



ELSEVIER

Contents lists available at ScienceDirect

The Lancet Regional Health - Western Pacific

journal homepage: www.elsevier.com/locate/lanwpc

Research paper

Large-scale epidemiological monitoring of the COVID-19 epidemic in Tokyo

Daisuke Yoneoka^{a,b,c,1}, Yuta Tanoue^{d,1}, Takayuki Kawashima^{e,1}, Shuhei Nomura^{a,c,1},
Shoi Shi^{f,g,1}, Akifumi Eguchi^{h,1}, Keisuke Ejimaⁱ, Toshibumi Taniguchi^j, Haruka Sakamoto^{a,c},
Hiroyuki Kunishima^k, Stuart Gilmour^b, Hiroshi Nishiura^l, Hiroaki Miyata^{a,*}

^a Department of Health Policy and Management, School of Medicine, Keio University, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan

^b Graduate School of Public Health, St. Luke's International University, Tokyo, Japan

^c Department of Global Health Policy, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

^d Institute for Business and Finance, Waseda University, Tokyo, Japan

^e Department of Mathematical and Computing Science, Tokyo Institute of Technology, Tokyo, Japan

^f Department of Systems Pharmacology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

^g Laboratory for Synthetic Biology, RIKEN Center for Biosystems Dynamics Research, Osaka, Japan

^h Center for Preventive Medical Sciences, Chiba University, Chiba, Japan

ⁱ Department of Epidemiology and Biostatistics, Indiana University School of Public Health-Bloomington, Bloomington, USA

^j Department of Infectious Diseases, Chiba University, Chiba, Japan

^k Department of Infectious Diseases, School of Medicine, St. Marianna University, Kanagawa, Japan

^l Graduate School of Medicine, Hokkaido University, Hokkaido, Japan

ARTICLE INFO

Article history:

Received 29 April 2020

Revised 5 August 2020

Accepted 19 August 2020

Available online 10 October 2020

ABSTRACT

Background: On April 7, 2020, the Japanese government declared a state of emergency regarding the novel coronavirus (COVID-19). Given the nation-wide spread of the coronavirus in major Japanese cities and the rapid increase in the number of cases with untraceable infection routes, large-scale monitoring for capturing the current epidemiological situation of COVID-19 in Japan is urgently required.

Methods: A chatbot-based healthcare system named COOPERA (COvid-19: Operation for Personalized Empowerment to Render smart prevention And AN care seeking) was developed to surveil the Japanese epidemiological situation in real-time. COOPERA asked questions regarding personal information, location, preventive actions, COVID-19 related symptoms and their residence. Empirical Bayes estimates of the age-sex-standardized incidence rate and disease mapping approach using scan statistics were utilized to identify the geographical distribution of the symptoms in Tokyo and their spatial correlation r with the identified COVID-19 cases.

Findings: We analyzed 353,010 participants from Tokyo recruited from 27th March to 6th April 2020. The mean (SD) age of participants was 42.7 (12.3), and 63.4%, 36.4% or 0.2% were female, male, or others, respectively. 95.6% of participants had no subjective symptoms. We identified several geographical clusters with high spatial correlation ($r = 0.9$), especially in downtown areas in central Tokyo such as Shibuya and Shinjuku.

Interpretation: With the global spread of COVID-19, medical resources are being depleted. A new system to monitor the epidemiological situation, COOPERA, can provide insights to assist political decision to tackle the epidemic. In addition, given that Japan has not had a strong lockdown policy to weaken the spread of the infection, our result would be useful for preparing for the second wave in other countries during the next flu season without a strong lockdown.

Funding: The present work was supported in part by a grant from the Ministry of Health, Labour and Welfare of Japan (H29-Gantaisaku-ippan-009).

© 2020 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

* Corresponding author.

E-mail address: hiroaki.miyata@gmail.com (H. Miyata).¹ Shared co-first authorship.

Research in context

Evidence before this study

In Japan, on February 1, 2020, novel coronavirus infectious disease (COVID-19) was added as a designated infectious disease under the Infectious Diseases Control Law, which requires physicians to immediately report diagnosed the infection cases to the public health center in their jurisdiction. In response, the National Institute of Infectious Diseases (NIID) launched National Epidemiological Surveillance of Infectious Diseases (NESID) system, as of March 23, 2000. Cases reported by local public health centers were linked to the daily follow-up data by the Ministry of Health, Labour and Welfare (MHLW), via the active epidemiological investigation program. They report the primary sources of data: e.g., lab-confirmed COVID-19 cases (including asymptomatic SARS-CoV-2 infection cases), demographic information such as age, sex and residential location, sign/symptom/comorbidities and medical intervention. Based on data reported as of March 23, 2020, the gender ratio was 1.2:1, higher for male and the median age was 60 years (min–max: 1–97). At the time of the case notification, 18% were reported as asymptomatic. The main signs/symptoms were fever (79%), cough (76%), pneumonia (63%), general malaise (47%). Among those who required invasive ventilation, the main comorbidities were diabetes (27%), hypertension (24%) and dyslipidemia (20%).

Added value of this study

This study is the first study to capture the current epidemiological situation of COVID-19 in Japan using the data of over 350,000 participants from Tokyo through a chatbot-based monitoring system (named COOPERA) between 27th March and 6th April 2020. We identified some geographical clusters of surrogate symptoms of COVID-19 (fever ≥ 37.5 °C) with high spatial correlation with the geographical distribution of PCR-confirmed COVID-19 cases ($r = 0.9$), especially in downtown areas in central Tokyo such as Shibuya and Shinjuku. This suggests that COOPERA partially succeeded in tracking the real-time COVID-19 pandemic situation in Tokyo, Japan.

Implications of all the available evidence

With the nation-wide spread of COVID-19 in major Japanese cities and the insufficient medical resources, a new system to monitor the epidemiological situation, COOPERA, is shown to work for tracking the infectious disease dynamics in communities by checking nondescript symptoms such as "fever" and "shortness of breath".

1. Introduction

The coronavirus disease 2019 (COVID-19), an emerging infectious disease reported in December in Wuhan, Hubei Province, China, has spread rapidly around the world [1,2]. Since the first deaths were reported in early January [3], as of 11 April, 213 countries and territories have been confirmed to have been infected, with 1,614,951 cases and 99,887 deaths reported [4,5].

The WHO has repeatedly recommended that early testing, thorough quarantine, and identification of infection routes are the primary ways to combat the COVID-19 pandemic [6]. However, it is difficult to implement this strategy perfectly and rapidly without enough preparation. In particular, this strategy becomes more difficult to implement as the number of infected people increases. A

few Asian countries (e.g. South Korea or Singapore) that experienced the Severe Acute Respiratory Syndrome (SARS) and Middle Eastern Respiratory Syndrome (MERS) pandemic [7,8] were well prepared, from large-scale testing and case isolation to tracking the behavior of infected individuals to understand the network of virus spread [9,10], and may have been successful in both maintaining economic activities and controlling further spread of the disease. With the exception of these special cases, governments around the world have taken painful and stringent measures, including lockdowns, to prevent the spread of the infection [11]. If the prevalence of COVID-19 in each region could be accurately assessed, the magnitude of the lockdown could be minimized by selecting the cities to close, thereby preparing for measures such as the redistribution of medical resources. Therefore, the actual data aggregated by each region are essential for capturing the situation of the current outbreak.

The Japanese capital, Tokyo, has entered a unique phase of the COVID-19 pandemic. By April 13, 2020, 2068 PCR-confirmed cases and 42 deaths have been identified in Tokyo [12]. Although the growth of cases of COVID-19 in Tokyo has been slower than in Europe and the United State the number is steadily increasing [12]. The Japanese government and the Expert Meeting on the Novel Coronavirus Disease Control have declared that they are close to an explosive surge in infections [13]. In other words, Tokyo is on the edge of COVID-19 control and a large scale monitoring in Tokyo may be able to provide valuable data on prevention strategy not only for COVID-19, but also (re)emerging infectious diseases to which we may be exposed in the future.

Covid-19: Operation for Personalized Empowerment to Render smart prevention And care seeking [COOPERA], a health care support system, has been launched in a collaboration between Kanagawa prefectural government and LINE Corporation [14], the provider of Japan's largest mobile messenger application with 83 million monthly active users, accounting for 65% of Japan's total population. In response to the information input by participants, COOPERA provided personal support such as telephone consultations for respondents who report serious symptoms. The support was based on the Ministry of Health, Labour and Welfare's guidelines for COVID-19 response [15]. COOPERA can be used to estimate the number of those who have COVID-19 related symptoms nationwide. In particular, revealing the number of fever cases per area, disaggregated by postcode, can provide information on infection hotspots by using a disease mapping approach, which is known to be a useful tool for identifying and controlling emerging hotspots not only in COVID-19 cases but also in other infectious diseases [16–19]. This may allow for a more accurate assessment of the spread of infection by region, and provide essential information to minimize the magnitude of lockdown and maximize redistribution of health care.

In this paper, we used the data collected by COOPERA and analyzed the number of geographic fever cases by postcode between March 27 to April 6, 2020, and estimated candidate hotspots in Tokyo. As the Japanese government declared a state of emergency on April 7, this study is one of the largest monitoring attempt that most closely represents the epidemiological situation in Tokyo just before the declaration of a state of emergency and will provide important information for comparison with the situation after the state of emergency was declared.

2. Methods*2.1. Data source and participants*

COOPERA uses a chatbot that asks the participants to provide information such as their physical condition (COVID-19 related symptoms, such as fatigue and fever), and their residential address.

Based on the information, COOPERA provides individualized information, such as telephone consultation for participants who report serious symptoms. LINE users could participate in the system either via the QR code page in the prefecture's website, or the banner at the top of the LINE app screen. The banner was displayed on the app users' phones from 11:30 a.m. on March 29 to 10:59 a.m. on March 30. Due to the policy of LINE Corporation, the users (and the COOPERA participants) are restricted to 13 years old or older. We used data of 353,010 participants who live in Tokyo from March 27 to April 6, 2020. Once the participants are enrolled in initial monitoring and answered that they have COVID-19 related symptoms, they received follow-up questions asking their physical conditions every other day. Note that the population of Tokyo is 13.9 million, as of March 2020. The number of populations was plotted in Appendix Fig. 1. In addition to the data collected through COOPERA, we extracted the daily reported COVID-19 cases (confirmed by PCR test) in each municipality in Tokyo, as of April 6, 2020 [12].

2.2. Outcomes

COOPERA asks (1) basic characteristics of individuals, including age, gender, occupation, comorbidities, preventive actions, and postcode, and (2) health conditions, including current and past month's symptoms (presence or absence of fever, strong feeling of weariness (fatigue) or shortness of breath) and duration of these symptoms. For those who report symptoms, COOPERA further asks additional questions about medical visits and clinical diagnoses, as well as mental conditions, which we do not report in this study. In this study, symptoms are categorized into 4 groups: fever above 37.5 °C (Condition 1), strong feeling of weariness (fatigue) or shortness of breath (Condition 2), both Conditions 1 and 2 (Condition 3), or either Condition 1 or 2 (Condition 4). This study used the initial response to the questionnaire, and follow-up data were not accounted for.

2.3. Statistical/geographical analysis

Basic characteristics were reported with a mean (standard deviation, SD) or proportion. According to National Land Numerical Information published from the Ministry of Land, Infrastructure, Transport and Tourism in Japan, there were 62 municipalities and 8623 postcodes at the time of the study in Tokyo. Geographical analysis in this study was conducted at the postcode level after standardizing the number of participants to the postal code population, stratified by age and sex. Geographical coordinates defined by the latitude and longitude of the center of the area were extracted from map data for each postcode. Empirical Bayes estimates of age-sex-standardized incidence ratio (EBSIR) were used to examine how those who have Conditions 1–4 during the study period were geographically distributed in Tokyo [20,21]. EBSIR incorporates information on the spatial neighborhood areas to smooth the estimate toward the local neighborhood mean, stabilizing estimates and minimizing the problem related to small population at risk [20–23]. Empirical Bayesian methods have been widely used in several studies that use small-area analysis techniques [23,24] and statistically validated [23,24]. The age-sex-standardization was based on the latest national census data in 2015 [25]. The spatial neighborhood was defined as the queen contiguity where spatial neighbors shared at least a common border. For checking the sensitivity of the results by the change of the definition of “spatial neighborhood”, we changed the local adjacency matrix calculated based on the k -nearest neighborhood method with $k = 10, 20$ and 60 [22]. To identify spatial correlation between two spatial data (the number of cases of each symptom and the number of confirmed COVID-19 cases, as of April 6), Tjostein's coefficient,

r , was used after aligning the spatial granularity of the two spatial datasets at the municipality level [26,27]. Lastly, to identify the exact locations of geographical hot spots of Condition 1 in Tokyo, the spatial scan statistic and its associated likelihood ratio-based test proposed by Kulldorff and Nagarwalla [28] and Neil et al. [29] were applied after adjusting for age and sex distribution at the postcode level. Under the assumption of a Poisson distribution for the occurrence of Condition 1, the top likely clusters with $p < 0.05$ were defined. To obtain the corresponding p -values, 10,000 Monte Carlo iterations were implemented. Statistical analyses were conducted with R (version 3.6.3) software.

3. Results

A total 353,010 participants from Tokyo were recruited from March 27 to April 6, 2020. The number of participants per postal code ranged from 0 to 1384 with a median of 128 (Appendix Fig. 2). Table 1 shows the basic characteristics of the participants. Most participants (95.6%) did not have any symptoms when enrolled in the system (i.e., No-symptom group). The distribution of symptomatic conditions was 1.62%, 3.63%, 0.82%, and 4.43% for Conditions 1–4, respectively. Mean and standard deviation (SD) of age at the baseline was 42.7 (SD: 12.3). More females participated: Female (63.4%), Male (36.4%) or Others (0.2%). The most popular preventive actions were covering mouth and nose (with masks, handkerchiefs, etc.) when coughing or sneezing (91.7%), washing hands with soap (93.2%), and hand disinfection with alcohol (68.8%) in the non-symptom group. Table 2 shows the proportion of participants with each preventive action stratified by symptomatic conditions at two time points (29th March and 6th April), showing that the preventive actions have been improving in Tokyo.

The timeline of the proportion of participants with each condition for each day from February 28 to April 3, 2020, is shown in Fig. 1. The gray bar plot depicts the proportions and the number of confirmed cases in Tokyo, showing the increasing number of confirmed cases and proportion of participants with COVID-19 related symptoms during the study period.

Fig. 2 shows the geographical distribution of the number of participants with each symptom (i.e. Conditions 1–4) in Tokyo. The areas are separated by postcodes and the boundaries are white lines. The black line represents railway lines. More symptom-positive participants are observed in central Tokyo, especially around downtowns such as Shibuya and Shinjuku. Appendix Fig. 3 indicates the map of the cumulative number of confirmed COVID-19 cases in Tokyo with the geographic resolution of the 23 Wards and cities. We calculated the spatial correlations (Tjostein's coefficient) between the number of cases of each symptom (Appendix Fig. 4) and number of confirmed COVID-19 cases in Appendix Fig. 3, resulting in the correlation index r of 0.91, 0.90, 0.89 and 0.91 for Conditions 1–4, respectively. Such high spatial correlation suggests that the map of incidence rates of symptoms, such as fever, might partially capture the real-time situation of COVID-19.

The EBSIR of each symptom by areas in Tokyo is shown in Fig. 3. The black line represents the railway line. For checking the sensitivity of the results, since the result are sensitive to the definition of “spatial neighborhood” (i.e., the parameter k), we showed the EBSIR of the group with fever (Condition 1) and the group with both fever and strong feeling of weariness or shortness of breath (Condition 3) during the study period, when k is 10, 20, or 60, respectively in Appendix Fig. 5.

Fig. 4 and Appendix Table 2 show the identified top 9 clusters of Condition 1. Large clusters were found in several populous cities, including Shinjuku (3rd cluster; Number of postcodes = 52; $p < 0.001$) and Shibuya (7th cluster; Number of postcodes=15; $p = 0.016$). In addition, there are some small clusters, including around Musashisakai station (1st cluster; Number of post-

Table 1
Basic characteristics of participants from Tokyo.

	Total (n = 353,010)				
	No symptom (n = 337,370) (%)	(A) Current fever \geq 37.5 °C (n = 5,724) (%)	(B) Current strong feeling of weariness or shortness of breath (n = 12,819) (%)	Both (A) and (B) (n = 2,903) (%)	Either (A) or (B) (n = 15,640) (%)
Age, years					
Mean (SD)	43.00 (12.22)	36.49 (13.02)	36.62 (11.39)	35.79 (12.67)	36.72 (11.77)
Range (Min–Max)	15–101	15–101	15–101	15–101	15–101
\leq 19	4,788 (91.4)	218 (4.2)	342 (6.5)	111 (2.1)	449 (8.6)
20–29	43,187 (90.9)	1,744 (3.7)	3,534 (7.4)	946 (2.0)	4,332 (9.1)
30–39	87,265 (94.5)	1,729 (1.9)	4,237 (4.6)	902 (1.0)	5,064 (5.5)
40–49	99,747 (96.5)	1,185 (1.2)	2,998 (2.9)	565 (0.6)	3,618 (3.5)
50–59	71,360 (97.8)	543 (0.7)	1,292 (1.8)	243 (0.3)	1,592 (2.2)
60–69	24,661 (98.5)	163 (0.7)	287 (1.2)	70 (0.3)	380 (1.5)
70–79	5,685 (97.6)	93 (1.6)	87 (1.5)	40 (0.7)	140 (2.4)
80–89	658 (92.4)	38 (5.3)	36 (5.1)	20 (2.8)	54 (7.6)
\geq 90	19 (63.3)	11 (36.7)	6 (20.0)	6 (20.0)	11 (36.7)
Sex					
Female	213,732 (95.4)	3,530 (1.6)	8,395 (3.8)	1,689 (0.8)	10,236 (4.6)
Male	123,099 (95.8)	2,174 (1.7)	4,379 (3.4)	1,202 (0.9)	5,351 (4.2)
Other	539 (91.1)	20 (3.4)	45 (7.6)	12 (2.0)	53 (9.0)
Occupation					
Self-employed and family employees in agriculture, forestry and fisheries	30,689 (89.2)	565 (1.6)	1,294 (3.8)	322 (0.9)	1,537 (4.5)
Self-employed and family employees in service and commerce	38,802 (91.7)	527 (1.2)	1,228 (2.9)	249 (0.6)	1,506 (3.6)
Self-employed	30,689 (95.2)	565 (1.8)	1,294 (4.0)	322 (1.0)	1,537 (4.8)
Employee (manager) in private company	38,802 (96.3)	527 (1.3)	1,228 (3.1)	249 (0.6)	1,506 (3.7)
Employee in private company	11,406 (91.6)	456 (3.7)	800 (6.4)	208 (1.7)	1,048 (8.4)
Government worker	14,085 (96.6)	176 (1.2)	404 (2.8)	79 (0.5)	501 (3.4)
Student	12,397 (94.0)	311 (2.4)	664 (5.0)	183 (1.4)	792 (6.0)
Part-time job	18,206 (95.7)	325 (1.7)	657 (3.5)	172 (0.9)	810 (4.3)
Unemployed	603 (96.0)	13 (2.1)	17 (2.7)	5 (0.8)	25 (4.0)
Others	39,090 (97.8)	346 (0.9)	717 (1.8)	182 (0.5)	881 (2.2)
Pregnant	4,146 (90.2)	136,206 (95.2)	2,436 (1.7)	5,731 (4.0)	1,238 (0.9)
Taking antifebrile medications (Loxonin, Caronal, etc.)					
Current	6,007 (59.3)	2,219 (21.9)	3,183 (31.4)	1,280 (12.6)	4,122 (40.7)
Past one month	18,692 (87.4)	922 (4.3)	2,318 (10.8)	540 (2.5)	2,700 (12.6)
Medical conditions at baseline					
Cardiovascular diseases	5,812 (94.7)	128 (2.1)	273 (4.5)	76 (1.2)	325 (5.3)
Chronic obstructive pulmonary disease	1,010 (86.6)	42 (3.6)	146 (12.5)	31 (2.7)	157 (13.5)
Kidney-related diseases	2,322 (93.9)	50 (2.0)	130 (5.3)	30 (1.2)	150 (6.1)
Diabetes mellitus	9,888 (95.6)	164 (1.6)	372 (3.6)	84 (0.8)	452 (4.4)
Malignant tumor with anticancer drugs	1,435 (94.9)	30 (2.0)	64 (4.2)	17 (1.1)	77 (5.1)
Malignant tumor without anticancer drugs	3,162 (95.6)	56 (1.7)	122 (3.7)	32 (1.0)	146 (4.4)
Treatment with immunosuppressant	3,210 (92.9)	84 (2.4)	215 (6.2)	55 (1.6)	244 (7.1)
Undergo dialysis	310 (94.8)	11 (3.4)	12 (3.7)	6 (1.8)	17 (5.2)
Preventive action					
Washing hands in runningwater (multiple times a day)	208,200 (95.6)	3,428 (1.6)	7,942 (3.7)	1,719 (0.8)	9,651 (4.4)
Washing hands with soap and water (multiple times a day)	314,399 (95.8)	4,986 (1.5)	11,388 (3.5)	2,477 (0.8)	13,897 (4.2)
Washing hands with alcohol (multiple times a day)	232,235 (95.8)	3,700 (1.5)	8,303 (3.4)	1,846 (0.8)	10,157 (4.2)
Etiquette (masks, handkerchiefs, etc.) in case of coughing or sneezing	309,387 (95.7)	4,960 (1.5)	11,341 (3.5)	2,471 (0.8)	13,830 (4.3)
Take time off from school or work when you have a fever or other symptoms	166,700 (95.8)	3,116 (1.8)	5,766 (3.3)	1,538 (0.9)	7,344 (4.2)
Gargling with water	209,314 (96.0)	3,093 (1.4)	7,142 (3.3)	1,493 (0.7)	8,742 (4.0)
Gargling with Iozine	53,131 (95.5)	830 (1.5)	2,102 (3.8)	453 (0.8)	2,479 (4.5)
Regular ventilation	201,519 (96.2)	2,844 (1.4)	6,478 (3.1)	1,406 (0.7)	7,916 (3.8)
Maintaining humidity	104,404 (96.6)	1,296 (1.2)	3,090 (2.9)	670 (0.6)	3,716 (3.4)
A well-balanced diet	178,354 (96.8)	2,067 (1.1)	4,736 (2.6)	964 (0.5)	5,839 (3.2)
Regular exercise	93,859 (97.5)	904 (0.9)	1,978 (2.1)	437 (0.5)	2,445 (2.5)
Plenty of rest	186,827 (96.6)	2,353 (1.2)	5,255 (2.7)	1,102 (0.6)	6,506 (3.4)
Telework	73,568 (96.3)	1,048 (1.4)	2,309 (3.0)	525 (0.7)	2,832 (3.7)
Staggered commuting hour	68,778 (96.3)	929 (1.3)	2,153 (3.0)	459 (0.6)	2,623 (3.7)
Avoidance of crowds other than staggered commuting hour	114,895 (96.8)	1,220 (1.0)	3,131 (2.6)	564 (0.5)	3,787 (3.2)
Obtain up-to-date coronavirus information	222,925 (96.2)	3,004 (1.3)	7,262 (3.1)	1,458 (0.6)	8,808 (3.8)
Other preventive action	6,249 (95.6)	85 (1.3)	258 (3.9)	51 (0.8)	292 (4.5)
No preventive action	676 (86.6)	44 (5.6)	89 (11.4)	28 (3.6)	105 (13.4)

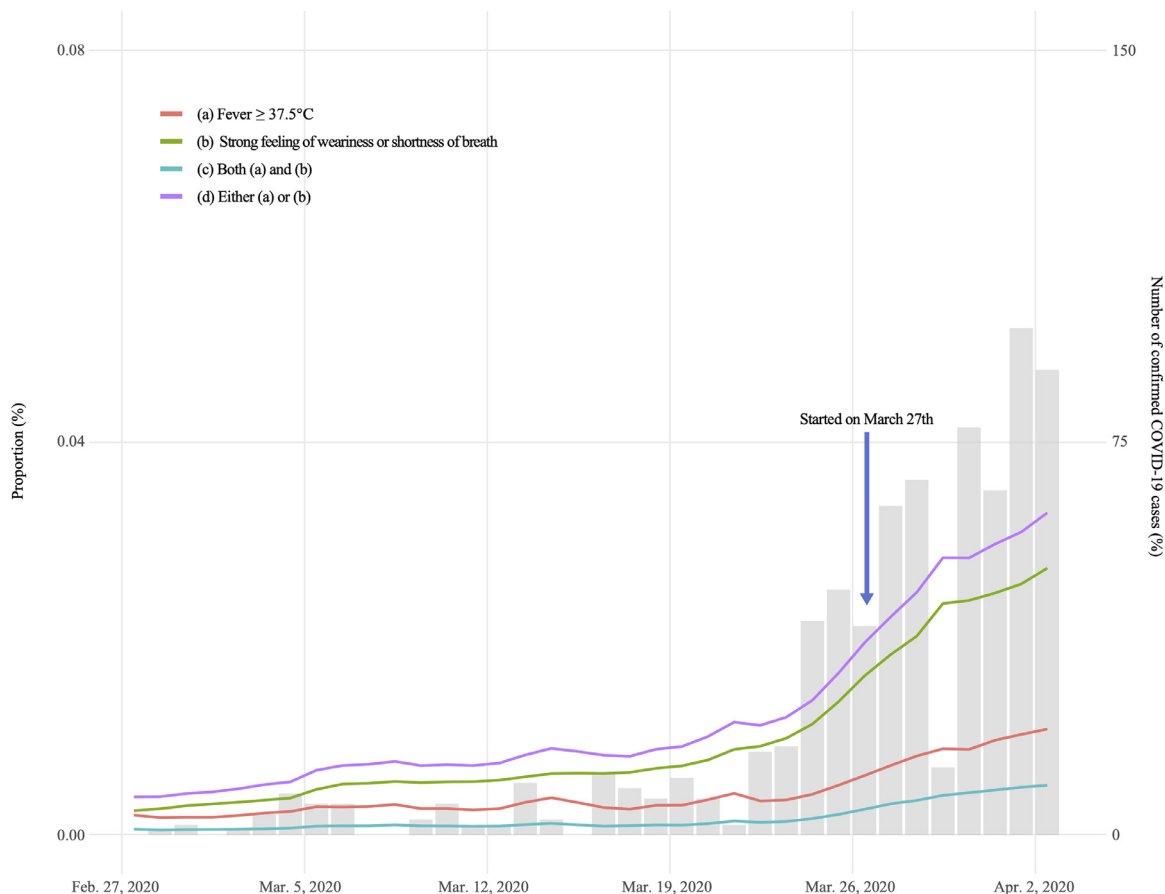


Fig. 1. Trend in the proportion of COVID-19 related symptoms in Tokyo. The red line, green line, blue line, and purple line indicate the proportion of participants who have (a) a fever (Condition 1), (b) a strong feeling of weariness or shortness of breath (Condition 2), both (a) and (b) (Condition 3), and either (a) or (b) (Condition 4), respectively. The gray bar indicates the reported number of confirmed COVID-19 cases in Tokyo. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

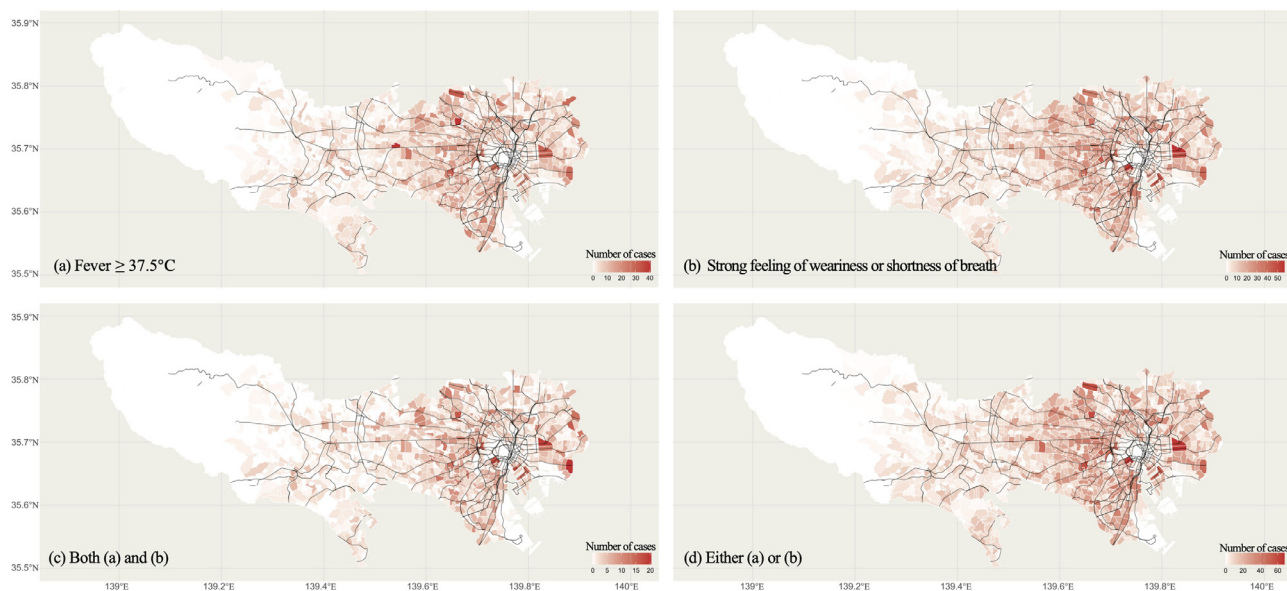


Fig. 2. Maps plotting the number of COVID-19 related symptoms in Tokyo. The black lines represents the railroad network. The block are divided by postcodes. White line indicate boundaries. Panels: (a) fever (Condition 1; upper left), (b) strong feeling of weariness or shortness of breath (Condition 2; upper right), both (a) and (b) (Condition 3; bottom left), and either (a) or (b) (Condition 4; bottom right).

Table 2
Trends in proportions of preventive actions taken in Tokyo.

	March 29	April 6
Washing hands in running water	60,648 (58.56)	36,115 (64.23)
Washing hands with soap and water	96,300 (92.99)	52,401 (93.19)
Hand disinfection with alcohol	68,700 (66.34)	40,040 (71.21)
Etiquette (masks, handkerchiefs, etc.) in case of coughing or sneezing	94,763 (91.50)	51,699 (91.94)
Take time off from school or work when you have a fever or other symptoms	50,938 (49.18)	27,830 (49.49)
Gargling with water	63,648 (61.46)	35,291 (62.76)
Gargling with Isozine	15,877 (15.33)	9,242 (16.44)
Regular ventilation	58,259 (56.25)	36,542 (64.99)
Maintaining humidity	33,366 (32.22)	16,854 (29.97)
A well-balanced diet	54,691 (52.81)	29,847 (53.08)
Regular exercise	28,113 (27.15)	16,436 (29.23)
Getting plenty of rest	57,624 (55.64)	31,100 (55.31)
Telework	21,018 (20.29)	13,808 (24.56)
Staggered commuting	22,502 (21.73)	10,733 (19.09)
Avoidance of crowds other than staggered commuting	33,858 (32.69)	19,556 (34.78)
Staying up-to-date on COVID-19	67,681 (65.35)	37,081 (65.95)
Other preventive measures	1,518 (1.47)	1,232 (2.19)
No preventive measures	224 (0.22)	120 (0.21)

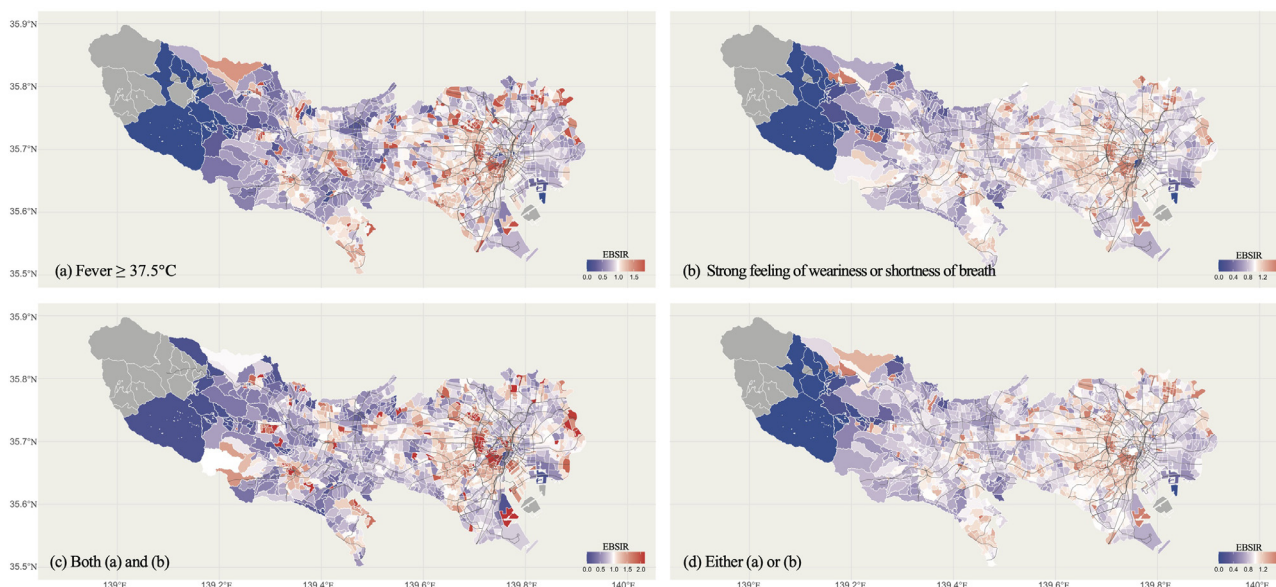


Fig. 3. Maps plotting the proportions of COVID-19 related symptoms in Tokyo. The black lines represent the railroad network. The blocks are divided by postcodes. White lines indicate boundaries. The maximum limit of heat map was set as 1.85, 1.59, 2.16, or 1.57 in Figure (a), (b), (c), or (d), respectively. Panels: (a) fever (Condition 1; upper left), (b) strong feeling of weariness or shortness of breath (Condition 2; upper right), both (a) and (b) (Condition 3; bottom left), and either (a) or (b) (Condition 4; bottom right).

codes = 1; $p < 0.001$) and Rokucho station (2st cluster; Number of postcodes=1; $p < 0.001$).

4. Discussion

In this study, we analyzed the data collected by the novel health care support system, COOPERA, to monitor epidemiological trends in COVID-19 related symptoms. We analyzed data from over 350,000 participants in the capital city, Tokyo, in Japan, which is 3.8% of the population (9.27 million) in the area. We found that 1.62% (5,724) of the participants had fever and 0.82% (2,903) had fever and fatigue during the study period, March 27th and April 6th. However, we should mention that the number of participants with fever includes both COVID-19 cases and other cases such as cold and influenza. In the sense that the estimated proportions were mixtures of COVID-19 and other disease cases, and there was a substantial number of asymptomatic cases have been observed, our estimation does not necessarily mean the prevalence of COVID-19. However, given that we have observed spatial correlation be-

tween the EBSIR of each symptom and the cumulative number of confirmed COVID-19 cases during the time period, we believe the system captured the part of the COVID-19 epidemic in Tokyo.

We illustrated the geographical distribution of COVID-19 related symptoms such as fever and strong feeling of weariness (fatigue) or shortness of breath, and showed the strong spatial correlation between such symptoms and the cumulative number of confirmed cases of COVID-19 ($r \approx 0.9$). Given that our symptom data is not lagged, and the symptoms can occur before a case is confirmed, it is noteworthy that we need to be careful to interpret such strong correlation because it is the correlation between non-lagged symptoms and COVID-19 cases. Several geographical clusters (i.e., the area with a large number of participants with symptoms) were observed, especially in downtown areas: for example, Shibuya and Shinjuku formed relatively larger geographic clusters. These areas, especially Shinjuku area, have reported relatively high rates of COVID-19 cases not only during but also after the study period [30,31]. Further, as we have previously shown [32], the proportion of participants with symptoms in COOPERA is significantly

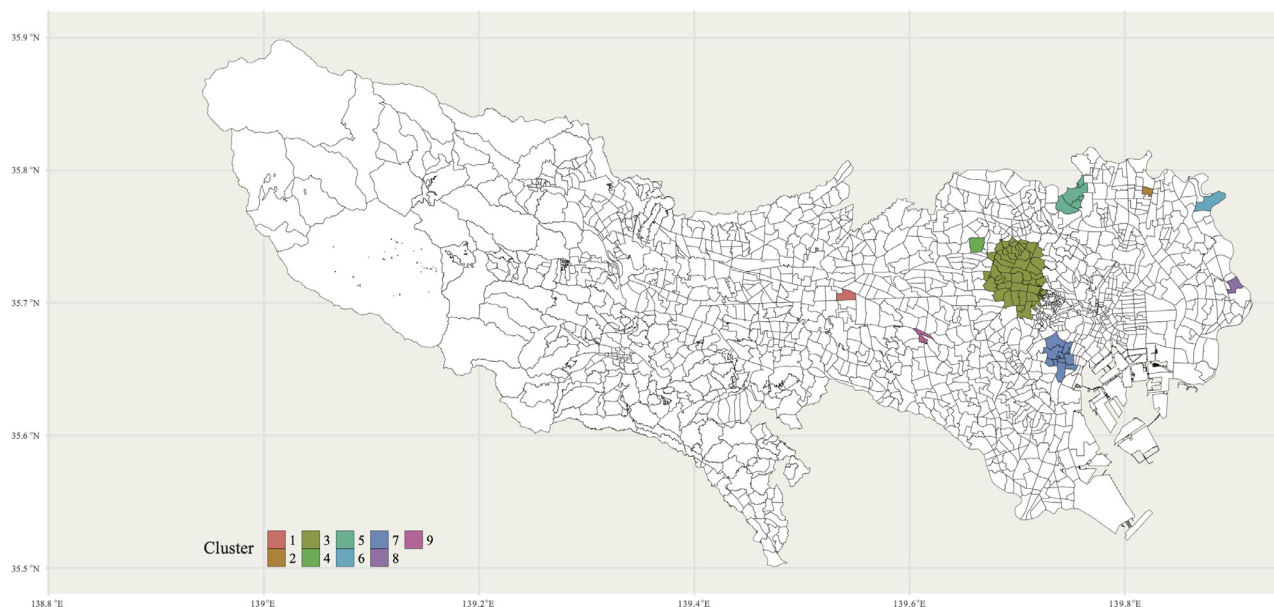


Fig. 4. Map plotting the identified top 9 clusters with $p < 0.05$. The black lines represent the postcode boundaries.

correlated with the number of COVID-19-positive patients after 0–3 days. These facts suggest that a chatbot-based large-scale epidemiological monitoring has a potential to capture the epidemiological situation of COVID-19 in real time. In terms of preventive action, cough etiquette and handwashing were commonly accepted preventive measures (over 90%) while about 20% of the participants executed telework and stagger commuting hours. Note that there was no noticeable change over the study period. This might indicate that most people are already successfully implementing individual-level preventive efforts. However, it is still difficult to execute society-level preventive efforts, which require governmental or community support.

This study has several limitations. The first one is the bias related to the characteristics of participants. Since the system was based on the non-random sampling scheme using the online tool LINE, our results might be influenced by selection bias: those who have no access to the internet or smartphone are underrepresented. In addition, only LINE users who had relatively stronger anxiety due to any symptoms or (subjectively perceived) high risk environment might be more likely to be enrolled in the system. Therefore, our estimation, especially the estimated proportions of symptoms might be overestimated. In addition, it should be noted that there might be “banner effect”: the participants who answered the questions on the day the banner was displayed might include more participants without symptoms than the other days, because people with any symptoms may not be able to check and tap their smartphone. Consequently, the percentage of participants with any symptoms was relatively low on that day. Second, our estimation could be influenced by recall bias: the system included questions regarding past symptoms a month before the enrollment. It might be difficult to accurately report past symptoms especially if they are mild. Thus, recall bias may have created or strengthened the upward trend of the proportions with symptoms (Fig. 1). Third, the male/female ratio in Tokyo residents aged 13 and over (the target age group in this study) in 2015 was 0.97, while the ratio in this study was 0.57, indicating that our data was skewed toward women. In addition, the study participants were relatively skewed toward the younger generation than the population in Tokyo: the majority of participants in the study were 40–49 years old (29.3%), followed by 30–39 years old (26.2%), while these age groups ac-

counted for 16.6% and 15.1% of the Tokyo population aged 13 or older. Further, while people aged 70 or older were only 1.9% of our sample, this age group accounted for 16.2% of the Tokyo population. Although we admit the difference between the population in Tokyo and the study participants, it should be noted that since our estimates were standardized based on the latest National population census in 2015 [25], the bias might be limited. Lastly, our data is based on self-reported answers, whose validation is ongoing. Unfortunately, it is difficult to examine the biases in the sense that it is not possible to compare the data of the participants in COOPERA and non-participants as well as the demographics of LINE users and non-users since there are no available (open) data that allow us to check the true background distribution of participants.

5. Conclusion

In summary, this study reported the analyses of the data collected through a large-scale (over 350,000 participants) health care support system, COOPERA in Tokyo. We observed a strong spatial correlation between the proportions with symptoms and confirmed COVID-19 cases in the areas, suggesting the system partially captured the real-time COVID-19 pandemic situation in Tokyo, Japan. Given the rapid spread of the COVID-19 and limited resources (i.e., PCR or serological test), the COOPERA system could be an alternative approach to monitor the real-time epidemic situation. In addition, given that Japan has not had a strong lockdown policy to weaken the spread of the infection, our result would provide useful insight for preparing for the second wave during the next flu season without a strong lockdown not only in the European countries and the U.S., but also in the low and middle income countries.

Declaration of Interests

Dr. Miyata reports grants from Ministry of Health, Labour and Welfare of Japan, during the conduct of the study.

Acknowledgment

We would like to thank Tokyo for installing the COOPERA system and providing us with data, LINE Corporation for develop-

ing and maintaining the system, and Amazon Web Services, Inc. for providing the data storage space. We are also grateful to the Japanese Society of Infectious Diseases for supervising the questionnaires and information provided to the participants from professional perspectives.

Contributors

All authors took responsibility for the integrity of the data and the accuracy of the data analysis. All the authors made critical revisions to the manuscript for important intellectual content and gave final approval of the manuscript. The opinions, results, and conclusions reported in this paper are those of the authors and are independent from the funding bodies.

Ethics statement

Ethical approval was granted by the ethics committee of Keio University School of Medicine, under authorization number 20190338.

Data statement

Data available on request due to privacy/ethical restrictions.

Role of the funding source

The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the paper. The authors had full access to all the data in the study and had final responsibility to submit for publication.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.lanwpc.2020.100016](https://doi.org/10.1016/j.lanwpc.2020.100016).

References

- [1] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020;382(13):1199–207.
- [2] Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 2020;323(13).
- [3] The New York Times. China reports first death from new virus. 2020 [cited 2020 13 April]; Available from: <https://www.nytimes.com/2020/01/10/world/asia/china-virus-wuhan-death.html>.
- [4] World Health Organization. Coronavirus disease (COVID-19) pandemic. 2020 [cited 2020 May 15]; Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.
- [5] Johns Hopkins Coronavirus Resource Center. Coronavirus COVID-19 global cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). 2020 [cited 2020 May 15]; Available from: <https://coronavirus.jhu.edu/map.html>
- [6] World Health Organization. Laboratory testing strategy recommendations for COVID-19. 2020 [cited 2020 13 April]; Available from: https://apps.who.int/iris/bitstream/handle/10665/331509/WHO-COVID-19-lab_testing-2020.1-eng.pdf.
- [7] Peiris JS, Yuen KY, Osterhaus AD, Stohr K. The severe acute respiratory syndrome. *N Engl J Med* 2003;349(25):2431–41.
- [8] Hui DS. Super-spreading events of MERS-CoV infection. *Lancet* 2016;388(10048):942–3.
- [9] CBS NEWS. South Korea is using fast food-style drive-thru to test for coronavirus. 2020 [cited 2020 13 April]; Available from: <https://www.cbsnews.com/news/coronavirus-south-korea-drive-thru-test-covid-19/>.
- [10] NIKKEI ASIAN REVIEW. Singapore urges citizens to sign up for COVID-19 tracking app. 2020 [cited 2020 13 April]; Available from: <https://asia.nikkei.com/Spotlight/Coronavirus/Singapore-urges-citizens-to-sign-up-for-COVID-19-tracking-app>.
- [11] BUSINESS INSIDER. A third of the global population is on coronavirus lockdown — here's our constantly updated list of countries and restrictions. 2020 [cited 2020 13 April]; Available from: <https://www.businessinsider.com/countries-on-lockdown-coronavirus-italy-2020-3>.
- [12] Tokyo COVID-19 Information. Latest updates on COVID-19 in Tokyo. 2020 [cited 2020 13 April]; Available from: <https://stopcovid19.metro.tokyo.lg.jp/en/>.
- [13] Tokyo Metropolitan Government. STOP COVID19 novel coronavirus latest information from governor Koike to the people of Tokyo. [cited 2020 13 April]; Available from: <https://www.youtube.com/watch?v=PNU7eYdlwts&t=1s>.
- [14] LINE Corporation. LINE corporation. 2020[cited 2020 8 April]; Available from: <https://linecorp.com/en/>
- [15] Ministry of Health Labour and Welfare. Dear people of Japan (New coronavirus infection). 2020 [cited 2020 April 6]; Available from: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000121431_00094.html.
- [16] Desjardins M, Hohl A, Delmelle E. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: detecting and evaluating emerging clusters. *Appl Geogr* 2020;118:102202.
- [17] Desjardins M, Hohl A, Delmelle EM. Space-time clusters and co-occurrence of chikungunya and dengue fever in Colombia from 2015 to 2016. *Acta Trop* 2018;185:77–85.
- [18] Owusu C, Desjardins R, Baker M, Delmelle E. Residential mobility impacts relative risk estimates of space-time clusters of chlamydia in Kalamazoo County, Michigan. *Geospat Health* 2019;14(2).
- [19] Whiteman A, Desjardins R, Eskildsen A, Loaiza R. Detecting space-time clusters of dengue fever in Panama after adjusting for vector surveillance data. *PLoS Negl Trop Dis* 2019;13(9):e0007266.
- [20] Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 1987;43(3):671–81.
- [21] Marshall RJ. Mapping disease and mortality rates using empirical Bayes estimators. *J R Stat Soc Ser C Appl Stat* 1991;40(2):283–94.
- [22] Lawson AB. Bayesian disease mapping: hierarchical modeling in spatial epidemiology. CRC Press; 2013.
- [23] Tsutakawa RK, Shoop GL, Marienfeld CJ. Empirical Bayes estimation of cancer mortality-rates. *Stat Med* 1985;4(2):201–12.
- [24] Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 1987;43(3):671–81.
- [25] Ministry of Internal Affairs and Communications. Portal site of official statistics of Japan. 2015 population census. 2020 [cited 2020 April 6]; Available from: https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001049104&cycle=0&tclass1=000001049105&tstat_infid=000031594311.
- [26] Tjøstheim D. A measure of association for spatial variables. *Biometrika* 1978;65(1):109–14.
- [27] Getis A, Ord JK. The analysis of spatial association by use of distance statistics. In: *Perspectives on spatial data analysis*. Springer; 2010. p. 127–45.
- [28] Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. *Stat Med* 1995;14(8):799–810.
- [29] Neill DB, et al. Detection of emerging space-time clusters. In: *Proceedings of the eleventh ACM SIGKDD international conference on knowledge discovery in data mining*; 2005.
- [30] SHINJUKU CITY. The first meeting for the novel corona in downtown Shinjuku. 2020 [cited 2020 July 31]; Available from: <https://www.city.shinjuku.lg.jp/whatsnew/pub/2020/0619-01.html>.
- [31] SHINJUKU CITY. How Shinjuku City is handling the novel coronavirus. 2020 [cited 2020 July 31]; Available from: http://www.foreign.city.shinjuku.lg.jp/en/kenko/kenko_88/.
- [32] Yoneoka D, Kawashima T, Tanoue Y, Nomura S, Ejima K, Shi S, et al. Early SNS-based screening system for the COVID-19 in Japan: a population-level observational study. *J Epidemiol* 2020;30 Accepted.