BMJ Public Health

Contact patterns between index patients and their close contacts and assessing risk for COVID-19 transmission during different exposure time windows: a large retrospective observational study of 450 770 close contacts in Shanghai

Yaxu Zheng,[1,2] Bo Zheng,[1] Xiaohuan Gong,[1,2] Hao Pan,[2] Chenyan Jiang,[2] Shenghua Mao,[2] Sheng Lin,[2] Bihong Jin,[2] Dechuan Kong,[2] Ye Yao ,[3] Genming Zhao,[1,4] Huanyu Wu,[2] Weibing Wang [1,4]

To cite: Zheng Y, Zheng B, Gong X, *et al*. Contact patterns between index patients and their close contacts and assessing risk for COVID-19 transmission during different exposure time windows: a large retrospective observational study of 450 770 close contacts in Shanghai. *BMJ Public Health* 2024;**2**:e000154. doi:10.1136/bmjph-2023-000154

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjph-2023-000154).

YZ, BZ and XG contributed equally.

Received 16 April 2023
Accepted 9 January 2024

Check for updates

© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Mr Huanyu Wu; wuhuanyu@scdc.sh.cn and Dr Weibing Wang; wwb@fudan.edu.cn

ABSTRACT

Introduction To characterise age-mixing patterns among index cases and contacts of COVID-19, and explore when patients are most infectious during the disease process.

Methods This study examined all initial 90 885 confirmed index cases in Shanghai and their 450 770 close contacts. A generalised additive mixed model was used to analyse the associations of the number of close contacts with different demographic and clinical characteristics. The effect of different exposure time windows on the infection of close contacts was evaluated using a modified mixed-effects Poisson regression.

Results Analysis of contacts indicated that 82 467 (18.29%; 95% CI 18.17%, 18.42%) were second-generation cases. Our result indicated the q-index was 0.300 (95% CI 0.298, 0.302) for overall contact matrix, and that assortativity was greatest for students (q-index=0.377; 95% CI 0.357, 0.396) and weakest for people working age not in the labour force (q-index=0.246; 95% CI 0.240, 0.252). The number of contacts was 4.96 individuals per index case (95% CI 4.86, 5.06). Contacts had a higher risk if they were exposed from 1 day before to 3 days after the onset of symptoms in the index patient, with a maximum at day 0 (adjusted relative risk (aRR)=1.52; 95% CI 1.30, 1.76). Contacts exposed from 3 days before to 3 days after an asymptomatic index case had a positive reverse transcriptase-PCR (RT-PCR) result had a higher risk, with a maximum on day 0 (aRR=1.48; 95% CI 1.37, 1.59).

Conclusions The greatest assortativity was for students and weakest for people working age not in the labour force. Contact in the household was a significant contributor to the infection of close contacts. Contact tracing should focus on individuals who had contact soon before or soon after the onset of symptoms (or positive RT-PCR test) in the index case.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Transmission of respiratory pathogens such as COVID-19 depends on patterns of contact and mixing across populations, and these patterns are crucial to predict pathogen spread and the effectiveness of control efforts.

⇒ Although SARS-CoV-2 can be detected one to 3 days before symptoms begin, detection of the virus does not necessarily mean that a person is infectious and able to spread the virus to others.

WHAT THIS STUDY ADDS

⇒ Contact patterns between index patients and their close contacts were similar to previous human social contact pattern.

⇒ Contact in the household increased the risk of infection of a close contact.

⇒ Individuals with COVID-19 were most infectious on day of the onset of symptoms (or positive reverse transcriptase-PCR (RT-PCR) test).

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Tracing of all possible index cases and close contacts is impossible due to limited available resources.

⇒ It suggested that surveillance should focus on individuals who had contact before or soon after the onset of symptoms (or positive RT-PCR test) in the index case, especially the symptomatic index case.

INTRODUCTION

The COVID-19 pandemic is a global outbreak of coronavirus and characterised by a high rate of transmission, with a high risk of exposure to and transmission of COVID-19 in indoor congregate settings,[1 2] high transmission mostly among close contacts and spread based on social contacts,[3] and social mixing patterns of host populations.[4] In late February 2022, Shanghai started a new wave of local Omicron epidemic. From that day on, altogether, Shanghai adopted a series of non-pharmaceutical interventions[5] to contain the spread of the SARS-CoV-2 virus. Contact tracing was a key part of the overall strategy of preventing transmission from individuals who are likely to be infectious.[6] In China, a close contact refers to a person who has effective contact with a suspected or confirmed case, typically by inhalation of infectious secretions from coughing, sneezing, laughing, singing, or talking, or by touching contaminated body parts or surfaces followed by ingestion of the pathogen.[7 8]

To understand the spread of infectious diseases, many previous studies focused on age-mixing patterns and measured the characteristics of general population and their contacts in different age groups.[9–11] However, there were few studies to report the contact patterns for index cases and close contacts. Data on the interactions of index cases and close contacts can provide early signals related to the identification of close contacts; supply important information regarding the number of close contacts among age cohorts for epidemiological modelling[12]; and increase awareness of the epidemic so that effective intervention measures can be developed and implemented.[13]

Additionally, there is limited data on how long Omicron infections last and the time when an infected individual is most infectious in comparison to other variants. A previous study suggested that in people who develop symptoms, the majority are not infectious before symptoms develop, but two-thirds of cases are still infectious 5 days after their symptoms begin,[14] but this study examined a small sample and had low statistical power. From the inferred distribution of the infectiousness profile, a study in Hong Kong estimated that the infectiousness peaked at 1 before the day of symptom onset.[15] Another study found that close contacts had the highest risk of infection if they were exposed from 2 days before to 3 days after the onset of symptoms in the index patient, with a peak at day 0.[16] Thus, additional epidemiological data are needed to understand the relationship of SARS-CoV-2 transmissibility with the timing of symptoms and diagnosis.

The aim of this study was to use data on index cases and contact tracing to quantify the contact patterns of index cases and contacts, and to investigate the association of the time of exposure with the development of disease in close contacts of index patients.

MATERIALS AND METHODS

Data sources and study design

Data on contact details for this study were sourced from the provincial Center for Disease Control and Prevention (CDC) in Shanghai, China, identifying patients newly diagnosed with COVID-19 and their close contacts between 1 March and 31 May 2022. Close contacts were quarantined for at least 14 days and received clinical examinations, including at least four times reverse transcriptase-PCR (RT-PCR) test during centralised quarantine period: on days 1, 4, 7 and 14.

This study was divided into two sections. The first section was an analysis of contact patterns. In this section, data were from 90 885 confirmed index cases in Shanghai and their 450 770 close contacts. The second section was an analysis of transmission risk at different exposure time windows. In this section, 409 close contacts were first excluded because of errors in the date of the onset or end of contact. An additional 36 110 contacts were excluded because their exposure times were not between 14 days before and 14 days after the positive RT-PCR result of the index patient. A previous systematic review reported the pooled incubation time of the Omicron variant was 3.42 days (95% CI 2.88, 3.96), slightly less than that of the Alpha variant.[17] Therefore, some contacts whose exposures occurred too early before or too late after the onset of symptoms (or positive RT-PCR result for the asymptomatic) in the symptomatic patient were also excluded. The final analytic sample consisted of 396 482 close contacts (figure 1).

Identification of cases and contact tracing

An index patient was defined as the first eligible patient with a diagnosis of COVID-19 who had one or more contacts, with confirmation by RT-PCR. The clinical severity of disease in each patient was defined as asymptomatic, mild, moderate, severe or critically ill.[16] An asymptomatic infection is defined as a PCR-confirmed individual who entry into hospital or quarantine (i) does not meet any of the following clinical criteria: fever, cough, sore throat, and other self-perceived and clinical-identifiable symptoms or signs; and (ii) has no radiographic evidence of pneumonia.[18] When a patient had a laboratory-confirmed infection, a thorough epidemiological investigation was administered to the index patient and contacts, and all demographic and clinical data (sex, age, occupation, residence, clinical severity, etc) and exposure-related characteristics were recorded by the Shanghai CDC. Health officials defined a close contact as an individual who was not wearing personal protection equipment and had contact with a suspected or confirmed case within 4 days before symptom onset or a positive RT-PCR result. To further prevent transmission, the window of investigation could be extended to 7 days before the date of symptom onset or the first positive RT-PCR result of an asymptomatic infected person to the day of entry into quarantine, especially for index cases who were from densely populated environments, such as factories, construction sites and farmers' markets.

Different exposure time windows

The exposure window was defined as the time between the first and last day of reported exposure to the index case, based on contact investigation.[19] According to the time from symptoms onset (or positive RT-PCR result for the asymptomatic) to exposure by the first day, contacts of index cases were classified into the following timing

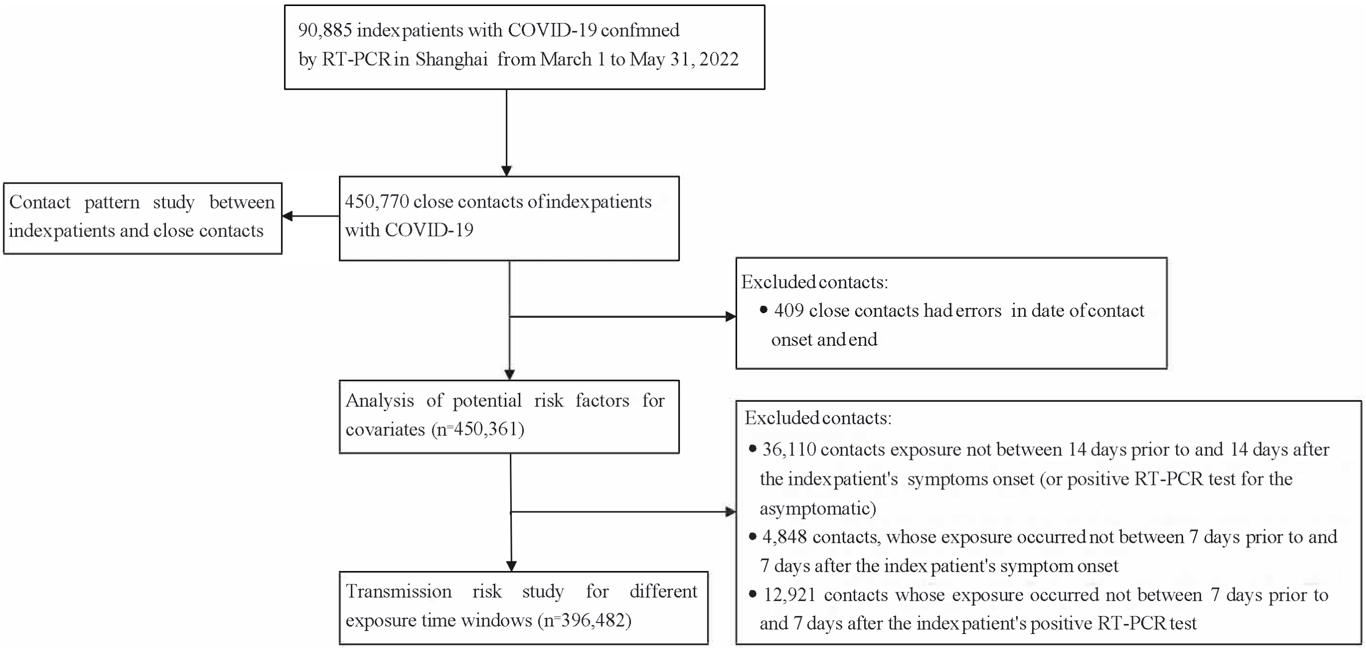

Figure 1 Disposition and inclusion of COVID-19 index patients and close contacts. The number of exposure events is not the same as the number of close contacts because of missing data in some close contacts (date of onset or end of exposure date) or index patients (symptom onset date and diagnosis date). RT-PCR, reverse transcriptase-PCR.

groups: more than 4 days before (−4+); 4 days before (−4); 3 days before (−3); 2 days before (−2); 1 day before (−1); 0 days before (0); 1 day after (1); 2 days after (2); and three or more days after (3+). Contacts of symptomatic cases were also divided into groups based on the interval from the onset of symptoms in the index case to diagnosis: day 0; day 1; day 2; day 3; and day 4 or more. Contacts of asymptomatic cases were grouped according to the time from a positive RT-PCR result to exposure and from a positive RT-PCR result to diagnosis, as indicators of different exposure time windows. Cumulative exposure duration was defined as the total number of days that a contact had exposure, from 4 days before to 4 days after the day of symptoms onset (or positive RT-PCR result for the asymptomatic) in the index patient.

Data processing and statistical analysis

Different age classes were used to develop age-specific contact matrices,[11] with stratification by occupation and confirmed disease duration of index cases, to estimate the number of age-specific close contacts per index case. The q-index, which represents departure from proportionate mixing and ranges from 0 (proportionate) to 1 (fully assortative), and the bootstrapped 95% CIs were used to assess the degree of age assortativity.[20]

A generalised additive mixed model (GAMM[21] with a negative binomial distribution was used to analyse the association of the number of close contacts with demographic and clinical parameters. A χ^2 test was used to compare the distribution of the clinical severity of index cases and secondary cases, and the duration from symptom onset to a positive RT-PCR result. The interval between a positive RT-PCR result and diagnosis was examined using a two-sample t-test.

The attack rate (AR) was defined as the total number of new cases diagnosed among contacts of index patients divided by the total number of exposed contacts.[22] The AR was estimated for all contacts, and then separately for each exposure-related characteristic. A modified mixed-effects Poisson regression with a robust error variance was used to analyse infection risk for categorical covariates and different exposure time windows. This model has a logarithmic link function that considers clustering of contacts and allows for direct estimation of relative risk in observational studies.[16] Univariate regression was used to analyse factors related to infection of contacts. Then, the significant independent variables in the univariate analysis (age, sex, confirmed time window and occupation of index cases, household contact) were included in a multivariable regression model. P values and 95% CIs were used to assess statistical significance in all models. All statistical analyses were performed using R.4.1.1 software.[23] A two-sided p value less than 0.05 was considered significant.

Patient and public involvement

Neither the patients nor the public were involved in the conceptualisation, design, conduct, reporting or dissemination of this study as it was a large retrospective study of 450 770 close contacts. Data analysis and article writing in our study was designed to be fully anonymous and no private information (eg, name of participant) was recorded.

RESULTS

Demographic characteristics of index cases, close contacts and secondary cases

The index patients consisted of 44 384 women (48.84%) and 46 501 men (51.16%) and the median age was 46.00 years (IQR: 32.00, 59.00 years), compared with 40.81 years (IQR: 28.70, 54.59) among the close contacts (online supplemental table S1). Among the 450 770 close contacts, 82 467 (18.29%, 95% CI 18.17%, 18.42%) received diagnoses of COVID-19. The mean interval from symptom (online supplemental table S2) onset to a positive RT-PCR result was 1.30 days (95% CI 1.26, 1.34) in index cases and 1.50 days (95% CI 1.46, 1.53) in secondary cases (p<0.001). Analysis of the duration between a positive RT-PCR result to diagnosis was 2.30 days (95% CI 2.29, 2.31) in index cases. The mean number of secondary cases per index patient was 0.91 (IQR: 0.89, 0.92) and the median number was 0.0 (IQR: 0.00, 1.00).

The mean number of secondary cases remained stable during April and May (figure 2), and was approximately 1 per index case. Analysis of the index cases indicated that 26.17% were confirmed on day 2 after symptom onset (online supplemental figure S1), but only 1.20%

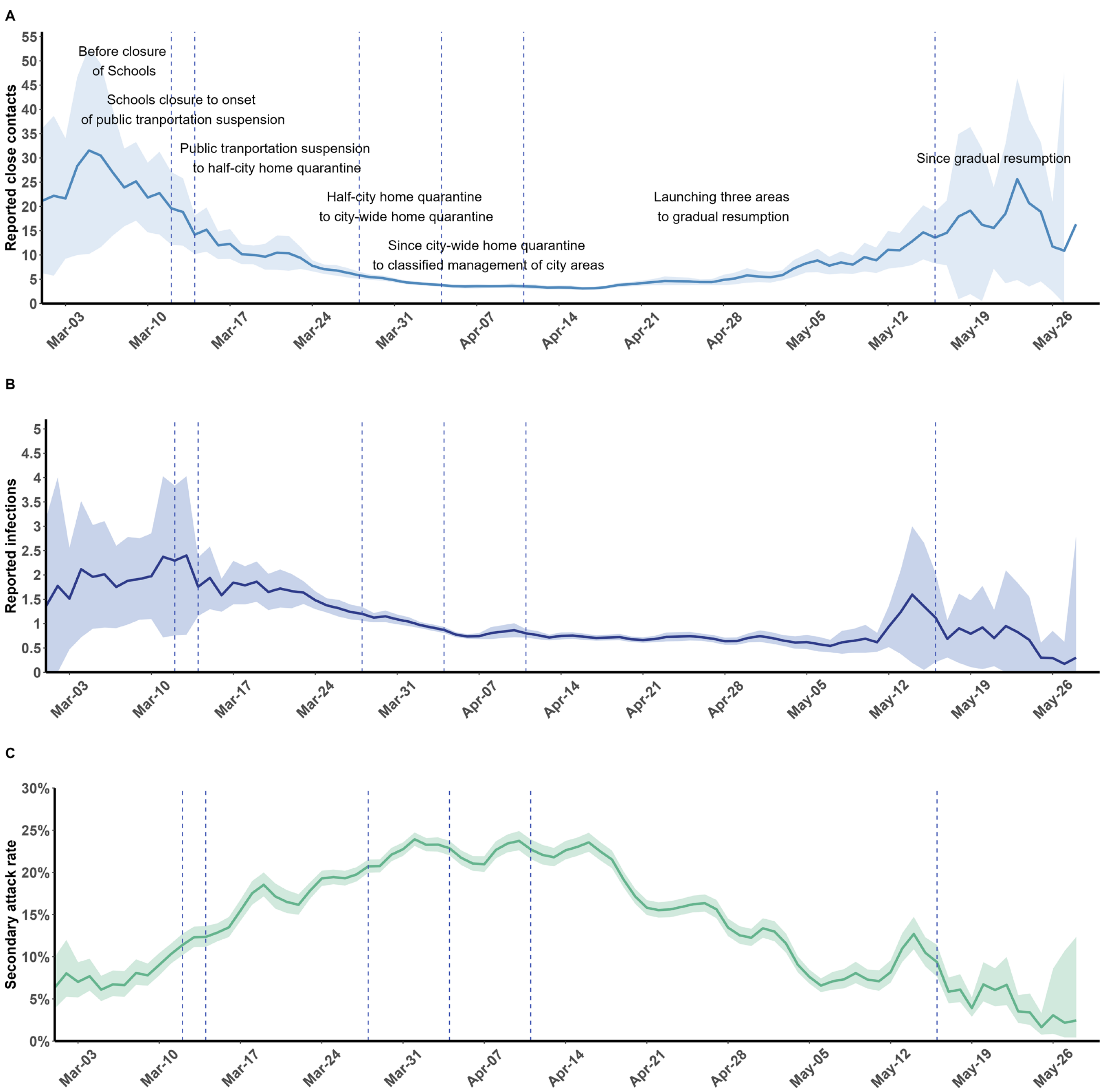

Figure 2 Changes in close contacts per index patient (A), secondary infections (B) and secondary attack rate (C). Shaded regions: 95% CIs.

on day 0 or before symptom onset. Most secondary cases were diagnosed 2 days (54.16%) or 1 day (23.12%) after symptom onset.

Contact patterns of index cases and close contacts

The distribution of the number of close contacts was positively skewed (online supplemental figure S2). Overall, only 1.01% of index cases had more than 50 close contacts, and most of index cases reported 1 close contact (38.14%) or 2 close contacts (21.00%). More than 48% of retired index cases reported 1 close contact, but 30.67% of school student index cases reported 1 close contact (online supplemental figure S2).

Among all 450 770 close contacts, most (50.32%) were at home (figure 2), and fewer were at the workplace (9.65%), farmers' market (4.17%) and hotels/restaurants (0.51%). As the contact duration increased, the proportion of contacts at home also increased (figure 3E). For 81.52% of the index cases with contact durations more than 1 day, the close contact was a daily contact (figure 3). Among daily contacts, 81.30% were at home (figure 3). Individuals who were family members, relatives or colleagues were less likely to be contacts (figure 3).

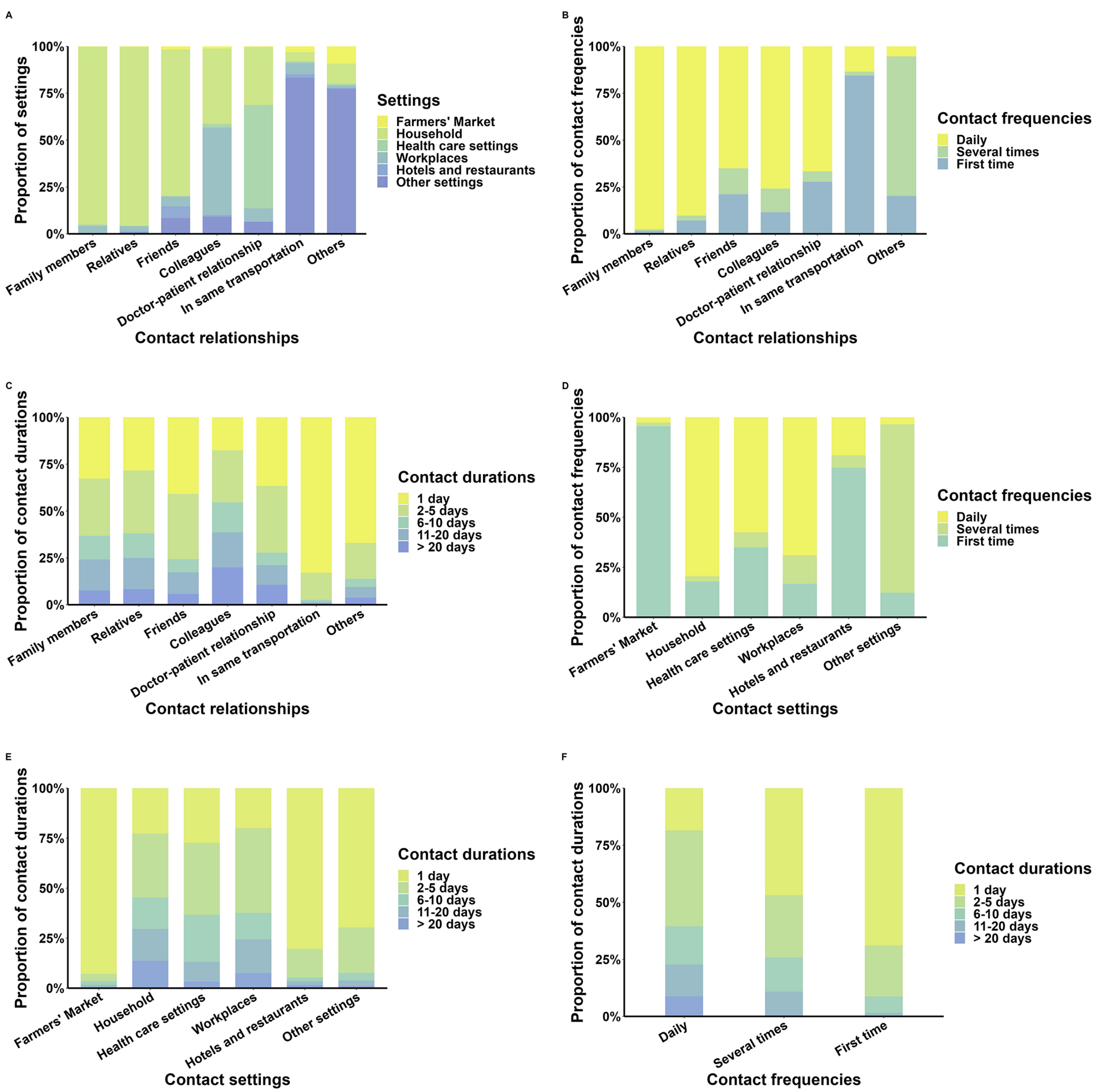

Figure 3 Patterns of contact of index patients and close contacts according to setting, frequency, relationship and duration. The proportions of contact relationships according to individual contact settings (A), frequencies (B) and durations (C). The proportions of contact settings according to contact frequencies (D) and durations (E). (F) Same as (A-E), but the proportions represent the contact frequencies according to contact durations. Farmers market is a public and recurring assembly of farmers or their representatives selling the food that they produced directly to consumers.

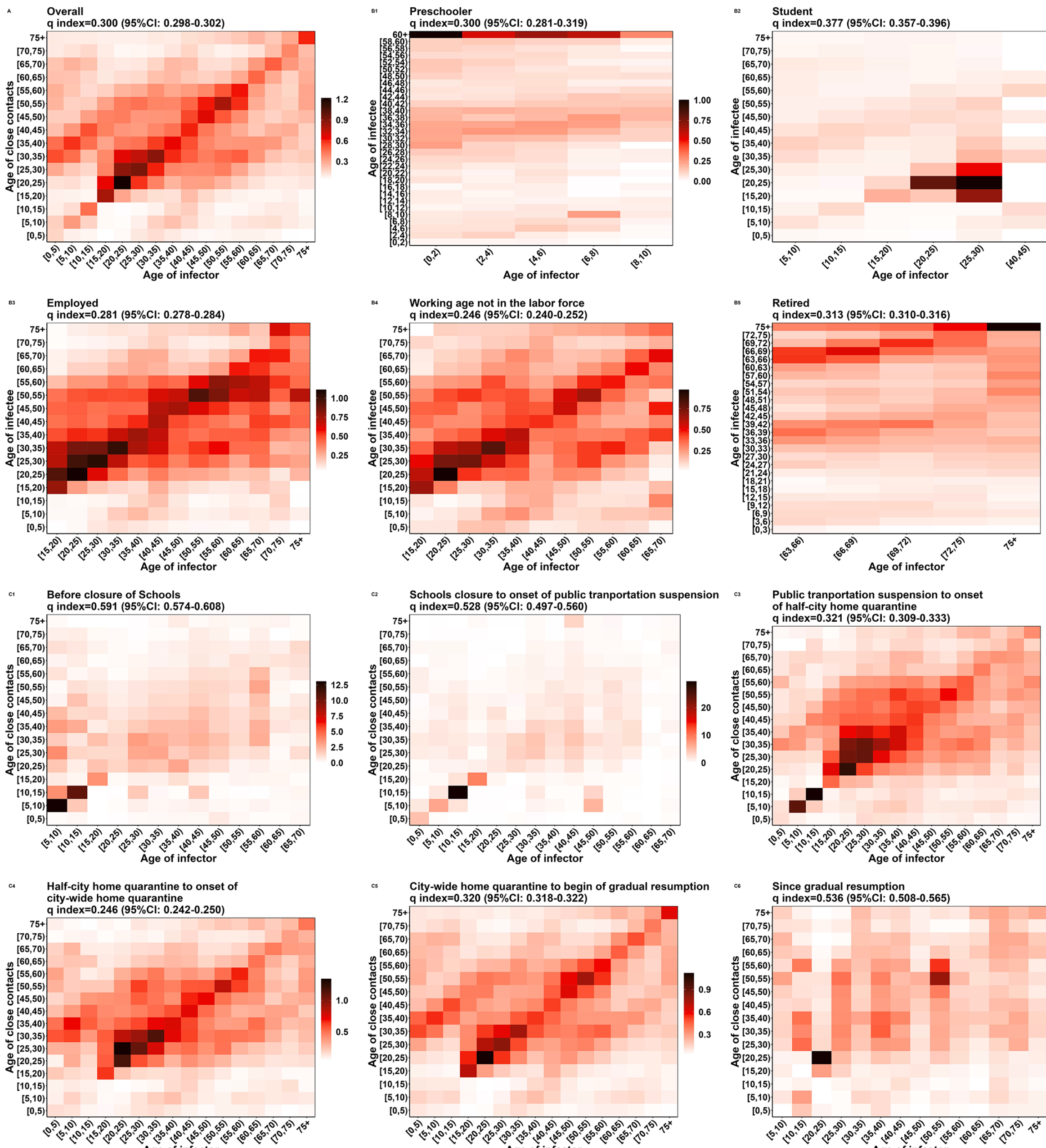

Figure 4 Contact matrices for index patients and close contacts according to age. Each cell of a matrix represents the mean number of close contacts that an index patient in an age group had with other close contacts of different age groups. Matrices are for close contacts overall (A), different epidemic phases (B1–B5), and different occupations (C1–C6). To construct the matrix, bootstrap sampling with replacement of index patients was performed with weighting by the age distributions of the index patients. Each cell of the matrix represents an average more than 100 bootstrapped results.

Contact matrix and assortativity of index cases and close contact

Analysis of the overall assortative mixing for the contact matrix indicated the q-index was 0.300 (bootstrapped 95% CI 0.298, 0.302) (figure 4). The epidemic phase with the lowest assortativity was the half-city home quarantine phase to the city-wide home quarantine phase (q-index=0.246; 95% CI 0.242, 0.250). The occupational contact matrix (figure 4B1,B2) showed that the contact matrix of index cases who were students had the greatest assortativity (q-index=0.377; 95% CI 0.357, 0.396), and assortativity was weakest for people working age not in the labour force (q-index=0.246; 95% CI 0.240, 0.252). Index cases who were preschoolers had fewer contacts with individuals aged 10–22 years, and more contacts with individuals who were aged 60 years or more. Students who were index cases tended to have contact with same-age individuals and those who were aged 30–60 years. Analysis of employed people (figure 4B3,) indicated a strong diagonal signal starting at about age 15–70 years old for contacts and index cases.

The pattern of different age groups of index cases tending have assortative mixing by age also occurred in the contact matrices that were stratified by time window and occupation (online supplemental figures S3-1 and S3-2). In particular, male index patients (q-index=0.301; 95% CI 0.298, 0.303) had more assortativity than female index patients (q-index=288; 95% CI 0.285, 0.290) (online supplemental figure S4). For contact matrices of index cases among registered residents in Shanghai or other provinces, the diagonal element was most pronounced in those who were aged 15–35 years (online supplemental figure S4, B1-B2). The diagonal element was evident in the contact matrix that was stratified by symptoms (online supplemental figure S4, C1-C2).

Factors associated with the number of close contacts of index cases

The GAMM regression mode indicated there were 4.96 (95% CI 4.86, 5.06) close contact per index case (table 1). Relative to male index cases, female index cases had fewer close contacts (OR=0.89, 95% CI 0.88, 0.90). There was also a nonlinear association between the number of close contacts and the age of index cases, with the greatest OR in those aged 20–39 years (OR=1.29, 95% CI 1.11, 1.48) and the smallest OR in those aged more than 60 years (OR=1.03, 95% CI 0.89, 1.19). The number of close contacts of employed index patients (6.35, 95% CI 6.16, 6.54) was significantly larger than that for student index patients (5.18, 95% CI 4.63, 5.72).

Compared with index cases who were residents in Shanghai, those with residences in other provinces had more contacts (OR=1.18, 95% CI 1.16, 1.20). The number of daily reported contacts of per case was slightly (although not significantly) greater before onset of the public transportation suspension phase, and declined significantly and progressively during the next four phases of the epidemic (OR=0.39; 95% CI 0.35, 0.44; OR=0.23; 95% CI 0.21, 0.26; OR=0.18; 95% CI 0.16, 0.20; OR=0.62; 95% CI 0.53, 0.73).

Analysis of categorical covariates and risk for different exposure time windows

The AR (table 2) peaked during the half-city home quarantine phase to the onset of city-wide home quarantine phase, and was lowest in the phase since gradual resumption (20.95%; 95% CI 20.62%, 21.29% vs 3.61%; 95% CI 3.09%, 4.19%). The 197 286 close contacts who were Shanghai residents had a higher risk for secondary infection (18.63%; 95% CI 18.46%, 18.80%) than the 25,3075 close contacts who were residents of other provinces (17.87%; 95% CI 17.69%, 18.06%), but this difference was not significant in the multivariate analysis (adjusted relative risk (aRR)=1.01; p=0.082). The multivariate (online supplemental table S3) analysis also indicated there were significant effects of age, occupation, status as a student and phase of the epidemic (all p<0.001). Analysis of contact setting indicated that contact in a household (aRR=1.32, 95% CI 1.29, 1.34) significantly increased the risk of infection in close contacts.

Our univariate analysis (online supplemental table S4) showed that a contact tended to have increased risk of infection with increases in the time from symptom onset to diagnosis of a symptomatic index case and with increases in the time from a positive RT-PCR test to diagnosis of an asymptomatic case (online supplemental figures S5 and S6). A mixed-effects model that adjusted for categorical covariates showed that contacts had a higher risk of COVID-19 if they were exposed from day −1 to day 3 of the index patient's symptom onset (figure 5), with a maximum at day 0 (aRR: 1.52; 95% CI 1.30, 1.76; AR: 23.37%; 95% CI 21.24%, 25.66%). The time from onset to diagnosis of an index case significantly increased the risk of infection risk for close contacts, and contacts had the highest risk of COVID-19 if the time from an index case's time from symptom onset to diagnosis was 4 days or more (AR: 21.44%; 95% CI 20.81%, 22.10%; crude risk ratio (CRR): 1.68; 95% CI 1.40, 2.01; aRR: 1.60; 95% CI 1.33, 1.91). For asymptomatic index cases, contacts had a higher risk of COVID-19 if they were exposed between day −3 and day 3 (figure 5) from the index patient's positive RT-PCR result, with a maximum at day 0 (aRR: 1.48; 95% CI 1.37, 1.59; AR: 22.20%; 95% CI 21.04%, 23.40%). Asymptomatic index cases tended to have an increased risk of infection of close contacts as time from a positive RT-PCR result to diagnosis increased; contacts had the highest risk of COVID-19 if the time of an index case's positive RT-PCR result to diagnosis was 4 days or more (AR: 23.46%; 95% CI 22.87%, 24.06%; CRR: 1.75; 95% CI 1.63, 1.87; aRR: 1.61; 95% CI 1.50, 1.72). Our results also showed that the cumulative duration of exposure significantly increased the risk of infection of close contacts (figure 5 and online supplemental figure S7). The ARs and infection risk of close contacts tended to increase as the cumulative exposure duration increased (0 days: 17.16%; 95% CI 16.17%, 18.20%; 7+ days: 19.98%–95% CI 19.45%, 20.53%; CRR: 1.16; 95% CI 1.09, 1.24; aRR: 1.13; 95% CI 1.06, 1.21).

DISCUSSION

In the present study, analysis of the overall degree of assortative mixing in the contact matrix indicated the q-index was 0.300 (95% CI 0.298, 0.302) and students had the greatest assortativity (q-index=0.377, 95% CI 0.357, 0.396). Several demographic and clinical factors of index cases (gender, age, residence, clinical severity) were associated with the number of close contacts. We also found that contact in the household increased the risk of infection of a close contact. Moreover, our results showed that the risk of COVID-19 transmission to a close contact was greater if the exposure time was from day −1 to day 3 of symptom onset in a symptomatic index case, and from day −3 to day 3 in an asymptomatic case with a positive RT-PCR result, with maximal risk on day 0 in both cases. As expected, the infection risk of close contacts tended to increase as the duration of cumulative exposure increased.

Similar to the findings of Prem *et al*,[24] we found that high assortativity of contacts was common in students,

Table 1 Regression coefficients used in the generalised additive model

Characteristics		Number of participants n (%)	Close contacts mean (95% CI)	OR (95% CI)	P value
Overall/intercept		90 885	4.96 (4.86, 5.06)	3.18 (2.99, 3.38)***	**<0.001**
Gender	Male	46 501 (51.2)	5.25 (5.09, 5.41)	1	–
	Female	44 384 (48.8)	4.65 (4.52, 4.78)	0.89 (0.88, 0.90)***	**<0.001**
Age group, years	0–6	1649 (1.8)	4.23 (3.48, 4.98)	1	–
	7–19	5215 (5.7)	4.72 (4.32, 5.12)	1.03 (0.9, 1.18)	0.649
	20–39	29 298 (32.2)	5.65 (5.47, 5.84)	1.29 (1.11, 1.48)***	**<0.001**
	40–59	32 545 (35.8)	5.32 (5.13, 5.5)	1.26 (1.09, 1.45)**	**0.002**
	≥60	22 178 (24.4)	3.63 (3.46, 3.8)	1.03 (0.89, 1.19)	0.687
Profession	Preschooler	1627 (1.8)	4.18 (3.43, 4.93)	1	–
	Student	4212 (4.6)	5.18 (4.63, 5.73)	1.00 (0.87, 1.14)	0.962
	Employed	31 750 (34.9)	6.35 (6.16, 6.54)	1.15 (1.03, 1.28)*	**0.013**
	Working age not in the labour force†	14 446 (15.9)	4.74 (4.45, 5.02)	1.11 (0.99, 1.25)	0.068
	Retired	12 089 (13.3)	3.53 (3.33, 3.73)	0.92 (0.84, 1.01)	0.067
	Unknown	26 761 (29.4)	4.1 (3.93, 4.26)	0.93 (0.83, 1.03)	0.163
Clinical severity	Asymptomatic	83 251 (91.6)	5.01 (4.9, 5.12)	1	–
	Mild	7250 (8.0)	4.45 (4.21, 4.69)	0.88 (0.85, 0.9)***	**<0.001**
	Moderate	304 (0.3)	3.88 (2.95, 4.82)	0.87 (0.77, 0.99)*	**0.044**
	Severe	64 (0.1)	3.92 (2.08, 5.77)	1.06 (0.8, 1.4)	0.676
	Critical	16 (0.0)	2.31 (0.8, 3.83)	0.63 (0.35, 1.14)	0.128
Epidemic phase	Before school closure	355 (0.4)	24.05 (19.38, 28.72)	1	–
	Schools closure to onset of public transportation suspension	256 (0.3)	26.83 (21.28, 32.39)	1.14 (0.97, 1.35)	0.119
	Public transportation suspension to onset of half-city home quarantine	6275 (6.9)	10.12 (9.4, 10.83)	0.39 (0.35, 0.44)***	**<0.001**
	Half-city home quarantine to onset of city-wide home quarantine	12 612 (13.9)	5.65 (5.38, 5.93)	0.23 (0.21, 0.26)***	**<0.001**
	City-wide home quarantine to onset of gradual resumption	71 067 (78.2)	4.16 (4.07, 4.26)	0.18 (0.16, 0.2)***	**<0.001**
	After onset of gradual resumption	320 (0.4)	14.46 (11.26, 17.66)	0.62 (0.53, 0.73)***	**<0.001**
Registered residence	Shanghai	57 095 (62.8)	4.44 (4.32, 4.55)	1	–
	Other province	33 790 (37.2)	5.84 (5.65, 6.04)	1.18 (1.16, 1.2)***	**<0.001**

(n) % indicates the proportion of index-patients.
*p<0.05.
**p<0.01.
***p<0.001.
†People of working age (≥18 years) unemployed or not in the labour force (disabled).

Table 2 Attack rate (AR) for COVID-19 among close contacts according to demographic and clinical characteristics of index cases

Characteristic		Contacts who developed COVID-19/total contacts (N/N)	AR (95% CI)
Overall		82 406/450 361	18.3 (18.17, 18.42)
Epidemic phase	Before school closure	581/8497	6.84 (6.3, 7.41)
	School closure to onset of public transportation suspension	570/6864	8.3 (7.64, 9.01)
	Public transportation suspension to onset of half-city home quarantine	10 641/63 434	16.77 (16.46, 17.1)
	Half-city home quarantine to onset of city-wide home quarantine	14 918/71 198	20.95 (20.62, 21.29)
	City-wide home quarantine to onset of gradual resumption	55 529/295 740	18.78 (18.62, 18.93)
	After onset of gradual resumption	167/4628	3.61 (3.09, 4.19)
Residence of index case	Shanghai	47 146/253 075	18.63 (18.46, 18.8)
	Other provinces	35 260/197 286	17.87 (17.69, 18.06)
Age group of index case, years	0–6	1365/6975	19.57 (18.55, 20.63)
	7–19	4138/24 605	16.82 (16.31, 17.34)
	20–39	26 995/165 488	16.31 (16.12, 16.51)
	40–59	32 403/172 799	18.75 (18.55, 18.96)
	≥60	17 505/80 494	21.75 (21.43, 22.07)
Gender of index case	Female	37 275/206 320	18.07 (17.88, 18.25)
	Male	45 131/244 041	18.49 (18.32, 18.66)
Occupation of index case	Preschooler	1530/7504	20.39 (19.39, 21.43)
	Student	2830/21 095	13.42 (12.93, 13.92)
	Employed	35 596/201 182	17.69 (17.51, 17.88)
	Working age not in the labour force	6459/37 266	17.33 (16.91, 17.76)
	Retired	15 525/73 760	21.05 (20.72, 21.38)
Severe/critical disease in index case	No	82 350/450 073	18.30 (18.17, 18.42)
	Yes	56/288	19.44 (14.84, 25.05)
Close contacts in household	No	24 676/139 254	17.72 (17.5, 17.94)
	Yes	32 801/141 031	23.26 (23.01, 23.51)

and less common in index cases working age not in the labour force. Compared with most individuals, those contacted by students and teenagers were of a similar age.[3 25] However, working age not in the labour force were more likely to have contact of diverse ages, which may provide a route for transmission from people working age not in the labour force and the rest of the population, and lead to a greater number of new infections.[26] Previous studies[3 27 28] reported that contact duration, frequency of contact, nature of the relationship and location of contact were associated with one another. For example, the less frequent contacts were less likely to be family members and colleagues. Importantly, intimate contacts had a greater risk of transmission, and these contacts typically occurred at home.[29] For instance, our results suggested that contact within the household (aRR=1.32, 95% CI 1.29, 1.34) was associated with a significantly higher risk of infection of close contacts than contact outside the household. We found that index patients who were residents in other provinces had more close contacts (OR=1.18; 95% CI 1.16, 1.20) than index patients from Shanghai. This might be due to the proximity of places used for gathering, working and living of these no-residents (9.78 million people (40.27% of the total population in Shanghai)).[30] Female index cases had fewer close contacts than male (OR=0.89; 95% CI 0.88, 0.90), possibly because men travel more outside the home with worse awareness of social distancing than female.[31]

Subgroup	Infection	Contacts	Attack rate	Estimate (95% CI)
Cumulative exposure duration, days				
1	1108	6456	17.16 (16.17, 18.20)	
2-3	7906	47656	16.59 (16.23, 16.96)	0.99 (0.93 to 1.06)
4-5	3671	20108	18.26 (17.67, 18.85)	1.06 (0.99 to 1.14)
6-7	3980	20125	19.78 (19.17, 20.40)	1.12 (1.05 to 1.20)
> 7	5216	26103	19.98 (19.45, 20.53)	1.13 (1.06 to 1.21)
Symptomatic:				
Timing from onset to exposure, days				
before -4	870	4896	17.77 (16.62, 18.98)	
-4	585	3504	16.70 (15.38, 18.09)	0.98 (0.87 to 1.10)
-3	313	2205	14.20 (12.69, 15.83)	0.92 (0.79 to 1.07)
-2	285	1540	18.51 (16.45, 20.75)	1.12 (0.95 to 1.31)
-1	355	1779	19.96 (17.96, 22.11)	1.28 (1.10 to 1.49)
0	431	1844	23.37 (21.24, 25.66)	1.52 (1.30 to 1.76)
1	447	1915	23.34 (21.25, 25.58)	1.48 (1.27 to 1.73)
2	284	1251	22.70 (20.18, 25.46)	1.37 (1.16 to 1.63)
3	202	1063	19.00 (16.52, 21.76)	1.21 (1.00 to 1.47)
≥ 4	357	2021	17.66 (15.90, 19.57)	1.04 (0.87 to 1.25)
Timing from onset to diagnosis, days				
0	122	954	12.79 (10.67, 15.21)	
1	1336	9329	14.32 (13.57, 15.10)	1.14 (0.95 to 1.37)
2	3074	15503	19.83 (19.14, 20.54)	1.45 (1.21 to 1.74)
3	2785	13267	20.99 (20.22, 21.78)	1.53 (1.27 to 1.84)
≥ 4	4266	19893	21.44 (20.81, 22.10)	1.60 (1.33 to 1.91)
Asymptomatic:				
Timing from positive RT-PCR test to exposure, days				
before -4	2977	16084	18.51 (17.85, 19.18)	
-4	962	5119	18.79 (17.63, 20.01)	1.07 (0.99 to 1.16)
-3	2373	11210	21.17 (20.33, 22.03)	1.11 (1.05 to 1.18)
-2	1104	5111	21.60 (20.35, 22.90)	1.30 (1.21 to 1.40)
-1	1128	5877	19.19 (18.10, 20.34)	1.18 (1.10 to 1.28)
0	1359	6123	22.20 (21.04, 23.40)	1.48 (1.37 to 1.59)
1	1878	9271	20.26 (19.36, 21.19)	1.28 (1.20 to 1.37)
2	1006	5230	19.24 (18.07, 20.45)	1.20 (1.11 to 1.30)
3	919	4177	22.00 (20.61, 23.46)	1.34 (1.23 to 1.45)
≥ 4	1170	8167	14.33 (13.52, 15.16)	0.78 (0.72 to 0.84)
Timing from positive RT-PCR test to diagnosis, days				
0	959	7152	13.41 (12.58, 14.28)	
1	9272	61722	15.02 (14.72, 15.33)	1.09 (1.02 to 1.17)
2	26478	131645	20.11 (19.87, 20.36)	1.44 (1.35 to 1.54)
3	8473	37312	22.71 (22.23, 23.20)	1.59 (1.48 to 1.70)
≥ 4	5947	25354	23.46 (22.87, 24.06)	1.61 (1.50 to 1.72)

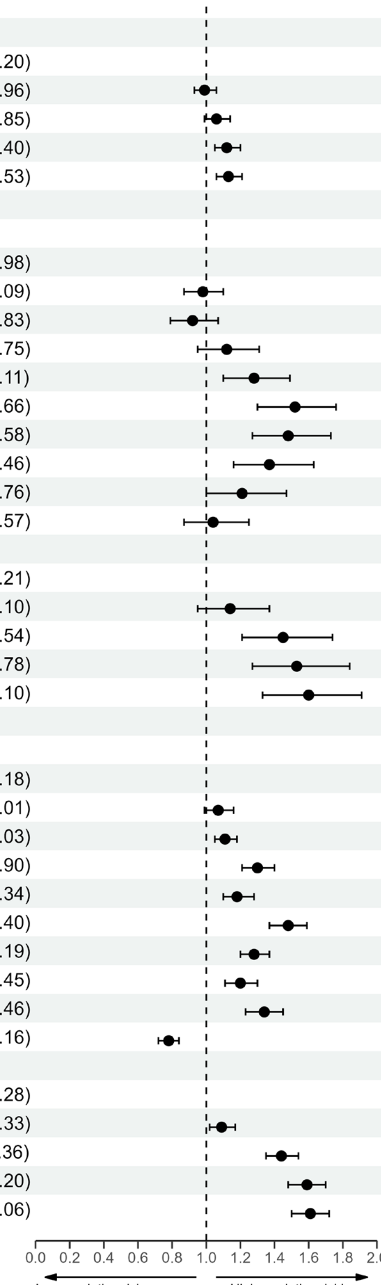

Figure 5 Attack rate (AR) for development of COVID-19 in close contacts of index patients according to cumulative exposure duration and time from symptom onset (or positive RT-PCR (reverse transcriptase-PCR) result for the asymptomatic) to exposure. For symptomatic patients, the 'time from onset to exposure' is the time from symptom onset in the index patient to exposure of the close contact; the 'time from onset to diagnosis' is the time from symptom onset in the index patient to diagnosis of the close contact. For asymptomatic patients, the 'time from a positive RT-PCR result to exposure' is the time from a positive RT-PCR result in the index patient to exposure of the close contact; the 'time from a positive RT-PCR result to diagnosis' is the time from a positive RT-PCR result in the index patient to diagnosis of the contact. Error bars (95% CIs) and dots represent AR for COVID-19 transmission to contacts. The vertical dotted line at 1.0 indicates no effect on risk. AR for each day was estimated using a multivariable Poisson regression with adjustment for household contacts, phase of the epidemic, and age, sex and occupation of the index patient.

Previous research reported that COVID-19 was more likely to be transmitted by symptomatic than by asymptomatic infected individuals (AR of contacts: 18% vs 13%).[32] In agreement, we found that symptomatic exposure was associated with a significantly higher AR of close contacts than asymptomatic exposure (aRR=0.93; 95% CI 0.92, 0.95). This result suggests there may be additional secondary benefits of reducing the symptoms or disease severity of infected individuals,[33] such as by vaccination or prompt diagnosis and treatment. Previous reports[16 19] concentrated on the risk of transmission in an earlier SARS-CoV-2 lineage, and only examined a limited number of cases (less than 800) and contacts. In contrast, the outbreak of COVID-19 in Shanghai was caused by the Omicron variant.[34] This led us to perform a large population-based study to investigate the association of the timing of exposure to the Omicron variant with the risk to close contacts. Our findings indicated that the risk of transmission to a close contact was greatest from 1 day before symptom onset to 3 days after symptom onset in the index patient. These results have important implications for understanding the transmission dynamics of COVID-19, and are consistent with other infectivity studies which suggested that viral load may be highest around the time of symptom onset,[35 36] with a gradual decrease in viral shedding at 1 week after symptom onset.[37] Similar to previous research,[38] the transmission risk of asymptomatic index cases was greatest from 3 days before to 3 days after a positive RT-PCR result, with a maximum at day 0 (aRR=1.48; 95% CI 1.37, 1.59). Our observation of a lower risk of transmission from 3 days after symptom onset or 3 days after a positive RT-PCR result during the Shanghai Omicron epidemic has important implications for optimising COVID-19 prevention and control measures, cutting quarantine periods for the close contacts. Our results suggest that contact tracing should focus on individuals who had contact soon before or soon after the onset of symptoms in the index case. This is an important consideration, because surveillance of all possible index cases and close contacts is impossible due to limited available resources.

There were some limitations in our study. First, there may have been some reporting bias if contacts or index patients did not accurately recall the details of their date of onset. Second, the contact times of all contacts were not traced and recorded by health authorities, presumably due to a severe shortage of staff related to the sudden surge of cases. This could have led to bias in the parameter estimates.[39] Third, there may be a proportion of initially diagnosed asymptomatic Omicron infections is in the pre-symptomatic stage.[40] The presence or absence of symptoms depends on self-reporting on admission, which could lead to recall bias, especially in distinguishing between asymptomatic and symptomatic infections. Fourth, we determined the directionality of transmission based on recall by the index patient and the sequence in which the index case and contact first developed symptoms, a widely used procedure.[19 41 42] However, this method could lead to misclassification if an index patient had an unusually long incubation time or the contact had an unusually short incubation time.

CONCLUSIONS

Our study of the relationships of index patients and their close contacts on the risk for COVID-19 transmission indicated that students had the greatest assortativity and people working age not in the labour force had the least assortativity. Individuals with COVID-19 were most infectious a few days before and a few days after symptom onset (or positive RT-PCR result for the asymptomatic), with a maximum at day 0. The duration of cumulative exposure significantly increased the risk of infection of a close contact. Contact in the household was a significant contributor to the infection of close contacts.

Author affiliations
[1]Department of Epidemiology, Fudan University School of Public Health, Shanghai, China
[2]Department of Infectious Disease Control and Prevention, Shanghai Municipal Center for Disease Control and Prevention, Shanghai, China
[3]Institute of Infectious Disease and Biosecurity, School of Public Health, Fudan University, Shanghai, China
[4]Key Laboratory of Public Health Safety of Ministry of Education, Fudan University, Shanghai, China

Acknowledgements We gratefully acknowledge all index cases and their close contacts in this study.

Contributors BZ and YZ designed the study. BZ, XG, HP, CJ, SM, SL, BJ and DK extracted the data and constructed the database. BZ and YZ analysed the data. BZ, YZ and YY drafted the manuscript—original draft. GZ, HW and WW conducted critical revisions of the manuscript—review and editing. All authors read and approved the final manuscript. WW is responsible for the overall content as guarantor.

Funding This study was granted by the Science and Technology Commission of Shanghai Municipality (Grant No. 20YF1441700, XHG); Shanghai Three-year Action Plan of Strengthen Public Health System (Grant No. GWV-1.1, HYW); the Bill & Melinda Gates Foundation, Seattle, WA (Grant No. INV-006277, WBW); the National Natural Science Foundation of China (Grant No. 82073612, WBW).

Competing interests The authors declare that they have no known competing financial interest or personal relationship that could influence the work reported in this paper.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable. The Shanghai Municipal Health Commission has the authority to collect and use the epidemiological data from patients with COVID-19 as part of the infectious disease outbreak investigation. This study was designed to be fully anonymous and was granted by the Shanghai Municipal Health Commission; therefore, individual patient consent was not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data are available upon reasonable request. The estimated contact matrices and analyzed code during the current study are available in the https://github.com/wwbgroup/Contact_patterns_COVID-19.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Ye Yao http://orcid.org/0000-0002-0878-8942
Weibing Wang http://orcid.org/0000-0002-4497-5251

REFERENCES

1 Loitz CC, Johansen S, Johnston JC, *et al*. Exploring the Alberta health services COVID-19 provincial public health integrated outbreak prevention, preparedness, management, and response interventions to support congregate and communal living sites in Alberta. *Can J Public Health* 2023;114:218–28.
2 Kotay SM, Tanabe KO, Colosi LM, *et al*. Building-level wastewater surveillance for SARS-Cov-2 in occupied University dormitories as an outbreak forecasting tool: one year case study. *ACS ES T Water* 2022;2:2094–104.
3 Zheng B, Zhu W, Pan J, *et al*. Patterns of human social contact and mask wearing in high-risk groups in China. *Infect Dis Poverty* 2022;11:69.
4 Leung WTM, Meeyai A, Holt HR, *et al*. Social contact patterns relevant for infectious disease transmission in Cambodia. *Sci Rep* 2023;13:5542.
5 Lou L, Zhang L, Guan J, *et al*. Retrospective modeling of the Omicron epidemic in Shanghai, China: exploring the timing and performance of control measures. *Trop Med Infect Dis* 2023;8:39.
6 Juneau C-E, Briand A-S, Collazzo P, *et al*. Effective contact tracing for COVID-19: A systematic review. *Glob Epidemiol* 2023;5:100103.
7 Yang K, Deng J, Tuo X, *et al*. Analysis of the close contact management mode and Epidemiological characteristics of COVID-19 in Chengdu, China. *One Health* 2022;15:100420.
8 Ringa N, Iyaniwura SA, David S, *et al*. Social contacts and transmission of COVID-19 in British Columbia, Canada. *Front Public Health* 2022;10:867425.
9 Mistry D, Litvinova M, Pastore y Piontti A, *et al*. Inferring high-resolution human mixing patterns for disease modeling. *Nat Commun* 2021;12.
10 Liang Y, Peng C, You Q, *et al*. Estimating changes in contact patterns in China over the first year of the COVID-19 pandemic: implications for SARS-Cov-2 spread — four cities. *China CDC Wkly* 2023;5:113–9.
11 Koltai J, Vásárhelyi O, Röst G, *et al*. Reconstructing social mixing patterns via weighted contact Matrices from online and representative surveys. *Sci Rep* 2022;12:4690.
12 McAloon CG, Wall P, Butler F, *et al*. Numbers of close contacts of individuals infected with SARS-Cov-2 and their association with government intervention strategies. *BMC Public Health* 2021;21:2238.
13 Yang K, Deng J, Wang L, *et al*. Tracing management and Epidemiological characteristics of COVID-19 close contacts in cities around Chengdu, China. *Front Public Health* 2021;9:645798.
14 Hakki S, Zhou J, Jonnerby J, *et al*. Onset and window of SARS-Cov-2 Infectiousness and temporal correlation with symptom onset: a prospective, longitudinal, community cohort study. *Lancet Respir Med* 2022;10:1061–73.
15 Guo Z, Zhao S, Mok CKP, *et al*. Comparing the incubation period, serial interval, and Infectiousness profile between SARS-Cov-2 Omicron and Delta variants. *J Med Virol* 2023;95:e28648.
16 Ge Y, Martinez L, Sun S, *et al*. COVID-19 transmission Dynamics among close contacts of index patients with COVID-19: A population-based cohort study in Zhejiang province, China. *JAMA Intern Med* 2021;181:1343–50.
17 Wu Y, Kang L, Guo Z, *et al*. Incubation period of COVID-19 caused by unique SARS-Cov-2 strains: A systematic review and meta-analysis. *JAMA Netw Open* 2022;5:e2228008.
18 Chen Z, Deng X, Fang L, *et al*. Epidemiological characteristics and transmission Dynamics of the outbreak caused by the SARS-Cov-2 Omicron variant in Shanghai, China: A descriptive study. *Lancet Reg Health West Pac* 2022;29:100592.
19 Cheng H-Y, Jian S-W, Liu D-P, *et al*. Contact tracing assessment of COVID-19 transmission Dynamics in Taiwan and risk at different exposure periods before and after symptom onset. *JAMA Intern Med* 2020;180:1156–63.
20 Keeling MJ, Rohani P. Modeling Infectious Diseases in Humans and Animals. Princeton University Press, 2008.
21 Zhang J, Klepac P, Read JM, *et al*. Patterns of human social contact and contact with animals in Shanghai, China. *Sci Rep* 2019;9.
22 Gebuis E, Vieyra B, Slegtenhorst R, *et al*. COVID-19 outbreaks among crew on commercial ships at the port of Rotterdam, the Netherlands, 2020 to 2021. *Euro Surveill* 2023;28:16.
23 R: A language and environment for statistical computing [program]. Vienna, Austria: R Foundation for Statistical Computing, 2021.
24 Prem K, Cook AR, Jit M. Projecting social contact Matrices in 152 countries using contact surveys and demographic data. *PLoS Comput Biol* 2017;13:e1005697.
25 Kadelka C. Projecting social contact Matrices to populations stratified by binary attributes with known Homophily. *Math Biosci Eng* 2023;20:3282–300.
26 Hâncean M-G, Lerner J, Perc M, *et al*. Occupations and their impact on the spreading of COVID-19 in urban communities. *Sci Rep* 2022;12:14115.
27 Zhao Y, O'Dell S, Yang X, *et al*. Quantifying human mixing patterns in Chinese provinces outside Hubei after the 2020 Lockdown was lifted. *BMC Infect Dis* 2022;22:483.
28 Mousa A, Winskill P, Watson OJ, *et al*. Social contact patterns and implications for infectious disease transmission – a systematic review and meta-analysis of contact surveys. *Elife* 2021;10:e70294.
29 Madewell ZJ, Yang Y, Longini IM, *et al*. Household transmission of SARS-Cov-2: A systematic review and meta-analysis. *JAMA Netw Open* 2020;3:e2031756.
30 Han K, Francis MR, Xia A, *et al*. Influenza vaccination uptake and its determinants during the 2019-2020 and early 2020-2021 flu seasons among migrants in Shanghai, China: a cross-sectional survey. *Hum Vaccin Immunother* 2022;18:1–8.
31 Eriksson K, Dickins TE, Strimling P. Global sex differences in hygiene norms and their relation to sex equality. *PLOS Glob Public Health* 2022;2:e0000591.
32 Methi F, Madslien EH. Lower Transmissibility of SARS-Cov-2 among asymptomatic cases: evidence from contact tracing data in Oslo, Norway. *BMC Med* 2022;20:427.
33 Baden LR, El Sahly HM, Essink B, *et al*. Efficacy and safety of the mRNA-1273 SARS-Cov-2 vaccine. *N Engl J Med* 2021;384:403–16.
34 Chen X, Yan X, Sun K, *et al*. Estimation of disease burden and clinical severity of COVID-19 caused by Omicron BA.2 in Shanghai, February-June 2022. *Emerg Microbes Infect* 2022;11:2800–7.
35 Puhach O, Meyer B, Eckerle I. SARS-Cov-2 viral load and shedding Kinetics. *Nat Rev Microbiol* 2023;21:147–61.
36 Boucau J, Marino C, Regan J, *et al*. Duration of shedding of Culturable virus in SARS-Cov-2 Omicron (BA.1) infection. *N Engl J Med* 2022;387:275–7.
37 Cevik M, Tate M, Lloyd O, *et al*. SARS-Cov-2, SARS-Cov, and MERS-Cov viral load Dynamics, duration of viral shedding, and Infectiousness: a systematic review and meta-analysis. *Lancet Microbe* 2021;2:e13–22.
38 Fu Y, Li Y, Guo E, *et al*. Dynamics and correlation among viral positivity, seroconversion, and disease severity in COVID-19: A retrospective study. *Ann Intern Med* 2021;174:453–61.
39 Mathur MB. The M-value: A simple sensitivity analysis for bias due to missing data in treatment effect estimates. *Am J Epidemiol* 2023;192:612–20.
40 Kang W, Yang P, Dang B, *et al*. Dynamics of disease characteristics and viral RNA decay in patients with asymptomatic and mild infections during the Omicron wave in Shanghai, China: A retrospective cohort study. *Int J Infect Dis* 2023;130:60–70.
41 Sun K, Wang W, Gao L, *et al*. Transmission Heterogeneities, Kinetics, and Controllability of SARS-Cov-2. *Science* 2021;371:eabe2424.
42 Bi Q, Wu Y, Mei S, *et al*. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect Dis* 2020;20:911–9.