Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

AGENT BASED SIMULATABLE CITY DIGITAL TWIN TO EXPLORE DYNAMICS OF COVID-19 PANDEMIC

Souvik Barat
Vinay Kulkarni

Tata Consultancy Services Research
Pune, 411013, INDIA

Aditya Paranjape

Imperial College London
London, SW7 2BX, UNITED KINGDOM

Ritu Parchure
Shrinivas Darak
Vinay Kulkarni

Prayas Health Group
Pune, 411004, INDIA

ABSTRACT

Predicting the evolution of Covid19 pandemic has been a challenge as it is significantly influenced by the characteristics of people, places and localities, dominant virus strains, extent of vaccination, and adherence to pandemic control interventions. Traditional SEIR based analyses help to arrive at a coarse-grained 'lumped up' understanding of pandemic evolution which is found wanting to determine locality-specific measures of controlling the pandemic. We comprehend the problem space from system theory perspective to develop a fine-grained simulatable city digital-twin for "in-silico" experimentations to systematically explore - Which indicators influence infection spread to what extent? Which intervention to introduce, and when, to control the pandemic with some a-priori assurance? How best to return to a new normal without compromising individual health safety? This paper presents a digital twin centric simulation-based approach, illustrates it in a real-world context of an Indian City, and summarizes the learning and insights based on this experience.

1 INTRODUCTION

The Covid19 pandemic has affected more than 500 million people globally over the past two years and the number is still going up in many parts of the world as we write this paper. More than 6 million people died, economy is severely impacted, and social wellbeing is heavily compromised. Over the last two years, we have witnessed multiple waves of varying amplitude as the virus mutated giving rise to new strains. WHO has reported tens of variants as *variants of concern* and *variants of interest* (WHO 2022). Some of them, such as *Alpha* and *Delta*, have stayed on for a prolonged period resulting in a large number of severe infections and deaths. Variants like *Omicron* and its descendent lineage have been quite infectious but significantly milder in comparison to *Alpha* and *Delta* variants. Some variants like *Gamma* and *Eta* have disappeared without much visible impact. For instance, India has witnessed three major waves caused by *Alpha*, *Delta* and *Omicron* variants respectively. The mutation of Covid19 will continue and may turn into more or less infectious and severe with other characteristics, such as higher reinfection propensity, in the future. Going forward, extended lockdown cannot be a preferred option from a socio-economic standpoint. But then the questions are: do we understand the dynamics of the infection spread for (existing/new)

978-1-6654-7661-4/22/$31.00 ©2022 IEEE

variants? How can we effectively deal with future situations? When do we focus on what aspect for effective control of unwanted situations?

Several data points across the world indicate that the impact of a wave has strong correlation with the virus characteristics, such as infectivity, severity and mortality rates along with the reinfection possibilities, of the dominant variant. However, the Key Indicators (KI) of the pandemic, such as infection spread, critical cases and mortality, are not solely dependent on virus characteristics. They also have deep correlations with other influencing factors (IFs) along multiple dimensions, such as citizen characteristics (i.e., age, gender, comorbidity, vaccination status, profession of the individual), contextual characteristics of the locality (i.e., places where people mingle and spend time – households, offices, schools, *etc*.) and people behavior in terms of movements and contacts. Moreover, these IFs interfere among themselves thus leading to non-linearities in pandemic evolution (*e.g.*, during the upward trend, during the plateau or downward trend of different waves). These stages are typically associated with the variables of the locality, which are often hidden and difficult to measure. Seroprevalence level, actual active infection (including undetected cases), distribution of dominant variant(s) in a locality are few more examples of such variables that exhibit uncertainty and volatility. Therefore, understanding the trajectory of the KIs in a locality unique demographic and socio-economic strata is equivalent to analyzing a large and complex system with significant volatility, uncertainty, complexity and ambiguity (VUCA). Understanding the influence of IFs on KIs is an advanced form of sensitivity analysis for a VUCA system. Keeping KIs along the desired trajectory by selecting the right intervention for IFs is essentially a multivariate optimization problem where the system is characterized by significant fuzziness and the objective function involves a tradeoff between economic situation of the locality and health of its inhabitants. Moreover, interventions may come with a large time constant. For instance, vaccinating a large population takes time and the vaccine is effective after a delay from the administration and has a waning effect over a time.

In last two years, many models for pandemic management have emerged – majority have attempted to predict the time, amplitude of the peak, and expected burden on public health infrastructure for the next waves with varying degree of success (Agrawal et al. 2021; Xu and Li 2020; Cacciapaglia et al. 2020; Mohan et al. 2022). In comparison, the literature on sensitivity analysis related issue (Zhang et al. 2021) and optimum control (Hussain et al. 2021) is relatively sparse. Moreover, as most of the reported experiments use coarse-grained statistical model and SEIR model (He et al. 2020), the analysis is not fine-grained enough for desired tradeoffs. While literature finds mention of use of fine-grained models, there is significant scope for improvement as regards quality and level of granularity of analysis. Moreover, their focus is also to predict the timeline and amplitude of the surge as opposed to understand evolution dynamics.

We adopt a method that combines *ex post* digital twin-based simulation experiment and the core concept of operational validity from simulation research (Robinson 1999; Sargent 2010) to understand complex infection spread dynamics of Covid19 and amplitude of IFs on KIs. At the heart of our solution is a purposive fine-grained hi-fidelity simulatable model i.e., digital twin of a city. In a bottom up modelling approach, we capture the demography details (i.e., age, gender, comorbidities, adherence to Covid Appropriate Behavior), business-as-usual behavior (i.e., *who* does *what*, *where*, and *when*), the place characteristics (i.e., *who all* congregate *where* and for *how long*), and the virus characteristics (i.e., infectivity, severity, affinity for a specific comorbidity, vaccine escape) in the form of fine-grained *agents* (Hewitt 2010). The pandemic control interventions are modelled as constraints to be adhered to by these agents. Moreover, we implement the agents as a configurable component so as to easily repurpose our city digital twin to evolving situations (*e.g.*, effective medical intervention) and/or a different context. We rely on probability modelling to capture known uncertainties, *e.g.*, vaccine efficacy, immunity loss over time, and compliance with Covid Appropriate Behaviors (CAB), as parameterized behaviors. These parameterized agents and their behaviors help in theory building through experimenting with various hypothetical IFs. We rely on simulation-led experimentation of sufficiently large set of hypotheses/anti-hypotheses. Simulation output is compared with real world observations, as suggested in operational

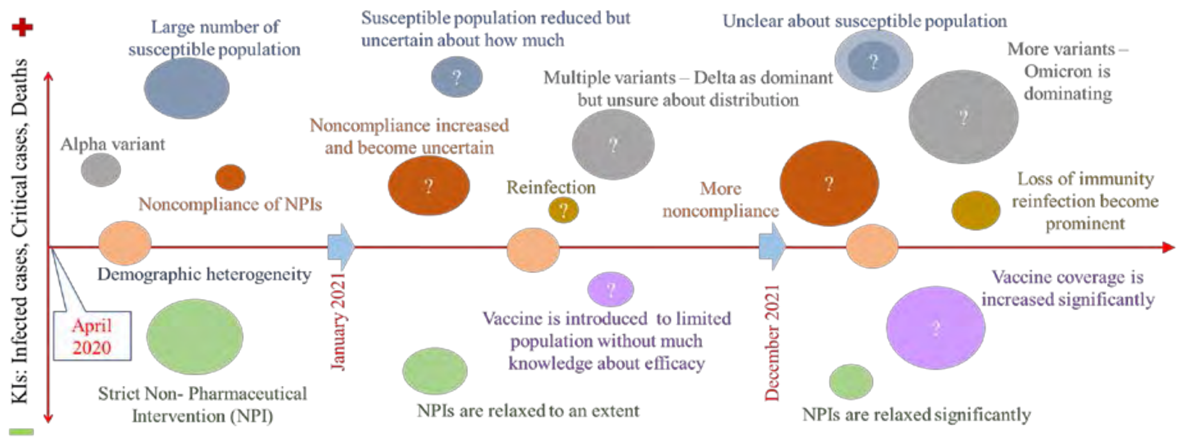

Figure 1: Known evolutions of situation and modelling complexity.

validity, to prove/disprove the hypotheses. Systematic exploration focusing on combination of IFs helps get an understanding of influence of IFs on KIs at various phases of pandemic.

The rest of the paper is structured as follows: Section 2 discusses the modelling and analysis challenges that make the problem complex. Section 3 presents a brief overview of the state-of-the-art techniques to predict evolution of Covid19 pandemic. Section 4 describes our city digital twin and simulation led experimentation approach. Section 5 presents a summary of our experiments and outcome therefrom. Section 6 concludes the paper by highlighting the learnings from 300+ simulation led experimentations.

2 MODELLING AND ANALYSIS CHALLENGES

Modelling the dynamics of the Covid19 pandemic and predicting emerging trajectories of KIs are becoming exceedingly complex as illustrated in Figure 1. The first wave dynamics of pandemic was limited to few factors – it was chiefly dominated by one variant (i.e., Alpha variant) across the world, people's movements were negligible with enforcement of strict lockdown and other Nonpharmaceutical interventions (NPIs), international and interstate movement was restricted, and compliance with Covid Appropriate Behaviors (CAB) was significantly high. The difference in KIs trends across the world was primarily due to the heterogeneity of demographic characteristics (*e.g.*, age, gender and comorbidity) & easy access to healthcare system, and minor variation in adopted NPIs. Several IFs have emerged and uncertainties around the IFs have also started becoming prominent after the first wave. People started violating administrative interventions and social norms, such as the use of face masks and social distancing. Testing uptake varied significantly with time and places across cities. Strict home quarantine and institutional quarantine norms faded away – city administrators had to close several institutional quarantine facilities due to low utilization. Waning of immunity/reinfection became a possibility and more significantly a new variant (i.e., *Delta*) with different characteristics emerged. Seroprevalence level and distribution of *Alpha* and *Delta* in a particular place became a topic for debate along with speculating the characteristics of *Delta* variant.

Subsequently adoption of vaccines with doses and efficacies, emergence of multiple new variants and their lineages with varying characteristics, different form of nonpharmaceutical interventions (*e.g.*, weekend lockdown, night curfew, restricted air travel, *etc*.), wide-ranging noncompliance of CAB started playing critical role in pandemic dynamics. More interestingly, these factors influenced each other in a non-linear way over time horizon. For example, an effective vaccine and a less severe variant (*e.g.*, as for the case of Omicron variant) led to less critical cases – an experience of such trend makes people reluctant to comply with administrative and social norms. People also started taking quarantine norms and testing less seriously over time. Therefore, available data about infections started becoming less relevant, uncertainty

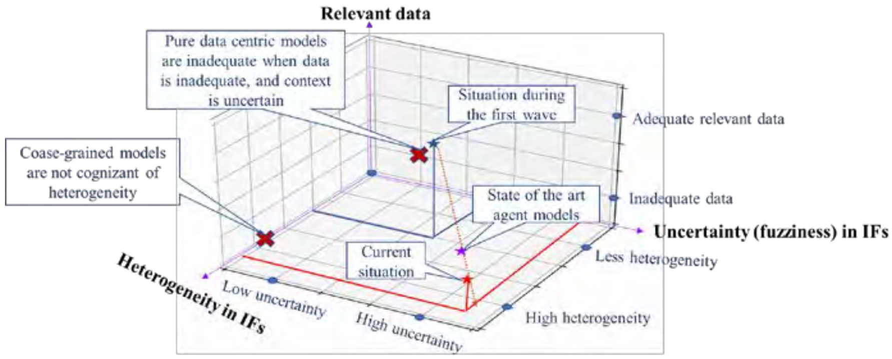

Figure 2: Complexities and state of the art modelling and analysis techniques.

increased significantly in all aspects over the time, and heterogeneity which was limited to demographic aspect during the first wave, became prominent for interventions, compliance of CAB, testing and so on as shown in Figure 1. All these factors have collectively contributed to increase the degree of heterogeneity and uncertainty manyfold and existing data inadequate as shown in Figure 2.

3 STATE OF THE PRACTICE AND OUR PRIOR WORK

A vast majority of predictive models for pandemics are coarse-grained. They adopt one of the two techniques to predict the future: (a) statistical modelling supported by historical data (Agrawal et al. 2021; Mohan et al. 2022; Zhu and Chen 2021), including those based on AI (Fayyoumi et al. 2020); or (b) compartmental models (*e.g.*, SEIR model (He et al. 2020)) that capture epidemiological progression in the form of differential equations. While coarse-grained models are usually computationally efficient and explainable (being based on mathematical techniques), they have several shortcomings that are particularly relevant in the context of predicting the evolution of a pandemic. These coarse-grained models ignore the heterogeneity of the population in terms of age, comorbidity and socio-economic factors that manifest in wide variance of individual behaviors of the population (Barat et al. 2021; Kerr et al. 2021), as they focus on aggregated movement of population from one cohort to other. Moreover, they fail to comprehend micro-causality and emergent behavior in a cohort, *e.g.*, super-spreader events from social gatherings.

In addition to these generic limitations, purely historical data centric coarse-grained models are vulnerable to both internal and external *threats to validity* (Winter 2000). External validity becomes prominent during the early phase of a new variant as one needs to rely on data collected from, for instance, an altogether different geographical region. For example, the Omicron related data collected from South Africa was used for predicting possible infection trends of other counties that differ in terms of vaccines administered, coverage of population vaccinated, demographic details of population and so on. Internal validity is a concern for infection prediction as the observed cases in a given area are not an accurate representation of the reality as observations depend not only on actual infection but also on the ratio of asymptomatic cases and the scale of random testing. For example, analysis of infection spread of Omicron based on the observed data might lead to inaccurate interpretation as asymptomatic cases are considerably high for Omicron, testing uptake is considerably low, and case reporting is a universal concern due to lower severity and wide (and largely undocumented) use of home-testing facilities. Overall, coarse-grained SEIR models and data centric AI/statistical models are becoming inadequate with increasing demand for heterogeneity and uncertainty in presence of inadequate relevant data as shown in Figure 1.

To overcome the limitations of coarse-grained models, fine-grained agent-based models have been employed as a competing approach. The key objective of these models is to capture the behaviour of micro-elements such as people, households and places (*e.g.*, office, school and shops) to predict KIs. However, these fine-grained models need to make a trade-off between richness and scale. Richness includes the ability to represent the heterogeneity of the people, households and places at a fine-grained level to take the model closer to reality, i.e., city, state or country. Many of the agent-based models (Cuevas 2020; Silva et al. 2020)

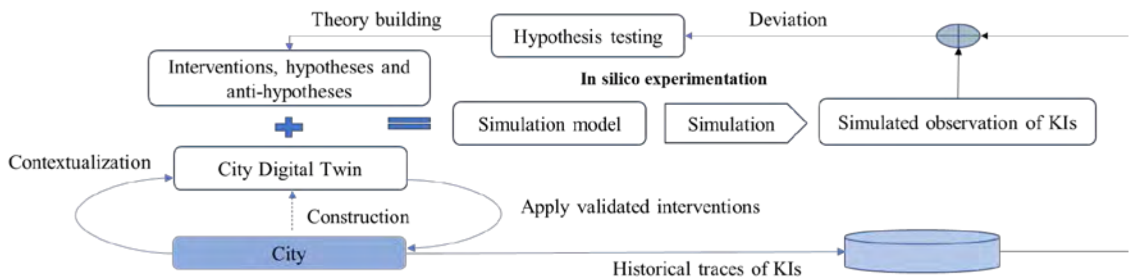

Figure 3: Simulation led experimentation environment.

consider high-level classifications of these entities as cohorts, where each cohort is internally represented using aggregated equations, and represent the whole system as a connected network of a limited number of cohorts. These models address scalability by aggregation and can estimate the impacts of fine-grained interventions, such as the impact of infection spread when all shops and/or offices are closed. However, they exhibit similar limitations as coarse-grained models for comprehending emergent behavior and micro-causalities of pandemic dynamics. Covasim (Kerr et al. 2021), on the other hand, uses an agent-based model to capture the individualistic behaviors of a wide range of micro-elements and their interactions. They linearly scale down the population (in the order of 10^3) to make the simulation manageable. From a richness perspective, they capture demographic variations in the population, a wide range of places and interventions. However, they encode the combined effect of a specific variant and vaccine on the individual as predefined equations within person agents. This limits the ability to understand the interplay of the effect of a vaccine and the characteristics of variant on the demographic factor of an individual.

Our earlier digital twin-based experimentation to understand the impact of nonpharmaceutical intervention (Barat et al. 2021) is similar to Covasim approach. Our trend analyses using earlier digital twin, starting from July 2020, closely resemble how the first wave unfolded in terms of KIs and the timeline of the peak in Pune. In our study published in January 2021, we also predicted that second wave may hit Pune city between March to April 2021 (*ref.* Figure 19 of Barat et al. 2021). While our predictions about second wave and its timeline matched closely with the reality, we were wrong in predicting the magnitude of the second wave peak – primarily our prediction was much milder than reality (around 70% of first wave peak versus around 130% of the first wave peak in reality).

Critical analyses of our earlier work along with other state-of-the-art agent based models for Covid19 pandemic indicate that they are vulnerable along two dimensions: a) they are not capable of addressing the increasing heterogeneity and uncertainties along various variants and their lineages, overlapping impacts of variants and vaccines over individuals, and wide range of noncompliance possibilities, and b) they are not designed to comprehend the implication and amplitude of various IFs on KI with growing complexities – in contrast they are made to predict future situation under known set of variants, vaccines and populations.

4 APPROACH

We visualize a city as a complex dynamic system and adopt a concept of digital twin to mimic the key elements of the system as an “in silico” experimentation environment as shown in Figure 3. In reality, the macro-behavior of the system (i.e., evolution of a pandemic in a city) emerges from stochastic and spatiotemporal micro behaviors of the constituent elements, such as people, virus and vaccines. We faithfully capture relevant constituent elements and their individualistic behaviors in a bottom-up manner using parameterized *agent* where the fully known behavior of agents is encoded as rules, and the uncertain behavior (i.e., *known unknown*) is specified as a probability distribution over a set of actions where the probability distribution is a configurable parameter. These agents are realized using the Enterprise Simulation Language (ESL) (Clark et al. 2017)

We populate the city digital twin with relevant information available with the municipal authority. We validate the city digital twin by simulating past situations and comparing the simulated results (i.e., KIs)

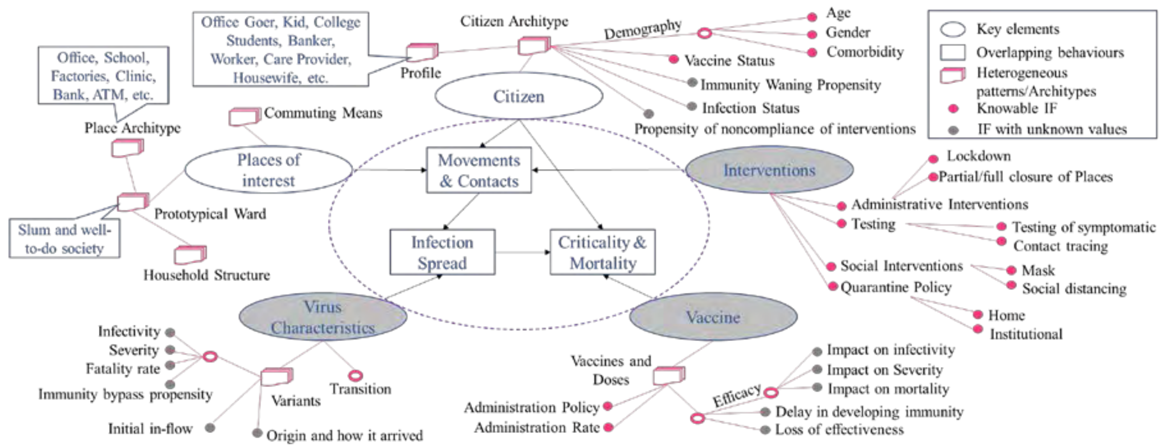

Figure 4: Schematic representation of City Digital Twin.

with real data. We simulate a candidate configuration (i.e., same set of parameter values) multiple times (typically 5 or more times) till the stochastic behaviors converge thus improving the confidence level of observed simulation results as suggested by (Robinson 1999). A contextualized and validated digital twin is then used for *ex post facto* analysis and future predictions. We set up the initial state of digital twin to validate a specific hypothesis (or anti-hypothesis) by assigning suitable values to the relevant configurable parameters. We then simulate for a specific what-if scenario multiple times till the stochastic behaviors converge. Human experts interpret the simulation results to prove (or disprove) the hypothesis. Examples of typical IFs are: relaxation of NPI, noncompliance of NPI, and introduction of a new variant with specific characteristics. Simulating the digital twin with this configuration produces results that can be interpreted to compute the desired KIs.

4.1 City Digital Twin

We extended our earlier agent based digital twin (Barat et al. 2021) along multiple dimensions to support evolving situations and complexities as highlighted in Figure 1. Extended agent topology of our new digital twin captures inherent heterogeneity, stochasticity and interactions of five key elements, namely citizen, places of interest, virus characteristics, vaccine and interventions as shown in Figure 4.

Citizen agent: It captures individualistic characteristics and behavior patterns including age, gender, comorbidity, profession, household structure, vaccination status, and infection status. Citizens agent is specialized into 25 citizen architypes, including Kid, College Student, Senior Citizen, and Office Goers, to represent professions with different behavioral patterns. Each architype has unique movements and contact patterns. For example, an office goer in back-end role interacts only with colleagues, a bank staff interacts with colleagues as well as with bank customers, a cab driver interacts with a large number of passengers each belonging to a different architype throughout the day; a housemaid interacts with a fixed set of house members for a fixed amount of time but visits multiple households in a day, a small shop keeper interacts with customers for short intervals but in a congested place. We also model behavioral patterns, such as propensity of compliance with CAB and NPIs.

Places of interest: We visualize a city as set of prototypical wards, where each word is a combination of well-to-do and slum localities with representative set of households (with different structure and area), citizens (with different age, gender, comorbidities and architype), places and commuting means. To precisely represent slums and well-to-do localities, 13 types of households and 20 place architypes including various commuting means (i.e., own car, bus or shared cab) are modelled. The household structures range from two-member family (1 Male and 1 Female) to twelve-member family (i.e., 3 Male, 3 Female, 4 Children and 2 Senior Citizens). Place architypes represent relevant places of a city where people frequently visit, spend time and make contacts, such as office, school, restaurants & pub, clinics, mall, market place, worship place and so on. For an illustration, a locality can be formed using ten offices, three schools, hundreds of local shops, tens of barber shops, hundreds of clinics and thousands of households with varying number of family members. Citizens from well-to-do localities may stay in relatively bigger houses with few family members as compared to slum area. Predominantly, the citizens from well-to-do area are office goers, bank employees, heath worker and from other white-collar professions. On the other hand, slum areas are densely populated and have smaller houses with bigger families.

Virus characteristics: We represent variants as configurable passive agents, which have their own (stochastic) characteristics but cannot act independently without a citizen agent. Essentially, they contribute to the citizen state/behavior when a citizen is exposed to a variant. Each variant agent defines (actual or hypothetical) virus characteristics including infectivity, severity, mortality, and probability of immune escape as parameters. One can introduce a new variant to a population of a city by specifying a start date and possible rate of introduction through in and out movement from and to other cities understand the impact of a new variant.

Interventions: We specify four types of interventions, namely administrative intervention, health care related intervention, social intervention, and vaccination as spatiotemporal stochastic agents. Administrative interventions are related to citizen movements, (partial) closure of places, allowed passengers in Cabs and Buses. Interventions from health care standpoint include testing of mildly infected citizens (in addition to severely infected citizen), contact tracing and isolation of detected mildly infected citizens. Social interventions include mask usage and social distancing. These intervention agents can be active or inactive for specific locality/citizen architype in a time bound manner – they influence the behavior of the place and citizen one when they are active. These spatiotemporal characteristics help to represent the emerging complexities along intervention dimension.

Vaccines: Similar to the variant agent, vaccines are visualized as configurable passive agent, which can influence citizens. As characteristics, they capture two aspects a) vaccine adoption, and b) vaccine efficacies in terms of reducing infection, severity and fatality probability. Vaccine adoption in a city is modelled via a parameterized administration rate. A specific vaccine (*e.g.*, Covishield (Rather et al. 2021) can be set to be introduced from a specific day of a month to the population with a set of criteria on age (*e.g.*, 60+ and or 45+), comorbidity (*e.g.*, diabetes and hypertension) and profession (*e.g.*, medical professionals). The dose intervals can also be configured based on the prevailing vaccination policies, such as a 90-day hiatus for the second dose of Covishield or 270 days for a third dose.

4.2 Simulation

We observe possible situations by simulating a contextualized digital twin with city-specific configuration and known facts about vaccines, variants, interventions and their compliances as shown in Figure 3. In a simulation, situations emerge through agent interactions and overlapping aspects as illustrated in Figure 4. For example, an individual moves within a place (*e.g.*, within office, school, and mall) and between places (*e.g.*, home to car, car to office, office to shop, and shop to home) for business-as-usual activities. The movements between the places are largely derived from the profession with certain random movements for living, socializing and entertainment. Movements within a place are inherently random, however, we adopt a rationalistic view to define movement patterns and movement frequency of an individual based on <citizen archetype, place archetype>. For example, a doctor stays in a specific location during the clinic hours whereas the patients must wait in a waiting area for a specific time before consultation in doctor's room. In addition to the intra city movements, we also consider movements into and out of city. A person with specific demographic characteristics, comorbidity, vaccine doses and infection history can move around across different places within and across cities. As a result, she may get exposed to a specific variant with a certain probability. The susceptible-to-exposed transition of a person (target) depends on the duration and frequency of proximal contact with an infectious person (source), infection history of the target person and variant characteristics. Essentially, susceptibility-to-exposed transmission dynamics is an interplay between two person agents (duration and proximity of source and target agents), place agent (characteristics of the place, i.e., open vs close) and infectivity of variant agent. Further, the progression of the infection in

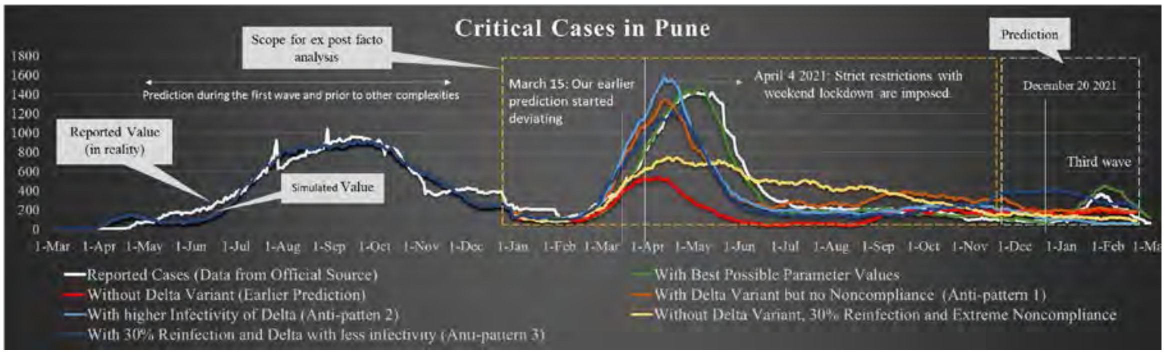

Figure 5: Summary of simulation experimentations.

an exposed person (i.e., exposed to infectious, infectious to asymptomatic, mildly symptomatic or severe, respective state to recovered or dead) and the possible degree of criticality of that person depend on the characteristics of the person, infection and vaccine histories (i.e., passive vaccine agent), as well as the infecting variant (i.e., passive variant agent). Thus, the progression dynamics is the interplay among person agent, vaccine agent and a variant agent and the dynamics depend on – (a) age and comorbidity of an individual; (b) vaccine effectiveness for that person; (c) immunity developed due to earlier infection, and (d) characteristics of the variant. Other interventions and their compliance also influence the overall dynamics. Movement-related restrictions of a person limit mixing. Testing helps in isolating infected people from the susceptible population. However, non-compliance of quarantine norm increases mixing possibilities within household members and close contacts. Multiple simulations with varying parameters help to comprehend the complexities.

5 SIMULATION-BASED EXPERIMENTATION

We contextualized the city digital twin for Pune (an Indian city with 4 million population) by configuring the demographic details and comorbidity distribution among Pune's citizens, household structures and prototypical areas that reflect socio-economic characteristics of the city, various professions and their movements, and prototypical places, such as offices, factories, schools, markets, worship places with official data collected from Pune city administrator. Our trend analyses, starting from March 2020, using extended digital twin (as discussed in section 4) closely resemble how the first wave unfolded in terms of KIs and the timeline as illustrated using white line and dark line in Figure 5. Moreover, simulation of a configuration represented as "Without Delta Variant" matches with the prediction from our earlier digital twin published in January 2021 (Barat et al. 2021) and observed trends. We considered this as an operational validity of our new extended digital twin.

We predicted a surge/wave in Pune during the month of March and April due to relaxation of NPIs, reduced testing uptake and noncompliance of CAB. However, we estimated the peak much milder than the first wave (around 70%) as shown using red line in Figure 5. Precisely our predictions started deviating from mid-March 2021 (as shown using red and white lines). To understand possible cause for unanticipated surge, we started exploring wide range of scenarios considering popular speculations at that point of time (important to note here that *Alpha* variant was the only variant of concern and dominance of *Delta* variant was not an accepted fact at that time). We simulated scenarios with - a) greater noncompliance of CAB and movement related restrictions, and b) reinfection & loss of immunity. The combination of all speculated factors with extreme possibilities could not explain the observed situation as shown using configuration "Without Delta Variant, 30% Reinfection and Extreme Noncompliance" (yellow line) City administrators imposed a strict lockdown in Pune from April 4, 2021 but surge of infections and critical cases continued to grow for a month as shown in Figure 5. The existence of Delta variant in Pune got detected in early April 2021 – we introduced it in the city digital twin and simulated results are as in "With Delta variant but no noncompliance" (i.e., orange line). The simulation matched the trend but details were still quite removed

Table 1: Parameters, their ranges (uncertainties) and derived parameter values.

Dimension	Parameter	Range of values considered	Derived values from ex-post facto analysis
1. Administrative interventions	Noncompliance of lockdown and relaxation as imposed in Pune	1. 75% to 100% compliance 2. 0 to 2 weeks delay	7-10 days delay and 10-15% noncompliance for all restrictions since March 2021. All relaxations are also delayed by 6-7 days.
	Noncompliance of Social gathering norms	10-25 % noncompliance	15-20% violation of restriction on social gathering is observed throughout all lockdowns during the second wave
2. Social norms	Mask usage	10-80%	Effective mask usage is as low as 15-20%
3. Quarantine policy and compliance	Institutional quarantine	5-100%	Institutional quarantine facility might not have availed by more than 10% population (most likely below 5%)
	Compliance of home quarantine	10-100%	50-60% home quarantined patients might have violated strict quarantine norms.
4. Vaccine	Administration policy	As per the policy	We considered a fixed rate starting from March 2021 with government policy.
	Overall efficacy	60-95%	Most possibly 30-40% after first dose and 80-85% after second dose.
	Delay in developing immunity	7- 30 days	Not very clear from the experiment but 14-21 days delay is most likely
	Lowering infectivity	20-80%	Explorations suggest 30-40% reduction in infectivity (possibly due to less viral load or less severity for vaccinated population), 80-85% reduction in severity and more that above 98-99% redaction in fatality.
	Lowering severity	50-95%	
	Lowering fatality	90-100%	
5. Variant of concern – Delta	Date of origin/import	Anytime between December 2020 to February 2021	The existence of Delta variant in Pune before February 2021 is questionable- most possibly it is imported in February
	Rate of initial inflow	0.001 to 0.01%	Significant inflow of infection with delta variant (around 100-200 cases/per day) might have happened in February and March 2021.
	Infectivity	1X – 2X	Infectivity of Delta variant is around 1.5-1.6X, severity is 1.3-1.4 X and fatality rate is 1.2-1.3 X of Alpha variant. Critical patients might have taken 3-6 days from hospitalization to become critical. No significant immunity bypass is observed.
	Severity	1X – 2X	
	Mortality Rate	1X-2X	
	Bypass immunity	0%	
6. Loss of immunity	Immunity due to infection	3-25 % after 115 – 240 days	Loss of immunity must be below 5% and not before 4 months of recovery.
	Immunity due to vaccine	0-50% after 120-150 days	No loss before 4 months.

from the reality thus indicating that a combination of several factors (IFs) is probably the cause. We set up an ex post facto analysis by considering reported data from January 2021 to November 2021 to understand the cause with an aim comprehend the dynamics of the pandemic in presence of multiple variants, vaccines, noncompliance related uncertainties and their overlapping impacts over demographically heterogeneous population. An overview of our ex post factor analysis is presented in the next subsection.

5.1 Post Ex Facto Analyses

We systematically explored a wide range of possibilities (i.e., configurations/hypothesis) along six dimensions, as shown in Table 1. Configurations are constructed, a theory building activity, by meaningfully selecting parameter values from Table 1 and evaluated through an iterative simulation of the digital twin. To make the iterations manageable without compromising on precision of the analysis, we divided the time from January 2021 to November 2021 into three phases namely, pre vaccine phase, low adoption of vaccine, and high vaccine adoption so as to exclude vaccine related configurations during the early phase of vaccination drive in Pune. For example, the effect of vaccine was negligible in Mar-Apr 2021 as less than 5% of people were vaccinated by then. Therefore, exploring vaccine efficacy during that time might not be a pragmatic consideration.

Multiple iterations focusing on a relevant set of parameters from Table 1 and the correlation of simulated values with actual reported values of the key indicators (KIs) helped to derive possible parameter values (i.e., hypothesis about a possible value of a parameter). Values derived from one simulation are further analysed in the context of subsequent simulations to prove or disprove the hypothesis under consideration. In addition, we also defined and simulated several anti-hypotheses to establish influences and to understand overall dynamics. Our hypotheses and anti-hypotheses are proved and disproved respectively through 345 scenario evaluations. To eliminate the threats to validity of simulation results and eliminate the possibility of extreme emergent situations of stochastic behaviours in a simulation run, we repeated each scenario 5 times. Normalized values of critical cases for best fitted parameter values and simulation results of selected anti-patterns are shown in Figure 5.

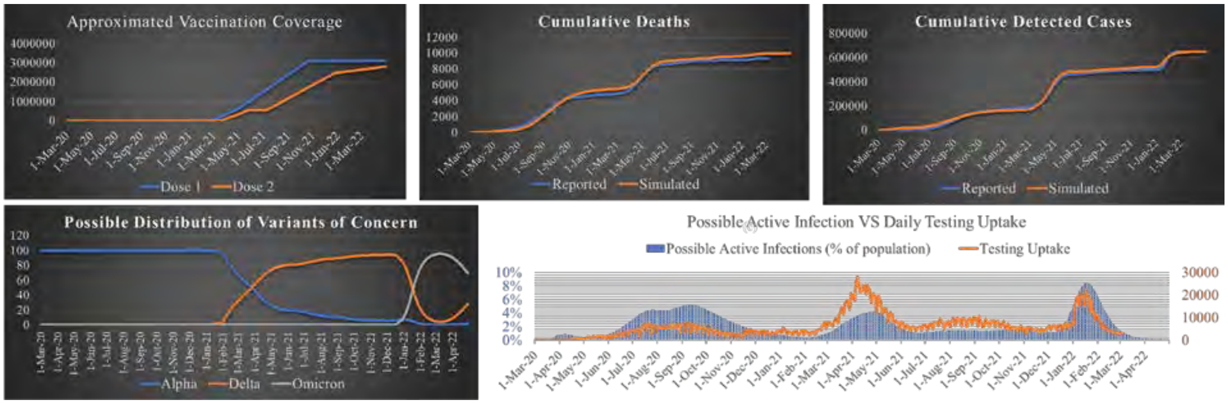

Figure 6: Predicted KIs and hidden factors for best possible parameter values.

5.2 Synthesis of Our Experimentation

Synthesis of 1725 simulation runs with varying parameter values from Table 1 guided by established facts (such as vaccines) and critical comparison with the real data helped to derive likely values of the parameters along six dimensions as summarized in column “Derived Values from Ex Post Facto Analysis” of Table 1. Predicted trends of the key indicators (KIs) for derived parameter values in comparison with actual reported data are shown as “With Best Possible Parameter Values” (i.e., green line) in Figure 5. Derived KIs for best possible parameter values (from ex post facto analysis considering data till November 2021) and projected trajectories (from December 2021) along with reported cases are shown in Figure 6. Predicted critical cases (shown in Figure 5), detected cases and death count (shown in Figure 6) from December 2021 match closely with reported counts (possible distribution of different variants, shown in Figure 6, is also at per with city-based healthcare experts). Synthesis of our experimentation’s outcome with respect to the official Covid19 pandemic data reported by Pune city authorities until now helps to understand the influences and pandemic dynamics as summarized in Figure 7. Our precise analysis insights used effectively to explain situation during the third wave in Pune that started from December 20, 2021 and continued till January 2022. As shown in Figure 5 and 6, there are several perplexing situations that emerged in Pune during the third wave (*e.g.*, significantly low critical cases in Pune as compared to general ratio that we observed in other Indian cities, sudden disappearance of cases, and low testing uptake). However, we were able to predict and justify them using derived insights from our experimentation. Our analyses and insights are used by Pune city administration as one of the input for arriving at pandemic control strategies.

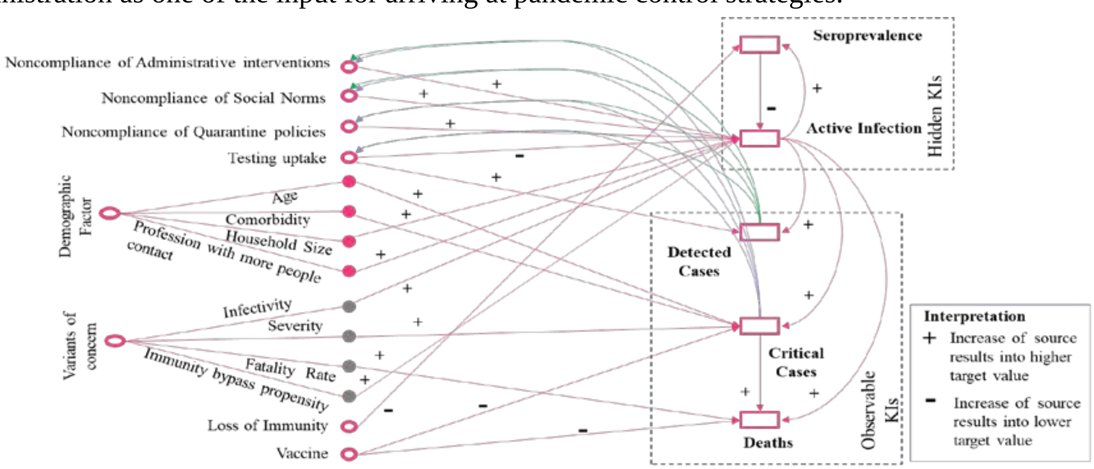

Figure 7: Qualitative analysis of possible influence.

6 CONCLUDING REMARKS

We presented a systematic simulation-based experimentation study to understand factors that influence the key indicators of Covid19 pandemic as it evolved in a city. Our analysis helps to justify the possible causes that could have led to perplexing situations observed during second and third waves. Analysis also identifies parameters that should be critically evaluated to predict future waves and their impacts. Our experiments indicate that characteristics of dominant variants is the principal contributing factor for infection spread. Variants with higher infectivity can potentially lead to a wave - as seen in last three major waves. People's movements along with the characteristics of variant play a major role to intensify/slowdown the pace at which a wave will unfold in a city during its initial phase. Interventions like lockdown that essentially restrict movements in open spaces can at best delay the onset of a wave and that too only to an extent. Instead, appropriate control measures that reduce mixing of people in closed places and limit household infections lead to better results. Therefore, isolation in the form of strict home / institutional quarantine backed by effective testing contribute the most toward reduced infection spread. Wearing face mask while in closed places helps to curb the infection spread to a large extent. Seroprevalence level, a hidden indicator, plays a role in deciding the peak value of active infection but only when the dominant variant is not capable of bypassing immunity - as we have seen for Omicron variant during the third wave. Infectivity and severity of a variant have a complex relationship with other factors that influence how quickly a variant can be the dominant variant and make impact to healthcare system in a city. A variant with low infectivity and low severity exponentially disappears from the community over time as we observed for *Alpha* vs *Delta*. Variant with low infectivity and high severity disappears much faster compared to a variant with low infectivity and low severity as severely infected person is likely to go for testing and subsequent isolation. Variant having high infectivity and low severity (*e.g.*, *Omicron* variant) quickly becomes the dominant. Variant with high infectivity and high severity can potentially be the most dangerous, however, high severity leads to early detection thus limiting the impact significantly - such a situation can be controlled through early testing, contract tracing and strict isolation. Therefore, a variant with high infectivity and high severity is unlikely to survive for long. Variant's ability to bypass immunity is another characteristic that needs to be considered carefully as it significantly contributes towards the magnitude of the peak of a wave.

Vaccine is the most critical factor towards controlling the severity of infection spread and subsequent fatality modulo the extent and quality of healthcare available. We have seen its impact during the later stage of the second wave and throughout the third wave. However, reduced severity typically leads to reduced testing and lax compliance to isolation norms - as witnessed during the third wave. These two factors lead to greater infection spread thus putting to risk those with comorbidities and/or without vaccine protection.

Rapid fading away of third wave is leading to quick return to normalcy. With sizeable population yet to receive full vaccination cover, there is a speculation about next wave. Based on simulation results, we think that a major surge in infections is possible only if a) a new variant characterized by moderate to high infectivity, moderate severity, and immunity bypassing capability emerges, and/or b) significant chunk of population becomes susceptible due to immunity waning over time. The latter factor can cause a wave even in absence of the former factor, as most of the variants are likely to be around for a while and can cause a serge if seroprevalence level drops significantly. As predicting the possibility of a new variant is out of the scope of our study, this paper urges administrators and policy-makers to be vigilant about new mutation with high severity and possibility of waning immunity while strategizing a safer return to new normal.

REFERENCES

Agrawal, M., M. Kanitkar, and M. Vidyasagar. 2021. "SUTRA: A Novel Approach to Modelling Pandemics with Applications to COVID-19". arXiv preprint arXiv:2101.09158.

Barat, S., R. Parchure, S. Darak, V. Kulkarni, A. Paranjape, M. Gajrani, and A. Yadav. 2021. "An Agent-Based Digital Twin for Exploring Localized Non-pharmaceutical Interventions to Control COVID-19 Pandemic". *Transactions of the Indian National Academy of Engineering* 6(2):323-353.

Cacciapaglia, G., C. Cot, and F. Sannino.. 2020. "Second Wave COVID-19 Pandemics in Europe: a Temporal Playbook". *Scientific Reports*, 10(1):1-8.

Clark, T., V. Kulkarni, S. Barat, and B. Barn. 2017. "ESL: An Actor-based Platform for Developing Emergent Behaviour Organisation Simulations". In *International Conference on Practical Applications of Agents and Multi-agent Systems*, June 21st -23rd, Porto, Portugal, 311-315.

Cuevas, E. 2020. "An Agent-based Model to Evaluate the COVID-19 Transmission Risks in Facilities". *Computers in Biology and Medicine*, 121, 103827.

Fayyoumi, E., S. Idwan,, and H. AboShindi. 2020. "Machine Learning and Statistical Modelling for Prediction of Novel COVID-19 Patients Case Study: Jordan". *International Journal of Advanced Computer Science and Applications* 11(5): 3-11.

He, S., Y. Peng, and K. Sun. 2020. "SEIR Modeling of the COVID-19 and Its Dynamics". *Nonlinear Dynamics* 101(3):1667-1680.

Hewitt, C. 2010. "Actor Model of Computation: Scalable Robust Information Systems". arXiv preprint arXiv:1008.1459.

Hussain, T., M. Ozair, F. Ali, S. Rehman, T. A. Assiri, and E. E. Mahmoud. 2021. "Sensitivity Analysis and Optimal Control of COVID-19 Dynamics Based on SEIQR Model". *Results in Physics* 22:103956.

Kerr, C.C., R.M. Stuart, D. Mistry, R.G. Abeysuriya, K. Rosenfeld, G.R. Hart, R.C. Núñez, J.A. Cohen, P. Selvaraj, B. Hagedorn, and L. George. 2021. "Covasim: An Agent-based Model of COVID-19 Dynamics and Interventions". *PLOS Computational Biology* 17(7): p.e1009149.

Mohan, S., A. K. Solanki, H. K. Taluja, and A. Singh. 2022. "Predicting the Impact of the Third Wave of COVID-19 in India Using Hybrid Statistical Machine Learning Models: A Time Series Forecasting and Sentiment Analysis Approach". *Computers in Biology and Medicine* 144:105354.

Rather, R. A., T. Islam, I. U. Rehman, and D. Pandey. 2021. "Development of Vaccine Against Coronavirus Disease 2019 (Covid-19) In India". *Asian Journal of Advances in Medical Science*:13-21

Robinson, S. 1999. "Simulation Verification, Validation and Confidence: A Tutorial". *Transactions of the Society for Computer Simulation* 16(2): 63-69.

Sargent, R. G. 2010. "Verification and Validation of Simulation Models". In *Proceedings of the 2010 Winter Simulation Conference*, edited by S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, 166-183. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Silva, P. C., P. V. Batista, H. S. Lima, M. A. Alves, F. G. Guimarães, and R. C. Silva. 2020. "COVID-ABS: An Agent-based Model of COVID-19 Epidemic to Simulate Health and Economic Effects of Social Distancing Interventions". *Chaos, Solitons & Fractals* 139:110088.

Winter, G. 2000. "A Comparative Discussion of the Notion of Validity in Qualitative and Quantitative Research". *The Qualitative Report* 4(3): 1-14.

World Health Organization (WHO). 2022. Tracking SARS-CoV-2 Variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants, accessed 20th April 2022.

Xu, S., and Y. Li. 2020. "Beware of the Second Wave of COVID-19". *The Lancet* 395(10233):1321-1322.

Zhang, Z., R. Gul, and A. Zeb. 2021. "Global Sensitivity Analysis of COVID-19 Mathematical Model". *Alexandria Engineering Journal* 60(1): 565-572.

Zhu, Y., and Y. Q. Chen. 2021. "On a Statistical Transmission Model in Analysis of the Early Phase of COVID-19 Outbreak". *Statistics in Biosciences* 13(1): 1-17.

AUTHOR BIOGRAPHIES

SOUVIK BARAT is a Principal Scientist at Tata Consultancy Services Research. He holds a PhD in Computer Science and his research interests include digital twin technology and modelling & simulation of complex systems, agent and actor technology, enterprise modeling, and business process modelling. His email address is souvik.barat@tcs.com.

VINAY KULKARNI is a Distinguished Chief Scientist at Tata Consultancy Services Research. An alumnus of Indian Institute of Technology Madras, Vinay is a Fellow of Indian National Academy of Engineering. His research interests include Digital Twins, Adaptive Enterprises, Model Driven Engineering, AI, and Software Engineering. His e-mail address is vinay.vkulkarni@tcs.com.

ADITYA A. PARANJAPE is a Lecturer in Aerial Robotics at Imperial College London. His research interests include control theory, nonlinear dynamics and multi-agent systems. His e-mail address is a.paranjape@imperial.ac.uk.

RITU PARCHURE, MBBS and Master of Public Health, is an epidemiologist and senior research fellow at Prayas Health Group. Her email address is ritu@prayaspune.org.

SHRINIVAS DARAK is a Public Health researcher at at Prayas Health Group. He holds a PhD in demography and population studies from University of Groningen, Netherlands. His email address is shirish@prayaspune.org.

VINAY KULKARNI, MBBS MD, FAMS(VIENNA), DNB(SKIN&VD), AAHIVS, is a practicing physician at Prayas' Amrita Clinic and Deenanath Mangeshkar Hospital, Pune. He is also a founder Trustee of Prayas Health Group. His e-mail address is vinay@prayaspune.org.