Geographic pattern of human mobility and COVID-19 before and after Hubei lockdown

T. Edwin Chow †, Yusik Choi †, Mei Yang ‡, David Mills ‡ and Ricci Yue †,‡,§

†Department of Geography, Texas State University, TX, USA; ‡Department of Geography and Resource Management, Chinese University of Hong Kong, HK SAR, China; §Department of Public Policy, City University of Hong Kong, Hong Kong SAR, China

ABSTRACT

This research investigates how travel restrictions affect the spatiotemporal pattern of human mobility and COVID-19 confirmed cases. Based on recorded movement and Baidu mobility index, in- and out-migration were estimated to examine the geographic pattern of human mobility across many Chinese cities from Jan 1 – Feb 11 of 2020. In addition to the baseline model of city lockdown, this study also explored the time lag effect of COVID-19 incubation period before/after Jan 28 (i.e. 5 days) and Feb 6 (i.e. 2 weeks) as well. Full factorial Analysis of Variance (ANOVA) tests reviewed significant differences of migration pattern by lockdown and origin/destination, which are also significantly associated with the confirmed cases of COVID-19 as well. Specifically, human mobility dropped proportionally after the lockdown regardless of origin location, but Hubei destination was significantly lower than non-Hubei destination. The model assuming an incubation period of 5 days differentiated the differences of COVID-19 cases better than the baseline and 14 days model. Spatiotemporal cluster analysis identified multiple space-time windows that were related to migration trajectory assuming a 5–14 days incubation period. The pre-lockdown clusters due to traveler’s outflow from Wuhan to those megacities were the pathways for international transmission of COVID-19, whereas the post-lockdown clusters were partially related to the migration pattern especially within the eastern part of Hubei around Wuhan. The geographic pattern revealed from this study confirmed the presence of super spreaders that were responsible for regional spreading at the early stage and caused local outbreaks in the latter stage.

Introduction

In December 2019, the first case of novel coronavirus disease (COVID-19) was reported in Wuhan, a megacity of Hubei province in China. Despite some ‘rumours’ about an outbreak of pneumonia cases that were associated with infected people who visited a local seafood market, the general public were not aware of any risk of this unknown virus of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). With a migratory floating population, of 2,243,700 people who live there for at least 6 months out of the year (Statistics Bureau of Wuhan, 2020), hundreds of thousands of migrant workers in Wuhan went back to visit their hometowns for celebration during the week before Lunar New Year on 24 January 2020. These events increased human mobility as well as the risk of disease transmission to many Chinese cities.

On 21 January 2020, the announcement of person-to-person transmission of COVID-19 has spawned a fear among the public and an outburst of out-migration from Wuhan. In fear of further dispersion of this infectious disease, Wuhan was locked down on 23 January, and 15 other cities in Hubei province followed on the next day. Nevertheless, millions of people have left Wuhan before the travel restriction (The Guardian 2020). Despite the drastic effort of locking down a megacity, the COVID-19 continued to spread across China and beyond. The World Health Organization (WHO) declared COVID-19 a pandemic on 11 March when the cumulative number of confirmed patients marked 118,319 worldwide (World Health Organization (WHO), 2020).

To mitigate the pandemic of COVID-19, both medical treatment and non-pharmaceutical intervention (NPI) strategies are needed. Besides treating the infected, the primary goal of NPI is to contain and control the transmission by early detection, contact tracing, and limiting human mobility to minimize the risk of exposure to the virus. Common NPI strategies include quarantine, travel restriction, social distancing, curfew, and a combination of all (i.e. city lockdown). A combination of these NPI strategies can be
regarded as various levels or stages of a city ‘lockdown’. The effectiveness of these transmission control measures in slowing down the spreading depends on the aggressiveness of these intervention strategies appropriate to the epidemiological nature of infectious disease (WHO Writing Group, 2006). As COVID-19 has infected millions of people worldwide, many cities around the globe have implemented varying magnitudes of multiple NPI.

Research leveraging migration data are helpful to understand spatial epidemiology and public health intervention. Spatial analysis of human mobility data can support decision making by modelling disease spread, detecting patterns and statistically significant hotspot, and predicting future risk (Dong et al. 2020). Geographic Information Systems (GIS) and spatial analysis are multidisciplinary toolsets that can be used to identify and understand the spatiotemporal patterns of geographic epidemiology (Rezaeian et al. 2007). As demonstrated by John Snow in advancing our understanding about cholera, simply mapping, geographic visualization and point pattern analysis can help exploring and confirming the disease origin and its relationship with environmental factors (Yuan et al. 2020). Lai et al. (2004) used cluster analysis and kernel density surfaces to examine and explain the hotspots of Severe Acute Respiratory Syndrome (SARS) in Hong Kong during 2003. In addition to clinical analysis in epidemiology, the geographic approach offers a supplementary framework to explore and identify not only where but also why the emergence, spread and subsidence of an acute infectious disease.

As COVID-19 remains lethal, highly contagious and poorly understood, the purpose of this study is to explore the geographic pattern of human mobility and confirmed cases of COVID-19. By mapping out the spatiotemporal clusters, this study also investigates their relationship, if there is any, to better understand the epidemiological traits of COVID-19.

**Literature review**

Human mobility is intrinsically tied with the spread of COVID-19 as it is highly contagious in person-to-person transmission. An infected person who moved around could easily infect others in close contact during the incubation period that ranges from a median of 5.1 days up to 2 weeks as revealed from 99% of COVID-19 patients (Lauer et al. 2020). Early cases of COVID-19 can often be asymptomatic or show very minor symptoms, which have been estimated to have been undocumented accounted for as much as 86% of all infections (Li et al. 2020b). As such, these epidemiological traits of COVID-19 are ideal for causing sudden and exponential growth in regional as well as international destinations (Cucinotta and Vanelli, 2020). Hence, the geographic pattern of the emergence of COVID-19 and its spread helps to better understand the epidemiological traits of this virus and evaluate existing non-pharmaceutical intervention (NPI) strategies to contain and control the infectious disease.

By comparing the recorded movement before and after travel restriction in Wuhan, out-migration from Wuhan had significantly reduced by as much as 90% (Tian et al. 2020). While Fang et al. reported a less dramatic drop in outflow at 50.2% from Wuhan, they also reported a reduced inflow of 76.6% as well as within-Wuhan movements by 54.2%. Using 8 days as one standard deviation of the average incubation period, Kraemer et al. (2020) reported that the number of confirmed cases with known history travelling to Wuhan before and after Jan 31 dropped from 515 cases to 39 cases. It was shown that human mobility correlates very well with COVID-19-confirmed cases at the early stage, indicating the effectiveness of the travel ban as a preventive control measure (Fang, Wang, and Yang 2020). However, as millions of people have left Wuhan before any travel restriction, COVID-19 has spread to other cities (Du et al. 2020). Researchers reported that the travel ban contributed a modest delay of 3 to 5 days to find the first reported cases in other cities in the mainland (Tian et al. 2020). Despite limited success in stopping the spread domestically (Jia et al. 2020), the travel ban was shown to be more effective in restricting the transmission across the international border (Chinazzi et al. 2020). Dependent on epidemiological traits like the length of incubation period and the proportion of asymptomatic patients, however, travel restrictions alone cannot completely eradicate the local outbreak of COVID-19 as evident from the declaration of pandemic in March.

As shown in previous studies, multiple measures of these NPIs constituting a ‘city lockdown’ can be effective in ‘flattening the curves’ over time (Fang, Wang, and Yang 2020; Jia et al. 2020; Pan et al. 2020; Tian et al. 2020; Hatchett, Mecher, and Lipsitch 2007; Matrajt and Leung 2020; Sanche et al. 2020). In Italy and Spain, the first city lockdown discouraged crowd gathering while recommended the keeping of social distancing in public space. Preliminary analysis in Italy & Spain has shown that restricting social contact in the first city lockdown has successfully reduced the daily confirmed cases, death rate as well as admission to the Intensive Care Unit (ICU) (Tobias 2020). Despite slowing down the rate
of infection, the upward trend of the epidemic remains. Hence, the second round of lockdown further restricts domestic travelling and shuts down all non-essential businesses. While these dramatic efforts were shown to ‘flatten the curve’, the confirmed cases at a time only reflect diagnosed cases after a certain incubation period, and hence its full effect has yet to be confirmed.

While the pandemic of COVID-19 is still in its full swing, some cities remain in lockdown while some are in various phases of re-opening. It is important to evaluate the impact of city lockdown on the geographic pattern of human mobility and spread of COVID-19. As the outbreak in China has been under control since March, it is critical to learn from the original epicentre of this outbreak with regards to how this infectious disease unfolded with human mobility and the effectiveness of relevant NPI strategies (e.g. travel restrictions). This research investigates how travel restrictions affect the spatiotemporal pattern of human mobility and COVID-19 confirmed cases. To answer the research question, this study identifies the following objectives:

1. Examine any significant difference of human mobility before and after the city lockdown
2. Extract the spatiotemporal patterns of confirmed cases of COVID-19
3. Explore the spatiotemporal relationship between human mobility and confirmed cases of COVID-19

Unlike previous studies that established the correlation between human mobility and COVID-19 at a spatially-invariant manner, a unique contribution of this study is that the authors analysed the geographic relationship between human mobility and COVID-19 in major Chinese cities. There are a couple key differences noted between this and previous studies.

First, this study reconstructed the real-time mobility data to all cities that have migration record, not just to/ from Wuhan. In previous mobility studies (e.g. Kraemer et al. (2020), Tian et al. (2020), and Jia et al. (2020)), only the out-mobility data leaving from Wuhan were used to correlate with COVID-19 cases. However, as COVID-19 spread to other places in Hubei province and all over China, inter-city travel outside of Wuhan or Hubei province are also critical in spreading the infectious disease. Similar to Fang, Wang, and Yang (2020), this study harvested the inter-city mobility data of top 100 origins and destinations for 364 major Chinese cities (i.e. the top 100 destinations leaving from city A as well as top 100 origins going into city A) from Baidu Qianxi data. This daily mobility matrix covers approximately 3 million origin-destination pairs. Moreover, the relative migration % directly obtained from Baidu Qianxi was further augmented into actual number of recorded movements by reconstructing and reconciliating the corresponding mobility transaction of the OD matrix based on travel volume data (Tian et al. 2020). By converting the relative migration % into actual number of recorded movements, inter-city inflow and outflow data can be meaningfully compared to evaluate the geographic distribution of human mobility and COVID-19 cases.

Second, this study used spatiotemporal analysis at the prefecture-level instead of a spatially-invariant statistical analysis. While Fang, Wang, and Yang (2020) employed a counterfactual exercise to examine the effect of NPI in the spread, this study identifies and delineates the spatiotemporal clusters of COVID-19 before and after the travel ban to better understanding of the space-time dispersion of the early-stage pandemic among the inter-connected network of major Chinese cities. Hence, this study extends previous studies and reconstruct a comprehensive mobility dataset to meaningfully examine the geographic relationship between human mobility and COVID-19 in major Chinese cities based on spatiotemporal analysis.

Methods

Data

Baidu is an Information Technology (IT) company in China that offers a variety of Internet applications with over 348 million users (Baidu, 2016). Baidu provides locational-based services (e.g. search, maps) and the ability to make train and flight reservations. Baidu’s applications gather user data and the company shares migratory patterns across China at the city level (https://qianxi.baidu.com). In this research, we created a script that scraped the top 100 origins/destinations of in- and out-migration for all major cities in China. The script utilized city identification numbers which were discovered within Baidu’s javascript file commentary to query the JavaScript Object Notation (JSON) directly from an exposed Hypertext Transfer Protocol Secure (HTTPS) frontend Baidu database access point. This method of web scraping resulted in a very reliable and clean copy of the Baidu database which was immediately useable for research with minimal data cleanup. In total, about 3 million records were gathered between January 1–11 February 2020. In addition to these data that represented human mobility across cities, this study also used the Baidu mobility index, a relative score that ranks the mobility of a city over time. Moreover, confirmed cases of COVID-19 data in China
were acquired through web crawling from local government reports, published papers from academia, WHO situation reports, and newspapers. The database synthesized and consolidated daily confirmed cases of COVID-19 across China at the prefecture-level since December 2019. For comparative analysis, the daily count of confirmed cases were normalized by the census population and converted to cases per 1,000.

Reconstruction of mobility count
The mobility data obtained from Baidu have in- and out-migration % for 364 major Chinese cities. The percentage is relative to the total migrants for each city and hence cannot be used to compare human mobility across cities and over time. Baidu also publishes a migration index, a dimensionless mobility indicator of how ‘active’ the residents were based on various locational-based services, that is standardized across cities over time. To reconstruct the human mobility movement, this research used mobile phone data that captured the outflow movement from Wuhan (Tian et al. 2020) during Jan 11–25, 2020. The Baidu migration % was regressed against the estimated recorded movement from Tian et al. (2020) to model the trend and extrapolate the Baidu mobility index to dates during Jan 1 – Feb 11. As the mobility index changes quite a bit around the Chinese New Year festival as well as the travel ban, this study explored the fitting of various functional relationships to the recorded movement from mobile data into two time periods of Jan 1–25 and Jan 11 – Feb 11 separately. By equating the actual recorded movement with in- and out-migration % of Wuhan to another city, it is possible to derive the recorded movement to/from Wuhan and other city-pairs on the same day as well. For example, if the 10% in-migration of Wuhan from city A equates to 10,000 people, 20% in-migration from city B would be 20,000. As the inter-city migration data has duplicated entries of in-/out-migration between an origin and destination city (i.e. in-migration % of city X from city Y should have the same migrants as out-migration % of city Y to city X), it is also possible to expand the actual movement calculation by cross-referencing the corresponding records in the Origin/Destination matrix (OD matrix). Hence, the approximate movement from an origin city to all other Chinese cities could be derived based on cross-referencing the migration %, Baidu mobility index and recorded movement from mobile data.

Figure 1 shows the workflow for this study.

Geographic pattern of human mobility before/after Wuhan lockdown
Using the mobility counts stated above, daily in- and out-migration of Wuhan can display fluctuating migration patterns on the timeline. Daily in-migration of Wuhan was calculated by summing up all cities’ mobility counts towards Wuhan city on each day, whereas daily out-migration has the daily sum of mobility counts from Wuhan to other cities. When expanded to the Hubei province scale, daily statistics of mobility also verifies the significance of Wuhan among all Hubei cities. Statistical tests were used to examine any significant differences in migration and COVID-19 cases at the destination as a function of lockdown (before or after), provinces at origin and destination (Hubei or Non-Hubei), recorded movement and/or confirmed cases of COVID-19 at origin, respectively. This research used two separate full factorial Analysis of Variance (ANOVA) tests to examine any significant differences in human mobility and confirmed cases at destination by the main and interaction effects of different combinations of the three and five variables, respectively. In considering the incubation period of COVID-19, this study also evaluated the significant difference before/after Jan 28 (i.e. 5 days) and Feb 6 (i.e. 2 weeks) to examine the time-lag of city lockdown on COVID-19 as well.

Space-time cluster analysis of COVID-19 cases
The spatiotemporal cluster of COVID-19 confirmed cases in China from Jan 1 – Feb 11 was analysed (SaTScan 2020). This research uses SaTScan, a software that can analyse and detect spatial, temporal and space-time cluster by using Kulldorff’s spatial scan statistics. It is designed to perform geographical surveillance of disease, detect statistically significant spatial or space-time disease clusters and model the disease outbreaks by performing prospective real-time or time-periodic disease dispersion. Among the different models available in SaTScan to perform discrete and continuous scan statistics, this study used Poisson model and applied both discrete (discrete Poisson-based model) and continues (continuous Poisson model) scan statistics.

In this research, the confirmed case number for each city was used as the case file, the population of each city in 2015, the latest population data for all prefecture-level cities in this study, was used as the population file (National Bureau of Statistics 2020), and the latitude and longitude for each city were geocoded and used as the coordinates file. The cluster analysis used the
Poisson probability model to identify the space-time hotspots of high infection rates by progressively comparing the number of observed cases with expected cases of COVID-19 cases. The log-likelihood ratio (LLR) tests the null hypothesis of a randomly distributed confirmed cases at the 0.05 level based on Monte Carlo randomization. The space-time window with high LLR indicates the likelihood of a cluster. The higher value of LLR, the more likely to be identified as clusters. In this study, a sensitivity analysis of the population at risk from 10 to 50% have been tested, and 10% performed the best with reasonable results and fit better with the migration trend from Wuhan to other cities.

**Contextual analysis of migration pattern and confirmed cases**

This study also investigated the relationship, if any, between migration patterns around Wuhan and any infectious clusters. However, given that COVID-19 has a typical 5–14 days of incubation period, migration data were also examined at 14 or 5 days before the clusters’ lifespan. Despite the time-lags, the present study overlaid statistically significant space-time clusters and relevant migration patterns for a better understanding of context. Many potential patients might have visited clusters-to-be, but they might also have travelled in the opposite direction. It is also possible that a person might repeat visiting more than twice within a certain period (e.g. daily commute, short trip, logistics). Therefore, this study used the contextual maps to examine interaction rather than one-way influence based on the sum of in- and out-migration.

**Results**

After preliminary exploration of the Baidu migration index and recorded mobile data, a polynomial trend was deployed to model the undulating migration pattern of Wuhan during Jan 1–12 ($r^2 = 0.85$), and a power function to simulate the downward trend after Jan 26 ($r^2 = 0.92$). By cross-referencing the in- and out-migration of the origin-destination matrix, an estimated movement of 163,326 origin-destination pairs were derived during the study period (Figure 2).

Figure 2 shows the effect of lockdown policy on basic migration counts. Calculated out-migration was greatly restricted right after its culmination on Jan 23, signifying the exodus on the lockdown day. Paralleled with immigration, in the meanwhile, the total migration before the lockdown peaked twice on Jan 8 and 22. A possible
explanation of the first surge can be found in China CDC’s official announcement on the same day, that COVID-19 is the causative pathogen of the Chinese outbreak (Li et al. 2020a). The cause of the second peak is presumed to be a homebound rush before the travel restriction. It was estimated that more than 450,000 people were leaving Wuhan on the day before lockdown, and it significantly decreased to less than 5% by the end of the month and stayed low.

The count of human mobility across cities and the confirmed case data were not normally distributed, and hence a log transformation was applied to these two variables in the statistical analysis. The results of a full factorial ANOVA revealed significant differences of human mobility caused by the main and interaction effect of lockdown and/or the origin/destination in Hubei vs non-Hubei province or not (Figure 3, Table 1). The overall model was significant (adjusted $R^2 = 0.53$) and all effects were significant at the 0.001 level. It was evident that human mobility was significantly lower after the city lockdown, and the non-Hubei province were higher than Hubei province in both origin/destination. The interaction effects indicated that human mobility (Mig) dropped proportionally after the lockdown (LD) between Hubei and non-Hubei provinces as origin (Orig), but Hubei as destination (Dest) was significantly lower than non-Hubei destination (Figure 3).

With an adjusted $R^2$ of 0.75, the five-variable ANOVA test reported significant differences of COVID-19 cases at destination by the main effect of LD (increased after Jan 23), Mig (positive relationship) and Orig (higher if it’s from non-Hubei province) (Figure 4, Table 2). However, the main effects of Dest and the number of confirmed cases in the origin (CaseOrig) were not significant. In fact, the latter variable was not significant in any of the interaction effect. The interaction effects of LD were significant with all other variables except those involving CaseOrig (Table 2). Interestingly, the full factorial model using Jan 28 assuming a 5-day incubation period (adjusted $R^2 = 0.77$) showed significant differences in the main effect of OrigCase and more factors associated with its interaction effects. However, the main effect of Orig became not significantly different and there were less interactive effects. Assuming a 14-days incubation period, the main effect of Mig became insignificant but more factors associated with OrigCase revealed significant differences (adjusted $R^2 = 0.75$).

During the pre-lockdown period between Jan 1 and 23, six significant clusters were found (Figure 5). Wuhan city was identified as the first cluster from Jan 17–23. The 23 cities in Chongqing municipality, Hubei, and Hunan provinces were identified as the second cluster from Jan 22–23. Other clusters were identified between Jan 20–23 with 13 cities in Shanghai municipality, Zhejiang, and Fujian province (3rd cluster); 38 cities in Guangdong, Guangxi, Hainan provinces (4th cluster); Beijing municipality (5th cluster); and 23 cities in Henan, Shandong, Anhui, Jiangsu, and Hubei provinces (6th cluster). A total of 99 cities have been identified as significant clusters in or around Wuhan and other major cities. These city clusters were found in megacities, such as Beijing, Shanghai and Guangzhou, and metropolises,
such as Changsha, Chongqing and Shenzhen, along with other secondary cities close to those megacities and major cities.

Seven significant clusters were found for the post-lockdown period, starting from Jan 24 to Feb 11 (Figure 6). A significant cluster was identified with ten cities in Hubei province from Feb 2 to 10. The secondary clusters were identified with 29 cities in Hubei, Henan, Shanxi, Shaanxi, Gansu, and Sichuan provinces from Jan 29 to Feb 6. Clusters in Wenzhou (Jan 28 – Feb 5), Shenzhen (Jan 31 to Feb 2), Bengbu (Feb 5 to 9), Shuangyashan (Feb 6), Sanya and Baoting (Feb 4 to 8) were identified as third, fourth, fifth, sixth and seventh clusters, respectively. Overall, the significant spatiotemporal clusters were identified with cities in or nearby Hubei Province, and cities along the coastline. There are altogether 45 cities identified as significant clusters. Compare to pre-lockdown time period, fewer cities have been identified as clusters in post-lockdown time period.

**Figure 3.** Least square mean plots of ANOVA examining the significant difference of migration as a function of lockdown (LD), origin province (Orig) and destination province (Dest).

**Table 1.** Statistical results of a full factorial ANOVA on log migration (mig).

| Effect Test | Estimate | t-Ratio | Probability |
|-------------|----------|---------|-------------|
| LD          | 0.85     | 162.60  | < 0.0001**  |
| Orig        | -0.37    | -71.25  | < 0.0001**  |
| Dest        | -0.22    | -41.93  | < 0.0001**  |
| Orig * Dest | 1.26     | 240.82  | < 0.0001**  |
| LD * Orig   | 0.02     | 3.37    | 0.0008***   |
| LD * Dest   | 0.23     | 44.36   | < 0.0001**  |
| LD * Orig * Dest | -0.15 | -29.44  | < 0.0001**  |
These city clusters were found in metropolises, like Wuhan and Xi’an. Besides, some secondary cities around Xi’an, Wuhan, as well as some prefecture-level cities in or close to east costal line were identified as significant clusters in this study.

Pre-lockdown clusters were affected by the traveller’s outflow from Wuhan. Assuming an incubation period of 14 (i.e. Jan 3–9) and 5 days (i.e. Jan 3–9), major out-migration from Wuhan were going to clusters 2, 3, 4 & 5 and 2 & 5, respectively (Figure 7). This spatiotemporal association was not prominent in nearby cities but distant ones like Beijing, Shanghai, Guangzhou, and Chongqing – they had the largest share of migration from Wuhan.

Regarding post-lockdown clusters, only clusters 1 and 2 were found to have some association between migration and confirmed cases (Figure 8). The first cluster centres around Wuhan and the eastern cities within Hubei where there was a high concentration of mobility during Jan 19–27. It was noted that the travel was much minimized from Jan 28 – Feb 5. The second cluster was anchored around Xi’an which did not have sizable migration with Wuhan and any other Hubei cities. In fact, Xi’an was well connected with Yulin, Beijing, and Shanghai during Jan 15–23 (i.e. 14-day prior), and nearby cities within Shaanxi province during Jan 24 – Feb 1 (i.e. 5-day prior). Otherwise, Xi’an had only a modest connection with other significant clusters of COVID-19 confirmed cases (Figure 9).

Figure 4. Least square mean plots of ANOVA examining the significant difference of COVID-19 cases at destination as a function of lockdown (LD), origin (Orig) and destination province (Dest), log of migration (Mig) and confirmed cases at origin (OrigCase). Only variables with least square mean plots can be shown here. More effects are documented in Table 2.
Table 2. Statistical results of a full factorial ANOVA on log confirmed cases of COVID-19 at destination (CaseDest).

| Effect Test       | Prob > F (Jan 23) | Prob > F (Jan 28) | Prob > F (Feb 6) |
|-------------------|--------------------|--------------------|-----------------|
| LD                | <.0001**           | <.0001**           | <.0001**        |
| Orig              | 0.0115*            | 0.2577             | <.0001**        |
| Dest              | 0.4625             | <.0001**           | <.0001**        |
| Mig               | 0.0032**           | 0.0052*            | 0.7164          |
| CaseOrig          | 0.5212             | <.0001**           | <.0001**        |
| Orig*Dest         | 0.0066**           | 0.0018             | 0.0077**        |
| LD*Dest           | 0.0007**           | <.0001**           | <.0001**        |
| LD*Dest*Mig       | 0.0003**           | <.0001**           | <.0001**        |
| LD*Mig            | 0.0007**           | <.0001**           | <.0001**        |
| LD*Orig           | 0.0268             | 0.1777             | <.0001**        |
| LD*Orig*Dest      | 0.0078*            | <.0001**           | 0.0775          |
| LD*Orig*Dest*Mig  | 0.0118*            | 0.4397             | 0.0122**        |
| LD*Orig*Mig       | 0.0255*            | 0.6111             | <.0001**        |
| Mig*Dest          | 0.0028**           | <.0001**           | <.0001**        |
| Mig*Orig          | 0.0753             | 0.6336             | <.0001**        |
| Mig*Orig*Dest     | 0.0111*            | 0.6463             | 0.8879          |
| Orig*CaseDest     | 0.6866             | 0.0101*            | 0.4653          |
| Orig*CaseDest*Mig | 0.2078             | 0.4104             | 0.6501          |
| Orig*Case*LD      | 0.9112             | <.0001**           | <.0001**        |
| Orig*Case*LD*Dest | 0.8704             | 0.0020**           | <.0001**        |
| Orig*Case*LD*Mig  | 0.8824             | <.0001**           | <.0001**        |
| Orig*Case*LD*Mig  | 0.6664             | 0.3248             | <.0001**        |
| Orig*Case*LD*Orig | 0.6468             | 0.0867             | 0.7500          |
| Orig*Case*LD*Orig*Dest | 0.6081 | 0.6434 | 0.0015** |
| Orig*Case*LD*Orig*Dest*Mig | 0.2014 | 0.4028 | <.0001** |
| Orig*Case*LD*Orig*Mig | 0.1097 | 0.1394 | 0.0050** |
| Orig*Case*Mig     | 0.3468             | 0.4449             | <.0001**        |
| Orig*Case*Orig    | 0.6749             | 0.6765             | 0.2104          |
| Orig*Case*Orig*Dest | 0.9771 | 0.0473* | 0.6204 |
| Orig*Case*Orig*Dest*Mig | 0.1510 | 0.0035** | 0.0686 |
| Orig*Case*Orig*Mig | 0.1264 | 0.0696 | 0.1426 |

1 lockdown (LD: Before, After]), origin province (Orig: [Hubei, Non-Hubei]), destination province (Dest: [Hubei, Non-Hubei]), log of migration (Mig: count), log of confirmed cases at origin (CaseOrig: case per 1,000)
2* Significant at the 0.05 level; ** Significant at the 0.01 level

Figure 5. Spatiotemporal cluster cities in pre-lockdown time period.

Figure 6. Spatiotemporal cluster cities in post-lockdown time period.

Figure 7. Relationship between pre-lockdown clusters and migration.

Discussion & conclusion

As indicated by the main effects, migration patterns significantly differed by when (i.e. city lockdown and where (i.e. origin and destination). There were also interaction effects where travel restrictions caused a significant reduction in human mobility (Tian et al. 2020), especially inside Hubei as both origin and destination (Figure 3, Table 1). Despite the lockdown, there were still some travelling in and out of Wuhan and the rest of Hubei, indicating the volume of essential travels during this sensitive time. In addition to these antecedent factors, it was also found that migration and confirmed cases at origin were significantly associated with COVID-19 cases at the destination by lockdown and location (Figure 4 and Table 2). As expected, there were significant increase of confirmed cases after the
lockdown (Du et al. 2020) and travel restrictions alone could not contain the infection (Dénes and Gumel 2019).

An interesting finding from this study was that the models assuming an incubation period of 5 days differentiated the differences of cases better than the baseline and 14 days models (Table 2). Although the adjusted $R^2$ only slightly improved in the time-lag model of 5 days, different variables were involved in the main and interaction effects, respectively. For example, it is noted that OrigCase became significant in the time-lag models while migration became less significant in the 14-days model, implying the likelihood of some super spreaders who might be responsible at early stage and the dissipating effect of migration but local outbreaks in the latter stage of regional spreading. Considering the temporal trait of COVID-19 and its variable incubation period, the reported cases at a specific time only reflects the cases confirmed at a previous time. While symptoms may appear after exposure to the virus, ongoing research suggests a median incubation period of

**Figure 8.** Relationship between post-lockdown cluster 1 and migration.

**Figure 9.** Relationship between post-lockdown cluster 2 and migration.
5.1 days and about 99% of patients would develop symptoms within 14 days (World Health Organization 2020; Lauer et al. 2020). It is noteworthy to state that the Chinese COVID-19 data do not include asymptomatic patients as the Chinese CDC do not release such data until April. As asymptomatic and pre-symptomatic carriers of COVID-19 could be attributed to the widespread of this infectious disease, the temporal trait of COVID-19 and time lag in the reported data contributes to the uncertainty of this epidemic infection. Future studies may examine how incubation period might vary in time and space as it evolves and their impacts in subsequent analyses.

In consistent with a previous study (Mo et al. 2020), the spatiotemporal clusters of confirmed cases were mainly found among Hubei cities nearby inland provinces and coastal cities within the study period. Assuming a 5–14 days incubation period, it was found that the migration among Wuhan and other cities were associated with the resulting spatiotemporal clusters of confirmed cases in Wuhan from Jan 17–23. The clustering during the pre-lockdown period could be attributed to the fact that most cities already have many infected travellers moving in and out before the Hubei lockdown (Du et al. 2020). Indeed, the pre-lockdown clusters due to traveller’s outflow from Wuhan to those megacities were the pathway for international transmission of COVID-19 (Chinazzi et al. 2020). It is also noteworthy that human mobility pattern contributed to the emergence of infectious clusters, more than Cartesian proximity from the pathologic epicentre did. Nanchang, the capital of Jiangxi province, did not exchange sizable travellers between Wuhan despite its close distance – approximately 260 kilometres. As a result, the city was not included in pre- nor post-lockdown clusters.

Moreover, clusters of confirmed cases identified within the post-lockdown period were partially related to the migration pattern (Du et al. 2020). This association is particularly profound with the eastern part of Hubei around Wuhan before the lockdown (Figure 8). As travel was minimized after Jan 23, this cluster may illustrate a long incubation period of COVID-19. The weak association between migration and COVID-19 clusters in other regions after the lockdown may indicate that there might be the 1) presence of ‘super spreaders’ in these clusters, 2) travel not reflected in the migration data, and/or 3) an uncertain incubation period. Nevertheless, this finding suggests that there have been local outbreaks in those clusters that the migration data alone may not be sufficient to explain the clustering of COVID-19. This does not, however, diminish the need to continue existing NPI already-in-place to minimize the intermixing and interaction of infected and uninfected populations across regions.

The spatiotemporal clusters results indicated the risk distribution and trends of COVID-19-confirmed cases, which helped us have a deep understanding of the diffusion of COVID-19 in terms of time and space. Besides, the finding of clusters for confirmed cases would provide vital information for health providers and policymakers to allocate the healthcare resources and take actions to control and mitigate the pandemic. That information is of great significance to the scientific prevention and control of infectious diseases and emergency management. However, this study did not differentiate various practices of NPI and distinguished their individual impacts on the spatiotemporal clustering of COVID-19 cases. It also did not quantify the functional relationship between mobility and COVID-19 that might be appropriate only with more data (e.g. hospitalization, healthcare resources) to support pathway analysis of COVID-19. Future studies can also be directed to explore the reasons related to those identified clusters through structural equation modelling and/or a mixed-method approach.

**Acknowledgements**

The authors appreciate those who have collected, prepared, and shared data throughout this outbreak. We are also grateful to the statistical consulting from Phil Vaughan at the Texas State University.

**Disclosure statement**

The authors declared no conflict of interest in the preparation and publication of this manuscript.

**ORCID**

T. Edwin Chow [http://orcid.org/0000-0002-0386-5902](http://orcid.org/0000-0002-0386-5902)

Yusik Choi [http://orcid.org/0000-0001-8021-2756](http://orcid.org/0000-0001-8021-2756)

Mei Yang [http://orcid.org/0000-0002-6168-0084](http://orcid.org/0000-0002-6168-0084)

David Mills [http://orcid.org/0000-0002-2494-9187](http://orcid.org/0000-0002-2494-9187)

Ricci Yue [http://orcid.org/0000-0002-8564-8556](http://orcid.org/0000-0002-8564-8556)

**References**

Baidu. 2016. “Report of Foreign Private Issuer Pursuant to Rule 13a-6 or 15d-16 Under the Securities Exchange Act of United States Securities and Exchange Commission.” 1 July 2016, ir.baidu.com/index.php/static-files/98d935f3-b39d-40ae-a514-a4e5e9bca424. Accessed 11 Sept. 2020

Chinazzi, M., J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. P. Piontti, et al. 2020. “The Effect of Travel Restrictions on the Spread of the 2019 Novel Coronavirus (COVID-19) Outbreak.” Science 368 (6489): 395–400. doi:10.1126/science.aba9757.
Cucinotta, D. and M. Vanelli. 2020. “WHO Declares COVID-19 a Pandemic.” *ACTA Biomedica* 91(1): 157–160. doi: 10.23750/abm.v91i1.9397.

Dénes, A., and A. B. Gumel. 2019. “Modeling the Impact of Quarantine during an Outbreak of Ebola Virus Disease.” *Infectious Disease Modelling* 4: 12–27. doi:10.1016/j.idm.2019.01.003.

Dong, E., H. Du, L. Gardner. 2020. “An interactive web-based dashboard to track COVID-19 in real time.” *The Lancet: Infectious Disease* 20(5): 533–534. at https://doi.org/10.1016/S1473-3099(20)30120-1.

Du, Z., X. Xu, Y. Wu, L. Wang, B. J. Cowling, and L. A. Meyers. 2020. “Serial Interval of COVID-19 among Publicly Reported Confirmed Cases.” *Emerging Infectious Diseases* 26 (6): 1341–1343. doi:10.3201/eid2606.200357.

Fang, H., L. Wang, and Y. Yang. 2020. “Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-ncov) in China (No. W26906),” *National Bureau of Economic Research*. doi:10.3386/w26906.

Hatchett, R. J., C. E. Mecher, and M. Lipsitch. 2007. “Public Health Interventions and Epidemic Intensity during the 1918 Influenza Pandemic.” *Proceedings of the National Academy of Sciences* 104 (18): 7582–7587. doi:10.1073/pnas.0610941110.

Jia, J. S., X. Lu, Y. Yuan, G. Xu, J. Jia, and N. A. Christakis. 2020. “Population Flow Drives Spatio-temporal Distribution of COVID-19 in China.” *Nature* 582 (7812):389-394. https://doi.org/10.1038/s41586-020-2284-y.

Kraemer, M. U. G., C. H. Yang, B. Gutierrez, C. H. Wu, B. Klein, D. M. Pigott, L. Du Plessis, et al., . 2020. “The Effect of Human Mobility and Control Measures on the COVID-19 Epidemic in China.” *Science* 368 (6490): 493–497. doi:10.1126.science.abb4218.

Lai, P. C., C. M. Wong, A. J. Hedley, S. V. Lo, P. Y. Leung, J. Kong, and G. M. Leung. 2004. “Understanding the Spatial Clustering of Severe Acute Respiratory Syndrome (SARS) in Hong Kong.” *Environmental Health Perspectives* 112 (15): 1550–1556.

Lauer, S. A., K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler. 2020. “The Incubation Period of Coronavirus Disease 2019 (COVID-19) from Publicly Reported Confirmed Cases: Estimation and Application.” *Annals of Internal Medicine* 172 (9): 577–583. doi:10.7326/M20-0504.

Li, Q., X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, et al. 2020a. “Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–infected Pneumonia.” *The New England Journal of Medicine* 382 (13): 1199–1207. doi:10.1056/NEJMoa2001316.

Li, R., S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, and J. Shaman. 2020b. “Substantial Undocumented Infection Facilitates the Rapid Dissemination of Novel Coronavirus (Sars-cov-2).” *Science* 368 (6490): 489–493. doi:10.1126/science.abb3221.

Matrajt, L., and T. Leung. 2020. “Evaluating the Effectiveness of Social Distancing Interventions to Delay or Flatten the Epidemic Curve of Coronavirus Disease.” *Emerging Infectious Diseases* 26 (8): 1740–1748. doi:10.3201/eid2608.201093.

Mo, C., D. Tan, T. Mai, C. Bei, J. Qin, W. Pang, and Z. Zhang. 2020. “An Analysis of Spatiotemporal Pattern for COVID-19 in China Based on Space-time Cube.” *Journal of Medical Virology* 92: 1587–1595.

National Bureau of Statistics. 2020. “Websites of Local Statistics.” Accessed 23 April 2020 http://www.stats.gov.cn/tjgz/wzlj/dftjwz/.

Pan, A., L. Liu, C. Wang, H. Guo, X. Hao, Q. Wang, J. Huang, et al. 2020. “Association of Public Health Interventions with the Epidemiology of the COVID-19 Outbreak in Wuhan, China.” *The Journal of the American Medical Association* 323 (19): 1915–1923. doi:10.1001/jama.2020.6130.

Rezaeeian, M., G. Dunn, S. St Leger, and L. Appleby. 2007 “Geographical Epidemiology, Spatial Analysis and Geographical Information Systems: A Multidisciplinary Glossary.” *Journal of Epidemiol Community Health* 61 (2): 98–102. doi:10.1136/jech.2005.043117.

Sanche, S., Y. T. Lin, C. Xu, E. Romero-Severson, N. Hengartner, and R. Ke. 2020. “High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2.” *Emerging Infectious Diseases* 26 (7): 1470–1477. doi:10.3201/eid2607.200282.

SaTScan. 2020. “SaTScan V9.6: Software for the Spatial, Temporal, and Space-time Scan Statistics.” Accessed 23 April 2020. https://www.satscan.org/.

The Guardian. 2020. “China Coronavirus: Mayor of Wuhan Admits Mistakes.” Accessed 30 May 2020 https://www.the guardian.com/science/2020/jan/27/china-coronavirus-who-to-hold-special-meeting-in-beijing-as-death-toll-jumps.

Tian, H., Y. Liu, Y. Li, C. H. Wu, B. Chen, M. U. G. Kraemer, B. Li, et al. 2020. “An Investigation of Transmission Control Measures during the First 50 Days of the COVID-19 Epidemic in China.” *Science* 368 (6491): 638–642. doi:10.1126/science.abb6105.

Tobías, A. 2020. “Evaluation of the Lockdowns for the SARS-Cov-2 Epidemic in Italy and Spain after One Month Follow Up.” *Science of the Total Environment* 725: 138539. doi:10.1016/j.scitotenv.2020.138539.

World Health Organization. 2020. “Coronavirus Disease 2019 (COVID-19) Situation Report – 73.” Accessed 23 May 2020. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf?sfvrsn=5ae25bc7_4#;text=The%20incubation%20period%20for%20COVID19%20range%20between%200%20and%202%20days%20after%20contact%20with%20infected%20individuals.

World Health Organization Writing Group. 2006. “Nonpharmaceutical Interventions for Pandemic Influenza, International Measures.” *Emerging Infectious Diseases* 12 (1): 81-87. doi:10.3201/eid1201.051370.

Yuan, Y., Y. Qiang, K. Bin Asad, and T. E. Chow. 2020. “Point Pattern Analysis.” Accessed 30 May 2020 https://gistbok. ucgis.org/bok-topics/point-pattern-analysis.