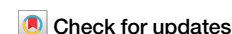Article

Evaluating the association between COVID-19 transmission and mobility in omicron outbreaks in China

Check for updates

Liping Peng[1], Kylie E. C. Ainslie [1,2], Xiaotong Huang[1], Benjamin J. Cowling [1,3], Peng Wu [1,3] & Tim K. Tsang [1,3] ✉

Abstract

Background Prior research has suggested a positive correlation between human mobility and COVID-19 transmission at national or provincial levels, assuming constant correlations during outbreaks. However, the correlation strength at finer scales and potential changes in relationships during outbreaks have been scarcely investigated.
Methods We gathered case and mobility data (within-city movement, inter-city inflow, and inter-city outflow) at the city level from Omicron outbreaks in mainland China between February and November 2022. For each outbreak, we calculated the time-varying effective reproduction number (R_t). Subsequently, we estimated the cross-correlation and rolling correlation between R_t and the mobility index, comparing them and identifying potential factors affecting these correlations.
Results We identify 57 outbreaks during Omicron wave 1 (February to June) and 171 outbreaks during Omicron wave 2 (July to December). Cross-correlation estimates vary between waves, with values ranging from 0.64 to 0.71 in wave 1 and 0.45 to 0.46 in wave 2. Oscillation models best fit the rolling correlation for almost all outbreaks, and there are significant differences between extreme values of rolling correlation and cross-correlation. Additionally, we estimate a positive relationship between the GRI and rolling correlation during the pre-peak stage, turning negative during the post-peak stage.
Conclusions Our findings suggest a positive relationship between Omicron transmission and mobility at the city level. However, significant fluctuations in their relationship, as demonstrated by rolling correlation, indicate that assuming a constant correlation between transmission and mobility may lead to inaccurate predictions or decisions when using mobility as a proxy for transmission intensity.

Plain language summary

It is a general belief that a higher rate of population mobility leads to an increase in transmission of COVID-19, but the underlying relationship can be more complex. To understand this relationship better, we analyzed data from various cities in mainland China that experienced Omicron outbreaks in 2022. Our findings showed a generally positive correlation between population mobility and transmission. However, this relationship can change over time and vary between different outbreaks. The level of government response has an impact on this relationship. Our findings can guide the potential of utilizing mobility as a predictor for nowcasting and forecasting epidemics and emphasize that such usage requires careful consideration of different factors.

Monitoring real-time transmissibility of infectious pathogens is critical for guiding control policies. Transmissibility is typically monitored by estimating the time-varying effective reproduction number (R_t)[1,2], a measure of the average number of secondary cases caused by each infected individual. Given that COVID-19 transmission occurs through close contact, population mobility is expected to strongly correlated with real-time transmissibility[3–5]. To capture this correlation, various types of mobility data have been utilized to generate different mobility index[6]. These include

data from cell phones and other electronic devices with location-tracking capabilities[7,8], GPS data from smartphones and mobile apps[9], as well as bus traffic and air traffic data[10–12].

Most studies examining the relationship between SARS-CoV-2 transmission and mobility have been conducted at the country or regional level[6] due to the limited availability of frequently used datasets, such as Google mobility[13] and Apple mobility[14]. However, it has been observed that the correlation between COVID-19 transmission (measured by the number

[1]WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China. [2]Centre for Infectious Disease Control, National Institute for Public Health and Environment (RIVM), Bilthoven, the Netherlands. [3]Laboratory of Data Discovery for Health Limited, Hong Kong Science and Technology Park, New Territories, Hong Kong Special Administrative Region, China. ✉e-mail: timtsang@connect.hku.hk

of COVID-19 cases) and community mobility was specific to individual cities[15].

Moreover, the correlation patterns between R_t and mobility have shown inconsistency across different studies[7]. One previous study analyzed the relationships between R_t and residential mobility in 125 countries, revealing a mix of positive (39), negative (58), and inconclusive (28) correlations among countries[16]. Several studies suggested that the relationship between R_t and mobility may be non-linear[16] or vary across different stages of the epidemic[17,18]. Additionally, the correlation between R_t and the reduction in population mobility might vary depending on the specific variants of the SARS-CoV-2 virus that are circulating[19]. On the other hand, the relationship between mobility and SARS-CoV-2 transmissibility was likely impacted by control measures, as some control measures were aimed at reducing mobility and hence transmissions in the population.

In summary, previous studies have highlighted dynamic, context-dependent relationships between SARS-CoV-2 transmission and population mobility, yet most epidemiological analyses either assumed static relationships or examined temporal variations only at provincial scales[20] or focused exclusively on one selected city[21]. Comprehensive studies examining the relationship between transmission and mobility at the city level in a country remain limited.

Here, we analyze the relationship between mobility and COVID-19 transmissibility at the city level using data from the Omicron outbreaks in mainland China in 2022. We estimate the time-varying effective reproduction number (R_t) for each outbreak and quantify mobility patterns using the Baidu mobility index. We then examine cross-correlation and rolling correlation between R_t and the mobility index. We also compare these correlation patterns and explore potential factors influencing the relationship between R_t and the mobility index. In this study, our analysis reveals that the association between human mobility and COVID-19 transmissibility is dynamic and context-dependent at the city level. We find that increased mobility is generally linked to higher transmission, yet the strength and direction of this relationship vary over time and with different outbreak phases. Moreover, our results indicate that government interventions modulate these associations, underscoring that static assumptions about the mobility–transmission relationship do not adequately capture the temporal fluctuations observed in outbreak dynamics. These findings highlight the importance of adopting dynamic, context-sensitive approaches for epidemic forecasting and public health decision-making.

Methods
Case data
Case data from 1 January 2022 to 27 November 2022 were obtained from the daily notification of COVID-19 on the National Health Commission of the People's Republic of China website[22]. Cases were reported based on the date of detection. It was anticipated that the report delay was minimal due to the stringent city-wide measures implemented in China during the outbreaks[23]. In China, cases were classified as either local (domestic) cases or imported cases. Local cases, including both symptomatic and asymptomatic cases, were used in our study. Cities were coded according to Baidu cityCode[24]. Case data for cities in Yunnan and Xinjiang provinces were missing and therefore not included in our analysis. In total, data from 336 cities were included in our study.

Mobility index data
The daily mobility index data used in this study were sourced from Baidu mobility Big Data[25,26], which is based on the widely used mapping service Baidu Maps in China, similar to Google Maps. Baidu mobility data is collected based on Baidu's location-based service technology, offering insights into city-specific and temporal migration patterns. Using the location-aware devices, Baidu mobility data captures the spatial-temporal trajectories of daily population movements within communities.

While it may not capture all migrations, it remains valuable for analyzing population flow patterns across different cities and times. To access the Baidu mobility data, hypertext markup language (HTML) requests were

sent to the Baidu mobility platform (http://qianxi.baidu.com/). To protect user privacy, the platform records daily travel flows for cities and aggregates this information into an index for cross-city comparisons.

For our study, Baidu mobility data were collected for 366 prefecture-level cities, including three mobility indices: within-city movement, inter-city inflow, and inter-city outflow. The inter-city inflow and outflow index reflects the magnitude of population migration between cities. The within-city movement is calculated as an index based on the ratio of daily intra-city trips to the resident population[27]. The mobility index can be compared across cities. To avoid weekly fluctuations induced by the work-leisure shift, the daily mobility index was smoothed using a moving average over a 7-day window.

Government response index (GRI)
The daily government response index (GRI) was obtained from the publicly available Oxford COVID-19 Government Response Tracker (OxCGRT)[28]. The OxCGRT is a comprehensive dataset that captures the diverse government policies implemented in response to the global COVID-19 pandemic, spanning across more than 180 countries. Within the OxCGRT, the GRI stands out as a reliable and thorough index, effectively portraying the wide range of policy modifications enacted by governments. The GRI comprised 13 indicators[29], including containment and closure indicators, economic response indicators, and health systems indicators (Supplementary Method and Supplementary Table 1). The GRI was constructed at the provincial level.

To construct the GRI at the city level, we first screened the textual notes about implementation and cessation of various public health measures at the city level, to ensure that those cities were included in the consideration of GRI. Then, on the days when the daily case count for a city accounted for 80% or more of the total cases in the corresponding province, the GRIs at the province level were considered as the GRI for that city. We conducted a sensitivity analysis with different thresholds, including 70, 90, and 100%.

Definition of outbreaks
An outbreak was defined as 20 or more cases occurring in a single day[30]. The start date of an outbreak was defined as the date on which the first case (symptomatic or asymptomatic) occurred, going backward from the date on which there were more than 20 cases in a single day. The end of an outbreak was defined as the day with no new cases for 7 consecutive days after the peak of the outbreak. Outbreaks with a duration longer than 14 days were included in the study.

Estimation of time-varying effective reproduction number (R_t)
Since there was pre-symptomatic transmission for SARS-CoV-2[31], reconstructing the epidemic curve by the date of infection could provide a more accurate estimation of R_t. As the case data was recorded based on the report date, we first reconstructed the epidemic curve by infection date based on the epidemic curve by the report date using a deconvolution approach[32], with the distribution of the delay from infection to report (Supplementary Method). Then, we estimated the R_t based on the Poisson framework in Cori et al.[1].

Relationship between transmission and mobility
We computed the cross-correlation between R_t and mobility indices for each identified outbreak using the Pearson correlation and selecting the optimal lag day based on the highest correlation. We combined correlation coefficients from each city and weighted the standard error of estimates to generate a weighted average (Supplementary Method).

Furthermore, we adopted rolling correlation between R_t and three mobility indices to measure and visualize short-term but potentially time-varying correlations[20,33]. A detailed comparison of cross-correlation and rolling correlation is presented in Supplementary Table 2. We performed biweekly rolling correlation analysis, where the correlation on day t was estimated based on the two time series data from day $t-13$ to t, covering a period of 14 days. We only included city outbreaks with a duration of at least

42 days to ensure a sufficient amount of data for estimation. We also conducted triweekly rolling correlation analyses to further explore the sensitivity of our results (Supplementary Method).

To determine if using rolling correlation was necessary, i.e., the magnitude is changing during outbreaks and not constant, we employed the non-linear least squares method to fit the rolling correlation for each outbreak. We fitted five models (constant, linear, quadratic, sine, and cosine) and chose the optimal one based on the smallest Akaike information criterion value. The inclusion of the constant model was based on its frequent adoption in prior analyses[34,35], reflecting situations where the model assumes a consistent relationship between R_t and mobility. We also compared the cross-correlation coefficient with the minimum and maximum values of the rolling correlation coefficient.

Factors affecting cross-correlation and rolling correlation

We aimed to investigate what factors may impact the cross-correlation and rolling correlation. Regarding cross-correlation, we conducted Pearson correlation tests on several potential factors, including outbreak duration, peak value of R_t, and GRI (Supplementary Method).

Regarding rolling correlation, we investigated whether rolling correlations were different by stages of outbreaks and level of GRI (Supplementary Method). Outbreaks were divided into two stages, namely the pre-peak and post-peak stages, based on the peak value of R_t. We employed a mixed-effect regression to assess the impact of different stages and GRI on the rolling correlation between R_t and the mobility index. Analysis stratified by outbreak stage was also conducted. The mixed-effect regression model included a random intercept term to account for variations among different outbreak cities. The rolling correlation for each outbreak was used as the outcome variable. The city-level GRI and different stages served as the predictor variables, respectively. We applied a Fisher transformation to the rolling correlation before fitting the models. In addition, we employed k-means clustering to identify patterns in rolling correlations across outbreaks and subsequently evaluated the associations between the resulting clusters and key urban characteristics (Supplementary Method and Supplementary Figs. 11–13).

Statistics analyses

Processing Python 3.8.6 (Python Software Foundation) and related libraries were utilized to capture the required data. All analyses were performed using R software version 3.6.3 (R Foundation for Statistical Computing, Austria). Uncertainty in effective reproduction number (R_t) was quantified as mean ± standard deviation. For correlation analyses, 95% confidence intervals were generated through 200 iterations of nonparametric bootstrap resampling. The significance level for all tests was set at $p < 0.05$.

Informed consent was not required for this study since the data used was obtained from publicly available data sources.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Results

Identifications of outbreaks

In 2022, mainland China experienced two major waves caused by the Omicron variant. Wave 1 (February to June) was caused by the Omicron BA.2 variant and resulted in a total of 766, 000 cases. We identified 57 outbreaks in 57 cities across 23 provinces during this period. Wave 2 (July to December) was caused by the more transmissible Omicron BA.5 variant. However, it is worth noting that the surveillance was gradually discontinued in different regions between November 27 and December 7, 2022[36], due to the cessation of the zero-COVID policy in mainland China[37]. As of November 27, there were a total of 554,000 cases, and we identified 171 outbreaks in 171 cities across 29 provinces during wave 2 (Supplementary Fig. 1).

For each outbreak in each city, we first estimated time-varying effective reproduction number (R_t). Then, we computed cross-correlation and

rolling-correlation, and tested models with different assumptions on variation of correlation during outbreaks. We used the outbreak in Shanghai during wave 1 as an example (Fig. 1). Details of each outbreak were available in Supplementary Data 1,2.

Cross-correlation

We found that a lag of zero had the highest correlation between R_t and mobility indices for all outbreaks in waves 1 and 2 (Supplementary Fig. 2). Overall, we observed positive cross-correlations between R_t and three mobility indices (Fig. 2), with weighted averages ranging from 0.64 to 0.71 in wave 1 and 0.45–0.46 in wave 2. However, 12–23% of wave 1 outbreaks and 22–26% of wave 2 outbreaks showed uncorrelated R_t and mobility indices, while 5–9% of wave 1 outbreaks and 19–24% of wave 2 outbreaks showed negative correlations. Among the 40 cities with outbreaks in both waves 1 and 2, we found no statistically significant correlation between the cross-correlation in two waves from the same cities (correlation: 0.08, $p = 0.61$).

Rolling correlation

The rolling correlation among 27 outbreaks in waves 1 and 64 outbreaks in wave 2 demonstrated frequent and substantial fluctuations during the outbreaks lasted at least 42 days, particularly those with prolonged outbreaks. The rolling correlation of using three different types of mobility indices for an outbreak showed a similar changing pattern (Fig. 3, Supplementary Data 1, 2). In particular, we observed an oscillating pattern between R_t and mobility indices.

Regression analysis showed that the constant model performed significantly worse than non-constant models in all outbreaks (Fig. 3d–f). Specifically, among 91 outbreaks (27 in waves 1 and 64 in wave 2), the sine/cosine models yielded the best fit in 68–75 (75–82%) outbreaks for within-city movement, inter-city inflow, and inter-city outflow, respectively. In the remaining outbreaks, the best fit was achieved using either a linear or quadratic model. sensitivity analysis using triweekly rolling correlation instead of biweekly rolling correlation showed smoother patterns (Supplementary Data 3, 4), but the constant model still performed significantly worse than all other models (Supplementary Fig. 3).

We also tested using the GAM model to fit the rolling correlation, and the results strongly support non-linearity (Supplementary Fig. 4). We also tested the relationship between the rolling correlations of waves 1 and 2 for the same city but found no statistically significant correlation between the rolling correlations in the two waves among 15 cities with outbreaks in both waves 1 and 2.

Comparison of cross-correlation and rolling correlation

No correlation was found between the magnitude of the cross-correlation and the extreme values of the rolling correlation (Fig. 4). There were negative values of minimum of rolling correlation among 86–90% of outbreaks (Fig. 4a, c, e). Moreover, the sign of cross-correlation and the minimum of rolling correlation were opposite among 68–77% of outbreaks. On the other hand, the maximum of rolling correlation was almost 1 in all outbreaks, but the cross-correlation is negative in 13–18% of outbreaks (Fig. 4b, d, f).

Factors affecting cross-correlation and rolling correlation

We found negative correlations between the duration of outbreaks and cross-correlations between R_t and three mobility indices during both waves 1 and 2, ranging from −0.37 to −0.21 (Table 1). In wave 2, the peak value of R_t was positively correlated with the cross-correlation between R_t and within-city movement and inter-city inflow, but not the inter-city outflow. Urban population density demonstrated significant associations with the cross-correlations (Supplementary Table 3). Notably, cross-correlation patterns showed no distinct regional variation (Supplementary Fig. 10).

Based on the 27 completed outbreaks in wave 1, mixed-effect regression analysis revealed that the rolling correlation for inter-city inflow and inter-city outflow was significantly positively correlated with the post-peak stage, relative to the pre-peak stage (Table 1). For inter-city inflow and inter-city outflow, the rolling correlation increased from 0.50 (95% CI: 0.26, 0.68)
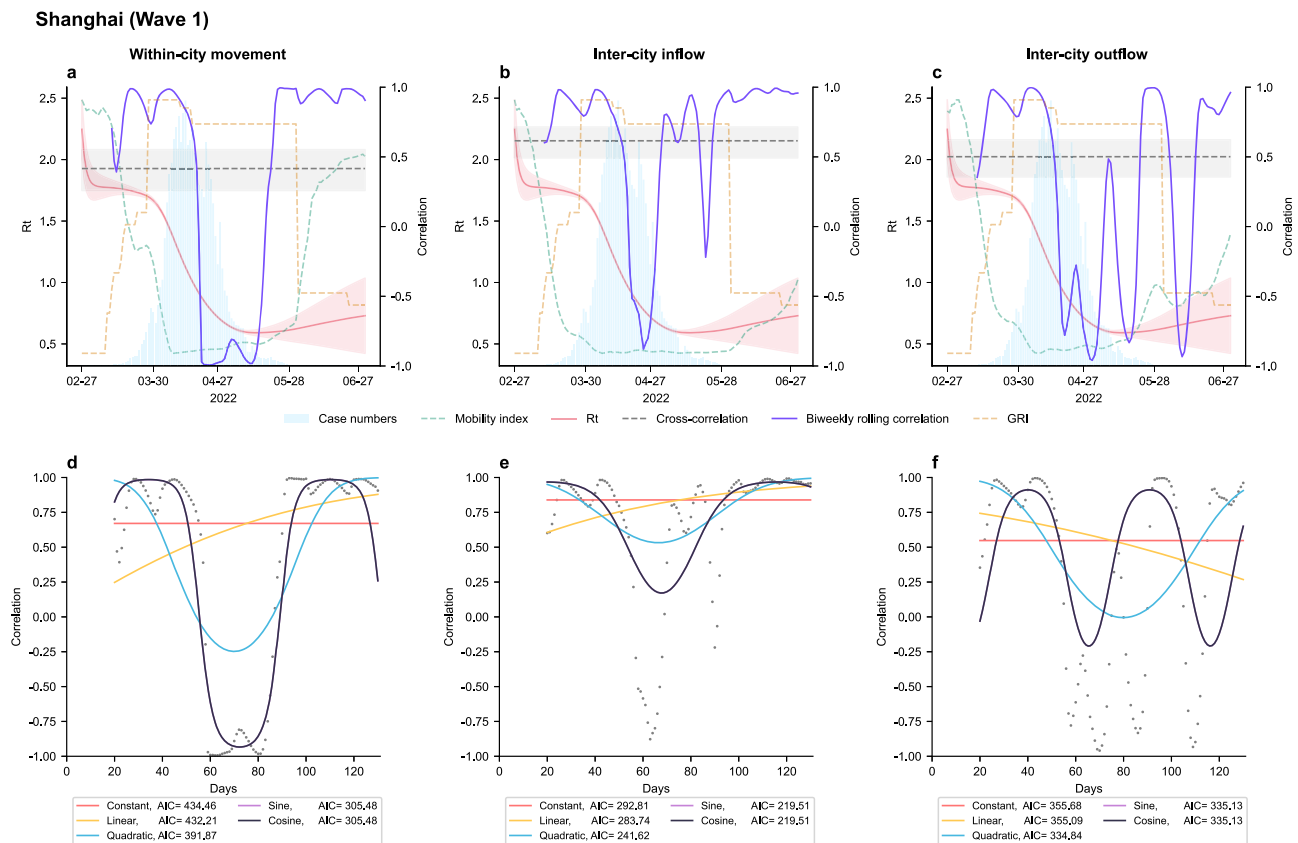
Fig. 1 | Data and analysis results of the Shanghai outbreak in wave 1. a–c Case number (blue), estimated reproduction number (R_t, red; 95% confidence intervals derived from parametric approximation as light red shading), mobility index (green), Government Response Index (GRI, orange), cross-correlation (gray; 95% confidence intervals derived from the bootstrap as light gray shading), and rolling correlation (purple): **a** within-city movement, **b** inter-city inflow, and **c** inter-city outflow. **d–e** Model fitting results of rolling correlations between mobility and transmission by mobility index: **d** within-city movement, **e** inter-city inflow, and **f** inter-city outflow, evaluated using the Akaike Information Criterion (AIC). Gray scatter points show rolling correlations; colored lines represent fitted models: constant (red), linear (orange), quadratic (blue), sine (purple), and cosine (black).

to 0.71 (95% CI: 0.57, 0.81) ($p < 0.01$), and from 0.50 (95% CI: 0.29, 0.66) to 0.66 (95% CI: 0.52, 0.77) ($p < 0.01$) respectively, when the stage changed from pre-peak to post-peak. However, there was no significant change in the within-city movement (Fig. 5a).

We explored the relationship between rolling correlation and GRI for 12 outbreaks in wave 1 due to availability of GRI data. We estimated a positive correlation between rolling correlation and GRI during the pre-peak stage, but a negative correlation during the post-peak stage (Fig. 5b, c). Specifically, during the pre-peak stage, when the GRI increased from 55 to 75, the rolling correlation was estimated to be increased from 0.04 to 0.99 for within-city movement ($p < 0.01$), and from 0.31 to 0.97 for inter-city outflow ($p < 0.01$), respectively, but no significant change for inter-city inflow (Fig. 5B). In contrast, during the post-peak stage, when the GRI increased from 55 to 75, the rolling correlation was estimated to be decreased from 0.86 to 0.33 for within-city movement ($p < 0.01$), from 0.89 to 0.45 for inter-city inflow ($p < 0.01$), and from 0.78 to 0.57 for inter-city outflow ($p = 0.01$) (Fig. 5c).

To assess the robustness of our findings, we conducted sensitivity analyses using: (1) alternative GRI extraction thresholds and (2) a modified outbreak phase classification based on peak case counts rather than peak R_t value (Supplementary Method). Both approaches produced results consistent with our primary analysis (Supplementary Figs. 6 and 9).

Comparison of cross-correlation and rolling correlation across three mobility indices

Inter-city inflow and inter-city outflow had high average correlations (0.88–0.96) in both waves, while within-city movement had lower

correlations with inter-city inflow/outflow (0.31–0.35) (Supplementary Fig. 7). However, the cross-correlations between R_t and mobility of within-city movement, inter-city inflow, and inter-city outflow were highly correlated with each other (ranging from 0.63 to 0.88). Additionally, mixed-effect regression analysis showed no significant differences in rolling correlations for these three mobility indices (Supplementary Fig. 8).

Discussion

Previous studies usually explored the relationship between mobility and transmission at the province or national level (Supplementary Method). Here, we analyzed this relationship at a finer scale, namely city level, using data from Omicron outbreaks in mainland China in 2022, and three types of mobility indices (within-city movement, inter-city inflow, and inter-city outflow). Overall, we found their relationship could be different among cities and outbreaks, suggesting that directly using mobility to proxy transmission intensity may be inaccurate. In particular, we found that the cross-correlation and rolling correlation between the waves for the same city could be different. Furthermore, the rolling correlations were also non-constant in all outbreaks, suggesting the relationship between transmission and mobility was time-varying during outbreaks. Despite these variations, we estimated that the rolling correlation was higher after the peak in outbreaks, and the intensity of control measured (measured by GRI) could modify the rolling correlation.

We found that the cross-correlation between R_t and mobility indices in outbreaks in different waves in the same cities could be different, particularly the cross-correlation during wave 2 was significantly lower than that of wave 1. This was consistent with previous studies reporting lower correlation

Fig. 2 | Cross-correlation between R_t and mobility.
a Cross-correlation between R_t and mobility of
outbreaks in wave 1 (n = 57). **b** Cross-correlation
between R_t and mobility of outbreaks in wave 2
(n = 171). Each data point represents the cross-
correlation between R_t and mobility for an indivi-
dual outbreak, with outbreaks lasting less than 42
days shown in yellow and those lasting 42 days or
more shown in blue. The distribution of the cross-
correlation is shown, with the central white dot
indicating the weighted average and the black bar
representing the 95% confidence intervals derived
from the bootstrap.
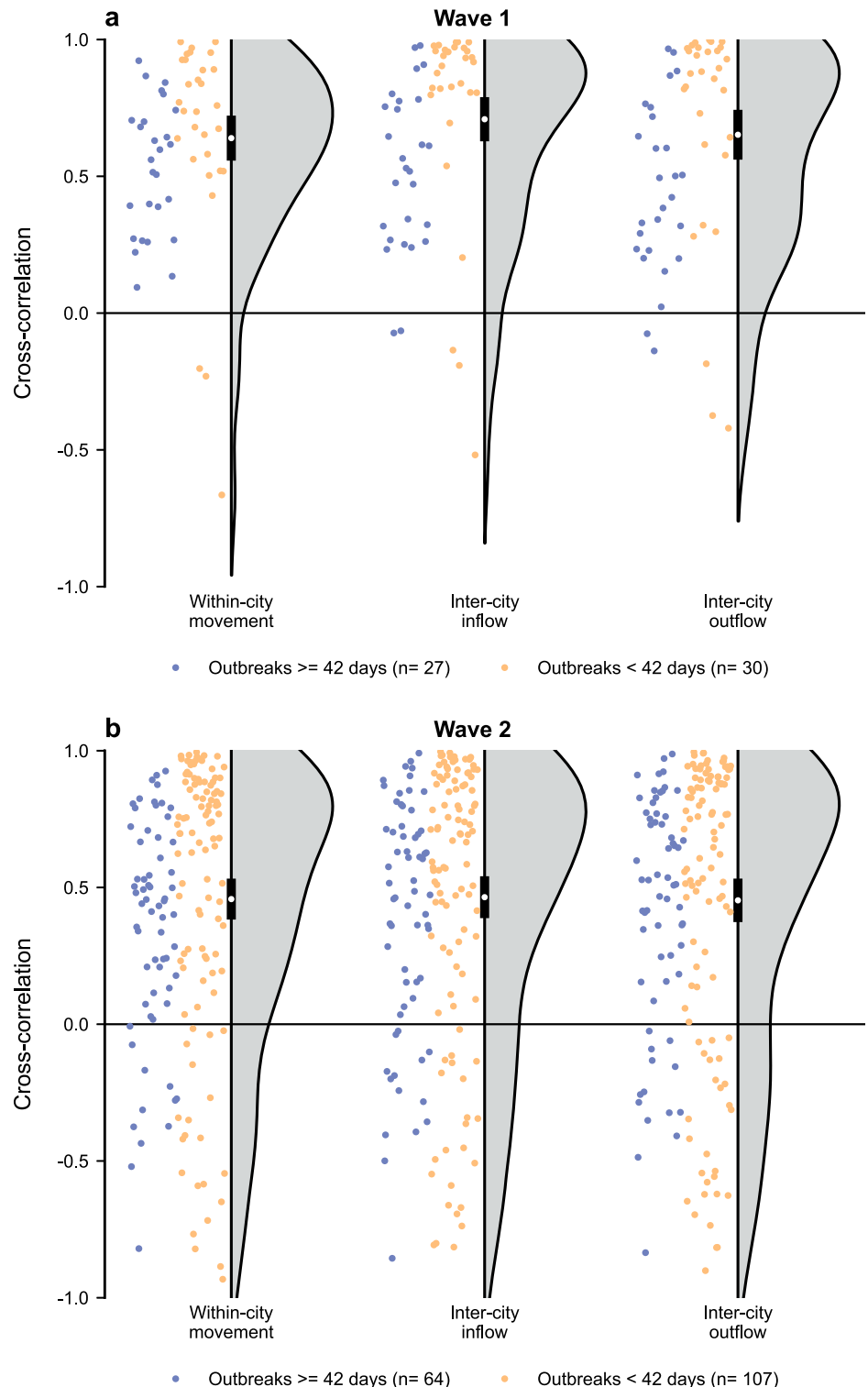
during later waves[3,34] than earlier waves. It may be explained by the increased transmissibility of the dominant virus, in which Omicron BA.5 in wave 2 was more transmissible than Omicron BA.2[38]. Another potential explanation was the pandemic fatigue[39] that self-protection behavior may have changed in the later wave, which cannot be fully captured by mobility indices. Therefore, the previous estimated relationship between mobility and transmission may not be generalizable to future outbreaks even in the same regions.

By using rolling correlation, we found that the relationship between mobility and transmission varied during outbreaks at the city level. Overall,

mobility was positively correlated with transmissions, but in some out-breaks, the minimum rolling correlation could be negative. This finding was consistent with a previous study that showed models allowing for different associations between R_t and mobility in subperiods outperformed models without subperiods[40]. Furthermore, prior studies conducted at both pro-vincial and city levels have demonstrated significant fluctuations in the rolling correlation between R_t and mobility indices throughout the entire epidemic period[20,33].

Increased mobility is expected to increase transmission due to more contacts in the community, but the magnitude of this effect can vary,
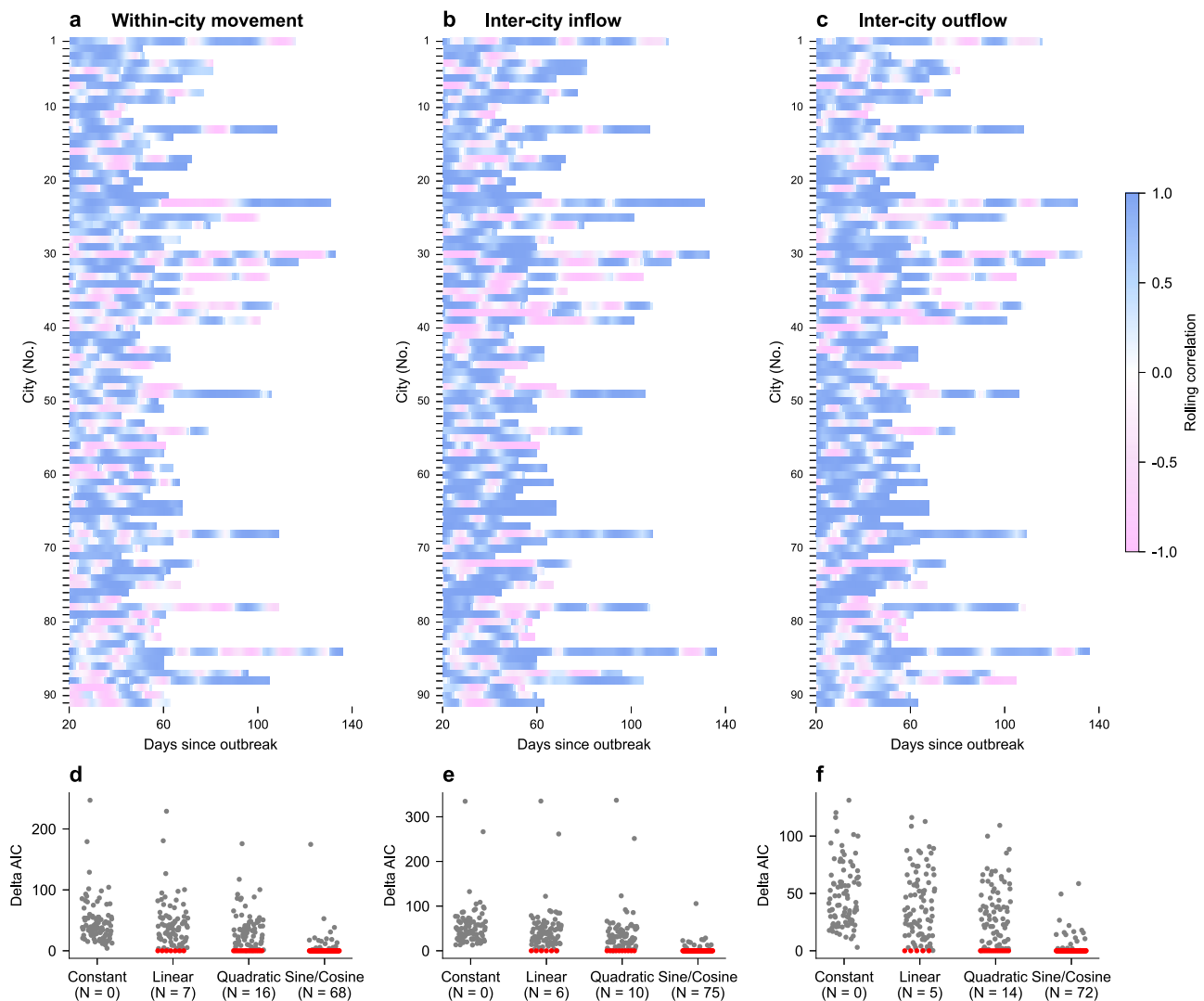
Fig. 3 | Rolling correlation between R_t and mobility. a–c Rolling correlation between R_t and three mobility indices of outbreaks in wave 1 ($n = 27$) and wave 2 ($n = 64$): **a** within-city movement, **b** inter-city inflow, and **c** inter-city outflow. City numbers (No.) are displayed, with corresponding city names provided in Supplementary Table 4. **d–f** Comparison of models fitting rolling correlation between R_t and three mobility indices: **d** within-city movement, **e** inter-city inflow, and **f** inter-city outflow. The gray scatter plots represent individual outbreaks. The red scatter plots represent the best model based on the Deviance Information Criterion (AIC). N denotes the number of outbreaks where each model demonstrated best fit (lowest AIC).

resulting in a time-varying association between R_t and mobility. This variability may be due to 1) factors that may affect transmission but not mobility, 2) factors that may disproportionally affect mobility and transmission, and 3) transmissions could have impact on mobility. First, the use of community mobility data may not capture the transmission intensity in smaller settings like households or neighborhoods. However, given that COVID-19 is more likely to spread indoors and in crowded spaces, the disease can still propagate even with minimal community mobility. Second, implementing various non-pharmaceutical interventions related to mobility restrictions and targeting specific destinations for reduced mobility can vary in their effectiveness in controlling disease transmission. For example, school closures may be more effective than broader measures like business closures or stay-at-home orders[41]. Focusing on reducing mobility to places such as bars and gyms could be more efficient than implementing general mobility restrictions[42]. Third, individual behaviors may change during outbreaks, including the degree of self-protection and contact patterns[43] that may have an impact on transmission intensity, but could not be captured by mobility indices. Also, higher transmissions may increase the individual self-protective behaviors, such as self-isolation or reducing going outside. Fourth, the intensity of self-protective behavior, and intensity and

adherence to implemented containment measures may also change due to pandemic fatigue[39], which can disproportionally affect both mobility and R_t[44,45]. On the other hand, prolonged and repeated outbreaks may also cause pandemic fatigue.

We estimated that the higher GRI was associated with higher rolling correlation before the epidemic peak but lower rolling correlation after the peak, which may explain the time-varying relationship between mobility and transmissions. As containment measures may reduce both mobility and transmission, causing positive and higher rolling correlation in the early stages of outbreaks. However, in later stages of outbreaks, there could be implementations of further measures, such as facial coverings, depending on the transmission intensity, that could weaken the relationship between R_t and mobility[46]. For example, higher transmission intensity may further trigger implementation of measures that increased GRI, resulting decreased rolling correlation. Additionally, when transmission began to decline, control measures were not immediately relaxed, causing a delay in the decline of GRI. Furthermore, a decrease in risk perception and compliance towards control measures among public could lead to an increase in mobility[47], resulting in a weaker relationship between R_t and mobility despite high GRI.
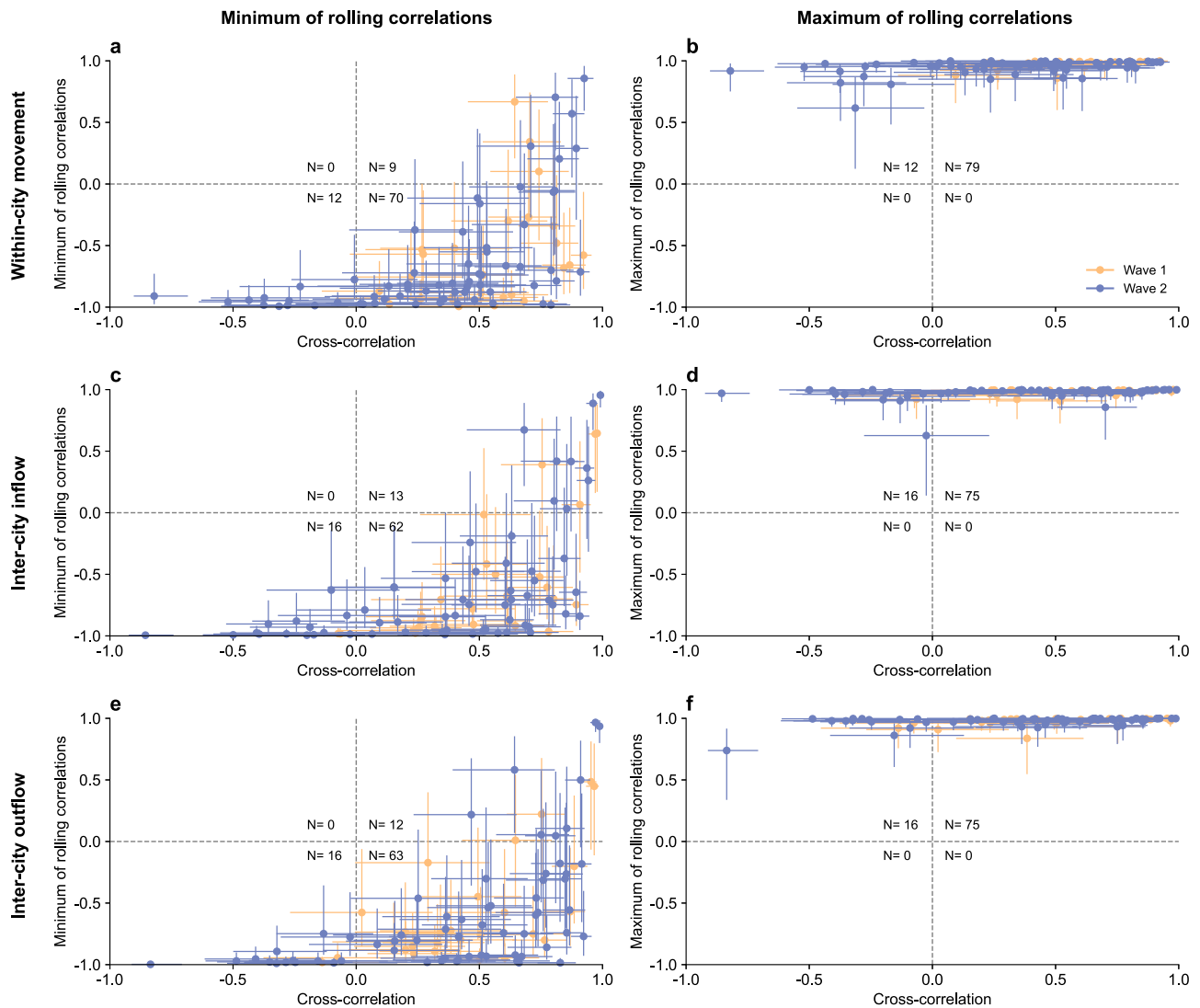
Fig. 4 | Comparison of cross-correlations and the minimum/maximum values of biweekly rolling correlations. a, c, e Comparison of cross-correlations and the minimum values of biweekly rolling correlations by mobility index: **a** within-city movement, **c** inter-city inflow, and **e** inter-city outflow. **b, d, f** Comparison of cross-correlations and the maximum values of biweekly rolling correlations by mobility index: **b** within-city movement, **d** inter-city inflow, and **f** inter-city outflow. Data points represent means with 95% percentile bootstrap confidence intervals. Outbreak waves are color-coded (Wave 1: $n = 27$, yellow; Wave 2: $n = 64$, blue).

Table 1 | Factors affecting cross-correlations and rolling correlation between R_t and mobity index

Correlation	Wave	Factor	Within-city movement	Inter-city inflow	Inter-city outflow
Cross-correlation	**Wave 1**	Peak value of R_t	−0.07 (−0.32, 0.20)	−0.09 (−0.34, 0.18)	−0.05 (−0.31, 0.21)
		Outbreak Duration	**−0.35 (−0.56, −0.10)**	**−0.37 (−0.58, −0.12)**	**−0.35 (−0.56, −0.10)**
		Maximum value of GRI	0.19 (−0.07, 0.43)	0.20 (−0.06, 0.44)	**0.27 (0.01, 0.49)**
	Wave 2	Peak value of R_t	**0.31 (0.15, 0.45)**	**0.24 (0.07, 0.39)**	0.12 (−0.04, 0.28)
		Outbreak Duration	**−0.22 (−0.37, −0.05)**	**−0.25 (−0.40, −0.09)**	**−0.21 (−0.36, −0.04)**
		Maximum value of GRI	−0.03 (−0.20, 0.13)	−0.04 (−0.21, 0.12)	−0.05 (−0.22, 0.12)
Rolling correlation(Fisher transformation)	**Wave 1: Pre-peak stage**	GRI	**0.13 (0.07, 0.20)**	0.06 (0.00, 0.11)	**0.09 (0.03, 0.15)**
	Wave 1: Post-peak stage	GRI	**−0.05 (−0.07, −0.03)**	**−0.05 (−0.06, −0.03)**	**−0.02 (−0.03, 0.00)**
	Wave 1	Pre-peak	ref	ref	ref
		Post-peak	0.05 (−0.12, 0.23)	**0.33 (0.17, 0.50)**	**0.25 (0.09, 0.42)**

GRI, Government Response Index.
For cross-correlation, Pearson correlation coefficients and 95% confidence intervals (CI) were displayed.
For rolling correlation, fixed effect results, including coefficients and 95% CI, were shown.
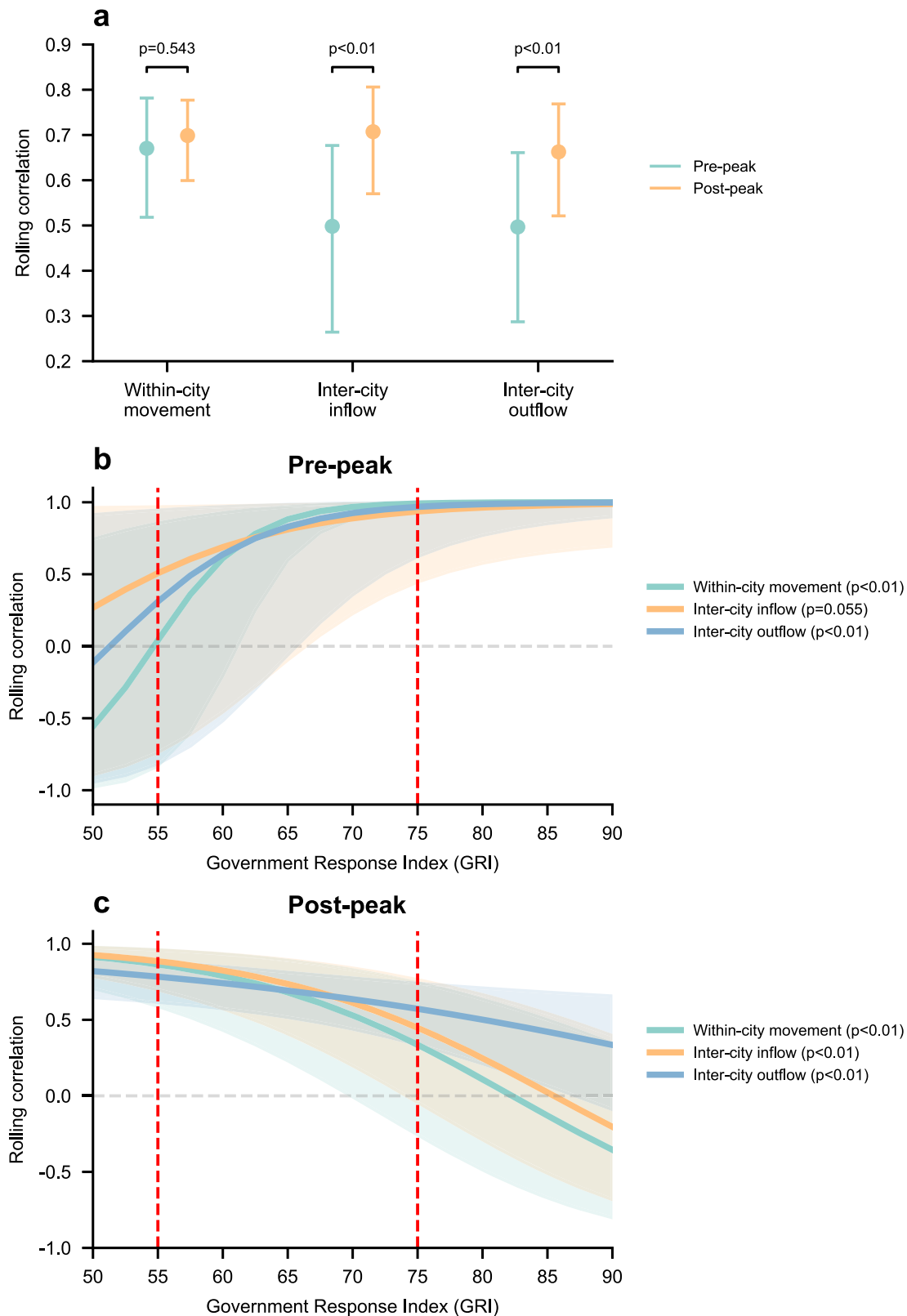
Fig. 5 | Factors affecting rolling correlation estimated by mixed-effect regression analysis. a Rolling correlation between R_t and mobility over two stages of outbreak (points: means; error bars: 95% bootstrap confidence intervals). **b** and **c** Rolling correlation between R_t and mobility in relation to the Government Response Index (GRI) during (**b**) pre-peak stage and (**c**) post-peak stage (shaded regions: 95% bootstrap confidence intervals).

The fluctuations observed in the rolling correlation between transmission and mobility have highlighted several factors that could make their relationship dynamic. Hence, predicting transmission based on mobility, assuming a constant relationship between these variables, could be inaccurate. In particular, for almost all outbreaks, the maximum of rolling correlation could reach one, suggesting that using cross-correlation may underuse the mobility data. More complicated models may be needed to utilize the mobility data to proxy transmission intensity.

When comparing rolling correlation calculated by different mobility indices, there were no significant differences in rolling correlation during the pre-peak stage and during the post-peak stage when using within-city movement as the mobility index. However, there were significant differences when inter-city inflow or outflow was used. These observations suggest that within-city movement may be a more sensitive indicator of outbreaks and capture outbreak information earlier than the other two mobility indices.

This study has several limitations. First, our analysis relied on the accuracy of the mobility data in measuring human movement within and between cities, and this may not capture all important flow changes of the public during the epidemic period[48]. Second, there were missing data, such as city-level case data for Yunnan and Xinjiang provinces. Third, extracting city-level GRI from the province-level data resulted in information loss for some outbreaks, as not all cities had city-level GRI available. Fourth, as the mobility data utilized in this study lacks directional information, we are unable to investigate the impact of city-to-city connectivity. Fifth, we could not validate the relationship between rolling correlation and GRI using data from outbreaks in wave 2, due to the presence of censored data. Lastly, all mobility indices were derived from a single platform, which may introduce platform-specific biases[49]. Further validation is necessary to ensure the strength and reliability of our findings.

In conclusion, using data from a finer scale (city level), our study revealed the dynamic relationship between mobility and transmission, varying across different waves, outbreaks, stages within an outbreak, and level of government response. These suggested that when employing mobility index as a proxy of real-time transmissibility, assuming a constant relationship between them throughout the entire stage may result in inaccurate evaluation of transmission intensity. Therefore, nowcasting and forecasting epidemic using mobility may require further consideration of other factors and development of methodology.

Data availability

The raw data analyzed in the study is sourced from publicly available data, as follows: (1) COVID-19 case data comes from National Health Commission of the People's Republic of China (http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml); (2) mobility data comes from Baidu mobility big data (http://qianxi.baidu.com/); and (3) government response data comes from the OxCGRT, https://github.com/OxCGRT/covid-policy-tracker. The processed datasets used to generate all figures and statistical analyses in this study are available on Zenodo at https://zenodo.org/records/15117352.

Code availability

The code developed in the study to perform the main analysis is available in the GitHub directory at https://github.com/Liping-Peng/COVID19_China and a snapshot of the code is provided on Zenodo[50].

References

1. Cori, A., Ferguson, N. M., Fraser, C. & Cauchemez, S. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am. J. Epidemiol.* **178**, 1505–1512 (2013).
2. Gostic, K. M. et al. Practical considerations for measuring the effective reproductive number, Rt. *PLoS Comput. Biol.* **16**, e1008409 (2020).
3. Lison, A., Persson, J., Banholzer, N. & Feuerriegel, S. Estimating the effect of mobility on SARS-CoV-2 transmission during the first and second wave of the COVID-19 epidemic, Switzerland, March to December 2020. *Euro. Surveill.* **27**, 2100374 (2022).
4. Kajitani, Y. & Hatayama, M. Explaining the effective reproduction number of COVID-19 through mobility and enterprise statistics: evidence from the first wave in Japan. *PLoS One* **16**, e0247186 (2021).
5. Buckee, C. O. et al. Aggregated mobility data could help fight COVID-19. *Science* **368**, 145–146 (2020).
6. Hu, T. et al. Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges. *Int. J. Digit. Earth* **14**, 1126–1147 (2021).
7. Zhang, M. et al. Human mobility and COVID-19 transmission: a systematic review and future directions. *Ann. GIS* **28**, 501–514 (2022).
8. Tan, S. et al. Mobility in China, 2020: a tale of four phases. *Natl. Sci. Rev.* **8**, NWAB148 (2021).
9. Unacast. *Location Data.* https://www.unacast.com/ (2023). Accessed June 26, 2023.
10. Eurocontrol. *Eurocontrol Dashboard.* https://www.eurocontrol.int/ServiceUnits/Dashboard/EnRouteMainDashboard.html (2023). Accessed June 26, 2023.
11. Leung, K., Lau, E. H. Y., Wong, C. K. H., Leung, G. M. & Wu, J. T. Estimating the transmission dynamics of SARS-CoV-2 Omicron BF.7 in Beijing after adjustment of the zero-COVID policy in November-December 2022. *Nat. Med.* **29**, 579–582 (2023).
12. Leung, K., Wu, J. T. & Leung, G. M. Real-time tracking and prediction of COVID-19 infection using digital proxies of population mobility and mixing. *Nat. Commun.* **12**, 1501 (2021).
13. Google. *COVID-19 Community Mobility Reports.* https://www.google.com/covid19/mobility/ (2023). Accessed December 10, 2022.
14. Apple. *Mobility Trends Reports,* https://covid19.apple.com/mobility (2022). Accessed April 1, 2022.
15. da Silva, T. T., Francisquini, R. & Nascimento, M. C. V. Meteorological and human mobility data on predicting COVID-19 cases by a novel hybrid decomposition method with anomaly detection analysis: a case study in the capitals of Brazil. *Expert Syst. Appl.* **182**, 115190 (2021).
16. Setti, M. O. & Tollis, S. In-depth correlation analysis of SARS-CoV-2 effective reproduction number and mobility patterns: three groups of countries. *J. Prev. Med. Public Health* **55**, 134–143 (2022).
17. Gottumukkala, R. et al. Exploring the relationship between mobility and COVID-19 infection rates for the second peak in the United States using phase-wise association. *BMC Public Health* **21**, 1669 (2021).
18. Nouvellet, P. et al. Reduction in mobility and COVID-19 transmission. *Nat. Commun.* **12**, 1090 (2021).
19. Zheng, J. X. et al. The rapid and efficient strategy for SARS-CoV-2 Omicron transmission control: analysis of outbreaks at the city level. *Infect. Dis. Poverty* **11**, 114 (2022).
20. Ainslie, K. E. C. et al. Evidence of initial success for China exiting COVID-19 social distancing policy after achieving containment. *Wellcome Open Res.* **5**, 81 (2020).
21. Ji, H. et al. The effectiveness of travel restriction measures in alleviating the COVID-19 epidemic: evidence from Shenzhen, China. *Environ. Geochem. Health* **44**, 3115–3132 (2022).
22. National Health Commission of the People's Republic of China. *Tracking the COVID-19 Epidemic in China* http://www.nhc.gov.cn/xcs/yqtb/list_gzbd (2024).
23. Gao, J. & Zhang, P. China's public health policies in response to COVID-19: from an "Authoritarian" perspective. *Front Public Health* **9**, 756677 (2021).
24. Baidu. *Baidu cityCode.* https://lbsyun.baidu.com/index.php?title=open/%E5%BC%80%E5%8F%91%E8%B5%84%E6%BA%90 (2023). Accessed 20 February, 2023.
25. Jiang, H. et al. Exploring the inter-monthly dynamic patterns of chinese urban spatial interaction networks based on Baidu migration data. *ISPRS Int. J. Geo-Inf.* **11**, 11090486 (2022).
26. Wei, S. & Wang, L. Examining the population flow network in China and its implications for epidemic control based on Baidu migration data. *Humanities Soc. Sci. Commun.* **7**, 145 (2020).
27. Baidu. *2022 Annual Report on Urban Transportation in China.* https://huiyan.baidu.com/reports/landing?id=138 (2022). Accessed 20 June, 2023.
28. Blavatnik School of Government, University of Oxford. *Oxford Covid-19 Government Response Tracker (OxCGRT).* https://github.com/OxCGRT/covid-policy-tracker (2023). Accessed December 20, 2022.

29. Hale, T. et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat. Hum. Behav.* **5**, 529–538 (2021).

30. Zha, H. et al. "Chinese Provincial Government Responses to COVID-19". Version 2. *Blavatnik School of Government Working Paper* (2022).

31. He, X. et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020).

32. Miller, A. C. et al. Statistical deconvolution for inference of infection time series. *Epidemiology* **33**, 470–479 (2022).

33. Ji, H. et al. The effectiveness of travel restriction measures in alleviating the COVID-19 epidemic: evidence from Shenzhen, China. *Environ. Geochem. Health* **44**, 3115–3132 (2021).

34. Dainton, C. & Hay, A. Quantifying the relationship between lockdowns, mobility, and effective reproduction number (Rt) during the COVID-19 pandemic in the Greater Toronto Area. *BMC Public Health* **21**, 1658 (2021).

35. Cintia, P. et al. The relationship between human mobility and viral transmissibility during the COVID-19 epidemics in Italy. arXiv preprint arXiv:2006.03141 (2020).

36. Peng, L. et al. Comparative epidemiology of outbreaks caused by SARS-CoV-2 Delta and Omicron variants in China. *Epidemiol. Infect.* **152**, e43 (2024).

37. Goldberg, E. E., Lin, Q., Romero-Severson, E. O. & Ke, R. Swift and extensive Omicron outbreak in China after sudden exit from 'zero-COVID' policy. *Nat. Commun.* **14**, 3888 (2023).

38. Xu, A. et al. Sub-lineages of the SARS-CoV-2 Omicron variants: characteristics and prevention. *MedComm* **3**, e172 (2022).

39. Gao, H. et al. Pandemic fatigue and attenuated impact of avoidance behaviours against COVID-19 transmission in Hong Kong by cross-sectional telephone surveys. *BMJ Open* **11**, e055909 (2021).

40. Kurita, J., Sugishita, Y., Sugawara, T. & Ohkusa, Y. Evaluating Apple Inc mobility trend data related to the COVID-19 outbreak in Japan: statistical analysis. *JMIR Public Health Surveill.* **7**, e20335 (2021).

41. Hunter, P. R., Colon-Gonzalez, F. J., Brainard, J. & Rushton, S. Impact of non-pharmaceutical interventions against COVID-19 in Europe in 2020: a quasi-experimental non-equivalent group and time series design study. *Euro. Surveill.* **26**, 2001401 (2021).

42. Chang, S. et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2021).

43. Zhang, J. et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* **368**, 1481–1486 (2020).

44. Huisman, J. S. et al. Estimation and worldwide monitoring of the effective reproductive number of SARS-CoV-2. *Elife* **11**, e71345 (2022).

45. Du, E., Chen, E., Liu, J. & Zheng, C. How do social media and individual behaviors affect epidemic transmission and control? *Sci. Total Environ.* **761**, 144114 (2021).

46. Bergman, N. K. & Fishman, R. Correlations of mobility and Covid-19 transmission in global data. *PLoS One* **18**, e0279484 (2023).

47. Petherick, A. et al. A worldwide assessment of changes in adherence to COVID-19 protective behaviours and hypothesized pandemic fatigue. *Nat. Hum. Behav.* **5**, 1145–1160 (2021).

48. Kishore, N. et al. Evaluating the reliability of mobility metrics from aggregated mobile phone data as proxies for SARS-CoV-2 transmission in the USA: a population-based study. *Lancet Digit. Health* **4**, e27–e36 (2022).

49. Gallotti, R., Maniscalco, D., Barthelemy, M. & De Domenico, M. Distorted insights from human mobility data. *Commun. Phys.* **7**, 421 (2024).

50. Peng, L. COVID-19 transmission-mobility relationships in China (2022 Omicron waves): Case data, mobility indices, and analysis code, https://doi.org/10.5281/zenodo.15117352 (2024).

Author contributions

T.K.T., K.E.C.A., and B.J.C. were responsible for study design. L.P. and X.H. were responsible for data collection. L.P. was responsible for data analysis. L.P., K.E.C.A., B.J.C., P.W., and T.K.T. were responsible for data interpretation. L.P. wrote the first draft of the manuscript. All authors contributed to the final draft.

Competing interests

BJC reports honoraria from AstraZeneca, Fosun Pharma, GSK, Haleon, Moderna, Pfizer, Roche, and Sanofi Pasteur. All other authors reported no competing interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43856-025-00906-7.

Correspondence and requests for materials should be addressed to Tim K. Tsang.

Peer review information *Communications Medicine* thanks S.I. and S.S. for their contribution to the peer review of this work. [Peer review reports are available].

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

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