



OPEN

Interregional mobility in different age groups is associated with COVID-19 transmission in the Taipei metropolitan area, Taiwan

Wei-Ming Jiang¹, Tzai-Hung Wen², Ying-Chi Huang³, Hung-Yi Chiou^{1,4,5}, Wei J. Chen^{6,7}, Chao A. Hsiung¹, Huey-Kang Sytwu³ & Hsiao-Hui Tsou^{1,8}✉

Before vaccines were introduced, mobility restriction was one of the primary control measures in the early stage of the coronavirus disease 2019 (COVID-19) pandemic. Because different age groups face disproportionate health risks, differences in their mobility changes affect the effectiveness of pandemic control measures. This study aimed to investigate the relationship between multiscale mobility patterns in different age groups and COVID-19 transmission before and after control measures implementation. Data on daily confirmed case numbers, anonymized mobile phone data, and 38 socioeconomic factors were used to construct negative binomial regression models of these relationships in the Taipei metropolitan area in May 2021. To avoid overfitting, the socioeconomic factor dimensions were reduced by principal component analysis. The results showed that inter-district mobility was a greater promoter of COVID-19 transmission than was intra-district mobility (coefficients: pre-alert, 0.52 and 0.43; post-alert, 0.41 and 0.36, respectively). Moreover, both the inter-district mobility of people aged 15–59 and ≥ 60 years were significantly related to the number of confirmed cases (coefficients: pre-alert, 0.82 and 1.05; post-alert, 0.48 and 0.66, respectively). The results can help agencies worldwide formulate public health responses to emerging infectious diseases.

From the end of December 2019, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) caused a coronavirus disease 2019 (COVID-19) pandemic worldwide. By November 2022, more than 632 million confirmed COVID-19 cases and more than 6 million deaths with COVID-19 had been reported across 193 countries and regions¹. Because COVID-19 is a respiratory disease, most countries quickly implemented control measures including mobility restrictions (e.g., travel) and non-pharmaceutical interventions (NPIs; e.g., social distancing, face mask use) to mitigate outbreaks before vaccines were distributed widely. Before the implementation of mobility restrictions, COVID-19 should spread relatively randomly and be uniformly distributed in population flows due to a relative lack of awareness². Because human mobility can result in early shifts of the epidemiological dynamics from epidemics driven by frequent importation into community transmission³, human mobility played an important role in COVID-19 transmission^{4,5}.

Several studies have analyzed the relationship between human mobility and COVID-19 transmission by spatial-temporal mobility data from various sources^{4,5}. These studies mainly relied on three types of mobility data. The first type of mobility data is public transit systems such as buses, trains, metro, ferries, and air flights. The mobility data including service capability, travel time, trip information, and boarding and alighting locations

¹Institute of Population Health Sciences, National Health Research Institutes, 35 Keyan Road, Zhunan 350, Miaoli County, Taiwan. ²Department of Geography, National Taiwan University, Taipei, Taiwan. ³National Institute of Infectious Diseases and Vaccinology, National Health Research Institutes, Zhunan, Miaoli County, Taiwan. ⁴School of Public Health, College of Public Health, Taipei Medical University, Taipei, Taiwan. ⁵Master's Program in Applied Epidemiology, College of Public Health, Taipei Medical University, Taipei, Taiwan. ⁶Center for Neuropsychiatric Research, National Health Research Institutes, Zhunan, Miaoli County, Taiwan. ⁷Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan. ⁸Graduate Institute of Biostatistics, College of Public Health, China Medical University, Taichung, Taiwan. ✉email: tsouhh@nhri.org.tw

can be retrieved from scheduled timetable data and actual travel records of passengers^{6–8}. However, this type of mobility data is only able to cover the mobility changes between boarding and alighting locations in public transit systems. The second type of mobility data is social media mobile apps such as Facebook, Twitter, Google Mobility Report, and Baidu Mobility Index^{3,9–13}. Social media mobile apps can offer aggregated and anonymized mobility data from users who have turned on the global positioning system (GPS) or posted content with geotagging. However, the limitations of each social media mobile app are different. For example, Baidu Mobility Index only contains mobility data in Mainland China. Although Google Mobility Report can provide mobility changes to the point of interest such as residential areas, workplaces, and transit stations, it cannot estimate inter-regional mobility. The third type of mobility data is mobile network operators. Mobile network operators decode signaling information between opening mobile phones and cell sites to provide anonymized and aggregated mobility data. The advantage of mobile phone data is that it is not limited to the analysis of mobility changes in boarding and alighting stations or particular locations¹⁴. Mobile phone data can offer extensive and representative mobility information at the population level to investigate intra- and inter-regional mobility changes between origin–destination geographic scales such as city and county^{2,15–17}.

According to the origin and destination locations of human mobility, mobility patterns can be classified into two types: intra-regional flows and inter-regional flows^{18,19}. The intra-regional flows are mainly responsible for local diffusion within specific areas such as workplaces and residential locations. For example, Nagata et al. (2021) found that the mobility changes at nightlife locations were related more significantly to COVID-19 outbreaks than were changes at workplaces and residential locations²⁰. On the other hand, the inter-regional flows contribute to the spatial spread of the epidemic from outbreak locales to other unaffected or weakly affected areas. For example, Jia et al. (2020) predicted the geographical distribution of infections according to population outflows from Wuhan to other provinces in China between 1 and 24 January 2020². Because the intra- and inter-regional flows play different roles in the spread of the epidemic, evaluating their impacts on transmission can help agencies formulate more specific prevention policies to mitigate COVID-19 outbreaks.

Given the age-dependent effects on COVID-19 transmission and health risk, researchers have investigated mobility changes in different age groups after control measures implementation^{21–23}. For instance, using anonymized and aggregated mobility indicators, Caselli et al. (2022) found that the mobility of people aged 18–24 years in Italy, Portugal, and Spain declined by > 20% after the 2020 introduction of stay-at-home orders, whereas that of people aged ≥ 65 years declined by only 11%, representing a significant difference²². To prevent bias and the mismeasurement of control measure effects due to the aggregation of data from all age groups, the heterogeneous responses of people in different age groups to control measures need to be clarified. However, there is still a lack of research to compare the impacts of intra- and inter-regional flows in different age groups on COVID-19 transmission.

This study aims to investigate the relationship between inter- and intra-regional population flows in different age groups and COVID-19 transmission before and after control measures implementation. There are two research questions. First, which one of inter- and intra-regional flows was related more significantly to COVID-19 transmission? Second, whether there was a significantly different relationship between mobility patterns in different age groups and COVID-19 transmission? To answer these two research questions, the study conducted negative binomial regression models to analyze the relationship between mobility patterns in different age groups and the number of confirmed COVID-19 cases before and after control measures implementation. In addition, socioeconomic factors are the important determinants of COVID-19 infection^{6,24,25}. Considering the influence of socioeconomic factors on the relationship between human mobility and COVID-19 transmissions, this study included 38 socioeconomic factors as covariates in regression models. To avoid overfitting and collinearity, socioeconomic factor dimensions were reduced by principal component analysis.

Due to the implementation of a strict quarantine policy, $\geq 90\%$ of COVID-19 cases recorded in Taiwan through April 2021 were imported from other countries²⁶. In May 2021, the B.1.1.7 variant of SARS-CoV-2 caused the first local outbreak in the Taipei metropolitan area. To control it, the National Health Command Center declared a level-3 COVID-19 alert in this area on 15 May 2021. All schools were closed, with teaching conducted online; members of the public were advised to avoid unnecessary travel; face mask use outside of the home was required at all times; indoor gatherings of > 5 people and outdoor gatherings of > 10 people were prohibited. Because this wave of the epidemic was independent and the timing of policy intervention is clear, we selected this context to compare the relationships between mobility patterns, overall and by age group, and COVID-19 transmission before and after control measures implementation.

Study area

The study area comprises 29 districts in the Taipei metropolitan area, which consists of Taipei City and New Taipei City and is the most densely populated region in Taiwan (Fig. 1). Most confirmed COVID-19 cases reported between 3 and 31 May 2021 were in urban areas. According to the definition of the Department of Statistics of Taiwan's Ministry of the Interior, this study selected 29 districts, with population densities ≥ 1000 people/km², as the study area²⁷. In Taiwan, districts are administrative subdivisions of special municipalities and provincial cities. Although most people's daily activities occur in their districts of residence, > 68% of people in the Taipei metropolitan area commute across districts daily²⁸.

Methods

Research designs

Figure 2a shows the flow chart of the overall analysis in this study. The dependent variable and the independent variable were the numbers of confirmed COVID-19 cases and population flows, respectively. Socioeconomic factors were included as covariates. There were three steps for the preprocessing of population flows. First,

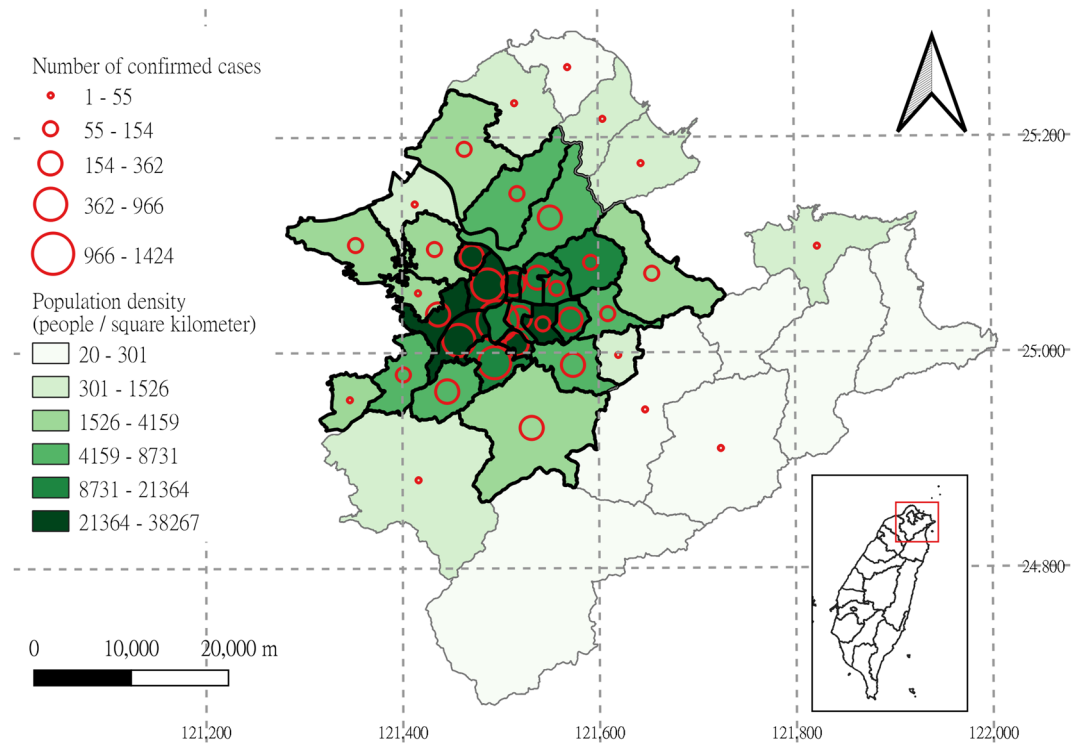


Figure 1. Study area. The study area comprises 29 districts in the Taipei metropolitan area (black outline). The numbers of confirmed COVID-19 cases were accumulated from 3 to 31 May 2021.

the population flows were classified into two categories: intra-district flows and inter-district flows. Second, the population flows were sorted into people aged 15–59 years and people ≥ 60 years by age groups. Third, the population flows were standardized by means and standard deviations to have unit variance which is beneficial to compare regression coefficients. On the other hand, this study conducted the principal component analysis to obtain principal components of the socioeconomic factors. To compare the changes in the relationship between human mobility and COVID-19 transmission, the study selected three periods before and after the COVID-19 alert implementation on 15 May 2021 (Fig. 2b). Because of the COVID-19 incubation period and reporting delay, the population flows a few days previously led to the number of new confirmed COVID-19 cases in each period.

Data

Daily district-level data on the numbers of confirmed COVID-19 cases in Taiwan were obtained from the Taiwan Centers for Disease Control's open data portal²⁹. Data on local (i.e., not imported) cases occurring in the study area between 5 and 31 May 2021 were extracted, according to patients' districts of residence and self-reported onset dates. During this period, COVID-19 cases were confirmed by the polymerase chain reaction. In the study area, the daily number of new confirmed COVID-19 cases increased steeply at the beginning of the outbreak, peaked on May 15, and then fluctuated dramatically between 16 and 31 May (Fig. 3a).

Population flow was examined using daily mobile phone data obtained from Far EasTone Telecommunications, a Taiwan-based company that held roughly 24.6% of the national market share in 2019³⁰. The company decoded information on signals between active mobile phones and cell sites to provide anonymized and aggregated mobility data (millions of data points) for the study area and period. The number of trips was weighted by poststratification using national census data³¹. Therefore, the mobile phone data were considered to comprise a representative sample of the population flow structure in the study area. This mobile phone data is origin–destination data which is measured as the total number of trips from origin to destination per day. The spatial resolution of the original mobile phone data was at the village level. According to the origin and destination villages, population flow was aggregated at the district level and classified as intra-district flows and inter-district flows. Intra- and inter-district flows decreased visually following control measure implementation on 15 May (Fig. 3b,c). Population flow was also determined separately for people aged 15–59 and ≥ 60 years. To eliminate the effects of daily fluctuations, standardized 7-day average population flows were calculated.

District-level data on 38 socioeconomic (demographic, education, economic, healthcare, equipment modernity, and land use) factors were collected from the government's open data platforms^{27,32,33} (Supplementary Table S1). Data pertaining to months as close as possible to May 2021 were extracted.

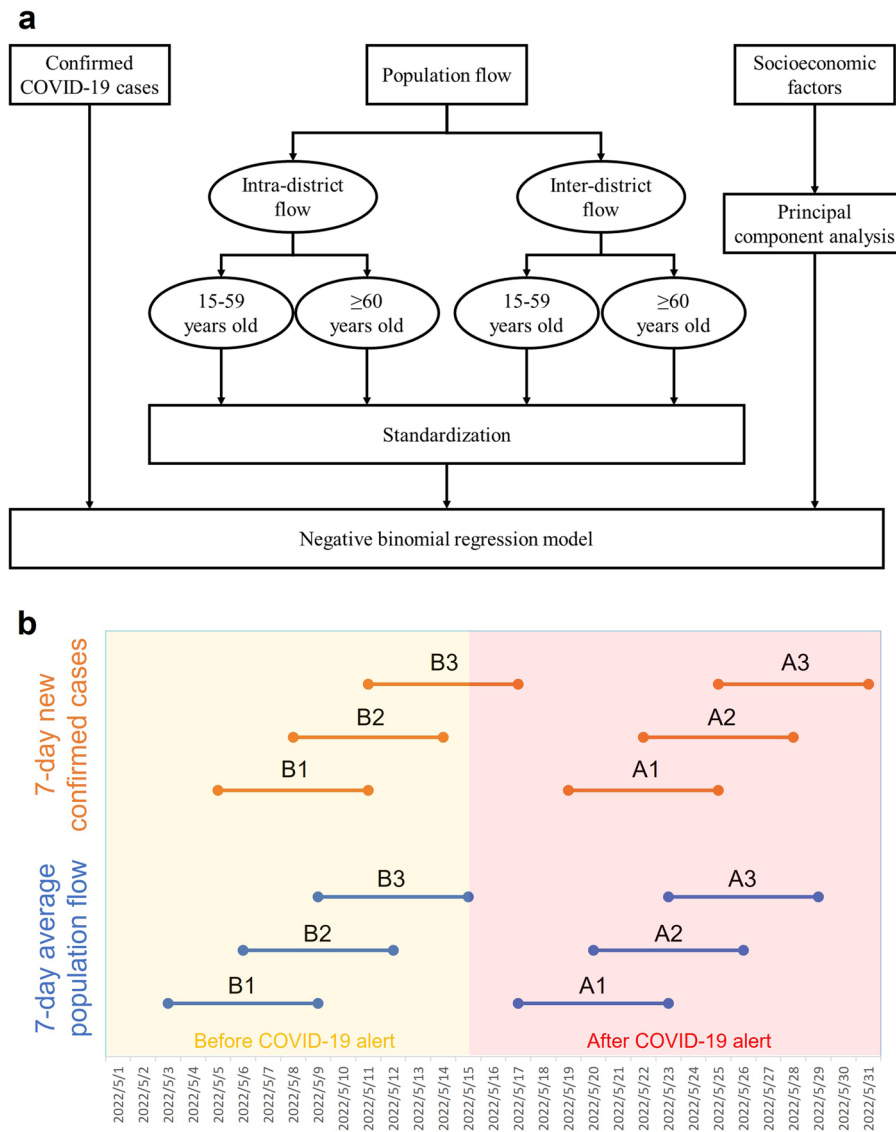


Figure 2. Flow chart and model structure. The flow chart of the overall analysis (a) and the model structure of regression models (b). B1 to B3 and A1 to A3 represent the three selected periods before and after the COVID-19 alert implementation, respectively.

Statistical analysis

Because the number of socioeconomic variables was too large for simultaneous model fitting, principal component analysis was conducted to resolve model overfitting and/or collinearity^{6,34}. If most of the total population variance could be attributed to the first one, two, or three principal components (PCs), those components could replace the original variables without much loss of information. Because the socioeconomic variables were measured at widely differing scales, they were standardized to avoid the domination of loadings by specific variables. PCs were obtained from the standardized variables using the following formula³⁵.

$$Z_i = e_i'(X - \mu)(V^{1/2})^{-1}, \quad i = 1, 2, \dots, p,$$

where Z_i represents the PCs, e_i' represents the transpose of eigenvectors, X represents the socioeconomic variables, μ represents the means of the socioeconomic variables, and $(V^{1/2})^{-1}$ represents the inverse of diagonal standard deviation matrices.

Because confirmed case variance is often greater than the mean, referred to as overdispersion^{8,36}, the data were analyzed using negative binomial regression models. These models can be used to analyze strictly positive response variables, such as counts and rates, and variable variance and means do not need to be equal. In these models, the dependent and independent variables were the numbers of confirmed COVID-19 cases and population flow, respectively, and the socioeconomic factors were included as covariates. Coefficients obtained before

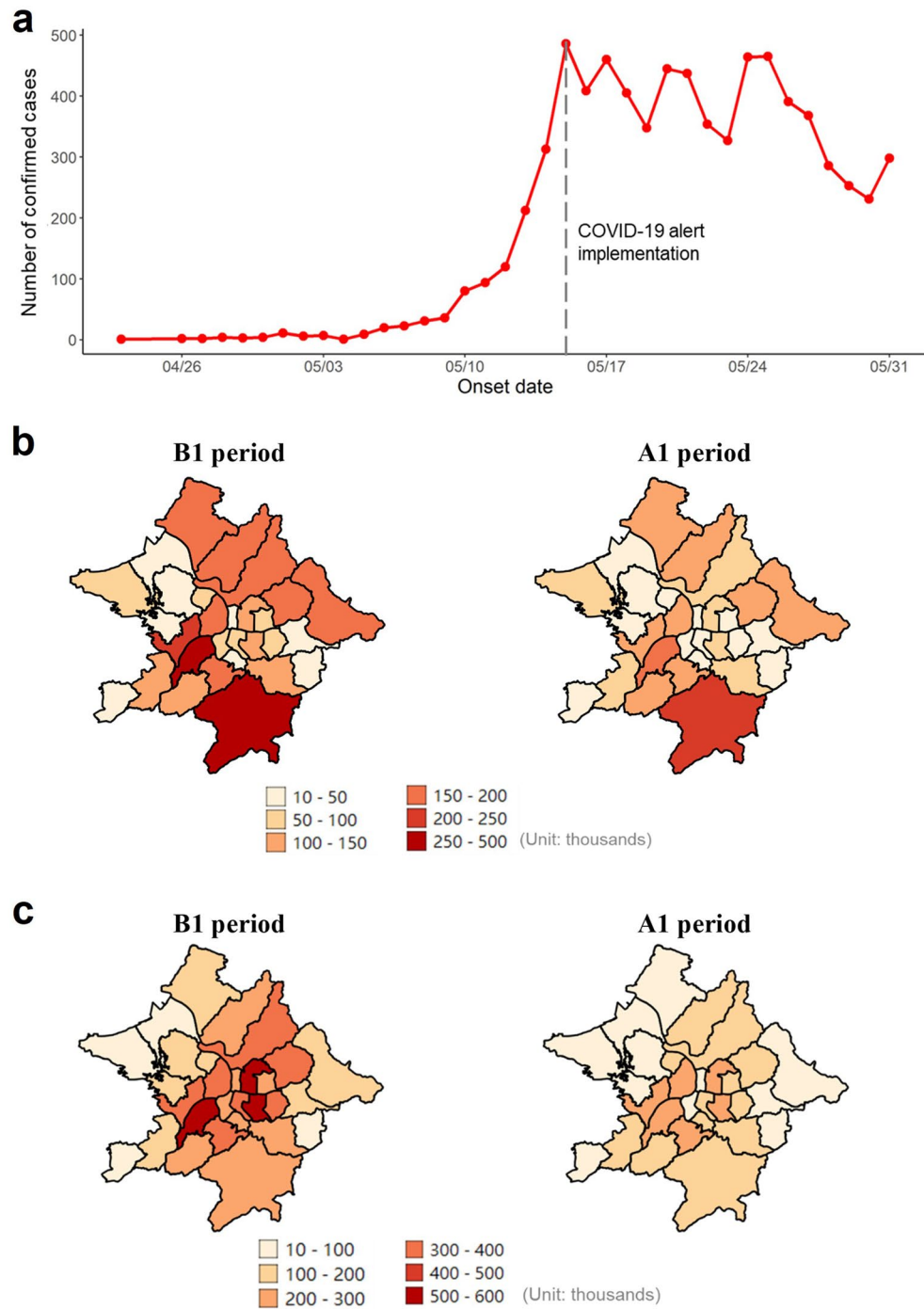


Figure 3. Change in the number of confirmed COVID-19 cases and different mobility patterns. The daily number of confirmed new COVID-19 cases (a), 7-day average of the total number of district-level intraflows (b), and interflows (c) following level-3 COVID-19 alert. B1 period, 3–9 May 2021; A1 period, 17–23 May 2021.

and after level-3 COVID-19 alert implementation on 15 May 2021 were compared. Negative binomial regression was conducted using the MASS package in R (version 3.6.3) and the following formula^{37,38}:

$$\ln E(y_i) = \alpha + \beta_1 \cdot \sum_{j=1, j \neq i}^m x_{ji} + \beta_2 \cdot x_{ii} + \sum_{k=1}^n \gamma_k \cdot z_k,$$

where $E(y_i)$ represents the expected number of confirmed COVID-19 cases in i district; α represents the intercept; x_{ji} represents the standardized 7-day average inflow from district j to district i ; m represents the number

of districts in the study area; x_{ji} represents the standardized 7-day average intra-district flow within district i ; z_k represents the PCs of the socioeconomic factors; n represents the number of PCs selected; and β_1 , β_2 , and γ_k are regression coefficients.

More specifically, the following formula was conducted when the inter- and intra-flows were separated into people aged 15–59 years and people ≥ 60 years by age groups.

$$\ln E(y_i) = \alpha + \beta_1^{Age15-59} \cdot \sum_{j=1, j \neq i}^m x_{ji}^{Age15-59} + \beta_1^{Age \geq 60} \cdot \sum_{j=1, j \neq i}^m x_{ji}^{Age \geq 60} + \beta_2^{Age15-59} \cdot x_{ii}^{Age15-59} + \beta_2^{Age \geq 60} \cdot x_{ii}^{Age \geq 60} + \sum_{k=1}^n \gamma_k \cdot z_k,$$

where $x_{ji}^{Age15-59}$ and $x_{ji}^{Age \geq 60}$ represents the standardized 7-day average inflows of people aged 15–59 years and people ≥ 60 years from district j to district i , respectively; $x_{ii}^{Age15-59}$ and $x_{ii}^{Age \geq 60}$ represents the standardized 7-day average intra-district flows of people aged 15–59 years and people ≥ 60 years within district i , respectively; and $\beta_1^{Age15-59}$, $\beta_1^{Age \geq 60}$, $\beta_2^{Age15-59}$, and $\beta_2^{Age \geq 60}$ are corresponding regression coefficients.

The maps were visualized using free and open geographic information system software QGIS (version 3.14), and the charts were plotted via the ggplot2 package (version 3.3.3) in R and Microsoft Excel (version 2210)^{39–41}.

Ethical approval

The institutional review board (IRB) of the National Health Research Institutes approved this study (EC1091110-E-R1). This study does not involve live animals and individual human participants. The requirement for informed consent from the study subjects was waived by the IRB of the National Health Research Institutes due to the reason that the mobile phone data analyzed in this study are de-identification, anonymized, and aggregated on a district level. No personally identifiable information, such as an individual's location, contacts, or movement, was made available at any point. This study confirms that all methods were performed in accordance with the relevant guidelines and regulations.

Results

Principal component analysis results

Figure 4 shows the proportion of variance explained by the first 10 PCs. Because the first 3 PCs explained roughly 70% of the total sample variance, they served as covariates in the models. PC1, PC2, and PC3 represented mainly the education level, demographic structure, and combined equipment modernity and economic development, respectively (Supplementary Table S2). PC1 was dominated by the percentages of the population that had completed senior high school (variable loading = 0.23) and obtained master's degrees (variable loading = - 0.24). PC2 was dominated by the aging index (variable loading = 0.37) and natural increase rate (variable loading = - 0.39). PC3 was dominated by the percentage of middle-low-income families (variable loading = 0.34) and the popularity of cable television (variable loading = - 0.32).

Relationships among mobility patterns, PCs, and confirmed case numbers

Pearson coefficients of correlation between the logarithm of the number of confirmed cases and standardized intra-district flows in periods B1 and A1 were close to 0.4, and the relationship between these variables was linear and positive (Fig. 5a). The coefficients of correlation between the logarithm of the number of confirmed

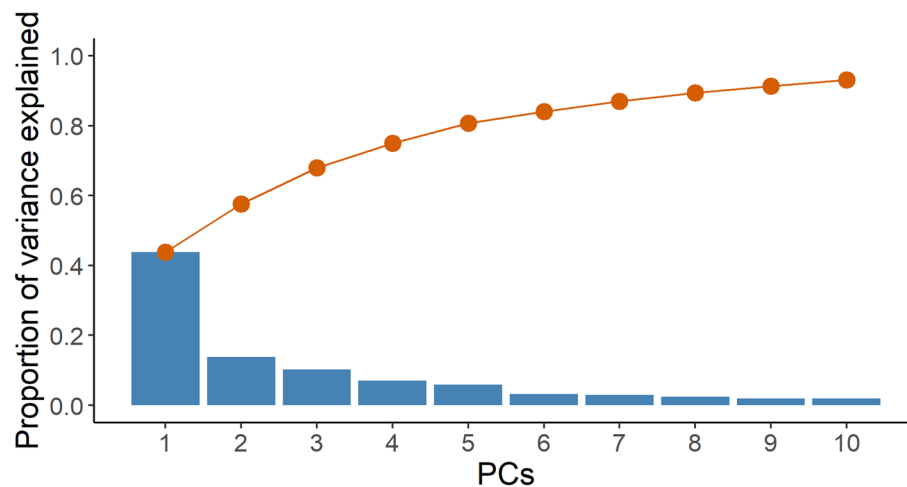


Figure 4. The proportion of variance explained by the first 10 PCs. The blue bars indicate the proportion of variance explained by PCs. The red lines represent the cumulative proportion of total variance.

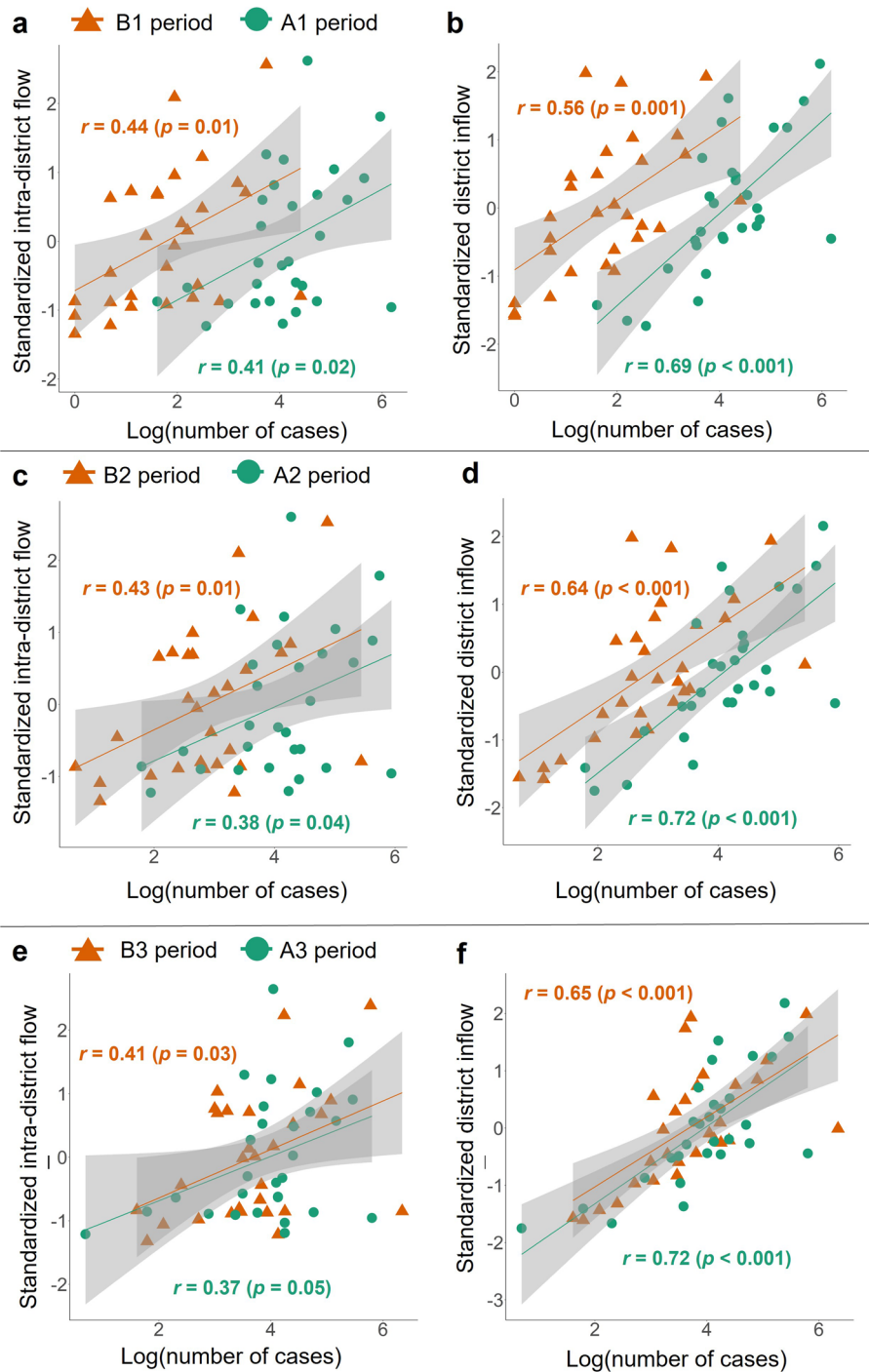


Figure 5. Pearson correlations among the number of confirmed COVID-19 cases, intra-district flow, and district inflow. The correlations among the natural logarithm of the number of confirmed COVID-19 cases per district, standardized intra-district flow (a, c, e), and standardized district inflow (b, d, f) before (B1, 2, 3 periods) and after (A1, 2, 3 periods) COVID-19 alert implementation. The regression lines were fitted with linear models. Gray shading represents 95% confidence intervals.

cases and standardized district inflows in periods B1 and A1 were 0.56 and 0.69, respectively, with a positive linear relationship (Fig. 5b). In addition, all relationships between the logarithm of the number of confirmed cases and standardized intra-district flows and district inflows in periods B2,3 and A2,3 were positive and linear (Fig. 5c,d,e,f).

The sum of intra-district flow and district inflow was associated positively with the number of confirmed cases, with coefficients 0.6, 0.59, and 0.52 in periods B1,2,3 and 0.44, 0.45, and 0.42 in periods A1,2,3 and

no significant difference between periods (Supplementary Table S3a, 4). At the district level, coefficients for intra-district flow were not significant in either period, whereas those for inflow were positive and significant (coefficients, 0.81, 0.81, 0.78 in periods B1,2,3; 0.59, 0.61, 0.59 in periods A1,2,3), with no significant difference between periods (Supplementary Table S3b,c, 5). The number of confirmed cases was associated more strongly with inflow (coefficients, 0.92, 0.92, 0.95 in periods B1,2,3; 0.65, 0.69, 0.66 in periods A1,2,3) than with intra-district flow (Supplementary Table S3d). No collinearity between intra-district flow and inflow was detected in any of the six periods (variance inflation factors < 5, $r = 0.45\text{--}0.50$; Supplementary Table S6).

In the model including the PCs, the sum of intra-district flows and district inflow was associated positively with the number of confirmed COVID-19 cases (coefficients: 0.84, 0.83, 0.81 in periods B1,2,3; 0.59, 0.68, 0.66 in periods A1,2,3; Table 1a, Supplementary Table S7a). The exponentiated coefficient, such as ($e^{0.84}$) in period B1, is the incidence rate ratio (IRR). In periods B1 and A1, the number of confirmed cases increased by a factor of 2.32 ($e^{0.84}$) and 1.80 ($e^{0.59}$) times, respectively, for every unit increase in mobility. The number of confirmed cases was related more strongly to PC2 and PC3 than to PC1.

In the model including the PCs, intra-district flow was associated positively with the number of confirmed COVID-19 cases pre- and post-alert, with coefficients 0.79, 0.78, 0.79 in periods B1,2,3 and 0.73, 0.73, 0.72 in periods A1,2,3 (Table 1b, Supplementary Table S7b). In periods B1 and A1, the number of confirmed cases increased by a factor of 2.20 ($e^{0.79}$) and 2.08 ($e^{0.73}$) times, respectively, for every unit increase in intra-district flow. The coefficients for PC2 and PC3 were positive and significant before and after alert implementation, whereas those for PC1 were negative and only significant post-alert.

Inflow was also associated significantly with the number of confirmed COVID-19 cases in the model including the PCs, with coefficients 0.89, 0.87, 0.84 in periods B1,2,3 and 0.49, 0.64, 0.60 in periods A1,2,3 (Table 1c, Supplementary Table S7c). The number of confirmed cases increased by a factor of 2.44 ($e^{0.89}$) times for every unit increase in inflow in period B1 and by a factor of 1.63 ($e^{0.49}$) times in period A1. The coefficients for PC2 were positive and significant before and after alert implementation, except in the A1 period, and those for PC1 were significant only before alert implementation.

In the model including the PCs, the number of confirmed COVID-19 cases was related positively to intra-district flow and district inflow, except in the B1 period (Table 1d, Supplementary Table S7d). Pre- and post-alert, this relationship was stronger for district inflow (coefficients, 0.52, 0.48, 0.48 in periods B1,2,3 and 0.41, 0.46, 0.43 in periods A1,2,3) than for intra-district flow (coefficients, 0.43, 0.45, 0.45 in periods B1,2,3 and 0.36, 0.33, 0.35 in periods A1,2,3). PC2 and PC3 were related more strongly than PC1 to the number of confirmed cases.

		B1 period		A1 period	
		Coefficient (std. error)	<i>p</i>	Coefficient (std. error)	<i>p</i>
a	Intercept	1.80 (0.15)	<0.001***	4.32 (0.12)	<0.001***
	Intraflow + Interflow	0.84 (0.18)	<0.001***	0.59 (0.14)	<0.001***
	PC1	0.07 (0.04)	0.12	- 0.009 (0.03)	0.77
	PC2	0.24 (0.07)	0.001**	0.05 (0.06)	0.36
	PC3	0.20 (0.08)	0.01*	0.13 (0.07)	0.06
	Pseudo r^2	0.61		0.55	
b	Intercept	1.84 (0.16)	<0.001***	4.32 (0.12)	<0.001***
	Intraflow	0.79 (0.19)	<0.001***	0.73 (0.15)	<0.001***
	PC1	- 0.03 (0.04)	0.48	- 0.07 (0.03)	0.03*
	PC2	0.26 (0.08)	0.001**	0.17 (0.05)	0.002**
	PC3	0.30 (0.10)	0.002**	0.23 (0.08)	0.002**
	Pseudo r^2	0.57		0.57	
c	Intercept	1.82 (0.16)	<0.001***	4.41 (0.14)	<0.001***
	Interflow	0.89 (0.21)	<0.001***	0.49 (0.17)	0.003**
	PC1	0.11 (0.05)	0.03*	0.02 (0.04)	0.61
	PC2	0.23 (0.08)	0.003**	0.01 (0.06)	0.85
	PC3	0.12 (0.08)	0.14	0.03 (0.07)	0.67
	Pseudo r^2	0.57		0.39	
d	Intercept	1.80 (0.15)	<0.001	4.28 (0.11)	<0.001***
	Intraflow	0.43 (0.24)	0.07	0.36 (0.17)	0.04*
	Interflow	0.52 (0.27)	0.05	0.41 (0.17)	0.01*
	PC1	0.05 (0.06)	0.37	- 0.01 (0.04)	0.70
	PC2	0.24 (0.07)	0.001**	0.11 (0.05)	0.03*
	PC3	0.23 (0.10)	0.02*	0.16 (0.07)	0.03*
	Pseudo r^2	0.61		0.64	

Table 1. Regression associations among mobility patterns, PCs, and the number of confirmed COVID-19 cases. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Relationships among mobility patterns in different age groups, PCs, and confirmed case numbers

The intra-district flows of people aged 15–59 years were related significantly to the number of confirmed COVID-19 cases before and after alert implementation (coefficients, 0.80, 0.79, 0.80 in periods B1,2,3; 0.75, 0.76, 0.76 in periods A1,2,3; Supplementary Table S8a). In addition, there was also a significant relationship between the intra-district flows of people aged ≥ 60 years and the number of confirmed COVID-19 cases before and after alert implementation (coefficients, 0.76, 0.77, 0.78 in periods B1,2,3; 0.67, 0.64, 0.61 in periods A1,2,3; Supplementary Table S8b). However, as shown in Fig. 6a based on Supplementary Table S8c, no such relationship was observed when both age groups were considered simultaneously. This is probably due to the issue of multicollinearity between the intra-district flows of people aged 15–59 and ≥ 60 years (Supplementary Table S10).

In the model including the PCs, the inflows of people aged 15–59 years were related significantly to the number of confirmed cases (coefficients: 0.82, 0.81, 0.78 in periods B1,2,3; 0.48, 0.60, 0.57 in periods A1,2,3; Supplementary Table S9a). Furthermore, the inflows of people aged ≥ 60 years were also related significantly to the number of confirmed cases (coefficients: 1.05, 1.01, 0.99 in periods B1,2,3; 0.66, 0.74, 0.70 in periods A1,2,3; Supplementary Table S9b).

However, as shown in Fig. 6b based on Supplementary Table S9c, when the inflows of people aged 15–59 and ≥ 60 years were considered simultaneously, the inflows of people aged 15–59 and ≥ 60 years showed no significant relationship, even though the inflows of people aged ≥ 60 years were weakly significant (p -values ranging from 0.01 to 0.05) in periods B3 and A2,3. This is probably because the results suffer from an issue on the multicollinearity, there are high correlations between the inflows of people aged 15–59 and ≥ 60 years (Supplementary Table S11).

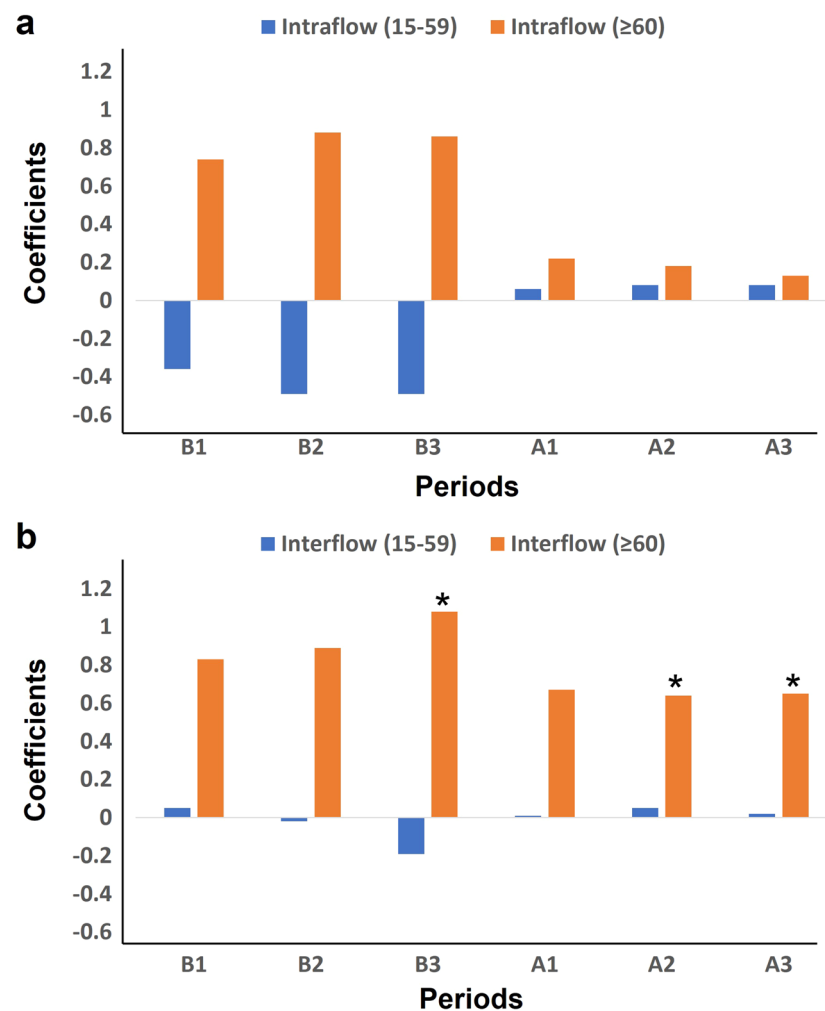


Figure 6. Regression coefficients between the number of confirmed COVID-19 cases and mobility patterns in different age groups. The regression coefficients between the intra- and inter-district flows (a, b) of people aged 15–59 and ≥ 60 years and the number of confirmed COVID-19 cases. * $p < 0.05$.

Discussion

The purpose of this study was to examine the relationship between mobility patterns in different age groups and COVID-19 transmission in the Taipei metropolitan area before and after the implementation of control measures. It showed that intra-district flow and district inflow were related positively to the number of confirmed COVID-19 cases before and after alert implementation, with the latter relationship stronger than the former and attenuating markedly post-alert. Moreover, both the inter-district flows of people aged 15–59 and ≥ 60 years were significantly related to the number of confirmed cases. Lastly, PC2 and PC3 were related positively to the number of confirmed COVID-19 cases before and after alert implementation.

The finding of the relationship between district inflow and COVID-19 transmission in this study may reflect the mobility behaviors in the Taipei metropolitan area. There is a tendency for people to congregate in districts with high population densities and a high percentage of cross-district mobility in the Taipei metropolitan area²⁸. On the other hand, the intra-district flow commonly represents the daily activities in neighborhoods such as grocery shopping, and it is less affected by travel restrictions. This finding is consistent with those of Jia et al. (2020) and Wei and Wang (2020), who observed that movement out of the COVID-19 outbreak center of Wuhan was related significantly to the number of confirmed cases in other Chinese cities^{2,13}. Inter-district flow is a greater promoter of COVID-19 transmission from outbreak locales to other locations than is intra-district flow.

In addition, the relationship between district inflow and the number of confirmed COVID-19 cases before and after alert implementation may be affected by several possible reasons such as NPIs and public awareness due to districts with more COVID-19 cases. The implementation of NPIs such as school closures and remote working was expected to reduce COVID-19 transmission due to a reduction in human mobility. Furthermore, public awareness and NPIs use likely reduced the risk of infection during movement⁴². This suggests that the variability in the number of confirmed COVID-19 cases among locations after control measure implementation is better explained by factors such as the public health response than by human mobility³.

The finding that older adults' mobility was significantly related to COVID-19 transmission may be explained by the following reasons. First, because older adults face more health risks related to COVID-19 infection, they likely contribute to large numbers of confirmed cases³⁴. Second, older adults may reduce their mobility less than do younger generations when outbreaks occur. In the absence of control measures, most members of younger generations leave their homes for work or school daily, whereas most older adults are retired and less mobile²². Regions with larger proportions of the population ranging from 24 to 59 years received the largest effects in mobility caused by the lockdown¹⁹. Moreover, members of younger generations use social media, where outbreak dangers are emphasized, more than do older adults²². Thus, younger generations' mobility likely decreases obviously, whereas older adults' mobility is not affected apparently when outbreaks occur. This inference is consistent with Liu et al.'s (2021) reporting that the presence of larger percentages of people aged > 60 years in Chinese cities was related to less mobility reduction after control measure implementation¹⁰.

The results of this study indicate that the district-level age structure (PC2), household income, and popularity of modern equipment (PC3) were related significantly to the number of confirmed COVID-19 cases before and after alert implementation. The PC2 results indicate that the district-level presence of larger percentages of older adults and fewer members of younger generations was related to a larger number of confirmed COVID-19 cases. This result is consistent with the finding of Mogi and Spijker (2021)³⁴, which inferred that the high rate of COVID-19 diagnosis in Italy was likely due to the country's aged population. The PC3 results indicate that the district-level presence of larger percentages of middle-low-income families and less cable television popularity was related to a larger number of confirmed COVID-19 cases. This result may reflect in part the emphasis on outbreak danger and awareness in the media to reduce exposure and comply with movement restrictions^{19,22,43}. On the other hand, socioeconomically disadvantaged segments of the population may be more likely to be exposed to infection due to their social contexts, which could influence infectious disease occurrence⁴⁴. This is consistent with the finding of Waku et al. (2022)⁴⁵, which concluded that death was negatively correlated with the signs of wealth. For instance, low-income status may be related to crowded living conditions, which have been related in turn to a greater risk of infection. People who have lower-income jobs may still need to go to workplaces, whereas people who have higher-income jobs could often work remotely¹⁹. Furthermore, population segments with lower income levels may not be able to obtain timely high-quality care and nutrition to combat the pandemic^{25,46}.

This study has several limitations. First, it was conducted with data from the Taipei metropolitan area, and the results may not be generalizable to other locations, especially rural areas, which have different demographic and regional characteristics. For instance, due to the small amount of district inflow to rural areas, the number of confirmed COVID-19 cases may not be proportional to the amount of human mobility. Thus, the relationships between human mobility and the number of confirmed COVID-19 cases in rural areas need to be clarified. Second, onset dates, rather than reporting dates, were used to determine the daily numbers of new confirmed COVID-19 cases in this study. This approach prevents the problem of delayed reporting, which may influence model results, but the onset dates were self-reported and thus may not be entirely accurate. Third, the current study only investigated the difference in the relationships between population flows and COVID-19 transmission in age groups before and after control measures implementation. However, the changes in the relationships between population flows and COVID-19 transmission were influenced by several possible reasons⁴⁷. The possible reasons influencing the difference in the relationships between population flows and COVID-19 transmission before and after control measures implementation still need to be examined clearly. Fourth, because of the restriction of data availability, the inter-district flows analyzed in this study do not include inflows from districts outside the study area. However, not only inflows from other districts within the study area but also inflows from those outside the study area are better to be included.

Based on the results of this study, we offer four policy suggestions. First, travel (and especially inter-regional travel) restrictions should be implemented as soon as possible in the initial stages of outbreaks of emerging

infectious diseases. Second, the rapid implementation of NPIs, such as face mask use and social distancing, contributed to the reduction of COVID-19 transmission caused by human mobility and is recommended in future outbreak situations. Third, the government should strengthen public awareness of epidemic prevention for older adults and provide an economic relief package to help those who have lost their jobs due to a pandemic, especially middle-low-income families. Fourth, the government should consider regional economic, societal, and cultural heterogeneity when formulating local public health responses to a pandemic.

Future research could further explore regional changes in intra-regional movement during COVID-19 outbreaks. In this study, the relationship between intra-district flow and the number of confirmed COVID-19 cases persisted after the implementation of control measures, unlike that of inflow movement. This may be due to people's maintenance of daily activities and work-related movement in their neighborhoods after alert implementation¹⁹. These results suggest that the existing epidemic prevention policies less control the intra-regional daily movement during a pandemic. A more complete understanding of local mobility behaviors would aid the formulation of public health responses to community transmission.

Conclusion

To conclude, this study showed that intra- and inter-regional mobility in age-specific groups had different relationships with COVID-19 transmission before and after the implementation of control measures. The inter-regional mobility was related more strongly to COVID-19 transmission than intra-regional mobility. Furthermore, both the inter-regional mobility of people aged 15–59 and ≥60 years were significantly related to COVID-19 transmission. These findings can aid governments' formulation of more specific epidemic prevention policies to reduce losses during the COVID-19 pandemic.

Data availability

R code is available at the GitHub repository: <https://github.com/JWMHub/Mobility>. The datasets of the numbers of confirmed COVID-19 cases and socioeconomic factors analyzed during this study are available from the government's open data platforms, <https://data.gov.tw/>. The data on anonymized and aggregated human mobility retrieved from mobile phone data are available from Far Eastone Telecommunications but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

Received: 9 November 2022; Accepted: 9 October 2023

Published online: 12 October 2023

References

1. Johns Hopkins Coronavirus Resource Center. COVID-19 Map. *Johns Hopkins Coronavirus Resource Center* <https://coronavirus.jhu.edu/map.html> (2022).
2. Jia, J. S. *et al.* Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* **582**, 389–394. <https://doi.org/10.1038/s41586-020-2284-y> (2020).
3. Kraemer, M. U. G. *et al.* The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497. <https://doi.org/10.1126/science.abb4218> (2020).
4. Hu, T. *et al.* Human mobility data in the COVID-19 pandemic: Characteristics, applications, and challenges. *Int. J. Digit. Earth* **14**, 1126–1147. <https://doi.org/10.1080/17538947.2021.1952324> (2021).
5. Zhang, M. *et al.* Human mobility and COVID-19 transmission: A systematic review and future directions. *Ann. GIS* **0**, 1–14. <https://doi.org/10.1080/19475683.2022.2041725> (2022).
6. Chan, C.-H. & Wen, T.-H. Revisiting the effects of high-speed railway transfers in the early COVID-19 cross-province transmission in Mainland China. *Int. J. Environ. Res. Public Health* **18**, 6394. <https://doi.org/10.3390/ijerph18126394> (2021).
7. Li, T., Rong, L. & Zhang, A. Assessing regional risk of COVID-19 infection from Wuhan via high-speed rail. *Transp. Policy* **106**, 226–238. <https://doi.org/10.1016/j.tranpol.2021.04.009> (2021).
8. Oztig, L. I. & Askin, O. E. Human mobility and coronavirus disease 2019 (COVID-19): A negative binomial regression analysis. *Public Health* **185**, 364–367. <https://doi.org/10.1016/j.puhe.2020.07.002> (2020).
9. Chang, M.-C. *et al.* Variation in human mobility and its impact on the risk of future COVID-19 outbreaks in Taiwan. *BMC Public Health* **21**, 226. <https://doi.org/10.1186/s12889-021-10260-7> (2021).
10. Liu, Y. *et al.* Associations between changes in population mobility in response to the COVID-19 pandemic and socioeconomic factors at the city level in China and country level worldwide: A retrospective, observational study. *Lancet Digit. Health* **3**, e349–e359. [https://doi.org/10.1016/S2589-7500\(21\)00059-5](https://doi.org/10.1016/S2589-7500(21)00059-5) (2021).
11. Paez, A. Using Google Community Mobility Reports to investigate the incidence of COVID-19 in the United States. *Findings* <https://doi.org/10.32866/001c.12976> (2020).
12. Sulyok, M. & Walker, M. Community movement and COVID-19: A global study using Google's Community Mobility Reports. *Epidemiol. Infect.* **148**, e284. <https://doi.org/10.1017/S0950268820002757> (2020).
13. Wei, S. & Wang, L. Examining the population flow network in China and its implications for epidemic control based on Baidu migration data. *Humanit. Soc. Sci. Commun.* **7**, 1–10. <https://doi.org/10.1057/s41599-020-00633-5> (2020).
14. Grantz, K. H. *et al.* The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nat. Commun.* **11**, 4961. <https://doi.org/10.1038/s41467-020-18190-5> (2020).
15. Badr, H. S. *et al.* Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. *Lancet Infect. Dis.* **20**, 1247–1254. [https://doi.org/10.1016/S1473-3099\(20\)30553-3](https://doi.org/10.1016/S1473-3099(20)30553-3) (2020).
16. Luo, M. *et al.* Population Mobility and the Transmission Risk of the COVID-19 in Wuhan, China. *ISPRS Int. J. Geo-Inf.* **10**, 395. <https://doi.org/10.3390/ijgi10060395> (2021).
17. Peixoto, P. S., Marcondes, D., Peixoto, C. & Oliva, S. M. Modeling future spread of infections via mobile geolocation data and population dynamics. An application to COVID-19 in Brazil. *PLOS ONE* **15**, e0235732. <https://doi.org/10.1371/journal.pone.0235732> (2020).
18. Balcan, D. *et al.* Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl. Acad. Sci.* **106**, 21484–21489. <https://doi.org/10.1073/pnas.0906910106> (2009).
19. Pullano, G., Valdano, E., Scarpa, N., Rubrichi, S. & Colizza, V. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: A population-based study. *Lancet Digit. Health* **2**, e638–e649. [https://doi.org/10.1016/S2589-7500\(20\)30243-0](https://doi.org/10.1016/S2589-7500(20)30243-0) (2020).

20. Nagata, S. *et al.* Mobility change and COVID-19 in Japan: Mobile data analysis of locations of infection. *J. Epidemiol.* **31**, 387–391. <https://doi.org/10.2188/jea.JE20200625> (2021).
21. Becher, H., Bönisch, S. & Wegscheider, K. Reduction of mobility during the COVID-19 pandemic in Germany according to age, sex, and federal state. *Dtsch. Arztebl. Int.* **118**, 536–537. <https://doi.org/10.3238/arztebl.m2021.0293> (2021).
22. Caselli, F., Grigoli, F., Sandri, D. & Spilimbergo, A. Mobility under the COVID-19 pandemic: Asymmetric effects across gender and age. *IMF Econ. Rev.* **70**, 105–138. <https://doi.org/10.1057/s41308-021-00149-1> (2022).
23. Nomura, S. *et al.* Mobility patterns in different age groups in Japan during the COVID-19 Pandemic: A small area time series analysis through march 2021. *J. Urban Health* **98**, 635–641. <https://doi.org/10.1007/s11524-021-00566-7> (2021).
24. Oshinubi, K., Rachdi, M. & Demongeot, J. Analysis of reproduction number R0 of COVID-19 using current health expenditure as gross domestic product percentage (CHE/GDP) across countries. *Healthcare* **9**, 1247. <https://doi.org/10.3390/healthcare9101247> (2021).
25. Abedi, V. *et al.* Racial, economic, and health inequality and COVID-19 infection in the United States. *J. Racial Ethnic Health Disparities* **8**, 732–742. <https://doi.org/10.1007/s40615-020-00833-4> (2021).
26. Taiwan Centers for Disease Control. <https://www.cdc.gov.tw/> (2021).
27. Socio-Economic Data Service Platform. https://segis.moi.gov.tw/STAT/Web/Portal/STAT_PortalHome.aspx (2021).
28. Ministry of Transportation and Communications R.O.C. Survey on the daily use of public transportation in Taiwan. <https://www.motc.gov.tw/ch/app/statistics101?lang=ch&folderName=ch&id=56> (2020).
29. Epidemic Intelligence Center. Area, age, and gender statistical table of daily COVID-19 cases. *Taiwan CDC Open Data Portal* <https://data.cdc.gov.tw/dataset/aagstable-day-19cov> (2021).
30. National Communications Commission. Statistics of mobile broadband service subscribers. *Statistics of Mobile Broadband Service Subscribers* https://www.ncc.gov.tw/chinese/chinese/news_detail.aspx?site_content_sn=5018&cate=0&keyword=&is_history=0&pages=0&sn_f=47658 (2022).
31. Dept. of Household Registration, M. of the Interior. R. of C. Population statistics. <https://www.ris.gov.tw/app/portal> (2021).
32. Government Opendata Platform. <https://data.gov.tw/> (2021).
33. New Taipei Statistical Database. *New Taipei Statistical Database* https://oas.bas.ntpc.gov.tw/BAS/pxweb/dialog/statfile9_n.asp (2021).
34. Mogi, R. & Spijker, J. The influence of social and economic ties to the spread of COVID-19 in Europe. *J. Popul. Res.* **39**, 495–511. <https://doi.org/10.1007/s12546-021-09257-1> (2021).
35. Johnson, R. A. & Wichern, D. W. *Applied Multivariate Statistical Analysis* (Pearson Prentice Hall, 2007).
36. Ibarra-Espinosa, S., de Freitas, E. D., Ropkins, K., Dominici, F. & Rehbein, A. Negative-Binomial and quasi-poisson regressions between COVID-19, mobility and environment in São Paulo, Brazil. *Environ. Res.* **204**, 112369. <https://doi.org/10.1016/j.envres.2021.112369> (2022).
37. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S* (Springer, 2002).
38. R Core Team. R: A Language and Environment for Statistical Computing. (2020).
39. QGIS Development Team. QGIS Geographic Information System. (2022).
40. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2016).
41. Microsoft Corporation. Microsoft Excel. (2021).
42. Nouvellet, P. *et al.* Reduction in mobility and COVID-19 transmission. *Nat. Commun.* **12**, 1090. <https://doi.org/10.1038/s41467-021-21358-2> (2021).
43. Funk, S., Salathé, M. & Jansen, V. A. A. Modelling the influence of human behaviour on the spread of infectious diseases: A review. *J. R. Soc. Interface* **7**, 1247–1256. <https://doi.org/10.1098/rsif.2010.0142> (2010).
44. Khalatbari-Soltani, S., Cumming, R. C., Delpierre, C. & Kelly-Irving, M. Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *J. Epidemiol. Community Health* **74**, 620–623. <https://doi.org/10.1136/jech-2020-214297> (2020).
45. Waku, J., Oshinubi, K. & Demongeot, J. Maximal reproduction number estimation and identification of transmission rate from the first inflection point of new infectious cases waves: COVID-19 outbreak example. *Math. Comput. Simul.* **198**, 47–64. <https://doi.org/10.1016/j.matcom.2022.02.023> (2022).
46. Oshinubi, K., Rachdi, M. & Demongeot, J. Modeling of COVID-19 pandemic vis-à-vis some socio-economic factors. *Front. Appl. Math. Stat.* **7**, 1–25. <https://doi.org/10.3389/fams.2021.786983> (2022).
47. Glaeser, E. L., Gorbach, C. & Redding, S. J. JUE Insight: How much does COVID-19 increase with mobility? Evidence from New York and four other U.S. cities. *J. Urban Econ.* **127**, 103292. <https://doi.org/10.1016/j.jue.2020.103292> (2022).

Acknowledgements

The authors thank Dr. Kung-Yee Liang and Dr. Shu-Chen Kuo of the National Health Research Institutes, Taiwan, for their insightful and constructive comments. The authors thank Ms. Fang-Jing Lee and Ms. Li-Yi Chen of the Institute of Population Health Sciences, the National Health Research Institutes, Taiwan, for their help with data management. Their expertise greatly assisted in the success of this research. The authors thank Far EasTone Telecommunications for collecting and aggregating the de-identification mobile phone signaling data used in this study.

Author contributions

All authors contribute significantly to the design and conduct of this study and have approved the submission of this paper for review by your Journal in its present form. The authors' contribution is listed below: W.M.J., T.H.W., Y.C.H., H.Y.C., C.A.H., and H.H.T. contributed significantly to the study's conception and design. W.M.J. contributed to the drafting of the manuscript. W.M.J. and H.H.T. contributed to data preparation and validation. W.M.J., T.H.W., and H.H.T. provided statistical analysis. T.H.W., Y.C.H., H.Y.C., W.J.C., C.A.H., H.K.S., and H.H.T. contributed to the interpretation of the data and/or provided revisions of the manuscript. H.Y.C., W.J.C., C.A.H., H.K.S., and H.H.T. provided administrative, technical, and material support. The research was supervised by H.H.T. H.H.T. is the corresponding author.

Funding

This study was supported by grants PH-111-PP-02, PH-111-GP-02, and PH-111-GP-12 from the National Health Research Institutes, a nonprofit foundation dedicated to medical research and improved healthcare in Taiwan.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44474-z>.

Correspondence and requests for materials should be addressed to H.-H.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023