MDPI

Article

Identifying High-Risk Events for COVID-19 Transmission: Estimating the Risk of Clustering Using Nationwide Data

Minami Ueda [ID], Katsuma Hayashi and Hiroshi Nishiura *[ID]

School of Public Health, Graduate School of Medicine, Kyoto University, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
* Correspondence: nishiura.hiroshi.5r@kyoto-u.ac.jp; Tel.: +81-75-753-4456

Abstract: The transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is known to be overdispersed, meaning that only a fraction of infected cases contributes to super-spreading. While cluster interventions are an effective measure for controlling pandemics due to the viruses' overdispersed nature, a quantitative assessment of the risk of clustering has yet to be sufficiently presented. Using systematically collected cluster surveillance data for coronavirus disease 2019 (COVID-19) from June 2020 to June 2021 in Japan, we estimated the activity-dependent risk of clustering in 23 establishment types. The analysis indicated that elderly care facilities, welfare facilities for people with disabilities, and hospitals had the highest risk of clustering, with 4.65 (95% confidence interval [CI]: 4.43–4.87), 2.99 (2.59–3.46), and 2.00 (1.88–2.12) cluster reports per million event users, respectively. Risks in educational settings were higher overall among older age groups, potentially being affected by activities with close and uncontrollable contact during extracurricular hours. In dining settings, drinking and singing increased the risk by 10- to 70-fold compared with regular eating settings. The comprehensive analysis of the COVID-19 cluster records provides an additional scientific basis for the design of customized interventions.

Keywords: severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2); epidemiology; observational study; disease outbreak; contact tracing; risk assessment

1. Introduction

As of March 2022, 470 million cases of coronavirus disease 2019 (COVID-19) and 6 million deaths had been recorded worldwide since the start of the pandemic [1]. During the early phases of this pandemic, many countries enforced radical countermeasures, such as lockdowns, to reduce the spread of COVID-19. Although such drastic interventions to reduce contacts have been proven effective in substantially reducing the reproduction number [2], they can have devastating effects on national economies [3] and people's overall health outcomes, including social well-being [4]. Due to the widespread vaccine-induced and naturally acquired immunity in populations worldwide, more countries are seeking a balance between controlling COVID-19 and boosting socioeconomic activities.

The COVID-19 epidemic levels in Japan have been maintained at relatively lower levels than those in Western countries, with a less than 2% cumulative risk of confirmed infection by the end of 2021. Japan is one of the few countries that has accumulated data on clusters since the early phases of the pandemic [5] by focusing on the prevention of clusters via retrospective contact tracing. Since the transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is highly dispersed [6–12], meaning that a small fraction of infected people produce secondary cases, the chain of transmission can effectively be contained by identifying these "super-spreading" individuals. Consequently, retrospective contact tracing, also known as the "backward tracing" approach, is an effective outbreak control measure that can be used to ascertain clusters involving asymptomatic infections [13]. The expert committee on COVID-19 prevention in Japan

analyzed the cluster surveillance data in early 2020 and proposed the concept of the "3Cs" (i.e., closed spaces, crowded places, and close contact settings), which are conditions where infections are likely to occur [14]. Public health and social measures in Japan initially requested that social contacts and movements be uniformly reduced in the form of a state of emergency (SoE). As an alternative measure to the SoE, pre-emergency measures (PEMs) were introduced in early 2021, in which interventions were more customized and focused on high-risk settings [15] that are likely to meet the 3Cs concept. This included settings such as eating and drinking establishments that serve alcohol. PEMs allowed the local governor of the target area to decide which sub-regional areas would be subject to these customized interventions [15].

In support of the 3Cs concept, published studies have demonstrated that SARS-CoV-2, which causes COVID-19, is primarily transmitted through exposure to respiratory fluids via inhalation, deposition, and contact with contaminated surfaces or objects [16–18]. Detailed environmental factors in COVID-19 transmission include enclosed spaces with inadequate ventilation and poor air quality [14,16,19–23], increased exhalation of respiratory fluids [16,24–31], crowded and close contact settings [14,19,23,32], and prolonged exposure [16,19,33]. Studies have also identified specific occasions and workplaces where infections are likely to occur: health care and long-term care facilities [34–37] and the hospitality and entertainment sectors (i.e., bars where eating and drinking take place) [22,32,35,38–41].

On the basis of evidence and expert opinions, the Infectious Diseases Society of America has qualitatively classified daily activities into high-, medium-, and low-risk categories [42]. However, such an assessment remains narrative. Studies on quantitative risk assessment have mostly been restricted to a specific activity category, such as eating and drinking [39], attending schools [43], and workplaces [44]. A test-negative study confirmed an elevated risk of infection in eating and drinking establishments [32], but the comparison was made between positives and negatives among those who attended healthcare establishments, and the sample size was limited. A modeling study showed a notably high infection risk in restaurants during the process of reopening [40], but the risk estimates relied on model fitting procedures using aggregated anonymized location data. A study analyzing surveillance and clinical data showed that entertainment settings, especially bars, have the longest cascade of cases [41]. However, this analysis was based on descriptive statistics and did not analyze the risk of clustering per event, which takes into account the number of potential chances of infection. To the best of our knowledge, no study has comprehensively quantified and compared the risk of clustering that accounts for the patterns of use (i.e., duration, frequency, and the number of users) in diverse daily activities using cluster records in which linked cases were identified through forward and backward tracing.

This study aimed to estimate the risk of COVID-19 clustering in a wide range of daily activities and to identify high-risk conditions of clustering by comparing the risk across different types of settings, including schools and eating and drinking establishments.

2. Materials and Methods

2.1. Cluster Surveillance Data

In this study, we used the incidence of COVID-19 clusters in Japan, which relies on the systematic collection of publicly available information on cluster events. A team commissioned by the Ministry of Health, Labour, and Welfare (MHLW) routinely collected published information on clusters from local health authorities and news articles in the media. This team consisted of full-time researchers from the National Institute of Infectious Diseases (NIID), Tohoku University, and Hokkaido University. The records were routinely updated to reflect the latest information on the identified clusters. We analyzed the entire period of cluster surveillance from June 2020 to June 2021. The SoE, the Japanese implementation of public health and social measures, was declared for some prefectures (a maximum of 11/47 prefectures) in the latter part of the study period. The number of reported clusters and cases by month starting in June 2020, along with the range of SoEs, is

shown in Figure S1. Only one variant of concern became predominant during the study period, and it occurred in the wave from March to June 2021 caused by the Alpha variant (B.1.1.7). When we compared the trend of the number of reported cases between the Alpha variant wave and the former wave (October 2020 to February 2021), we did not observe any significant difference in the trend of the number of reported clusters. We believe that the effect of the variant's predominance on the risk of clustering is negligible because the transmission of concerned variants has remained overdispersed. Additionally, a proportion of the population was vaccinated in the latter part of the study period. The vaccination program in Japan began on 17 February 2021 and prioritized healthcare workers, followed by individuals aged 65 years and older from 12 April. The national vaccination uptake rate was only 10.62% at the end of June 2021 (Figure S2). Therefore, we believe that the effects of herd immunity on transmission patterns are negligible.

We defined a *cluster* as "a group of infected individuals who were identified through one or more links of infection and tested positive", with usually five or more cases that emerged from the same environmental setting [45,46]. Following a guideline presented by the MHLW and the NIID, clusters were identified through contact tracing—combining both forward and backward tracing—by local health authorities with the legally mandated cooperation of infected individuals. A total of 468 main offices of local health authorities are located throughout Japan in every prefecture. Those who were acknowledged as close contacts from contact tracing were additionally tested. Positive cases were confirmed via a reverse transcription-polymerase chain reaction (RT-PCR) test or an antigen test (including both a qualitative test based on immunochromatography and a quantitative test based on the chemiluminescence enzyme immunoassay [CLEIA] method). Reporting of positive cases to the MHLW was mandated by the infectious disease control law in Japan [47]. Even when an antigen test resulted negative, an RT-PCR test was additionally conducted at a physician's discretion for confirmation if the patient was suspected of being infected. Although COVID-19 screening (which includes testing for people without symptoms) in elderly care facilities and health institutions was approved by the government health authority in 2021, it has not been conducted in a systematic and regular manner. However, the members of facilities were prioritized for testing if a case was reported in elderly care facilities. Based on publicly available data from the MHLW, the majority of the tests conducted were performed using the RT-PCR method (Figure S3).

The original dataset included the following information for each record of a cluster: type of setting (primary and secondary), prefecture, name of the cluster, number of cases in the cluster, occurrence date, and notes. Information on the size and area of the location was not available in the dataset.

Prior to the analysis, we additionally annotated the type of setting in each cluster record by reclassifying these into primary, secondary, tertiary, and quaternary categories. This classification was based on the following information: the original classification, the unique name of the cluster, and additional information found in the notes (Figure S4). We used information obtained from search engines, online maps, and news articles to judge the type of setting when the information provided in the original record was inadequate. After the annotation process was completed, we finalized the dataset so that one record corresponded to one specific type of setting. We selected records with the following types of settings for analysis: elderly care facility, welfare facility for people with disabilities, hospital, clinic, dental clinic, serviced entertaining bar and escort club, alcohol-serving eating and drinking establishment, eating and drinking establishment with karaoke, restaurant, university, high school, middle school, elementary school, child welfare facility, kindergarten, cram school, workplace, police station, local government office, sports facility, fire station, theater, and beauty salon or barber (Figure S5). "Serviced entertaining bars and escort clubs" included facilities such as pubs, hostess bars, and snack bars, where staff primarily entertained customers while drinking in close proximity. These facilities are registered as "types 1, 2, and 3" under the Entertainment Business Act. An "alcohol-serving eating and drinking establishment" included facilities categorized

as a "late-night liquor service restaurant", which comprise casual bars such as "Izakaya" (Japanese-style bars) where customers enjoy eating and drinking by themselves or with friends. A "restaurant" encompassed the remaining restaurants and coffee shops used for usual dining occasions other than the preceding two types of setting. With regard to educational settings, "child welfare facilities" included nursery schools, where children aged 0 to 6 years are looked after during the day while their parents work. A kindergarten is a facility similar to a nursery school but with a shorter duration of use and limited age groups from 3 to 6 years old. We calculated the sum of cluster reports for each setting type, represented as c_x, where x represents the type of setting ($x = 1, 2, \cdots, 23$) (Table S1).

2.2. Event Frequency and Users

Additional data on events and types of settings were used to calculate the risk of clustering per event (Table 1). For each type of setting, we defined a concept named *event*, which is a unit of use of the facility each day. An example of this concept is that the definition of an event for a high school was "attending school" for a duration of 8 h per day, which occurs every weekday except during vacations.

We calculated the "total number of events" during the observation period by multiplying the number of days of event d_x by the number of facilities f_x, representing how many days the facilities were open or used during the corresponding period. The number of facilities f_x was obtained or calculated using vital statistics and survey-based statistical information. We used the latest statistics available at the point of analysis. The number of event days d_x was calculated by summing the days for every month according to business and academic calendars.

Two indices—the average number of users per event $\bar{u}_x$ and the duration of the event t_x—were used as the denominator to standardize the activity-dependent risk of clustering between different setting types. We lacked information on the changes in the number of users. Therefore, we used the average value from the pre-COVID period, which was obtained or calculated from statistics, for most of the setting types (other than sports facilities, serviced entertainment bars and escort clubs, alcohol-serving eating and drinking establishments, and restaurants). We have made the assumption that the number of users in other types of settings (i.e., schools, workplaces, and welfare facilities) did not significantly change owing to the nature of establishments. The SoE implemented during the study period was not as powerful as the first SoE in April 2020, which included uniform school closure in the nation. We set assumptions on the duration of the event, assuming a typical length of use if relevant statistics were unavailable. Details of the calculation for each parameter are available as part of the Supplementary Material (Text S1).

Table 1. Demographic details of each setting type. The parameters $\bar{u}_x$, d_x, f_x, and t_x represent the average number of users per event [persons], the number of days of the event [days], the number of facilities [facilities], and the duration of the event [hours]. The "total number of events", which indicates how many days the facilities were open/used during the observation period, was calculated by multiplying d_x by f_x. With regard to theaters, the "total number of events", was calculated as 46,986 [sessions] using statistics rather than calculating from d_x and f_x, for which we could not find reliable data sources.

Category	Establishment Type	Definition of an Event	Type of Users	$\bar{u}_x$ [Users]	d_x [Days]	f_x [Facilities]	t_x [Hours]
Recreational facilities	Theater	Attend a performance	Audiences	341.2		46,986	3.0
	Eating and drinking establishments with Karaoke facility	Visit stores/facilities	Users	110.4	395.0	25,108	2.6
	Sports facility			30.8	268.0	71,271	1.5
	Beauty salon/Barber			6.4	327.0	371,688	1.0
Welfare facility	Welfare facility for persons with disabilities	Stay for a night	Residents, Users	27.5	395.0	5636	16.0
	Elderly care facility			32.2	395.0	28,352	16.0
	Child welfare facility			66.8	268.0	44,616	11.0
Health institute	Hospital	Stay for a night	Inpatients	165.8	395.0	8300	16.0
	Clinic	Visit a doctor	Outpatients	41.1	323.0	102,616	0.5
	Dental clinic			19.7	326.0	68,500	1.0
Eating and drinking establishment	Restaurant	Visit stores	Users	127.4	323.0	1,279,784	1.0
	Alcohol serving eating and drinking establishment			42.9	323.0	264,359	2.0
	Serviced entertaining bars and escort clubs			57.0	323.0	61,857	2.0
Educational institution	University	Attend for a day	Students	3667.4	187.0	795	5.0
	High school			634.4	237.0	4874	8.0
	Middle school			316.6	237.0	10,142	8.0
	Elementary school			322.7	237.0	19,525	7.0
	Kindergarten			111.2	237.0	9698	5.0
	Cram school			122.8	124.7	46,734	1.5
Public office and enterprises	Police station	Work for a day	Staffs	246.2	268.0	1204	8.0
	Local government office			474.1	268.0	1971	8.0
	Fire station			68.2	268.0	2445	8.0
	Workplace			10.6	268.0	5,340,783	7.8

2.3. Estimating the Risk of Clustering

We calculated three types of the "activity-dependent cluster infection risk index" by dividing the sum of cluster reports by three types of denominators. Given that the patterns of use are different for every type of setting (e.g., people stay longer at schools and workplaces than at restaurants), the three indices have different values in the denominator to allow for the comparison of risks from multiple aspects. γ_x (reports/event)—representing the risk of clustering per event per facility, without adjustment for the number of users per event and duration per use—was calculated by dividing the total number of cluster reports by the total number of events. γ_x' (reports/[person·event])—representing the risk per event user—was calculated by adjusting γ_x for the number of users per event. This index identifies the risk of clustering if the same number of people engage in respective events in each establishment, which is likely to be the most common interest for the general public as it represents the risks they are exposed to for an ordinary usage occasion. γ_x'' (reports/[person·event·hour])—representing the risk of clustering per event user hours—was calculated by adjusting γ_x for the number of users per event and duration per use. This index identifies the risk of clustering if the number of users and the duration of usage were equivalent across all establishment types.

We calculated these risk estimates as:

$$\gamma_x = \frac{c_x}{\sum_{j=1}^{f_x}\left(\sum_{i=m}^{n} d_{x_i}\right)}$$
$$\gamma_x' = \frac{c_x}{\sum_{j=1}^{f_x}\left(\sum_{i=m}^{n} d_{x_i}\right)\cdot \overline{u}_x}$$
$$\gamma_x'' = \frac{c_x}{\sum_{j=1}^{f_x}\left(\sum_{i=m}^{n} d_{x_i}\right)\cdot \overline{u}_x \cdot t_x}$$

In the equations, i indicates the month, where the lower limit m and upper limit n correspond to June 2020 and June 2021, respectively. The numerator and denominator for each calculation are shown in Table S2. We calculated the 95% confidence interval (CI) using the Agresti–Coull method [48].

We also calculated the rate ratio (RR) of γ_x', the activity-dependent risk of clustering (adjusted for the number of users) among two categories consisting of similar types of setting—educational and dining facilities—to understand the relative chance of clustering. γ_x' was used for the evaluation of RR for two reasons: (a) it represents the risks that users are exposed to for an ordinary usage occasion; (b) the study aimed to provide the best robust estimate possible without introducing an additional assumption for the duration of usage, as we did in γ_x''. In educational settings, the types of settings included an elementary school, middle school, high school, university, child welfare facility, kindergarten, and cram school. Elementary school was set as the baseline for calculating the RR. With regard to dining, the types of settings included serviced entertaining bars and escort clubs, eating and drinking establishments with karaoke, alcohol-serving eating and drinking establishments, and restaurants. The restaurant was set as the baseline. The CI of the RR was derived using the Wald method. We also calculated the risk difference using the same baseline as a supplemental analysis. All calculations were performed using R software version 3.6.3 [49].

2.4. Sensitivity Analysis on the Estimates

We conducted a sensitivity analysis of the estimates by varying the parameters within the possible range. The range of each parameter is shown in Table S3. Using the Latin hypercube sampling method, we randomly generated 10,000 pairs of parameters, assuming that they followed a uniform distribution in the defined range. We then calculated the risk index (γ_x'), the activity-dependent risk of clustering adjusted for the number of users, for each type of setting using each pair. This process resulted in the distribution of the risk index for each type of setting. Additionally, by using the denominators of the types of settings that were derived from 10,000 parameter pairs, we calculated the RR for each

type of setting in educational and dining facilities. The process similarly resulted in the distribution of RR for each type of setting.

2.5. Simulation-Based Bias Assessment

We performed a simulation-based assessment to evaluate the potential reporting bias in the cluster surveillance data. Specifically, we evaluated our estimate on the rate ratio of the risk of clustering by comparing it with an estimate based on the values from an external study that is not likely to be influenced by reporting biases. We used values from a reference study that assessed the risks of reopening using a fine-grained dynamic mobility network using geolocation data constructed from mobile phones in the USA [40]. They estimated the additional infections for each setting (point of interest; POI) when a particular POI was reopened while other POIs continued to experience a limited level of activities. Among the POIs in the reference study, we focused on "full-service restaurants" and "limited-service restaurants" since we could consider that they respectively correspond to "alcohol serving eating and drinking establishment" and "restaurant" in our study, according to the definitions of the North American Industry Classification System.

Using our cluster surveillance data, we first calculated a rate ratio, separate from the one described earlier, that can be compared with the rate ratio estimated using the values from the reference study. For such convenience, we calculated a provisional index by dividing the number of clusters by the number of facilities. In other words, the index indicated the number of clusters normalized by the number of facilities, which is equivalent to POI in the reference study. By dividing this index for "alcohol serving eating and drinking establishment" by that for "restaurant", we derived a rate ratio that is subject to the bias assessment. We computed the 95% CI using the Agresti–Coull method.

We then calculated the rate ratio of the risk of clustering using the values extracted from the reference study. From Extended Data Figure S5c of the reference study [40], we manually extracted the estimates of additional infections per 100 k patients compared with not reopening, normalized by the number of POIs, using WebPlotDigitizer [50] and fitted the estimates to normal distributions from the median, 25th percentile, and 75th percentile. For each setting, we obtained the distribution of the number of clusters through Monte Carlo simulation by dividing the fitted distribution of the number of additional infections by the distribution of the size of clusters observed in our cluster surveillance data. The mean size of clusters observed in our cluster surveillance dataset was 8.91 and 6.67 for "alcohol serving eating and drinking establishment" and "restaurant", respectively. Just as the provisional index that we calculated from our data indicates, this index indicates the number of clusters normalized by the number of POIs. We used 20,000 samples, respectively, in the process of numerical computation. From the simulated distributions of the number of clusters, we obtained the distribution of rate ratio by dividing the value for "full-service restaurant" by "limited-service restaurant". We conducted parametric bootstrapping from 20,000 samples and calculated the mean and its 95% CI using the bootstrap percentile method.

2.6. Evaluation on the Consistency of Reporting througout the Study Period

Additionally, we conducted a separate analysis to evaluate the consistency of reporting throughout the study period. We first calculated the proportion of the number of clusters for each type of setting until and after the second wave and determined the difference in proportions between the periods. Here, we defined the end of the second wave as the end of September 2020.

3. Results

A total of 7002 clusters were reported in 23 types of settings from June 2020 to June 2021. The average size of clusters included in the dataset was 13.31 people per cluster. Overall, the incidence of cluster reports was proportional to the incidence of confirmed cases in Japan (correlation coefficient of 0.92, Figure S1). Generally, clusters predomi-

nantly occurred in elderly care facilities, workplaces, and hospitals, with 1674, 1160, and 1085 reports, respectively (Figure 1). Among all types of settings, elderly care facilities had the largest number of clusters, with 1674 total reports and a monthly peak of 370 reports in January 2021 (Table S1).
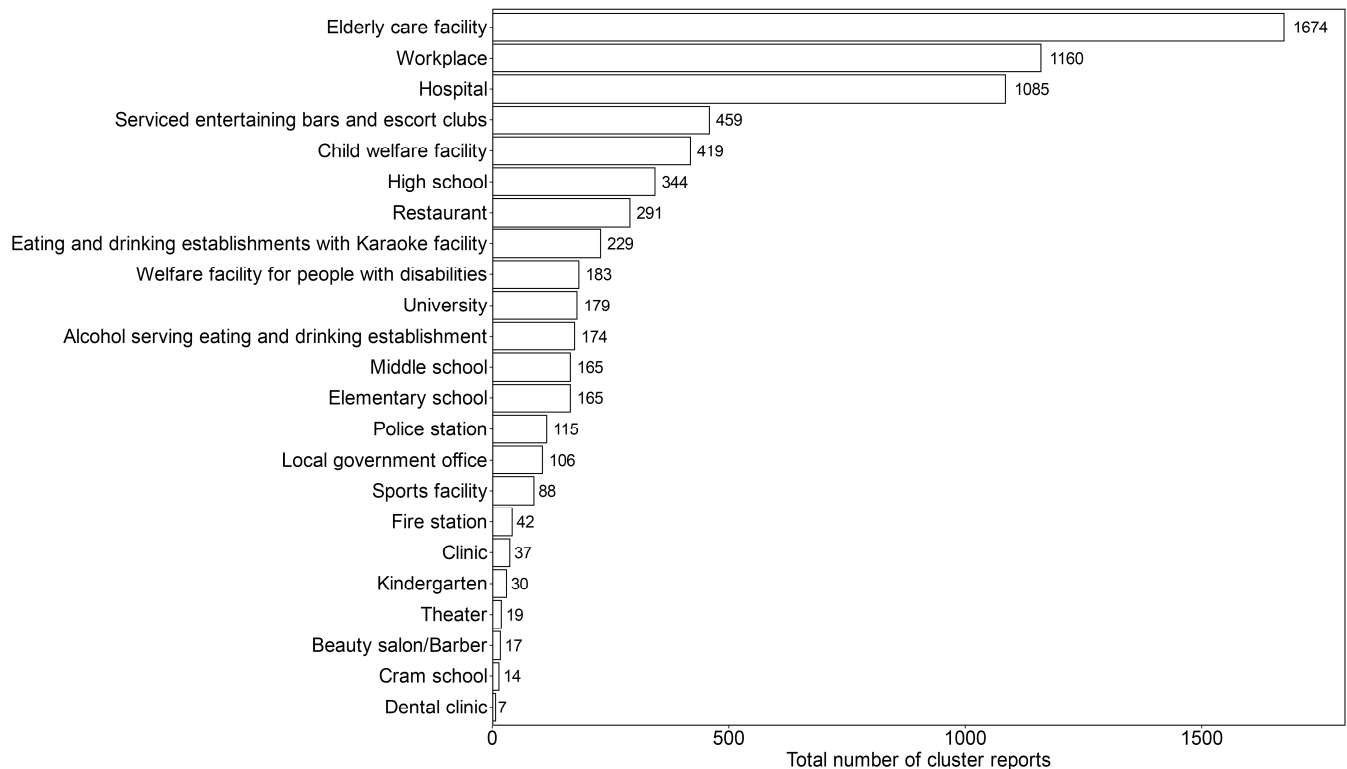

Figure 1. Total number of cluster reports by the type of setting from June 2020 to June 2021. The types of settings were annotated for each record of clusters. The figure shows the total number of cluster reports from June 2020 to June 2021.

Among university clusters, in-depth exploration, in which we examined the settings further in detail based on the name of the cluster and notes, showed that 46.4% were sports-related (i.e., athletics clubs and other extracurricular activities). Only 0.56% of school clusters occurred within the classroom or during classes. Among high school clusters, 8.1% were identifiable as sports-related. Apart from sports-related clusters, another 8.7% of clusters in high schools were related to extracurricular activities; however, details regarding the types of activities were unknown.

3.1. Activity-Dependent Cluster Infection Risk Index

With regard to the activity-dependent risk of clustering without any adjustment (γ_x), university settings showed the highest risk, with 1204 reports per million events (95% CI: 1039–1394) (Figure 2A). This was followed by theaters (404; 95% CI: 254–636), police stations (356; 95% CI: 297–428), hospitals (331; 95% CI: 312–351), high schools (298; 95% CI: 268–331), and local government offices (201; 95% CI: 166–243).

When adjusted for the number of users per event (γ_x'), the risk of clustering at elderly care facilities, welfare facilities for people with disabilities, and hospitals were prominent, with 4.65 (95% CI: 4.43–4.87), 2.99 (95% CI: 2.59–3.46), and 2.00 (95% CI: 1.88–2.12) reports per million event-users, respectively (Figure 2B). This was followed by police stations, theaters, and fire stations. Restaurants had the lowest risk, with 0.01 reports per million event users.
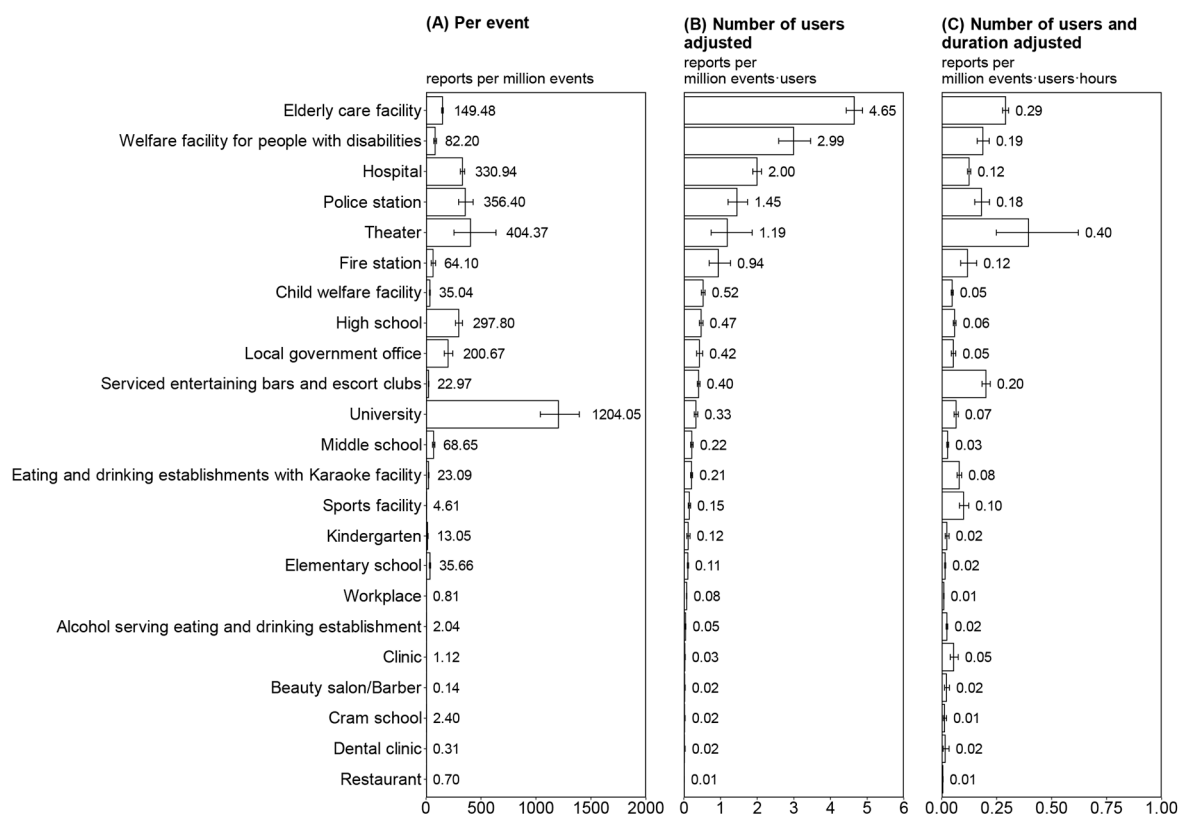
Figure 2. Activity-dependent risk of clustering from June 2020 to June 2021. In each type of setting, we defined event, which represented a typical daily use unit, (i.e., going to school, working at the office, and eating at a restaurant). (**A**) Plot of γ_x, which represents the risk per event without adjustment for the number of users and duration per event. (**B**) Plot of γ_x', which represents the risk when controlling for the number of users. (**C**) Plot of γ_x'', which represents the risk, controlling for the number of users and duration of use. The 95% confidence interval was calculated using the Agresti–Coull method.

When activity-dependent risk was adjusted for the number of users and duration of the event (γ_x''), elderly care facilities, welfare facilities for people with disabilities, and serviced entertainment bars and escort clubs showed comparably high levels of risk (Figure 2C). Theaters had the highest risk with 0.40 reports per million event-user-hours, and the 95% CI ranged broadly from 0.25 to 0.62. When the same number of individuals was assumed to have used each facility for the same duration, these facilities were regarded as having a high risk.

3.2. Rate Ratios of Activity-Dependent Risk

Among educational facilities, the RR of clustering was frequently higher in older age groups, excluding child welfare facilities (Figure 3). Child welfare facilities had the highest RR of 4.75 (95% CI: 3.96–5.68), followed by high school (4.25; 95% CI: 3.53–5.11), university (2.97; 95% CI: 2.40–3.67), middle school (1.96; 95% CI: 1.58–2.43), and kindergarten (1.06; 95% CI: 0.72–1.57). Cram school was the only type of setting with a lower risk than elementary school.

With regard to dining, serviced entertainment bars and escort clubs had a high risk of clustering with a RR of 72.9 (95% CI: 62.95–84.44), followed by eating and drinking establishments with karaoke (37.85; 95% CI: 31.83–45.00) and alcohol-serving eating and drinking establishments (8.60; 95% CI: 7.12–10.37) compared with ordinary restaurants (baseline). The risk differences for types of settings in both categories is shown in Tables S4 and S5.
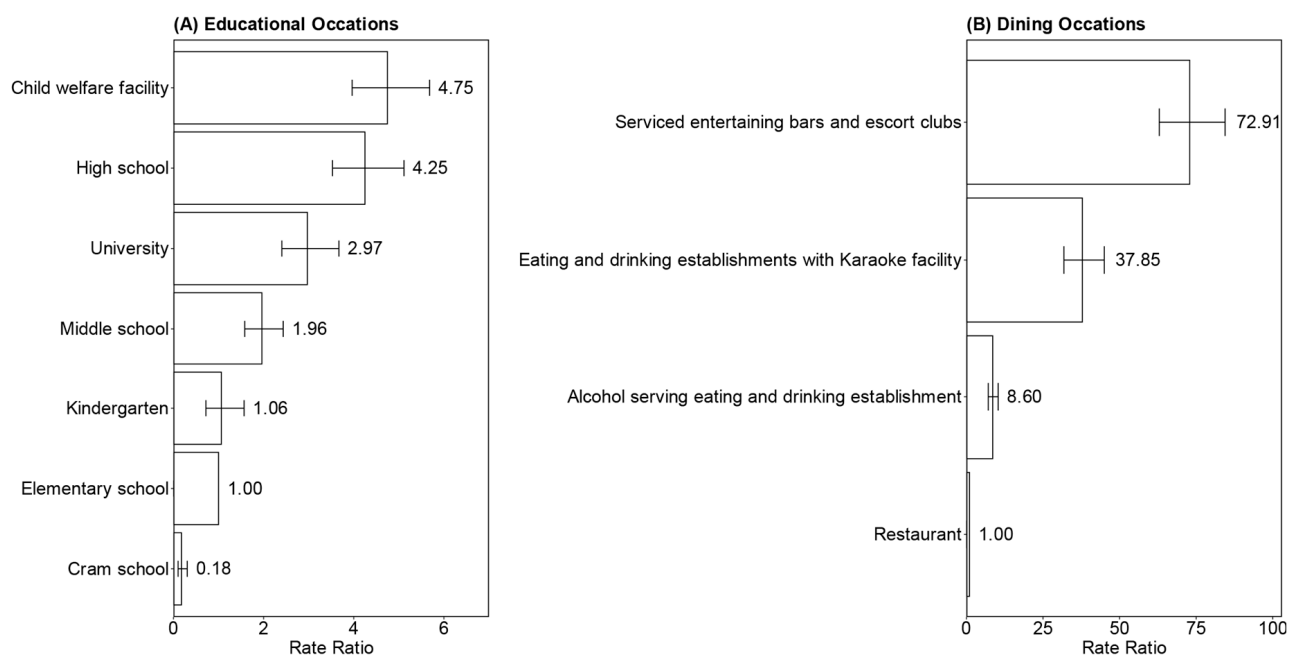
Figure 3. Rate ratio of the activity-dependent risk of clustering (adjusted for the number of users) in educational and dining settings. The rate ratio of the activity-dependent risk of clustering (adjusted for the number of users) γ_x' was calculated among two categories consisting of similar types of settings: (**A**) educational and (**B**) dining settings. Elementary school and restaurants were considered the baseline, respectively. The 95% confidence interval was calculated using the Wald method.

3.3. Sensitivity Analysis of the Risk Index and RR

The results of the sensitivity analysis of the risk index and RR are shown in Figures S6–S8. Each of the figures shows the distribution of each index when the parameters were altered within the predefined possible range. The results of the sensitivity analysis were generally consistent with the original estimates shown in Figures 2 and 3. The mean of the randomly generated 10,000 samples is shown in the top right corner of each panel. The original estimates were within a reasonable range of the estimated distribution.

3.4. Simulation-Based Bias Assessment

The provisional index for our cluster surveillance data, used in the process of bias assessment, was 658.20 and 227.38 for "alcohol serving eating and drinking establishment" and "restaurant", respectively. This resulted in the rate ratio, subject to the bias assessment, for "alcohol serving eating and drinking establishment" compared with "restaurant" to be 2.89 (95% CI: 2.40–3.49).

The fitted distributions of the number of additional infections (per 100 k) compared with not reopening, normalized by the number of POIs, for "full service restaurant" and "limited-service restaurant" are shown in Figure 4A. Using the simulated distribution of the number of clusters in the reference study data, we estimated the mean of the rate ratio for "full service restaurant" compared with "limited-service restaurant" to be 2.76 (95% CI: 2.71–2.81). Figure 4B shows the distribution for the bootstrapped mean of the rate ratio. The estimate from the reference study data provided a consistent value for our estimate, with overlaps in the 95% CI.
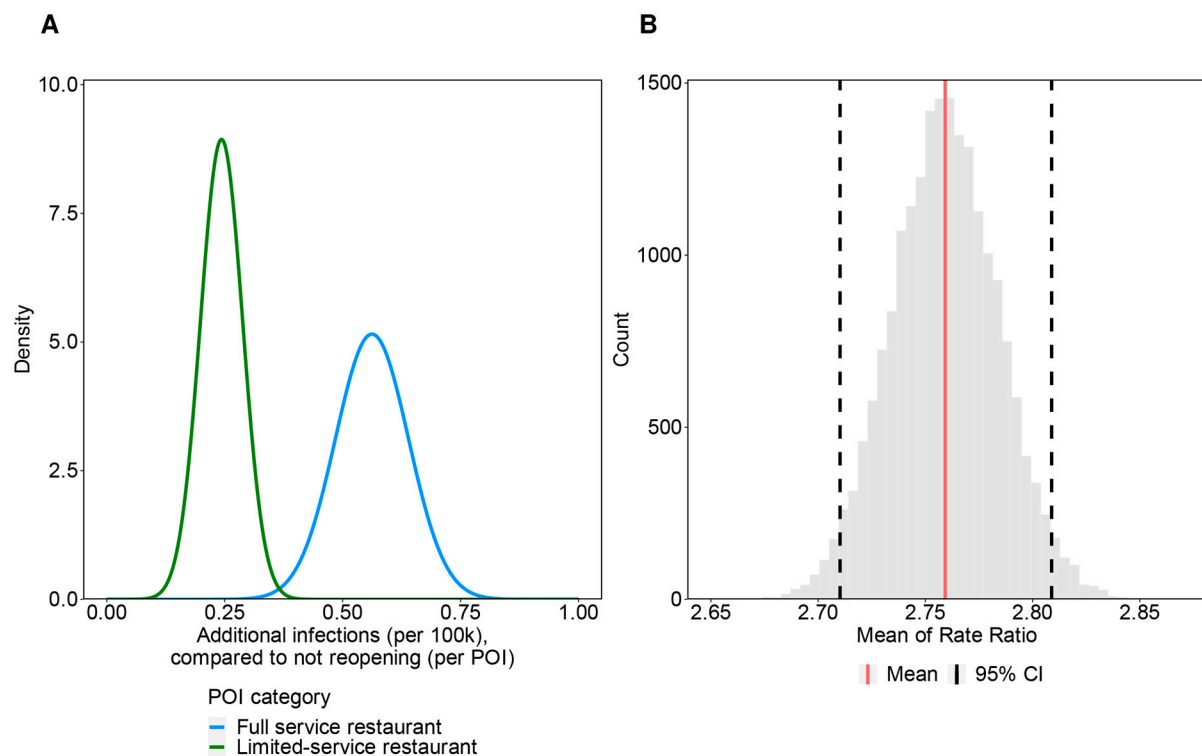
A

B

Figure 4. Result for simulation-based bias assessment: (**A**) Fitted distribution of additional infections compared to not reopening per POI; (**B**) Bootstrapped mean of the rate ratio between "full service restaurant" and "limited-service restaurant" in the reference study. (**A**) shows the fitted distribution of additional infections (per 100 k) compared to not reopening, normalized by the number of POI. The distributions were fitted from median, 25th percentile, and 75th percentile in the original figure of the reference study assuming that they follow normal distribution. (**B**) is the distribution of the bootstrapped mean of rate ratio. Red line shows the mean (2.76) and dotted line shows the 95% CI (2.71–2.81) computed by bootstrap percentile method.

3.5. Evaluation on the Consistency of Reporting throughout the Study Period

In terms of the difference in the proportion of the number of clusters for each type of setting until and after the second wave, schools and care facilities experienced a particular increase in their proportions: elderly care facilities (+11.71 percentage point [pp]), high schools (+2.24 pp), elementary schools (+1.12 pp), and welfare facilities for people with disabilities (+1.07 pp). In contrast, eating and drinking facilities and other settings in the general community observed a particular decrease in their proportions: serviced entertainment bars and escort clubs (−4.58 pp), restaurants (−3.54 pp), alcohol-serving eating and drinking establishments (−2.92 pp), workplaces (−1.86 pp), and eating and drinking establishments with karaoke (−1.19 pp). The values indicate the transition of the major settings where clustering has occurred: settings with positive values experienced a surge of clusters in the later half of the epidemic. The trends generally matched with preceding reports suggesting that clusters initially occur in high-risk settings such as alcohol-serving eating and drinking establishments, propagate to the general community, including schools, and eventually reach the elderly care facilities that are located at the end of the local transmission chain [24,37]. Consistent reporting of clusters is also supported by the comparison of the number of reported clusters and cases by month, as shown in Figure S1.

4. Discussion

Using unique cluster surveillance data from Japan, we investigated the activity-dependent risk of COVID-19 clustering by analyzing the cumulative frequency of cluster reports. Our risk calculation indicated that care facilities posed the highest risk of clustering

per event-user among all examined facilities. A substantial risk with drinking and singing (karaoke) was also observed in dining settings, with an RR 10- to 70-fold greater than usual eating occasions at restaurants. In educational settings, the risk of clustering in child welfare facilities and high schools was nearly four times that of elementary schools.

Our findings could help provide a better understanding of the underlying factors that lead to the clustering of COVID-19 cases. The different levels of clustering among diverse types of settings—primarily caused by differences in characteristics, environments, duration, and frequency of the activities held in each setting—further support the scientific notion of "focused interventions". These interventions have partly been implemented in Japan as PEMs. In particular, the focus on countermeasures on eating and drinking establishments in PEMs is quantitatively supported by our results. With presentations of the risks as numerical estimates, interventions, and resource allocations can be further optimized.

The increased risk in care facilities (i.e., elderly care facilities, welfare facilities for people with disabilities, and hospitals), which is in line with existing studies [34–37], could reflect differences in the duration and quality of contact. As we have assumed in Table 1, many users *live* in these facilities, meaning that there would be prolonged contact. Moreover, owing to the need for physical services provided by caregivers, close contact occurs more often than in other settings. Our findings showing that care facilities are vulnerable to clusters imply that prevention measures for these establishment types should be strengthened. Furthermore, our findings also support the prioritization of vaccination among older people and healthcare professionals, in line with previous evidence [51].

With regard to educational facilities, an overall elevation in the risk of clusters among older age groups is consistent with prior evidence indicating that the spread of SARS-CoV-2 can occur more easily in high schools than in primary schools [52]. The risk of infection could have increased owing to the tendency that schools with older students have larger class sizes, and the frequency and range of social contacts are higher in older age groups [53]. Individuals in older age groups commonly take classes at cram schools with students from other schools, commute to school by train, and participate in extracurricular and volunteer activities in the community outside of school. In-depth analysis of school clusters suggests that the principal causes of infection involve sports- and extracurricular-related activities. As prior research has identified a low risk of SARS-CoV-2 transmission through interactions in close proximity during contact sports [54], settings outside of actual sports practice, such as locker rooms, dormitories, or eating and drinking after the activity, may be the primary source of clustering. While further research is required to pinpoint the exact cause of COVID-19 infections in educational facilities, classes themselves may not necessarily present a high risk of clustering. The difference in risks between child welfare facilities and kindergarten, where the age range of users is similar, could be due to several factors. First, the duration of use is generally longer in child welfare facilities, with children remaining at the facility until their parents come to pick them up after work. In contrast, children in kindergarten usually go home much earlier. Furthermore, the parents' working status could be related to the increased risk in child welfare facilities. For the reason that parents who leave their children in child welfare facilities work outside the home, in contrast to the parents of children in kindergarten, the chance of the former parents becoming infected could be higher. This situation could ultimately lead to the virus spreading at home and to facilities where the child spends time. While the extent is unknown, the wider range of ages in child welfare facilities (0 to 5 years old compared with 3 to 5 years old for kindergarten) could also contribute to the increased infection risk.

Our findings provide evidence that serviced entertainment bars and escort clubs and eating and drinking establishments with karaoke have a significantly greater risk of clustering than regular eating and drinking settings. The content of activities determines the quality of contact and could thus characterize the risks as follows. First, intoxication can induce behaviors that are likely to result in a higher rate of infection. As reported as a concern among staff working at serviced entertainment bars and escort clubs, individuals are likely to be less conscientious about infection prevention measures in an intoxicated

state [55]. Proper wearing of masks can easily be forgotten, and individuals are likely to speak in a louder voice, which causes more emission of respiratory droplets than usual [56,57]. Additionally, singing (karaoke) can further intensify the risks. Although our results should be interpreted with caution because of the population-level study method, they suggested that singing was five times more likely to cause clustering than situations involving alcohol consumption. There are numerous reports of clusters arising from group singing events, and singing is known to cause more dispersion of aerosols [56]. Notably, clusters in eating and drinking establishments with karaoke have most frequently occurred on a specific occasion called "*Hiru-kara*" in Japanese. Unlike usual karaoke settings in karaoke boxes, *Hiru-kara* is a unique form of karaoke in which older adults remain for an extended period during the daytime and into the evening to enjoy singing, eating, and talking with each other. On these occasions, ventilation, which is considered an environmental factor in infection [14,16,19,22,23], could have been insufficient compared with the usual karaoke boxes, which have their own guidelines for infection prevention that were decided under the supervision of public health experts. Poor ventilation is also often observed in serviced entertainment bars and escort clubs. The design structure of these venues is stipulated by law so that the interior cannot be easily seen from the outside, which prevents a sufficient flow of air even when ventilation is attempted [55]. Although our RR results were significantly high and could be interpreted as an overestimation, there have been numerous COVID-19 cases from settings consisting of the above-mentioned factors (i.e., singing and drinking) [24], which support the validity of our estimates. The results of the simulation-based bias assessment, in which we compared the risk ratio using values from a mobility-data-based study that was unlikely to have been influenced by reporting bias, indicated that our estimates possess a comparable level of limited reporting bias within the best extent that we could assess. The result of the difference in the proportions of the clusters until and after the second wave indicates that the reporting was conducted consistently, even amid changes in major settings where clusters frequently occur. The risk for "serviced entertaining bars and escort clubs" could potentially be higher than our estimate because people are likely to be more hesitant to disclose that they have been to such establishments.

We must acknowledge four major limitations of this study. First, while we have shown that the data has unlikely been influenced by the reporting bias as far as we could assess, we still cannot deny the possibility for further reporting and ascertainment biases in the results, owing to the design of data collection. The study investigated clusters that were reported to the Japanese health authorities and were subsequently reported publicly (e.g., via the media). The results regarding types of settings, such as schools and public offices (police, fire stations, and local governments), are more likely to reflect the actual risk than other types of settings because more clusters in these settings are likely to be reported and covered in the media. The results regarding care facilities can be regarded similarly because the testing frequency is commonly more intensive than that in other types of settings. Children are likely to be asymptomatic or mildly symptomatic even when infected with COVID-19. Therefore, the number of clusters in schools and child welfare facilities could have been underreported, resulting in the underestimation of risks. Additionally, the transmission dynamics have altered over time owing to multiple factors after the study period (i.e., emergence of variants, progress of immunization, changes in policies). Especially in the early stages of 2022, child welfare facilities and schools have become one of the major settings of clustering (Figure S2), because progress in immunization has decreased the risk in health care and elderly facilities [51]. Furthermore, the limited testing capacity in Japan [58] may have affected our estimates on the risk of clustering. Taking into consideration the results of a serological survey in Japan [59], we believe that the risk of clustering could be four times that of the original estimate. Even with such potential biases, the present study provided a comprehensive overview of differences in the risk of clustering for most activities in which risk is a concern. While the absolute value of the estimates may need further confirmation from additional studies, we consider that the

relative levels of risks in this study can contribute to real-world decision making during the pandemic to effectively allocate limited resources. It has to be noted that risks for each setting may change due to attributable factors such as policy restrictions, interventions, vaccination rates, participant risk behaviors, and compliance with the preventive measures. Second, limitations in the denominator of risks should be discussed, especially for the use of average values from census data and the time dependence of these values. Due to the SoE's declaration to reduce contact on a voluntary basis, the actual activity patterns might have varied substantially in a dynamic manner over time. We also note that the assumptions for constant school attendance throughout the period lacked quantitative support. Third, most clusters in theaters did not involve audiences but were mainly workplace clusters. Most infected people within theater clusters were performers and staff, and only a few reports have indicated infections among audience members. Fourth, owing to our definition of clusters, local settings with frequent transmission but involving a small number of people (e.g., family events or activities among a small group of friends) were not captured in the dataset. Therefore, our dataset could not identify household transmission risks. The findings from our study should be interpreted with caution because of the inherent limitations listed above.

5. Conclusions

Through a systematic assessment of the activity-dependent risk of COVID-19 case clustering, this study shows that care facilities pose the highest risk of clustering per event user among all facilities. Additionally, drinking and singing in settings accompanied by eating significantly increases the risk of infection by 10- to 70-fold compared with eating in an ordinary restaurant. Among educational facilities, the infection risk is proportionately greater in older age groups and frequently associated with sports and extracurricular activities. The present study is the first to comprehensively analyze COVID-19 cluster records accumulated in Japan, in which clusters were identified via backward and forward contact tracing. By conducting similar surveys during the rest of the pandemic, we can accumulate further quantitative evidence regarding the risk of clustering to consider when designing customized and focused interventions.

References

1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. Available online: https://covid19.who.int (accessed on 23 March 2022).
2. Lau, H.; Khosrawipour, V.; Kocbach, P.; Mikolajczyk, A.; Schubert, J.; Bania, J.; Khosrawipour, T. The Positive Impact of Lockdown in Wuhan on Containing the COVID-19 Outbreak in China. *J. Travel Med.* **2020**, *27*, taaa037. [CrossRef] [PubMed]
3. International Monetary Fund. World Economic Outlook, April 2020: The Great Lockdown. Available online: https://www.imf.org/en/Publications/WEO/Issues/2020/04/14/weo-april-2020 (accessed on 28 December 2021).
4. Sibley, C.G.; Greaves, L.M.; Satherley, N.; Wilson, M.S.; Overall, N.C.; Lee, C.H.J.; Milojev, P.; Bulbulia, J.; Osborne, D.; Milfont, T.L.; et al. Effects of the COVID-19 Pandemic and Nationwide Lockdown on Trust, Attitudes Toward Government, and Well-Being. *Am. Psychol.* **2020**, *75*, 618–630. [CrossRef] [PubMed]
5. National Institute of Infectious Diseases, Japan. Guideline for Execution of Active Epidemiological Investigation on Patients Infected to COVID-19 (Edition as of 12 March 2020). Available online: https://www.niid.go.jp/niid/images/epi/corona/2019nCoV-02-200312.pdf (accessed on 20 December 2021).
6. Endo, A.; Abbott, S.; Kucharski, A.J.; Funk, S. Estimating the Overdispersion in COVID-19 Transmission Using Outbreak Sizes Outside China [Version 3; Peer Review: 2 Approved]. *Wellcome Open Res.* **2020**, *5*, 67. [CrossRef] [PubMed]
7. Lee, H.; Han, C.; Jung, J.; Lee, S. Analysis of Superspreading Potential from Transmission Clusters of COVID-19 in South Korea. *Int. J. Environ. Res. Public Health* **2021**, *18*, 12893. [CrossRef]
8. Tariq, A.; Lee, Y.; Roosa, K.; Blumberg, S.; Yan, P.; Ma, S.; Chowell, G. Real-Time Monitoring the Transmission Potential of COVID-19 in Singapore, March 2020. *BMC Med.* **2020**, *18*, 166. [CrossRef]
9. Wang, L.; Didelot, X.; Yang, J.; Wong, G.; Shi, Y.; Liu, W.; Gao, G.F.; Bi, Y. Inference of Person-to-Person Transmission of COVID-19 Reveals Hidden Super-Spreading Events during the Early Outbreak Phase. *Nat. Commun.* **2020**, *11*, 5006. [CrossRef]
10. Nakajo, K.; Nishiura, H. Transmissibility of Asymptomatic COVID-19: Data from Japanese Clusters. *Int. J. Infect. Dis.* **2021**, *105*, 236–238. [CrossRef]
11. Adam, D.C.; Wu, P.; Wong, J.Y.; Lau, E.H.Y.; Tsang, T.K.; Cauchemez, S.; Leung, G.M.; Cowling, B.J. Clustering and Superspreading Potential of SARS-CoV-2 Infections in Hong Kong. *Nat. Med.* **2020**, *26*, 1714–1719. [CrossRef]
12. Kwok, K.O.; Chan, H.H.H.; Huang, Y.; Hui, D.S.C.; Tambyah, P.A.; Wei, W.I.; Chau, P.Y.K.; Wong, S.Y.S.; Tang, J.W.T. Inferring Super-Spreading from Transmission Clusters of COVID-19 in Hong Kong, Japan, and Singapore. *J. Hosp. Infect.* **2020**, *105*, 682–685. [CrossRef]
13. Endo, A.; Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group; Leclerc, Q.J.; Knight, G.M.; Medley, G.F.; Atkins, K.E.; Funk, S.; Kucharski, A.J. Implication of Backward Contact Tracing in the Presence of Overdispersed Transmission in COVID-19 Outbreaks [Version 3; Peer Review: 2 Approved]. *Wellcome Open Res.* **2021**, *5*, 239. [CrossRef]
14. Expert Meeting on COVID-19 Prevention, Japan. Opinion on COVID-19 Prevention. Available online: https://www.mhlw.go.jp/content/10900000/000606000.pdf (accessed on 20 December 2021).
15. Hayashi, K.; Kayano, T.; Anzai, A.; Fujimoto, M.; Linton, N.; Sasanami, M.; Suzuki, A.; Kobayashi, T.; Otani, K.; Yamauchi, M.; et al. Assessing Public Health and Social Measures against COVID-19 in Japan From March to June 2021. *Front. Med.* **2022**, *9*, 937732. [CrossRef]

16. Centers for Disease Control and Prevention. Scientific Brief: SARS-CoV-2 Transmission. Available online: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/sars-cov-2-transmission.html (accessed on 3 December 2021).

17. Setti, L.; Passarini, F.; De Gennaro, G.; Barbieri, P.; Perrone, M.G.; Borelli, M.; Palmisani, J.; Di Gilio, A.; Piscitelli, P.; Miani, A. Airborne Transmission Route of COVID-19: Why 2 Meters/6 Feet of Inter-Personal Distance Could Not Be Enough. *Int. J. Env. Res. Public Health* **2020**, *17*, 2932. [CrossRef]

18. Stadnytskyi, V.; Bax, C.E.; Bax, A.; Anfinrud, P. The Airborne Lifetime of Small Speech Droplets and Their Potential Importance in SARS-CoV-2 Transmission. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 11875–11877. [CrossRef]

19. Department for Digital, Culture, Media & Sport. Science Note-Emerging Findings from Studies of Indicators of SARS-CoV-2 Transmission Risk at the Events Research Programme: Environment, Crowd Densities and Attendee Behaviour. Available online: https://www.gov.uk/government/publications/events-research-programme-phase-ii-and-iii-findings/science-note-emerging-findings-from-studies-of-indicators-of-sars-cov-2-transmission-risk-at-the-events-research-programme-environment-crowd-densi (accessed on 22 December 2021).

20. Leclerc, Q.J.; Fuller, N.M.; Knight, L.E.; Group, C.C.W.; Funk, S.; Knight, G.M. What Settings Have Been Linked to SARS-CoV-2 Transmission Clusters? [Version 2; Peer Review: 2 Approved]. *Wellcome Open Res.* **2020**, *5*, 83. [CrossRef]

21. Nishiura, H.; Oshitani, H.; Kobayashi, T.; Saito, T.; Sunagawa, T.; Matsui, T.; Wakita, T.; MHLW COVID-19 Response Team; Suzuki, M. Closed Environments Facilitate Secondary Transmission of Coronavirus Disease 2019 (COVID-19). Available online: http://medrxiv.org/lookup/doi/10.1101/2020.02.28.20029272 (accessed on 6 December 2021).

22. Lu, J.; Gu, J.; Li, K.; Xu, C.; Su, W.; Lai, Z.; Zhou, D.; Yu, C.; Xu, B.; Yang, Z. COVID-19 Outbreak Associated with Air Conditioning in Restaurant, Guangzhou, China, 2020. *Emerg. Infect. Dis.* **2020**, *26*, 1628–1631. [CrossRef]

23. Liu, Y.; Ning, Z.; Chen, Y.; Guo, M.; Liu, Y.; Gali, N.K.; Sun, L.; Duan, Y.; Cai, J.; Westerdahl, D.; et al. Aerodynamic Analysis of SARS-CoV-2 in Two Wuhan Hospitals. *Nature* **2020**, *582*, 557–560. [CrossRef]

24. Furuse, Y.; Sando, E.; Tsuchiya, N.; Miyahara, R.; Yasuda, I.; Ko, Y.K.; Saito, M.; Morimoto, K.; Imamura, T.; Shobugawa, Y.; et al. Clusters of Coronavirus Disease in Communities, Japan, January 2020. *Emerg. Infect. Dis.* **2020**, *26*, 2176–2179. [CrossRef]

25. Hamner, L. High SARS-CoV-2 Attack Rate Following Exposure at a Choir Practice Skagit County, Washington, March 2020. *MMWR Morb. Mortal. Wkly. Rep.* **2020**, *69*, 606–610. [CrossRef]

26. Atrubin, D. An Outbreak of COVID-19 Associated with a Recreational Hockey Game Florida, June 2020. *MMWR Morb. Mortal. Wkly. Rep.* **2020**, *69*, 1492–1493. [CrossRef]

27. Atherstone, C. SARS-CoV-2 Transmission Associated with High School Wrestling Tournaments Florida, December 2020–January 2021. *MMWR Morb. Mortal. Wkly. Rep.* **2021**, *70*, 141–143. [CrossRef]

28. Jang, S.; Han, S.H.; Rhee, J.-Y. Cluster of Coronavirus Disease Associated with Fitness Dance Classes, South Korea. *Emerg. Infect. Dis.* **2020**, *26*, 1917–1920. [CrossRef] [PubMed]

29. Bae, S.; Kim, H.; Jung, T.-Y.; Lim, J.-A.; Jo, D.-H.; Kang, G.-S.; Jeong, S.-H.; Choi, D.-K.; Kim, H.-J.; Cheon, Y.H.; et al. Epidemiological Characteristics of COVID-19 Outbreak at Fitness Centers in Cheonan, Korea. *J. Korean Med. Sci.* **2020**, *35*, e288. [CrossRef] [PubMed]

30. Alsved, M.; Matamis, A.; Bohlin, R.; Richter, M.; Bengtsson, P.-E.; Fraenkel, C.-J.; Medstrand, P.; Löndahl, J. Exhaled Respiratory Particles during Singing and Talking. *Aerosol. Sci. Technol.* **2020**, *54*, 1245–1248. [CrossRef]

31. Dougherty, K. SARS-CoV-2 B.1.617.2 (Delta) Variant COVID-19 Outbreak Associated with a Gymnastics Facility Oklahoma, April 2021. *MMWR Morb. Mortal. Wkly. Rep.* **2021**, *70*, 1004–1007. [CrossRef] [PubMed]

32. Fisher, K.A. Community and Close Contact Exposures Associated with COVID-19 among Symptomatic Adults ≥ 18 Years in 11 Outpatient Health Care Facilities United States, July 2020. *MMWR Morb. Mortal. Wkly. Rep.* **2020**, *69*, 1258–1264. [CrossRef]

33. Ge, Y.; Martinez, L.; Sun, S.; Chen, Z.; Zhang, F.; Li, F.; Sun, W.; Chen, E.; Pan, J.; Li, C.; 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*, 1343–1350. [CrossRef]

34. Nguyen, L.H.; Drew, D.A.; Graham, M.S.; Joshi, A.D.; Guo, C.-G.; Ma, W.; Mehta, R.S.; Warner, E.T.; Sikavi, D.R.; Lo, C.-H.; et al. Risk of COVID-19 among Front-Line Health-Care Workers and the General Community: A Prospective Cohort Study. *Lancet Public Health* **2020**, *5*, e475–e483. [CrossRef]

35. Hiironen, I.; Saavedra-Campos, M.; Panitz, J.; Ma, T.; Nsonwu, O.; Charlett, A.; Hughes, G.; Oliver, I. Occupational Exposures Associated with Being a COVID-19 Case; Evidence from Three Case-Controls Studies. Available online: https://www.medrxiv.org/content/10.1101/2020.12.21.20248161v1 (accessed on 2 December 2021).

36. European Centre for Disease Prevention and Control. COVID-19 Clusters and Outbreaks in Occupational Settings in the EU/EEA and the UK. Available online: https://www.ecdc.europa.eu/en/publications-data/covid-19-clusters-and-outbreaks-occupational-settings-eueea-and-uk (accessed on 2 December 2021).

37. Pray, I.W.; Kocharian, A.; Mason, J.; Westergaard, R.; Meiman, J. Trends in Outbreak-Associated Cases of COVID-19 Wisconsin, March 2020. *MMWR Morb. Mortal. Wkly. Rep.* **2021**, *70*, 114–117. [CrossRef]

38. Environmental Modelling Group Transmission Group. Insights on Transmission of COVID-19 with a Focus on the Hospitality, Retail and Leisure Sector. Available online: https://www.gov.uk/government/publications/emg-transmission-group-insights-on-transmission-of-covid-19-with-a-focus-on-the-hospitality-retail-and-leisure-sector-8-april-2021 (accessed on 20 December 2021).

39. National Institute of Infectious Diseases, Japan. Case-Control Study of the Social Activity and Behavioral Risks of COVID-19 Infection in Unvaccinated Individuals (Provisional Report). Available online: https://www.niid.go.jp/niid/ja/2019-ncov/2484-idsc/10692-covid19-59.html (accessed on 20 December 2021).

40. Chang, S.; Pierson, E.; Koh, P.W.; Gerardin, J.; Redbird, B.; Grusky, D.; Leskovec, J. Mobility Network Models of COVID-19 Explain Inequities and Inform Reopening. *Nature* **2021**, *589*, 82–87. [CrossRef]

41. Wong, N.S.; Lee, S.S.; Kwan, T.H.; Yeoh, E.-K. Settings of Virus Exposure and Their Implications in the Propagation of Transmission Networks in a COVID-19 Outbreak. *Lancet Reg. Health West. Pac.* **2020**, *4*, 100052. [CrossRef]

42. Infectious Disease Society of America. What the Experts Say about COVID-19 Risks. Available online: https://www.idsociety.org/globalassets/idsa/public-health/covid-19/activity-risk.pdf (accessed on 20 December 2021).

43. Fontanet, A.; Tondeur, L.; Grant, R.; Temmam, S.; Madec, Y.; Bigot, T.; Grzelak, L.; Cailleau, I.; Besombes, C.; Ungeheuer, M.-N.; et al. SARS-CoV-2 Infection in Schools in a Northern French City: A Retrospective Serological Cohort Study in an Area of High Transmission, France, January to April 2020. *Euro Surveill.* **2021**, *26*, 2001695. [CrossRef]

44. Swedish National Board of Health. Prevalence of COVID-19 in Different Occupational Groups. Available online: https://www.folkhalsomyndigheten.se/publicerat-material/publikationsarkiv/f/forekomst-av-covid-19-i-olika-yrkesgrupper/ (accessed on 20 December 2021).

45. National Institute of Infectious Diseases, Japan. Guideline for Execution of Active Epidemiological Investigation on Patients Infected to COVID-19 (Edition as of 29 November 2021). Available online: https://www.niid.go.jp/niid/ja/2019-ncov/2559-cfeir/10800-covid19-02.html (accessed on 20 December 2021).

46. Ministry of Health, Labor, and Welfare, Japan. Prevention of the Spread of Infection and Preparation of Medical Care Provision System. Available online: https://www.mhlw.go.jp/stf/covid-19/kansenkakudaiboushi-iryouteikyou.html (accessed on 20 December 2021).

47. Ministry of Health, Labor, and Welfare, Japan. Testing of COVID-19. Available online: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000121431_00132.html (accessed on 20 December 2021).

48. Agresti, A.; Coull, B.A. Approximate Is Better Than "Exact" for Interval Estimation of Binomial Proportions. *Am. Stat.* **1998**, *52*, 119–126. [CrossRef]

49. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.

50. Rohatgi, A. Webplotdigitizer: Version 4.6. Available online: https://automeris.io/WebPlotDigitizer (accessed on 10 October 2022).

51. Sasanami, M.; Kayano, T.; Nishiura, H. The Number of COVID-19 Clusters in Healthcare and Elderly Care Facilities Averted by Vaccination of Healthcare Workers in Japan, February 2021. *MBE* **2022**, *19*, 2762–2773. [CrossRef]

52. Goldstein, E.; Lipsitch, M.; Cevik, M. On the Effect of Age on the Transmission of SARS-CoV-2 in Households, Schools, and the Community. *J. Infect. Dis.* **2021**, *223*, 362–369. [CrossRef]

53. Bi, Q.; Wu, Y.; Mei, S.; Ye, C.; Zou, X.; Zhang, Z.; Liu, X.; Wei, L.; Truelove, S.A.; Zhang, T.; 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*, 911–919. [CrossRef]

54. Jones, B.; Phillips, G.; Kemp, S.; Payne, B.; Hart, B.; Cross, M.; Stokes, K.A. SARS-CoV-2 Transmission during Rugby League Matches: Do Players Become Infected after Participating with SARS-CoV-2 Positive Players? *Br. J. Sport. Med.* **2021**, *55*, 807–813. [CrossRef]

55. National Institute of Infectious Diseases, Japan. Investigation Research on the Infection Risk of COVID-19 in So-Called "Serviced Eating and Drinking Establishments" in the Downtown of Shinjuku-Ward (Interim Report). Available online: https://www.niid.go.jp/niid/ja/diseases/ka/corona-virus/2019-ncov/2488-idsc/iasr-news/10081-491p04.html (accessed on 5 January 2022).

56. Echternach, M.; Gantner, S.; Peters, G.; Westphalen, C.; Benthaus, T.; Jakubaß, B.; Kuranova, L.; Döllinger, M.; Kniesburges, S. Impulse Dispersion of Aerosols during Singing and Speaking: A Potential COVID-19 Transmission Pathway. *Am. J. Respir. Crit. Care. Med.* **2020**, *202*, 1584–1587. [CrossRef]

57. Asadi, S.; Wexler, A.S.; Cappa, C.D.; Barreda, S.; Bouvier, N.M.; Ristenpart, W.D. Aerosol Emission and Superemission during Human Speech Increase with Voice Loudness. *Sci. Rep.* **2019**, *9*, 2348. [CrossRef]

58. Shimizu, K.; Tokuda, Y.; Shibuya, K. Japan Should Aim to Eliminate COVID-19. *BMJ* **2021**, *372*, n294. [CrossRef]

59. Sanada, T.; Honda, T.; Yasui, F.; Yamaji, K.; Munakata, T.; Yamamoto, N.; Kurano, M.; Matsumoto, Y.; Kohno, R.; Toyama, S.; et al. Serologic Survey of IgG against SARS-CoV-2 among Hospital Visitors without a History of SARS-CoV-2 Infection in Tokyo, 2020. *J. Epidemiol.* **2022**, *32*, 105–111. [CrossRef]

60. Live and Entertainment Investigation Committee. White Paper on Live and Entertainment 2021. Available online: https://live-entertainment-whitepaper.jp/pdf/summary2021.pdf (accessed on 28 December 2021).

61. Japan Karaoke Industrialist Association. *White Paper on Karaoke 2021*; Japan Karaoke Industrialist Association: Tokyo, Japan, 2021.

62. Ministry of Economy, Trade, and Industry, Japan. Current Survey of Selected Service Industries, Long-Term Data. Available online: https://www.meti.go.jp/statistics/tyo/tokusabido/result/result_1.html (accessed on 5 January 2022).

63. Ministry of Economy, Trade, and Industry, Japan. Indices of Tertiary Industry Activity. Available online: https://www.meti.go.jp/statistics/tyo/sanzi/result-2.html (accessed on 5 January 2022).

64. Ministry of Education, Culture, Sports, Science and Technology, Japan. Social Education Survey 2018, Survey on Physical Exercise Facilities. Table 207. Number of Users of Public Physical Exercise Facilities. Available online: https:

//www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00400004&tstat=000001017254&cycle=0&tclass1=0000011 38486&tclass2=000001138488&tclass3=000001138498&stat_infid=000031924537&tclass4val=0 (accessed on 5 January 2022).

65. Ministry of Education, Culture, Sports, Science and Technology, Japan. Social Education Survey 2018, Survey on Physical Exercise Facilities. Table 221. Number of Users of Private Physical Exercise Facilities. Available online: https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00400004&tstat=000001017254&cycle=0&tclass1=0000011 38486&tclass2=000001138488&tclass3=000001138499&stat_infid=000031924551&tclass4val=0 (accessed on 5 January 2022).

66. Ministry of Health, Labor, and Welfare, Japan. Fact-Finding Survey on Economic Conditions of Environmental Health Industry 2015. Hair-Dressing and Beauty Salons. Available online: https://www.mhlw.go.jp/content/10900000/000501315.pdf (accessed on 4 January 2022).

67. Ministry of Health, Labor, and Welfare, Japan. Fact-Finding Survey on Economic Conditions of Environmental Health Industry 2015. Barbershops. Available online: https://www.mhlw.go.jp/content/10900000/000501312.pdf (accessed on 4 January 2022).

68. Ministry of Health, Labor, and Welfare, Japan. Survey of Social Welfare Institutions. 2019. Available online: https://www.mhlw.go.jp/toukei/saikin/hw/fukushi/19/dl/tyosa.pdf (accessed on 5 January 2022).

69. Senior Housing Association. Registration Status of Residences with Health and Welfare Services for the Elderly (as of the End of August 2021). Available online: https://www.satsuki-jutaku.jp/doc/system_registration_01.pdf (accessed on 10 October 2021).

70. Ministry of Health, Labor, and Welfare, Japan. Overview of Survey of Medical Institutions 2019. Available online: https://www.mhlw.go.jp/toukei/saikin/hw/iryosd/19/dl/02sisetu01.pdf (accessed on 28 December 2021).

71. Ministry of Health, Labor, and Welfare, Japan. Overview of Patient Survey 2017. Available online: https://www.mhlw.go.jp/toukei/saikin/hw/kanja/17/index.html (accessed on 4 January 2022).

72. Ministry of Health, Labor, and Welfare, Japan. Fact-Finding Survey on Economic Conditions of Environmental Health Industry 2013. Restaurant Business (Eating Places). Survey A. Table 13. Available online: https://www.mhlw.go.jp/file/06-Seisakujouhou-10900000-Kenkoukyoku/0000169420.xls (accessed on 5 January 2022).

73. Ministry of Health, Labor, and Welfare, Japan. Fact-Finding Survey on Economic Conditions of Environmental Health Industry 2013. Coffee Shop Business. Survey A. Table 13. Available online: https://www.mhlw.go.jp/file/06-Seisakujouhou-10900000 -Kenkoukyoku/0000171258.xls (accessed on 5 January 2022).

74. National Police Agency, Japan. White Paper on Police 2020. Chapter 7, Section 1, Item1. Figure 7-1. Available online: https://www.npa.go.jp/hakusyo/r02/honbun/html/w7711000.html (accessed on 4 January 2022).

75. Fire and Disaster Management Agency. 2020 White Paper on Fire Service. Available online: https://www.fdma.go.jp/publication/hakusho/r2/56707.html (accessed on 4 January 2022).

76. Ministry of Internal Affairs and Communications, Japan. Survey on the Total Number Management of Civil Servants in Local Governments. 2021. Available online: https://www.soumu.go.jp/main_sosiki/jichi_gyousei/c-gyousei/teiin/ (accessed on 4 January 2022).

77. Ministry of Economy, Trade, and Industry, Japan. Economy Census 2016, Cross-Industrial Survey on Office. Available online: https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200553&tstat=000001095895&cycle=0&tclass1 =000001106235&tclass2=000001106275&tclass3=000001114495&stat_infid=000031724997&tclass4val=0 (accessed on 4 January 2022).

78. Ministry of Education, Culture, Sports, Science and Technology, Japan. School Basic Survey 2020. Elementary, Secondary Educational Institutions; Specialized Training College; Miscellaneous Schools. School Survey. Summary. Available online: https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00400001&tstat=000001011528&cycle=0&year= 20200&tclass1=000001148386&tclass2=000001148404&tclass3=000001148405&tclass4=000001148406&stat_infid=000032040244 &tclass5val=0 (accessed on 4 January 2022).

79. Ministry of Education, Culture, Sports, Science and Technology, Japan. School Basic Survey 2020. Higher Education Institutions. School Survey. Summary. Available online: https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=004000 01&tstat=000001011528&cycle=0&year=20200&tclass1=000001148386&tclass2=000001148387&tclass3=000001148388&tclass4=0 00001148389&stat_infid=000032040259&tclass5val=0 (accessed on 4 January 2022).

80. National Institute for Educational Policy Research. National Survey on Academic Abilities and Learning Conditions 2017, Questionnaire Survey. Available online: https://www.nier.go.jp/17chousakekkahoukoku/report/data/17qn_05.pdf (accessed on 4 January 2022).

81. Benesse Educational Research and Development Institute. Report on 5th Basic Survey of Learning. Chapter 3. Learning Opportunities Outside School. Available online: https://berd.benesse.jp/up_images/research/3_chp3.pdf (accessed on 4 January 2022).

82. Japan Karaoke Industrialist Association. *White Paper on Karaoke 2020*; Japan Karaoke Industrialist Association: Tokyo, Japan, 2020.

83. Ministry of Education, Culture, Sports, Science and Technology, Japan. Overview of the Result of Survey on Physical Education and Sports Facilities. 2018. Available online: https://www.mext.go.jp/sports/content/20200422-spt_stiiki-1368165.pdf (accessed on 28 December 2021).

84. Yano Research Institute. Execution of Survey on Fitness Facilities (2020). Available online: https://www.yano.co.jp/press-release/show/press_id/2657 (accessed on 28 December 2021).

85. Ministry of Health, Labor, and Welfare, Japan. Report on Public Health Administration and Services. 2019. Available online: https://www.mhlw.go.jp/toukei/saikin/hw/eisei_houkoku/19/dl/gaikyo.pdf (accessed on 28 December 2021).

86. National Police Agency, Japan. White Paper on Police 2021. Chapter 2. Figure 2-54. Available online: https://www.npa.go.jp/hakusyo/r03/honbun/html/x2217000.html (accessed on 5 January 2022).

87. National Police Agency, Japan. Current Status of Amusement Businesses and Crackdown on Crimes Related to Amusement Business in 2020. Available online: https://www.npa.go.jp/hakusyo/r03/pdf/09_dai2sho.pdf (accessed on 28 December 2021).

88. Ministry of Economy, Trade, and Industry, Japan. Current Survey of Selected Service Industries. 2018. Available online: https://www.meti.go.jp/statistics/tyo/tokusabizi/result-2/h30/pdf/h30outline.pdf (accessed on 4 January 2022).

89. National Police Agency, Japan. Overview of National Police Agency. Available online: https://www.npa.go.jp/about/overview/index.html (accessed on 4 January 2022).

90. Japan Geographic Data Center. National Database of Prefecture and Local Govenment Office. Available online: https://www.kokudo.or.jp/database/007.html (accessed on 4 January 2022).

91. Ministry of Health, Labor, and Welfare, Japan. Summary of Statistical Surveys. D: Working Time. Table 1: Actual Working Hours. Available online: https://www.mhlw.go.jp/toukei/youran/indexyr_d.html (accessed on 4 January 2022).

92. Benesse Educational Research and Development Institute. Report on Survey of Learning and Living Conditions of College Students. Chapter 2, Section 2. Living Conditions of College Students. Available online: https://berd.benesse.jp/berd/center/open/report/daigaku_jittai/hon/daigaku_jittai_2_2_1.html (accessed on 4 January 2022).

93. Cabinet Office, Government of Japan. Ordinance for Enforcement of Child and Child Care Support Act. Available online: https://elaws.e-gov.go.jp/document?lawid=426M60000002044 (accessed on 4 January 2022).