Secondary transmission of SARS-CoV-2 during the first two waves in Japan: Demographic characteristics and overdispersion

Yura K Ko [1,2], Yuki Furuse [3,4], Kota Ninomiya [5,6], Kanako Otani [1], Hiroki Akaba [2], Reiko Miyahara [1], Tadatsugu Imamura [7,8], Takeaki Imamura [2], Alex R Cook [9], Mayuko Saito [2], Motoi Suzuki [1], Hitoshi Oshitani [2,*]

[1] *Center for Surveillance, Immunization, and Epidemiologic Research, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo, Japan*
[2] *Department of Virology, Tohoku University Graduate School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai, Miyagi, Japan 980-8575*
[3] *Institute for Frontier Life and Medical Sciences, Kyoto University, 53 kawaramachi, Shogoin, Sakyo-ku, Kyoto, Japan*
[4] *Nagasaki University Graduate School of Biomedical Sciences, 1-12-4 Sakamoto, Nagasaki, Japan*
[5] *Graduate School of Pharmaceutical Sciences, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[6] *National Institute of Public Health, 2-3-6 Minami, Wako-shi, Saitama 351-0197 Japan*
[7] *Japan International Cooperation Agency, 5-25 Niban-cho, Chiyoda-ku, Tokyo 102-8012, Japan*
[8] *Center for Postgraduate Education and Training, National Center for Child Health and Development, 2-10-1 Okura, Setagaya-ku, Tokyo, Japan*
[9] *Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, 12 Science Drive 2, #10-01, Singapore*

ARTICLE INFO

ABSTRACT

Objectives: Super-spreading events caused by overdispersed secondary transmission are crucial in the transmission of COVID-19. However, the exact level of overdispersion, demographics, and other factors associated with secondary transmission remain elusive. In this study, we aimed to elucidate the frequency and patterns of secondary transmission of SARS-CoV-2 in Japan.

Methods: We analyzed 16,471 cases between January 2020 and August 2020. We generated the number of secondary cases distribution and estimated the dispersion parameter (k) by fitting the negative binomial distribution in each phase. The frequencies of the secondary transmission were compared by demographic and clinical characteristics, calculating the odds ratio using logistic regression models.

Results: We observed that 76.7% of the primary cases did not generate secondary cases with an estimated dispersion parameter k of 0.23. The demographic patterns of primary-secondary cases differed between phases, with 20–69 years being the predominant age group. There were higher proportions of secondary transmissions among older individuals, symptomatic patients, and patients with 2 days or more between onset and confirmation.

Conclusions: The study showed the estimation of the frequency of secondary transmission of SARS-CoV-2 and the characteristics of people who generated the secondary transmission.

© 2022 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Introduction

Epidemiologic findings on COVID-19 have accumulated at an unprecedented rate. A key finding is the heterogeneity in the num-

ber of secondary transmissions, which is often characterized as overdispersion. Super-spreading events are essential factors contributing to the sustained transmission of SARS-CoV-2 (Endo et al., 2020; Xu et al., 2020) similar to severe acute respiratory syndrome (SARS) (Leo et al., 2003) and other infectious diseases (Kucharski and Althaus, 2015; Lau et al., 2017; Wong et al., 2015). A study in Hong Kong revealed that only 19% of infected persons with COVID-19 generated 80% of all transmissions. Contrastingly, 69% of cases did not lead to secondary transmissions, indicating substantial transmission heterogeneity (Adam et al., 2020). In addition, a study in Korea reported temporal changes in the level of transmission overdispersion during different epidemic periods (Lim et al., 2021). A study conducted in Japan at the beginning of the COVID-19 outbreak, which examined 110 cases, including 11

* Corresponding author. Hitoshi Oshitani, MD, MPH, PhD, Department of Virology, Tohoku University Graduate School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai, Miyagi, Japan 980-8575. Telephone number: +81-22-717-8211, FAX number: +81-22-717-8212.

E-mail addresses: yurako0603@gmail.com (Y.K. Ko), yukifuruse@gmail.com (Y. Furuse), kninomiya-mope@g.ecc.u-tokyo.ac.jp (K. Ninomiya), otanik@niid.go.jp (K. Otani), zh44893@gmail.com (H. Akaba), rmiyahara@niid.go.jp (R. Miyahara), imamura-t@ncchd.go.jp (T. Imamura), timamura@med.tohoku.ac.jp (T. Imamura), ephcar@nus.edu.sg (A.R. Cook), msaitop@med.tohoku.ac.jp (M. Saito), mosuzuki@niid.go.jp (M. Suzuki), oshitanih@med.tohoku.ac.jp (H. Oshitani).

clusters, identified that most infected individuals did not generate any secondary transmission (Nishiura et al., 2020b). The characteristics of clusters at the early stage of the outbreak were also analyzed in Japan (Furuse et al., 2020b). On the basis of these findings, Japan developed a cluster-based approach to COVID-19 management that focuses on identifying and preventing clusters (superspreading events) to suppress transmission (Oshitani, 2020).

COVID-19 vaccinations are rapidly progressing in many countries, and real-world data indicate that the impact of COVID-19 could be significantly reduced by vaccination (Haas et al., 2021). However, reduced vaccine efficacy against the delta variant (Lopez Bernal et al., 2021) and waning of vaccine immunity resulting in breakthrough infections have recently been reported (Rosenberg et al., 2021). Furthermore, vaccine coverage remains low in many countries, especially in the low- and middle-income countries (Ritchie et al., 2021). Therefore, developing more effective public health measures to suppress transmission is critical. To achieve this, the exact level of overdispersion and the demographic characteristics that are likely to generate secondary transmissions need to be defined. It is also necessary to identify whether the level of overdispersion changes during the different phases of COVID-19 transmission.

This study aimed to analyze the frequency and demographic patterns of secondary transmission of SARS-CoV-2 during the first 2 waves in Japan using the data of individual cases available from local governments in Japan. We also analyzed the characteristics of cases that generated secondary transmissions to define the factors associated with secondary transmission.

Methods

Data collection

The Japanese government included COVID-19 as a designated infectious disease on January 2020. As a consequence, physicians have a legal mandate to report all confirmed cases to local public health centers (Furuse et al., 2020a). Public health nurses in public health centers interview confirmed cases on the basis of the National Institute of Infectious Diseases guidelines (Imamura et al., 2021a). They collect demographic data, clinical information, history of high-risk activities such as traveling to affected areas, contact with confirmed COVID-19 cases, and visits to high-risk venues using standardized forms. These data are collected for the 14 days before the onset of symptoms. Local governments compile the data on all confirmed cases, and most release this information online on a daily basis. The data are different between local governments, but most local governments include age, gender, identification number, and epidemiologic link. We collected and collated these data for the analysis.

Definition of phases

We analyzed the data of confirmed cases between January 15, 2020 and August 31, 2020. This period included 2 waves, which were divided into 4 phases on the basis of the epidemic curve, namely the increasing (phase 1: January 15 to April 13) and decreasing (phase 2: April 14 to May 24) phases of the first wave and the increasing (phase 3: May 25– to August 6) and decreasing (phase 4: August 7 to August 31) phases of the second wave (Figure 1). The levels of interventions differed between the phases. During phase 1, the government implemented limited control measures, such as school closures and cancellations of mass gathering events. In response to the surge in cases, the government declared

a state of emergency for 7 prefectures, including Tokyo and Osaka, on April 7, 2020, and extended it to all prefectures on April 16, 2020. During the state of emergency, the government requested people to stay at home unless necessary, and people's compliance was high during this period until the state of emergency was lifted from all prefectures on May 25, 2020 (phase 2). The number of newly confirmed cases dropped in May; however, the transmission of infection continued in nightlife areas in metropolitan cities such as Tokyo. The number of cases increased in June and July until it peaked on August 7, 2020 (phase 3) (Nagata et al., 2021). Although the government did not declare another state of emergency at that time, specific interventions targeting nightlife areas were implemented, and newly confirmed cases decreased until August 31, 2020 (phase 4).

Definition of primary and secondary cases

We defined the primary and secondary cases as follows: a primary case had the earliest date of onset among epidemiologically linked cases if the duration between the onset dates of primary and secondary cases was less than 15 days based on the serial interval reported in previous studies (Nishiura et al., 2020a). If the onset date was unavailable or the case had no symptoms, we used the confirmation date or reporting date instead of the onset date. If 2 or more cases had the same earliest onset date, we defined the case with an earlier confirmed date as the primary case. For clusters comprising several cases with primary exposure reported at a common event or venue, we considered the case that had the earliest onset date as a primary case and other cases as secondary cases if the onset date of the cases was ≤7 days after the onset of the primary case. If the onset date was unknown or the cases in a cluster were asymptomatic, we regarded those cases as secondary cases if the confirmed date was ≤7 days after the confirmation of the primary case.

Inclusion and exclusion criteria

We analyzed the frequency of secondary transmission in primary cases reported between January 15, 2020, and August 31, 2020. However, we included those cases identified as secondary cases in September 2020 if their corresponding primary cases occurred by August 31, 2020.

We only included primary cases with no identified exposure to calculate the frequency of secondary transmission for 2 reasons. First, it was difficult to identify the primary and secondary cases for the cases in a cluster that could have had more than 1 generation of transmission. Second, individuals with a contact history with the previously identified cases were likely to have practiced precautionary measures, such as self-quarantine, before testing positive for COVID-19.

Although there was a certain degree of consistency in the data released by the local governments, some variations were observed, especially regarding the contact history, which was primarily a result of privacy concerns. Thus, restrictions on prefectures and phases were imposed on the basis of contact history reporting details. Specifically, we used 2 criteria to evaluate the level of the released contact history: 1) the proportion of cases with an identified source of infection (cut-off 1) and 2) the proportion of cases with a disclosed source of infection among cases with an identified source of infection (cut-off 2). We used ≥25% for cut-off 1 and ≥75% for cut-off 2 for the principal analysis (Supplementary Figure 1). We also analyzed the data using different cut-off values for the sensitivity analysis. We also excluded cases of international arrivals to avoid underestimating the secondary transmission fre-
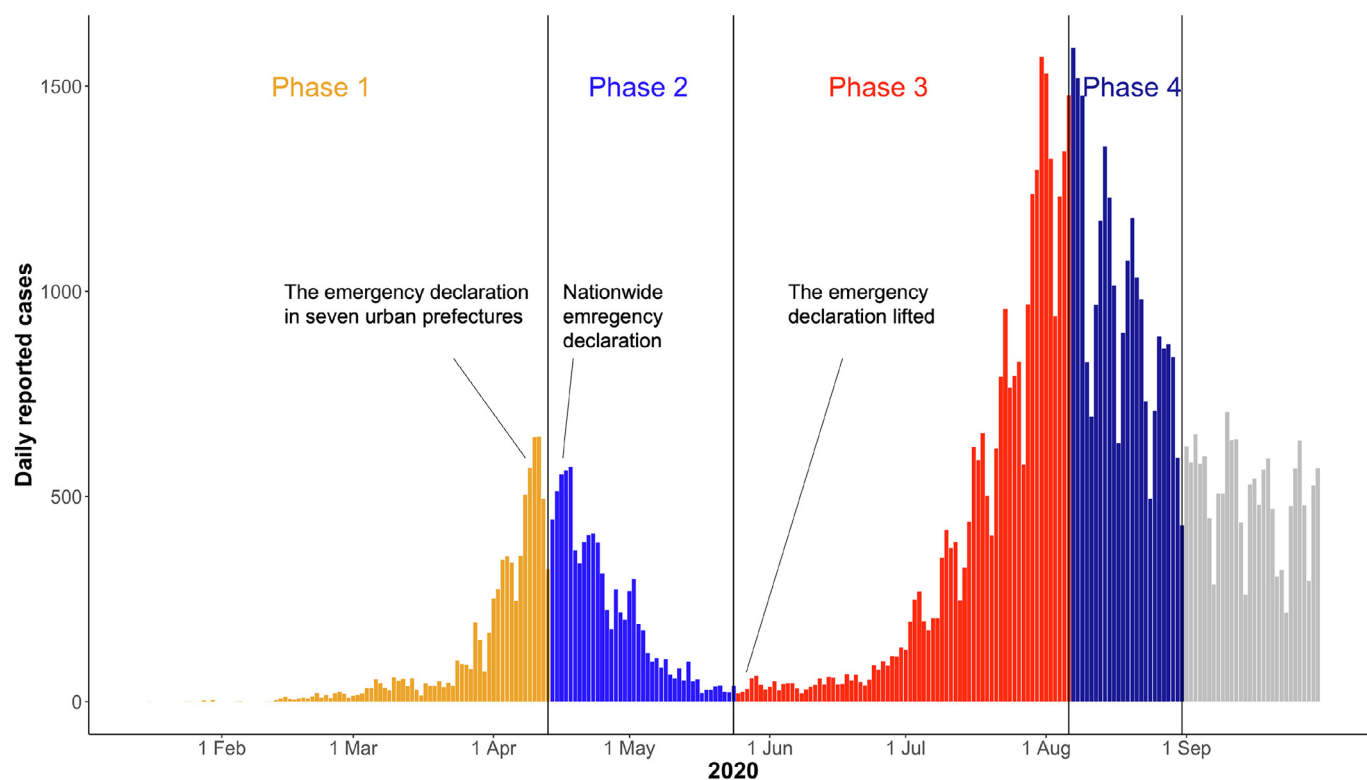
Figure 1. The epidemiologic curve of COVID-19 cases in Japan by reported date between January 2020 and August 2020. The whole period was divided into 4 phases (phases 1–4) on the basis of the epidemic trend.

quency, considering the requirement of 14-day quarantine after arrival.

Statistical analysis

We analyzed the observed distribution of secondary cases by calculating the number of secondary cases per primary case by fitting a negative binomial distribution by maximum likelihood. We calculated the reproductive number (R) from the mean of the negative binomial distribution fit to the observed distribution of secondary cases and the degree of transmission heterogeneity from the corresponding dispersion parameter (k) as described elsewhere (Lloyd-Smith et al., 2005). Furthermore, we employed Markov chain Monte Carlo method to obtain the joint density estimates and marginal density estimates for R and k.

We analyzed the number of secondary cases stratified by phase, age, gender, presence of symptoms at confirmation, and days from onset to confirmation. We calculated the odds ratios (ORs) by 1) comparing the proportion of cases that caused secondary cases with that of cases that did not cause secondary transmission and 2) comparing the proportion of cases that caused 1–4 secondary cases with that of cases that generated ≥5 secondary cases by a logistic regression model. The cut-off was ≥5 because superspreading events are generally considered to be transmission to more than 4–6 people. In each stratified group, the group with the largest number of cases was selected as the reference group, such as the group that belonged to phase 3, group of those aged 20–29 years, group comprising men, group of those who were symptomatic at confirmation, and group of those had 2 days between onset and confirmation. All analyses were performed using R (R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/).

Results

Overdispersion of secondary transmission

We collected data on 67,761 confirmed COVID-19 cases in Japan between January 15, 2020, and August 31, 2020, from local government websites, including demographic and epidemiologic information. Among them, 46,481 cases had no identified exposure. From these 46,481 cases, we excluded the 30,010 cases reported by prefectures during the specific phases in which the contact history information was limited. The remaining 16,471 cases were included as primary cases for the principal analysis of secondary transmission. We found that 76.7% (12,638/16,471) of the cases did not generate secondary transmissions (Figure 2). Contrastingly, 21.9% (3,604/16,471) of the cases generated 1–4 secondary transmissions, and 1.4% (229/16,471) generated ≥5 secondary transmissions. We fitted the negative binomial distribution and obtained an R of 0.47 (95% CI, 0.45–0.49) and k of 0.23 (95% CI, 0.22–0.25). The expected proportion of cases accounting for 80% of all COVID-19 transmission was 13.3% (95% CI, 12.8%–13.9%). While k was low for all age groups throughout the entire period, R varied for each age group in each phase. Particularly, the R for those aged 0–19 years was low in phases 1 and 2 (0.24 and 0.20, respectively) before increasing in phase 3 (0.45). In individuals aged ≥70 years, R fell from phase 1 to phase 2 (from 0.83 to 0.39) and increased from phase 3 to phase 4 (from 0.49 to 0.70) (Table 1). When we excluded cases associated with health care and other facilities, k increased in all phases except for phase 3, and the R decreased, particularly in those aged ≥70 years. However, the change was minimal in phase 3, suggesting a limited contribution of health care and other facilities in phase 3. The overall patterns of R and k in each phase were the same after excluding cases associated with health care and other facilities (Supplementary Table 2). For the sensitivity analysis, we
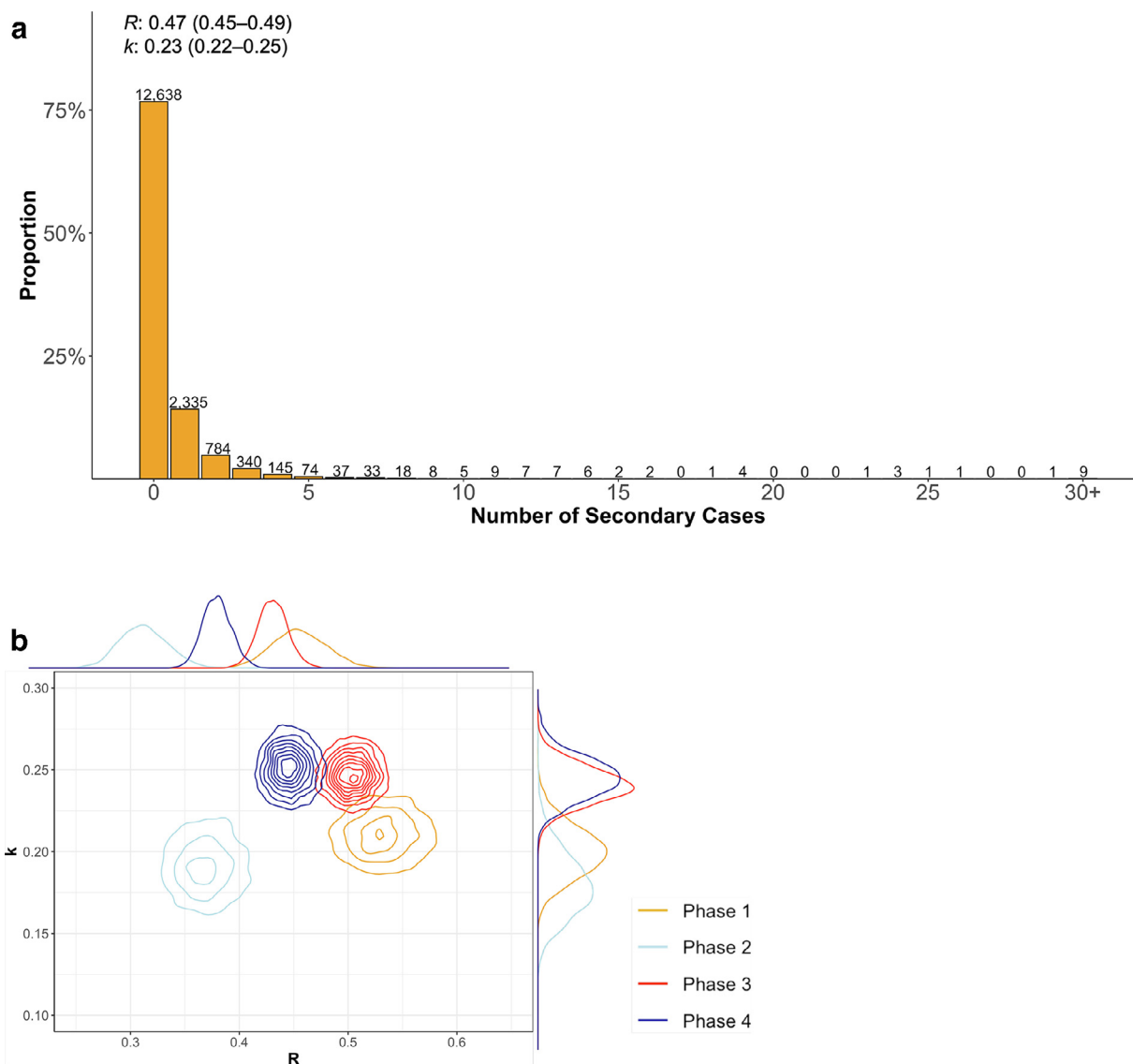
Figure 2. a) The observed distribution of the number of secondary cases from total 16,471 primary cases for the entire period. The number above each bar indicates the absolute number of secondary cases. Estimated reproductive number (*R*) and overdispersion parameter (*k*) by fitting the negative binomial distribution are also shown. b) Joint estimates of *k* and *R* in each phase. The posterior distributions of each parameter are plotted in the outer margin of each axis.

calculated *k* using different inclusion criteria and found that it was not higher than 0.5 in all phases under different criteria (Supplementary Table 3).

Changes in demographic patterns of primary and secondary cases

The demographic characteristics of the primary and secondary cases differed significantly by phase (Figure 3). In phase 1, the most common transmissions were from males aged 50–59 years to females in the same age group and between males aged 20–29 years. In phase 2, the most common transmissions were also from males aged 50–59 years to females in the same age group; however, transmissions from males aged 60–69 years to females in the same age group increased. In phase 3, transmission between males aged 20–29 years was significant. In phase 4, the proportion of transmissions from various age groups, especially from the older adults corresponding to both genders, increased. When age and gender were analyzed separately, most of the primary cases were identified between those aged 20–39 years and those aged 40–69 years. However, the proportion of patients aged 40–69 years

was the highest in phases 1 and 2 (55.0% and 58.0%, respectively), whereas the proportion of patients aged 20–39 years was the highest in phase 3 (57.4%). In phase 4, the proportions of those aged 20–39 years and 40–69 years were approximately the same (38.2% and 38.3%, respectively). Regarding gender, males were more likely to be primary cases throughout the entire period (60.7%) (Supplementary Tables 4 and 5).

Characteristics of cases generating secondary transmissions

Compared with phase 3, the proportion of cases generating secondary transmissions was lower in other phases, whereas there was no significant difference in the proportion of cases generating ≥5 transmissions between phases 1 and 3 (Figure 4, Supplement Table 7). Those aged 40–49 to 80–89 years were more likely to cause a secondary transmission than those aged 20–29 years. However, there was no significant difference in transmission to 5 or more people between the age groups. Children aged 0–9 years were less likely to generate secondary transmissions but were more likely to generate ≥5 secondary transmissions; how-

Table 1

The estimated effective reproductive number (R), overdispersion parameter (k), proportion of infectious cases responsible for 80% of all cases, and proportion of cases that did not generate any secondary transmissions stratified by the phase and age group.

		N	R (95%CI)	k (95%CI)	*Prop80% (95%CI)	**Prop0% (95%CI)
Whole Period	All	16471	0.47 (0.45–0.49)	0.23 (0.22–0.25)	13.3% (12.8%–13.9%)	77.2% (76.3%–78.1%)
	0-19	847	0.36 (0.30–0.43)	0.22 (0.17–0.29)	12.1% (9.9%–14.6%)	80.7% (76.8%–84%)
	20-39	7647	0.41 (0.38–0.43)	0.21 (0.19–0.23)	12.1% (11.4%–12.9%)	79.7% (78.4%–80.9%)
	40-69	5832	0.52 (0.49–0.55)	0.29 (0.26–0.32)	15.4% (14.4%–16.3%)	74.1% (72.6%–75.8%)
	70-	1996	0.62 (0.55–0.69)	0.21 (0.18–0.24)	13% (11.6%–14.5%)	74.9% (72.1%–77.5%)
Phase1	All	2415	0.53 (0.48–0.58)	0.21 (0.18–0.24)	12.6% (11.4%–13.9%)	76.9% (74.4%–79.1%)
	0-19	62	0.24 (0.11–0.45)	0.42 (0.12–*Inf*)	12.6% (5.4%–27.3%)	82.6% (63.7%–92.4%)
	20-39	762	0.35 (0.29–0.42)	0.24 (0.18–0.33)	12.4% (10%–15.3%)	80.6% (76.3%–84.2%)
	40-69	1233	0.57 (0.50–0.65)	0.22 (0.19–0.27)	13.3% (11.6%–15.3%)	75.3% (71.6%–78.4%)
	70-	343	0.83 (0.62–1.06)	0.20 (0.15–0.28)	13.9% (10.7%–17.6%)	72% (64.7%–78.3%)
Phase2	All	1738	0.37 (0.32–0.42)	0.19 (0.16–0.23)	11.2% (9.7%–12.9%)	81.5% (78.8%–83.9%)
	0-19	44	0.2 (0.07–0.43)	0.53 (0.08–*Inf*)	11.8% (3.5%–26.4%)	84.1% (64.9%–95.1%)
	20-39	471	0.24 (0.17–0.31)	0.16 (0.10–0.24)	8.8% (6.2%–11.9%)	86.4% (81.9%–90.4%)
	40-69	823	0.44 (0.36–0.53)	0.20 (0.15–0.26)	11.9% (9.7%–14.5%)	79.2% (74.8%–83%)
	70-	364	0.39 (0.29–0.5)	0.24 (0.16–0.36)	12.9% (9.5%–16.9%)	79.3% (73.1%–84.7%)
Phase3	All	6397	0.50 (0.47–0.53)	0.25 (0.23–0.27)	13.9% (13.1%–14.9%)	76% (74.5%–77.4%)
	0-19	443	0.45 (0.36–0.57)	0.26 (0.18–0.38)	14% (10.7%–18.1%)	76.9% (70.5%–82.1%)
	20-39	3795	0.47 (0.44–0.51)	0.21 (0.19–0.24)	12.5% (11.5%–13.6%)	78% (76.1%–79.8%)
	40-69	1625	0.58 (0.52–0.65)	0.32 (0.27–0.38)	16.5% (14.9%–18.5%)	71.8% (68.6%–74.6%)
	70-	455	0.49 (0.39–0.59)	0.34 (0.25–0.5)	16.4% (13%–20.5%)	73.8% (67.6%–79.1%)
Phase4	All	5921	0.44 (0.42–0.48)	0.25 (0.23–0.28)	13.7% (12.8%–14.7%)	77.4% (75.8%–78.9%)
	0-19	298	0.27 (0.17–0.38)	0.16 (0.09–0.28)	9.1% (5.9%–13.6%)	85.5% (78.9%–90.7%)
	20-39	2619	0.36 (0.32–0.39)	0.23 (0.19–0.27)	12.1% (10.8%–13.6%)	80.7% (78.5%–82.7%)
	40-69	2151	0.48 (0.44–0.53)	0.39 (0.33–0.46)	17.3% (15.6%–19.2%)	73.1% (70.3%–75.7%)
	70-	834	0.70 (0.58–0.82)	0.19 (0.16–0.24)	12.9% (10.9%–15.2%)	74.6% (70.1%–78.5%)

* Prop80% is the proportion of cases responsible for 80% of all cases.

** Prop0% is the proportion of cases that did not generate any secondary transmissions.

ever, this finding was not statistically significant. If cases had no symptoms at confirmation, they were less likely to generate secondary transmissions. Cases diagnosed within 2 days from onset were significantly less likely to generate secondary transmissions, although the likelihood of generating ≥5 secondary transmissions was lower in those diagnosed later. We did not find any gender-related differences in the frequency of generating secondary transmission.

Discussion

We found that just over three-quarters of the included cases did not generate any secondary transmissions during the entire study period (Table 1), which is in line with previous reports from other countries (Adam et al., 2020; Laxminarayan et al., 2020; Sun et al., 2021). All sensitivity analyses showed relatively low values of k (<0.5) in all phases, suggesting a high possibility of extinction for most transmission chains (Lloyd-Smith et al., 2005). In contrast, a study in Korea showed that k was larger in the later epidemic period (Lim et al., 2021). There are 2 possible reasons for the difference between Japan and Korea. First, in Japan, cluster-based approach was implemented from the beginning and reflected k in all phases. Second, in Korea, religious gatherings, which may have contributed to the overdispersion in the early stages, were banned in later phases. We believe that Japan's cluster-based approach focusing on super-spreading can suppress transmission because most infected individuals do not contribute to transmission, and sustained transmission is unlikely to occur without super-spreading events (Endo et al., 2020) (Sneppen et al., 2021).

Analysis showed that R was estimated to be <1 for all periods, even during the increase phase (phase 1 and 3). This is inconsistent with the effective reproduction number estimated from the epidemic curve because our estimation indicates how many secondary transmissions were observed on average from the primary cases with no identified exposure.

Our data indicated that COVID-19 transmission in Japan during the first 2 waves was driven mainly by individuals aged 20–69 years; 80% of primary cases causing secondary transmission belonged to these age groups. However, the predominant age group comprising primary cases shifted from middle-aged adults (40–69 years) in phase 1 to younger adults (20–39 years) in phase 3. Studies in other countries have shown a similar demographic transition (Monod et al., 2021; Oster et al., 2020). Although the proportion of primary cases generating secondary transmissions was highest among those aged 20–39 years, secondary transmission was more common in those aged 40–89 years (Figure 4 and Supplementary Table 6). A higher rate of secondary transmission in older age groups has been reported in other countries, which may be due to the high viral load in these age groups (Hu et al., 2021; Jones et al., 2021; To et al., 2020). In addition, we found that the absence of symptoms at the confirmation was associated with a low secondary transmission rate, as reported in other studies (Buitrago-Garcia et al., 2020; Heavey et al., 2020; Sayampanathan et al., 2021). People in older age groups are more likely to develop typical clinical symptoms, such as fever and cough (Davies et al., 2020). This may be another possible reason for higher secondary transmission in older age groups, as suggested by a modeling study (Chen et al., 2021).

Our data showed that most transmissions occurred between younger age groups in earlier phases of each wave (phases 1 and 3) and that increased transmission between different age groups occurred in later stages (phases 2 and 4). This resulted in more cases among older individuals (≥70 years) during the later stages (phases 2 and 4).

Our data also indicated that children aged 0–9 years comprised only a small proportion of primary cases (4%). However, R for those aged 0–19 years was higher in phases 3 and 4 than in phases 1 and 2, indicating that the transmission from this age group increased during this period. These changes were possibly caused by the reopening of schools in phase 3 after the universal closure of schools during almost all of phases 1 and 2. However, our data showed
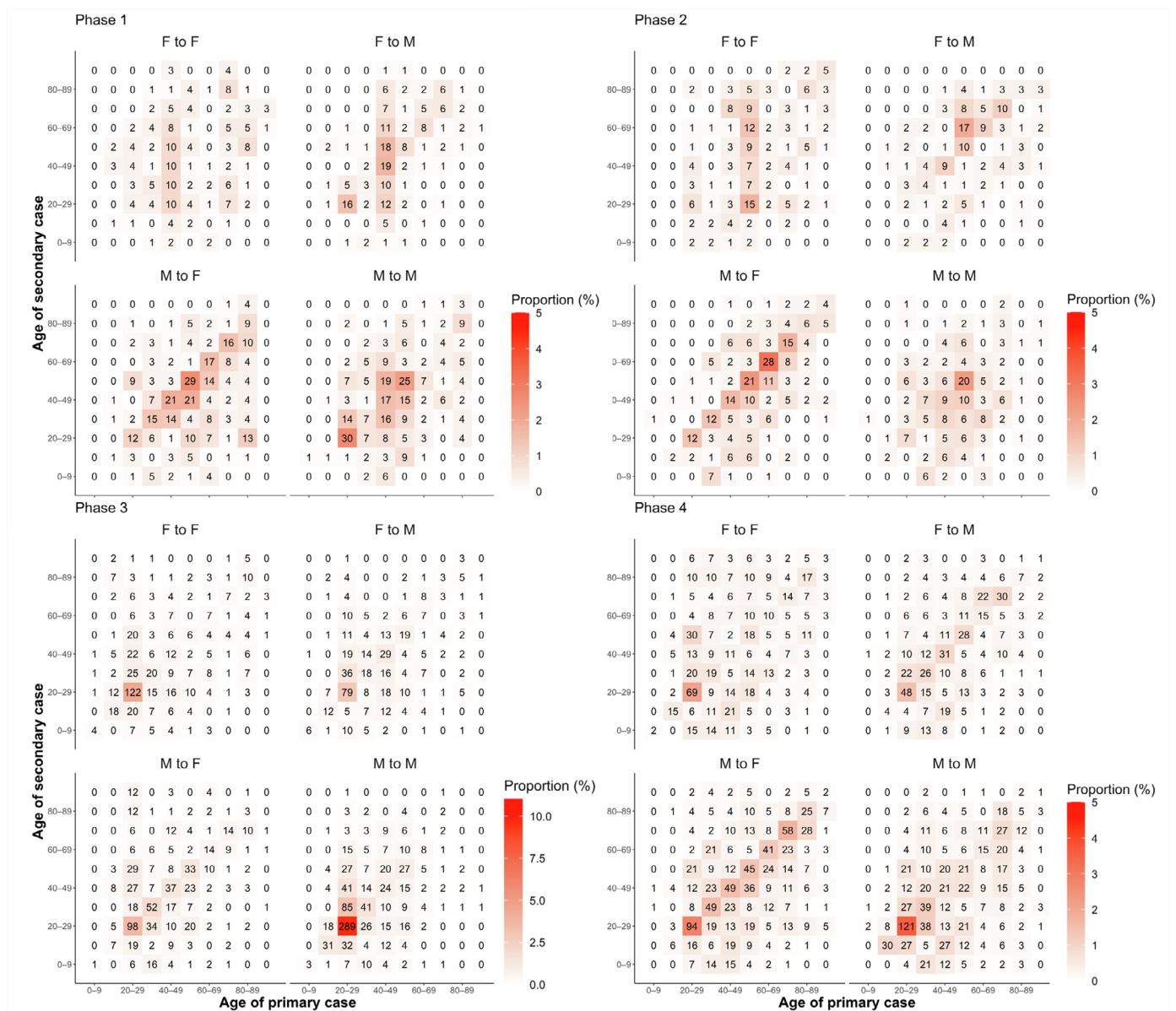
Figure 3. Relationship between primary and secondary cases by age group and gender for each phase. The color scale shows the proportion of observed total transmission pairs in each phase. Note that the color scales are different between phases 1, 2, and 4 and phases 3.
F, female; M, male.

that children did not play a major role in increasing community transmission even after schools reopened. We also found that secondary transmission was less common in children aged 0–9 years. A low transmission rate of COVID-19 in children has been reported in our previous study (Imamura et al., 2021) and other studies (Zhu et al., 2021). These data are consistent with other studies that have not found children to be major drivers of community transmission of SARS-CoV-2 (Davies et al., 2020).

Identifying and monitoring the age group that principally drives transmission is vital to implement more specific interventions. Although our data clearly indicated that individuals aged 20–39 years had a predominant role in transmission during the second wave in Japan, implementing effective measures targeting these young adults is challenging; they are less likely to develop typical or severe symptom (Davies et al., 2020), which may result in underdiagnosis. Notably, some seroepidemiologic studies have shown that children and young adults have greater under-ascertainment

rates (Havers et al., 2020; Pollán et al., 2020; Yoshiyama et al., 2021). Moreover, implementing health promotion strategies to induce behavioral change is particularly challenging in this age group (Kim and Crimmins, 2020), and data from other countries indicate that it is also difficult to achieve high vaccine coverage in this age group (Murphy et al., 2021). Our data also showed that males, particularly those aged 20–59 years, had a more significant role in transmission even though there was no observed difference in the frequency of generating secondary transmission by gender. The dominant role of males in COVID-19 transmission has been reported in other studies (Galasso et al., 2020); biological, social, and behavioral factors likely contribute to this difference (Capraro and Barcelo, 2020). There were more transmissions to the older individuals from younger age groups as well as older-to-older transmission in the later phases of each wave. This is consistent with a previous study that reported community cluster outbreaks among the older adults in Japan (Furuse et al., 2021). Because most severe
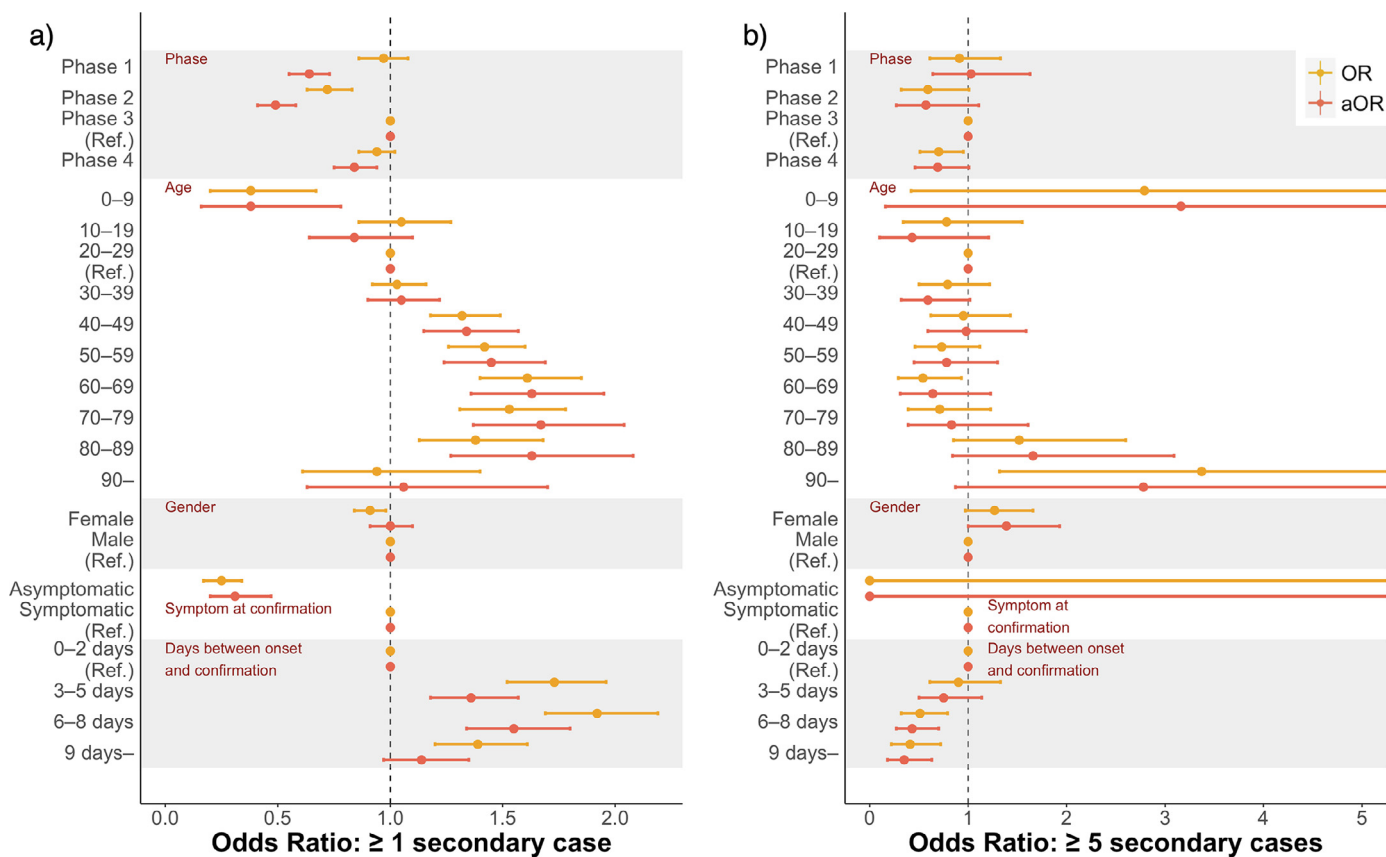
Figure 4. OR and aOR by comparing between a) the proportion of cases that generated ≥1 secondary transmissions and that of cases with no secondary transmission and comparing between b) the proportion of cases that generated ≥5 secondary transmissions and that of cases that generated 1–4 secondary transmissions. ORs were calculated between phases, age group, gender, presence of symptoms, and days from onset to confirmation, with phase 3, those aged 20–29 years, male gender, symptomatic at confirmation, and within 2 days from onset to confirmation as the reference for each category.

OR, odds ratio; aOR, adjusted odds ratio; adjusted for phase, age, presence of symptoms at confirmation, and days between onset and confirmation.

cases occur in the older people (Driscoll et al., 2021), interventions focusing on preventing infection in older age groups are essential to minimize the impact of COVID-19.

When cases were detected within 2 days after illness onset, they were less likely to cause secondary transmission (Figure 4). A previous study showed that infectiousness peaked around illness onset (He et al., 2020) and that the possibility of SARS-CoV-2 isolation decreased a few days after illness onset (Wölfel et al., 2020). Our data indicate that the early detection and isolation of cases are useful in suppressing transmission, as reported previously (Pung et al., 2020). However, delays in case isolation after 5 days of illness onset would not substantially increase the chance of secondary transmission.

Our study has several limitations. First, because we used contact tracing data on public domain, case, contact, and cluster ascertainment might have affected the results. In particular, before May 29, 2020, testing was only done when a close contact presented symptoms; later, testing of all close contacts was recommended regardless of whether they presented symptoms. Therefore, it is possible that transmission pairs were underestimated especially during phase 1 and 2 in our study. In contrast, during phases 3 and 4, with high case numbers, there may have been underestimation owing to the reduction in contact tracing caused by the increased workload of public health centers. Under-reporting generally leads to an underestimation of R but overestimation of k, especially when super-spreading events are missed (Blumberg and Lloyd–Smith, 2013; Endo et al., 2020). Thus, the overdispersion of secondary transmission would remain unchanged even after accounting for the effects of these biases. Second, the data released by the

local governments were inconsistent, particularly the data on close contacts, and might be subject to certain biases depending on the quality of the interviews. We addressed this as much as possible by limiting the cases to those from municipalities where the quality of data on epidemiologic links was considered to maintain a certain level. Still, under-ascertainment of transmission chains was likely to exist. In this case, as mentioned earlier, R is generally considered to be underestimated, and k is overestimated. In addition, we excluded many cases from prefectures without sufficient information. In particular, all cases in Tokyo were excluded; however, other prefectures with large populations, such as Aichi, Osaka, and Fukuoka, and many small prefectures were included (Supplementary Table 1, Supplementary Figure 4), suggesting that our study has some generalizability to all of Japan. Third, primary and secondary transmissions decided by our criteria may not reflect valid infector-infectee pairs. However, if we had used different definitions of primary and secondary transmissions with different inclusion criteria, the results would not have changed substantially. The consistency of our findings regarding the multiple sensitivity analyses strengthens the robustness of this study. Fourth, because our study period included only the first and second waves of the outbreak in Japan, it does not reflect the transmission characteristics of newly emerged variants of SARS-CoV-2, such as the alpha and delta variants. The new variants are now predominant in the community transmission of SARS-CoV-2 in Japan, potentially changing transmission patterns from those observed in this study. A Korean study reported no significant difference in delta compared with pre-delta in terms of the likelihood of super-spreading events (Ryu et al., 2022). However, there is still limited empirical

evidence using country-level data on the delta variant, and future studies are warranted.

In conclusion, this study estimates the frequency of secondary transmission of SARS-CoV-2 in Japan, and its overdispersion was observed in the whole period up to the first and second waves. We also revealed the temporal changes in the predominant population in the transmission chain and the characteristics of people who generated the secondary transmissions; further interventions should be performed on the basis of these characteristics.

Conflict of Interest

The authors declare no competing interests.

Author's Contributions

Y.K.K. and H.O. designed the research. Y.K.K, Y.F., K.N., K.O., H.K., R.M., and T.I. managed and curated the data. Y.K.K. performed the research. Y.K.K., Y.F., and H.O. analyzed the data. T.I., A.R.C., M.S., and M.S. interpreted the data. Y.K.K wrote the first draft of the manuscript, which was revised by all authors. All authors approved the final version of the manuscript.

Ethical Approval statement

This study did not require ethical approval because the data were public, anonymous, and collected as part of the outbreak response.

Funding Source

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijid.2022.01.036.

References

Adam DC, Wu P, Wong JY, Lau EHY, Tsang TK, Cauchemez S, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. Nat Med 2020. doi:10.1038/s41591-020-1092-0.

Blumberg S, Lloyd-Smith JO. Inference of R0 and Transmission Heterogeneity from the Size Distribution of Stuttering Chains. PLoS Comput Biol 2013;9:1–17. doi:10.1371/journal.pcbi.1002993.

Buitrago-Garcia D, Egli-Gany D, Counotte MJ, Hossmann S, Imeri H, Ipekci AM, et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARSCoV-2 infections: A living systematic review and meta-analysis. PLoS Med 2020;17:1–25. doi:10.1371/journal.pmed.1003346.

Capraro V, Barcelo H. The effect of messaging and gender on intentions to wear a face covering to slow down COVID-19 transmission. J Behav Econ Policy 2020. doi:10.31234/osf.io/tg7vz.

Chen PZ, Bobrovitz N, Premji Z, Koopmans M, Fisman DN, Gu FX. Heterogeneity in transmissibility and shedding SARS-CoV-2 via droplets and aerosols. Elife 2021;10:1–32. doi:10.7554/ELIFE.65774.

Davies NG, Klepac P, Liu Y, Prem K, Jit M, Eggo RM. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nat Med 2020;26. doi:10.1038/s41591-020-0962-9.

Driscoll MO, Ribeiro G, Santos D, Wang L, Cummings DAT, Azman AS, et al. Age-specific mortality and immunity patterns of SARS-CoV-2. Nature 2021;590. doi:10.1038/s41586-020-2918-0.

Endo A, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Res 2020;5:67. doi:10.12688/wellcomeopenres.15842.3.

Furuse Y, K Ko Y, Saito M, Shobugawa Y, Jindai K, Saito T, et al. Epidemiology of COVID-19 Outbreak in Japan, January–March 2020. Jpn J Infect Dis 2020a JJID.2020.271. doi:10.7883/yoken.JJID.2020.271.

Furuse Y, Sando E, Tsuchiya N, Miyahara R, Yasuda I, Ko YK, et al. Clusters of Coronavirus Disease in Communities, Japan, January–April 2020. Emerg Infect Dis J 2020b;26. doi:10.3201/eid2609.202272.

Furuse Y, Tsuchiya N, Miyahara R. COVID-19 case-clusters and transmission chains in the communities in Japan. J Infect 2021. doi:10.1016/j.jinf.2021.08.016.

Galasso V, Pons V, Profeta P, Becher M, Brouard S. Gender differences in COVID-19 attitudes and behavior : Panel evidence from eight countries. Proc Natl Acad Sci U S A 2020;117. doi:10.1073/pnas.2012520117.

Haas EJ, Angulo FJ, McLaughlin JM, Anis E, Singer SR, Khan F, et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. Lancet 2021;397:1819–29. doi:10.1016/S0140-6736(21)00947-8.

Havers FP, Reed C, Lim T, Montgomery JM, Klena JD, Hall AJ, et al. Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020. JAMA Intern Med 2020;180:1776–86. doi:10.1001/jamainternmed.2020.4130.

He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 2020;26:672–5. doi:10.1038/s41591-020-0869-5.

Heavey L, Casey G, Kelly C, Kelly D, McDarby G. No evidence of secondary transmission of COVID-19 from children attending school in Ireland, 2020. Eurosurveillance 2020;25:1–4. doi:10.2807/1560-7917.ES.2020.25.21.2000903.

Hu S, Wang W, Wang Y, Litvinova M, Luo K, Ren L, et al. Infectivity, susceptibility, and risk factors associated with SARS-CoV-2 transmission under intensive contact tracing in Hunan, China. Nat Commun 2021;12:1–11. doi:10.1038/s41467-021-21710-6.

Imamura T, Saito M, Ko YK, et al. Roles of children and adolescents in COVID-19 transmission in the community: A retrospective analysis of nationwide data in Japan. Front Pediatr 2021. doi:10.3389/fped.2021.705882.

Imamura T, Saito T, Oshitani H. Roles of Public Health Centers and Cluster-Based Approach for COVID- 19 Response in. Japan 2021;19:1–3. doi:10.1089/hs.2020.0159.

Jones TC, Biele G, Mühlemann B, Veith T, Schneider J, Beheim-Schwarzbach J, et al. Estimating infectiousness throughout SARS-CoV-2 infection course. Science 2021;5273:1–21. doi:10.1126/science.abi5273.

Kim JK, Crimmins EM. How does age affect personal and social reactions to COVID-19: Results from the national Understanding America Study. PLosOne 2020. doi:10.1371/journal.pone.0241950.

Kucharski AJ, Althaus CL. The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission. Eurosurveillance 2015;20:14–18. doi:10.2807/1560-7917.ES2015.20.25.21167.

Lau MSY, Dalziel BD, Funk S, McClelland A, Tiffany A, Riley S, et al. Spatial and temporal dynamics of superspreading events in the 2014-2015 West Africa Ebola epidemic. Proc Natl Acad Sci U S A 2017;114:2337–42. doi:10.1073/pnas.1614595114.

Laxminarayan R, Wahl B, Dudala SR, Gopal K, Mohan C. Epidemiology and transmission dynamics of COVID-19 in two Indian sates. Science 2020(80-):7672. doi:10.1126/science.abd7672.

Leo YS, Chen M, Heng BH, Lee CC, Paton N, Ang B, et al. Severe acute respiratory syndrome - Singapore. MMWR 2003;52:405–11. 289:3231–4 https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5218a1.htm.

Lim JS, Noh E, Shim E, Ryu S. Temporal Changes in the Risk of Superspreading Events of Coronavirus Disease 2019. Open Forum Infect Dis 2021;8:2019–21. doi:10.1093/ofid/ofab350.

Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the ef-

fect of individual variation on disease emergence. Nature 2005;438:355–9. doi:10.1038/nature04153.

Lopez Bernal J, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et al. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. N Engl J Med 2021;385:585–94. doi:10.1056/nejmoa2108891.

Monod M, Blenkinsop A, Xi X, Hebert D, Bershan S, Tietze S, et al. Age groups that sustain resurging COVID-19 epidemics in the United States. Science 2021;371(80-). doi:10.1126/SCIENCE.ABE8372.

Murphy J, Vallières F, Bentall RP, Shevlin M, Mcbride O, Hartman TK, et al. Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. Nat Commun n.d 2021:1–15. doi:10.1038/s41467-020-20226-9.

Nagata S, Nakaya T, Adachi Y, Inamori T, Nakamura K, Arima D, et al. Mobility change and covid-19 in japan: Mobile data analysis of locations of infection. J Epidemiol 2021;31:387–91. doi:10.2188/jea.JE20200625.

Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. Int J Infect Dis 2020a;93:284–6. doi:10.1016/j.ijid.2020.02.060.

Nishiura H, Oshitani H, Kobayashi T. Closed environments facilitate secondary transmission of coronavirus disease 2019. MedRxiv 2020b:1–5. doi:10.1101/2020.02.28.20029272.

Oshitani H. Cluster-based approach to Coronavirus Disease 2019 (COVID-19) response in Japan—February–April 2020. Jpn J Infect Dis 2020:2019. doi:10.7883/yoken.JJID.2020.363.

Oster AM, Caruso E, DeVies J, Hartnett KP, Boehmer TK. Transmission Dynamics by Age Group in COVID-19 Hotspot Counties. MMWR 2020;69:1494–6 https://www.cdc.gov/mmwr/volumes/69/wr/pdfs/mm6941e1-H.pdf.

Pollán M, Pérez-Gómez B, Pastor-Barriuso R, Oteo J, Hernán MA, Pérez-Olmeda M, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet 2020;396:535–44. doi:10.1016/S0140-6736(20)31483-5.

Pung R, Cook AR, Chiew CJ, Clapham HE, Sun Y, Li Z, et al. Effectiveness of Containment Measures against COVID-19 in Singapore: Implications for Other National Containment Efforts. Epidemiology 2020;32:79–86. doi:10.1097/EDE.0000000000001257.

Ritchie H, Mathieu E, Lucas Rodés-Guirao, Cameron Appel CG, Ortiz-Ospina E, Hasell J, Macdonald B, et al. Coronavirus Pandemic (COVID-19). Our World Data 2020. https://ourworldindata.org/covid-vaccinations (accessed August 27, 2021).

Rosenberg ES, Holtgrave DR, Dorabawila V, Conroy M, Greene D. New COVID-19 Cases and Hospitalizations Among Adults, by Vaccination Status — New York, May 3 – July 25, 2021. MMWR 2021;70:1150–5 https://www.cdc.gov/mmwr/volumes/70/wr/mm7037a7.htm.

Ryu S, Kim D, Lim J-S, ST Ali, Cowling B. Serial Interval and Transmission Dynamics during SARS-CoV-2 Delta Variant Predominance, South Korea. Emerg Infect Dis J 2022;28. doi:10.3201/eid2802.211774.

Sayampanathan AA, Heng CS, Pin PH, Pang J, Leong TY, Lee VJ. Infectivity of asymptomatic versus symptomatic COVID-19. Lancet 2021;397:93–4. doi:10.1016/S0140-6736(20)32651-9.

Sneppen K, Nielsen BF, Taylor RJ, Simonsen L. Overdispersion in COVID-19 increases the effectiveness of limiting nonrepetitive contacts for transmission control. Proc Natl Acad Sci U S A 2021;118:1–6. doi:10.1073/pnas.2016623118.

Sun K, Wang W, Gao L, Wang Y, Luo K, Ren L, et al. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. Science 2021;371(80-). doi:10.1126/science.abe2424.

To KKW, Tsang OTY, Leung WS, Tam AR, Wu TC, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis 2020;20:565–74. doi:10.1016/S1473-3099(20)30196-1.

Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological assessment of hospitalized patients with COVID-2019. Nature 2020;581. doi:10.1038/s41586-020-2196-x.

Wong G, Liu W, Liu Y, Zhou B, Bi Y, Gao GF. MERS, SARS, and Ebola: The Role of Super-Spreaders in Infectious Disease. Cell Host Microbe 2015;18:398–401. doi:10.1016/j.chom.2015.09.013.

Xu X, Liu F, Wu Y, Ali T, Du Z, Bosetti P, et al. Reconstruction of Transmission Pairs for Novel Coronavirus Disease 2019 (COVID-19) in Mainland China : Estimation of Superspreading Events, Serial Interval, and Hazard of Infection. Clin Infect Dis 2020;2019:3163–7. doi:10.1093/cid/ciaa790.

Yoshiyama T, Saito Y, Masuda K, Nakanishi Y, Kido Y, Uchimura K, et al. Prevalence of SARS-CoV-2 – Specific Antibodies, Japan, June 2020. Emerg Infect Dis 2021;27 June 2020. doi:10.3201/eid2702.204088.

Zhu Y, Bloxham CJ, Sinclair JE, Tong ZWM, et al. A Meta-analysis on the Role of Children in Severe Acute Respiratory Syndrome Coronavirus 2 in Household Transmission Clusters. Clin Infect Dis 2021. doi:10.1093/cid/ciaa1825.