Impact of population density on Covid 19 infected and mortality rate in India

Arunava Bhadra[1*], Arindam Mukherjee[1] and Kabita Sarkar[2]

[1] High Energy & Cosmic Ray Research Centre, University of North Bengal, Siliguri, WB, 734013, India,

[2] Department of Mathematics, Swami Vivekananda Institute of Science Technology, Dakshin Gobindapur, Kolkata-700145

- Email: aru_bhadra@yahoo.com

Abstract:
The residents living in areas with high population density, such as big or metropolitan cities have higher probability to come into close contact with others and consequently any contagious disease are expected to spread rapidly in dense areas. However, recently after analyzing Covid-19 cases in US researchers at the Johns Hopkins Bloomberg School of Public Health, London school of economics and IZA – Institute of Labor Economics conclude that spread of Covid-19 is not linked with population density. Here we investigate the influence of population density on Covid-19 spread and related mortality in the context of India. We find some correlation between Covid-19 spread and population density which becomes more pronounced as statistics improves.

Keywords: Covid-19, infection and mortality rate, Population density, India

Introduction:

The emergence and continuous spreading of the highly contagious disease Covid-19 leads the World in a very distressing stage. The disease has severe adverse effects on the world economy, and on many aspects of human life like employment, physical and mental health of individuals, education etc.

Since coronavirus (SARS-COV-2) transmits via human contact the common perception is that covid-19 spread rapidly in dense areas whereas the probability of getting infected is low in areas with low population density. However, after analyzing Covid-19 infection and death rates of 913 urban counties in the United States a recent investigation by researchers at the Johns Hopkins Bloomberg School of Public Health claims that the infection rate is not linked with population density whereas death rate is inversely related with population density [Hamidi et al, 2020] except for metropolitan areas where higher infection and higher mortality rates have been noted. The inverse mortality relation with density has been attributed by the authors on the availability of better healthcare systems at higher density locations. Another very recent study by the researchers from London school of economics and IZA – Institute of Labor Economics**,** based on urban Covid-19 cases in US concluded that the timing of outbreak depends on population density with denser regions led an early breakout but both Covid-19 infection and death rates are unrelated with urban population density [Carozzi et al, 2020].

India is a large country with a total population of more than 1.3 billion. It has diverse demographic features. On the other hand as of today (17th August 2020), India is the third maximum covid-19 affected country by total number of infected people after USA and Brazil. The Covid-19 pandemic situation in India thus allows us to cross-check the recent findings [Hamidi et al, 2020, Carozzi et al, 2020] on the impact of population density on Covid-19 spread and related mortality. Probably the first study on impact of population density (along with several other geographical factors) on spread of Covid-19 was done by Gupta et al (2020) in the context of India but their analysis was based on the Covid data of Indian states up to 27th April 2020 when total reported Covid cases was only 29458 and they considered population density at the state level only.

In India, major part of total infections and mortality have, at least apparently, been attributed to metropolitan cities such as Mumbai, Delhi, Bangalore, Ahmedabad, Chennai and Kolkata. The infection and death rates appear to be much lower at remote districts. However, because of the large population of the metropolitan cities the total infections and mortality are normally higher in those places. Moreover, a strong presence of media lead reporting the Covid-19 cases in metropolitan cities in greater detail which might give a false impression that only these cities are dominantly contributing Covid-19 infection and related death.

To have a concrete and reliable idea on the issue in the present work a detailed analysis of Covid-19 infected and death cases as a function of population density of Indian districts is performed. Our findings suggest that till 17th August 2020 Covid-19 infections and mortality in India are correlated with population density; both infected and death cases are higher for large population density.

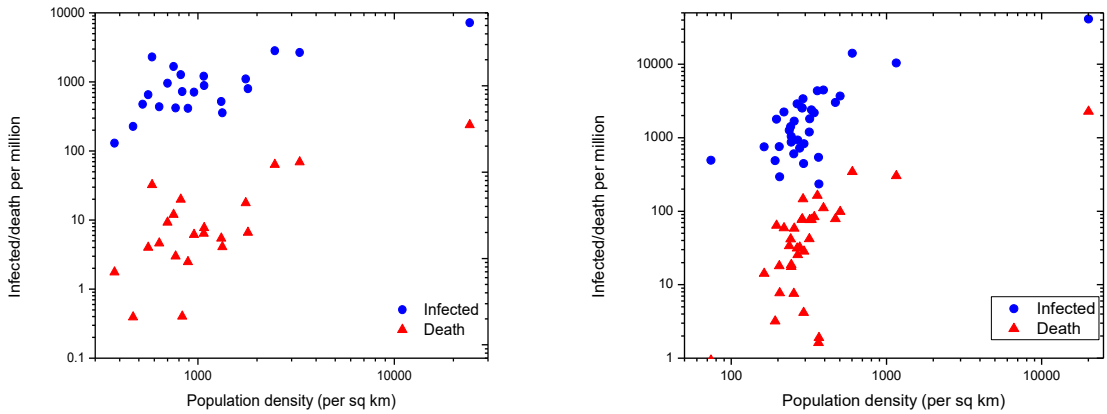
Data:

The first case of COVID-19 was reported in India on 30th January 2020. By 17th August 2020 the total 27,03517 people are infected in India by coronavirus out of which 1978436 have recovered, while it is fatal for 51936 Indians. The updated corona infected and death data are not available for a few districts such as the districts of Delhi, Goa, Andaman & Nicobars. Besides that the cumulative infected and mortality data for districts of Telangana are not reported**.** We, therefore could not include those districts in our analysis.

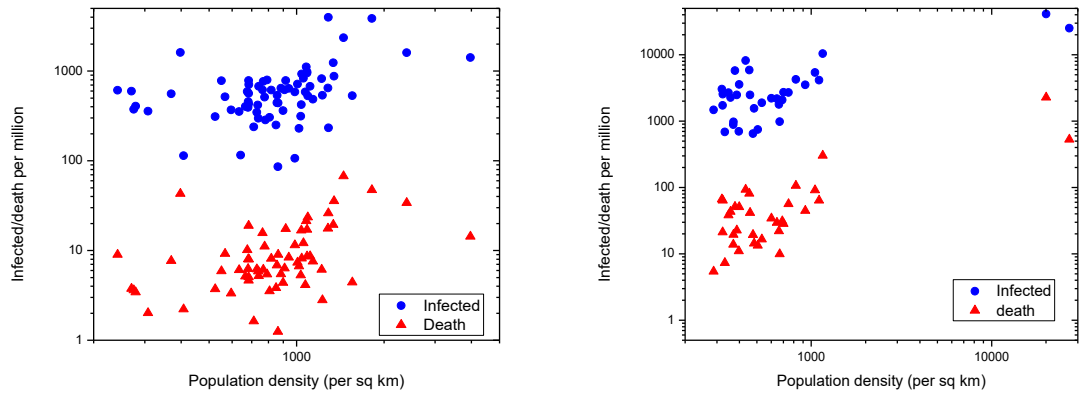
We have taken the district wise Covid-19 cumulative infection and death cases data from api.covid19india.org**.** Since the data is dynamic we use the data till 17th August, 2020 (the data was downloaded on 18th August 2020 at 09:53 hrs). For the district wise population and population density we have used the census 2011 data. A few new districts have been created since 2011 by splitting, rearranging some districts. For such cases we have taken the population and density data from the web site of the concerned districts though we could not include all such districts because of lack of data. We consider a total six hundred districts of India in our analysis.

Results:

At present there are twenty eight states and eight union territories in India. Many of the states are also quite large and populous. First we address the states separately. Since the population of different districts are vastly different we have considered the total infected and death cases per one million people for each of the districts. Here we shall show the results for four major states – West Bengal, Maharashtra, Uttar Pradesh and Tamil Nadu, which are from the East, West, North and South part of the country respectively. The scatter plot of the total (cumulative) infected and death cases per one million people against population density for the stated states are depicted in Figure 1. It is found that the infection and death rate are higher in metropolitan cities having huge population density in all the cases except Uttar Pradesh where no districts have very large population density. On the other hand it is noticed that Covid-19 spread and related death are low in districts with low population density. The situation is not very clear in districts with moderate density; the data points spread all over the graphical area with a slight trend to rise of infection and death with density.



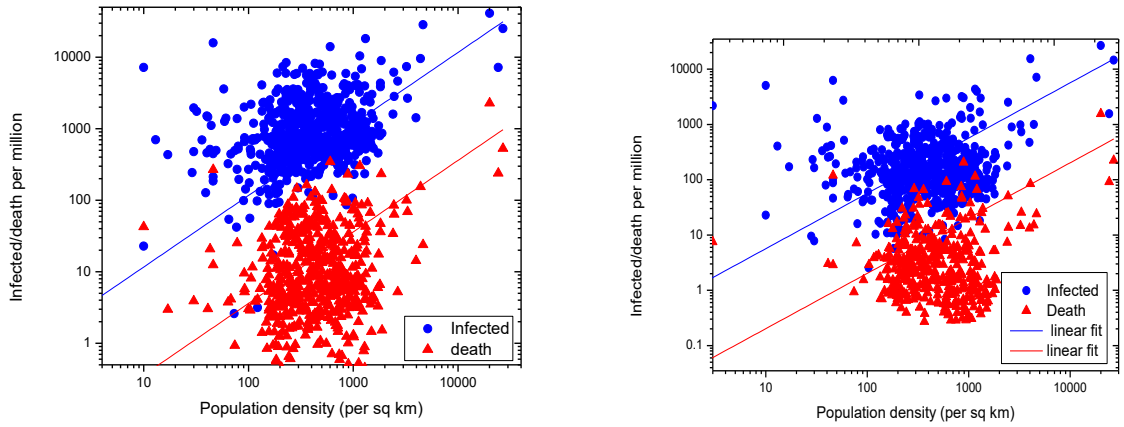
(a) (b)



(b) (d)

Figure 1: Variation of infected and death rate due to Covid-19 with population density of districts of (a) West Bengal, (b) Maharashtra (C) Uttar Pradesh and (D) Tamil Nadu.

Finally we consider India as a whole. The scatter plot of the total (cumulative) infected and death cases per one million people against population density for 600 districts of the country is shown in Figure 2 (a). Though there are large fluctuations, particularly in mid population density regions, over all there is a tendency that both infection and death cases rise with population density.



(a) (b)

Figure 2: Same as Figure 1 but for the whole India, (a) up to ~~10th~~ **17th** August 2020 (b) till 5th July2020. The lines describe the linear fit of the data.

To verify the claim of researchers that denser regions led an early breakout we plot the same Covid-19 infection/death variation with population density in figure 2(b) but for the data till 5th

July 2020. It appears that with the time the number of infection and death cases are increased but the overall trend remains unaltered. To quantify the correlation of infection and mortality cases we perform both linear (proportional) fit ($y_i = ad$, where y_i denote infected/death cases, d is the population density and a is the fitting parameter which is essentially the proportionality constant) and power law ($y_i = x^{\alpha}$) fit of the data using reduced-χ^2 technique. We also consider the fitting excluding the megacities. The results are given in Table 1**.**

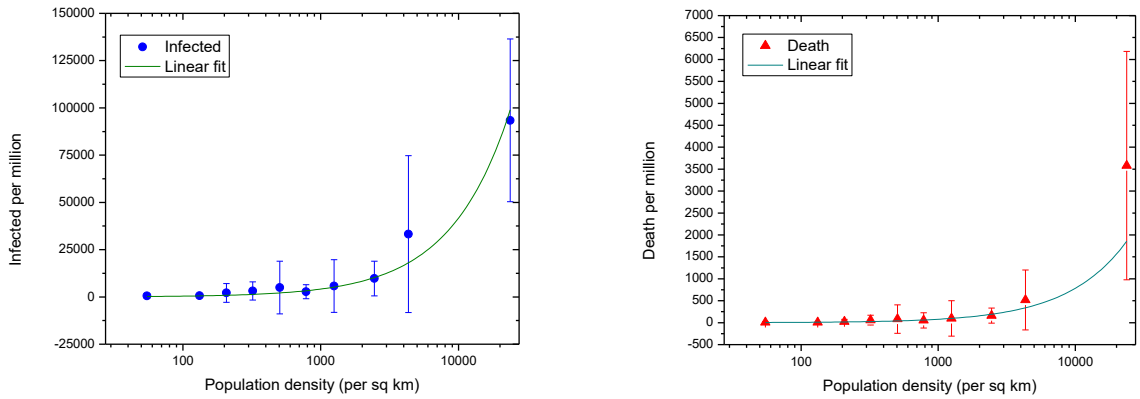
Table 1: Linear and power fitting of Infection and death rate.

data	Infection				mortality			
	Linear fit		Power fit		Linear fit		Power fit	
	a	Reduced -R^2	α	Reduced -R^2	a	Reduced -R^2	α	Reduced -R^2
Up to 17th August 2020	1.16 ± 0.06	0.45	1.00± 0.005	0.41	0.036±0.002	0.41	0.606±0.008	0.21
Up to 5th July 2020	0.56 ±0.02	0.43	0.941 ±005	0.42	0.02±0.001	0.32	0.52±0.01	0.10
Up to 17th August 2020 but excluding the megacities	2.1±0.1	0.36	1.001 ±0.007	0.05	0.024±0.001	0.21	0.47±0.01	0.03
Up to 5th July 2020 but excluding the megacities	0.50±0.04	0.22	0.93± 0.01	0.11	0.0069± 0.0007	0.12	0.27±0.02	0.02

For linear fit the adjusted-R^2 is found to be around 45% that decreases to 36% when megacities are excluded for fitting. For power law fit the adjusted-R^2 is found much lower compare to the linear fit. The proportionality constant a and the adjusted-R^2 are found to increase with time which implies that impact of density is becoming prominent as statistics (data size) improves.

The correlation between population density and infected/death cases is clearer when we consider small bin in population density and corresponding average infection/death cases are taken. In the dataset of districts of India, the population density per sq km is found to vary from 3 to 26903. We bin the population density in log scale with varying widths; particularly larger bin widths at high population density are considered so as to have at least a reasonable count of districts (at least 3) in each bin. The results are shown in Figure 3. A linear (proportional) fit describes both the infection and mortality data well with proportionality constant 4.17 ± 0.33 and 0.08 ± 0.01

and adjusted R^2 0.94 and 0.76 respectively. The Figure 3 thus indicates that the infected and death rate correlate with population density.



(a) (b)

Figure 3. Variation of averaged infected and death rate due to Covid-19 with population density for the whole country.

Discussion:

So far there is no established model about the role of population density on the spread of highly contagious diseases. The findings of the impact of population density on mortality during the 1918 influenza pandemic are debatable. A few researchers find no link between population density and death in 1918 influenza [Chowell et al., 2008, Nishiura, & Chowell 2008] whereas a few others claim for a significant positive correlation [Garrett 2010]. The Covid-19 pandemic allows us to study the issue in more detail and reliably. The issue is also important in connection of planning of future towns/cities.

In the present work we have considered population density obtained dividing the total population of a district by its total area which implicitly assumes a uniform spatial distribution of populations within a district. The Covid spread and death cases are also considered as uniform. Such assumptions lead to some uncertainty in estimating the correlation between Covid spread and population density. For instance, the population density in Darjeeling district of West Bengal is among the lowest in the state though the infection and mortality cases are relatively high there. This is because Siliguri, which is one of the congested cities of the state, falls in the Darjeeling district. Most of the infection and mortality cases in the Darjeeling district (and also in the Jalpaiguri district which jurisdiction wise contain a part of Siliguri) are from Siliguri. Whereas major geographical part of Darjeeling district is hill area where population density is quite low and thereby the resulting population density of the district is on the lower side. Further study is needed to overcome the stated issues, say by considering weighted population mean which is a measure of average "experienced" density or considering sub-division as a basic unit for the study. However, the limited availability of necessary data in public domain is a major hindrance for such improvements.

.

The Covid-19 spread and related mortality in a district may depend on various other factors including geographical features, economic conditions, prevailing health conditions, genetic factors, health infrastructure, policies adopted by the regulating authorities, average age of the residents of the district, number of testing etc. It is already known that Covid-19 has a larger impact on the older population (and population with comorbidity). The share of the older population is not the same in all the districts, neither the total testing number. However, there is no prevailing straightforward and unambiguous method of entangling different probable dependent factors. More research on these aspects is necessary to have a proper correlation between Covid-19 spread/ related mortality and population density.

In conclusion the present analysis indicates a positive correlation between Covid-19 infection and related mortality with population density, in contrast to the findings by the researchers at the Johns Hopkins Bloomberg School of Public Health and London school of economics and IZA – Institute of Labor Economics, based on US data. There is a vast difference in the living conditions of people in the USA and in India. The (large) density in India is reflected through the pressing of people against each other in the street, public vehicles, trains, queue for ration etc. The average area occupied by a family for leaving in cities of India is also much smaller. The Covid-19 cases were found higher in metropolitan cities even in the USA [Hamidi et al, 2020]. So containing a highly infectious disease like Covid-19 is a serious challenge for the country unless by any means (may be by mutation also) the corona virus (SARS-COV-2) is transformed to a non-lethal variety or an effective vaccine/medicine is discovered.

References:

Carozzi, F., Provenzano, S. and Roth, S. (2020) Urban Density and COVID-19, Discussion Paper Series, IZA Institure of Labor Economics, IZA DP No. 13440

Chowell, G., Bettencourt, L. M., Johnson, N., Alonso, W. J., & Viboud, C. (2008). The 1918–1919 influenza pandemic in England and Wales: Spatial patterns in transmissibility and mortality impact. Proceedings of the Royal Society B: Biological Sciences, 275(1634), 501–509. https://doi.org/10.1098/rspb.2007.1477

Garrett, T. A. (2010). Economic effects of the 1918 influenza pandemic: Implications for a modern-day pandemic. Working paper CA0721.2007. Federal Reserve Bank of St. Louis. https://www.stlouisfed.org/_/media/files/pdfs/community-development/research-reports/pandemic_flu_report.pdf

Gupta, A., Banerjee,S., Das, S. (2020); Significance of geographical factors to the COVID-19 outbreak in India, Modeling Earth Systems and Environment https://doi.org/10.1007/s40808-020-00838-2

Hamidi, S. Sabouri, S. & Ewing, R. (2020): Does Density Aggravate the COVID-19 Pandemic?, Journal of the American Planning Association, DOI:10.1080/01944363.2020.1777891

Nishiura, H., & Chowell, G. (2008). Rurality and pandemic influenza: Geographic heterogeneity in the risks of infection and death in Kanagawa, Japan (1918–1919). The New Zealand Medical Journal, 121, 18–27