ORIGINAL RESEARCH ARTICLE
pISSN09763325 | eISSN22296816
Open Access Article (CC BY-SA)
www.njcmindia.com
DOI: 10.55489/njcm.130720221346

Covid19 Hazard Index: A Spatiotemporal Risk Forecast Tool

Manvendra Singh Rathore[1], UC Samudyatha[2], JK Kosambiya[3]

[1]Government Medical College, Surat
[2]Sri Devraj Urs Medical, College, Tamaka, Kolar, Karnataka
[3]Government Medical College, Surat

ABSTRACT

Background: In each geographic region, risk of new cases of COVID19 are driven by internal factors such as agent, host, and environment characteristics, as well as external factors, such as population mobility and cross border transmission of disease. COVID19 control measures are best implemented when local governments and health teams are well aware of these internal and external risks. These risks are dynamic in nature and hence need to be reviewed at regular intervals. The study conducted to develop a composite spatiotemporal Hazard Index comprising of three factors – presence of susceptible population, population density and presence of active cases with corresponding growth rates, to rank areas within an administrative boundary by their fortnightly risk of active COVID19 cases.

Methods: Using Principal Component Analysis, the weights of each of these factors were determined and applied to transformed values of factors in the districts of Gujarat state for months of January to July 2021. Hazard Index thus obtained was used to rank the districts.

Results: Spearman correlation between the Hazard Index and number of active cases 15 days later was moderate and significant ($p<0.01$) throughout the study period.

Conclusion: Hazard Index can predict Districts at highest risk of active cases in the given time period. These districts with high Hazard Index would require different control measures, depending on the factor that resulted in higher index value.

Keywords: Geo mapping, Spatial analysis, Principal Component Analysis, Weekly Growth rate, Population density, COVID19 vaccine coverage

INTRODUCTION

The second wave of COVID19 in India is largely attributed to laxity in the practice of COVID appropriate behavior such as wearing masks and social distancing, opening up of societies to gatherings and emergence of Delta variant of the virus.[1] As with most other Indian States, Gujarat also had two waves of COVID19 pandemic, with the first wave peaking between June-September 2020 and the second in March- May 2021. Three public health measures have been used in the control of COVID19 epidemic – testing, contact tracing and containment of cases, travel restrictions and lockdown, and vaccination. The Indian States are divided into Districts, which play a crucial role in the implementing these control measures. District level epidemic response for COVID19 has unique challenges. Challenges intrinsic to the administrative domain of a district include effective case detection, containment, contact tracing and vaccination within the district, coordinated by the district level health team. Growth rates of COVID19 cases and compartmental models can help in disease projection and planning. However, in order to tackle challenges extrinsic to the district, such as intra and interstate import of cases, multilevel coordination and involvement of multiple stakeholders

How to cite this article: Rathore MS, Samudyatha UC, Kosambiya JK. Covid-19 Hazard Index: A Spatiotemporal Risk Forecast Tool. Natl J Community Med 2022;13(7):424-429. DOI: 10.55489/njcm.130720221346

Financial Support: None declared
Conflict of Interest: None declared
Date of Submission: 09-04-2022
Date of Acceptance: 24-05-2022
Date of Publication: 31-07-2022

Correspondence: JK Kosambiya (Email: jkkosambiya@gmail.com)

are required. Regional and temporal variations in the design, implementation and uptake of COVID19 regulations can also pose a challenge to this coordination system. Hence the State and its Districts, as a team, need to be informed regarding their relative position in terms of risk due to intrinsic and extrinsic factors.

In this study, we have aimed at creating a composite Hazard Index, based on intrinsic and extrinsic factors, that can be computed fortnightly using readily available parameters. Using this Hazard Index, the districts at higher risk of new cases can be earmarked beforehand to optimally implement various public health measures to combat COVID19.

METHODOLOGY

Study area: Gujarat State is located in Western coastal India (23˚13'N, 72˚41'E).[2] It has a population of about 6.4 crores (2011 census) and a population density of 308 persons per sq.km.[3] It is administratively divided into 33 districts. It is surrounded by Indian states of Rajasthan in the Northeast, Madhya Pradesh in the East, Maharashtra in the South East and Dadra and Nagar Haveli and Daman and Diu in the South. Its North and Northwest borders are shared with Pakistan and to its West is the Arabian Sea. Gujarat has recorded above 8 lakh cases of COVID19 as of date, accounting for 2% of total cases recorded in India.[4]

Methods: The hazard of new COVID19 cases in an area would primarily depend on population characteristics (internal factors) within the area and cases that could be imported from elsewhere (external factors). Several internal characteristics, such as population density, age distribution of the population, housing facilities, overcrowding, air quality index, current number of active cases and variants etc. were initially considered while designing the study, based on literature review.[5–7] Similarly, external influences, like emergence of new variants, population cross border mobility, active cases in the rest of the country were also considered as influencers. After peer review, it was decided to construct a composite hazard index based on factors, the data of which were available for all districts concerned – active cases of COVID and weekly growth rate within the district and neighboring districts, number of COVID19 susceptible people in the district and population density of the district.

The map of 33 districts of Gujarat and 120 districts of neighboring States (52 districts of Madhya Pradesh, 35 districts of Maharashtra and 33 districts of Rajasthan) was plotted in QGIS ver 3.10.[8] The distance from centroid of each of the 33 districts to the remaining 32 districts of the State and 120 districts of neighboring States was estimated. Union territories were not considered, because their relatively smaller population. The number of active cases in each of these districts, as on 15th of each month, from January 2021 to July 2021 was collected from the aggregated crowd sourced data.[9] The weekly growth rate corresponding to these dates was also calculated from this data for all the 153 districts. The number of active cases within 200 km radius (A) from the centroid of each of the 33 districts of Gujarat and corresponding weekly growth rate (G) were determined separately for each district. The cutoff of 200 km was decided, since it ensured inclusion of at least two neighboring districts for each Gujarat district. The sum-product of active cases and corresponding weekly growth rate was calculated as $\sum(A*G)$.

The number of 'Protected' people in the 33 districts was estimated as the sum of vaccinated people and reported cumulative number of infected people (natural infection) as on 15th of each month from January- July 2021. It was noted that a certain proportion of 'protected people' could be missing, due to unreported, unrecognized cases. Similarly, a certain proportion of 'protected people' could be a duplication, due to the number of recovered patients receiving the vaccine. Nevertheless, number of vaccinated people, has largely exceeded the reported number of cumulative cases,[4] to mask this error. The number of vulnerable people (V) was calculated as the difference between total population estimate of the district and the estimate of 'vaccinated' people.

The population densities (D) of the districts (persons per square kilometer) were obtained from census data.[3]

The three factors, viz., $\sum(A*G)$, V and D were transformed individually using the following formula[10], treating each month separately (Jan- July):

$$X_t^n = \frac{X_t}{\text{mean}(X_t)} \qquad \text{.......} (1)$$

Where X_t^n is the normalized value of the factor [$\sum(A*G)$, V and D], X_t is the factor value at time t (month) and mean (X_t) is the average of the factor values at time t within the sample.

In order to determine the weights for each factor, the normalized values of the three factors were run through Principal Component Analysis using correlation matrix approach using SPSS ver 23.[11] Inputs were transformed values of the three factors, analyzed separately for each of the seven months. The resulting component coefficient score of each factor represents the weight (W_t) of each factor in the month (t). The weights were rescaled to 1 by dividing each weight by the sum of weights of all districts for the month. These were applied to the transformed values of the factors to obtain the Hazard score of each district (Formula 2).

$$HI_t = \sum W_t^n \,.\, X_t^n \qquad \text{.......} (2)$$

Where, HI_t is the Hazard Index of the district in month t, W_t^n is the rescaled weight for the month t, and X_t^n is the transformed values of the factor [$\sum(A*G)$, V and D] in month t,

To make the Hazard scores comparable across the months, they were normalized using the formula:

$$HI_t^n = \frac{HI_t - \min(HI_t)}{\max(HI_t) - \min(HI_t)} \quad \ldots\ldots (3)$$

Where HI_t^n is the normalized Hazard Index, HI_t is the Hazard Index of the district in the month t, min (HI_t) and max (HI_t) are the minimum and maximum values of Hazard Index among the 33 districts for month t.

Thus, the normalized Hazard Index ranged between 0 and 1, where proximity to 1 suggested higher hazard of new cases of COVID than the rest of the districts at the same time-period and proximity to 0 suggested lower hazard. The districts were ranked accordingly, and comparative maps were prepared on Day 1 of each month. The results were correlated using Spearman Rank correlation with the reported number of active cases 15 days later (Day 15) in each district.

The study was approved by the institutional review board.

RESULTS

The outputs of Principal Component analysis and rescaled weights of the factors derived from the analysis are given in Table 1.

The normalized Hazard Index is given in Table 2. The normalized Hazard Index was found to be positively and significantly correlated with number of active cases in the districts 15 days later when analyzed for the months separately.

Table 1: PCA output of month wise data (original)

Month →	Jan 2021	Feb 2021	March 2021	April 2021	May 2021	June 2021	July 2021	Avg. weights	Rescaled weights
∑(A*G)	0.31	0.10	0.01	0.27	0.26	0.03	0.03	0.14	0.12
Vulnerability	0.45	0.53	0.53	0.46	0.48	0.53	0.53	0.50	0.43
Density	0.47	0.54	0.52	0.50	0.50	0.55	0.56	0.52	0.45
Total								1.16	1.00

Table 2: Normalized Hazard Index of districts of Gujarat (Red represents top one third districts, yellow the middle one third and blue represents the bottom one third districts in terms of Hazard Index, representing High, Moderate and Low Risk Districts for the particular month) (original table)

Region	District	Jan-21	Feb-21	Mar-21	Apr-21	May-21	Jun-21	Jul-21
South Gujarat	Surat	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	Valsad	0.28	0.30	0.39	0.33	0.28	0.63	0.80
	Navsari	0.26	0.24	0.28	0.24	0.25	0.59	0.75
	Bharuch	0.17	0.14	0.17	0.15	0.16	0.12	0.11
	Tapi	0.08	0.07	0.21	0.07	0.07	0.05	0.05
	Narmada	0.07	0.03	0.06	0.04	0.05	0.01	0.00
	Dang	0.00	0.06	0.34	0.01	0.00	0.37	0.52
Central Gujarat	Ahmedabad	0.98	0.97	0.92	0.96	0.95	0.89	0.88
	Vadodara	0.48	0.46	0.47	0.46	0.46	0.41	0.39
	Dahod	0.41	0.38	0.36	0.41	0.43	0.38	0.42
	Anand	0.36	0.33	0.31	0.33	0.36	0.32	0.32
	Kheda	0.34	0.31	0.30	0.34	0.36	0.31	0.32
	Panchmahal	0.26	0.23	0.21	0.26	0.27	0.22	0.23
	Mahisagar	0.20	0.17	0.15	0.20	0.21	0.15	0.15
	Chhota Udaipur	0.16	0.13	0.16	0.15	0.16	0.12	0.12
North Gujarat	Gandhinagar	0.33	0.30	0.28	0.32	0.34	0.27	0.27
	Banaskantha	0.29	0.26	0.25	0.29	0.31	0.28	0.33
	Mehsana	0.28	0.25	0.23	0.27	0.29	0.23	0.23
	Sabarkantha	0.16	0.12	0.11	0.15	0.18	0.11	0.11
	Aravalli	0.15	0.11	0.09	0.13	0.15	0.08	0.09
	Patan	0.11	0.08	0.08	0.10	0.13	0.09	0.10
Saurashtra – Kutch	Rajkot	0.34	0.32	0.30	0.32	0.34	0.30	0.31
	Bhavnagar	0.26	0.23	0.21	0.23	0.24	0.21	0.21
	Surendranagar	0.13	0.09	0.08	0.10	0.14	0.10	0.10
	Junagadh	0.13	0.11	0.10	0.11	0.13	0.12	0.12
	Amreli	0.12	0.10	0.08	0.10	0.11	0.09	0.10
	Gir Somnath	0.12	0.09	0.09	0.10	0.12	0.11	0.12
	Jamnagar	0.11	0.08	0.08	0.08	0.09	0.06	0.06
	Botad	0.09	0.05	0.03	0.06	0.07	0.03	0.03
	Morbi	0.08	0.05	0.04	0.06	0.09	0.05	0.06
	Kutch	0.08	0.05	0.05	0.06	0.07	0.06	0.08
	Porbandar	0.04	0.01	0.01	0.01	0.03	0.01	0.01
	Devbhumi Dwarka	0.03	0.00	0.00	0.00	0.02	0.00	0.00
Spearman Correlation with absolute number of active cases 15 days later		0.50 (p=0.003)	0.49 (p=0.003)	0.40 (p=0.01)	0.55 (p=0.001)	0.4 (p=0.01)	0.4 (p=0.01)	0.52 (p=0.01)

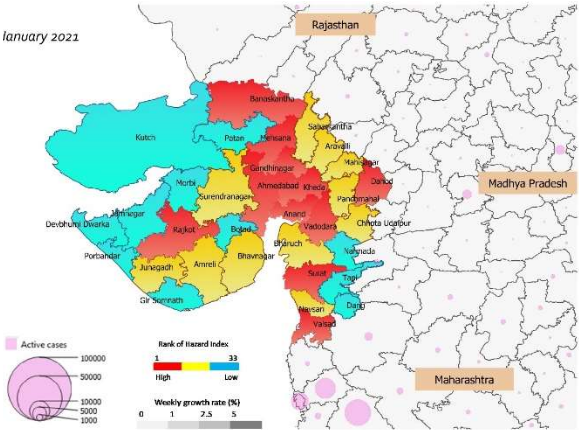

Fig 1a

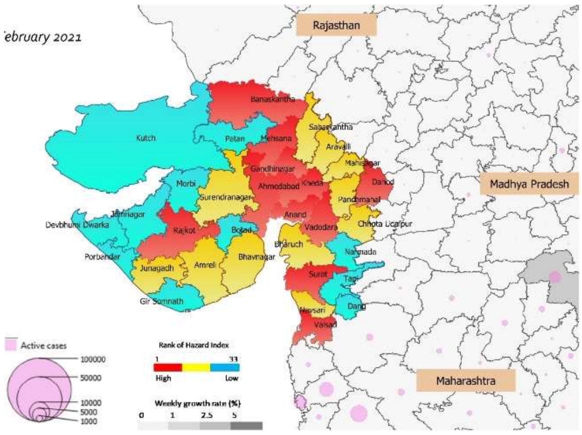

Fig 1b

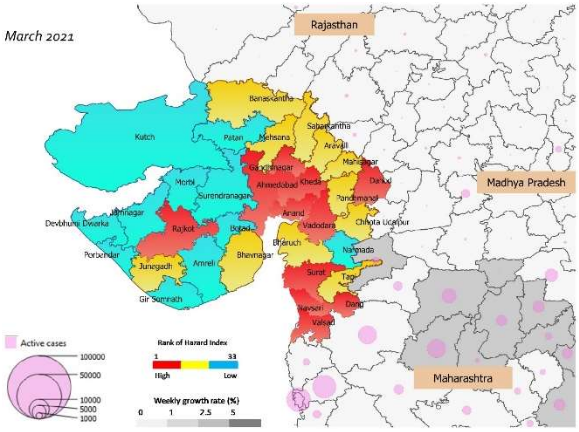

Fig 1c

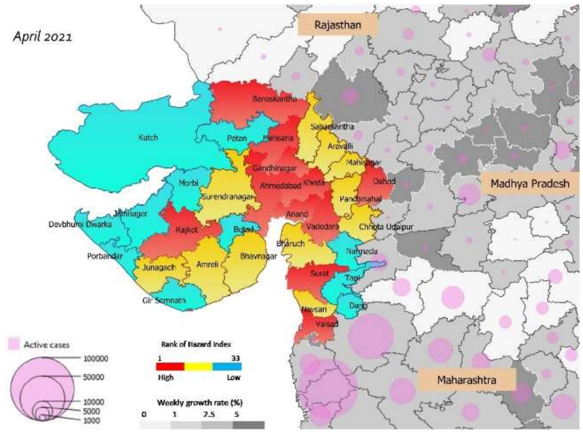

Fig 1d

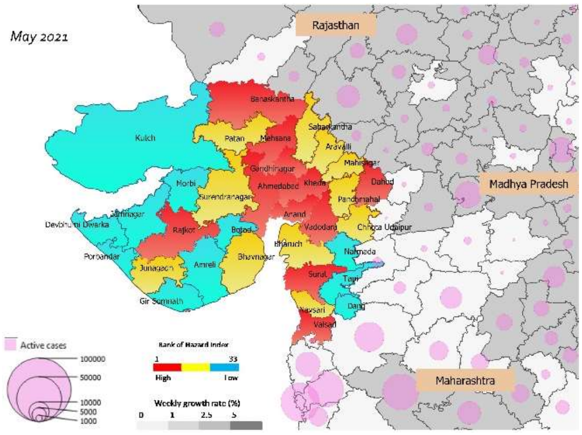

Fig 1e

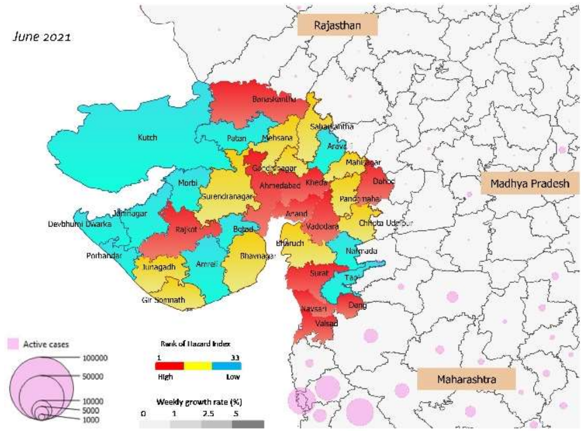

Fig 1f

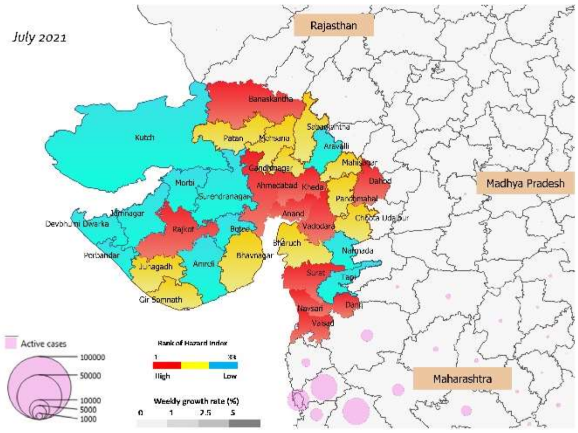

Fig 1g

Figure 1(a-g). Districts of Gujarat State, ranked by Hazard Index as High (red), Moderate (Yellow) and Low (Blue) risk of new cases (January-July 2021). Neighbouring States with active cases and weekly growth rates are shown in grayscale (original figures)

However, the correlation was weak (R_s=0.2, p=0.00) when all the months were considered together. Thus, the Index is a relative measure of Hazard in the given month and does not compare the magnitude of cases every month. The districts ranked based on Hazard Index have been mapped in Figure 1.

In South Gujarat region, Hazard Index was comparatively high throughout the study period in some districts such as Surat and Valsad, while in Tapi, Narmada and Dang districts the Hazard Index were comparatively low. In the Central Gujarat, throughout the study period, the Index values were either comparatively high or moderate. In Saurashtra-Kutch region of the State, most districts showed either moderate or low risk.

DISCUSSION

To the best of our knowledge, this is the first of its kind study, where a composite hazard index has been developed including the active cases in the neighboring areas as a risk factor in a temporal manner. Gujarat State, shares its borders with one of the worst affected States in the country – Maharashtra.[4] Gujarat State has a high share of migrant population[12], especially from neighboring states of Maharashtra and Madhya Pradesh. In the presence of access to individual level data of confirmed cases of COVID19 from Gujarat and the neighboring states, spatial and spatiotemporal risk can be estimated using Kernel Density estimation.[13] However, the current study used aggregated data, defining spatiotemporal risk considering Districts as units.

In this study, weights of three factors contributing to COVID19 risk were considered simultaneously – population density of the district, number of susceptible population in the district and the active cases with their growth rate in the district and in a radius of 200km. High population density is known to be associated with high reproduction number in COVID19.[14] On the other hand, the active cases and weekly growth rate in the district and neighboring districts have been considered as contributing factors, assuming the current conditions of community mobility to prevail.

All the three above mentioned factors were found to have a positive loading over the component, thus confirming the hypothesis that these factors had a positive influence on the hazard. The weight of the factor sum-product of Active cases and Growth rate, $\sum(A*G)$, was found to differ over the months, nearly reaching zero during months of March, June and July. These months corresponded to the beginning and ending of second wave in the region, suggesting that in the beginning of a wave, the hazard of new cases was predominantly loaded over population vulnerability and density. Similarly, when active cases were low in the neighboring districts, the hazard was again loaded over internal factors (vulnerability and density). It may be observed that the hazard index represents the relative comparison of risk between the districts for the given month, and does not compare the magnitude of absolute cases between months. Thus, a lower Index in a given District would not mean absence of cases, but a comparatively lower number of cases than in other Districts.

The risk of each District must be read in the context of each of the factors. Accordingly, specific control measures have to be intensified in specific districts. For instance, as of July 2021, districts in the South Gujarat border such as Valsad, Navsari, Dang and Surat are at higher risk, due to their proximity to active cases in Maharashtra. Districts higher up along the border are also at high to moderate risk (Red/yellow category) except Narmada and Tapi. This could be because of the lower population density and a smaller number of susceptible population in Narmada and Tapi. Dang District, despite having lower population density and susceptible population, is at high risk due to its proximity to active cases and being a transit point for travelers. This warrants the need of stepping up surveillance among interstate travelers in these districts. In Central Gujarat, majority of the districts are in high risk, because of their proximity to active cases within the region and high number of susceptible populations. Hence, effective contact tracing and containment measures should be strengthened in these regions. In North Gujarat, the districts are at moderate to high risk, due to their proximity to districts of Central Gujarat and presence of high number of susceptible populations. Here, surveillance among intrastate travelers has to be established effectively. In West Gujarat or Saurashtra-Kutch region, the districts are at moderate to low risk. These districts must strengthen their internal surveillance system to further reduce the number of active cases. Vaccination drive must be intensified in high-risk regions, followed by regions of moderate and low risk.

The continued use of this Index would help in identifying districts that require particular attention for public health strategy implementation. The parameters have to be reviewed at regular intervals, depending on the vaccine efficacy to circulating strains, arrival of susceptible population and travel restriction patterns. We suggest that these estimates be done yearly. Addition of other factors such as population mobility, age structure[15], distribution of co-morbid population[16] and other factors can be considered for a sub-district level analysis.

LIMITATIONS

Despite several other possible factors that can be added to the model, only three factors were considered because of non-availability of reliable longitudinal data for these factors. As mentioned in discussion, addition of these factors after capturing the large scale, longitudinal data might improve the model.

CONCLUSION

The allocation of resources and designing public health strategies for control of COVID19 must be informed by the risk that districts face in terms of population density, susceptible population and their proximity to active cases. Accordingly, different approaches, viz., travel restrictions and surveillance, internal active surveillance and containment and intensifying vaccination will have to be enacted in different parts of the State. The Hazard Index given by this study provides a bird eye view of these factors, which helps in forecasting situations and multi-level decision making for control of COVID19. The use of Hazard Index can be expanded to other States as well as to a sub-district level.

ACKNOWLEDGEMENT

We thank the Gujarat Government of India and the crowd sourced database covid19.india.org for making the data required for the study freely available in Internet.

REFERENCES

1. Chakraborty C, Ranjan A, Bhattacharya M, Agoramoorthy G, Lee S-S. The current second wave and COVID-19 vaccination status in India. Brain Behav Immun. 2021;(Online ahead of print).
2. Wikimedia Cloud Services. GeoHack-Gujarat [Internet]. 2021 [cited 2021 Jul 15]. Available from: https://geohack.toolforge.org/geohack.php?pagename=Gujarat&params=23_13_N_72_41_E_region:IN-GJ_type:city(60439692)
3. Registrar General and Census Commissioner of India. Census of India 2011 [Internet]. 2011 [cited 2021 Jul 15]. Available from: https://www.census2011.co.in/census/state/districtlist/gujarat.html
4. Government of India. myGov [Internet]. [cited 2021 Jul 15]. Available from: https://www.mygov.in/covid-19
5. Pierce JB, Harrington K, Mccabe ME, Petito LC, Kershaw KN, Pool LR, et al. Racial/ethnic minority and neighborhood disadvantage leads to disproportionate mortality burden and years of potential life lost due to COVID-19 in Chicago, Illinois. Health and Place. 2021;68(February).
6. Subramanian S V., Karlsson O, Zhang W, Kim R. Geo-mapping of COVID-19 Risk Correlates Across Districts and Parliamentary Constituencies in India. Harvard Data Science Review. 2020;
7. Ramírez IJ, Lee J. COVID-19 emergence and social and health determinants in Colorado: A rapid spatial analysis. International Journal of Environmental Research and Public Health. 2020;17(11):1–15.
8. QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project; 2019.
9. COVID19 India [Internet]. [cited 2021 Jul 16]. Available from: https://www.covid19india.org
10. Dharmawardena JSNP, Thattil RO, Samita S. Adjusting variables in constructing composite indices by using principal component analysis: illustrated by Colombo district data. Tropical Agricultural Research. 2016;27(1):95.
11. IBM Corp. IBM SPSS Statistics for Windows, ver 23. Armonk, N.Y: IBM Corp., Armonk, N.Y., USA; 2015. p. Armonk, NY: IBM Corp.
12. International Institute of Population Sciences. Census of India 2001 Migration Tables. In Mumbai; 2001. p. 15.
13. Elson R, Davies TM, Lake IR, Vivancos R, Blomquist PB, Charlett A, et al. The spatio-temporal distribution of COVID-19 infection in England between January and June 2020. Epidemiology and infection. 2021;149(June 2020):e73.
14. Rubin D, Huang J, Fisher BT, Gasparrini A, Tam V, Song L, et al. Association of Social Distancing, Population Density, and Temperature With the Instantaneous Reproduction Number of SARS-CoV-2 in Counties Across the United States. JAMA network open [Internet]. 2020 Jul 1;3(7):e2016099–e2016099. Available from: https://pubmed.ncbi.nlm.nih.gov/32701162
15. Monod M, Blenkinsop A, Xi X, Hebert D, Bershan S, Tietze S, et al. Age groups that sustain resurging COVID-19 epidemics in the United States. Science. 2021;371(6536).
16. Gémes K, Talbäck M, Modig K, Ahlbom A, Berglund A, Feychting M, et al. Burden and prevalence of prognostic factors for severe COVID-19 in Sweden. European Journal of Epidemiology [Internet]. 2020;35(5):401–9. Available from: https://doi.org/10.1007/s10654-020-00646-z