Statistics and Stories:
A multidimensional risk dashboard for COVID-19

Sanjana Krishnan, Peiling Yap, Sahil Deo, and Amandeep Gill.
Statistics and Stories: A multidimensional risk dashboard for COVID-19

Micro-narratives and hyper local data to inform multidimensional risk dashboarding for decision makers in Indian cities, with a focus in Pune, Maharashtra

Abstract:
Numerous digital health solutions and dashboards have emerged in response to the COVID-19 pandemic. However, several of them operate in extremely data scarce environments and fail to consider the realities of the systems they are embedded in, prompting the need for a new data imagination.

This ‘stories and statistics’ report explores the value of synthesising qualitative and quantitative data in digital-health dashboarding for COVID and other emerging infectious diseases (EIDs). It proposes a framework for creating a multidimensional pandemic risk dashboard for Pune city that considers spread of disease, health infrastructure capacity, vulnerabilities and expert opinions. It tries to understand how qualitative information in the form of micronarratives could provide missing data and inform COVID treatment, help understand transmission patterns of the virus and formulate policy responses.

The findings of this study reveal the advantages and limitations of our method. It highlights the need for representative, granular, qualitative data and data pipelines and demonstrates how well-designed dashboards could be useful for decision making, designing health interventions and for sharing learnings.

This report hopes to contribute to the knowledge base on digital-health dashboarding for emerging infectious diseases (EIDs), and demonstrate the use of ‘micronarratives’ in improving the effectiveness of dashboards.

The COVID-19 pandemic has brought the world to a halt and created a ‘new normal’ like never before in our living memory. While there have been several emerging infectious diseases (EIDs) over the last few decades, they have either been geographically confined (Ebola), contained early (2002 SARS-Cov), or had fewer hospitalizations (2009 H1N1 pandemic1) and a lower fatality rate. In contrast, the SARS-CoV-2 or the ‘novel’ coronavirus

---
1 Requirement for intensive care: For H1N1, 1 in 104,000 required intensive care as opposed to SARS-CoV-2 where 1 in 16000 require intensive care (Petersen, Koopmans, et all, 2020)
is extremely infectious, has a much longer incubation period, and is acute in its progression, resulting in a faster epidemic formation and a rapid spread globally.

This posed an exceptional challenge to healthcare systems across the world, forcing several counties to ‘lockdown’, suspend mass transportation and economic activity in order to arrest the spread. Despite governments efforts to limit the economic shock, the World Bank estimates one of the deepest global recession in decades, with a 5.2% contraction of the global GDP in 2020. But the health and socio-economic fallout of this pandemic is not uniform. Older people and those with comorbidities are facing a greater risk. Certain regions in the world appear to have a higher fatality rate than others. Some livelihoods, sectors and informal workers have been more adversely impacted. There is a staggering impact on the education of millions of children. And overall, the pandemic has disproportionally affected the already disadvantaged communities, further exacerbating inequalities in society.

However, this unprecedented disaster has also elicited an extraordinary global response. Apart from responses from governments, communities and philanthropies, a multitude of scientific and technological efforts have helped accelerate the response to the virus. Test kits were developed and manufactured rapidly, vaccines are being developed at record speed and numerous technological interventions emerged to trace, monitor and track the spread of the virus. Digital health interventions - from data-decision dashboards to AI-powered voice-based detection of COVID - have proliferated and aided the fight against the COVID pandemic as never before in our battle against microbes.

Unfortunately, several of these interventions are unable to bridge the digital divide, and as often observed, the socioeconomic status is directly linked to the health status and access and use of digital health technology. Contact tracing apps for instance, work only for those with smart phones and internet connectivity. Apart from this, ‘system-level challenges’ are common in several LMICs. Interventions formulated on the bedrock of inadequate data can result in a suboptimal development and use of these digital health interventions may even deepen the digital divide. For example, given that development-related data is often riddled with issues of scarcity, accessibility, granularity, and timeliness, data-driven digital health decisions often neglect those who are left out of the purview of traditional systems. Modelling and decision making in the absence of representative datasets could lead to inaccurate results.

In this context, the current study tries to bridge the data-divide while creating a digital health dashboard for COVID. It does so by examining two distinct sources - statistics and

---

2 SARS-CoV-2 has the highest average R0 of 2.5, greater than 1918 Spanish influenza (R0 of 2) and the 2009 H1N1 pandemic (R0 of 1.7). (Petersen, Koopmans, et al, 2020)

3 SARS-CoV-2 has an incubation period of 4-12 days compared to 2 days for H1N1 and 2-7 days for SARS-COV. (Petersen, Koopmans, et al, 2020)

4 Interval between symptom onset and maximum infectivity is 0 days for SARS-CoV-2, 5-7 days for SARS-CoV and 2 days for H1N1. This means that it is harder to contain. (Petersen, Koopmans, et al, 2020)
stories. The first uses hyper local, multidimensional data from the lens of ‘risk’. The second source relies on ‘narratives’ from various stakeholders, providing a grounded view of ‘missing data’ to plug the gaps in our knowledge. This report hopes to contribute to learning & knowledge sharing and to the literature on EID dashboarding and digital health.

The background: How we got here, what are the broad aims?

In April 2020, the International Digital Health and AI Research Collaborative (I-DAIR) started to pilot projects in cities around the world with a dual objective of exploring (i) Real time epidemiology and dashboards (RTED) and (ii) use of ‘micronarratives’ as a qualitative method to help plug data gaps, validating policy utility of AI models and provide policy insights.

RTED aims to do scenario modelling, real-time data integration and develop dashboards to predict pandemic spread and system burden. It attempts to decipher transmission networks and population vulnerabilities, taking sociological and anthropological factors into account. This is crucial to help policy makers answer questions such as: What is the expected burden of testing, hospitalization, and critical care? Where should we focus limited enforcement and problem-solving resources? How effective/ineffective have restrictions been on movement? How can we take differential action in easing these restrictions?

Jelc and Fabiszak (2019) define micronarratives as “one or more clauses on an identifiable topic, embedded in a social context, where meaning is created or co-created by one or more speakers”. Operating in a data scarce environment necessitates the need to employ innovative methods to plug the knowledge gaps. Everyday experiences of millions of frontline workers fighting the virus generate a huge amount of fungible knowledge. However, this knowledge is produced and shared in silos and is rarely systematically captured and widely disseminated. Micronarratives could provide the basis for a methodology to verbalize and formalize the learning and tacit knowledge co-produced by expert stakeholders.

I-DAIR proposes to build such a methodology, with the use of AI for analyzing of micronarratives at scale for collaborative research and learning. The questions sought to be answered are as follows. Can effective knowledge sharing mechanisms be instituted so that knowledge, best practices and learnings across geographies can be captured and disseminated in an efficient and quick manner? Could this qualitative information be summarized and made intelligible through a dashboard? Further, could this be used to validate or improve statistical models and provide a guide for data collection?
With these two broad aims, the project started with a pilot in Pune City\(^5\) to understand how local data and micronarratives can contribute to the development of EID dashboards, models and knowledge base, with the aim of delivering useful and actionable results to decision makers.

**Our approach**

Managing the pandemic in India has been more than managing the public health crisis. It also involves managing the indirect effects of the pandemic on the socio-economic front. India has severe differences in economic and living conditions, and access to healthcare and other basic amenities. This implied judicious planning and use of limited health system resources. Multiple efforts were therefore made at a sub-national and national level to assist policy makers plan health infrastructure capacity based on advanced quantitative modelling. These do add value to the understanding of health-related capacity indicators especially at the macro-level. However, as of now, they are of limited use at the micro-level at the ward, which is the smallest administrative level in Indian cities.

One of the reasons that these models do not perform well when we zoom in geographically is driven by the heterogenous nature of India, the stark contrasts that exist begin to play out statistically. Further, these models and data systems often do not include several vulnerable and often marginalized population groups, reflecting the digital divide and the problem of missing data. As we go forward in dealing with this pandemic, decisions need to be made rapidly in a decentralized manner, and digital support tools need to become granular in parallel.

With the longer term I-DAIR objective to nowcast and forecast using multidimensional real time data, dashboards and scenario models in hyper local settings, the project started out with the limited aim of understanding how stories and statistics can help create a multidimensional pandemic dashboard for Pune city that considers spread of disease, health infrastructure capacity, vulnerabilities and expert opinions.

More specifically, the aims were to:

- Understand and demonstrate the effects and use of hyperlocal and vulnerabilities data in city level dashboards and models
  - How can local and vulnerabilities data give a more comprehensive view of the city?
  - How can this data be made useful for policy/ decision makers? (spread of disease, estimating the need for healthcare infrastructure, infrastructural asymmetries)

---

\(^5\) Pune city is located in western Maharashtra and has a population of 3.12 million. It reported the first confirmed cases in Maharashtra, and at the time of writing the report, Pune district had the highest number of cases in India.

www.cpc-analytics.com |
mail@cpc-analytics.com
• Understand and demonstrate how stories can be used to complement and supplement the statistics
  o What treatment protocols, epidemiology relevant insights and policy recommendations can doctors and COVID-19 survivors provide?
  o How can such information be curated, processed and presented through a dashboard?

The study uses mixed methods, integrating quantitative and qualitative research methods. The following two sections of the report detail the findings from the ‘statistics’ and from the ‘stories’.
Statistics: Using multi-dimensional indicators to assess the pandemic risk at a hyperlocal scale

The COVID-19 hazard has affected populations around the world. A hazard is any source of potential damage, harm or adverse health effects on something or someone. COVID-19 is a biological hazard viii caused by the SARS-CoV-2 virus that has been detrimental to development and resulted in a negative impact on lives, economy and society. The magnitude of the COVID-19 hazard on the health of populations at a hyperlocal scale can be measured by the incidence of cases or deaths in various localities within the city.

When the same exogenous hazard confronts a population, the ability of the system to manage and contain it becomes extremely important. This ability to plan and build the necessary infrastructure and services to deal with the hazard is known as the ‘capacity’ of the system or the ‘resources’ that the system has. For instance, a healthcare system capable of providing adequate and accessible care to the entire population is more capable of dealing with the surge coming from a pandemic. A governance system capable of providing the requisite relief to those facing economic hardships during an economic downturn has a greater capacity to mitigate the social and economic effects of the pandemic.6

Apart from the exogenous factors, the inherent, intrinsic ability of the system to deal with a hazard is an important risk factor. Vulnerability7 is related to “predisposition, susceptibilities, fragilities, weaknesses, deficiencies, or lack of capacities” that results in adverse effects on the elements exposed to a hazard.ix This could arise due to physical, social, economic, environmental or political factors. The inherent nature and spatial characteristics of cities makes them susceptible to a pandemic. As the hubs of globalization and economic activity, they have been the first victims of COVID-19 given their density of living, networks of public transportation and their characteristics as nodes of domestic and international travel. This has also been true for Indian cities, which face additional challenges of overcrowding, significant percentage of populations residing in slum settlements and inadequate health infrastructure.x In the context of COVID-19, patients with comorbidities are more vulnerable because they have a lower resilience to the virus. Studies have also shown that the populations with higher living density and lower access to adequate WASH infrastructure have a higher exposure to infectious diseasesxii,xii, making them more vulnerable. A study by the Lancet for example created a vulnerability index to identify vulnerable regions in India that could be strongly impacted by the pandemic.xiii

These three components; the hazard, vulnerability and capacity of a system; give the concept of risk. In literaturexiv,xv, risk is defined as:

---

6 Germany’s budget surpluses over the preceding years have given it greater margin of manoeuvre in stimulating its economy than its European neighbours.
7 Vulnerability is a sum of the exposure to hazard, the resistance (measures taken to prevent, avoid or reduce loss) against the hazard and resilience (Ability to recover prior state or achieve desired post-disaster state) to the hazard.
Risk = \frac{Hazard \times Vulnerability}{Capacity}

Flanagan et. al (2011)\textsuperscript{xvi} use the following formula in disaster management research to quantify the risk

\begin{equation}
Risk = Hazard \times (Vulnerability - Resource)
\end{equation}

Indicating that the ‘risk’ that a system is exposed to is directly proportional to the magnitude of the hazard and vulnerability, and inversely proportional to the capacity or resources that the system has to mitigate the hazard.

With this framing of risk, we collated and analysed data on the above-mentioned three indicators to arrive at a multidimensional pandemic risk assessment for wards in Pune city. In our study, we have selected 15 sub-indicators that are described in Table 1. Given the increase in EIDs and other natural disasters like floods, the need for such ‘risk response dashboards’ that considers multiple facets of the risk (beyond the magnitude of hazard) gains increasing significance.

Hazard mapping

Figure 1: COVID Hazard in Pune wards. Total number of cases in wards (left) and 2-week growth rate of cases (right)

Other hazard indicators include active cases, recovered and deaths.

![Hazard mapping](source: PMC | Data as of 23rd September 2020)

Figure 1 shows two hazard indicators for Pune at a ward level. As the maps show, the case load is highest in Hadapsar and Dhanakwadi-Sahakarnagar. While the cases were initially concentrated in the city centre, outer Pune shows the fastest rate of growth now. As in many cities, the virus is spreading outwards from dense city centres to the peri urban and then rural areas. Other hazard indicators, including the number of active and recovered cases and number of deaths, can also be seen on our dashboard.
An index for the hazard indicating the magnitude or impact of COVID was constructed by taking the average of the normalized values of the number of active cases and the 14-day average growth rate of cases in each ward. Variables were normalized using the formula:

$$x_{\text{normalized}} = x_N = \frac{x_i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}; x_N \in \{0,1\}$$

$$\text{Hazard Index } (HI) = \frac{(\text{Active cases})_N + (\text{Growth of cases})_N}{2}$$

The normalized values were used to get an index between 0 and 1. A higher value indicates a larger hazard. The index ‘0’ would indicate that the ward has the lowest number of active cases and the lowest rate of growth. The index ‘1’ would indicate largest number of active cases and the fastest rate of growth of cases.

Vulnerabilities mapping

Figure 2: Spatial, Water Sanitation and Hygiene (WASH) and health vulnerabilities in Pune wards. (i) Spatial vulnerability- density of living in slum settlements, other spatial vulnerabilities include population density; (ii) WASH vulnerability- Average value of the number of shared toilet/latrine facilities per capita in slum settlements and the number of tap points/ public hydrants installed for supply of protected water per capita in slum settlements; (iii) Health vulnerability- Number of people with comorbidities, other health vulnerabilities include the number of pregnant women in the ward; (iv) Composite Vulnerability Index- index capturing the spatial, WASH and health vulnerabilities for the wards.
While vulnerabilities are due to various factors, we identify and map three types of vulnerabilities—spatial, health and infrastructural vulnerabilities. Figure 2 maps the number of comorbid patients in each ward. This has important governance implications for decision makers to ensure that there is greater vigilance, sanitization and screening in wards where there is a larger proportion of vulnerable population. Figure 2 also maps the WASH vulnerability in wards. This shows the lack of water, sanitation and hygiene infrastructure in slum settlements, with a large number of people sharing an inadequate number of community toilets and public taps.

Additionally, information on spatial vulnerabilities including population and slum density and health vulnerabilities (the number of pregnant women) was identified and mapped. At a decision-making level, these areas require additional resources and mitigating measures to ensure that the spread of the disease is arrested, and the number of serious infections and deaths is reduced.

The index for vulnerability was calculated by taking an equally weighted average of the normalized values of the six parameters use to estimate the vulnerability—population density, slum population density, toilets per capita, public taps per capita, number of comorbid patients and number of pregnant women.

Variables were normalized using the formula

$$x_{\text{normalized}} = x_N = \frac{x_i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}; \ x_N \in \{0,1\}$$

Spatial vulnerability = \(\frac{(\text{Population density})_N + (\text{Slum population density})_N}{2}\)

Health vulnerability = \(\frac{(\text{People with comorbidities})_N + (\text{Pregnant women})_N}{2}\)

WASH vulnerability = \(\frac{(\text{Public taps per capita in slums})_N + (\text{Shared toilets per capita in slums})_N}{2}\)

Vulnerability Index (VI)

= \(\frac{(\text{spatial vulnerability})_N + (\text{Health vulnerability})_N + (\text{WASH vulnerability})_N}{3}\)

The vulnerability index ranges between 0 and 1. Index ‘0’ would indicate that the ward scores lowest on all six indicators related to vulnerabilities, an index of ‘1’ would indicate that the ward scores the highest on all six indicators. A higher value indicates a higher vulnerability.

**Capacity mapping**
Pune had 3179 Oxygen beds, 405 ICU beds and 515 ventilator beds as of 1st October 2020.\textsuperscript{a, b, c} Analysis of hospital data in Pune from May to October 2020 shows that of all the active cases, 17.6\% require oxygen, 4.04\% require ICU and 2.13\% need ventilator assistance. Weekly moving average were used to arrive at short term forecasts of the number of active cases. Based on this and the percentage of active cases requiring hospitalization, the need for healthcare infrastructure was predicted.

**Figure 3: Healthcare infrastructure capacity and projections**

There would be 18490 active cases on 31st October. Vertical lines show the lockdown.

![Graph showing healthcare infrastructure capacity and projections](source)

We estimated the number of active cases in three scenarios- status quo, best-case and worst-case. Figure 3 shows the availability (blue line) and requirement (red line) for oxygen beds, ICUs and ventilators. In Pune, the capacity was progressively raised with time even though a shortage of beds forced Pune city to go into a ‘re-lockdown’ in July.

Apart from this, data from Pune showed that the number of active patients requiring ICUs and ventilators decreased initially, but later increased. The criticality rate (patients requiring ICUs and ventilators) is currently at 5.5\%.

**Figure 4: Age wise distribution of cases and deaths**

x-axis shows the age-group

![Graph showing age wise distribution of cases and deaths](source)

Source: PMC | Data and modelling as of 27th July 2020.
Pune city has a case fatality rate of 2.4% (reduced from ~6% at the start of the pandemic). 77% of those who died of COVID also had some comorbidity. Those older than 60 made up ~15% of the cases but resulted in about 62% of the deaths. CFR is 17% in age group 70+, 10.2% in 60-70 age group and 5.3% in the 50-60 age group. CFR is less that 0.2% in age groups below 30.

Risk mapping

The Risk Index was calculated by taking the average value of the Hazard Index and the Vulnerability Index. As the healthcare infrastructure is distributed across the city and beds are centrally allocated, locations within the city do not have any particular advantage.

\[
Risk \, Index \, (RI) = \frac{Hazard \, Index + Vulnerability \, Index}{2}
\]

Figure 5: Mapping the Hazard, Vulnerability and Risk Index in Pune wards

The risk index is a dynamic indicator and summarises 12 variables that are important for understanding and benchmarking the performance of wards. Figure 5 maps the vulnerabilities and hazard in Pune wards. The chart has been divided into four quadrants based on the median values of hazard and vulnerabilities. As of 23rd September, two wards (Dhankwadi- Sahakarnagar and Hadapsar-Mundhwa) faced a high risk because of a high case load and high vulnerability. Six wards had a high vulnerability and could face a high
risk if the number of cases increase. This risk-based distribution of administrative units allows public officials to plan and monitor the impact of the interventions on the hazard in near real time. For instance, if a ward with high vulnerability registers a high growth rate of cases, ward-level partial lockdowns can be planned till the cases are under control. Public officials can also add to the complexity of the risk-based index by adding more near-real time indicators.
<table>
<thead>
<tr>
<th>Category</th>
<th>ID</th>
<th>Indicator</th>
<th>Definition</th>
<th>Year/ source</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>COVID Hazard</td>
<td>1</td>
<td>Total Cases&lt;sup&gt;xx&lt;/sup&gt;</td>
<td>Reported total cumulative count of detected and laboratory confirmed positive cases</td>
<td>Daily, PMC</td>
<td>Ward</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Active Cases</td>
<td>Represents the current number of people detected and confirmed to be infected with the virus.</td>
<td>Daily, PMC</td>
<td>Ward</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Recoveries</td>
<td>Definition used by government unclear- WHO recommends following the criteria of [symptoms resolve + 2 negative tests within 24 hours] or [symptoms resolve + additional 14 days]. Could also be when a patient is discharged from the hospital// anyone who was diagnosed with COVID-19 three or more weeks ago and has not died</td>
<td>Daily, PMC</td>
<td>Ward</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Deaths</td>
<td>Cumulative number of deaths among detected cases.</td>
<td>Daily, PMC</td>
<td>City</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Average Growth rate of cases</td>
<td>The average increase in the number of cases over the period of n days.</td>
<td>Daily, PMC</td>
<td>Ward</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Test positivity rate</td>
<td>That’s the percentage of people who test positive for the virus of those overall who have been tested</td>
<td>Daily, PMC</td>
<td>District</td>
</tr>
<tr>
<td>Health System Capacity and performance</td>
<td>7</td>
<td>Case Fatality Rate</td>
<td>The proportion of deaths among the total diagnosed individuals</td>
<td>Daily, PMC</td>
<td>City</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Criticality Rate</td>
<td>The number of patients currently being treated in Intensive Care Unit (ICU) or ventilator</td>
<td>Daily, PMC</td>
<td>City</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Hospital Beds available and occupied</td>
<td>Total number of oxygen beds, ICU beds and ventilators available and occupied</td>
<td>Daily, PMC</td>
<td>District and city</td>
</tr>
<tr>
<td>Spatial vulnerabilities</td>
<td>10</td>
<td>Population Density</td>
<td>Number of people living per unit area</td>
<td>2011, Census</td>
<td>Ward</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Slum population density</td>
<td>Number of people living in slum settlements in per unit area occupied by the settlement</td>
<td>2011, Census</td>
<td>Slum</td>
</tr>
<tr>
<td>WASH vulnerabilities</td>
<td>12</td>
<td>Toilets per capita</td>
<td>Number of shared toilet/latrine facilities per capita in slum settlements</td>
<td>2011, Census</td>
<td>Slum</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Public taps per capita</td>
<td>No. of tap points/ public hydrants installed for supply of protected water per capita in slum settlements</td>
<td>2011, Census</td>
<td>Slum</td>
</tr>
<tr>
<td>Health vulnerabilities</td>
<td>14</td>
<td>Comorbid population</td>
<td>Number of people with comorbidities</td>
<td>June 2020, PMC survey</td>
<td>Ward</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Pregnant Women</td>
<td>Number of pregnant women</td>
<td>June 2020, PMC survey</td>
<td>Ward</td>
</tr>
</tbody>
</table>
Stories: Co-producing knowledge and data by collating the experiences of frontline workers and COVID survivors

As mentioned, in this study, we employ ‘micronarratives’ as a method to try and plug data gaps, validate models, provide policy insights, supplement and compliment existing quantitative data. To do this, we bring together experiences of doctors and COVID survivors to get a deeper, more nuanced understanding of the virus with the aim of contributing to the COVID knowledge base on three aspects:

- **Treatment protocols** - compiles the learnings that doctors have had by observing patients and reflecting on what has helped them treat patients in a more effective manner. Consequently, it provides insights on the protocols and best practices that need to be replicated and shared. It could also provide insights for modelling, for instance, knowing the kind and number of patients that require interventions could help build more accurate projections for healthcare capacity.

- **Epidemiology insights** - the ‘novel’ nature of the coronavirus took the world by surprise. Scientists and doctors knew very little about the virus and its effects, and research to understand it is still a work in progress. While the scientific approach to understanding the virus is paramount, there is value in understanding the virus through the experience of doctors treating patients and survivors who have recovered from the virus. For example, what is the observed effects of the disease on people, what is the effect on death rates of particular modalities of treatment, how is the infectivity dependent on factors such as living density in the eyes of the health workers on the ground? This could in turn provide information for modelling and policy. For instance, understanding the relation of infectivity with living density could help modellers have a more accurate projection of cases. This could also help policy makers plan interventions in such areas.

- **Governance/Policy recommendations** - As the situation evolves, there is a need for rapid decision making on various aspects like locking down, allocating human resource, purchasing hospital equipment etc. There is also a need for proper governance responses to the issues faced by doctors and frontline workers in order to provide all necessary resources to them to arrest the spread and ensure their wellbeing in the process. Micronarratives from frontline workers present data on aspects of resource crunch, high patient load, mental health of doctors and other issues. In addition, stories from COVID survivors could also help to highlight how the infection and lockdown had impacted on their longer term health needs and socio-economic status. Such qualitative data would allow decision makers to understand the ground situation their constituents are facing and make adjustments to their policies in a timely manner.

**Methodology in a nutshell**

For the micronarratives, we conducted in-depth interviews with 32 doctors. We conducted snowball sampling to find the doctors but tried to ensure representation across gender, experience, specialization and the type of their engagement with COVID patients. Ethics
guidelines and consent documents were sent to the interviewees before the interview. Semi structured questions across 5 themes (health infrastructure burden, treatment and testing protocols, understanding of the virus, effects of drugs and clinical interventions and personal experience) were asked in interviews ranging from 30-90 minutes. The interviews were audio recorded and conducted in English, Hindi and Marathi. We coded the answers based on common themes that emerged from the interviews and used methods from Qualitative Content Analysis (the process of decontextualization, recontextualization, categorization and compilation) to analyse the interviews. xxi

These interviews were supplemented with a structured questionnaire-based surveys of doctors, nursing staff, paramedics and with COVID survivors. This included 1512 healthcare workers and 2282 people who tested positive for COVID and recovered, across cities in Maharashtra. This was conducted by a leading Marathi daily newspaper.

**Table 2: Framework for analysing micronarratives**

<table>
<thead>
<tr>
<th>What it means</th>
<th>Treatment protocols</th>
<th>Epidemiology Insights</th>
<th>Policy recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>What insights do doctors and patients have on the effects of the disease and effectiveness of treatment based on their experience.</td>
<td>What epidemiological insights do micronarratives reveal about the incidence, spread, and possible control of diseases.</td>
<td>What insights about hospital management, infrastructure burden do doctors provide that could help guide decision makers.</td>
</tr>
<tr>
<td><strong>Insights on</strong></td>
<td>Need for medical assistance</td>
<td>Risk factors of COVID</td>
<td>Resources required/ interventions needed</td>
</tr>
<tr>
<td></td>
<td>Effectiveness of medical assistance</td>
<td>Assessment of virus over time</td>
<td>Patient management</td>
</tr>
<tr>
<td></td>
<td>Long term effects of virus</td>
<td>Relation between spread and living density</td>
<td>Wellbeing of frontline workers</td>
</tr>
<tr>
<td></td>
<td>Causes of death</td>
<td>Impact on different age groups</td>
<td>Impact on non-COVID care</td>
</tr>
<tr>
<td></td>
<td>Treatment strategies</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Relevance for decision making</strong></td>
<td>Purchase of equipment, sharing learnings/best practices across geographies</td>
<td>Reducing risk factors and targeting interventions on vulnerable populations</td>
<td>Addressing needs of frontline workers</td>
</tr>
<tr>
<td><strong>Possible insights for modelling</strong></td>
<td>Projecting need for healthcare infrastructure</td>
<td>Understanding infectivity at a granular level, projections of deaths</td>
<td>Need for lockdown based on capability of system to handle load, ability to meet hospital needs and effectiveness of interventions</td>
</tr>
</tbody>
</table>

**Treatment protocols**

- Patients requiring assistance: Detailed interviews with doctors revealed the percentage of mild, moderate and severe patients requiring oxygen, ICU and ventilator assistance. 7-10% of mild patients, 50-60% of moderate and 100% of severe patients required oxygen. Only about 3% of the mild case progressed to requiring ICU or ventilator. Of the moderate cases, 10% of the young and 30-40% of the old patients deteriorated to require ICU and ventilators. Doctors also reported that 60% of severe patients died and 90% of those on ventilators died. Doctors reflected on using methods like the 6-minute walk test to understand oxygen saturation and monitor patients.
- Long term effects: While all doctors agreed that it is very early to determine the long term effects, they observed some effects after the patient tested negative. Nearly 20% of the doctors interviewed observed some form of lung fibrosis and they believed that it was rare to restore normal lung function. Apart from this, compromised cognitive functions, difficulty and pain while breathing, and cardiovascular damage were also observed. Out of the COVID-19 survivors, the survey showed that 28% of those above 60, 20% of those between 40 and 59 and about 11% of those younger than 40, reported tiredness and fatigue. 4% of those above 60 and 2% between 40 and 59 reported having difficulty in breathing. Looking at research publications on fatigue, it was observed that 16.3% of survivors in China had persistent symptoms of fatigue.

- Leading cause of death apart from comorbidities: While comorbidities pose the highest risk and about 75% of those in Pune who died also had comorbidities, doctors observed that the leading cause of death apart from comorbidities was late admissions. This raises the need for early testing and admissions to arrest the spread of complications and maintain oxygen levels. Apart from this, stress, lung diseases and any other condition which makes the patient immunocompromised could result in complications and deaths.

- Effectiveness of medical assistance and treatment strategies: Doctors found oxygen to be the most effective intervention, especially when it is given early. There is a narrow window when treatment works, and if drugs are given in the right setting, the number of complications could be reduced. Invasive ventilators are seen as the very last resort, as very few of the patients survive thereafter. Some doctors felt that it was tough to pinpoint what exactly works, as they offered all treatment early and in full to prevent patients from becoming serious, given the shortage of ICU beds. Only a few doctors reported that they counselled and motivated patients, especially those who were scared or older and faced loneliness in the wards.

**Epidemiology relevant insights**

- Understanding the risk factors of COVID: Most doctors felt that the high infectivity of Sars-Cov-2 is what makes the virus very dangerous. Apart from this, there is a lack of knowledge because of the ‘novel’ nature of the disease and the unpredictability in its symptoms and outcomes. This combined with a longer recovery period and higher mortality especially for those with comorbidities increase the risk associated with COVID.

- Reasons for a reducing case fatality rate: Over time, globally and in India, data shows that the CFR has reduced. Most doctors observed the same. In their assessment, one of the main reasons for this is that they had more knowledge and drugs with time. Seeing more cases and observing the effects of treatment, they feel more confident and get better at handling patients. Their gut instinct as a clinician gets better. They also observed that increased testing, early detection and treatment helped reduce deaths, along with preparedness and protocols that have improved with time.

- Assessment of virus over three months: As compared to the start, doctors feel more confident and prepared, the initial fear has reduced. They have observed several new and diverse symptoms that has made it both easier and tougher to diagnose the disease without testing.
▪ How the infection has spread: About 40% of COVID survivors positive do not know the source of their infection. 30% reported that family members contracted the virus before them. Around 8% reported that the source was through a gathering.

▪ Relation between infectivity and living density: COVID survivors were asked about the number of people they share a room with and the number of them who were infected before or after they tested positive. It was observed that of those who did not share a room, 76% did not infect anyone and just 5.2% infected more than five people. In contrast, among those who shared a room with more than five people, 14.6% infected more than 10 people, 15.5% infected 5-10 people and 32% infected between 1 to 4 people. This provides insights on the relation between infectivity and living density.

▪ Observed effects on children: Doctors who interacted with paediatric COVID cases observed that children recovered very well and encountered no complications. Gynaecologists also observed that pregnant women with COVID mostly did not transmit the virus to their fetus.

Policy relevant recommendations

▪ Resources required/ interventions needed: Most doctors reported a need for more equipment and beds. Several doctors also raised the issue of access and affordability of drugs. A few hospitals did not have the required specialized drugs like Remdesivier. Others reported the high cost of each drug and their limited availability as reasons to pause before administering the drugs, as they had to consider the ability of the family to afford the treatment and the probability of the drug saving the patient's life. The need for adequate secondary health workers and staff, like nurses and ward boys, along with adequate facilities for them (like resting rooms, shorter shifts) was raised. Doctors working in private hospitals felt the financial strain that hospitals were under due to reduction in non-COVID care and regularization of beds and hospital rates, and raised the need for monetary support and compensation.

▪ Issues with patient management: A few doctors observed that the fear, anxiety and depression that patients felt was pressing. Additionally, patients felt stressed because of the fear of high medical bills and loss of livelihood; this hindered the recovery from COVID. Commonly, patients did not wear masks, and a few even hid symptoms. Further, doctors often found it very difficult to communicate with patients when they were in their PPE suits.

▪ Status of other healthcare workers: Survey of healthcare workers showed that paramedics had the lowest access to PPE kits and the highest rate of infection, with about 20% of paramedics contracting the virus. In contrast, very few doctors faced a shortage of PPE kits and very few of them contracted the virus. This raises the need to ensure that other healthcare workers also receive the protective gear that is required.

▪ Mental health of frontline workers: Most doctors interviewed reported that their mental health was either not very affected or not affected at all. Most of those who were worried feared carrying the infection to their family. Few doctors were affected by seeing the way patients suffered and the lack of facilities for them. Resident doctors were worried about their examinations. Other issues reported included isolation and managing hospital finances and politics.
The micronarratives in our study provide qualitative nuances to the government and hospital statistics and data. For instance, data shows trends in CFR and the number of comorbidities related deaths. But conversations with doctors and those infected reveal how the stigma and fear surrounding COVID deter those infected from accessing timely treatment and how late admission increases the number of avoidable death.

Micronarratives also help produce new knowledge to bridge missing information. Even though the government does not record COVID infection rate among frontline health workers, a sample survey indicates that that paramedics have the highest percentage of infections, followed by nursing staff and then doctors. Similarly, reports have revealed long term effects of COVID. 

Our survey was able to identify the proportion of patients who experienced fatigue in various age groups in Maharashtra.

Finally, a sole reliance on numbers takes away the inherently human elements of a world dealing with a deadly and novel virus. Qualitative inputs have an emotive power and the ability to humanize data. The experiences of frontline workers and risk taken by them, even if anecdotal, could evoke responsible behaviour from people and expedite governance decisions that eases the burden on health systems.
‘Micro-narratives’ - quotes from interviews

**Why a private doctor joined the governments COVID duty?**

I went to the OPD in Kasturba hospital (one of the major governments COVID hospitals in Mumbai city) because I had symptoms and saw a huge shortage of doctors. I was anxious because of my weakness, fever, and bronchitis, and worried about “What will happen if the doctor is not there?”. Thankfully, I tested negative. After recovery I immediately joined the government duty. Doctors need to take a ‘calculated risk’ and come out to work. “Satark rehna chahiye, but kaam bhi karna chahiye” (should stay cautious but should also work).

**Difficulties faced while working during COVID**

Medical examinations are hampered. Patient care is hampered because doctor to patient contact has reduced. It is difficult to communicate and understand patients to get their medical history. Also, for tests you cannot use stethoscopes.

Being a doctor, patients really want to hear words of comfort from you. With the PPE suit, it is very tough to stay for enough time or speak to them. They cannot even see your face.

**Difficulties faced while working during COVID**

Sometimes PPEs have tears, we just put cello tape and use it. Some batches are short in length and because I’m very tall, there is a gap between the gown and the boots. I wrap plastic sheet when this happens.

**Impact on mental health care**

With the lack of regular check-ups symptoms might be worse now. There will be a lot of relapses for those without medication, and those who are not coming to refill their prescriptions. Mood and psychotic symptoms have returned for many who were earlier doing well. Those well-off have other alternatives. Private practice are running. But it is not affordable for poor. The poor and people from lower socioeconomic strata are hit harder.

**Effect on non COVID care**

700-800 patients in cardiology have reduced to 25-30 patients, especially because majority of the cardiology patients came via public transport. Cardiac arrest patients have suddenly dropped. We are not sure what the reason is.
Elective surgeries can be postponed, but not indefinitely. And emergencies cannot be avoided. People are not reaching hospital as ambulance is not available. Dialysis patients are dying as it cannot be postponed more than a week. For a few weeks we barely had stroke patients.

Some patients seem okay and in a matter of 2-3 hours they deteriorate and become serious. This could be due to some undiagnosed comorbidity, obesity. If there is breathlessness, sometimes it can go bad very soon. Pneumonia keeps increasing sometimes and the patient needs intubation or more aggressive treatment. There is a window of treatment where it works. Sometimes the body does not accept it. After a point, you can understand a patient cannot survive.

Remdesivir costs Rs.6000/injection and needs to be given for 5 days. Tocilizumab costs Rs.60,000 and the government has subsidised it to Rs. 45,000. This is very, very expensive, and also the stock in limited. As doctors, we have to think twice before giving critical patients the drugs. We have to have some certainty that the patient will definitely improve. We ask the patient’s family and they are desperate to save them but they also don’t have the financial capacity to pay for the treatment. It is very tough to act and make decisions, it is a question of the patients life.

At some point, we thought we will contain the virus in a few months, that there are only a few cases. It was definitely worse than a few, but it is also definitely not as bad as the paranoia that first came, resulting in very strict restrictions and a lot of stigma. The truth about the virus is somewhere in between. It is difficult to communicate that. People find the paradox tough to accept.
The first time I wore a PPE, my heart rate went through the roof. My parents were also very worried about me. Now, after a few months, I am comfortable donning and docking the PPE, and wearing it for 7 hours, or even 12 hours if required.

I was scared—every healthcare worker should be worried about getting it. We should always take the necessary precautions, and a little anxiety is normal.

I have a 2-year-old kid that I haven’t met for 3 months. I go home to tell him hi from the gate and come back to the hospital.

Yes, I am worried about spreading it to my father as he is also a police officer on duty and is mildly diabetic. There is this constant worry, which drains you.

Yes, I am prepared to work. If not us, who will?

Very tough to say. At this stage, probably an astrologer can say with more confidence. It is like the game ‘chain reaction’. Everything can seem good and suddenly a hotspot or super spreader can come up and it can become very tough to control.

Yes, I am prepared to work for another year or two if this continues?
Conclusions

In this study, we put forth an approach to organize relevant information for digital health dashboards in the context of COVID-19 and other EIDs. We propose the use of multidimensional risk parameters for a comprehensive understanding of the hazard, capacity and vulnerabilities of a system. We also demonstrate how qualitative information in the form of micronarratives could provide missing data and inform COVID treatment, help understand the severity of the viral infection and guide policy response.

By combining quantitative and qualitative information, this approach could help provide actionable insights useful for quick and effective response and decision making. It could help frame policy, give direction to governance, prioritize action, target interventions, inform resource planning and allocation, and assess the efficacy of interventions. It would also be useful for modelling and creating better digital health dashboards. In the course of this pandemic, it has been observed that several models have a high bias, poor performance and are not fit for use. Further, these models do not account for several variables that are endemic to cities and play a huge role in transmission. Having such data could help modellers create more localized, accurate and representative models. Benchmarked micronaratives could also help validate and improve these models, consequently increasing the trust in them.

This approach could also contribute to learning and preparedness. The virus is spreading from metros and tier 1 cities to smaller cities and rural areas. It is also possible that a ‘second wave’ could emerge in several cities. In such a case, the learnings, coordination efforts and knowledge generated across geographies could be documented. Guidelines could be created to provide other cities with a framework for approaching the problem and preparing multiple defences and systems against the virus. This would also help prepare for future pandemic like situations and other emerging infectious diseases.

However, to implement sustainable and reproducible data solutions, we need to address several data issues that we encountered in our study, which could also be endemic to data systems in other cities in India or elsewhere.

The data used needs to be representative and reflect the diversity, realities and complexities of the systems that it is generated from. Data is very often not granular and not available or collected and aggregated at the required geographic unit, which hinders our ability to make localized decisions. Several data points are also not collected and updated at a regular frequency, which implies that the data could reflect the city as it looked a decade ago. In terms of access to this data, it is often observed that the data produced is not released in a usable format, very often data is released as images or as graphs which necessitate manual intervention to prepare the data for analysis. The metadata and definitions of variables is often incongruent or missing. Software interface and data models used in different government departments need to be standardised and
Interoperable. The lack of this makes collating and comparing data from different departments difficult.

Apart from this, several important indicators are often missing or scarce. This is true especially at the start of a pandemic or disaster when the data that emerges is constantly evolving, is often noisy, inadequate and riddled with errors. There are aspects of the impact and of needs that cannot be quantified with current data; simple models cannot accurately capture and account for them. This raises the need to incorporate qualitative data. Finally, all of this needs to be tied together with data pipelines that connect siloed data sources and ensure smooth flow of qualitative and quantitative data from source to system, and further to decision makers and policy researchers.

We have adopted several measures to address some of these shortcomings and gaps. Our proposed risk response dashboard includes data from several sources that are hard to access and rarely used in mainstream governance. To streamline statistics, we undertook a painstaking process of piecing together hard-to-access, siloed datasets. We geocoded data and other information to understand the spatial and hyperlocal trends. We also relied on micronarratives to plug gaps in traditional data and get a more nuanced understanding of the effects of the virus.

However, our approach has limitations and there is lot of scope to improve and expand on it. Several more data points need to be incorporated into the ward level risk model. Datasets such as transport networks or formal and informal resource networks can be identified, cleaned and mainstreamed in decision making. Micronarratives in their current form are tedious to obtain and analyse. The process needs to be streamlined and rapid data collection methods need to be standardized to ensure that the method is replicable and robust. Further, we also need to develop methods to validate and benchmark these results with scientific studies. Finally, methods to regularly update, quickly summarize and efficiently represent such stories in dashboards requires research.

In the backdrop of this pandemic, there is immense scope for reimagining data systems in Indian cities such that they mainstream representative, granular, qualitative and timely data into the planning imagination. Strengthening and maintaining such data systems using time-tested technology could be instrumental for responding to such events and prove as useful as contact tracing apps, drone-based surveillance and other such digital health interventions.

**Next steps and taking this forward**

There is an urgent need for a response that also considers multiple emerging dimensions of the virus and its effects on individuals and society. With the progression of the disease, a possible second wave and availability of vaccines, these dashboards also need to adapt to consider the upcoming future challenges such as access to vaccines. As the disease progresses to rural areas that have limited health infrastructure, could we reallocate resources?
resources and health workers from cities that have crossed their peak? If yes, how do we plan it? What is the immunity level of the population as indicated by antibody/seroprevalence surveys, can normalcy be returned as we approach herd immunity thresholds? Based on this, how should the delivery of vaccines be targeted? These are pertinent questions that require ex-ante planning and dashboarding.

Based on the findings from this report, we are developing a real time dashboard on multidimensional risk factors for Pune city. Through this, we hope to contribute to the larger literature on pandemic and EID dashboarding as a digital health intervention. Taking the micronarrative methodology forward, we are also benchmarking the process.

Bhatt, M. writing about disaster management in India reflects on how post-disaster response still means ‘rebuilding what has been destroyed’ instead of ‘developing affected areas and communities’ to what it could and should be. The pandemic has exposed fault lines in our ailing public health and data systems, but also provides us an impetus and opportunity to reimagine and reform them. Overcoming this pandemic, we will hopefully emerge with stronger systems that are rich with learnings and capacitated with multi-modal data and interoperable digital infrastructures to support the road to recovery and the fight against EIDs.

The multidimensional risk dashboard can be found here.
Appendix

The data and method used for statistical analysis for the study can be found in the supplementary appendix.

Acknowledgements

We thank all the doctors and patients who participated in the study, and specifically the 32 doctors who gave detailed interviews and shared their experiences (this includes 6 doctors who requested anonymity): Dr. Aditya Mapari, Dr. Ajinkya Deshmukh, Dr. Amit Mane, Dr. Anuja Borkar, Dr. Atul Tat, Dr. Avinaskulkarni, Dr. Bharat Purandare, Dr. Chaitanya Chiplunkar, Dr. Gauri Mankar, Dr. Kalia, Dr. Kapil Zippe, Dr. Lancelot Pinto, Dr. Meeta Nakhere, Dr. Payal Fenders, Dr. Pradyumna Singh, Dr. Pranaya Prabhu, Dr. Pratik Patil, Dr. Sagar Thote, Dr. Sarmath Thakkar, Dr. Sanjeev Palta, Dr. Shailesh Ingle, Dr. Sneha, Dr. Sonia S Ingle, Dr. Sushant VP, Dr. Uma Arulkar, Dr. Yashendu Sarda

References


xix Data for Pune district as per “COVID CARE SOFTWARE- PUNE DIVISION”. Accessed on October 1, 2020, from: https://www.dlvcommunepunevid.com/ccsbedashboard/hsr

xx Definitions of total, active, recovered, deaths and critical cases taken from “Worldometer COVID-19 Data” on Sep 2, 2020. URL: https://www.worldometers.info/coronavirus/about/#definitions


