the incubation period, which was reported to be 5.2 days (95% CI: 4.1–7.0) in.[1] Finally, we examined the association with τ varying at 3, 4, 5, 6 and 7 days in sensitivity analysis.

The LR test yielded statistically significant outcomes, which suggests that the model in Equation (1) is more reasonable than the baseline form in Equation (2). Through the goodness-of-

fit in terms of the McFadden's pseudo-R-squared, we found that $\tau = 5$ days attains the best fitting performance to explain the patterns of the cases offsite detected. This well matched the estimate of the mean incubation period of the infection at 5.2 in.[1] We found a statistically significant positive association between the load of passengers multiplied by the local infectivity in Wuhan and the number of cases reported outside Wuhan, see Table 1. We estimated that per 100 cases increase in the daily number of newly reported cases in Wuhan together with per 100 persons increase in the daily number of passengers departed Wuhan were likely to cause a 16.25% (95% CI: 14.86–17.66%) increase in the daily number of cases offsite detected on average. The sensitivity analysis by varying τ suggested this fundamental relationship holds, and the details can be found in Table 1 and the Supplementary Data S2.

This analysis has limitations. The data used in this study are in the early phase of the outbreak and may not represent subsequent waves. We aim at quantifying the association between domestic travel and 2019-nCoV exportation in China. Although the differences in case ascertainment in different cities clusters are considered in the model by the dummy term 'province$_i$', the association could also bear heterogeneity. Temporal and spatial correlations were not addressed in this simple modelling analysis due to lack of data. The 2019-nCoV surveillance data were too scattered and short at this early stage to consider temporal and spatial correlation. The correlation between population flow, number of cases offsite detected, and source infection prevalence was addressed in this work. Our modelling framework would be extended to a more complex and realistic form for exploring the potential spatial correlations, and benefit from more detailed disease surveillance and travel population flow data.

Supplementary data

Supplementary data are available at *JTM* online.

Funding

Acknowledgements

The authors would like to acknowledge colleagues for helpful comments.

Conflict of interest

The authors declared no competing interests.

Authors' contributions

All authors conceived the study, carried out the analysis, discussed the results, drafted the first manuscript, critically read and revised the manuscript and gave final approval for publication.

Disclaimer

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.

Ethics approval and consent to participate

The ethical approval or individual consent was not applicable.

Availability of data and materials

All data and materials used in this work were publicly available and also available based on request.

Consent for publication

Not applicable.

References

1. 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.
2. Zhao S, Zhuang Z, Ran J *et al*. The association between domestic train transportation and novel coronavirus outbreak in China, from 2019 to 2020: a data-driven correlational report. *Travel Med Infect Dis* 2020; **33**:101568.
3. 'Situation report of the pneumonia cases caused by the novel coronavirus', released by the National Health Commission of the People's Republic of China (in Chinese). http://www.nhc.gov.cn/yjb/s3578/202001/a3c8b5144067417889d8760254b1a7ca.shtml.
4. Novel Coronavirus (2019-nCoV) situation reports, released by the World Health Organization (WHO). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.
5. 'News press and situation reports of the pneumonia caused by novel coronavirus', from December 31, 2019 to Janunary 21, 2020 released by the Wuhan municipal health commission, China. http://wjw.wuhan.gov.cn/front/web/list2nd/no/710.
6. Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases: Interim guidance, released on January 17, 2020 by the World Health Organization (WHO). https://www.who.int/publications-detail/laboratory-testing-for-2019-novel-coronavirus-in-suspected-human-cases-20200117.
7. Bogoch II, Watts A, Thomas-Bachli A, Huber C, MUG K, Khan K. Pneumonia of unknown Etiology in Wuhan, China: potential for international spread via commercial air travel. *J Travel Med* 2020; **pii**:taaa011. doi: 10.1093/jtm/taaa011.
8. Leung K, Wu JT, Leung GM: *Nowcasting and forecasting the Wuhan 2019-nCoV outbreak Preprint published by the School of Public Health of the University of Hong Kong*, 2020. https://files.sph.hku.hk/download/wuhan_exportation_preprint.pdf.
9. Imai N, Dorigatti I, Cori A *et al*. *Estimating the potential total number of novel Coronavirus (2019-nCoV) cases in Wuhan City, China Preprint published by the*. *Imperial College London*, 2020.
10. 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. doi: 10.1016/S0140-6736(20)30260-9.