An analysis of COVID-19 clusters in India
Two case studies on Nizamuddin and Dharavi

Pooja Sengupta · Bhaswati Ganguli · Sugata SenRoy · Aditya Chatterjee

Received: date / Accepted: date

Abstract

Background- With the COVID-19 pandemic wreaking havoc across nations, several research projects are being carried out to study the propagation of the virus. In this study we have made an endeavour to analyse the spread of COVID-19 in the districts of India.

Methods- Some districts in India have been much more affected than the others. A cluster analysis of the worst affected districts in India provide insight about the similarities between them. The effects of public health interventions in flattening the curve in their respective states is studied using the individual contact SEIQHRF model.

Results - The clustering of hotspot districts in India provide homogeneous clusters of districts that stand out in terms of number of positive COVID-19 cases and covariates like population density and number of COVID-19 special hospitals. The cluster analysis reveal that distribution of number of COVID-19 hospitals in the districts vary from the distribution of confirmed COVID-19 cases. The distribution of hospitals is much less skewed than the population density and COVID-19 cases. From the SEIQHRF model for Nizamuddin we
observe in the second phase the number of infected individuals had seen a
multitudinous increase in the states where Nizamuddin attendees returned.
Thereby, increasing the risk of the disease spread in the respective states.
However, the simulations reveal that the administrative interventions, if im-
plemented strictly, flatten the curve of disease spread. In Dharavi however, as
claimed by the Brihanmumbai Municipal Corporation officials, through trac-
ing, tracking, testing and treating, massive breakout of COVID-19 was also
brought under control.

Conclusions - The study rounds up with two important case studies on Niza-
muddin basti and Dharavi slum to illustrate the growth curve of COVID-19
in two very densely populated regions in India. In the case of Nizamuddin,
the attendees of the religious events who went back to their respective states,
increased the risk of infection manifold. However, Dharavi was one of the few
COVID-19 success stories. Through strict testing, treating, tracking and trac-
ing large-scale COVID-19 infection was brought under control.

Keywords SARS-COV2 · SEIQHRF model · Cluster analysis

1 Background

As COVID-19 pandemic wreaks havoc across continents, vigorous public health
responses are now being put in place in all the countries hit by the virus. Ar-
ticles are appearing in the the mainstream media explaining the importance
of such public health interventions in flattening the curve. Several research
groups are working on the evaluation of such containment measures by predict-
ing the course of COVID-19 cases in India, under various scenarios (Chatterjee
et al. [2020], Das et al. [2020], Das [2020], Ghosh et al. [2020]). We have made
an endeavour to look at some simple simulations of COVID-19 spread using
R, and used those simulations to illustrate how the various public health in-
terventions worked. To that end, we have used the SEIQHRF model which fits
a stochastic individual contact model (ICM) to allow the effects of various
public health interventions, specifically social distancing and number of hospi-
tal beds available, on the spread of an infectious agent such as the COVID-19
virus.

In the current study, we apply cluster analysis to classify real groups of
COVID-19 confirmed cases in different districts in India according to their
high similarity to each other. The results we obtain permit us to have a sense
of clusters of affected Indian districts. The main objective of clustering in this
study is to optimize monitoring of the affected areas which will be useful in
understanding seriousness of the spread of novel coronavirus (COVID-19) to
improve government policies, decisions, medical facilities (ventilators, testing
kits, masks etc.), treatment etc. to reduce number of infected and deceased
individuals.

In addition to clustering, a scenario analysis of the respective areas facil-
itate following the growth curve in the respective states. For that purpose,
individuals have been divided into various groups, such as, susceptible, exposed, infected, infectious (but self isolated), hospitalized, recovered and dead (death but not hospitalized, from COVID-19). Various parameters used in the stochastic model are, the number of exposure events (act) between infectious individuals and susceptible individuals, per day; probability of passing on infection at each exposure event between infectious people and susceptible people; rate at which symptomatic people self isolate themselves; the rate per day of symptomatic people requiring hospitalisation; rate per day at which people needing hospitalisation recover; and mortality rate per day for people needing hospitalisation but could not due to the hospitals being full. Each of these groups of individuals are segregated into compartments, and the transition rates of individuals between these compartments are also taken into consideration in the model. The intention is to look at different intervention experiments and follow the prevalence of COVID-19 (in terms of number of persons) with each passing day since the beginning of the epidemic. The study has been carried out for select Indian states where we find the maximum number of districts with high number of confirmed COVID-19 cases. The act parameter for each state has been adjusted based on the population density of the state.

The rest of the paper is organized as follows: Section 2 describes the SEIQHRF epidemic model and explains the parameters involved and explains the two clustering techniques used to classify districts. Section 3 reports the district-wise cluster analysis results. Section 4 discusses the case of Nizamuddin basti in Delhi, where a religious gathering exposed thousands to this virus and caused a second wave of infection in states when Nizamuddin attendees returned home. Here we discuss the results from fitting the SEIQHRF model. Section 5 discusses the case of Dharavi, one of Asia’s largest slums with very high population density. Section 6 enumerates the findings of this study and concluding remarks.

2 Methods

2.1 Clustering

Among the most commonly used algorithms, the hierarchical and the k-means algorithm are widely used in clustering. In this study we have used both the hierarchical and k-means clustering technique, mainly due to their great visualization power and simple and intuitive interpretation. One of the most relevant characteristics of hierarchical clustering is its generality; the user does not need to provide the number of clusters a priori. Based on the variable upon which clustering is done, the observations are clustered into optimum number of groups. As a disadvantage however, hierarchical clustering has a high computational complexity when the number of observations to classify increases. Thus using it on all 640 districts in India is quite cumbersome. So we have chosen the top 50 worst affected districts with maximum number of
confirmed COVID-19 cases. The methodology used to build the hierarchical clustering is the following:

1. Calculate the pairwise distance between all the hotspot districts using $D_i(S_A, S_B)$, for a certain value of the parameter $\alpha$. This distance matrix, symmetrical and with the null diagonal, will be essential to analyze the similarity between the behavior of the different districts.
2. Search through the distance matrix in order to select the two most similar elements (in our case of two districts).
3. Join (linkage) these two districts to produce a new group that now have at least two objects (districts).
4. Update the distance matrix by calculating the distances between the new cluster and all other clusters.
5. Repeat step 2 until all districts belong to a group.

The k-means algorithm is a popular clustering technique that aims to divide the data samples into $k$ pre-defined distinct non-overlapping subgroups, where each data point belongs to only one group. It keeps inter-cluster data points similar to each other and also tries to maximize the distance between two clusters. A cluster refers to a collection of data points aggregated together because of certain similarities. Unlike hierarchical clustering in k-means the number of cluster is prefixed. The optimum number of clusters, say $k$, is obtained by using an elbow plot. From the dataset the algorithm identifies $k$ number of centroids, and then allocates every data point to the nearest cluster, keeping the centroids as small as possible. The clustering stops when either:

1. The centroids have stabilized and there is no change in their values.
2. The defined number of iterations has been achieved.

In our study, data on confirmed COVID-19 cases, population density and number of COVID-19 hospitals in each districts is used to get the set of centroids. Although before the k-means clustering is performed the variables were standardized so that they are all on the same scale.

2.2 SEIQHRF model

Susceptible-Infectious-Recovered (SIR) models are popular in predicting the course of epidemics (Smith and Moore [2004], Allen et al. [2008], Korobeinikov [2009]), where the population is divided into 3 compartments of Susceptible (S), Infectious (I) and Removal (R) (through recovery or death) groups with defined rates of transition between them. Figure 1 is a pictorial representation of the three compartments in SIR model. Two differential equations are used to describe the rates of change into and out of each compartment.
Variation of SIR model is the SEIR model which includes another compartment as Exposed (E) population in the model (Stehlé et al. [2011], Li et al. [1999], Röst [2008]). Recently an extension of the SEIR model has been proposed (Churches [2020], Ghosh et al. [2020]), where apart from the above mentioned 4 stages, two other stages viz. Quarantine (Q) and Hospitalization (H) are considered to take into account the healthcare capability and the R stage is segregated to usual Recovery (denoted by R) and fatality (denoted by F). The model is dynamic in the sense that the number of individuals in each compartment may change over time. The flowchart of the compartments and the transition into and out of them is shown in figure 2 (Churches [2020]). The description of individual compartments are as follows.

| State       | Indicator | Functional Definition                                                                 |
|-------------|-----------|---------------------------------------------------------------------------------------|
| Susceptible | S         | Susceptible to COVID-19                                                                |
| Exposed     | E         | Exposed and infected, not yet symptomatic but potentially infectious                    |
| Infected    | I         | Infected symptomatic and infectious                                                    |
| Quarantined | Q         | Infectious and self-isolated (individuals who are isolated and hence do not come in contact with the susceptible population) |
| Hospitalized| H         | Requiring hospitalization (would normally be hospitalized if capacity is available)    |
| Recovered   | R         | Recovered, assumed to be immune from further infection. However, repeat infection is possible, but chances are low and for the time being assumed ignorable. |
| Fatality    | F         | Case fatality (death exclusively due to COVID-19 and of no other cause.                 |
With the introduction of more compartments into the SIR and SEIR models, the systems of equations become very complicated. Thus, in case of SEIQHRF, a stochastic compartment models were developed. Stochastic models help simulate individual members of a population, which allows the introduction of probability. Stochastic models make it relatively easy to specify different probabilities of passing on infection between the individuals in the different compartments like infectious to susceptible and infectious quarantined individuals and susceptible. The stochastic model can be made sufficiently sophisticated to adequately model real-life scenarios. We have thus used the model to explore the effects of different types of public health interventions and policies through scenario analyses.

![SEIQHRF model flowchart](image)

**Fig. 2 SEIQHRF model flowchart**

In the present study we will fit the model mentioned in 2 compartment model to COVID 19 outbreak data for select states in India and use simulations
to explore the impact of lockdown on the epidemic and perform a what-if analysis for strategies of imposition and relaxation of the lockdown.

3 District-wise analysis

3.1 Hotspot districts

With the recent updates on COVID-19 cases in India, it has become evident that some parts of the country are more affected than others. For instance, Maharashtra, Tamil Nadu, Delhi, Gujarat, Telangana, Karnataka and Uttar Pradesh are on the top of the list of most affected states in India. At a more disaggregate level, few districts of these states are more badly affected. The common trend in COVID-19 infection is that mostly the urban regions have been more affected than their rural counterparts. This is primarily due to the fact that the virus was introduced by travellers from abroad. In order to focus on the heavily affected areas the government has demarcated some of the districts as hotspots based on the confirmed cases of COVID-19. We have used the last updated number of confirmed cases on 28 June 2020. The variables used in this study are described in detail in table 2 below. The summary measures of population density and number of COVID-19 hospitals reveal a large variation between the districts. The minimum number of COVID hospital shows that there are districts without any dedicated facility like in Baramulla district of Jammu. Whereas some districts have as many as 62 like in Thane district of Maharashtra. Again, going by population density districts like Central Delhi have a density of 27730 persons per sq. km. when in Jodhpur district of Rajasthan the figure is only 161 persons per sq. km. These variations necessitate performing a cluster analysis of the districts worst hit by COVID-19. A hierarchical clustering followed by k-means clustering of the districts, based on only confirmed cases and confirmed cases, population density and number of COVID hospitals respectively, help us understand the similarities between them.

Table 2

| Variables       | Description                | Min | Mean  | Median | Max  | St. Dev |
|-----------------|----------------------------|-----|-------|--------|------|---------|
| Confirmed cases | No. of confirmed cases     | 122 | 5651  | 1332   | 74252| 13545.7 |
| COVID hospitals | No. of COVID-19 hospitals  | 0   | 7.14  | 4      | 62   | 11.08   |
| Density         | Population density         | 161 | 3746.4| 838.5  | 27730| 7320.3  |
| Active          | No. of active cases        | 3   | 2274.5| 491    | 27631| 5508.5  |
| Deceased        | No. of deceased cases      | 0   | 211.9 | 29.5   | 4284 | 643.7   |

Using hierarchical clustering we have obtained the clusters of districts as shown in figure 3. From figure 3 it is seen that one of the first few clusters comprises of Mumbai, Pune, Nashik, Nagpur, Thane, Yavatmal, Aurangabad
and Ujjain. Seven out of these eight districts are from Maharashtra which so far has been the worst hit among the Indian states. This particular cluster has an average of 17196.5 confirmed cases which is highest among the other five clusters.

Next we used k-means clustering technique to find groups of districts that are similar in terms of number of COVID special hospitals, population density and number of confirmed COVID-19 cases. In figure 4 we represent three distinct clusters of districts based on number of COVID-19 confirmed cases and the two covariates that directly or indirectly impact the spread of COVID-19. Clusters 1 comprises of Mumbai, Thane and Chennai. These three districts, have similarities in the number of confirmed COVID-19 cases, their population density and number of COVID-19 special hospitals. From the data we could see that these three districts hold the top three positions in terms of
confirmed cases, 74252, 32735 and 53762 respectively. Mumbai, Thane and Chennai also are few districts with many COVID-19 dedicated hospitals, 44, 62 and 20 respectively. Hence these three districts may be considered outliers compared to the rest. The second cluster comprises of Kolkata, West Delhi, Central Delhi and Hyderabad. Upon inspection of the data, it is seen that most similarity in these districts is in terms of population density; although Kolkata and Hyderabad have much more confirmed cases than West and Central Delhi. According to [India.com][3] as on 7th June 2020, the four metro cities, Mumbai, Delhi, Kolkata, Chennai and three other cities, Ahmedabad, Indore and Pune accounted for almost 60% of COVID-19 positive cases in India.

So far most of the hotspot districts have been either the metro cities or the area in their vicinity. According to [Desai][2020] although the urban residencies appear more sustainable in terms of the economies of scale provided by their population, they also tend to be almost defenceless in the face of a pandemic outbreak. If we look at the cases of megacities like New York and London, their healthcare system was practically crippled by the sudden spread of COVID-19. The list of hotspots in India includes some of the densest metro cities of the nation like, Delhi, Mumbai, Chennai, Kolkata, Hyderabad. [Rocklov and Sjodin][2020] state that high population densities act as an impetus to the spread of the virus. In many locations the main strategy to control COVID-19 spread is by contact tracing. One major difficulty in the method is the detection of asymptomatic carriers. To add to that if we consider a densely population area where it is difficult to keep up social distancing practices, it is a nightmare to mitigate the proliferation of the virus. The COVID-19 outbreak has also placed an unparalleled demand on the healthcare system. Our health system and healthcare officials have their hands full with more patients coming into the hospitals than they can handle. This has also compromised other essential health services that people expect from the facilities. Also the lack of COVID-19 special hospitals in some regions have forced patients to travel to other places increasing the risk that they might infect other people on their way. Keeping the effects they might have on the escalation in the number of COVID-19 infection, we are looking at the number of COVID-19 special hospitals in the hotspot districts and their population densities. In figures 5 to 9 the panels represent the districtwise distribution of COVID-19 cases, number of COVID-19 special hospitals and population densities. The selected states are Maharashtra, Tamil Nadu, Uttar Pradesh, West Bengal and Gujarat.
Fig. 4 Clusters of districts

Fig. 5 Maharashtra
Fig. 6  Tamil Nadu

Fig. 7  Uttar Pradesh

Fig. 8  West Bengal
In most of the states, the metropolitan areas are most affected. For Maharashtra (5), the district marked in red is in the extreme west coastal region, Mumbai city and Mumbai Suburban. Out of around 1,92,000 cases in Maharashtra, almost 87,000 are in Mumbai city. Similarly, in Gujarat, the most affected district is Ahmedabad. In Uttar Pradesh and Tamil Nadu the main hotspot districts are Agra and Chennai with 1225 and 2082 cases respectively. In West Bengal the main focus is on three districts, i.e. Kolkata, Howrah and North 24 parganas. The population density map shows a similar pattern across districts in the states. However, the distribution of number of COVID-19 hospitals in the districts vary from the distribution of confirmed COVID-19 cases. The distribution of hospitals is much less skewed than the population density and COVID-19 cases. The state government and the central government has been proactively trying to augment hospital facilities. The almost even distribution of the covid-19 special hospitals in districts is a result of those initiatives.

4 Nizamuddin : A case study

In the wake of the global crisis due to COVID-19, almost all affected countries, have seen an exponential growth in the number of confirmed case count. In most of such countries, health workers discovered a group of people who got infected at one place and mostly at the same time. Such groups are termed clusters. For example, in South Korea, close to 56% of the infections started from a church frequented by an infected woman. In Singapore, a dinner party was found responsible for 10% of the cases. In India one such case has been identified at the Nizamuddin basti in Delhi. Nizamuddin is a crowded, busy neighborhood of narrow lanes lined with market stalls and tiny shops, known for two important historic sites. The map, figure 10, shows just how densely populated the locality is. It is this locality that has been identified as a massive cluster, after a religious congregation held in mid-March in this locality led to
COVID-19 spread among the attendees; at least 130 cases have been identified as having originated from this cluster.

As a special case study, we have used the SEIQHRF model to simulate the spread of the virus through infected individuals from the Nizamuddin cluster. A stochastic individual contact model (ICM) is used to simulate the baseline projections of the timeline of incubation period, illness duration and survival time of the case fatalities. In the five panels of figure 11 we provide general statistical properties of the model in relation to some of the major parameters estimated.

1. The incubation period for the virus has a median of about 10 days and in few cases could reach 20 days or more.
2. For some, isolation started much later, could be as long as 10 days. However, many people took less than a couple of days before they started to isolate themselves because of some COVID-19 like symptoms.
3. Illness duration seems span around 25 days. Although this may vary based on the individuals comorbid conditions.
4. Hospital care duration is about 10 to 15 days mostly, which seems to match actual observations.
5. Survival time of case fatalities is seen to be mostly between two to five days. However few fatal cases have had a survival time as high as 20 days.
Prevalence is the number of individuals in each compartment at each point in time (each day). The next plot shows the prevalence of COVID-19 among the Nizamuddin population. Given the dense nature of the population in the Nizamuddin basti the act parameter (average number of exposure events between infectious individuals and susceptible individuals) for this area is quite high making it a hotbed for the virus spread in the community. From figure [12] we can see how peaked the distribution of exposed individuals is in the neighbourhood. Also due to the delayed signs of this viral infection among individuals, the ones living in these densely populated areas unknowingly pass on the virus to others. The prevalence of cases where self-isolation by individuals is undertaken a comparatively low peak, thus isolating oneself when symptoms are visible reduce the risk for others. The model estimate that from 15,000 susceptible people, almost as many as 15000 other people are potentially exposed during the 3 days event, and from this almost 15000, about 4000 are potentially infected and require hospitalization. However, due to the delayed signs of the infection causes possible delays in the people getting to the hospital for checking, which causes about 20 days delay in the infected
person to be hospitalized. Furthermore, the number of people who went to self-isolation is predicted to be very small.

![Baseline simulation](image)

**Fig. 12** Prevalence simulation of COVID-19.

In figure 13 we have isolated the infected, hospitalized and fatal compartments to look at them more closely. It shows that of the almost 15000 exposed individuals, more than 4000 required hospitalisation, although with a delay of almost 15 days. There were more than 2500 people infected. However, the number of fatal cases remained low, around 150.
In India, to fight COVID-19 the most important strategy so far has been to implement social distancing. However, important steps need to be taken like increasing hospital capacity. The administration and the medical community have recommended lockdown and social distancing at various stages and phases with a combination of home confinement of population, suspension of all natural human activities and movements barring a few emergencies. Some of these administrative interventions are implemented in the simulation of the Nizamuddin cluster and represented in figures 12 and 13. A side-by-side comparison of similar implementation in Delhi is in figure 14. The baseline plot shows that the number of infected/infectious individuals in the Nizamuddin cluster is close to 10,000, almost double of that in Delhi. This can mostly be due to the fact that in a densely populated area there is a higher propensity for people to mingle with more number of individuals, thus increasing the prevalence of the spread of COVID-19 manifold.
In both Nizamuddin cluster and Delhi, the curve of infected/infectious individuals have flattened with implementation of the above interventions. Among the administrative mediation, increasing both social distancing and
self-isolation starting at day 15 after the start of the disease produces the best results in flattening the disease spread curve. However, a comparison shows that the Nizamuddin cluster increased the number of confirmed cases and also the number requiring hospitalisation. The numbers have almost doubled from that of Delhi.

The Nizamuddin cluster was discovered in mid-March and at least 130 cases in India have been identified as having originated from this cluster. There was then a second wave of infectious individuals identified in several other states that were linked to the Nizamuddin cluster. Tamil Nadu, Delhi, Telengana, Gujarat, Maharashtra are a few such states.
An analysis of COVID-19 clusters in India

Fig. 15 Comparison of baseline prevalence with those with interventions
Fig. 16 Comparison of baseline prevalence with those with interventions
In each of the states above (Tamil Nadu, Delhi, Telangana) the number of infected individuals linked to Nizamuddin cluster are 72, 24 and 6 respectively. The comparison in figures 15, 16 and 17 show that in the second phase the
number of infected individuals had seen a multitudinous increase. Thereby, increasing the risk of the disease spread in the respective states. However, the simulations reveal that the administrative interventions, if implemented strictly, flatten the curve of disease spread. Specifically, the best results are seen if social distancing is strictly practised by day 15 since the disease outbreak.

5 Dharavi: A case study

Dharavi is considered Asia’s largest slum, one of the most densely populated areas in the world (with 2.6 lakh people per sq. km) and, now, also marked as a containment zone for COVID-19. When driving through the lanes and bylanes of Dharavi, thronging with people, it is evident that the one necessary norm to prevent the spread of COVID-19, social distancing, is practically impossible to implement here. Most of the houses in the area are merely 10X10 feet, it is a challenge to keep people confined to such small area. In addition, there is a major problem with the common toilets that inhabitants have to use, which makes the containment of the virus an impossible task. Although the first case of COVID-19 in India was reported on the 30th of January, by middle of March only mere eight positive cases of COVID-19 were identified in Mumbai. All of these cases had a travel history abroad. However, by the end of March, the number of positive cases started growing exponentially. In Dharavi the first case was observed on the first of April. [Times 2020]. The index patient was a 56 year-old garment shop owner complained of high fever and cough. When his symptoms worsened he was admitted to the hospital where he succumbed to the disease before civic officials could talk to him about the people he might have come in contact with. So the officials began their contact-tracing exercise. It came to light during their investigation that a few days before the garment shop owner started showing the signs of COVID-19 infection, he had hosted a party with some people in his house, and all these guests were attendees of the religious event that took place in mid-March in Nizamuddin area of Delhi.

Other than the contraints of space that make the inhabitants more exposed to the virus, most COVID-19 positive cases in Dharavi have been found to be asymptomatic. The silent carriers living in the area were unaware of how many people they were infecting with the virus. A simulation run using the SEIQHRF model produces the number of infected/ asymptomatic, infectious, self-isolated, hospitalised and fatal cases in Dharavi (figure 18).
Due to the very high $act$ parameter value for Dharavi the SEIQHRF model algorithm is used to simulate the scenario for a period of 30 days. The results illustrate the predicted course of COVID-19 and the impacts of various
experimental interventions. Dharavi has especially been a challenge for the Maharashtra government, slowly turning out to be the COVID-19 capital of the state. From figure [18] it is seen that the prevalence rate among the residents was very high since the beginning of the outbreak. The lack of space to practice social distancing and the use of common toilets, exposed them more to the risk of acquiring the virus. In the baseline model the number of infectious/asymptomatic patients is as high as almost 10,000. However, measures like increased social distancing and self-quarantine is expected to help in flattening the curve to some extent. From the panel on the left we can see that with increase in social distancing and quarantining the number of infected/asymptomatic has come down by almost 2500. There is also a delay in the prevalence number reaching it’s peak, which is almost 10 to 15 days. Also the panel on the right reveals predicted number of people requiring hospitalization. Dharavi alone predicts a requirement of 200 hospital beds for hospitalization. As, of 11th July 2020 the total number of positive cases are 2359. The fatality rate in Dharavi increased from about 3 % to 4.1 % over a period of two weeks between May 5th and May 20th.

Amidst all the challenges, Dharavi has emerged as one of the few successes in containing the spread of COVID-19 with 1952 recovered cases of its 2359 confirmed and only a handful 166 remaining active. According to the Brihanmumbai Municipal Corporation officials, the steps taken in Dharavi can be defined by four T’s- tracing, tracking, testing and treating. Almost as many as 47500 houses were screened, while 14970 people were screened in mobile vans. Quarantine facilities were ramped up with the use of schools, marriage halls and sports complexes. The strict implementation of lockdown and an endeavour to treat all COVID-19 patients in Dharavi itself proved successful in containing the infection. A proof that these interventions were fruitful is the fact that while in April, the doubling rate in Dharavi was 18 days, it gradually improved to 43 days in May and slowed down to 108 and 430 days in June and July respectively. In a recent press meet officials from WHO cited Dharavi as one of the few cases around the world where a massive breakout of the infection could still be brought under control, [India.com] [4].

6 Conclusion

The various scenarios analysed in this paper introduce the various interventions undertaken and compare the course of COVID-19 under each of them with a baseline model. To summarise the outcomes some important observations made are as follows: In the baseline model, we see the result of our simulation on a hypothetical world of 1000 people. Our observations are;

— The epidemic subsides in about two months.
— The population density of Delhi being very high, very high number of people are infected, although asymptomatic.
We see typical exponential behaviour, although these are prevalence numbers, not incidence. Prevalence tends to start with exponential growth then tapers off.

- The number requiring hospitalisation is not too large.
- The number in the case fatality compartment is monotonically increasing, as expected, but at a much lower rate.

But the various interventions have varied effect on the spread of COVID-19. In most of the states in India the best results are seen in implementation of early social distancing and quarantine of the infectious individuals. We see typical exponential behaviour, although these are prevalence numbers, not incidence. Prevalence tends to start with exponential growth then tapers off. That’s what we are seeing here, so that’s good. The number in the case fatality compartment is monotonically increasing, as expected, but at a much lower rate; which is good.

In case of Nizamuddin, the highest frequency is seen between day 5 and 7, the individuals for whom the disease turned fatal mostly survived for five to seven days. The attendees of the gathering in Nizamuddin then exposed some of the states/UTs, like, Delhi, Tamil Nadu and Telangana to a second wave of COVID-19 infection. A side-by-side comparison of prevalence in these states during the first and second wave show a significant growth in number of COVID-19 positive cases. In contrast to Nizamudding, Dharavi slum in Maharashtra, in spite of being one of the most densely populated slums in Asia, managed to control the massive outbreak of COVID-19 majorly through strict implementation of government policies of testing as many as possible, treating the ones already tested positive in hospitals. The healthcare facilities were ramped up to cater to the huge population in the area. Also back tracking and tracing of contacts who were in close proximity of the infected were also proactively undertaken. These measures made it possible to reduce the reproduction number significantly in the area.

### 7 Abbreviations

Not applicable.

### 8 Declarations

#### 8.1 Ethics approval and consent to participate

Not applicable

#### 8.2 Consent to publish

Not applicable
8.3 Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. All data generated or analysed during this study are included in this published article.

8.4 Competing interest

The authors declare that they have no competing interests.

8.5 Funding

LISA 2020 grant of USAID and University of Colorado Boulder.

8.6 Authors’ Contributions

Pooja Sengupta - drafted the work, analysis, and interpretation of data
Bhaswati Ganguli - contributions to the conception, substantively revised it
Sugata SenRoy - contributions to the conception, substantively revised it
Aditya Chatterjee - contributions to the conception, substantively revised it

8.7 Acknowledgements

Not applicable

References

Allen LJ, Brauer F, Van den Driessche P, Wu J (2008) Mathematical epidemiology, vol 1945. Springer
Chatterjee K, Chatterjee K, Kumar A, Shankar S (2020) Healthcare impact of covid-19 epidemic in India: A stochastic mathematical model. Medical Journal Armed Forces India
Churches T (2020) Tim churches health data science blog: Modelling the effects of public health interventions on covid-19 transmission using r - part 2. URL https://timchurches.github.io/blog/posts/2020-03-18-modelling-the-effects-of-public-health-interventions-on-covid-19-transmission-part-2/
Das S (2020) Prediction of covid-19 disease progression in India: Under the effect of national lockdown. arXiv preprint arXiv:200403147
Das S, Ghosh P, Sen B, Mukhopadhyay I (2020) Critical community size for covid-19—a model based approach to provide a rationale behind the lockdown. arXiv preprint arXiv:200403126
Desai D (2020) Urban densities and the covid-19 pandemic: Upending the sustainability myth of global megacities. ORF Occasional Paper 244(4)

Ghosh P, Basheer S, Paul S, Chakrabarti P, Sarkar J (2020) Increased detection coupled with social distancing and health capacity planning reduce the burden of covid-19 cases and fatalities: A proof of concept study using a stochastic computational simulation model. medRxiv

Indiacom(a) (2020) Dharavi flattens its covid curve: After who’s praise, bmc official’s insight on efforts, approach & challenges for curbing spread.[internet]. 2020 [cited 11th july, 2020]

Indiacom(b) (2020) Coronavirus in india: These 7 cities contributing 60 per cent of overall cases.[internet]. 2020 [cited 7th june, 2020]

Korobeinikov A (2009) Global properties of sir and seir epidemic models with multiple parallel infectious stages. Bulletin of mathematical biology 71(1):75–83

Li MY, Graef JR, Wang L, Karsai J (1999) Global dynamics of a seir model with varying total population size. Mathematical biosciences 160(2):191–213

Rocklöv J, Sjödin H (2020) High population densities catalyse the spread of covid-19. Journal of travel medicine 27(3):taaa038

Röst G (2008) Seir epidemiological model with varying infectivity and infinite delay. Mathematical Biosciences and Engineering 5(2):389–402

Smith D, Moore L (2004) The sir model for spread of disease: The differential equation model. Loci(originally Convergence) https://www.maa.org/press/periodicals/loci/joma/the-sir-model-for-spread-of-disease-the-differential-equation-model

Stehlé J, Voirin N, Barrat A, Cattuto C, Colizza V, Isella L, Régis C, Pinton JF, Khanafer N, Van den Broeck W, et al. (2011) Simulation of an seir infectious disease model on the dynamic contact network of conference attendees. BMC medicine 9(1):87

Times E (2020) Asia’s largest slum dharavi reports first covid-19 casualty.[internet]. 2020 [cited 2nd april, 2020]