Contents lists available at ScienceDirect

Public Health

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

Short Communication

Modeling the effect of area deprivation on COVID-19 incidences: a study of Chennai megacity, India

A. Das [a], S. Ghosh [b, *], K. Das [a], T. Basu [a], M. Das [a], I. Dutta [a]

[a] Department of Geography, University of Gour Banga, Malda, India
[b] Department of Geography, Kazi Nazrul University, Asansol, India

A B S T R A C T

Objectives: Socio-economic inequalities may affect coronavirus disease 2019 (COVID-19) incidence. The goal of the research was to explore the association between deprivation of socio-economic status (SES) and spatial patterns of COVID-19 incidence in Chennai megacity for unfolding the disease epidemiology.
Study design: This is an ecological (or contextual) study for electoral wards (subcities) of Chennai megacity.
Methods: Using data of confirmed COVID-19 cases from May 15, 2020, to May 21, 2020, for 155 electoral wards obtained from the official website of the Chennai Municipal Corporation, we examined the incidence of COVID-19 using two count regression models, namely, Poisson regression (PR) and negative binomial regression (NBR). As explanatory factors, we considered area deprivation that represented the deprivation of SES. An index of multiple deprivations (IMD) was developed to measure the area deprivation using an advanced local statistic, geographically weighted principal component analysis. Based on the availability of appropriately scaled data, five domains (i.e., poor housing condition, low asset possession, poor availability of WaSH services, lack of household amenities and services, and gender disparity) were selected as components of the IMD in this study.
Results: The hot spot analysis revealed that area deprivation was significantly associated with higher incidences of COVID-19 in Chennai megacity. The high variations (adjusted R^2: 72.2%) with the lower Bayesian Information Criteria (BIC) (124.34) and Akaike's Information Criteria (AIC) (112.12) for NBR compared with PR suggests that the NBR model better explains the relationship between area deprivation and COVID-19 incidences in Chennai megacity. NBR with two-sided tests and $P < 0.05$ were considered statistically significant. The outcome of the PR and NBR models suggests that when all other variables were constant, according to NBR, the relative risk (RR) of COVID-19 incidences was 2.19 for the wards with high housing deprivation or, in other words, the wards with high housing deprivation having 119% higher probability (RR = $e^{0.786}$ = 2.19, 95% confidence interval [CI] = 1.98 to 2.40), compared with areas with low deprivation. Similarly, in the wards with poor availability of WaSH services, chances of having COVID-19 incidence was 90% higher than in the wards with good WaSH services (RR = $e^{0.642}$ = 1.90, 95% CI = 1.79 to 2.00). Spatial risks of COVID-19 were predominantly concentrated in the wards with higher levels of area deprivation, which were mostly located in the northeastern parts of Chennai megacity.
Conclusions: We formulated an area-based IMD, which was substantially related to COVID-19 incidences in Chennai megacity. This study highlights that the risks of COVID-19 tend to be higher in areas with low SES and that the northeastern part of Chennai megacity is predominantly high-risk areas. Our results can guide measures of COVID-19 control and prevention by considering spatial risks and area deprivation.

Introduction

Coronavirus disease 2019 (COVID-19) is an epidemic illness that was discovered in Wuhan of China at the end of 2019. Shortly after, it rapidly spread worldwide to emerge as a 'Public Health

* Corresponding author. Tel.: +91 8001413335.
E-mail addresses: Arijit3333@gmail.com (A. Das), sasankaghoshsg@gmail.com (S. Ghosh), geokinkar@gmail.com (K. Das), basu.tirthankar89@gmail.com (T. Basu), dasmanob631@gmail.com (M. Das), ipsitadutta25@gmail.com (I. Dutta).

Emergency of International Concern.' Subsequently, the World Health Organization declares COVID-19 as a global pandemic. As of May 22, 2020, COVID-19 has affected about 4.99 million people and claimed more than 327,738 deaths globally,[1] and these figures are increasing every day. In India, the first COVID-19 case was reported on January 30, 2020, in Kerala, and then, the number of cases increased gradually in all states and union territories except Daman and Diu, Nagaland, and Lakshadweep.[2] COVID-19 has mostly affected the urban areas, particularly the megacities of India, which became the epicenters of the COVID-19 spread. The geographic distribution of COVID-19 incidences showed that the disease was not uniformly affecting all parts of the Indian megacities but led to spatial clustering of cases.[3]

The existence of an inverse relationship between the socio-economic status (SES) of populations and higher incidences of lower respiratory tract infection among populations in society was well recognized.[4] Despite the preliminary evidence of social inequality in COVID-19 incidences and the possibility of SES deprivation as a major contributor to COVID-19, research on the relationship between SES deprivation and COVID-19 incidences is inadequate. It has, therefore, become imperative that the role of SES deprivation in COVID-19 incidences be evaluated. To fill up the existing research gap, this study attempted to provide scientific evidence about the influence of SES deprivation on spatial clustering of COVID-19 hot spots in Chennai megacity.

In the absence of individual data, we have used ecological (or contextual) measures of SES deprivation to describe the inequalities in COVID-19 incidences. In this article, we have modified and improved the index of multiple deprivations (IMD) developed by Baud et al.[5] for Chennai megacity by incorporating specific indicators of SES deprivation that affect COVID-19, such as non-availability to a drinking water source within premises or not having a toilet inside houses. The households (HHs) without having the availability of a drinking water source within premises were compelled to collect drinking water from community standposts and tube wells. Similarly, the HH not having a toilet inside houses also was compelled to use a community toilet. These situations will certainly reduce compliance with social distancing. Limited availability of community latrines, standposts, and tube wells certainly increases the chance of community transmission of the virus. Examining the spatial inequalities in COVID-19 incidences in Chennai megacity will provide important insights into the spatial pattern of the disease. The development of an IMD representing SES disadvantage would explore the linkages between living environment deprivation and COVID-19 incidences. Ultimately, the outcome of this study will provide useful insights to policymakers for targeted interventions to combat the COVID-19 pandemic.

Methods

The study area is Chennai megacity (13.04°N—80.17°E), which is the fourth largest metropolis in India (after Mumbai, New Delhi, and Kolkata) with a population of 10.2 million. It is the most important urban center in the southeast coastal region of India, which has a typical subtropical, hot, humid, monsoon climate classified as Aw (tropical wet and dry) as per the Köppen climate classification. With mild and moderate winters and very hot summers, the average air temperature ranges from 21 to 35 °C (70—95 °F), and relative humidity varies from 45% to 95%.

The first COVID-19 case in Chennai was detected on March 9, 2020, and later, community transmission has taken place rapidly. The number of confirmed COVID-19 cases in 155 wards of Chennai megacity reported in this study was collected from the official

website of Greater Chennai online database releases from May 15, 2020, to May 21, 2020. Ward-wise confirmed cases during this same period were also obtained from The News Minute coronavirus data repository.[6] The two data sets were compared to ensure consistency of COVID-19 incidences before executing the statistical analysis.

Because SES is a complex and multidimensional phenomenon, in the absence of individual data, which were not available at this pandemic situation, we developed well-recognized ecological (or contextual) measures of SES in the form of an index that represents area-based deprivation for 155 electoral wards (subcities) of Chennai megacity. Table S1 in the supplementary material summarizes the domains of selected widely used IMDs developed earlier, and most IMDs include income, employment, SES, education, housing quality, and ownership of goods or items.[7,8] The dimensions and indicators selected to devise an IMD for 155 electoral wards of Chennai megacity are slightly different from IMDs developed earlier because information on income is not available in the Census of India (see Table S1). The IMD is devised by using geographically weighted principal component analysis (GWPCA). GWPCA is now recognized as a very effective tool for detection of local non-stationary effects of variance in a data structure.[8] The local principal components (PCs) and local variance derived from GWPCA are suitable in devising the IMD.[8]

Mathematically, the local eigen decomposition of GWPCA transformation can be written in its algebraic expression as follows:

$$LVLT\ (u, v) = \Sigma\ (u, v) = XT\ W\ (u, v)\ X \qquad (1)$$

where W (u,v) is a diagonal matrix obtained from optimal bandwidths (here adaptive) based on the 'bi-square' kernel weighting scheme. The detailed description of GWPCA is given in the appendix section. To reduce noise and locate important factors of the IMD, the first 3 PCs with eigenvalues higher than 1 (i.e., $\lambda i \geq 1$) were retained.

The GWPCA-derived dimension weights were computed by multiplying the squared component loads and the proportion of variance explained by the corresponding PC and summing up across PCs. Weights are therefore derived using Equation 2.

$$W_k = \sum_{k=1\ldots3} PC_{k,i}^2 \times \frac{\sqrt{\lambda_k}}{\sum_{j=1\ldots3}\sqrt{\lambda_j}} \qquad (2)$$

where W_k is the weight given to IMD dimension i (either poor housing condition, lack of HH amenities and services, low asset possession, poor availability of WaSH services, or gender disparity), $PC_{k,i}$ is the component load in the kth PC (column of L), k is the eigenvalue of the kth PC (in V), and j is the number of PCs retained (here 3). The initial IMD is developed as a weighted aggregation of component scores of 3 PCs. The initial IMD was standardized using the minimum-maximum normalization method to obtain the final IMD.

The hot spot analysis tool of ArcGIS 10.2 software (Getis-Ord Gi*) was used to explore the spatial clustering of COVID-19 incidences and high IMD values (mathematical expression given in the Supplementary Material). The distribution of COVID-19 cases in Chennai megacity was negative binomial because its variance was higher than the means (see Table S2). Therefore, we used Poisson regression (PR) and negative binomial regression (NBR) to analyze the impact of individual domains of the IMD on COVID-19 incidences in Chennai megacity (see Supplementary Material for details).

Results

The descriptive analysis of COVID-19 incidences is reported in Table S2, in which it is shown that variability (σ=54.52) of COVID-19 incidences is higher than the mean ($\mu = 49.49$) and that it is following the negative binomial distribution. The first three components with eigenvalues higher than 1 (i.e., $\lambda i \geq 1$) accounted for 80.7% of the total variance in the data, and the first component alone explained more than 47% variance in the data. The product of the proportion of local variance explained by three components and the component score was summed up to devise the initial IMD. The IMD score of 0 stands for the least deprived ward (ward no. 125) and 100 for the most deprived wards (ward no. 40).

Fig. 1e shows the spatial distribution of COVID-19 and IMD hot spots. Hot spot areas for both COVID-19 and the IMD were mainly located in the northern and central parts. This area is crowded with a higher concentration of slums; the majority of the HHs with poor housing conditions and lack of HH services are located in this zone of the megacity. Fig. 1f shows the correlation coefficient between the ward-level prevalence of COVID-19 per lakh population, and the IMD showed a high correlation ($R^2 = 0.755$, r = 0.86, $P < 0.001$). A total of 34 wards of 155 wards have shown a prevalence of COVID-19 per lakh population of 50 and more cases (see Table S3). It was observed that 21 such wards (61.7% of the total) belong to the high and very high IMD category.

Estimated coefficients and other test statistics from the PR and NBR models are presented in Table S4, and the highest variations (adjusted R^2: 72.2%) with the lowest BIC and AIC for the NBR model compared with the PR model suggest that the NBR model better explains the relationship between area deprivation and COVID-19 incidences in Chennai megacity. NBR with two-sided tests and $P <0.05$ were considered statistically significant. The outcome of the PR and NBR models suggests that when all other variables were constant, according to NBR, the relative risk (RR) of COVID-19 incidences was 2.19 for the wards with high housing deprivation or, in other words, the wards with high housing deprivation having 119% higher probability ($RR = e^{0.786} = 2.19$, 95% confidence interval [CI] = 1.98 to 2.40), compared with areas with low deprivation. Similarly, in the wards with poor availability of WaSH services, chances of having COVID-19 incidence were 90% higher than in the wards with good WaSH services ($RR = e^{0.642} = 1.90$, 95% CI = 1.79 to 2.00). Spatial risks of COVID-19 were predominantly concentrated in the wards with a high IMD, which were mostly located in the northeastern parts of Chennai megacity.

Discussion and conclusion

We formulated the IMD using multiple socio-economic indicators at the electoral ward level in Chennai megacity and examined the relationship between the IMD and COVID-19 incidences. The hot spot analysis indicates that the formulated IMD was significantly related to higher incidences of COVID-19 in areas with a higher IMD.

The critical matters of the study were the selection of domains and the method of index formulation. The selection of domains was based on the previous IMDs, data availability in the Indian context, and factors that directly or indirectly affect the transmission of COVID-19. The indicators of the IMD were drawn from
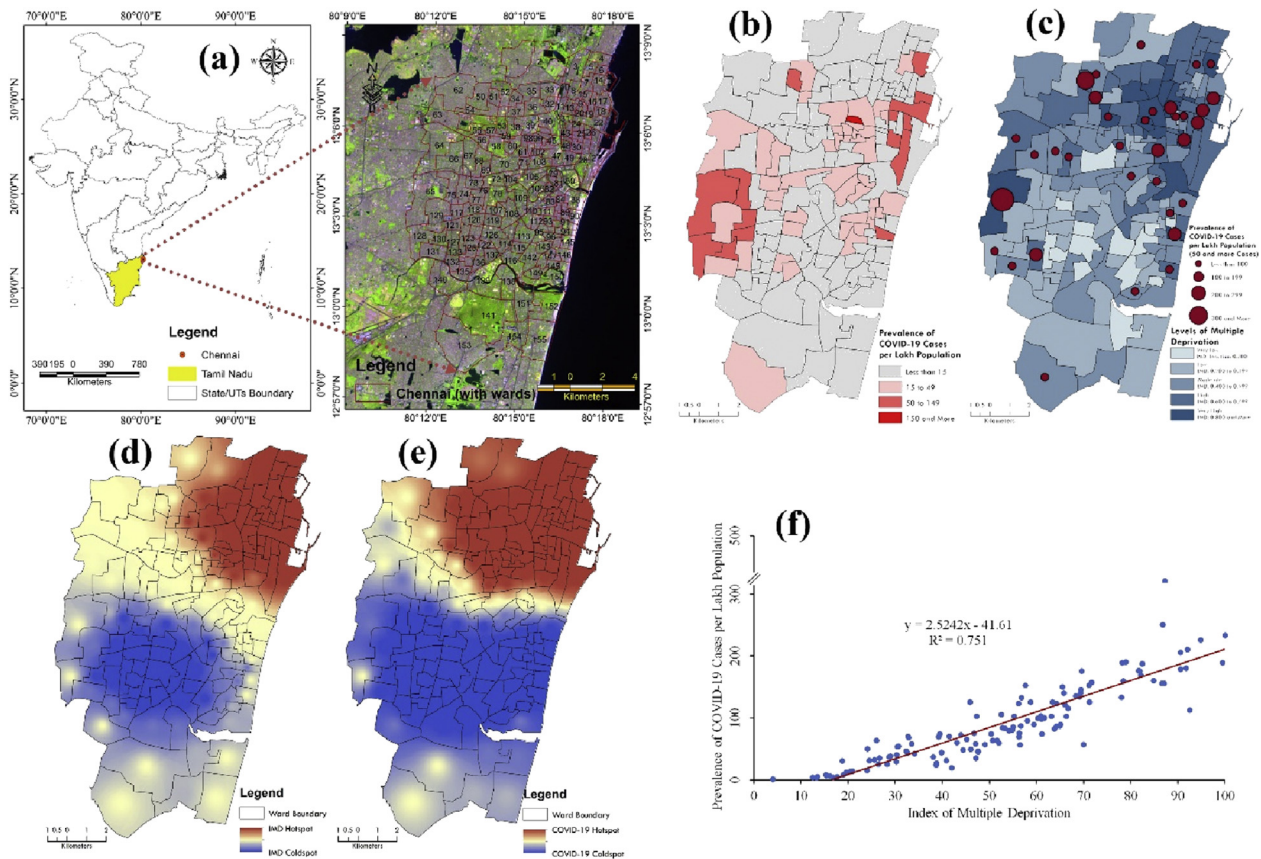

Fig. 1. (a) Study area; (b) ward-wise prevalence of COVID-19 incidences per 1,00,000 population; (c) distribution of wards with the prevalence of COVID-19 incidences of 50 and more cases per 1,00,000 population; (d) area deprivation hot spots and cold spots; (e) COVID-19 hot spots and cold spots; and (f) scatter plot showing the relationship between the IMD and prevalence of COVID-19 incidences per 1,00,000 population. COVID-19 = coronavirus disease 2019; IMD =Index of Multiple Deprivations.

the Census of India, the most reliable data source on SES in India. To formulate a composite index of SES deprivation, we used GWPCA, which is preferred over the currently widely used normal PCA. GWPCA as an emerging and promising tool[7] has a certain advantage over PCA such that it provides covariance structure, component scores, loadings, and explained variance for each electoral wards.[8] Therefore, we were able to devise the IMD by using the local component weights of IMD indicators, which was not possible in normal PCA.[8]

The regression results of this study support the findings of the earlier study that addressed the possible impact of SES on COVID-19 incidences.[9] To the best of our knowledge, this study was the first attempt on COVID-19 that quantitatively establishes the influence of area deprivation on COVID-19 incidences. The results of PR and NBR suggested that area deprivation has both positive and inverse associations with the incidences of COVID-19 in Chennai megacity. This further strengthens the findings of Ahmed et al.[10] that the socio-economic disadvantages and inequalities have a profound role in the spread of COVID-19. The findings of this study suggest that approaches to combat the COVID-19 pandemic must incorporate SES dynamics to develop a mitigation strategy.

In conclusion, this study formulated an IMD for area deprivation measures of SES disadvantage, and the index showed a substantial relation to COVID-19 incidences, especially for poor availability of WaSH services and poor housing conditions. This study highlights that the risks of COVID-19 tend to be higher in areas with low SES and that the northeastern part of Chennai megacity is predominantly high-risk areas. Although a more contextual discussion on IMD formulation is needed, the proposed index based on a common set of SES indicators may be readily applicable for research on the relationship between COVID-19 and SES deprivation. Our results can guide measures of COVID-19 control and prevention by considering spatial risks and area deprivation.

Author statements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2020.06.011.

References

1. World Health Organization (Who). *Coronavirus disease 2019 (COVID-19) situation report-103*. 2020. Retrieved from, https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200502-covid-19 sitrep103.pdf?sfvrsn=d95e76d8_4.
2. Hindusthan Times. *Uneven distribution of covid-19 among megacities*. Retrieved, https://www.hindustantimes.com/india-news/why-the-state-was-not-able-to-check-rapid-covid-spread/story-RakrhoQUlNLLg1pc2jMe1N.html; May 16 2020.
3. Tomar A, Gupta N. Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *Sci Total Environ* 2020;**728**:138762. https://doi.org/10.1016/j.scitotenv.2020.138762.
4. Cardoso MRA, Cousens SN, de Góes Siqueira LF, Alves FM, D'Angelo LAV. Crowding: risk factor or protective factor for lower respiratory disease in young children? *BMC Publ Health* 2004;**4**(1):19.
5. Baud ISA, Pfeffer K, Sridharan N, Nainan N. Matching deprivation mapping to urban governance in three Indian mega-cities. *Habitat Int* 2009;**33**(4):365–77. https://doi.org/10.1016/j.habitatint.2008.10.024.
6. *The News Minute, coronavirus data repository*. May 16 2020. Retrieved, https://www.thenewsminute.com/article/chennai-s-pulianthope-ward-has-highest-number-covid-19-cases-city-124901.
7. Fukuda Y, Nakamura K, Takano T. Higher mortality in areas of lower socio-economic position measured by a single index of deprivation in Japan. *Publ Health* 2007;**121**:163–73.
8. Mishra Swasti Vardhan. Urban deprivation in a global south city-a neighborhood scale study of Kolkata, India. *Habitat Int* 2018;**80**:1–10. https://doi.org/10.1016/j.habitatint.2018.08.006.
9. Patel JA, Nielsen FBH, Badiani AA, Assi S, Unadkat VA, Patel B, et al. Poverty, inequality & COVID-19: the forgotten vulnerable. *Public Health* 2020;**13**:110–1.
10. Ahmed F, Ahmed NE, Pissarides C, Stiglitz J. Why inequality could spread COVID-19. *The Lancet Publ Health* 2020;**5**(5). e240.