

# Ensemble forecast of COVID-19 in Karnataka for vulnerability assessment and policy interventions

Sashikumaar Ganesan<sup>1</sup>, Deepak Subramani<sup>1</sup>, Thivin Anandh<sup>1</sup>, Divij Ghose<sup>1</sup>, Giridhara R Babu<sup>2</sup>

**1** Department of Computational and Data Sciences, Indian Institute of Science, Bangalore, India

**2** Life-course Epidemiology, Indian Institute of Public Health-Bangalore, Public Health Foundation of India, Bangalore, India

sashi@iisc.ac.in, deepakns@iisc.ac.in, giridhar@iiph.org

## Abstract

We present an ensemble forecast for Wave-3 of COVID-19 in the state of Karnataka, India, using the IISc Population Balance Model for infectious disease spread. The reported data of confirmed, recovered, and deceased cases in Karnataka from 1 July 2020 to 4 July 2021 is utilized to tune the model's parameters, and an ensemble forecast is done from 5 July 2021 to 30 June 2022. The ensemble is built with 972 members by varying seven critical parameters that quantify the uncertainty in the spread dynamics (antibody waning, viral mutation) and interventions (pharmaceutical, non-pharmaceutical). The probability of Wave-3, the peak date distribution, and the peak caseload distribution are estimated from the ensemble forecast. Our analysis shows that the most significant causal factors are compliance to Covid-appropriate behavior, daily vaccination rate, and the immune escape new variant emergence-time. These causal factors determine when and how severe the Wave-3 of COVID-19 would be in Karnataka. We observe that when compliance to Covid-Appropriate Behavior is good (i.e., lockdown-like compliance), the emergence of new immune-escape variants beyond Sep '21 is unlikely to induce a new wave. A new wave is inevitable when compliance to Covid-Appropriate Behavior is only partial. Increasing the daily vaccination rates reduces the peak active caseload at Wave-3. Consequently, the hospitalization, ICU, and Oxygen requirements also decrease. Compared to Wave-2, the ensemble forecast indicates that the number of daily confirmed cases of children (0-17 years) at Wave-3's peak could be seven times more on average. Our results provide insights to plan science-informed policy interventions and public health response.

## 1 Introduction

The COVID-19 pandemic necessitates forecasts to frame science-informed policies. An accurate forecast of the size and timing of future waves could help public health officials and governments to plan appropriate responses. An ensemble forecast by aggregating different scenarios and models makes the prediction robust and reliable.

COVID cases were reported in India from February 2020, and the first nationwide lockdown was imposed on 25 March 2020. A graded reopening started in June 2020. After that, in the state of Karnataka, the Wave-1 began gradually building up and reached a maximum daily confirmed caseload of 10K (7-day average) around 11 October 2020. Subsequently, the daily confirmed caseload reduced to less than 1K by December

2020 and stayed below 1K until mid-March 2021. From the 3rd week of March 2021, the caseload started increasing again, and a Wave-2 started raging, reaching a peak of 47.5K daily confirmed cases on 9 May 2021 [1]. Additionally, since 16 January 2021, vaccination against Covid-19 started in India with a rationing policy based on the recipient's age [2]. Genome sequencing studies have shown that Wave-1 was caused by the B.1.1.7 ('Alpha') variant and Wave-2 was caused by the B.1.617.2 ('Delta') variant [3–7]. Despite timely NPIs (decentralized state-wise, city-wise, and community-wise lockdown), there was a widespread concern due to the high load on the medical infrastructure with reports of ICU and oxygen shortages.

Unfortunately, almost no computational model for Covid-19 spread in India predicted a Wave-2, drawing widespread flak and criticism from the Government and general public [8]. The forecast failure could have been due to the uncertainties in the spread dynamics, intervention effectiveness, virus mutation, anti-body waning, and vaccination effectiveness. As such, a comprehensive model incorporating the above uncertainties is essential to predict new waves.

Furthermore, the COVID guidelines and policies on non-pharmaceutical interventions (NPIs) and pharmaceutical interventions, including vaccination and planning, differed for Wave-1 and Wave-2 in India. Specifically, a uniform national policy was implemented during the first wave. In contrast, each state government imposed individual localized policies during the second wave. Hence, the ensemble forecast must be done at a state level to take appropriate decentralized action for containment. Here, we present an ensemble forecast of COVID-19 in Karnataka for 2021-22, focusing on predicting a new Wave-3 and quantifying its impact on vulnerable populations.

An ensemble forecast is built by aggregating predictions from different scenarios and models. Such an approach has been shown to consistently perform well in different fields, see for e.g., [9–12]. Uncertainties in human behavior, mobility, local government policies, mutation, immune response, and anti-body waning make COVID spread forecasting a challenging problem [13, 14]. An ensemble method incorporates these uncertainties as scenarios and makes the forecast more robust. Appropriate trade-offs can be made by policymakers by analyzing the probability of occurrence of each scenario being modeled.

Several authors have analysed the data of COVID Wave-2 in India and ascribed causal factors [15]. The importance of mutation in the Wave-2 and [16] the effect of immunity waning have been characterized [17]. The effect of non-pharmaceutical interventions like lockdowns [18, 19] and vaccine allocation strategies [2] has been studied. Predictions for third wave based on immunity waning, emergence of an immune-escape new variant and social distancing [20] have also been made. However, no comprehensive ensemble forecast considering all known causal factors have been reported in literature. In what follows, we first describe in Sect. 2 our IISc-COVID Model, the causal control variables that we have considered and the kind of analysis that we perform. Next, we discuss our results drawing insights from the analysis done on our ensemble forecast (Sect. 3). Finally we summarize our findings and raise a call for action for effective intervention (Sec. 4).

## 2 Materials and methods

### 2.1 IISc-COVID Model

The IISc COVID model proposed in [21] is employed to compute all estimates using our in-house finite element package [22, 23]. The model consists of an unknown scalar function describing the dynamics of the infected population in a six-dimensional space. In particular, the active infected population is distributed in space, infection severity,

duration of the infection, and age of the infected people. Let  $T_\infty$  be a given final time,  $\Omega := \Omega_x \otimes \Omega_\ell$  be the spatial domain, and  $\Omega_\ell := L_v \times L_d \times L_a$  be the internal domain. Here,  $L_v$ ,  $L_d$  and  $L_a$  denote the infection severity, duration of infection, age of the people, respectively. Then, the dynamics of the infected population  $I(t, \mathbf{x}, \ell) \in (0, T_\infty] \times \Omega_x \otimes \Omega_\ell$  is described by the population balance equation

$$\begin{aligned} \frac{\partial I}{\partial t} + \nabla \cdot (\mathbf{u}I) + \nabla_\ell \cdot (\mathbf{G}I) + CI &= F && \text{in } (0, T_\infty] \times \Omega_x \times \Omega_\ell, \\ I(t, \mathbf{x}, \ell) &= g_n && \text{in } (0, T_\infty] \times \partial\Omega_x^- \times \Omega_\ell, \\ I(t, \mathbf{x}, (\ell_v, 0, \ell_a)) &= B_{nuc} && \text{in } (0, T_\infty] \times \Omega_x \times L_v \times L_a, \\ I(t, \mathbf{x}, (0, \ell_d > 0, \ell_a)) &= 0 && \text{in } (0, T_\infty] \times \Omega_x \times L_d \times L_a, \\ I(0, \mathbf{x}, \ell) &= I_0 && \text{in } \Omega_x \times \Omega_\ell. \end{aligned}$$

The spatial movement of the population within the spatial neighbourhood can be incorporated through  $\mathbf{u}$ . The internal growth  $\mathbf{G} = (G_{\ell_v}, G_{\ell_d}, G_{\ell_a})^T$  is described by

$$G_{\ell_v} = \frac{d\ell_v}{dt} \quad G_{\ell_d} = \frac{d\ell_d}{dt} = 1, \quad G_{\ell_a} = \frac{d\ell_a}{dt} = 1.$$

The recovery,  $C_R$ , and infectious death,  $C_{ID}$ , rates are included as  $C = C_R + C_{ID}$ . The transport/mobility in the continuous model is modelled through the source term  $F$  defined by

$$F(t, \mathbf{x}) = \int_{\Omega_x} D(t, \mathbf{x}, \mathbf{s}) \left( \int_{\Omega_\ell} I(t, \mathbf{s}, \ell) d\ell \right) d\mathbf{s}.$$

Here, the mobility function,  $D(t, \mathbf{x}, \mathbf{s}) : (0, T_\infty] \times \Omega_x \times \Omega_x \mapsto \mathbb{R}$  is skew-symmetric with respect to  $\mathbf{x}$  and  $\mathbf{s}$  and satisfies  $-1 \leq D(t, \mathbf{x}, \mathbf{s}) \leq 1$ . In particular, there will be no mobility when  $D(t, \mathbf{x}, \mathbf{s}) = 0$  and the entire infected active population move from the place  $\mathbf{s}$  to the place  $\mathbf{x}$ , when  $D(t, \mathbf{x}, \mathbf{s}) = 1$ . Further,  $B_{nuc}$  and  $I_0$  are the nucleation function that quantifies the newly infected population and the initial active infected population, respectively. More details on the parameter modeling and on the numerical scheme can be found in [21]. COVID forecasting models should include antibody waning, vaccination, breakthrough infections, new immune escape variants, case-to-infection ratio (CIR) (unreported cases), social distancing, people interaction, comorbidity, travel, lockdown/unlock, etc. Eventually, all these factors affect the new infection, that is, the nucleation of COVID cases.

Let us first consider the modeling of antibody waning. Let  $N$  be the total number of people and  $N_{ab}$  be the number of people with antibody. Suppose there is no antibody waning, then the the susceptible ratio is defined as

$$S_R = \frac{N - N_{ab}}{N}.$$

To introduce antibody waning, we use the cumulative distribution function for the Weibull distribution

$$W(t) = 1 - \exp(-t/\lambda)^k, \quad t > 0.$$

Now define the antibody waning susceptible ratio as

$$S_R^w(t) = \frac{N - N_{ab}^w}{N}, \quad N_{ab}^w = (1 - W(t)) * N_{ab}.$$

Here,  $k = 5.67$  defines the shape of the curve,  $\lambda$  is the duration of antibody retention, and  $t$  is the time. Next, the vaccinated population must be removed from the susceptible population. Moreover, the efficacy of the vaccine should also be considered.

Suppose  $N_v$  is the total number of vaccinated people with 70% efficacy, then the vaccinated population is added to  $N_{ab}^w$ , i.e.,

$$N_{ab}^w = (1 - W(t)) * N_{ab} + 0.7N_v.$$

Serological surveys indicate that  $CIR$  varies between 20 and 60 or even more [24]. For instance, the first serological survey report of Karnataka shows  $CIR = 40$  in Karnataka [25]; that is, there are 39 unreported cases for every reported case in Karnataka during the first wave.  $CIR$  has to be incorporated into the model since  $CIR$  will significantly influence the nucleation. In our model, we included the  $CIR$  as follows:

$$N_{ab}^w = CIR N_{ab}^w + 0.7N_v.$$

Vaccine-dependent fatality is another critical feature in our model. The straightforward approach is to define the infectious rate as a function of the vaccinated population ratio. Alternatively, vaccine-dependent fatality can also be incorporated in the nucleation when the model accounts for the severity of the infection. In particular, the newly infected but vaccinated population must be added into the model with less severity. Since one of the internal variables in our model is infection severity  $\ell_v$  and the infectious death rate is zero for  $\ell_v < 0.64$ , the newly infected but vaccinated population is distributed within  $\ell_v < 0.64$ .

Next, we discuss the modeling of new immune escape variants, which are the primary source for new waves. Suppose  $N_{ab}(t)$  be the total antibody population on the day  $t$  of the introduction of a new immune escape variant. Let us assume a fraction of the antibody population,  $F_{nv} * N_{ab}$ , where  $0 \leq F_{nv} < 1$ , is still immune to the new variant of Covid. Then, the antibody population with a new immune-escape variant is redistributed as

$$N_{ab}^{nv}(t) = ((1 - F_{nv}) + F_{nv}(1 - W(t))) N_{ab}(t)$$

with  $k = 2.67$  and  $\lambda = 15$ . Incorporating all these models, the nucleation is given by

$$B_{nuc}(t, \mathbf{x}, \ell) = R(t) \int_{\Omega_\ell} [1 - \gamma_Q] I(t, \mathbf{x}, \ell) d\ell,$$

where

$$R(t) = R_0 S_R^w(t) f_1(t, S_D) f_2(\ell_v),$$

$$\gamma_Q = \frac{1}{1 + \exp(-(\ell_v - 0.4)/0.1)} \frac{1}{1 + \exp(-(\ell_d - 5.1)/2)}.$$

Here,  $\gamma_Q \in [0, 1]$  is the fraction of the infected population in quarantine, and it depends on testing, isolation, and comorbidity of the susceptible population. Further,

$$f_1(t, S_D) = \left[ 1 - \frac{1}{1 + \exp(-(S_D(t) - 0.5)/0.1)} \right],$$

$$f_2(\ell_v) = \begin{cases} 3\sqrt{\frac{2}{\pi}} \exp\left(\frac{-(\ell_v - 0.4)^2}{2(0.4/3)^2}\right) & 0 \leq \ell_v < 0.4, \\ 3\sqrt{\frac{2}{\pi}} \exp\left(\frac{-(\ell_v - 0.4)^2}{2V_R((1-0.4)/3)^2}\right) & 0.4 \leq \ell_v \leq 1. \end{cases}$$

The  $S_D \in [0, 1]$  in the function  $f_1(t, S_D)$  is a social distancing parameter, where  $S_D = 1$  implies a perfect social distancing and  $R \rightarrow 0$ . Hence,  $S_D$  is a key parameter to forecast lockdown and unlock phases. Next, the function  $f_2(\ell_v)$  is used to distribute the newly infected population as a function of vaccine-dependent severity. Here, the vaccinated ratio is defined as

$$V_R = \frac{N - N_{Vac}}{N},$$

where  $N_{Vac}$  is the total number of vaccinated people.

## 2.2 Control Variables

To fit the model parameters with the dynamics of new waves and to build scenarios based on them, computations are performed from Jul 1, 2020, to Jun 30, 2022, where the actual data of the emergence of the second wave, that is, the data until Jul 4, 2021, is used. Moreover, there are many uncertainties in terms of the COVID spread [20], and a few data is not accessible to accurately fit the parameters. To account for these uncertainties, we introduce the following control variables in the model. Further, we vary these influencing parameters to build a total of 972 scenarios. Finally, an ensemble forecast is made by combining the results of these scenarios.

### 2.2.1 Case-to-Infection Ratio (CIR)

The reported district-wise CIR values in the first serological survey for Karnataka [7, 26, 27] are used in our model. It has been reported that the average CIR in Karnataka was 40, and we denote this test case as CIR40. Further, to account for the uncertainty in the district-wise CIR estimate, we consider CIR50 as well, where the district-wise CIR is multiplied with a factor of 1.25.

### 2.2.2 Emergence of Immune-Escape New viral Variant (IENV)

The emergence of immune-escape new viral variants is critical for the emergence of the new waves [28]. The emergence of new variants and their transmission capability depends on the circulation of the existing variants and the antibody of the infected people. Further, we can observe that the duration between the peaks across the world is between nine and thirty-five weeks. In this study, we introduce the immune-escape new viral variant after (i) eight weeks (IENV-Jul21), (ii) sixteen weeks (IENV-Sep21), and (iii) twenty-four weeks (IENV-Nov21) of the second wave (May 2021) in Karnataka. Furthermore, we consider the transmission of immune-escape new viral variants in 33% (IENV-33P), 66% (IENV-66P), and 100% (IENV-100P) of the recovered population.

### 2.2.3 Antibody Waning (ABW)

The antibody waning could be a factor for reinfections and consequently induce new waves. A few studies show that the COVID antibody wains between five to twelve months. However, the COVID virus is new, and we do not have unambiguous evidence. Therefore, six scenarios by varying the duration of the antibody waning and the number of waning antibody populations are considered. More specifically, the antibody waning in 33% (ABW-33P), 66% (ABW-66P), and 100% (ABW-100P) of the recovered population in 150 (ABW-150) and 180 (ABW-180) days are considered.

### 2.2.4 Vaccination Rate

Vaccination is the key to contain the spread of the virus. In the absence of treatment, preventing the spread by vaccinating the entire population is the only solution to eradicate the virus. Nevertheless, with the present vaccination rate (as of Jul 4, 2021), people in the age group of 60 and above, 45-59 Yrs., 18-44 Yrs., would be fully vaccinated by Jul 9, 21, Jul 21, 2021, and Dec 13, 2021, respectively. In addition to the availability of the vaccines, the vaccine-hesitancy is also delaying the vaccination drive. Hence to study the effect of the vaccination rate on the emergence of new waves, three scenarios, (i) present rate, which is 280K vaccinations per day (VR-280K), (ii) 50% increase in the present rate (VR-420K), and (iii) double the present rate (VR-560K), are being considered.

### 2.2.5 COVID-appropriate behavior (CAB)

COVID-appropriate behavior (social distancing) plays a significant role in COVID spread, and it varies from region to region. Though the behavior of individuals determines the social distancing, NPIs such as lockdowns, restrictions on mass gatherings, curfew, etc., will also affect the social distancing. To study the effect of different compliance of CAB on new waves, three scenarios (i) Good compliance (Good CAB): similar to the behavior between Mar-May'21, (ii) Bad compliance (Partial CAB), and (iii) Worse compliance (No CAB), are considered.

## 2.3 Ensemble Forecast and Analysis Methods

We build an ensemble with 972 equi-probable scenarios by varying the above seven parameters. For each scenario we estimate the date of peak (if any) and the caseload at peak. A scenario without a peak is declared as a No-Wave scenario. The conditional probability of Wave-3 given a particular value for a parameter is then estimated, from only those scenarios with that chosen parameter value, as the ratio of scenarios leading to a wave 3 to the total number of scenarios. Conditional probability given two parameter values are estimated by considering only those scenarios in which those chosen parameter values are used. All statistical quantities (mean, median, confidence interval, inter-quartile range, probability distribution function and scatter) are estimated from this ensemble forecast.

## 3 Results

Our primary interest is to quantify the probability of a new COVID wave (Wave-3) in Karnataka, India, and to determine the distribution of the date and age-wise caseload of Wave-3's peak. Most importantly, we want to determine the causal factors for a new wave and recommend suitable interventions to reduce its likelihood and impacts.

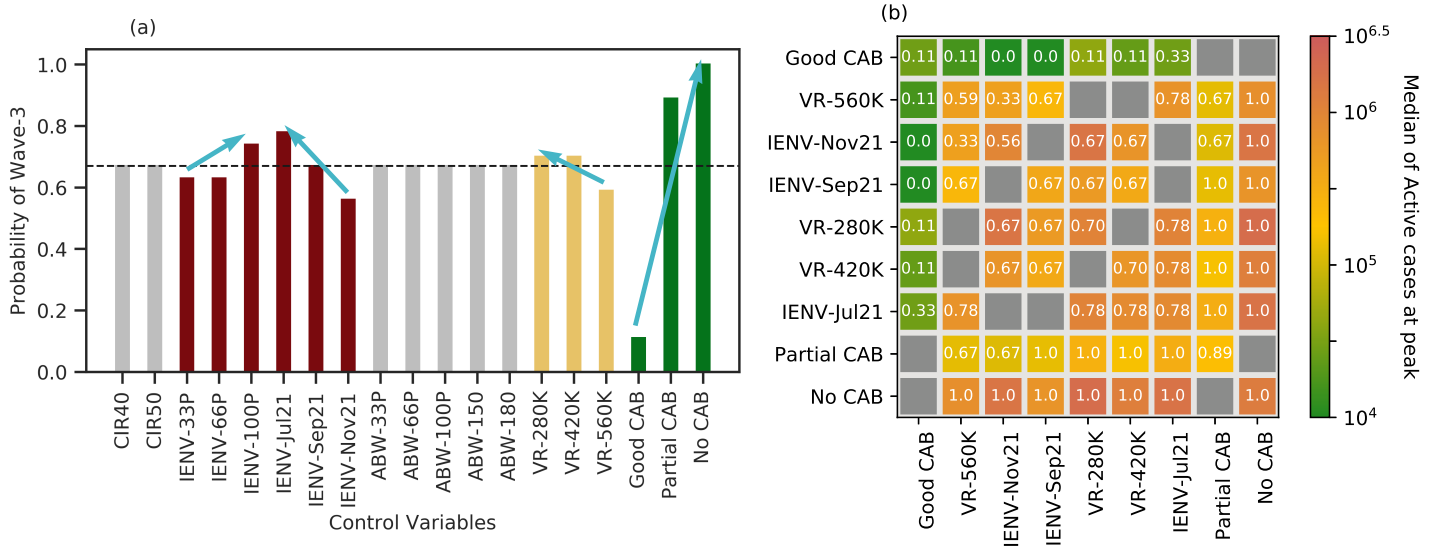
### 3.1 Causal factors for a new COVID wave in Karnataka, India

We first present the influence of the control variables (Sect. 2.2) on the emergence of a new wave and identify the significant causal factors. Among the considered 972 scenarios, Wave-3 is observed in 648 scenarios. In the event of Wave-3, the conditional probability (likelihood) of Wave-3 given a control variable is presented in Fig. 1 (a). Even though the Case-to-Infection ratio and the antibody waning could impact the emergence of new COVID waves, these factors are not very sensitive, that is, variations in these control variables are not influencing the probability of Wave-3. Nevertheless, we can observe that the variations in the new immune-escape variant, vaccination rate, and Covid-appropriate behavior significantly influence the emergence of a new COVID wave. Note that the length of the arrow indicates the sensitivity of the control variables. Moreover, we can see that the timing of the new immune-escape variant influence the probability of Wave-3 more than its reinfection percentage. Therefore, we characterize the timing of the new immune-escape variant, vaccination rate, and Covid-appropriate behavior as the primary causal factors and further analyze its influence on the Wave-3.

To investigate the pair-wise combined effect of these identified causal parameters on Wave-3, the conditional probability of Wave-3 given any two of these causal parameters is presented in Fig. 1 (b). The values in the cells represent the conditional probability of Wave-3 given two causal parameters, and the colors indicate the median of active cases at peak.

This analysis reveals that the CAB is the most responsible factor for the third wave. Further, no compliance of CAB guarantees the third wave even with a doubled





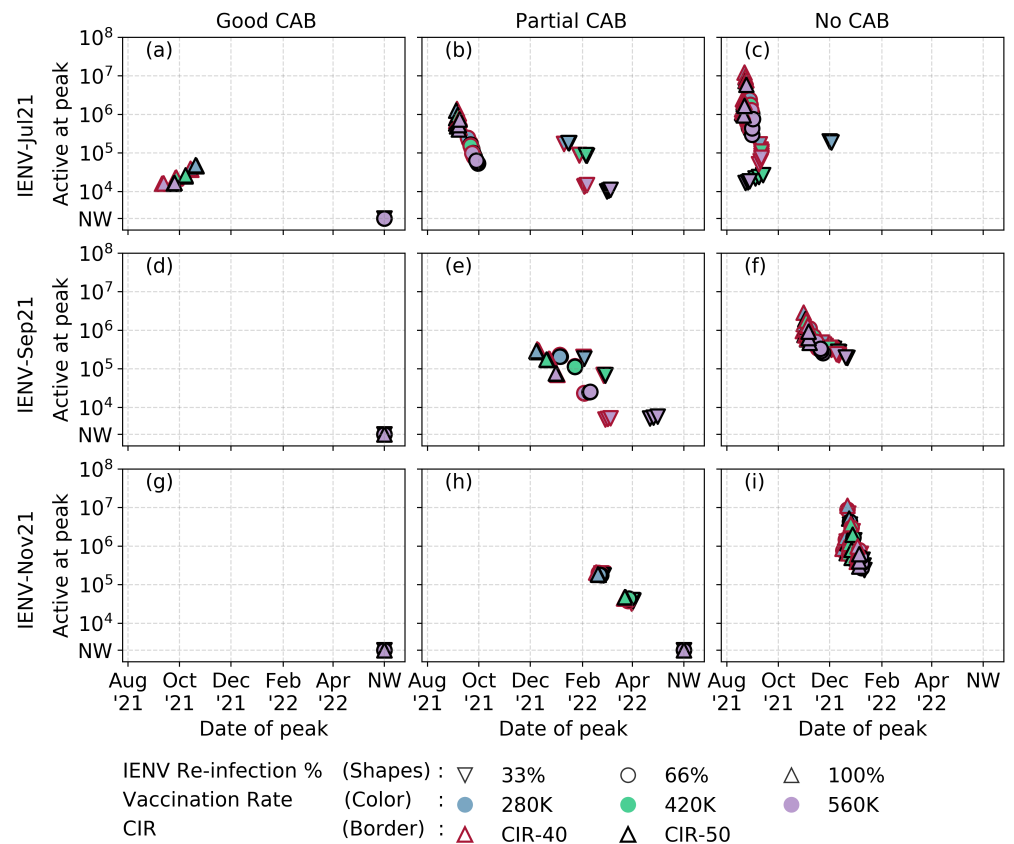
**Fig 1.** The conditional probability of Wave-3 given a control variable (a) and two control variables (b) are depicted. The values in the cells represent the conditional probability of Wave-3 given two control variables, whereas the color indicates the median of active cases at peak.

vaccination rate and more active cases at the peak of Wave-3. On the contrary, good compliance of CAB and doubling the present vaccination rate effectively reduce the probability of Wave-3 and the number of active cases in the event of Wave-3.

### 3.2 Wave-3 peak date and active caseload

The quantities of interest for policy-making are the peak date and active caseload at peak. Now, we proceed to analyze these two quantities in different scenarios corresponding to the primary causal factors identified above. Fig. 2 shows a scatter plot between the date of Wave-3's peak and active caseload at peak for different scenarios. A special axis entry NW is used to depict the scenarios that do not show a new wave. The arrangement and color of scatter markers in the figure help us visualize the joint-effect of adherence to Covid-appropriate behavior by the public, IENV emergence and re-infection fraction, vaccination rate, and CIR in the population on caseload and date of peak. Each panel (a-i) in Fig. 2 shows only those scenarios in which the variable mentioned in the column (CAB) and the row (IENV emergence time) are active. For example, panel (a) shows all the scenarios where the CAB is set to "Good" level and IENV emergence time is set to July 2021. Within the panel, each point represents a particular scenario. Each scenario is colored by the vaccination rate, marked by a marker that represents the IENV re-infection fraction, and the border is colored by the CIR.

First, we see that when Good CAB is active (panels a,d,g) as opposed to the other two values for CAB, the active case at the peak is lower. Significantly, scenarios with Good CAB and IENV emergence time beyond September 2021 (panels d,g) do not see a new wave. In panel (a), we see a peak only when 100% IENV re-infection occurs, and the other two re-infection fractions do not lead to a new wave. This result shows that Good CAB is the primary intervention strategy for preventing COVID waves, a well-known fact. However, in reality, the financial and livelihood requirements to re-open economies will result in a relaxation of CAB, and only a Partial CAB is likely to be followed. In those scenarios with partial CAB (panels b,e,h), if an immune escape



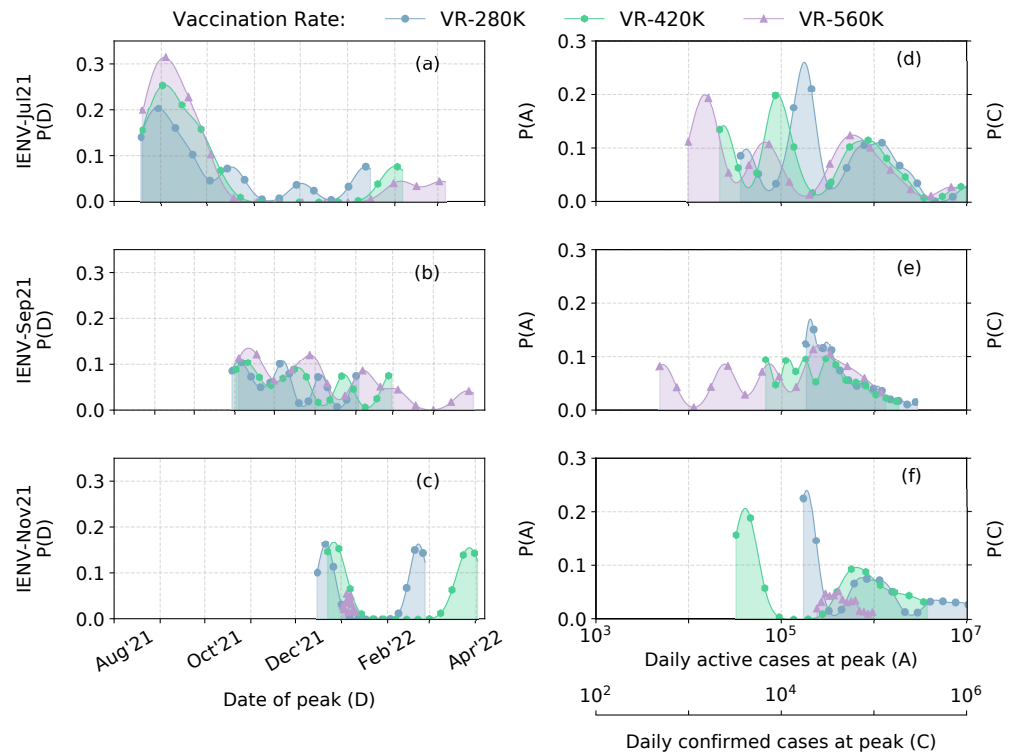
**Fig 2.** Scatter plot of active cases at the peak of Wave-3 v/s the date of peak for different scenarios of CAB and timing of the immune-escape variant with different vaccine rates.

new variant emerges in July 2021 or September 2021, Wave-3 is inevitable in Karnataka. Moreover, we see that vaccination rate is the key governing factor that reduces the active caseload at peak. Crucially, reaching a vaccination level of 580k per day in Karnataka will ensure no Wave-3 if no immune escape new variant starts spreading in the population before November 2021. A higher CIR (meaning lower reported cases during Wave-2) will result in a delayed peak but at the same active caseload as seen by the clustering among the same color and marker shapes.

To complete the analysis, we look at the scenarios with no CAB. Under these scenarios, we expect the highest levels of the caseload at peak and an assured emergence of new COVID waves in Karnataka. Fortunately, we see that vaccination reduces the caseload but doesn't prevent the inevitable Wave-3. The re-infection fraction significantly affects the caseload in the No CAB, IENV-Jul'21 scenarios (panel c). We see wide distribution in the caseload at the peak but a narrow distribution in the peak date.

To further understand Wave-3 properties from the perspective of vaccination, Fig. 3 shows the probability distribution function of peak date and active cases at peak for different vaccination rates. Here, we only look at the scenarios resulting in a Wave-3 and ignore those that do not lead to a Wave-3. We see that almost all the distribution curves are multi-modal, indicating the cross effects of CAB and IENV re-infection rates. It is evident from Fig. 3 that the emergence of a new immune-escape variant is likely to be followed by a third wave within 45 to 60 days. In most cases, a higher vaccination





**Fig 3.** The probability of the predicted date of peak, the number of active cases and the number of confirmed cases on the day of peak for different vaccination rates and timing of emergence of immune-escape new variants, i.e., in Jul'21, Sep'21 and Nov'21.

rate helps to bring down the active caseload at peak. Doubling the daily vaccination rate (VR-560K, Fig. 3(c),(f)) results in fewer scenarios with Wave-3 despite the emergence of an immune-escape new variant in November.

Table 2 shows a summary of the median peak date, inter-quartile range of peak date and the minimum and maximum active caseload during this period of inter-quartile range for scenarios with specific combinations of IENV emergence time and vaccination rates. This table can be used as a ready reference for policy making.

### 3.3 Effect of Wave-3 on different age-groups

The data until 10 June 2021 shows that 4.2%, 5%, 53.2%, 23% and 14.645 % of the people in the age groups 0–11, 12–17, 18–44, 45–59, and above 60 years were infected with COVID-19. However, India started an age-dependent vaccination drive for 18+ population only. As such, for Wave-3, we expect a higher contribution to the daily confirmed caseload from children (0–17 years). To estimate this impact on different age groups, we analyze the age-wise daily confirmed caseload at Wave-3 peak and compare it with the corresponding numbers at Wave-2. Table. 2 shows the statistics of the ratio between the 7-day average of daily confirmed cases of Wave-3 and Wave-2 peaks. The first line in each cell shows the ensemble mean with the 95% confidence interval and the second line shows the ensemble median. We see that the children (0-17 years) are the worst affected due to unavailability of vaccine. In the worst case, with VR-280K and IENV-Nov21, the daily confirmed cases of children (0–11 years) at the Wave-3 peak will be on an average twenty times more than that of the daily confirmed cases at Wave-3 peak. On an average, the impact of Wave-3 on children (0–17 years) would be seven

**Table 1.** Statistics for total active cases and the date of peak at various vaccine rates, and COVID-appropriate behavior. The first value in the cell is the the median date of peak and the mean value of active cases on the median date of peak, followed by the first and third quartile values for the date of peak, and the minimum and maximum number of active cases in the inter-quartile range.

Vaccination rate per day	Emergence of IENV	Good CAB	Partial CAB	No CAB
280K	Jul'21	15-Oct'21 (36K) 14-Oct'21, 21-Oct'21 36K – 48K	21-Sep'21 (173K) 05-Sep'21, 10-Jan'22 144K – 207K	27-Aug'21 (451K) 22-Aug'21, 10-Sep'21 451K – 2M
	Sep'21	No Wave-3	05-Jan'22 (220K) 10-Dec'21, 02-Feb'22 207K – 229K	08-Nov'21 (674K) 02-Nov'21, 24-Nov'21 527K – 766K
	Nov'21	No Wave-3	21-Feb'22 (189K) 19-Feb'22, 24-Feb'22 179K – 197K	22-Dec'21 (7M) 20-Dec'21, 24-Dec'21 616K – 11M
420K	Jul'21	27-Sep'21 (23K) 26-Sep'21, 08-Oct'21 23K – 26K	23-Sep'21 (119K) 06-Sep'21, 28-Jan'22 89K – 134K	29-Aug'21 (1M) 23-Aug'21, 08-Sep'21 370K – 1M
	Sep'21	No Wave-3	23-Jan'22 (114K) 23-Dec'21, 26-Feb'22 112K – 117K	13-Nov'21 (479K) 05-Nov'21, 02-Dec'21 377K – 529K
	Nov'21	No Wave-3	27-Mar'22 (38K) 24-Mar'22, 31-Mar'22 37K – 43K	27-Dec'21 (1M) 26-Dec'21, 29-Dec'21 460K – 3M
560K	Jul'21	12-Sep'21 (16K) 11-Sep'21, 24-Sep'21 15K – 16K	25-Sep'21 (79K) 08-Sep'21, 04-Feb'22 57K – 88K	28-Aug'21 (225K) 23-Aug'21, 31-Aug'21 17K – 656K
	Sep'21	No Wave-3	03-Feb'22 (23K) 02-Jan'22, 03-Mar'22 23K – 25K	17-Nov'21 (340K) 06-Nov'21, 11-Dec'21 288K – 406K
	Nov'21	No Wave-3	No Wave-3	05-Jan'22 (529K) 04-Jan'22, 08-Jan'22 296K – 792K

times more than that of Wave-2.

### 3.4 Active Cases During May'21 - Jun'22

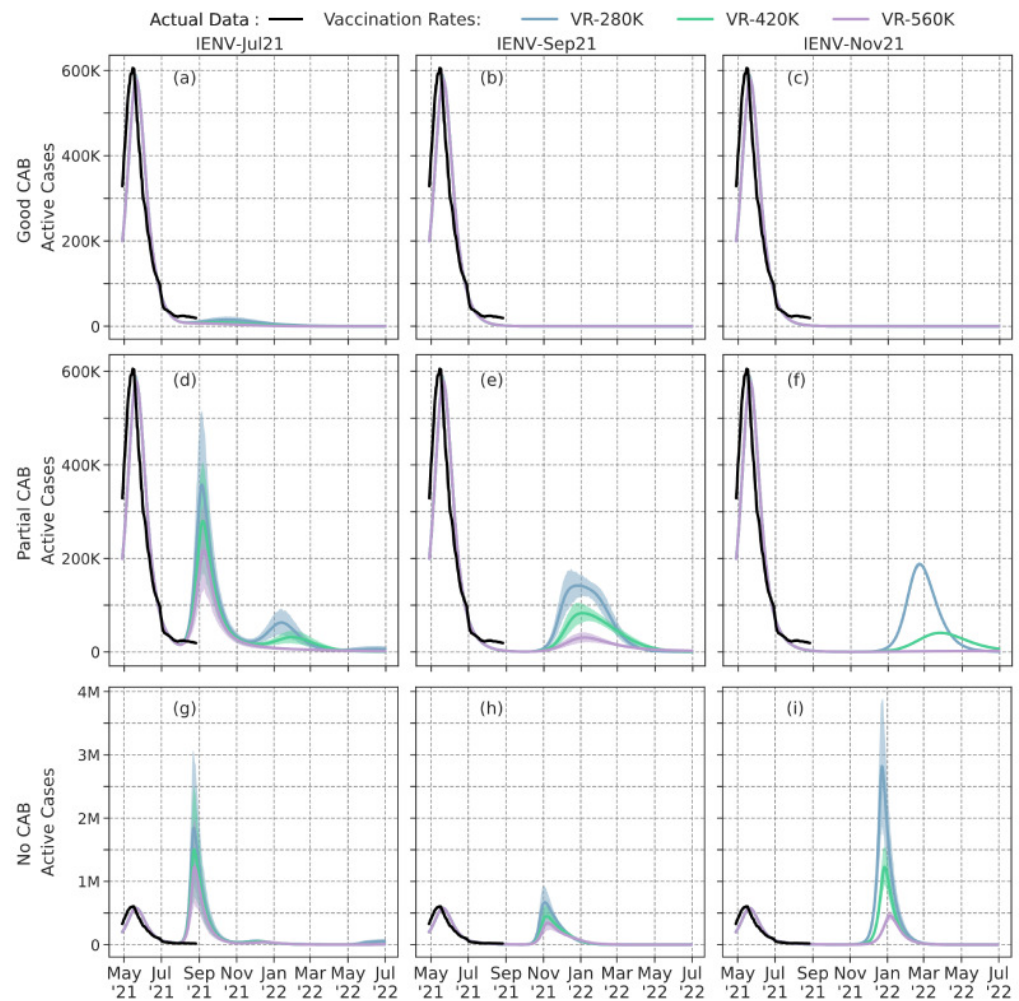
Fig. 4 shows the ensemble forecast of active cases in Karnataka from May 2021 to June 2022. Each panel shows the scenarios given a pair of CAB and IENV emergence-time control variables. The colored solid line is the mean value of the ensemble of active cases, and the shaded zone is the region of uncertainty (each color corresponds to a particular vaccination rate). The rows in Fig. 4 represent the good, partial, and no CAB, whereas the columns represent the emergence of IENV in Jul'21, Sep'21, and

**Table 2.** Predicted impact of Wave-3 on different age groups with different vaccination rate. The first value in the cell is the mean with the confidence interval of ratios of the 7-day average of confirmed cases at Wave-3 peak and Wave-2 peak, followed by the median.

Vaccine Rate	Emergence of IENV	0-11 Years	12-18 Years	18+ Years
280K	Jul'21	10.17 ± 5.1 1.92	7.88 ± 3.96 1.29	2.36 ± 1.27 0.24
	Sep'21	6.18 ± 1.64 3.22	4.66 ± 1.26 2.33	0.96 ± 0.29 0.44
	Nov'21	21.12 ± 8.72 4.68	15.78 ± 6.57 3.34	2.76 ± 1.18 0.54
420K	Jul'21	8.69 ± 4.42 1.06	6.75 ± 3.42 0.85	1.77 ± 0.96 0.13
	Sep'21	5.01 ± 1.3 2.88	3.66 ± 0.97 2	0.51 ± 0.16 0.23
	Nov'21	11 ± 3.73 3.65	8.06 ± 2.79 2.55	0.88 ± 0.31 0.27
560K	Jul'21	7.4 ± 3.76 0.79	5.79 ± 2.96 0.61	1.34 ± 0.75 0.07
	Sep'21	4.05 ± 1.21 3.11	2.91 ± 0.88 1.98	0.3 ± 0.1 0.15
	Nov'21	8.85 ± 1.32 7.61	6.42 ± 1.05 5.41	0.38 ± 0.06 0.32

Nov'21, respectively.

These plots confirm the observations from Sect. 3.2, but now showing a time series of the active caseload distribution. As mentioned earlier, the scenarios with Partial CAB (middle row in Fig. 4) are likely to be followed due to the re-opening of economy including offices and educational institutions. The impact of the new wave will be more when it emerges early (Fig. 4 (d)). At the peak of the new wave, the active caseload will be very high with the present vaccination rate, around 350K, when a new variant emerges before September 2021. If we assume that there is no antibody waning at least for 180 days, then, a new wave is predicted only at the beginning of next year, even if a new variant emerges before September 2021. In this case, the impact will be minimal. The benefits of increasing the vaccination rate can clearly be seen in Fig. 4 (e). Though a new wave is predicted in a few scenarios, the active caseload is very minimal with the doubled vaccination rate. If a new variant does not emerge until the end of October 2021, then a new wave can be suppressed by doubling the vaccination rate (Fig. 4 (f)). In this case, the study shows that the active caseload will be as high as 200K with the present vaccination rate even if the new variant emerges after October 2021. This finding emphasizes the necessity of vaccinating the entire population by December 2021.



**Fig 4.** Time series of Mean of active cases with 95% Confidence Interval

## 4 Discussion and Summary

We have undertaken a comprehensive study by quantifying several uncertain causal factors to provide modeling evidence on the emergence of a new COVID wave in Karnataka, India. A total of 972 ensemble members were formed by varying seven key causal factors in our Population Balance PDE model for infectious disease spread [21]. This first-principles based model incorporates the nonlinear dynamics between all the causal factors and the confirmed, active, recovered, and deceased caseload of Covid-19 in Karnataka, India. The major findings of this study are:

1. A new wave is most sensitive to mutation (emergence of immune-escape new variant) timing, compliance of COVID Appropriate Behavior, and vaccination rate.
2. Practically, the critical causal factor is compliance to CAB, which can be regulated with appropriate NPI such as mobility restriction, masking and physical distancing mandates, and crowd control measures.
3. Emergence of new variants beyond Sep'21 is not likely to induce new waves when the social distancing is good (lockdown-like restrictions).

4. Among scenarios with Wave-3, increasing the vaccination rate reduces the peak active caseload and consequently decreases hospitalization, ICU and Oxygen requirements.
5. Scenarios with a vaccination rate of 480k per day (2x the average vaccination rate as on July 4, 2021) should have a lower probability of Wave-3 than other scenarios
6. The daily confirmed cases of children (Age 0-11 and 12-17 years) at peak could be on an average seven times more than the corresponding daily confirmed cases at Wave-2 peak.

The above findings raise an important call for action to the state's public health response. First, policies to promote CAB so as to ensure at least partial compliance from the population is required. Next, increasing the vaccination rate so as to fully vaccinate the entire population by mid-Dec is necessary to counter new variants and antibody waning, if any. Based on the predicted increase in daily confirmed case load of children in a potential wave 3 compared to data of wave 2, the policymakers can plan pediatric COVID care centers.

Though the estimated Wave-3 numbers are for the state of Karnataka, the dynamics of the ensemble forecast will be applicable to all other states of India. Therefore, the above recommended policy interventions can be imposed nation-wide.

One limitation of the ensemble forecast is that we have not considered the possibility of new variants infecting the vaccinated population, but we have considered a fixed efficacy of 70% for the vaccinated population that can partially offset the limitation. Further, the recovery rate is fitted to the data from July 2020 to May 2021; however, it could be more due to vaccination effect or less due to new variants.

## Acknowledgments

We thank members of the Karnataka Technical Advisory Committee, and Dr. Vinod Paul of Niti Aayog, Govt. of India for insightful questions and discussion. All reported COVID data is obtained from [www.covid19india.org](http://www.covid19india.org). We are thankful to the MHRD Grant No. STARS-1/388 (SPADE) for partial support. Thivin is supported by a Graduate Fellowship from the Ministry of Education, Govt. of India. Deepak acknowledges the Arcot Ramachandran Young Investigator award for partial support.

## References

1. Government of India. COVID-19 dashboard. Last accessed on Aug 26, 2021. <https://www.mygovin/covid-19;>
2. Foy BH, Wahl B, Mehta K, Shet A, Menon GI, Britto C. Comparing COVID-19 vaccine allocation strategies in India: A mathematical modelling study. *International Journal of Infectious Diseases*. 2021;103:431–438. doi:<https://doi.org/10.1016/j.ijid.2020.12.075>.
3. Samarasekera U. India grapples with second wave of COVID-19. *The Lancet Microbe*. 2021;2(6):e238.
4. Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, Das M, et al. Convergent evolution of SARS-CoV-2 spike mutations, L452R, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. *BioRxiv*. 2021;.

5. Budhiraja S, Indrayan A, Aggarwal M, Jha V, Jain D, Tarai B, et al. Differentials in the characteristics of COVID-19 cases in Wave-1 and Wave-2 admitted to a network of hospitals in North India. *medRxiv*. 2021;.
6. Asrani P, Eapen MS, Hassan MI, Sohal SS. Implications of the second wave of COVID-19 in India. *The Lancet Respiratory Medicine*. 2021;.
7. Babu GR, Sundaresan R, Athreya S, Akhtar J, Pandey PK, Maroor PS, et al. The burden of active infection and anti-sars-cov-2 igg antibodies in the general population: Results from a statewide sentinel-based population survey in karnataka, india. *International Journal of Infectious Diseases*. 2021;108:27–36.
8. <https://pib.gov.in/PressReleasePage.aspx?PRID=1715485>. Press Information Bureau. 2021;.
9. Viboud C, Sun K, Gaffey R, Ajelli M, Fumanelli L, Merler S, et al. The RAPIDD ebola forecasting challenge: Synthesis and lessons learnt. *Epidemics*. 2018;22:13–21.
10. Johansson MA, Apfeldorf KM, Dobson S, Devita J, Buczak AL, Baugher B, et al. An open challenge to advance probabilistic forecasting for dengue epidemics. *Proceedings of the National Academy of Sciences*. 2019;116(48):24268–24274.
11. Reich NG, Brooks LC, Fox SJ, Kandula S, McGowan CJ, Moore E, et al. A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States. *Proceedings of the National Academy of Sciences*. 2019;116(8):3146–3154.
12. Leutbecher M, Palmer TN. Ensemble Forecasting. *J Comput Phys*. 2008;227(7):3515–3539. doi:10.1016/j.jcp.2007.02.014.
13. Babu GR, Ray D, Bhaduri R, Halder A, Kundu R, Menon GI, et al. COVID-19 Pandemic in India: Through the Lens of Modeling. *Global Health: Science and Practice*. 2021;.
14. Mandal S, Das H, Deo S, Arinaminpathy N. Combining serology with case-detection, to allow the easing of restrictions against SARS-CoV-2: a modelling-based study in India. *Scientific reports*. 2021;11(1):1–9.
15. Ranjan R, Sharma A, Verma MK. Characterization of the Second Wave of COVID-19 in India. *medRxiv*. 2021;.
16. Chen J, Wang R, Wang M, Wei GW. Mutations strengthened SARS-CoV-2 infectivity. *Journal of molecular biology*. 2020;432(19):5212–5226.
17. Adiga A, Athreya S, Lewis B, Marathe MV, Rathod N, Sundaresan R, et al. Strategies to Mitigate COVID-19 Resurgence Assuming Immunity Waning: A Study for Karnataka, India. *medRxiv*. 2021;doi:10.1101/2021.05.26.21257836.
18. Salvatore M, Basu D, Ray D, Kleinsasser M, Purkayastha S, Bhattacharyya R, et al. Comprehensive public health evaluation of lockdown as a non-pharmaceutical intervention on COVID-19 spread in India: national trends masking state-level variations. *BMJ open*. 2020;10(12):e041778.
19. Ray D, Salvatore M, Bhattacharyya R, Wang L, Du J, Mohammed S, et al. Predictions, role of interventions and effects of a historic national lockdown in India's response to the COVID-19 pandemic: data science call to arms. *Harvard data science review*. 2020;2020(Suppl 1).



20. Mandal S, Arinaminpathy N, Bhargava B, Panda S. Plausibility of a third wave of COVID-19 in India: A mathematical modelling based analysis. *Ind J Med Res* DOI. 2021;10.
21. Ganesan S, Subramani D. Spatio-temporal predictive modeling framework for infectious disease spread. *Sci Rep*. 2021; p. 6741. doi:10.1038/s41598-021-86084-7.
22. Ganesan S, John V, Matthies G, Meesala R, Abdus S, Wilbrandt U. An object oriented parallel finite element scheme for computations of PDEs: Design and implementation. 2016 IEEE 23rd International Conference on High Performance Computing Workshops (HiPCW). 2016; p. 2–11. doi:10.1109/HiPCW.2016.023.
23. Wilbrandt U, Bartsch C, Ahmed N, Alia N, Anker F, Blank L, et al. ParMoon – a modernized program package based on mapped finite elements. *Computers and Mathematics with Applications*. 2016;74:74–88. doi:10.1016/j.camwa.2016.12.020.
24. Murhekar MV, Bhatnagar T, Selvaraju S, Saravanakumar V, Thangaraj JWV, Shah N, et al. SARS-CoV-2 antibody seroprevalence in India, August–September, 2020: findings from the second nationwide household serosurvey. *The Lancet Global Health*. 2021;9(3):e257–e266.
25. Murhekar MV, Bhatnagar T, Selvaraju S, Rade K, Saravanakumar V, Thangaraj JWV, et al. Prevalence of SARS-CoV-2 infection in India: Findings from the national serosurvey, May–June 2020. *Indian Journal of Medical Research*. 2020;152(1):48.
26. Lau EH, Tsang OT, Hui DS, Kwan MY, Chan Wh, Chiu SS, et al. Neutralizing antibody titres in SARS-CoV-2 infections. *Nature communications*. 2021;12(1):1–7.
27. Muecksch F, Wise H, Batchelor B, Squires M, Semple E, Richardson C, et al. Longitudinal Serological Analysis and Neutralizing Antibody Levels in Coronavirus Disease 2019 Convalescent Patients. *The Journal of Infectious Diseases*. 2020;223(3):389–398. doi:10.1093/infdis/jiaa659.
28. Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, Das M, et al. SARS-CoV-2 Spike Mutations, L452R, T478K, E484Q and P681R, in the Second Wave of COVID-19 in Maharashtra, India. *Microorganisms*. 2021;9(7):1542.