Received: 4 June 2021 | Revised: 25 August 2021 | Accepted: 3 September 2021

DOI: 10.1111/tbed.14313

Transboundary and Emerging Diseases WILEY

ORIGINAL ARTICLE

Geo-clusters and socio-demographic profiles at village-level associated with COVID-19 incidence in the metropolitan city of Jakarta: An ecological study

Pandji Wibawa Dhewantara[1] | Tities Puspita[1] | Rina Marina[1] | Doni Lasut[1] | Muhammad Umar Riandi[2] | Tri Wahono[2] | Wawan Ridwan[2] | Andri Ruliansyah[2]

[1] Centre for Research and Development of Public Health Efforts, National Institute of Health Research and Development (NIHRD), Indonesian Ministry of Health, Jakarta, Indonesia

[2] Pangandaran Unit for Health Research and Development, National Institute of Health Research and Development (NIHRD), Indonesian Ministry of Health, West Java, Indonesia

Correspondence
Pandji Wibawa Dhewantara, Centre for Research and Development of Public Health Effort, National Institute of Health Research and Development (NIHRD), Indonesian Ministry of Health, Jakarta 10560, Indonesia.
Email: p.dhewantara@gmail.com

Funding information
National Institute of Health Research and Development; Indonesian Ministry of Health

Abstract

The Special Capital Region of Jakarta is the epicentre of the transmission of COVID-19 in Indonesia. However, much remains unknown about the spatial and temporal patterns of COVID-19 incidence and related socio-demographic factors explaining the variations of COVID-19 incidence at local level. COVID-19 cases at the village level of Jakarta from March 2020 to June 2021 were analyzed from the local public COVID-19 dashboard. Global and local spatial clustering of COVID-19 incidence was examined using the Moran's *I* and local Moran analysis. Socio-demographic profiles of identified hotspots were elaborated. The association between village characteristics and COVID-19 incidence was evaluated. The COVID-19 incidence was significantly clustered based on the geographical village level (Moran's I = 0.174; p = .002). Seventeen COVID-19 high-risk clusters were found and dynamically shifted over the study period. The proportion of people aged 20–49 (incidence rate ratio [IRR] = 1.016; 95% confidence interval [CI]: 1.012–1.019), proportion of elderly (≥50 years) (IRR = 1.045; 95% CI = 1.041–1.050), number of households (IRR = 1.196; 95% CI = 1.193–1.200), access to metered water for washing, and the main occupation of the residents were village level socio-demographic factors associated with the risk of COVID-19. Targeted public health responses such as restriction, improved testing and contact tracing, and improved access to health services for those vulnerable populations are essential in areas with high-risk COVID-19.

KEYWORDS
COVID-19, Indonesia, inequality, spatial analysis, socio-demographics

1 | INTRODUCTION

The coronavirus disease 2019 (COVID-19) is caused by a new type of virus namely severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It has been declared as a public health emergency of international concern by the World Health Organization (WHO) since it first emerged in Wuhan, China, in late December 2019 (WHO, 2020; Zhu et al., 2020). As of 19 August 2021, more than 100 million cases and 2.3 million deaths have been reported worldwide (WHO, 2021). Fever, cough, and shortness of breath are the clinical symptoms that occur at the beginning of the infection and may develop into dyspnoea and lead to severe complications and death (C. Huang, Wang, et al., 2020; Yan et al., 2020). Asymptomatic COVID-19 is common, making the public health and disease control difficult (Mizumoto et al., 2020). COVID-19 is mainly transmitted through respiratory droplets and human-to-human contact (Chan et al., 2020; Kutti-Sridharan et al., 2020).

Indonesia has also encountered the impacts of the COVID-19 pandemic. The first confirmed COVID-19 case was reported in March 2020. As of 19 August 2021, more than 3.9 million people were tested positive and 123,981 people died due to COVID-19 (Indonesia COVID-19 Task Force, 2021). Indonesia is one of the Southeast Asian countries with a high number of COVID-19 cases. The Special Capital Region of Jakarta has reported more than 800,000 cases, accounting for approximately 20% of the total COVID-19 cases nationally. To control the transmission, the government of Indonesia is implementing extensive public health measures, which are outlined in the Decree of the Indonesian Ministry of Health No. 9 of 2020 (Cabinet Secretariate of Republic of Indonesia, 2020). In addition, vaccination programmes have now been rolled out since January 2021, targeting healthcare workers, the elderly, and the general population.

Jakarta, the capital city of Indonesia, is the epicentre of COVID-19 transmission. The first COVID-19 case in Jakarta was confirmed on 3 March 2020. As of 19 August 2021 in Jakarta, a total of 822,399 people were tested positive and 13,132 people died due to COVID-19 (Jakarta COVID-19 Taskforce, 2021). In order to control the transmission, several public health responses are being implemented such as stay-at-home order, travel bans, promoting face masks use, public places closure, contact tracing, and extensive testing (Government of The Special Capital Region of Jakarta, 2020). Prompt and effective responses are, therefore, crucial to reduce and eliminate COVID-19 incidence. As of 19 August 2021, up to 9.2 million Jakartans had received at least one dose of COVID vaccines.

The spatiotemporal variations of COVID-19 incidence have been studied extensively as knowledge regarding the epidemiology of COVID-19 is growing. Understanding patterns of the disease, spatially and temporally, will help researchers and policymakers to identify and target areas and populations that are most at risk, supporting better resource allocation. Much remains unknown about variations of COVID-19 incidence and factors related to this heterogeneity across Jakarta. The use of spatial analysis to support disease outbreak control has been discussed in other studies (Fletcher-Lartey & Caprarelli, 2016). A recent paper reviewed the use of geographic information systems (GIS) and spatial analysis on COVID-19 cases (Franch-Pardo et al., 2020). For instance, a study in China analyzed the spatial-temporal patterns of COVID-19 using Moran statistics and found that COVID-19 incidence was spatially clustered in China (R. Huang, Liu, et al., 2020). A study in England demonstrated that spatial analyses can be used to assess inequalities of environmental and socio-economic attributes that contributed to spatial heterogeneity of COVID-related deaths (Sun et al., 2021). Further, studies revealed significant associations between COVID-19 spatial heterogeneities as well as environmental and socio-demographic factors (Chakraborty, 2020; Kwok et al., 2020; Lakhani, 2020; Luo et al., 2020; Nicodemo et al., 2020; Nakada & Urban, 2020). The GIS techniques can be applied to locate hotspots and to understand factors associated with the disease distribution. These techniques help improve our knowledge regarding the epidemiology of COVID-19 to help design public health responses at various levels (i.e. national, sub-national, and local).

Although there have been studies on the transmission pattern of COVID-19 in Indonesia (Aisyah et al., 2020; Eryando et al., 2020), the variation and dynamic of COVID-19 incidence at the local level, especially in Jakarta, and socio-demographic factors associated with COVID-19 distribution, are poorly investigated. An investigation on the transmission patterns at a finer scale (i.e. village level) and the characteristics of high-risk neighbourhoods could help local health authorities design more effective and targeted interventions to control COVID-19 transmission. This paper explored the spatial pattern of COVID-19 transmission in Jakarta from March 2020 to June 2021 to understand the geographical diffusion of COVID-19 at a fine scale, locate and profile high-risk areas, and examine the association between COVID-19 incidence and its socio-demographic factors at the village level.

2 | MATERIALS AND METHODS

2.1 | Study site

An ecological study was performed in the Special Capital Region of Jakarta (6.208°S, 106.846°E). It has an area of 623,33 square kilometres and is inhabited by about 10 million people. Administratively, Jakarta has six districts, 44 sub-districts, and 267 villages. One district (Kepulauan Seribu or Thousand Islands) is off the coast of the Java Sea, comprising of 105 small islands with a population of 21,000 people. The population density ranges from 2423 to 18,761 people per square kilometres.

2.2 | Data collection

2.2.1 | COVID-19 data

Aggregated village level data on daily confirmed COVID-19 cases from 1 March 2020 to 30 June 2021 were retrieved from the government's COVID-19 Taskforce website (https://corona.jakarta.go.id/en). COVID-19 case definitions were applied based on the guidelines of the Indonesian Ministry of Health (Ministry of Health of Indonesia, 2020). A total of 458,837 confirmed SARS-CoV-2 positive cases were analyzed.

2.2.2 | Socio-demographic data

Socio-demographic data included the proportion of people aged 20–49 years, the proportion of people aged 50 and above, population density, number of households, the proportion of villages that had access to safe water infrastructure (metered water), and type of occupation for most residents in a village. All of these are the known risk factors of COVID-19 transmission (Jefferies et al., 2020; Nicodemo et al., 2020; Russell et al., 2020; White & Hébert-Dufresne, 2020). Data on the population by age (i.e. the proportion of people aged 20–49 years and over 50 years) at a village level were obtained from the Jakarta Demographic and Civil Registration Office. Data for population density, number of households, access to metered water (categorized into metered, not metered, and borehole), and occupation of residents (categorized into employment in manufacture, trade and retail, transportation and communication, and service sectors) for each village were collected from the Provincial Bureau of Statistics and 2018 Village Survey or Potensi Desa (PODES) conducted by the National Bureau of Statistics. Additionally, data on road density (km per square kilometres) collected from each village were estimated using the GIS. Vector data on the road network were retrieved from the GIS database. Road density was then determined by dividing the length of the road by the area of each polygon (village). Road density is one of the urban factors correlated with the spread of COVID-19 (Kwok et al., 2020; Li et al., 2020).

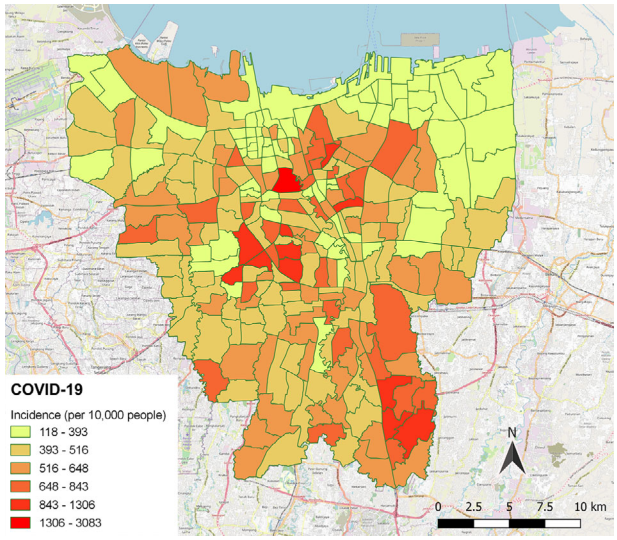

FIGURE 1 Cumulative incidence of COVID-19 in Jakarta (March 2020 to June 2021)

2.3 | Data analysis

2.3.1 | Mapping COVID-19 incidence and socio-demographic characteristics at the village level

In this analysis, villages were used as the spatial unit of analysis. Monthly and cumulative COVID-19 incidence for each village was calculated. Maps of cumulative incidence, monthly incidence of COVID-19, and socio-demographic characteristics were produced using QGIS version 3.2.0-Bonn.

Global spatial clustering

Global Moran's *I* analysis (Moran, 1950) was carried out to examine spatial clustering of the COVID-19 incidence in Jakarta from March 2020 to June 2021. A queen-based spatial contiguity weight matrix was constructed prior to the analysis. The municipality of the Thousand Islands was excluded in the analysis as it influenced spatial weight estimates. The Moran's *I* coefficient ranges from −1 to 1. A positive coefficient indicates a positive spatial autocorrelation, while a negative coefficient means negative spatial autocorrelation. When the coefficient is zero, that means there is no spatial autocorrelation. The significance of Moran's *I* of the COVID-19 incidence was assessed using the Monte Carlo randomization with 999 permutations. A significance value was set at $p < .05$, meaning that the incidence is geographically clustered or dispersed.

2.3.2 | Local spatial clustering

Local Moran's analysis was performed to locate significant high-high (HH) clusters, low-low (LL) clusters, and spatial outliers (low-high and high-low clusters) (Anselin, 1995). HH clusters (later defined as high-risk areas) are areas with high rates surrounded by more areas with high rates and so on. Both global and local analyses were performed using the GeoDA version 1.8 software (Anselin et al., 2006).

Comparison of socio-demographic characteristics of the spatial clusters of COVID-19

Socio-demographical characteristics between identified spatial clusters (HH and LL) were compared. To explore the differences in socio-demographic profiles between clusters, the non-parametric Mann–Whitney *U* test (numerical variables) or Chi-square test (categorical variables) was conducted appropriately. A significance value was set at $p < .05$.

Poisson regression of socio-demographic factors and COVID-19 incidence

To examine the association between COVID-19 cases and socio-demographic characteristics of the neighbourhood, bivariate and multivariate Poisson regressions were modelled with the log of the population of the village as an offset. Five numerical socio-demographic variables (i.e. population density, proportion of people aged 20–49, proportion of people aged above 50, road network density, and number of households) were standardized using *z* (SD) scores. This process was done by rescaling the original values to have a mean of 0 and SD of 1, which enabled for easy interpretation of the model coefficients. The incidence rate ratio (IRR) and 95% confidence interval (CI) were then calculated. A type-1 error rate of 0.05 was used to identify statistically significant results in the analysis. All statistical analyses were performed using the SPSS 21 (IBM Corp).

3 | RESULTS

3.1 | Mapping incidence and socio-demographic factors

The cumulative COVID-19 incidence varied spatially at a village level across Jakarta (Figure 1). Villages with high incidence were identified in the central and south Jakarta. The spatial-temporal maps of monthly incidence showed that the dynamic shift in incidence of COVID-19 at the village level varied across Jakarta over the period studied (Figure 2). An increase in incidence per 10,000 people was gradually observed at village level since September 2020. the highest incidence was observed in Jagakarsa (857/10,000 people) in January 2021. In June 2021, high COVID-19 incidence was observed in north Jakarta including Kapuk (1330/10,000 people), East Cengkareng (1123/10,000 people), and Pejagalan (1098/10,000 people).

The maps of socio-demographic factors demonstrated a variation at the village level (Figure 3). The proportion of the population aged 20–49 years was likely to be randomly distributed across the city. Although the proportion of people aged over 50 was relatively higher in central villages compared to northern villages. The number of households living in slums was relatively high and clustered in the northern villages of Jakarta. Villages in Northern and Central Jakarta had a relatively higher population density and road network density.

3.2 | Global and local measures in clustering of COVID-19 incidence

In general, COVID-19 incidence was significantly clustered across Jakarta over the period studied (Moran's $I = 0.174$; $p = .002$). Monthly observations revealed significant Moran's I coefficients, except in October ($I = 0.013$; $p = .233$) (Table 1). The Moran's I coefficient fluctuated over time, ranging from 0.120 in November up to 0.320 in July.

The local Moran's analysis identified 17 significant HH spatial clusters during the research period, with total population at risk was 330,599 people. However, the analysis also found 41 LL clusters across Jakarta. The significant HH clusters belonged to 10 sub-districts including Cipayung (7 villages), Kebayoran Baru (1), Makassar (1), Cempaka Putih (1), Setia Budi (2), Sawah Besar (1), Gambir (1), Tanah Abang (1), Kramat Jati (1), and Senen (1). Spatial clusters of HH were dynamically shifted over time (Figure 4). A small set of HH clusters were steadily concentrated in central Jakarta during the period studied. In addition, the HH clusters dynamically emerged in the south of Jakarta. These HH clusters newly emerged and expanded significantly in the south of Jakarta during December 2020 and January 2021. This HH cluster re-emerged in June 2021. A high number of HH clusters were observed in July 2020 (23 villages) and January 2021 (24 villages) with a total population at risk was 617,055 and 752,726 people, respectively. A low number of HH clusters was reported in October 2020 (six villages) with a total of 107,395 people at risk (Table 2).

3.3 | Socio-demographic profiles of spatial clusters of COVID-19

Table 3 summarizes the socio-demographic differences of the identified HH and LL clusters of COVID-19 in Jakarta. Compared to the LL clusters, the HH clusters were likely to have lower people density ($p < .001$), have less road network density ($p < .05$), have fewer households ($p < .001$), a small proportion of the village had access to metered water (17.6%, $p < .001$), and most of its residents had a primary occupation in service sectors (47%, $p = .001$). While it is not statistically different ($p > .05$), the HH cluster was likely to have a high proportion of people aged above 50 years (20.6%) and a low proportion of people aged 20–49 years (48.7%), compared to the LL clusters.

3.4 | The association between COVID-19 incidence and socio-demographic characteristics at the village level

The bivariate and multivariate Poisson regression analyses are presented in Table 4. In the bivariate analysis, an increased SD of the proportion of people aged 20–49 was associated with a 5.9% (95% CI: 5.6%–6.2%, $p < .001$) increase in the risk of COVID-19. Each SD increase in the number of households was associated with a 15.3% (95% CI: 15%–15.6%, $p < .001$) increase in the COVID-19 incidence. The risk of COVID-19 incidence was 8.7% (95% CI: 8%–9.3%, $p < .001$) higher in villages where most households used borehole water for bathing or washing. A village where most people worked at service industries was likely to have higher incidence by 23.5% (95% CI: 22.4%–24.6%, $p < .001$) compared to the village where people predominantly worked at manufacturing.

In the multivariate analysis, an increased SD of the population aged 20–49 and above 50 was associated with a 1.6% (95% CI: 1.2%–1.9%) and 4.5% (9% CI: 4.1%–5%) increase in the risk of COVID-19, respectively. Each SD increase in number of households was associated with a 19.6% increase in the incidence of COVID-19 (95% CI: 19.3%–20%, $p < .001$). The COVID-19 incidence was 3.7% (95% CI = 36.4%–39.1%, $p < .001$) higher in the communities that used borehole water compared to metered water. Further, the risk was 1.2 and 1.4 times higher in the community where most people worked in transportation, communication, and service industries, respectively, after controlling for other socio-demographic factors.

4 | DISCUSSION

All provinces in Indonesia have been impacted by the pandemic, of which Jakarta is the most hit region by COVID-19 (Eryando et al., 2020). In this ecological study, we explored spatial variation in the incidence of COVID-19 at the village level across the city of Jakarta and summarizes socio-demographical features of the high-risk villages by using spatial statistics approach. This study is the first one investigating spatial variations of COVID-19 incidence in relation to socio-demographic factors across Jakarta. The study demonstrates dynamic patterns in the spatial clustering and distribution of the COVID-19 hotspots over the 16 months (March 2020 to June 2021). Further, our study reveals that socio-demographical variation at neighbourhood level explained the heterogeneity in risk of COVID-19 across the city. These findings could help inform better strategies for effective targeted intervention to curb the transmission of COVID-19.

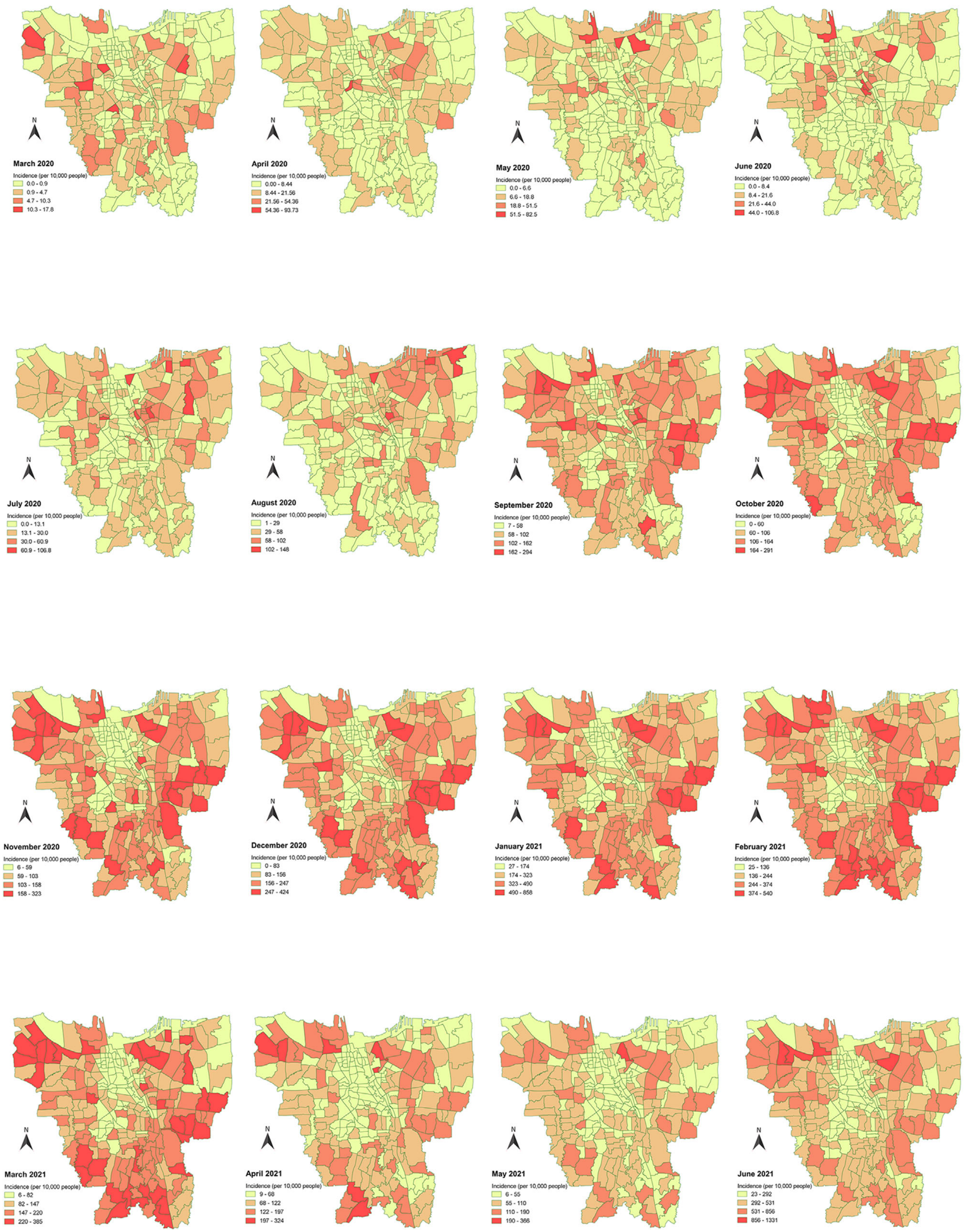

FIGURE 2 Monthly incidence of COVID-19 at the village level in Jakarta (March 2020 to June 2021)

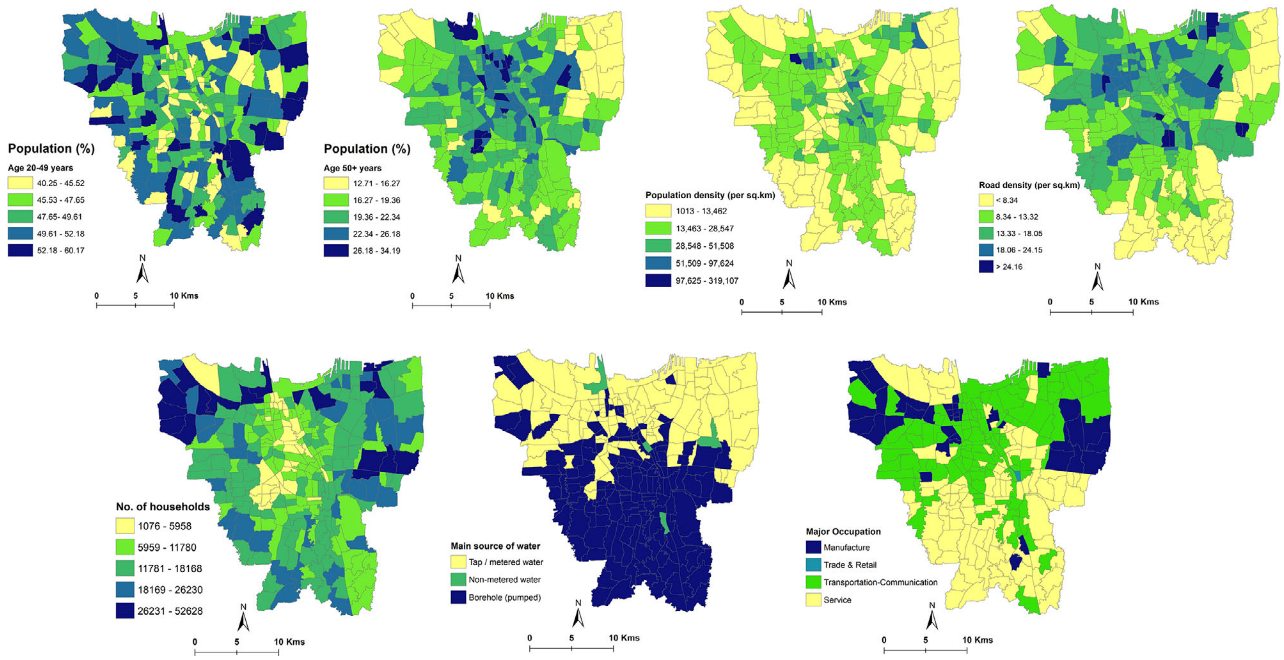

FIGURE 3 Map of socio-demographic characteristics at the village level in Jakarta. Maps were produced using QGIS version 3.2.0-Bonn

TABLE 1 Monthly spatial clustering of COVID-19 incidence in Jakarta, March 2020 to June 2021

Month	*I*	SD	*z*-Values	*p*-Values
2020				
March	0.139	0.022	6.5	.001
April	0.262	0.036	7.472	.001
May	0.165	0.036	4.638	.001
June	0.174	0.025	6.978	.002
July	0.320	0.035	9.103	.001
August	0.155	0.032	4.916	.001
September	0.158	0.033	4.863	.001
October	0.013	0.029	0.610	.233
November	0.120	0.035	0.349	.002
December	0.243	0.035	6.953	.001
2021				
January	0.283	0.036	7.885	.001
February	0.238	0.035	6.900	.001
March	0.226	0.036	6.243	.001
April	0.138	0.036	3.942	.003
May	0.132	0.034	3.934	.005
June	0.153	0.033	4.598	.002
Overall	0.174	0.035	5.040	.002

The heterogeneity in incidence over space and time was likely due to the impact of behavioural changes (e.g. face mask use and stay-at-home order) and the implementation of control measures (e.g. travel ban and testing). It is important to note that the trends in COVID-19 incidence was fluctuated over the period. We observed an increased incidence in December to February, which was likely driven by Christmas and New Year holidays (Figure S1). Further, an increased incidence might indicate the emergence of new SARS-CoV-2 variants in the population, yet more evidence is needed. In addition to promoting adequate health protocols, lockdowns or large-scale social restrictions seem to have contributed to reducing the transmission. Our visual examination of the trends in people's mobility (indicated by changes in travel distance) and in the number of daily cases indicated the effect of mobility restriction on COVID-19 cases (Supporting information). People's mobility was reduced significantly as lockdown was implemented. This has helped flatten the epidemic curve at some points (October 2020 and March 2021). The impact of such intervention on the spatial distribution incidence is reflected by the reduced power and significance of spatial clustering (as indicated by a reduced *I* coefficient). However, further investigation is necessary to examine the impact of public health interventions and change in behaviour and mobility due to the spread of COVID-19 at the population level.

This study identified 17 high-risk spatial clusters of COVID-19. A small set of high-risk villages was found in central Jakarta over the period of study. Interestingly, the study reveals the emergence of high-risk villages on the south border of Jakarta from December 2020 to March 2021. This could be partly driven by underlying socio-demographical factors (i.e. population structure, mobility, behavioural, and activities). The comparative analysis showed that these high-risk villages were different from low-risk villages. High-risk areas are characterized by a mixture of commercial and office activities which allow higher mobility and interactions, thus leading to a greater risk of SARS-CoV-2 transmission. Similar clustering phenomena of COVID-19 cases

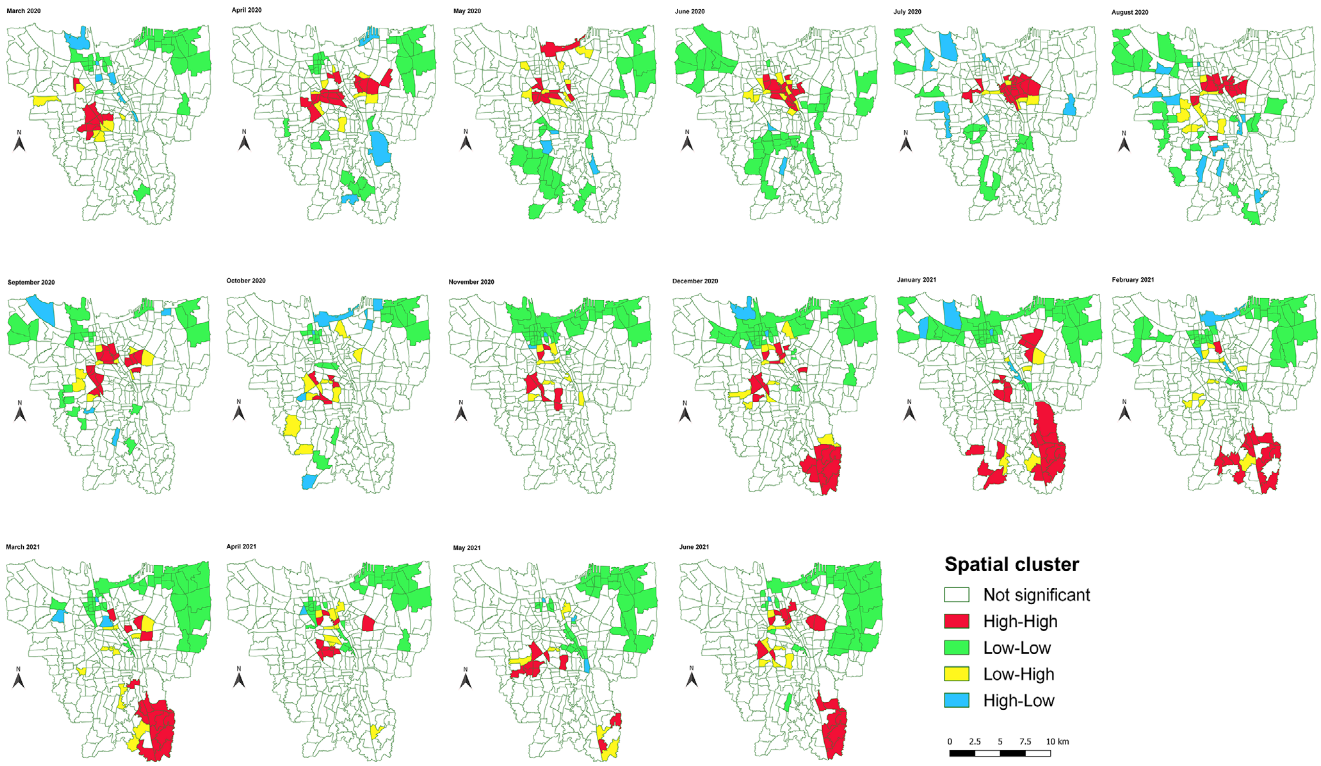

FIGURE 4 Spatiotemporal spatial clusters of COVID-19 in Jakarta as identified by local Moran analysis (March 2020 to June 2021). Maps were created using QGIS version 3.2.0-Bonn

have also been reported in other countries (Kim & Castro, 2020; Yang et al., 2020).

This study suggested that area-level incidence was associated with population composition higher incidence was likely occurred in that area with a high proportion of senior residents. The elderly, particularly those with comorbidities, were reported to be at higher risk of contracting severe illnesses (Jefferies et al., 2020; Liu et al., 2020; Russell et al., 2020). Our finding also indicates that area-level incidence was associated with the proportion of people aged 20–49 years and COVID-19 incidence. Young and productive populations are relatively highly mobile (Kronbichler et al., 2020; Lai et al., 2020; Tan et al., 2020) that can facilitate transmission. The findings highlight that appropriate non-pharmaceutical intervention (NPI) should be targeted towards these populations (e.g. promoting face masks and social distancing). As COVID vaccines are now available, putting the elderly and active population on the top of the list is recommended. Besides, routine health monitoring towards these elderly population should be improved. Local health authorities can utilize digital tools such as telehealth and telemedicine services to facilitate access to care and continue providing care for the community during the pandemic.

These spatial variations of COVID-19 incidence appear to be driven by different socio-demographic profiles at the village level. Villages, where most residents work in the transportation, communication, and public service sectors, were likely to have a higher incidence of COVID-19; this finding is consistent with other studies (Koh, 2020; Kwok et al., 2020). Some occupations, such as healthcare workers (HCWs) and public staff, have been more at risk of COVID-19 infection (Gholami et al., 2021; Koh, 2020). These workers are likely to have frequent contact with people and thus potentially increase the risk of being exposed to the circulating virus. Local epidemiological reports have documented that most of the COVID-19 transmission occurred in workplaces and households. Thus, robust health and safety protocols should be applied.

In this study, we included road network density to reflect connectivity. This is a known factor associated with the transmission of respiratory infection (Vu et al., 2013). Our finding is inconsistent with a study in Hong Kong that reported a positive correlation between road network density and COVID-19 incidence (Kwok et al., 2020). One possible reason is that while in some way dense road networks may likely promote higher people's mobility, several unmeasured factors might have also indirectly contributed to lowering the risk of transmission (e.g. mobility restrictions, face mask use, stay-at-home order and traffic manipulation).

A higher number of families within the village were associated with the increased COVID-19 incidence. The effect of overcrowding could be the reason for this situation, which has also been reported by other studies (Ahmad et al., 2020; Daras et al., 2021). Further, in areas (e.g. the southern part of Jakarta) where access to tapped water is limited, the incidence were likely higher. Studies showed that poor housing and inequality in access to basic services are known social factors associated with COVID-19 (Ahmad et al., 2020; Das et al., 2020). This suggests that improving access to basic services is essential to facilitate a proper hygiene behaviour in the community.

Unexpectedly, an inverse association between COVID-19 incidence and population density was found. This is contradictory to previous

TABLE 2 Summary of monthly spatial clusters detected by local Moran analysis and estimated population at risk of COVID-19 in Jakarta from March 2020 to June 2021

Months	Spatial cluster	COVID-19 cases	Number of villages	Estimated population-at-risk
2020				
March	High-high	37	10	109,244
	Low-low	16	27	1,258,650
	Low-high	7	7	193,355
	High-low	15	6	121,357
April	High-high	388	18	500,588
	Low-low	141	27	1,031,725
	Low-high	49	9	225,048
	High-low	45	4	128,130
May	High-high	220	12	331,142
	Low-low	91	20	909,340
	Low-high	22	8	234,708
	High-low	18	3	68,598
June	High-high	470	15	332,141
	Low-low	197	34	1,636,460
	Low-high	62	10	242,277
	High-low	20	2	44,109
July	High-high	912	23	617,055
	Low-low	172	15	492,610
	Low-high	68	7	161,792
	High-low	211	7	296,492
August	High-high	1072	19	467,808
	Low-low	728	23	1,197,522
	Low-high	232	11	298,632
	High-low	508	10	324,561
September	High-high	1755	17	349,721
	Low-low	1916	25	1,016,095
	Low-high	644	9	306,827
	High-low	373	3	118,115
October	High-high	400	6	107,395
	Low-low	1477	21	814,962
	Low-high	549	8	253,210
	High-low	641	6	171,584
November	High-high	503	9	142,087
	Low-low	2763	38	1,581,976
	Low-high	324	7	174,004
	High-low	108	2	37,514
December	High-high	2775	19	408,031
	Low-low	4235	39	1,748,645
	Low-high	894	10	274,986
	High-low	436	3	80,896

(Continues)

TABLE 2 (Continued)

Months	Spatial cluster	COVID-19 cases	Number of villages	Estimated population-at-risk
2021				
January	High-high	8285	24	752,723
	Low-low	9281	44	1,899,924
	Low-high	1614	5	222,719
	High-low	1265	5	144,483
February	High-high	4004	14	442,744
	Low-low	7213	39	1,655,714
	Low-high	1244	10	221,546
	High-low	474	3	60,918
March	High-high	3086	9	525,929
	Low-low	3739	31	1,549,970
	Low-high	1203	8	361,464
	High-low	340	1	62,981
April	High-high	620	9	174,477
	Low-low	1869	31	1,307,569
	Low-high	289	8	153,581
	High-low	67	1	21,771
May	High-high	592	10	167,750
	Low-low	1764	37	1,570,404
	Low-high	226	5	130,524
	High-low	159	3	65,569
June	High-high	6074	17	368,813
	Low-low	11,138	31	1,490,754
	Low-high	1657	9	179,574
	High-low	140	1	9642

TABLE 3 Socio-demographic profiles of the identified high- and low-risk spatial clusters of COVID-19 in Jakarta

	Spatial cluster		
Characteristics	HH ($n = 17$)	LL ($n = 41$)	*p*-Values
Population density per km^2 #	9,138 (5104–11251)	30,799 (12625–46,095)	<.001
% people aged 20–49 years#	48.71 (47.76–50.15)	49 (47.37-50.70)	.704
% people aged 50+ years#	20.65 (17.58-22.65)	19.63 (15.13–22.29)	.413
Road network density (in km/km^2) #	11.66 (6.53–15.62)	15.12 (10.08–19.23)	.048
Number of households#	7,597 (3038–9779)	15211 (8,294–22,552)	<.001
% of village had access to metered water for bathing/washing	17.6	70.7	<.001
% of village where most of residents working in:			.001
Manufacture	5.9	36.6	
Trade and retail	0	2.4	
Transport and communication	47.1	53.7	
Service	47.1	7.3	

Abbreviations: 95% CI, 95% confidence interval; HH, high-high; LL, low-low.
#Results expressed as mean (95% CI).

TABLE 4 Associations between COVID-19 incidence and socio-demographic factors at the village level

	Bivariate			Multivariate		
		95% CI			95% CI	
	IRR	Lower	Upper	IRR	Lower	Upper
Population density (person/km^2)	0.905***	0.901	0.908	0.937***	0.933	0.940
People aged 20–49 years (%)	1.059***	1.056	1.062	1.016***	1.012	1.019
People aged 50+ years (%)	0.947***	0.944	0.950	1.045***	1.041	1.050
Road network density (in km/km^2)	0.952***	0.950	0.955	0.986***	0.982	0.989
Number of households	1.153***	1.150	1.156	1.196***	1.193	1.200
Main source of water for bathing/washing						
Metered water	1			1		
Non-metered water	1.015	0.992	1.038	0.824***	0.805	0.843
Borehole	1.087***	1.080	1.093	1.037***	1.030	1.044
Village where most of residents working in						
Manufacture	1			1		
Trade and retail	0.696***	0.650	0.746	0.921*	0.859	0.986
Transport and communication	1.044***	1.035	1.053	1.239***	1.228	1.251
Service	1.235***	1.224	1.246	1.378***	1.364	1.391

Abbreviations: 95% CI = 95% confidence interval; IRR = incidence rate ratio.
*$p < .1$.
**$p < .05$.
***$p < .01$.

studies (You et al., 2020; Whittle and Diaz-Artiles, 2020) although one recent study in China suggested insignificant association between both variables (Xiong et al., 2020). This inconsistent finding may be partly explained by the local behaviours and the difference in spatial extent. However, this finding needs further investigation.

This study has some limitations. First, the analysis was based on the cases reported by health facilities, which are prone to many uncertainties. For example, surveillance and testing capacity had not been optimally established during the early phase of the pandemic. Delayed and incomplete epidemiological reports due to lack of technical and human resources had also challenged data accuracy (Aisyah et al., 2020). Furthermore, a substantial proportion of asymptomatic cases might have not been detected at the beginning of the pandemic, and thus the real burden of COVID-19 might have been underreported. Second, we did not consider meteorological variables in the analysis despite its importance on COVID-19 (Tosepu et al., 2020). Given Jakarta is a relatively small area, we assumed that meteorological variability at the village level might not be significantly heterogeneous. Thus, it might be not a good predictor for area-level COVID incidence. Also, limited by data availability, we could not include individual and behavioural risks (e.g. chronic conditions, face mask use, and smoking). Lastly, all available socio-demographic data at the village level were obtained from past censuses and reports, which may not reflect the actual condition. Further local scale population-based study is encouraged to better understand the epidemiology of COVID-19 in Jakarta. Despite these limitations, this study highlights the benefit of GIS and available administrative data in understanding the distribution and factors associated with COVID-19 incidence in Jakarta.

Mapping the disease spread and understanding the underlying drivers would help locate the source of transmission to immediately suppress COVID-19 transmission. Public health measures, such as localizing hotspots, identifying, and instructing high-risk populations for quarantine, promoting improved NPIs, as well as improving health systems, could be appropriately applied to the identified high-risk areas. Moreover, it is also important to regularly monitor the LL clusters in the city as they have the potential to become new hotspots for COVID-19. They require attention for detection and control of COVID-19 transmission, for example, by conducting more frequent early detection, contact tracing, clinical management, and healthcare delivery.

5 | CONCLUSIONS

To sum up, the distribution of COVID-19 incidence was spatially heterogenous at the village level throughout Jakarta. From March 2020 to June 2021, 17 significant hotspots existed mainly in the central part of the city. The geographical variations were significantly related to the social inequalities that exist in the city. Area-specific and targeted intervention measures to those high-risk villages and vulnerable communities, in addition to NPIs, adequate testing strategies, comprehensive contact tracing, equal access to treatment and social protection, could help mitigate the spread of the virus and reduce the risk amongst vulnerable communities.

ACKNOWLEDGEMENT

The authors want to thank the Jakarta Provincial Government for providing the data. The study was supported by the National Institute of Health Research and Development, Indonesian Ministry of Health.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ETHICS STATEMENT

The research protocol of this study was reviewed and approved by the Health Research Ethics Committee of National Institute of Health Research and Development (number: LB.02.01/2/KE.345/2020). All COVID-19 cases were anonymized and aggregated to village level. No personal identifiers were presented, and maps presented in this paper did not identify respondents' precise addresses.

AUTHOR CONTRIBUTIONS

Study conceptualization: Pandji Wibawa Dhewantara and Tities Puspita. *Statistical methodology*: Pandji Wibawa Dhewantara, Rina Marina, Tities Puspita, and Doni Lasut. *Formal analysis*: Pandji Wibawa Dhewantara and Tities Puspita. *Project administration*: Pandji Wibawa Dhewantara, Tri Wahono, and Andri Ruliansyah. *Data curation and validation*: Pandji Wibawa Dhewantara, Tities Puspita, Wawan Ridwan, Andri Ruliansyah, and Muhammad Umar Riandi. *Funding*: Pandji Wibawa Dhewantara and Tri Wahono. *Writing (original draft)*: Pandji Wibawa Dhewantara and Tities Puspita. *Writing (review and editing)*: all authors.

DATA AVAILABILITY STATEMENT

The data of daily village-level cases are free available on the COVID-19 website published by Jakarta's COVID-19 Taskforce Team (https://corona.jakarta.go.id/en).

ORCID

Pandji Wibawa Dhewantara https://orcid.org/0000-0002-9461-0648

REFERENCES

Ahmad, K., Erqou, S., Shah, N., Nazir, U., Morrison, A. R., Choudhary, G., & Wu, W.-C. (2020). Association of poor housing conditions with COVID-19 incidence and mortality across US counties. *Plos One*, *15*, e0241327–e0241327.

Aisyah, D. N., Mayadewi, C. A., Diva, H., Kozlakidis, Z., SISWANTO, & ADISASMITO, W. (2020). A spatial-temporal description of the SARS-CoV-2 infections in Indonesia during the first six months of outbreak. *Plos One*, *15*, e0243703.

Anselin L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, *27*, 93–115.

Anselin L., Syabri, I., & Kho, Y. (2006). GeoDa, an introduction to spatial data analysis. *Geographical Analysis*, *38*, 5–22.

Cabinet Secretariate of Republic of Indonesia. (2020). *Health minister regulation on guidelines to propose large-scale social restrictions amid COVID-19 pandemic*. https://setkab.go.id/en/health-minister-signs-regulation-on-guidelines-to-propose-large-scale-social-restrictions-amid-covid-19-pandemic/.

Chakraborty, J. (2020). Social inequities in the distribution of COVID-19: An intra-categorical analysis of people with disabilities in the U.S. *Disability and Health Journal*, *14*, 101007.

Chan, J. F.-W., Yuan, S., Kok, K.-H., To, K. K.-W., Chu, H., Yang, J., Xing, F., Liu, J., Yip, C. C.-Y., & Poon, R. W.-S. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. *The Lancet*, *395*, 514–523.

Daras, K., Alexiou, A., Rose, T. C., Buchan, I., Taylor-Robinson, D., & Barr, B. (2021). How does vulnerability to COVID-19 vary between communities in England? Developing a Small Area Vulnerability Index (SAVI). *Journal of Epidemiology and Community Health*, *75*(8), 729–734.

Das, A., Ghosh, S., Das, K., Basu, T., Das, M., & Dutta, I. (2020). Modeling the effect of area deprivation on COVID-19 incidences: A study of Chennai megacity, India. *Public Health*, *185*, 266–269.

Eryando, T., Sipahutar, T., & Rahardiantoro, S. (2020). The risk distribution of COVID-19 in Indonesia: A spatial analysis. *Asia-Pacific Academic Consortium for Public Health*, *32*(8), 450–452.

Fletcher-Lartey, S. M., & Caprarelli, G. (2016). Application of GIS technology in public health: Successes and challenges. *Parasitology*, *143*, 401–415.

Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F., & Billa, L. (2020). Spatial analysis and GIS in the study of COVID-19. A review. *The Science of the Total Environment*, *739*, 140033.

Gholami, M., Fawad, I., Shadan, S., Rowaiee, R., Ghanem, H., Hassan Khamis, A., & Ho, S. B. (2021). COVID-19 and healthcare workers: A systematic review and meta-analysis. *International Journal of Infectious Diseases*, *104*, 335–346.

Government of the Special Capital Region of Jakarta. (2020). *Governor decree no. 380/2020 on implementation of large-scale social restriction to control COVID-19 in Jakarta*.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., & Gu, X. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, *395*, 497–506.

Huang, R., Liu, M., & Ding, Y. (2020). Spatial-temporal distribution of COVID-19 in China and its prediction: A data-driven modeling analysis. *Journal of Infection in Developing Countries*, *14*, 246–253.

Indonesia COVID-19 Task Force. (2021). *Peta Sebaran COVID-19 Indonesia*. https://covid19.go.id/peta-sebaran

Jakarta COVID-19 Taskforce. (2021). *Jakarta COVID Dashboard*. https://corona.jakarta.go.id/id/data-pemantauan

Jefferies, S., French, N., Gilkison, C., Graham, G., Hope, V., Marshall, J., Mcelnay, C., Mcneill, A., Muellner, P., Paine, S., Prasad, N., Scott, J., Sherwood, J., Yang, L., & Priest, P. (2020). COVID-19 in New Zealand and the impact of the national response: A descriptive epidemiological study. *Lancet Public Health*, *5*(11), e612–e623.

Kim, S., & Castro, M. C. (2020). Spatiotemporal pattern of COVID-19 and government response in South Korea (as of May 31, 2020). *International Journal of Infectious Diseases*, *98*, 328–333.

Koh, D. (2020). Occupational risks for COVID-19 infection. *Occupational Medicine*, *70*, 3–5.

kronbichler, A., Kresse, D., Yoon, S., Lee, K. H., Effenberger, M., & Shin, J. I. (2020). Asymptomatic patients as a source of COVID-19 infections: A systematic review and meta-analysis. *International Journal of Infectious Diseases*, *98*, 180–186.

Kutti-Sridharan, G., Vegunta, R., Vegunta, R., Mohan, B. P., & Rokkam, V. R. P. (2020). SARS-CoV2 in different body fluids, risks of transmission, and preventing COVID-19: A comprehensive evidence-based review. *International Journal of Preventive Medicine*, *11*, 97.

Kwok, C. Y. T., Wong, M. S., Chan, K. L., Kwan, M. P., Nichol, J. E., Liu, C. H., Wong, J. Y. H., Wai, A. K. C., Chan, L. W. C., Xu, Y., Li, H., Huang, J., & Kan, Z. (2020). Spatial analysis of the impact of urban geometry and socio-demographic characteristics on COVID-19, a study in Hong Kong. *The Science of the Total Environment*, *764*, 144455.

Lai, C.-C., Liu, Y. H., Wang, C.-Y., Wang, Y.-H., Hsueh, S.-C., Yen, M.-Y., Ko, W.-C., & Hsueh, P.-R. (2020). Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARSCoV-2): Facts and myths. *Journal of Microbiology, Immunology and Infection*, *53*(3), 404–412.

Lakhani, A. (2020). Which Melbourne metropolitan areas are vulnerable to COVID-19 based on age, disability and access to health services? Using spatial analysis to identify service gaps and inform delivery. *Journal of Pain and Symptom Management*, *60*(1), e41–e44.

Li, X., Jia, T., Peng, R., Fu, X., & Zou, Y. (2020). Associating COVID-19 severity with urban factors: A case study of Wuhan. *International Journal of Environmental Research and Public Health*, *17*, 6712.

Liu, J., Liu, Z., Jiang, W., Wang, J., Zhu, M., Song, J., Wang, X., Su, Y., Xiang, G., Ye, M., Li, J., Zhang, Y., Shen, Q., Li, Z., Yao, D., Song, Y., Yu, K., Luo, Z., & Ye, L. (2020). Clinical predictors of COVID-19 disease progression and death: Analysis of 214 hospitalized patients from Wuhan, China. *Clinical Respiratory Journal*, *15*(3), 293–309.

Luo, Y., Yan, J., & Mcclure, S. (2020). Distribution of the environmental and socioeconomic risk factors on COVID-19 death rate across continental USA: A spatial nonlinear analysis. *Environmental Science and Pollution Research International*, 1–13.

Ministry of Health of Indonesia (2020). *Pedoman Pencegahan dan Pengendalian Coronavirus Disease (COVID-19) Revisi 4 [Guidelines on COVID-19 Prevention and Control Rev.4]*. Ministry of Health of Indonesia.

Mizumoto, K., Kagaya, K., Zarebski, A., & Chowell, G. (2020). Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance*, *25*, 2000180.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*, 17–23.

Nakada, L. Y. K., & Urban, R. C. (2020). COVID-19 pandemic: environmental and social factors influencing the spread of SARS-CoV-2 in São Paulo, Brazil. *Environmental Science and Pollution Research International*, *28*(30), 40322–40328.

Nicodemo, C., Barzin, S., Cavalli, N., Lasserson, D., Moscone, F., Redding, S., & Shaikh, M. (2020). Measuring geographical disparities in England at the time of COVID-19: results using a composite indicator of population vulnerability. *BMJ Open*, *10*, e039749.

Russell, F. M., Wang, A., Ehrman, R. R., Jacobs, J., Croft, A., & Larsen, C. (2020). Risk factors associated with hospital admission in COVID-19 patients initially admitted to an observation unit. *American Journal of Emergency Medicine*, *46*, 339–343.

Sun, Y., Hu, X., & Xie, J. (2021). Spatial inequalities of COVID-19 mortality rate in relation to socioeconomic and environmental factors across England. *The Science of the Total Environment*, *758*, 143595.

Tan, F., Wang, K., Liu, J., Liu, D., Luo, J., & Zhou, R. (2020). Viral transmission and clinical features in asymptomatic carriers of SARS-CoV-2 in Wuhan, China. *Frontiers in Medicine*, *7*, 547.

Tosepu, R., Gunawan, J., Effendy, D. S., Ahmad, O. A. I., Lestari, H., Bahar, H., & Asfian, P. (2020). Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia. *The Science of the Total Environment*, *725*, 138436.

Vu, V. H., Le, X. Q., Pham, N. H., & Hens, L. (2013). Application of GIS and modelling in health risk assessment for urban road mobility. *Environmental Science and Pollution Research International*, *20*, 5138–5149.

White, E. R., & Hébert-Dufresne, L. (2020). State-level variation of initial COVID-19 dynamics in the United States. *Plos One*, *15*, e0240648.

Whittle, R. S., & Diaz-Artiles, A. (2020). An ecological study of socioeconomic predictors in detection of COVID-19 cases across neighborhoods in New York City. *BMC Medicine*, *18*, 271–271.

WHO. (2020). *Novel Coronavirus (2019-nCoV) situation reports-22*. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf2020

WHO. (2021). *Weekly Epidemiological Update - 16 February 2021*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

Xiong, Y., Wang, Y., Chen, F., & Zhu, M. (2020). Spatial statistics and influencing factors of the COVID-19 epidemic at both prefecture and county levels in Hubei Province, China. *International Journal of Environmental Research and Public Health*, *17*, 3903.

Yan, Y., Shin, W. I., Pang, Y. X., Meng, Y., Lai, J., You, C., Zhao, H., Lester, E., Wu, T., & Pang, C. H. (2020). The first 75 days of novel coronavirus (SARS-CoV-2) outbreak: Recent advances, prevention, and treatment. *International Journal of Environmental Research and Public Health*, *17*, 2323.

Yang, W., Deng, M., Li, C., & Huang, J. (2020). Spatio-temporal patterns of the 2019-nCoV epidemic at the county level in Hubei province, China. *International Journal of Environmental Research and Public Health*, *17*, 2563.

You, H., Wu, X., & Guo, X. (2020). Distribution of COVID-19 morbidity rate in association with social and economic factors in Wuhan, China: Implications for urban development. *International Journal of Environmental Research and Public Health*, *17*, 3417.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., & Lu, R. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*, *382*(8), 727–733.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Dhewantara, P. W., Puspita, T., Marina, R., Lasut, D., Riandi, M. U., Wahono, T., Ridwan, W., & Ruliansyah, A. (2022). Geo-clusters and socio-demographic profiles at village-level associated with COVID-19 incidence in the metropolitan city of Jakarta: An ecological study. *Transboundary and Emerging Diseases*, *69*, e362–e373. https://doi.org/10.1111/tbed.14313