COVID-19 epidemic scenarios into 2021 based on observed key superdispersion events

Mario Santana-Cibrian[1*], M. Adrian Acuña-Zegarra[2], Carlos E. Rodriguez Hernandez-Vela[3], Ramsés H. Mena[3], and Jorge X. Velasco-Hernández[4]

[1]CONACyT-Instituto de Matematicas UNAM, Queretaro, 76230, Mexico
[2]Departamento de Matematicas, Universidad de Sonora, Sonora, 83000, Mexico
[3]Instituto de Investigacion en Matematicas Aplicadas y Sistemas UNAM, CDMX, 01000, Mexico
[4]Instituto de Matematicas UNAM, Queretaro, 76230, Mexico
[*]Corresponding author: msantana@im.unam.mx

ABSTRACT

Key high transmission dates for the year 2020 are used to create scenarios to model the evolution of the COVID-19 pandemic in several states of Mexico for 2021. These scenarios are obtained through the estimation of a time-dependent contact rate, where the main assumption is that the dynamic of the disease is heavily determined by the mobility and social activity of the population during holidays and other important calendar dates. First, changes in the effective contact rate on predetermined dates of 2020 are estimated. Then, using the instantaneous reproduction number to characterize the status of the epidemic ($R_t \approx 1$, $R_t > 1$ or $R_t < 1$) this information is used to propose different scenarios for the number of cases and deaths for 2021. The main assumption is that the effective contact rate during 2021 will maintain a similar trend to that observed during 2020 on key calendar dates. All other conditions are assumed to remain constant in the time scale of the projections. The objective is to generate a range of scenarios that could be useful to evaluate the possible evolution of the epidemic and its likely impact on incidence and mortality.

1 Introduction

The effective contact rate for Mexico, along the year 2020, is estimated using a standard generalization of the SIR model without age structure. This effective contact rate is then used as a baseline to explore scenarios for 2021. Our approach is similar to the one presented in[1]. However, to determine changes in the effective contact rate, besides looking at the actual effects of governmental mitigation measures, we look at particular events on dates related to civic, religious and official vacation periods that are known in advance each year. These events triggered, in 2020, increases in transmission and upward changes in positivity, hospitalization and mortality rates as we show in Section 2. In Mexico, reported cases are very sensitive to actions that are not directly related to disease transmission. Reporting delays vary in a rather random form not only within the same hospital, but also depending on the state or municipality, there exists sub-reporting of cases and deaths, and there is a pronounced deficiency in testing. In this work, the epidemic developing in Mexico is the main focus of interest, taking as examples two particular ones (Mexico City and Queretaro state). Here, one primordial aspect of human behavior that is an important factor and has driven the evolution of the pandemic is explored: superspreading events occurring in particular calendar dates associated with religious or civic holidays on the one hand and commercial incentives on the other. During or after these events (depending on their length), the effective contact rate changed and therefore, we use them as change points in the spreading rate[1]. The effect of these change points or key calendar dates on the contact rate occurs at a much faster scale than changes due to the implementation or lifting of non-pharmaceutical interventions (NPIs)[1,2] (see Section 4). One central reason is that all these dates are known in advance and activities for them can be planned (short vacations, family visits, etc). As a consequence of the COVID-19 epidemic impact on the regional and global economy, one of the most relevant problems that each and every country faces is to decide how and when businesses, public centers, tourism, schools and universities can safely reopen[3]. For decision-makers it is of the highest importance to count with plausible scenarios for the evolution of the SARS-CoV-2 pandemic in order to design effective reopening strategies. This knowledge is even more pressing in countries that lack the full infrastructure to acquire a more precise or, perhaps we should say, a less uncertain idea of the behavior of the pandemic on days or, ideally, weeks into the future.

During 2020 much effort was centered on projecting the fate of the COVID-19 pandemic and evaluating the efficacy of the mitigation strategies adopted to contain it[4,5]. Around the world, the implementation of these measures has varied in strength. Examples include lockdowns, use of face masks, regulations on the number of people allowed in meetings and so forth, and

they range from strict and mandatory enforcement by the state or government to a voluntary personal decision. The rationale behind a particular version of the mitigation strategy followed in each country or region is mainly based on local factors that combine public health, economic and political conditions and perspectives[6]. However, one factor that has shown to be decisive in the success or failure of a given containment or mitigation strategy is human behavior. Modeling human behavior is elusive, although recently, some efforts in this direction have been reported[7].

The transmission dynamics of the pandemic has been modeled using many different methodologies, several of them centered on estimating the effective reproduction number R_t or using some version of the well-known Kermack-McKendrick model to account for the different infectious stages and to measure the impact of control and mitigation measures, like vaccination or NPIs on incidence and mortality[8–11]. We are interested in the characterization of the transmission dynamics of the disease with a view to develop a methodology of epidemic scenario forecasting that uses the known history of the disease recorded on the year 2020. This history is synthesized in the time series of incidence and mortality, in the known heightened transmission events reported during that year and in a mathematical model that approximates the disease dynamics.

The onset of an epidemic outbreak is not only a function of increases in the effective contact rate at a given time, $\beta(t)$, but also on the proportion of susceptible available for infection. Both of these parameters should be large enough to bring $R_t > 1$. During 2020 in Mexico the effective reproduction number fluctuated around 1 roughly from June to mid-September. This implies that an increase in $\beta(t)$ had the potential to increase R_t above 1, as was indeed the case for most of the key dates listed in Table 2. The situation is not the same in 2021. In March and April 2021, R_t in Mexico has stayed on average below 1. This implies that the observed increases in R_t for 2020 may not be large enough to generate an epidemic outbreak in 2021. However, this year immunity loss is playing a role that was not significant for most of 2020. Immunity loss replenishes the susceptible pool and, therefore may rise the proportion of susceptible above threshold and trigger and epidemic ($R_t > 1$) for a sufficiently large increase of $\beta(t)$.

Figure 1 shows the epidemic curves of Mexico City and Queretaro state up to March 2021. There are two interesting features in these curves that we want to point out. One is the initial slow growth of the epidemic that seems almost linear in both places with a shorter growth period in Queretaro state; the other is the long plateau that develops in Mexico City and a much shorter one in Queretaro state, right after the epidemic reached the day of maximum incidence in late May[12,13]. This plateau was associated with the effective reproduction number taking the value $R_t \approx 1$ from June to mid-September in both places.

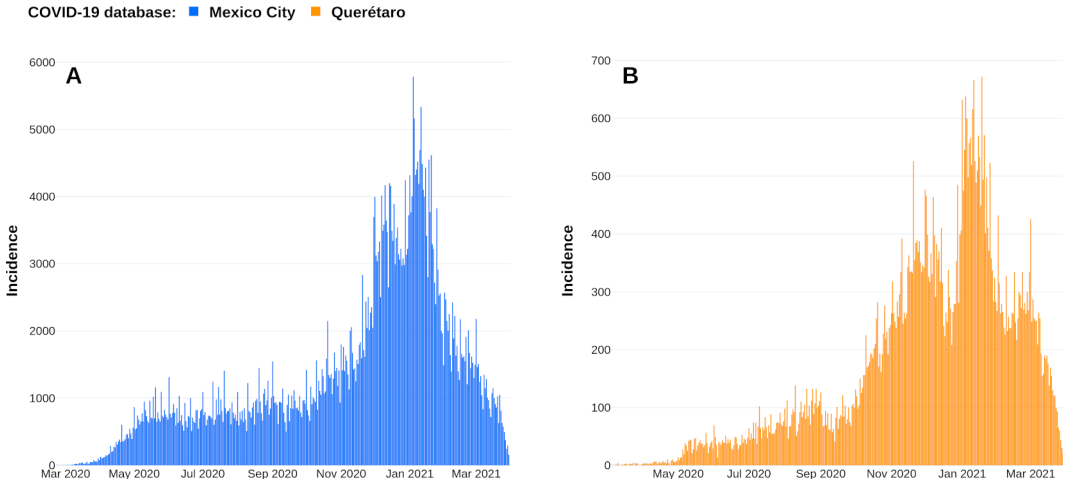

Figure 1. Epidemic curves of A) Mexico City and B) Queretaro state. Data shown are from the outset of the pandemic until March 31, 2021. In both places the quasi-linear initial growth can be appreciated. In Mexico City a long plateau covering from June to mid September is evident. Note the different scales of the epidemic curves.

The initial quasi-linear growth of the epidemic cannot be explained by classical models. Previously we have argued[12], that the initial slow growth was a consequence of an early application of largely voluntary mitigation measures. The first case in Mexico occurred by the end of February 2020, and the first mitigation measures were applied 25 days later, on March 23, 2020. Moreover, there were superdispersion events in Mexico City during Easter celebrations (April 6-12) and early May (April 30 to May 10) that shifted the day of maximum incidence to the end of May[12,14], and pushed the epidemic into a quasi-stationary state characterized by values of $R_t \approx 1$. This behavior, observed around the world, has been explored in[15–17]. In particular,[17] claim that this quasi-linear growth and the maintenance of the effective reproduction number around $R_t \approx 1$ for sustained periods of time, involves critical changes in the structure of the underlying contact network of individuals.

The characteristic plateau of Mexico City and epidemics elsewhere has been marked by a succession of events that, we

postulate, were essentially superspreading events[12,14]. Superspreading is an important phenomenon that has shown to be determinant for many infectious diseases[18]. The epidemic dynamics observed in Mexico City and, in general, in the country, has been driven by events associated with heightened mobility and increased social activity during holidays and other important calendar key dates, where formal job attendance is diminished. Superspreading events are associated to heightened population mobility and social activity[19–27].

At a very broad level, the effective contact rate ($\beta(t)$) changes when behavioral patterns modify the underlying contact structure of the populations[15,17,18]. These changes have not occurred at random but, we postulate, have been centered on pre-defined calendar key dates associated with variations of the strength of the compliance with mitigation measures (e.g., changes in the traffic light risk monitoring system), national holidays (e.g., Virgin of Guadalupe day) or commercial opportunities (Buen Fin, Mexican equivalent to the Black Friday in the USA).

The key dates used in the estimation process and the estimated changes in the contact rate are listed in Table 2. The statistical details of the estimation can be found in the SI.

2 Results

In order to identify changes in daily positivity time series trends, hospitalization, and mortality rates in Mexico City and Queretaro during the key calendar dates listed in Table 2, we start by desegregating the population by age groups: children $0-18$, young adults $19-40$, adults $41-60$ and senior >60 and looking at the tendencies. We are interested in increments of these indicators as evidence of superspreading events and therefore, changes in the effective contact rate. With this simple analysis we have evidence to support our hypothesis of the existence of superspreading events in the key calendar rates of interest and we can proceed to estimate a time-dependent effective contact rate for the epidemic. The idea is simple: subdivide last year 2020 into time segments defined by the key calendar dates (Table 2) and estimate the effective contact rate for each of these segments. As a result, we obtain a piece-wise linear function for $\beta(t)$.

To support our key calendar dates superdispersion hypothesis, we look at three main indicators in 2020: positivity, hospitalizations and mortality. When analyzing raw counts of the number of cases by day, the variability between days makes it difficult to determine the behavior of the curve. To deal with this problem, we have chosen to use a simple moving average alternative where the window size is set to seven days thus making it easier to identify overall trends. In every case, our starting date is Monday March 2, 2020. We define the positivity rate as the ratio of positive tests to suspected cases, the hospitalization rate as the ratio of hospitalized to positive cases and the mortality rate as the ratio of deceased to positive cases. For example,the mortality rate for children in age group $0-18$ at time t is calculated via

$$(\text{Mortality rate age group 0-18})_t = \frac{(\text{Deceased age group 0-18})_t}{(\text{Positive cases age group 0-18})_t}.$$

The examples we have chosen represent a large city (Mexico City or CDMX) of about 10 million inhabitants and a medium-small city (Santiago de Queretaro or Qro) of about two and a half million inhabitants. At the beginning of an epidemic the uncertainty generated by scarce data is very large and trends and patterns are not clearly observed. This implies that, for the first quarter of 2021, the $\beta(t)$ estimate has to be adjusted to control this large uncertainty (see Section 4). In Mexico, testing rate has been limited to symptomatic cases that seek medical assistance at hospitals and health centers and, thus, it has stayed, on average, well above 20% for the duration of the epidemic. The positivity rate (Figure 2) is noisy but still we can appreciate that, depending on the length of the holiday, there is an increase either immediately after or even during the event. As an example, on June the 1st the government ordered the end of the mitigation measures, resulting in an initiation of the reopening of the economy in the country. Figure 2 shows that in Mexico City and in Queretaro (although less clearly here), the positivity rate increased in the days immediately following this date. The increase in positivity starts before June 1 probably because the government announced with several days of anticipation, the end of the partial lockdown. We must point out that changes in these indicators are not exclusive to the key dates we have selected (Table 2). However, significant changes are indeed associated to these dates as mentioned before.

The hospitalization rate follows a similar trend as the positivity rate. Figure 3 shows this for the same two cities. The epidemic in Queretaro started over a month later than in Mexico City, where the first case was recorded on February 22, 2020. The initial growth of the epidemic in Mexico was impacted by Holy Week vacations and the period of April 30 to May 10, which correspond to children's to mother's day, respectively. The impact of periods comprising September 12-16 (Independence Day), November 1-2 (Day of the Dead and All Saints Day) and November 9-20 ("Buen Fin", Mexican equivalent of Black Friday in the US), clearly show an increment in hospitalizations, particularly for the two age groups above 40 years (Figure 3).

In Queretaro (Figure 3B), this same trend is also apparent but admittedly less clear due to the smaller sample size associated with the smaller population size. An increase in hospitalization is also visible in the winter holiday period that spans from December 12, 2020 (Virgin of Guadalupe Day) to January 5, 2021 (Wise Men Day).

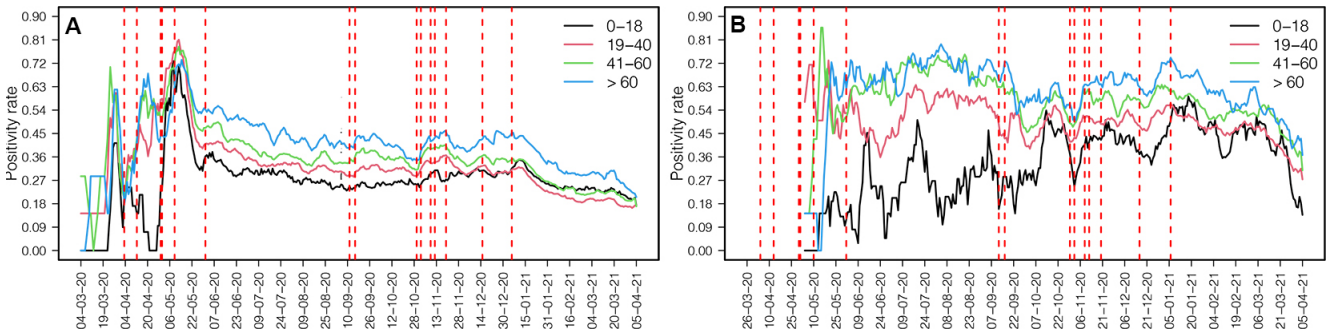

Figure 2. Positivity rate by age group for A) Mexico City and B) Queretaro. Dashed red vertical lines give the dates listed in Table 2.

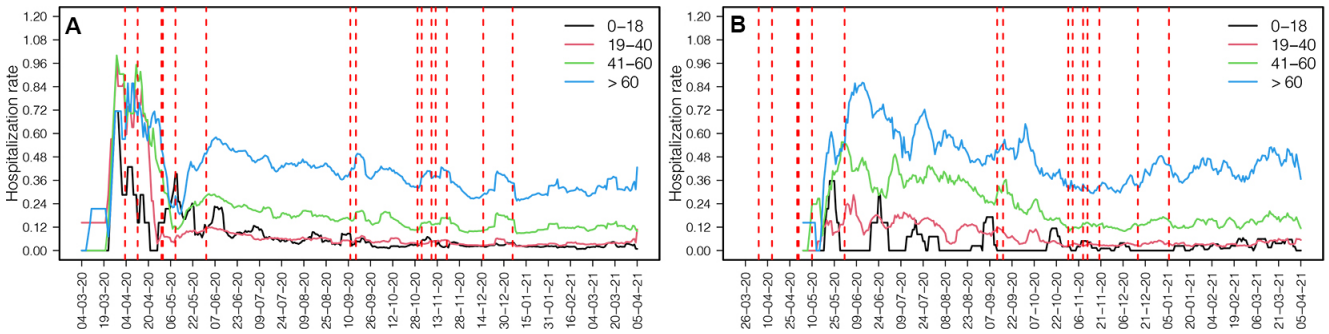

Figure 3. Hospitalization rate by age group for A) Mexico City and B) Queretaro. Dashed red vertical lines give the dates listed in Table 2.

The mortality rate shows even more clearly the impact of these same periods both in Mexico City and in Queretaro (Figure 4). Particularly evident is the jump in the mortality rate occurring four days after the Virgin of Guadalupe Day (December 12, 2020). This is a very important religious festivity in Mexico and is characterized by very large pilgrimage crowds walking into Mexico City and other places of sanctuary, interspersed with multitudinary masses and lively economic activity mainly in the informal sector.

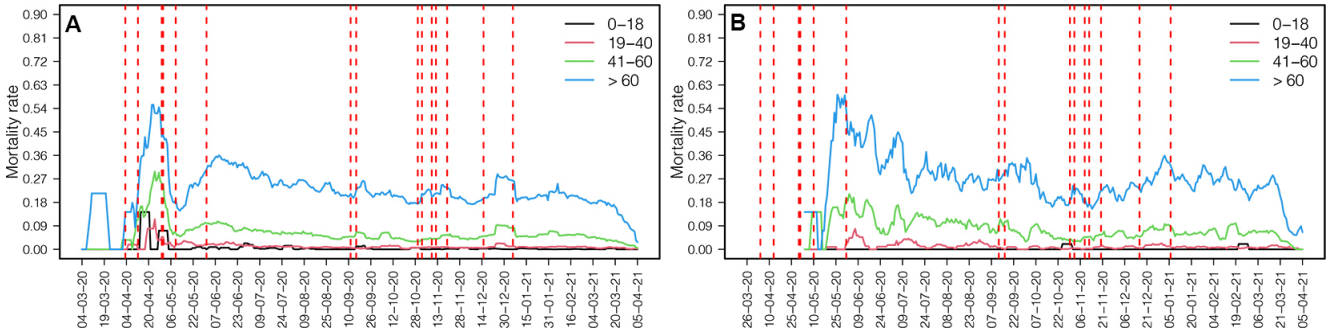

Figure 4. Mortality rate by age group for A) Mexico City and B) Queretaro. Dashed red vertical lines give the dates listed in Table 2.

As described above, we modify the estimate of the effective contact rate, $\beta(t)$, corresponding to February 2021, with the purpose of generating five scenarios supported on the effective contact rate patterns observed on the same dates in 2020. We show scenarios for the rest of 2021 and the beginning of 2022. However, the scenarios intend to be plausible only for the first semester of 2021 since, afterwards vaccination and environmental or climatic factors will start to play a role not incorporated into our basic model. The basic assumption is that the social behaviour and mobility patterns of the Mexican population will remain invariant on the key calendar dates identified in 2020. Some key dates may change, but these changes are predictable. For example, in June 6, 2021, there will be Federal Elections in the country, therefore, we can adjust our

estimates to accommodate this date. There were no Federal elections in 2020. Figure 5 shows some scenarios for Mexico City and the state of Queretaro, that we will use to illustrate the rationale behind our methodology. Details concerning the key calendar dates and the associated changes in contact rates can be found in Table 3. Results for other states are shown in the Supporting information A.

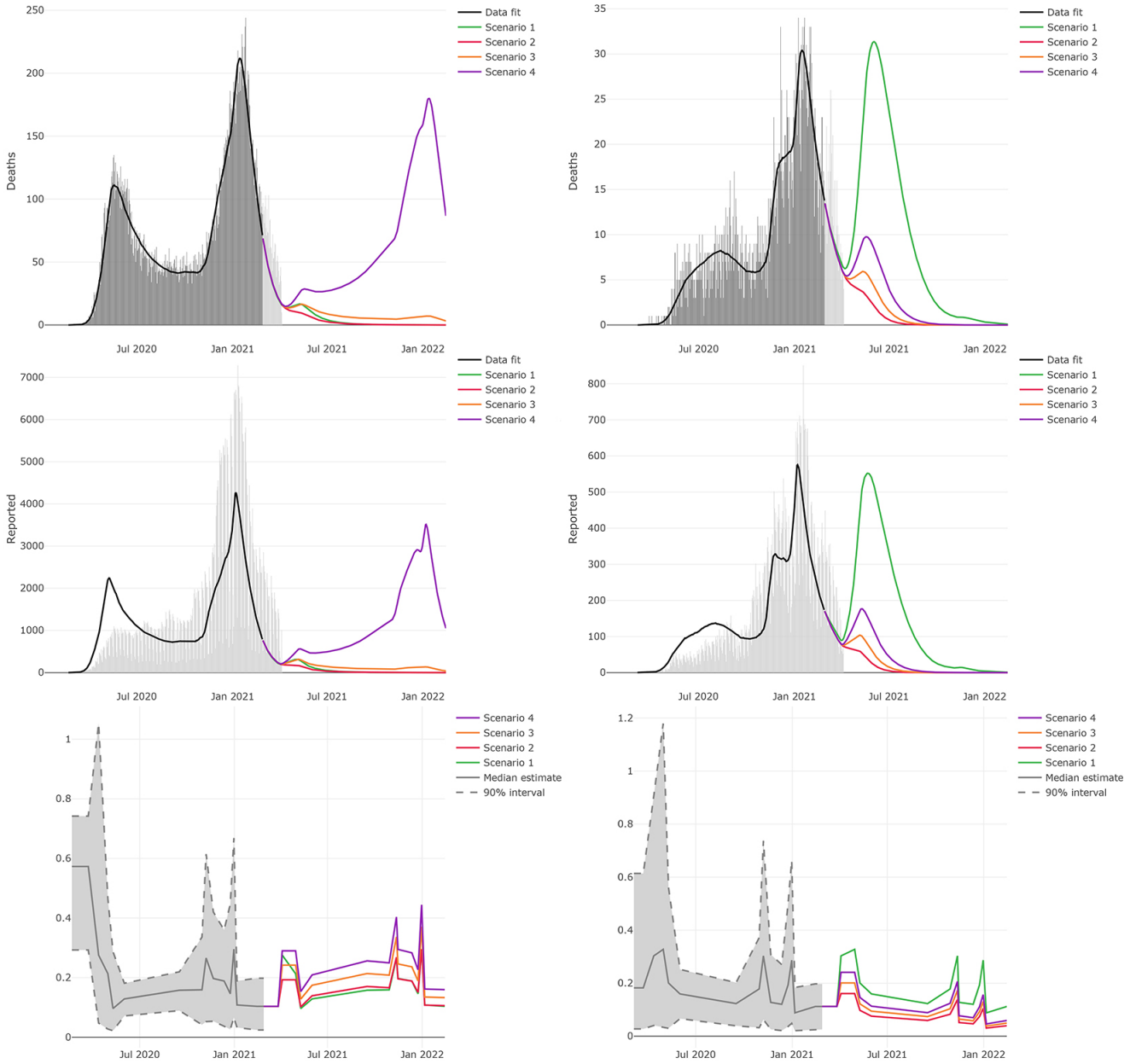

Figure 5. Projections of the evolution of the epidemic in 2021. Left side column shows results for Mexico City, while the right side column shows results for Queretaro state. Upper and mid panels show the daily number of deaths and reported COVID-19 cases from February 20, 2020 to February 14, 2022, respectively. The lower panel shows the effective contact rates for the same period. Dark gray bars show mortality data used in the estimation of the effective contact rate $\beta(t)$. Light gray bars show available data (daily deaths and reported cases) that was not used for estimation purposes. Black lines show the median posterior estimates for the effective contact rate, daily deaths and reported cases from February 20, 2020 to February 27, 2021. Colored lines show the predicted scenarios from February 28, 2021 to February 14, 2022, as were described in the main text.

Projected scenarios For all the scenarios below we assume demographic dynamics, i.e., birth and death (non-COVID-19 related) rates are positive, and a constant environment. The lack of seasonality is important and its effect should start acting on

the epidemic by October 2021. However, since we are using the pattern of changes of the observed contact rate for 2020, any seasonal effect with impact on last year's contact rate will be inherited to the projected contact rates of 2021. All scenarios include vaccination starting on January 1, 2021. For each case we assume that half of the population will be vaccinated during 2021 which might be an optimistic scenario given the present vaccination strategy and vaccine availability in the country. All other conditions (social, behavioral, etc) are assumed to remain constant in the time scale of the projections. Our projections thus provide a baseline for plausible epidemic trends.

To set the scenarios, two main ideas are used. The first is to calculate the changes (percentage) in the contact rate observed during 2020 between a pair of consecutive key dates for each state. These changes are used, in each case, as the basis to predict the evolution of the disease. We assume that similar relative changes will occur in 2021. The second idea is to use the effective reproduction number R_t (equation 2) to define the magnitude of the contact rate increase for the first key date in the prediction period which is April 4, 2021, the end of Easter celebration. The scenarios cover a plausible range of values for the contact rates according to the general trends observed in this parameters in 2020. Table 3 shows the specific values of the contact rates for Mexico City and Queretaro state.

1. The first scenario assumes that the contact rates will be the same as they were in 2020 for the same key dates in 2021. For the case of Mexico City and Queretaro, this will create a very small rise in cases and deaths, but after that the number of cases will decrease until the disease disappears at the end of 2021. In this scenario it is assumed that compliance of NPIs is as good as it was in 2020, which is, perhaps, not very realistic because, on the one hand, the economic situation of the country is causing exhaustion in the general population and on the other, the fact that vaccines are now available, even when applied at a very low rate, is creating a false sense of security in the general population. In Mexico City, this behaviour may cause a very small outbreak and the epidemic would end or enter an inter-epidemic period during 2021. However, in the case of Queretaro state, this same situation wold lead to a new outbreak of the same magnitude of the one observed in December 2020.

2. Scenario 2 sets the contact rate to $\beta(t) = 0.193$ from April 4 to April 30, 2021. Then, the contact rate will change as it did in 2020. This value is obtained from setting the effective reproduction number $R_t = 1$ in equation 2 and using $S(t)/N$, the proportion of susceptible individuals, estimated on February 27, 2021, the start of the prediction period. Solving the equation of $R_t = 1$ for $\beta(t)$, at a given t, gives a threshold value for $\beta(t)$: any increase above it will generate an outbreak. In this scenario for Mexico City and Queretaro, there will be no more outbreaks and the disease will end even sooner than in Scenario 1.

3. Set up is similar to Scenario 2, but with a contact rate $\beta(t)$, $t fixed$, such that it corresponds to effective reproduction numbers of 1.25 (scenario 3) and 1.5 (scenario 4), respectively (Table 3). Figure 1, shows this case to project larger outbreaks than in Scenario 2. In the case of Mexico City, only scenario 4 creates an important outbreak which would occur during December 2021, comparable to the on seen in the same period of 2020. For Queretaro state, $R_t = 1.25$ o $R_t = 1.25$ will cause small outbreaks but the epidemic still would end or enter an inter-epidemic period during 2021.

3 Discussion

Using holidays and other civic and religious days as change points in the contact rate and the magnitude of R_t, we explore the expected COVID-19 dynamics in situations where strict lockdowns and other non-pharmaceutical interventions are not mandatory such as is the case in Mexico. In several states, the effective contact rate significantly decreased upon the enforcing of governmental mitigation measures on March 23-30, 2020. However, this was not a general trend. In other states (Guanajuato, Mexico, Jalisco (see SI)), the effective contact rate increased right after this date, perhaps indicating a slow adhesion to the mitigation policies. Interestingly in all the states analyzed, the effective contact rate sharply decreased during the month of May. In some states the contact rate remained practically stable until September when a steady increase is observed (Figure 7); others like Guanajuato show a decreasing trend until September. A common feature to all states is that the effective contact rate had two fast periods of growth around November 9-20, 2020 and on the winter holidays (December 12 to December 31, 2020). After this latter period, the effective contact rate decreased below the average level shown before, namely, the November peak. This situation may change with the recently passed Easter Holidays period, from March 19 to April 4, 2021, that, at the time of writing this work it is only beginning to show in the data. The magnitude and changes of the contact rates estimated using key dates on 2020 together with the magnitude of R_t constitute two indicators that give a plausible approximation of what could happened during 2021. Our basic assumption is that these dates will constitute superspreading events. Their effect on the epidemic curve depends, however, on R_t being above or below the threshold $R_t \approx 1$. Our idea is simple but we think that key calendar dates provide a practical basis to anticipate changes in the transmission of the disease that may be able provide mid to long term outbreak risk projections. There is no doubt that superspreading events will still be important for the COVID-19 pandemic however, it is not clear what the magnitude of their impact will be. We believe that the methodology presented here

is a first step towards the solution of this problem. Our study highlights the importance for Mexico and countries of similar characteristics, of enforcing the application of NPIs. In all our scenarios, the reduction in $\beta(t)$ can be achieved only through them, given the low vaccination rate that exists at the end of March. This is not surprising, a recent study[28] has shown that even with high efficacy vaccines and a greater coverage than current one in Mexico, vaccination alone cannot achieve a significant reduction in cases if NPIs are relaxed too quickly.

Finally, the objective of this modelling effort is to generate a range of scenarios using the known history of the epidemic, that could be useful to evaluate its possible evolution and its likely impact on incidence and mortality.

4 Methods

A compartmental model (1) is used to estimate the effective contact rate. The model considers three classes of infected individuals: Asymptomatic (I), Symptomatic (Y) and Reported (T). Once reported, infected individuals are effectively isolated and no longer participants in the transmission process. Figure 6 shows the corresponding model diagram. See the SI for the model equations. It is important to stress that the model allows Susceptible (S) individuals to be Vaccinated (V) with an effective vaccination rate ψ. During 2020, there was no vaccination and ψ is set to zero for that period. Vaccine roll-out in Mexico has been slow. By the end of March, 2021, only about 6% of the population had been vaccinated. For this reason, the model assumption that the vaccine is perfect (100% efficacy) is not crucial. Vital dynamics are also included since this work is aimed to produce mid to long-term scenarios for the epidemic. Table 1 shows a description of all the model's parameters and their values.

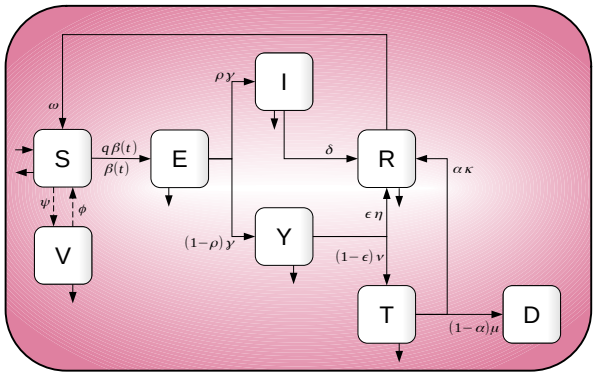

Figure 6. Diagram for the mathematical model. There are three types of infectious individuals: asymptomatic, symptomatic and reported. Reported cases do not play a role in transmission.

Parameter	Description	Value	Units
ψ	Vaccination rate	9.6×10^{-4}	days^{-1}
ϕ	Vaccination immunity rate	0.024	days^{-1}
ω	Natural immunity rate	0.024	days^{-1}
γ	Incubation rate	1.37	days^{-1}
δ	Asymptomatic recovery rate	1.0	days^{-1}
η	Symptomatic recovery rate	0.5	days^{-1}
κ	Reported recovery rate	0.5	days^{-1}
ν	Screening rate	1.75	days^{-1}
μ	Death rate	0.7	days^{-1}
ρ	Proportion of asymptomatic	0.2	%
ε	Proportion of symptomatic recovered	0.85	%
α	Proportion of reported recovered	0.94	%
q	Asymptomatic infectiousness reduction	0.45	%

Table 1. Description of the parameters of the Kermack-McKendrick model (see[12,14] for sources).

The effective contact rate $\beta(t)$ is estimated using the reported COVID-19 mortality from the start of the pandemic (late February to early March, depending on the specific location) until February 27, 2021. We use mortality instead of reported cases

since it is more accurately recorded. We allow $\beta(t)$ to linearly change between pre-defined dates associated with mitigation measures implemented by the government, and holidays of civic or religious nature. We have found that it is enough to estimate the effective contact rate in all of these dates (or some of them, depending on the specific location) to get a good fit for the number of daily deaths. The overall perspective of the epidemic management in Mexico can be fairly summarized by saying that a) all non-pharmaceutical interventions are, to date, non-mandatory but voluntary except some policies for closure or reopening of business, and b) the population increased mobility and social activities during certain periods strongly associated to our key calendar dates. This behavior can be seen in Figure 7, where the estimated effective contact rates of several Mexican states is presented. Regardless of the differences between states, especially at the beginning of the epidemic, it is clear that, for example, after the celebration of the Day of the Dead and Christmas, there is an increase in $\beta(t)$. Changes in the effective

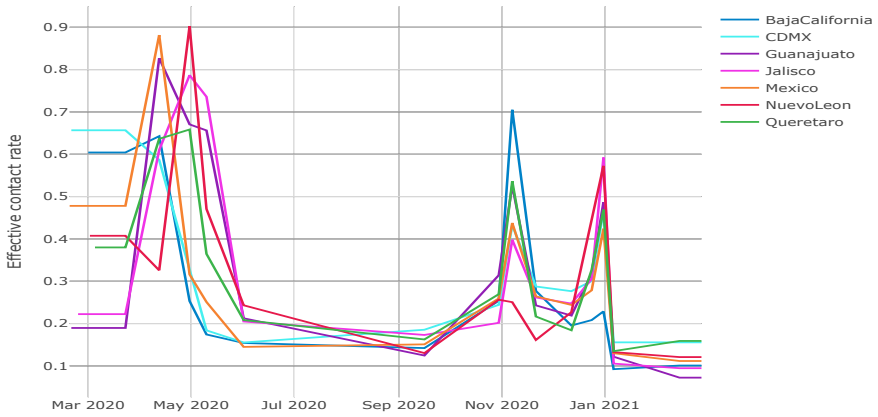

Figure 7. Effective contact rates estimated for several Mexican states from the beginning of the pandemic (late February to early March) in 2020 until February 27, 2021. The first peak is associated with the initial epidemic outbreak characterized by high reproductive numbers and large uncertainties. A roughly downward trend in the effective contact rate can be seen between early 2020 to January 2021. The two flare-ups in the effective contact rate at the end of the year 2020, are associated with the Buen Fin and winter holidays. Horizontal axis is calendar days and the vertical axis is the effective contact rate in the number of contacts per capita per unit time.

contact rate during the year 2020 are used as a basic guideline for the year 2021 scenarios presented in this work. A complete list of the dates and the estimated values of the contact rate can be found in Table 2.

Many methodologies exist to deal with ways of estimating the effective reproduction number (e.g.,[29–34]). Although we use statistical methods to estimate the effective contact rate, our approach for making projections relies of the past evidence of changes of the contact rate. It is essentially qualitative in nature. Our statistical methods can be found in the SI.

Acknowledgements

JXVH and MSC acknowledge support from UNAM PAPIIT DGAPA IN115720 and IV100220 grants. CERHV and RHM are grateful for the support of UNAM PAPIIT grant IG100221.

Author contributions statement

MSC, MAAZ and JXVH conceived the idea; CERHV and RHM analyzed data, MSC and MAAZ performed the model projections and parameter estimation, all authors discussed, reviewed and wrote the paper.

Additional information

Competing interests The authors report no competing interests.

References

1. Dehning, J. *et al.* Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Sci. (New York, N.Y.)* **9789**, 1–15, DOI: 10.1126/science.abb9789 (2020).

Date	Description	Contact rate
2020-02-20	First COVID-19 case (symptoms onset)	0.572
2020-03-23	Mitigation measures start	0.572
2020-04-12	Easter celebrations	0.290
2020-04-30	Children's day	0.198
2020-05-10	Mother's day	0.095
2020-06-01	End of mitigation measures	0.129
2020-09-16	Independence day	0.157
2020-10-30	Day of the Dead	0.163
2020-11-07	Buen Fin starts	0.268
2020-11-20	Buen Fin ends	0.196
2020-12-12	Day of the Virgin of Guadalupe	0.190
2020-12-24	Christmas	0.145
2020-12-31	New year's eve	0.282
2021-01-06	Wise men day	0.110
2021-02-14	Valentine's day	0.102
2021-02-27	Last day of available data	0.102

Table 2. Effective contact rates posterior median estimates for each of the key dates previously identified for Mexico City. Each state analyzed in this work starts its epidemic curve at a different date between late February and early March; all the other key dates are the same.

2. Li, Y. *et al.* The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. *The Lancet Infect. Dis.* **21**, 193–202, DOI: 10.1016/S1473-3099(20)30785-4 (2021).

3. May, B. Research Briefing — Global World: GDP revisions. Tech. Rep., Oxford Economics, Oxford (2021).

4. Askitas, N., Tatsiramos, K. & Verheyden, B. Estimating worldwide effects of non-pharmaceutical interventions on COVID-19 incidence and population mobility patterns using a multiple-event study. *Sci. Reports* **11**, 1–13, DOI: 10.1038/s41598-021-81442-x (2021).

5. INSP. Recomendaciones. In *Reflexiones sobre la respuesta de Mexico ante la pandemia de COVID19 y sugerencias para enfrentar los próximos retos.* Instituto Nacional de Salud Pública (Instituto Nacional de Salud Pública, Mexico, Mexico City, 2020).

6. Allain-Dupre, D., Chatry, I., Kornprobst, A. & Michalun, M.-V. The territorial impact of COVID-19: Managing the crisis across levels of government. Tech. Rep., OECD (2020).

7. Agusto, F. B. *et al.* To isolate or not to isolate: The impact of changing behavior on COVID-19 transmission. *medRxiv* DOI: 10.1101/2020.08.30.20184804 (2020).

8. Bubar, K. M. *et al.* Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* **371**, 916–921, DOI: 10.1101/2020.09.08.20190629 (2020).

9. Kermack, W. & McKendrick, A. A Contribution to the Mathematical Theory of Epidmics. *Proc. Royal Soc. A* **115**, DOI: 10.2307/2304150 (1927).

10. Subramanian, R., He, Q. & Pascual, M. Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology and testing capacity. *PNAS* **118**, e2019716118, DOI: 10.1073/pnas.2019716118 (2020).

11. Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan , China : a modelling study. *The Lancet* **S0140-6736**, 30260–9, DOI: 10.1016/S0140-6736(20)30260-9 (2020).

12. Acuña-Zegarra, M. A., Santana-Cibrian, M. & Velasco-Hernandez, J. X. Modeling behavioral change and COVID-19 containment in Mexico: A trade-off between lockdown and compliance. *Math. biosciences* **10**, 108370, DOI: 10.1016/j.mbs.2020.108370 (2020).

13. Mena, R. H. R. *et al.* Using the posterior predictive distribution to analyse epidemic models: COVID-19 in Mexico City. *Phys. biology* 1–11, DOI: 10.1088/1478-3975/abb115 (2020). 2005.02294.

	Mexico City				Queretaro state			
Dates	Scenario1	Scenario2	Scenario3	Scenario4	Scenario1	Scenario2	Scenario3	Scenario4
2021-03-27	0.104	0.104	0.104	0.104	0.112	0.112	0.112	0.112
2021-04-04	0.276	0.193	0.242	0.290	0.303	0.161	0.201	0.241
2021-04-30	0.213	0.193	0.242	0.290	0.328	0.161	0.201	0.241
2021-05-10	0.097	0.102	0.128	0.154	0.201	0.098	0.122	0.146
2021-06-01	0.129	0.139	0.175	0.209	0.160	0.076	0.094	0.113
2021-09-16	0.158	0.171	0.214	0.256	0.123	0.059	0.074	0.089
2021-10-29	0.159	0.166	0.208	0.250	0.178	0.083	0.104	0.124
2021-11-12	0.266	0.269	0.337	0.404	0.303	0.138	0.172	0.207
2021-11-15	0.197	0.196	0.246	0.295	0.128	0.052	0.064	0.077
2021-12-12	0.189	0.189	0.236	0.283	0.121	0.047	0.058	0.070
2021-12-24	0.146	0.151	0.189	0.227	0.193	0.074	0.092	0.111
2021-12-31	0.296	0.296	0.371	0.444	0.287	0.105	0.131	0.157
2022-01-06	0.108	0.108	0.135	0.162	0.088	0.031	0.039	0.046
2022-02-14	0.104	0.106	0.133	0.160	0.112	0.040	0.049	0.059

Table 3. Effective contact rates used for prediction of reported cases and deaths in Mexico City and the state of Queretaro during 2021 and the beginning of 2022 for each of the five scenarios described in the text.

14. Santana-Cibrian, M., A. Acuña-Zegarra, M. & X. Velasco-Hernandez, J. Lifting mobility restrictions and the effect of superspreading events on the short-term dynamics of COVID-19. *Math. Biosci. Eng.* **17**, 6240–6258, DOI: 10.3934/mbe.2020330 (2020).

15. Komarova, N. L., Schang, L. M. & Wodarz, D. Patterns of the COVID-19 pandemic spread around the world: Exponential versus power laws: Patterns of the COVID-19 pandemic spread around the world: Exponential versus power laws. *J. Royal Soc. Interface* **17**, DOI: 10.1098/rsif.2020.0518 (2020).

16. Singer, H. M. The COVID-19 pandemic: growth patterns, power law scaling, and saturation. *Phys. biology* **17**, 055001, DOI: 10.1088/1478-3975/ab9bf5 (2020). 2004.03859.

17. Thurner, S., Klimek, P. & Hanel, R. A network-based explanation of why most COVID-19 infection curves are linear. *Proc. Natl. Acad. Sci. United States Am.* **1**, 22684–22689, DOI: 10.1073/pnas.2010398117 (2020).

18. Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359, DOI: 10.1038/nature04153 (2005).

19. Adam, D. C. *et al.* Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nat. Medicine* **26**, 1714–1719, DOI: 10.1038/s41591-020-1092-0 (2020).

20. Althouse, B. M. *et al.* Superspreading events in the transmission dynamics of SARS-CoV-2: Opportunities for interventions and control. *PLoS Biol.* **18**, 1–13, DOI: 10.1371/journal.pbio.3000897 (2020).

21. Chang, S. *et al.* Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87, DOI: 10.1038/s41586-020-2923-3 (2021).

22. Liu, Y., Eggo, R. M. & Kucharski, A. J. Secondary attack rate and superspreading events for SARS-CoV-2. *The Lancet* **395**, e47, DOI: 10.1016/S0140-6736(20)30462-1 (2020).

23. Popa, A. *et al.* Genomic epidemiology of superspreading events in Austria reveals mutational dynamics and transmission properties of SARS-CoV-2. *Sci. Transl. Medicine* **12**, DOI: 10.1126/scitranslmed.abe2555 (2020).

24. Rader, B. *et al.* Crowding and the shape of COVID-19 epidemics. *Nat. Medicine* **0.1038/s41**, DOI: 10.1038/s41591-020-1104-0 (2020).

25. Sun, K. *et al.* Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science* **371**, eabe2424, DOI: 10.1126/science.abe2424 (2021).

26. Wong, F. & Collins, J. J. Evidence that coronavirus superspreading is fat-tailed. *Proc. Natl. Acad. Sci.* 202018490, DOI: 10.1073/pnas.2018490117 (2020).

27. Zhang, Y., Li, Y., Wang, L., Li, M. & Zhou, X. Evaluating transmission heterogeneity and super-spreading event of COVID-19 in a metropolis of China. *Int. J. Environ. Res. Public Heal.* **17**, DOI: 10.3390/ijerph17103705 (2020).

28. Moore, S., Hill, E. M., Tildesley, M. J., Dyson, L. & Keeling, M. J. Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *The Lancet Infect. Dis.* DOI: 10.1016/s1473-3099(21)00143-2 (2021).

29. Collective. Reproduction number (r) and growth rate (r) of the covid-19 epidemic in the uk: methods of estimation, data sources, causes of heterogeneity, and use as a guide in policy formulation. *The Royal Soc.* (2020).

30. Ives, A. R. & Bozzuto, C. Estimating and explaining the spread of covid-19 at the county level in the usa. *Commun. Biol.* **4**, DOI: 10.1038/s42003-020-01609-6 (2021).

31. Britton, T. Basic estimation-prediction techniques for covid-19, and a prediction for stockholm. *medRxiv* DOI: 10.1101/2020.04.15.20066050 (2020).

32. Kiamari, M. *et al.* Covid-19 risk estimation using a time-varying sir-model. *medRxiv* (2020).

33. Wan, H., Cui, J. A. & Yang, G. J. Risk estimation and prediction of the transmission of coronavirus disease-2019 (covid-19) in the mainland of china excluding hubei province. *Infect. Dis. Poverty* **9**, DOI: 10.1186/s40249-020-00683-6 (2020).

34. Maruotti, A., Ciccozzi, M. & Divino, F. On the misuse of the reproduction number in the covid-19 surveillance system in italy. *J. Med. Virol.* DOI: 10.1002/jmv.26881 (2021).

35. Christen, J. A. & Fox, C. A general purpose sampling algorithm for continuous distributions (the t-walk). *Bayesian Analysis* **5**, 263–281, DOI: 10.1214/10-BA603 (2010).

A Supporting information

Mathematical model. The mathematical model represented in Figure 6 is the following. Parameters and a general description are given in Section 4:

$$
\begin{aligned}
S' &= \omega R + \phi V - \psi S - \beta(t)(Y + qI)\frac{S}{N} + \chi M - \chi S, \\
V' &= \psi S - \phi V - \chi V, \\
E' &= \beta(t)(Y + qI)\frac{S}{N} - \gamma E - \chi E, \\
I' &= \rho\gamma E - \delta I - \chi I, \\
Y' &= (1-\rho)\gamma E - \varepsilon\eta Y - (1-\varepsilon)\nu Y - \chi Y, \\
T' &= (1-\varepsilon)\nu Y - \alpha\kappa T - (1-\alpha)\mu T - \chi T, \\
R' &= \delta I + \varepsilon\eta Y + \alpha\kappa T - \omega R - \chi R, \\
D' &= (1-\alpha)\mu T.
\end{aligned} \tag{1}
$$

with $N = S + V + E + I + Y + R$ and $M = N + T$. The basic reproduction number given by

$$
\mathscr{R}_0 = \beta_0\Big(\frac{q\rho\gamma}{(\delta+\chi)(\gamma+\chi)} + \frac{(1-\rho)\gamma}{(\varepsilon\eta + (1-\varepsilon)\nu + \chi)(\gamma+\chi)}\Big).
$$

The parameter β_0 is the constant contact rate of the first month of the epidemic. After this period of time $\beta(t)$ is the time-varying effective contact rate estimated in this work. Let

$$
\mathscr{R}(t) = \beta(t)\Big(\frac{q\rho\gamma}{(\delta+\chi)(\gamma+\chi)} + \frac{(1-\rho)\gamma}{(\varepsilon\eta + (1-\varepsilon)\nu + \chi)(\gamma+\chi)}\Big),
$$

then the effective reproduction number is

$$
R_t = \mathscr{R}(t)S(t)/N(t). \tag{2}
$$

Statistical inference In order to create predictions for the evolution of the COVID-19 pandemic during 2021, we focus on the estimation of the time dependent contact rate $\beta(t)$ during 2020 and the start of 2021 as was described in Section 2. To simplify the estimation process, $\beta(t)$ is assumed to be a piecewise linear function that only changes at preset times $t_1, t_2, \ldots, t_k$, and it is constant before t_1 and after t_k. Then, instead of estimating a continuous function, we only need to estimate the values of the effective contact rate $b_1, b_2, \ldots, b_k$ at $t_1, t_2, \ldots, t_k$.

The dates where the effective contact rate changes are described in Table 2. Those dates are converted into a numeric scale to get $t_1, t_2, \ldots, t_k$ simply by calculating the number of days from the first reported case (by symptoms onset). For each Mexican state, the starting date for the analysis is different.

The initial number of exposed (E_0), asymptomatic (I_0) and symptomatic (Y_0) individuals will also be estimated as these are important unknown quantities. Let $\theta = (E_0, I_0, Y_0, b_1, b_2, \ldots, b_k)$ the vector of parameters that will be estimated using a Bayesian inference approach. All the other parameters needed to solve model (1) are fixed and their values can be found in Table 1.

Let X_j be the random variable that counts the number of daily COVID-19 deaths at time t_j, for $j = 1, 2, ...n$, where t_j represent the number of days since the first reported case by symptoms onset. We assume that the probability distribution of X_j, conditional on the vector of parameters θ, is a Poisson distribution such that $E[Y_j] = mu(t_j|\theta) = D(t_j|\theta) - D(t_{j-1}|\theta)$, with $D(t|\theta)$ being the cumulative number of deaths according to model (1). Assuming that variables $X_1, X_2, \ldots, X_n$ are conditionally independent, then the likelihood function is given by

$$\pi(x_1, \ldots, x_n|\theta) = \prod_{j=1}^{n} \frac{\mu(t_j|\theta)^{x_j} \exp\{-\mu(t_j|\theta)\}}{x_j!}. \tag{3}$$

The joint prior distribution for vector θ is a product of independent Uniform distributions. For all the initial conditions E_0, I_0, Y_0, the prior distribution is Uniform(0,20) and for all the contact rates $b_1, \ldots, b_k$ the prior is Uniform(0,5). Then

$$\pi(\theta) = \pi(E_0)\pi(I_0)\pi(Y_0)\prod_{i=1}^{k}\pi(b_i) \propto 1,$$

The posterior distribution of the parameters of interest is

$$\pi(\theta|y_1, \ldots, y_n) \propto \pi(y_1, \ldots, y_n|\theta)\pi(\theta),$$

and it does not have an analytical form since the likelihood function depends on the numerical solution of the ODE system 1. We analyze the posterior distribution using an MCMC algorithm called *t-walk*[35]. For each state, eight chains of 500,000 iterations are run, from which 25,000 are discarded as burn in. At the end, only 5000 iterations are retained to create the estimations presented in this work.

Projections for other Mexican states. In this section we present several other projections for other Mexican states following the same methodology applied for Mexico City and Queretaro discussed in the main text.

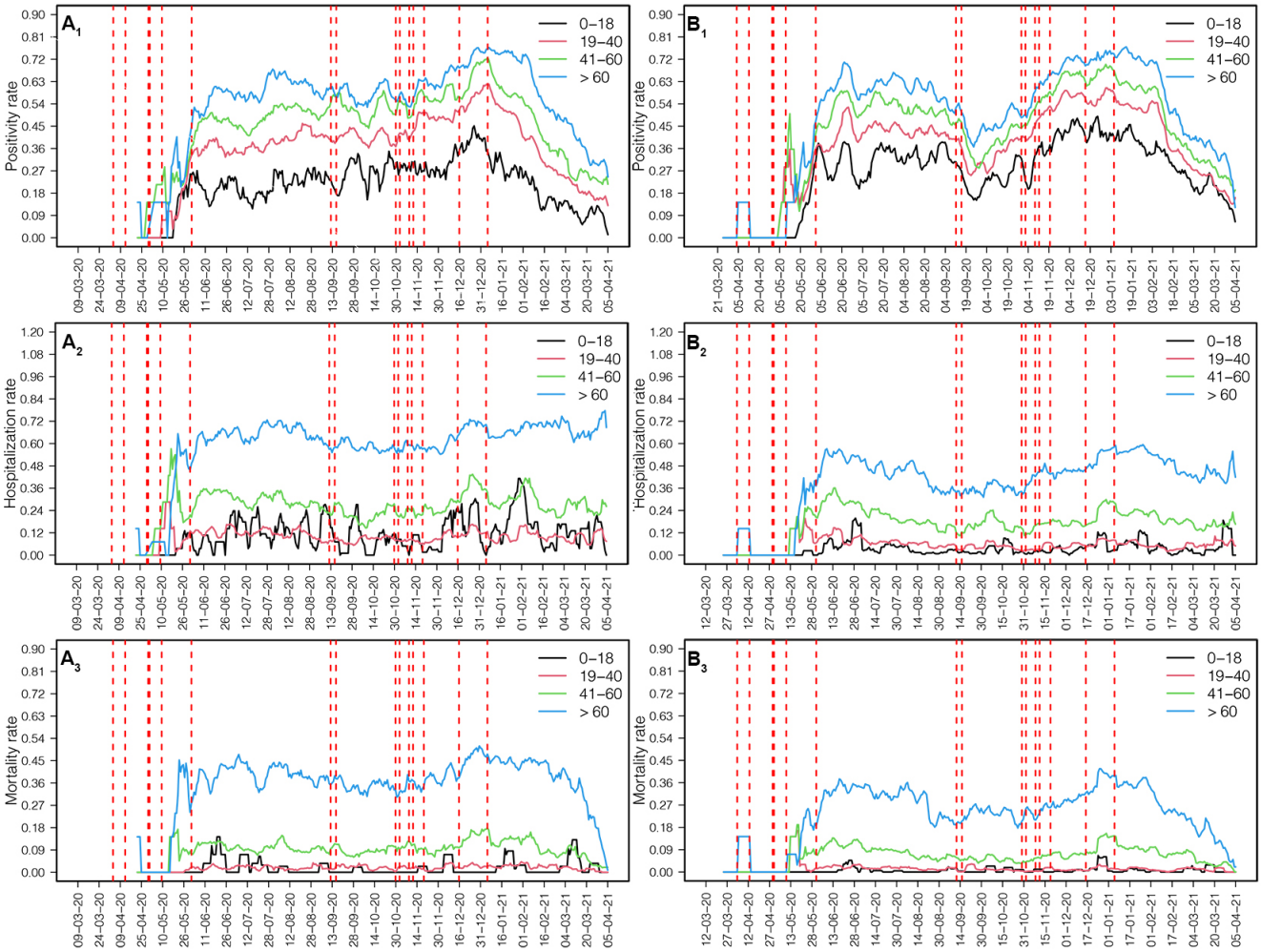

Figure 8. Changes in rates for different states. A_i represents Jalisco rates. B_i illustrates Guanajuato rates. Subindex i denotes: positivity (1), hospitalization (2) and mortality (3).

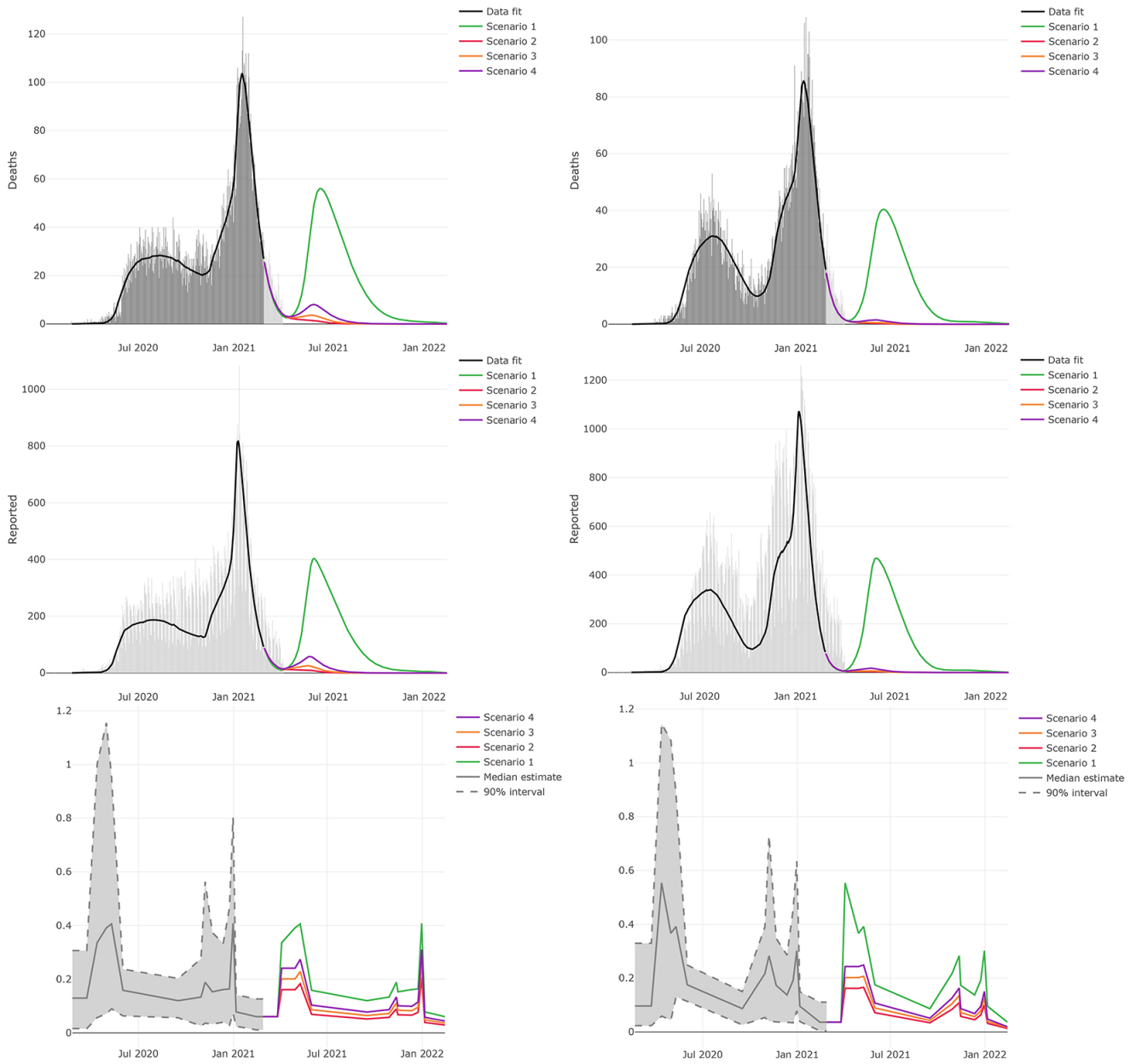

Figure 9. Projections of the evolution of the epidemic in 2021. Left side column shows results for Jalisco state, while the right side column shows results for Guanajuato state. Upper and mid panels show the daily number of deaths and reported COVID-19 cases from February 20, 2020 to February 14, 2022, respectively. The lower panel shows the effective contact rates for the same period. Dark gray bars show mortality data used in the estimation of the effective contact rate $\beta(t)$. Light gray bars show available data (daily deaths and reported cases) that was not used for estimation purposes. Black lines show the median posterior estimates for the effective contact rate, daily deaths and reported cases from February 20, 2020 to February 27, 2021. Colored lines show the predicted scenarios from February 28, 2021 to February 14, 2022.

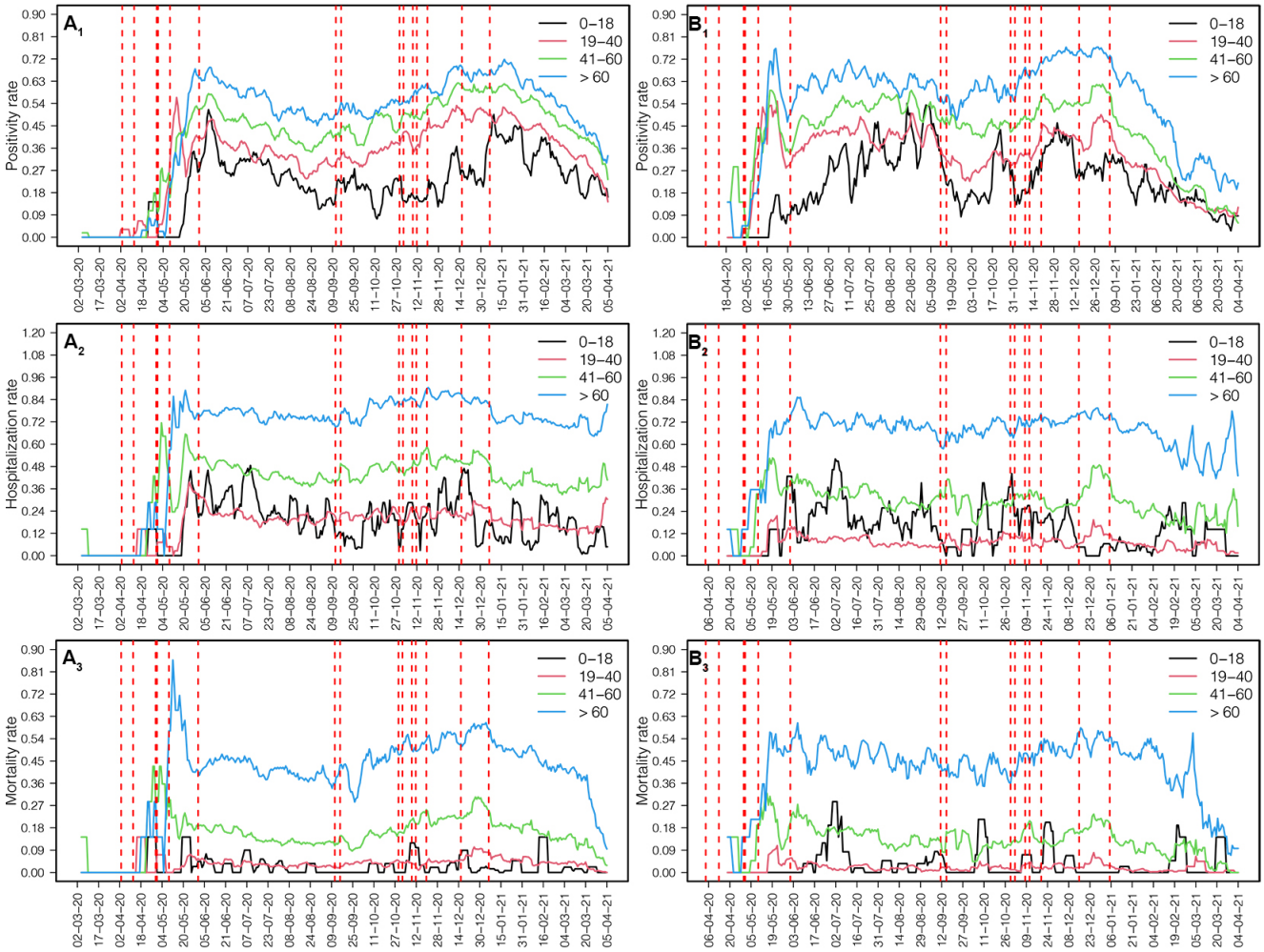

Figure 10. Changes in rates for different states. A_i represents State of Mexico rates. B_i illustrates Baja California rates. Subindex i denotes: positivity (1), hospitalization (2) and mortality (3).

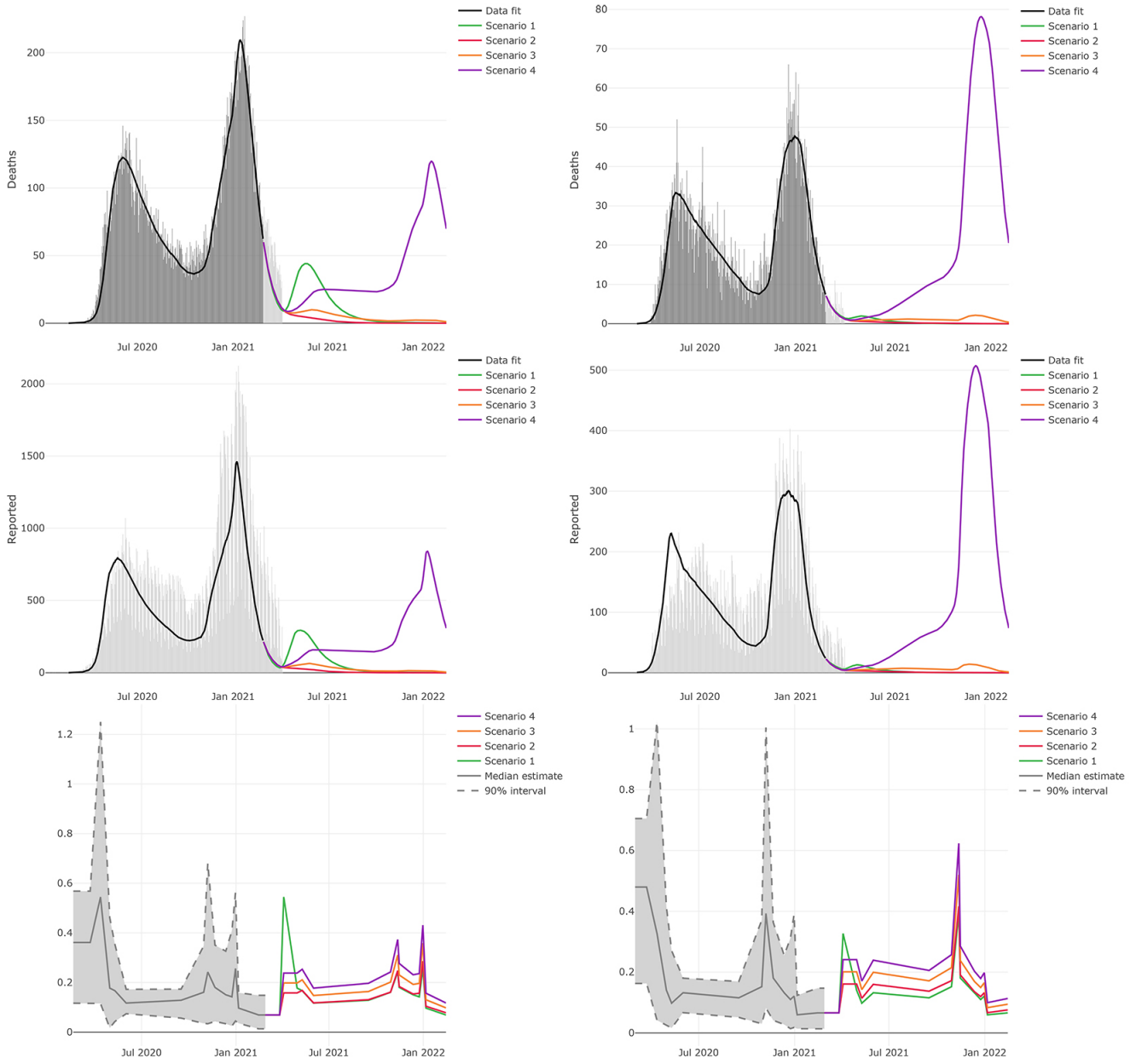

Figure 11. Projections of the evolution of the epidemic in 2021. Left side column shows results for Mexico state, while the right side column shows results for Baja California state. Upper and mid panels show the daily number of deaths and reported COVID-19 cases from February 19, 2020 to February 14, 2022, respectively. The lower panel shows the effective contact rates for the same period. Dark gray bars show mortality data used in the estimation of the effective contact rate $\beta(t)$. Light gray bars show available data (daily deaths and reported cases) that was not used for estimation purposes. Black lines show the median posterior estimates for the effective contact rate, daily deaths and reported cases from February 19, 2020 to February 27, 2021. Colored lines show the predicted scenarios from February 28, 2021 to February 14, 2022.

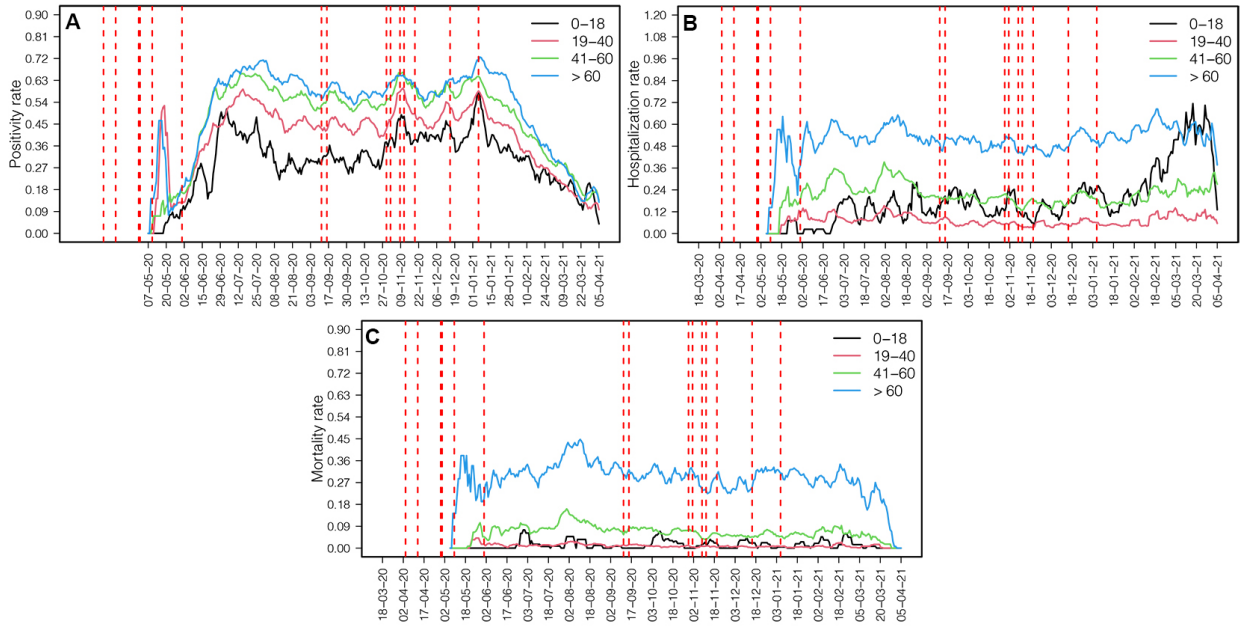

Figure 12. Changes in rates for Nuevo Leon state. A) positivity rate. B) hospitalization rate. C) mortality rate.

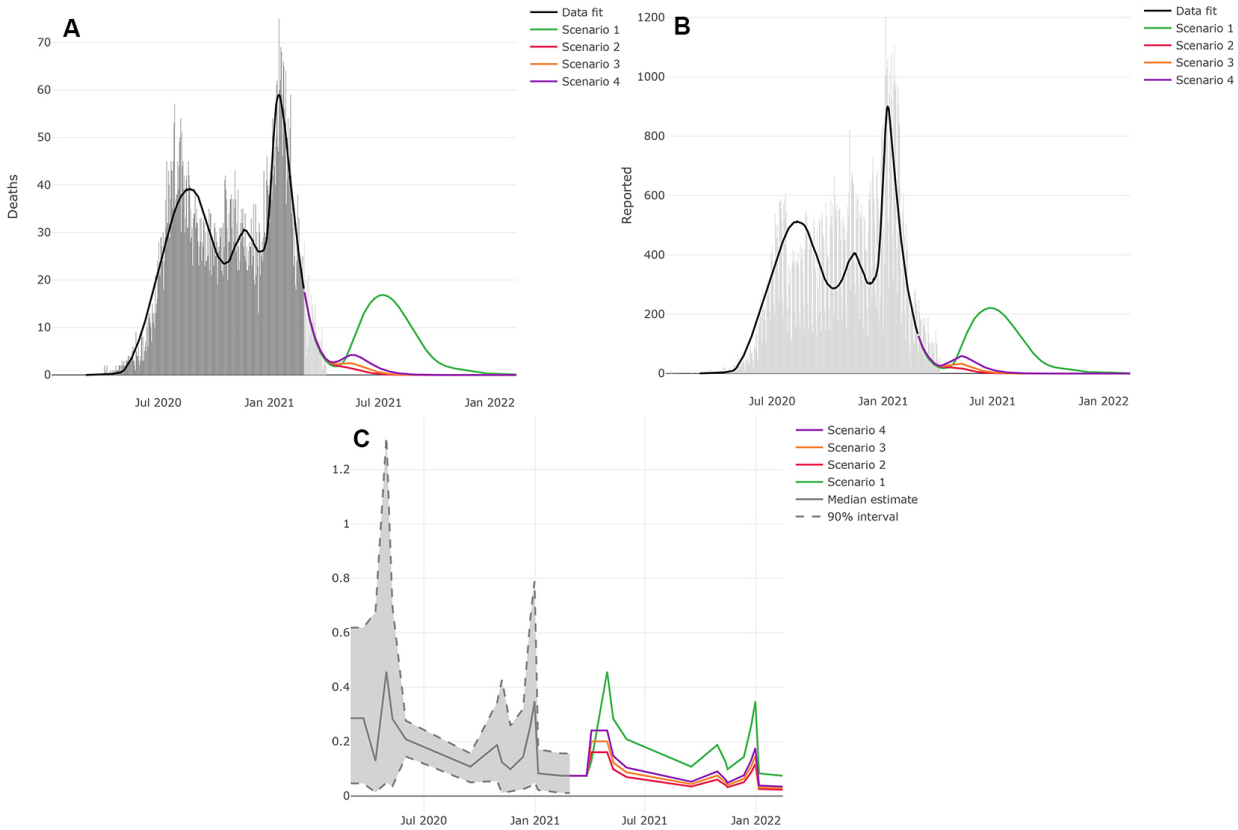

Figure 13. Scenarios for Nuevo Leon state from March 2, 2020 to February 14, 2022. Panel A shows daily number of deaths by COVID-19. Panel B illustrates daily reported COVID-19 cases. Panel C shows the contact rate $(\beta(t))$. Black lines show estimates from March 2, 2020, to February 27, 2021. Colored lines correspond to the four scenarios described in the main text. Dark gray bars show mortality data used in the estimation of the effective contact rate $\beta(t)$. Light gray bars show available data (daily deaths and reported cases) that was not used for estimation purposes.