

Forecasting hospital demand in metropolitan areas during the current COVID-19 pandemic and estimates of lockdown-induced 2nd waves

Marcos A. Capistrán^{1*}, Antonio Capella², and J. Andrés Christen¹.

¹ Centro de Investigación en Matemáticas, CIMAT-CONACYT,

² Instituto de Matemáticas, UNAM, Circuito Exterior, CU, CDMX, Mexico.

These authors contributed equally to this work.

* marcos@cimat.mx

Summary

We present a forecasting model aim to predict hospital occupancy in metropolitan areas during the current COVID-19 pandemic. Our SEIRD type model features asymptomatic and symptomatic infections with detailed hospital dynamics. We model explicitly branching probabilities and non-exponential residence times in each latent and infected compartments. Using both hospital admittance confirmed cases and deaths, we infer the contact rate and the initial conditions of the dynamical system, considering breakpoints to model lockdown interventions and the increase in effective population size due to lockdown relaxation. The latter features let us model lockdown-induced 2nd waves. Our Bayesian approach allows us to produce timely probabilistic forecasts of hospital demand. We have applied the model to analyze more than 70 metropolitan areas and 32 states in Mexico.

Introduction

The ongoing COVID-19 pandemic has posed a major challenge to public health systems in many countries with the imminent risk of saturated hospitals and patients not receiving proper medical care. Although the scientific community and public health authorities had insight regarding the risks and preparedness measures required at the onset of a zoonotic pandemic, our knowledge of the fatality and spread rates of COVID-19 remains limited [1–4]. In terms of disease handling, two leading issues determining the current situation are the lack of pharmaceutical treatment and our inability to estimate the extent of the asymptomatic infection of COVID-19 [5–7].

Under current circumstances, control measures reduce new infections by limiting the number of contacts through mitigation and suppression [1]. Mitigation includes social distancing, testing, tracing, and isolating infected individuals, while suppression imposes temporary cancellation of non-essential activities. Mitigation and suppression pose a burden on the economy, affecting more individuals living in low-income conditions, challenging the population’s capacity to comply with control measures. As lockdown measures are eased, more people become in contact with the outbreak, and there is a risk of induced 2nd waves that may increase healthcare system pressure.

Broadly speaking, data-driven epidemiological models are built out of the necessity of making forecasts. There are many lessons learned on emergency preparedness and epidemic surveillance from previous pandemic events: AH1N1 influenza [8], MERS [9],

SARS [10], Zika [11], Ebola [12], etcetera. However, surveillance data during a pandemic event often suffer from serious deficiencies such as incompleteness and backlogs. Another critical issue is the design of data acquisition, taking into account geographical granularity [13]. Epidemic surveillance of COVID-19 is no different since there is an unknown percentage of asymptomatic infections, and susceptibility is related to economic vulnerability.

Undoubtedly, one key task during the early pandemic response efforts is using epidemiological records and mathematical and statistical modeling to forecast excess hospital care demand with formal quantified uncertainty.

In this paper, we pose a compartmental SEIRD model that considers both asymptomatic and symptomatic infection, including hospital dynamics. We model the residence time in each latent and infected compartments explicitly [14, 15], and we use records of daily confirmed cases and deaths to pose a statistical model that accounts for data overdispersion [16, 17]. Furthermore, we use Bayesian inference to estimate the initial state of the governing equations, the contact rate, and a proxy of the population size to make probabilistic forecasts of the required hospital beds, including the number of intensive care units. The model output has been used by Mexican public health authorities to assist decision making during the COVID-19 pandemic in more than 70 metropolitan areas and the country's 32 states.

Contributions and limitations

- We developed a model to produce accurate midterm (several weeks) probabilistic forecasting of COVID-19 hospital pressure, namely hospital beds and respiratory support or mechanical ventilation demands, using confirmed records of cases at hospital admittance and deaths.
- Our model accounts for policy changes in control measures, such as school closures [18] and lockdowns, as breakpoints in the transmission rates.
- Assuming a given fraction of asymptomatic individuals, we infer changes in the transmission rate and the effective population size before and after a given lockdown-relaxation day.
- Inferred changes in effective population size allows us to produce a forecast of lockdown-induced 2nd waves.

Since asymptomatic infection is not fully understood so far [19], the fraction of asymptomatic individuals is yet unknown. Therefore:

- The effective population size is only a proxy, and its absolute value is not meaningful, but only its relative value before and after a relaxation day.
- Without serological studies in the open population – ideally after an outbreak – it is impossible to forecast the population fraction that will be in contact with the virus by the end of the current outbreak.
- At this point, our model does not address next pandemic outbreaks beyond lockdown-induced 2nd waves.
- Finally, although the model does not account explicitly for biases due to behavioral changes [20, 21], population clustering and super spreading events [22], we argue that our approach to lockdowns and relaxation events is a proxy model of these more general events.

Related work

There are many modeling efforts aimed at forecasting the number of cases, deaths and hospital occupancy during the ongoing COVID-19 pandemic [23–28]. Broadly speaking, models are informed with evolving information about COVID-19 cases, clinical description of the patient residence time in each compartment, fraction of cases per age group, number of deaths, hospital bed occupancy, etc. Columbia University metapopulation SEIR model [23] forecasts are based on assumptions relating an effective contact rate with population density at a metropolitan area and social distancing policies. The COVID Act Now model [24] forecasts the replacement number R_t and the fraction of infections requiring hospitalization using the Bayesian paradigm to fit a SEIR model to cases, hospitalization, death, and recovery counts. The Imperial College response team mathematical model [25] uses an unweighted ensemble of four models to produce forecasts of the number of deaths in the week ahead for each country with active transmission. The IHME model [26] combines a mechanistic model of transmission with curve fitting to forecast the number of infections and deaths. Moghadas *et al.* [27] pose a mechanistic model parametrized with demographic data to project hospital utilization in the United States during the COVID-19 pandemic. The main goal of Moghadas *et al.* is to estimate hospital pressure throughout.

Other COVID-19 models have been used to explore exit strategies [29,30], the role of recovered individuals as human shields [31], digital contact tracing [32], break points in the contact rate to account for changes in suppression and mitigation policies [18] and lockdown-induced 2nd COVID waves [33] under the assumption that population is temporally geographically isolated.

Materials and Methods

“Models should not be presented as scientific truth” [34]. Indeed, models are tools intended to serve a specific purpose, evaluate or forecast particular aspects of phenomena and ideally should be developed following the processes of predictive science [35]. Namely, identify the quantities of interest (QoI), verify the computational and mathematical approximation error, including their implication in the estimation of QoI, and calibrate the parameters to adjust the model in light of data to bring it closer to experimental observation. When considering uncertainty, Bayesian inference may be used to calibrate some key model features given data. Finally, a validation process must be used to build confidence in the accuracy of the QoI predictions. Our model is built out of three interrelated components; a law for dynamics, a law for uncertainty, and the choice of parameters.

Dynamical model

As a proxy of hospital pressure, the quantities of interest in our model are the evolving demand of ICU/respiratory-support beds and no-ICU hospital beds. We developed a full compartmental SEIRD model featuring several compartments to describe hospital dynamics (see Figure 1 and supporting materials, SM) with sub-compartments to model explicitly residence rates as Erlang distributions [14,15]. The model has two variants, one with age structure and one that assumes age-independent dynamics. Here we describe the latter (see supporting material for some additional comments on the age-dependent model).

Succinctly our model goes as follows: once the susceptible individuals (S) become infected, they remain in the incubation compartment (E) for mean time of $1/\sigma_1$ days (i.e. residence rate σ_1). After the incubation period, exposed individuals become

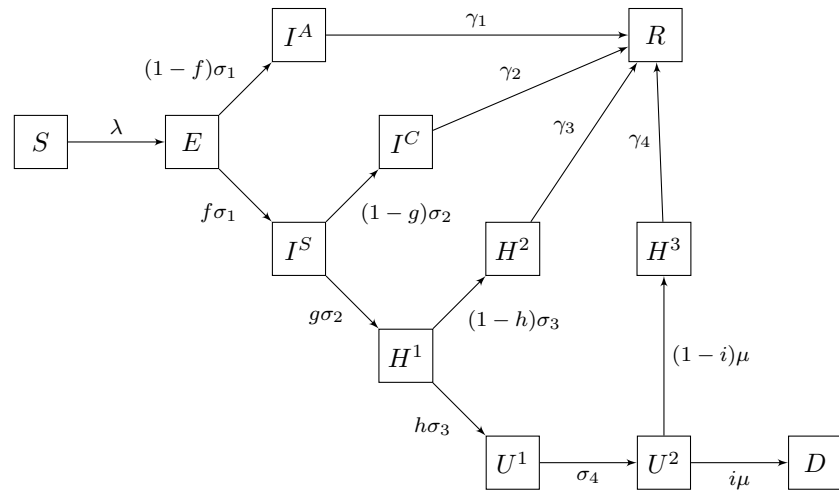


Fig 1. Schematic diagram of the dynamical model. Erlang sub-compartments not shown. For a precise definition of parameters, see supplementary material.

infectious and a proportion f of them become sufficiently severe symptomatic cases (I^S) to approach a hospital, while the remaining cases become mild-symptomatic to asymptomatic (I^A). The asymptomatic/mild-symptomatic cases remain infectious a mean time of $1/\gamma_1$ days and eventually recover. For the symptomatic cases (I^S) we assume that after an average time of $1/\sigma_2$ days a proportion g of infected individuals will need hospitalization (H^1), while the rest (I^C) will receive ambulatory care, recovering after an average convalescent time of $1/\gamma_2$ days in quarantine. The hospitalized patients (H^1) remain an average time of $1/\sigma_3$ days until a fraction h will need assisting respiratory measures or ICU care such as mechanical ventilation (U^1). The remaining fraction $1 - h$ of hospitalized patients (H^2) will recover after $1/\gamma_3$ days in average. Respiratory-assisted/ICU patients (U^1) remain in that state an average of $1/\sigma_4$ days until a critical day is reached when a proportion i of them will die (D) and the remaining proportion $1 - i$ will recover (H^3) after an average period of $1/\gamma_4$ days. Similar models have been proposed by [2, 31, 32, 36]. For the infection force (λ) we assume that only mild-symptomatic/asymptomatic (I^A) and symptomatic (I^S) individuals spread the infection, that is

$$\lambda = \frac{\beta_A I^A + \beta_S I^S}{N_{eff}}$$

where β_A and β_S are the contact rates of asymptomatic/mild-symptomatic and symptomatic individuals, respectively.

Parameters and observational model

The model has two kinds of parameters that have to be calibrated or inferred; the ones related to COVID-19 disease and hospitalization dynamics (such as residence times and proportions of individuals that split at each bifurcation of the model) and those associated with the public response to mitigation measures such as the contact rates β 's and the effective population size N_{eff} during the outbreak. Some of these parameters can be estimated from hospital records or found in recent literature or inferred from reported cases and deaths, but some remain mostly unknown. In the latter category, we have N_{eff} and the fraction $1 - f$ of asymptomatic/mild-symptomatic infections.

Reported values of the proportion of asymptomatic/ mild-symptomatic infections cases $1 - f$ range from 10% to 75%, and even 95% in children population [6, 7, 37]. The number N_{eff} is lower than the full population of a metropolitan area and depends on different aspects. Still, it is likely to be a consequence of unequal observance of social distancing public policies among the population that, in turn, yield some clustering effects. As lockdown measures are relaxed, more people become in contact with the outbreak, and N_{eff} may change. This change is a proxy of the fact that the connectivity between clusters increases, and new paths become open to the virus to colonize the full population. Notice that in our model the total number of patients that will visit a hospital is given roughly (bounded) by the product $N_{eff} \times f$ and the total number of patients admitted to the hospital is given by $N_{eff} \times f \times g$, where g is the portion of infected persons that need hospitalization.

Since our QoI are concerned with hospital pressure, we choose from the available data two sources of information for the observational model: The registered confirmed COVID-19 patients at hospitals, with or without hospitalization, and B deceased patients. Even under an outbreak, these data are reasonably consistent and systematic information on the inflow A and outflow B, that “hedge” the hospital dynamics. We have evidence (see SM) that given our choice of observation model, the inference of our QoI only depends on the product $N_{eff} \times f \times g$, and not on the value of their factors. The fraction g is easy to estimate from hospital records (see SM) of admissions and ambulatory patients. Thus we are only required to postulate a value for the product $N_{eff} \times f$.

Lockdowns, relaxation and lockdown-induced 2nd waves

The model features lockdown intervention and lockdown relaxation as discontinuities in the transmission rate and the effective population size N_{eff} . To model lockdown interventions, a breakpoint is established at which $\beta = \beta_1$ before and $\beta = \beta_2$ after the intervention day. This creates a non-linear time-dependent $\beta(t)$ [18, 38]. In the same fashion, further intervention days may be included by adding more change points and β parameters. These additional parameters are then included in the inference.

Assuming that effective population size N_{eff} and transmission rates are fixed, SEIRD type models converge to the attractor $E = 0, I = 0$, i.e. the system models an epidemic that dies out after one single acme. Even for sensible non-constant transmission rates, these kinds of models are only able to produce single epidemic outbreaks. Therefore, in order to be able to estimate secondary outbreak waves after lockdown relaxation measures, one necessarily needs to estimate a different N_{eff} before and after relaxation days. An increase of N_{eff} is aimed as a proxy to the rise in the average connectivity between population clusters after relaxation days. Our approach here is as follows: we include the new parameter(s) $\omega_i \in (0, 1)$ and set $N_{eff} = \omega_i N$, where N is the total population of the metropolitan area or region under study. Next, we postulate a value of f fixed and estimate ω_1 and ω_2 before and after the relaxation day, respectively. To add flexibility to our method and model possible changes in population behavior, we also force a shift from β_1 to β_2 before and after a relaxation day. With this procedure, given f we are able to estimate N_{eff} in terms of the ω 's. Nevertheless, due to the confounding effect of the product $N_{eff} \times f = N \times \omega \times f$, we are still unable to estimate the real N_{eff} 's until the real value of f is known. Relaxation days may include both lockdown relaxation day or mayor changes in the public response to mitigation measures (see SM for a methodology on how to choose relaxation days).

Setting lockdown and relaxation days

As explained before, we model interventions and relaxation days as discontinuities in β and (β, ω) , respectively. We consider a lockdown day on 22 March 2020, where a country-wide lockdown started. In Mexico City, we include a second intervention day to model a further local intervention in early April. The methodology to set the relaxation days is as follows: we computed R_t following [39] and look for days where local minimums occur. We choose the relaxation days judiciously, keeping them at least three weeks apart to have enough data to perform the inference. A rise in R_t can be a consequence of an increase in contact rates, effective population size, or both. In either case, our model captures these changes, that serve as a model for the connectivity changes in the outbreak cluster structure.

Observational model and data

To make our inferences, we use both confirm cases and deceased counts. In some regions, sub reporting of COVID-19 related deaths may become relevant, especially in places hit by a severe outbreak [40]. Nonetheless, deaths are a more reliable data source to estimate a COVID-19 outbreak, especially in the forecast of hospital demand. The problem here is that the number of confirmed cases depends heavily on local practices, particularly with the intensity of testing, adding a level of complication if testing intensity has varied due to ambiguous policies. Regarding data from Mexico, patients are tested when arriving at hospitals with probable COVID-19 symptoms and limited testing is done elsewhere; accordingly, most confirmed COVID-19 cases are limited to A as described above. Therefore, for our inferences, we use both confirm cases A and deceased counts B, as explained in the previous section. Regarding data availability for our observational model, we use the patient's reported onset of symptoms date. Due to administrative reporting delays, we discard the last 11 days of reporting and add four days as the time stamp for hospital admittance A. We use the registered deceased date as the timestamp for B.

We consider daily deaths counts d_i and its theoretical expectation that is estimated in terms of the dynamical model as $\mu_D(t_i) = D(t_i) - D(t_{i-1})$ for the metropolitan area or region being analyzed. Analogously, we consider daily cases c_i and its corresponding $\mu_c(t_i)$ given by the daily flux entering the H^1 compartment, which may be calculated as in [17], namely

$$\mu_c(t_i) = \int_{t_{i-1}}^{t_i} g\sigma_2 I_m^S(t) dt,$$

where $I_m^S(t)$ is the last state variable in the I^S Erlang series. We calculate the above integral using a simple trapezoidal rule with 10 points (1/10 day).

Bayesian Inference

In order to carry out a likelihood-based analysis, we assume that epidemic data has more variation than implied by a standard Poisson process, as is the case in other ecological studies. Following [16], we postulate that the number of both the confirmed cases and deaths follows a negative binomial distribution NB . Denoting the mean and variance as μ and σ^2 , and requiring that $\sigma^2 = \omega\mu + \theta\mu^2 > \mu$ we enforce overdispersion for suitable chosen parameters ω and θ . For data y_i , namely c_i and d_i , we reparametrize the negative binomial distribution and let $y_i \sim NB(p\mu(t_i), \omega, \theta)$, with fixed values for the overdispersion parameters ω, θ and an additional reporting probability p (see SM for further details).

We assume conditional independence in the data, and therefore from the NB model, we obtain a likelihood. Our parameters are the contact rate parameter β 's, the ω 's and

crucially we also infer the initial conditions $E(0), I^A(0), I^S(0)$. Letting $S(0) = N - (E(0) + I^A(0) + I^S(0))$ and setting the rest of the parameters to zero, we have all initial conditions defined and the model may be solved numerically to obtain μ_D and μ_c to evaluate our likelihood.

Finally, regarding the elicitation of the parameters prior distribution, we use *Gamma* distributions with scale 1 and shape parameter 10 to model the initial conditions $E(0), I^A(0), I^S(0)$ of the community transmission. Following the argumentation of Cori *et al.* [39], we assume that local transmission starts when there are 10 confirmed cases. The rationale for using Gamma distribution priors is that we can specify the distribution by prescribing its first two moments, and the resulting distribution verifies a maximum entropy condition. Namely, we obtain the less informative distribution that has the prescribed mean and (log) variance [41]. The prior for the first transmission rate β_0 , is a long tail, log Normal with $\sigma^2 = 1$ and scale parameter 1; that is $\log(\beta_0) \sim N(0, 1)$. For the subsequent β 's, we use autoregressive priors to impose some coherence from one change point to the next with $\log(\beta_i) \sim N(\log(\beta_{i-1}), 1)$. The prior on ω_i is a *Beta*(1 + 1/6, 1 + 1/3) restricted to $\omega_i > \omega_{i-1}$. This beta distribution is a fairly flat near uniform density in $[0, 1]$, touches zero in 0 and 1, and is slightly skewed to lower values. We model here the unlikely values $\omega_i = 0, 1$ and that under current social distancing measures we expect smaller rather than larger N_{eff} . Otherwise, the prior for the ω_i 's is rather diffuse and non-informative.

To sample from the posterior we resort to MCMC using the “t-walk” generic sampler [42]. The MCMC runs semi-automatic, with consistent performances in most data sets. For any state variable V , the MCMC allows us to sample from the posterior predictive distribution for $V(t_i)$. By plotting some of its quantiles sequentially, we may produce predictions with quantified probabilistic uncertainty, as seen in Figure 3.

As in the case of climate forecasting, due to the stochastic nature of a pandemic outbreak point-wise estimates such as the maximum a posteriori estimate (MAP) does not provide good descriptions of the outbreak evolution. No single trajectory of the SEIRD model provides a good description of the outbreak evolution, nor give accurate forecasts.

Results

Local transmission started at different dates in each Mexican metropolitan area given the different communicability with Mexico City and with the rest of the world. On the other hand, a country-wide general lockdown started on May 22nd until June 1st where each state started differently local control measures. Figure 3 (a) shows the model forecast, with quantified uncertainty, of the daily records of COVID-19 confirmed cases in Mexico City. Gray bars correspond to two weeks of trimmed data to assess the model performance.

Figure 3 (b) depicts records and forecasts incidence of deaths. In Figure 3 (c) and (d) we compare the model forecasts with hospital bed and ICU occupancy obtained from a secondary official source of epidemiological surveillance depicted as red bars. Notice that the forecast begins after June 20th, and an uncertainty cone opens to the right the next 3 to 4 weeks. However, the attractor of the dynamical system closes the cone for longer times, and the predictive power of our forecast decreases.

The estimate of total hospital bed occupancy corresponds to the daily integral of H^1 and H^2 in the model and the ICU occupancy corresponds to the U^1 daily integral. We consistently overestimated the total number of hospital beds and ICU units and shifted to earlier times. We calibrate residence times from reports on daily demand of hospital beds and intensive care unit records from *Instituto Mexicano del Seguro Social* or Mexican Social Security Institute (IMSS) at the early stages of the outbreak. Given the

virus's emerging nature and evidence of an excess of hospital pressure in places like Spain, Italy, and New York. We deliberately assumed value parameters to forecast higher demands in both types of hospital occupancy.

In Figure 3 (a) and (b), black and red vertical lines represent lockdown and relaxation days, respectively. Our model forecasts three different bump-shape regions where the effective population size increases.

Mexico city, metropolitan area

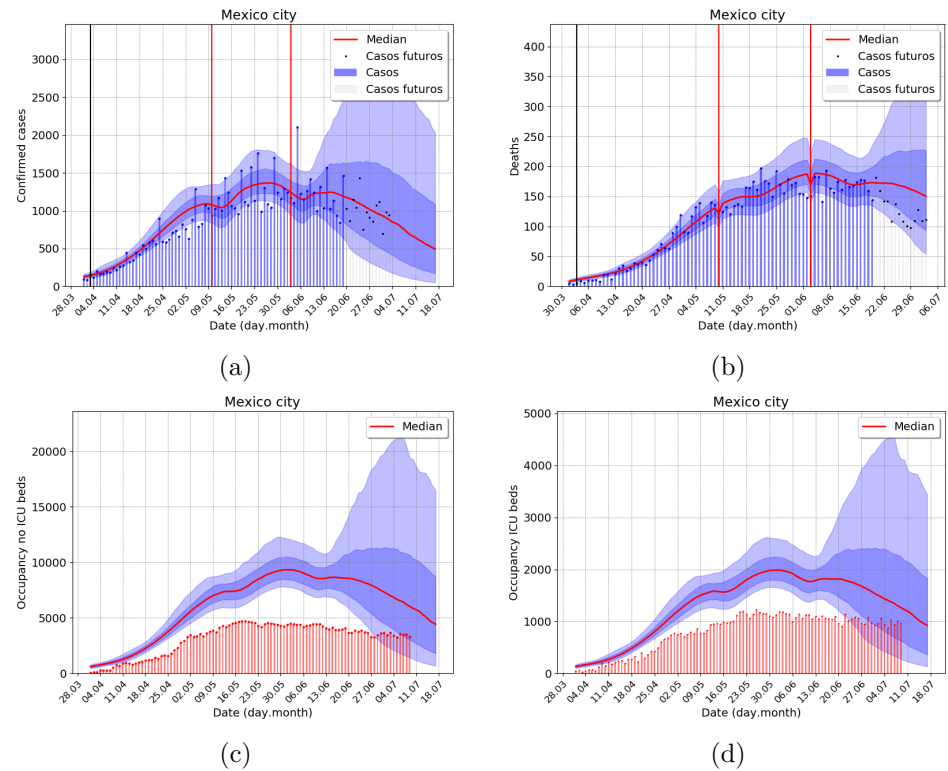


Fig 2. Outbreak analysis for Mexico city metropolitan area, using data from 7 July 2020. Gray bars correspond to two weeks of trimmed data. (a) Incidence of confirmed cases, (b) Incidence of deaths (c) No ICU, and (d) ICU demand of hospital beds. Total population 21,942,666 inhabitants

Cancun metropolitan area

We present the case of Cancun's metropolitan area since it is a medium-sized city with considerable international connectivity was among the first ones with an outbreak in Mexico. The forecast shows a clear first wave, with a long decreasing tail and a lockdown-induced 2nd wave after a lockdown easing and reopening of touristic activities.

In Figure 4 we show the three posterior distributions for ω . In (a), we show the Mexico City case, and in (b), we show the Cancun metropolitan area case. In the supplementary material, we show the outbreak analysis for some other cities to illustrate different aspects of our forecasting model's performance. Besides this paper's examples, we apply our model to 70 metropolitan areas and the 32 states in Mexico ("ama" model; <https://coronavirus.conacyt.mx/proyectos/ama.html>, in Spanish).

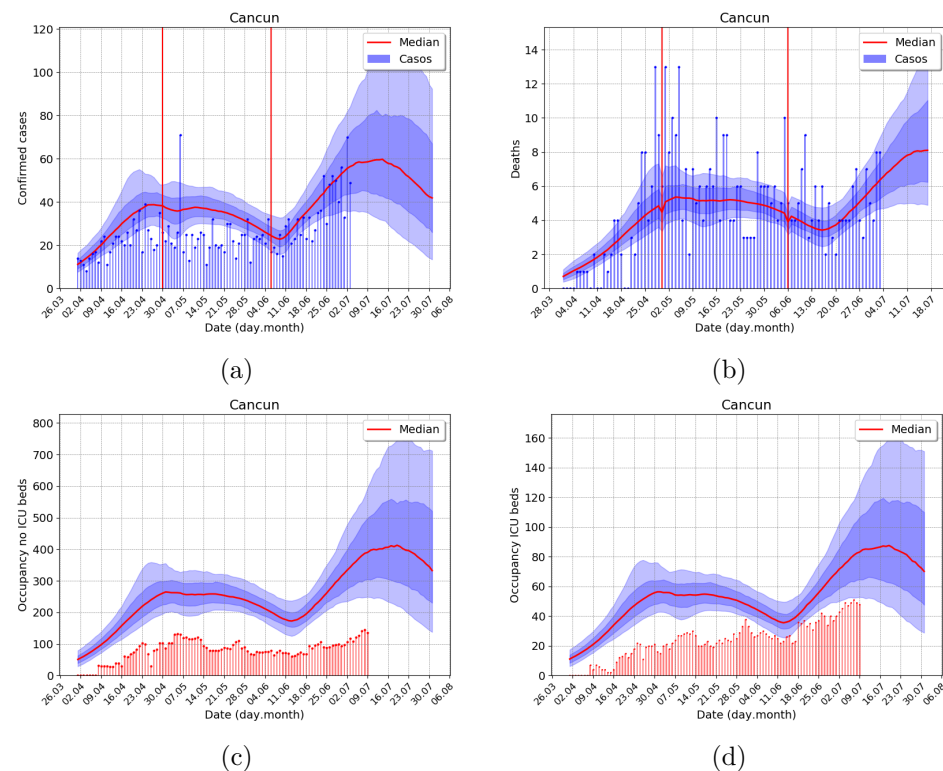


Fig 3. Outbreak analysis for Cancun metropolitan area, using data from 7 July 2020. (a) Incidence of confirmed cases, (b) Incidence of deaths (c) No ICU, and (d) ICU demand of hospital beds. Total population 891,843 inhabitants.

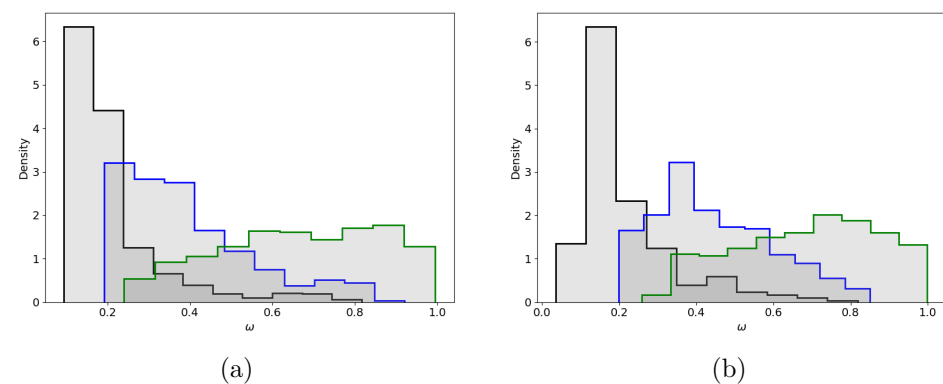


Fig 4. Posterior distribution for ω 's ($N_{eff} = N\omega_i, f = 0.4$) (a) Mexico city (b) Cancun.

Discussion

We present a compartmental SEIRD model to make probabilistic forecasts of hospital pressure during COVID-19 outbreaks in metropolitan areas. In this work, we also consider lockdowns and lockdown-relaxations as two different kinds of interventions. We model the former as a change in transmission rates while the latter also allows for changes in the effective population size. These changes in effective population size and

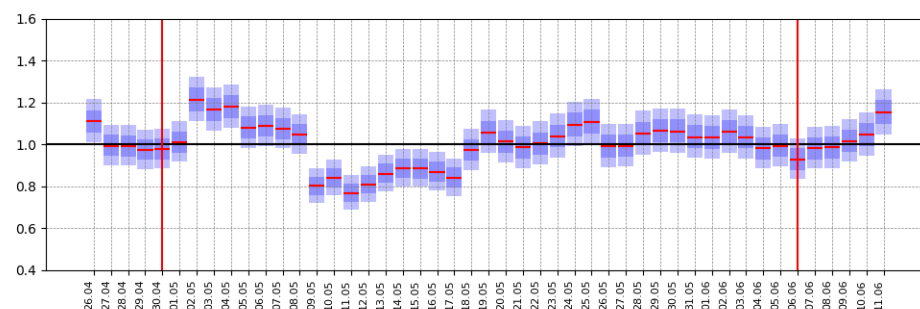


Fig 5. Daily R_t 's calculated as in [39] for Cancun metropolitan area. Two relaxation days, marked with red vertical lines, were judiciously included at local minimums, allowing for a minimum gap of 3 weeks.

transmission rates are used as a proxy of behavioral changes, changes in the connectivity between population clusters and super spreading events.

We in-line monitor exogenous changes in SEIRD parameters to inform the model of interventions and relaxation days, in the same spirit of our inferences in the transmission rate and initial conditions. Our analysis showed that we account for the outbreak evolution in many metropolitan areas by setting one lockdown on 22 May 2020, the beginning of country-wide lockdown, and two more relaxation days. One around 10 May, mother's day, a popular celebration in Mexican culture where families gather together. The other relaxation day is set around four to eight days after 1 June, the announced date for the country-wide lockdown relaxation measures. In some cases, we also impose somewhat different relaxation days to account for local changes, such as the opening of tourist activity. Of note, all relaxation days were set exclusive by our R_t based methodology. The interpretation, as above, came afterward.

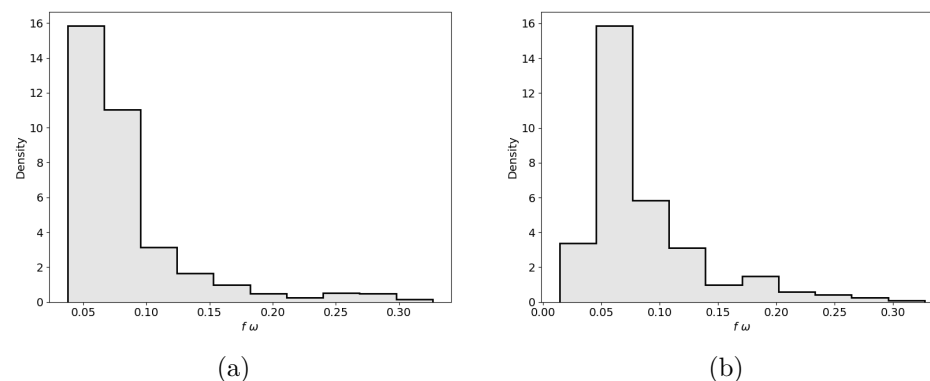


Fig 6. Posterior distribution of $f \times \omega_0$ for (a) Mexico city and (b) Cancun metropolitan areas, as explained in Section . Note how, using the proxy value of $f = 0.4$, the proportion $f \times \omega_0$ is estimated close to 0.05. In a previous version of this model we postulated $\omega = 1$ and $f = 0.05$.

In a prior version of this ama model –before reaching the first acme– we did not infer the effective population size, namely ω . Instead, by modeling several countries that already had passed the acme, we postulated $\omega = 1$ and $f = 0.05$. This yields some controversy that even reached the news media [43]. With the current model we can show that the product $\omega \times f$, that is the quantity that affects the estimates, takes

maximum a posteriori values between 0.04 and 0.12 in most of our studied cases, see Figure 6. This fact explains the reason why our forecast weeks ahead on the acme have been rather accurate. It is also important to notice that this is evidence of a confounding effect between ω , β , and the transmission rates β on the early stages of the outbreak. Many forecasting models that fail to recognize this fact will consequently fail on their acme estimates.

We note that after the acme SEIRD type models converge to the attractor $E = 0, I = 0$, and long term estimates tend to overestimate the decline of an outbreak. The underlying assumption that social contact and other conditions remain constant is not reasonable for most societies over long periods (e.g., more than eight weeks). Conservative short to mid-term forecasts must be preferred and change points added when necessary.

The age-independent model has proven to be adequate to produce accurate forecasts for the hospitalization dynamics during current outbreaks. Therefore the added complexity of the age-structure model may not justify its use at this point. However, we continue to work in our age structure model.

Since our QoI are related to the hospital pressure, we choose all parameters conservatively. However, Erlang densities can not correctly approximate residence times of some cases real distributions and more general distributions should be considered. Moreover, as health professionals learn to treat the disease, hospital residence times also change. Both these effects should also be considered to obtain more accurate estimates.

Our observation model is designed to integrate data after the nonlinear term in the flow diagram of the dynamic model (see Figure 1) and the rest of the dynamics is proportional to the hospital occupancy curves, therefore the model forecasts can be used as a proxy of the full outbreak. We follow this idea to predict the date when an outbreak will reach its first peak as well as lockdown induced second waves.

The confounding effect between the population size, namely ω , and the fraction of asymptomatic/mild-symptomatic infections $1 - f$ makes it impossible to forecast the population that will be in contact with the virus at the end of an outbreak reliably. Likewise, although it is possible to make a model based analyses of scenarios of lockdown exit strategies, scenario estimation is limited due to the lack of information regarding population viral seroprevalence. Therefore, without serological studies in the open population after a COVID-19 outbreak it is not possible to assess the final outbreak size.

Up to our knowledge, there are very few models that can produce forecast lockdown-induced 2nd waves with quantified uncertainty. Although more elaborate models can be considered, our model is simple and flexible enough to deliver reliable and useful forecasts.

Acknowledgments

The authors wish to thank Elena Álvarez-Buylla, Paola Villarreal (CONACYT) and Hugo López-Gatell (Secretaría de Salud) for their support and comments for improving this forecast model and also Instituto Mexicano del Seguro Social (IMSS) for sharing data to adjust the hospitalization dynamics. Special thanks to CIMAT technicians Judith Esquivel-Vázquez and Oscar González-Vázquez (CIMAT-CONACYT), for their help in retrieving data and running our programs at a CIMAT cluster. We also thank the CONACYT COVID-19 response group, for additional comments in the development of our model. The authors are partially founded by CONACYT CB-2016-01-284451 grant. AC was partially supported by UNAM PAPPIT-IN106118 grant.

Supporting information

348

List of Supplementary Materials

349

S1 Other examples: Toluca, Merida, Cuernavaca, Acapulco, Culiacan

350

Figures S1-S6

351

Tables S1-S4

352

S2 Model

353

S3 Data and observation model

354

S4 Modeling interventions and Bayesian Inference

355

S5 Adding age structure

356

S6 Confounding effect of $N_{eff} \times f$

357

S7 Data Sources

358

S1 Other examples

S1.1 Toluca

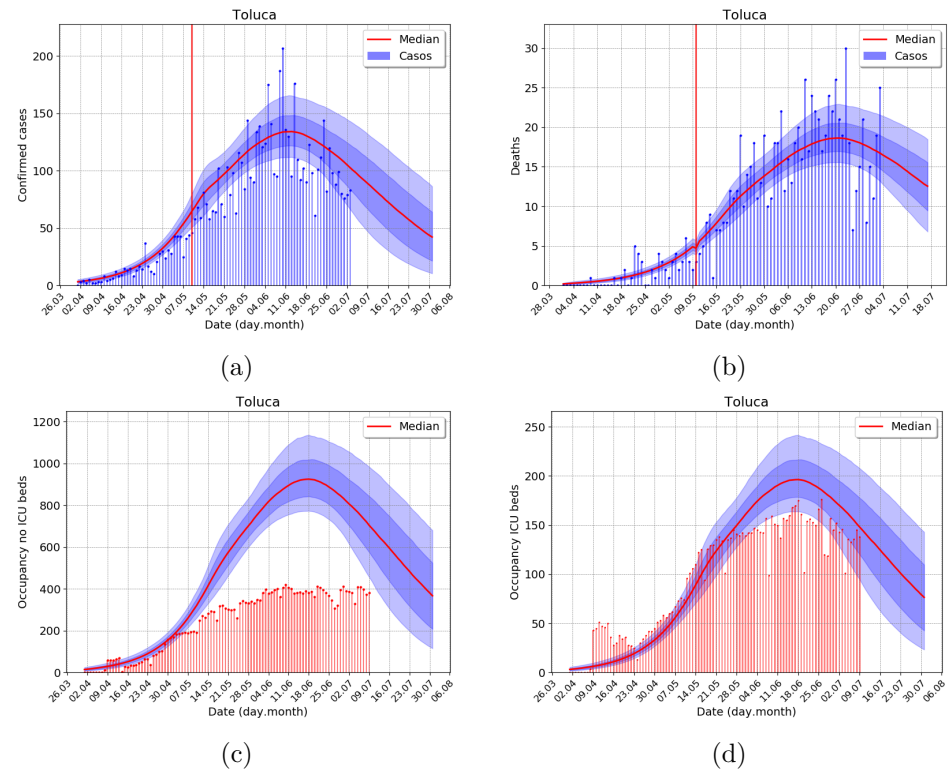


Fig S1. Outbreak analysis for Toluca (the state of Mexico, Mexico central highlands) metropolitan area, using data from 7 July. (a) Incidence of confirmed cases, (b) Incidence of deaths (c) No ICU, and (d) ICU demand of hospital beds. Total population 2,377,828 inhabitants. Forecast for the city of Toluca is in course with one relaxation event.

S1.2 Merida

361

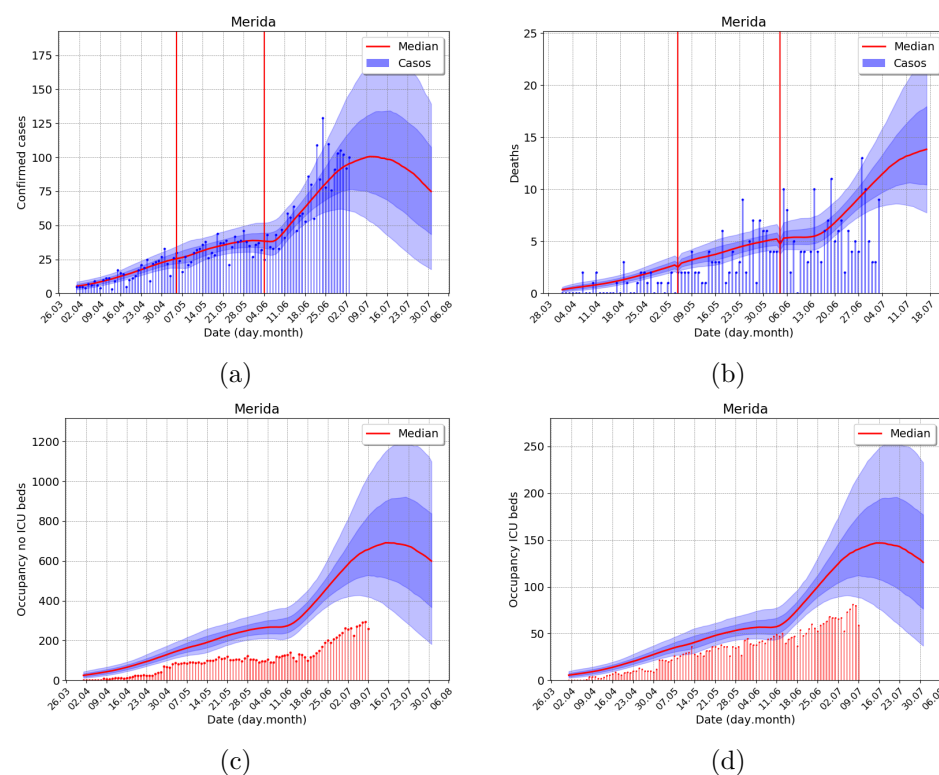


Fig S2. Outbreak analysis for Merida (state of Yucatan, Yucatan peninsula) metropolitan area, using data from 7 July. (a) Incidence of confirmed cases, (b) Incidence of deaths (c) No ICU, and (d) ICU demand of hospital beds. Total population 1,237,697 inhabitants. There is an evident ongoing second outbreak in the city of Merida.

S1.3 Cuernavaca

362

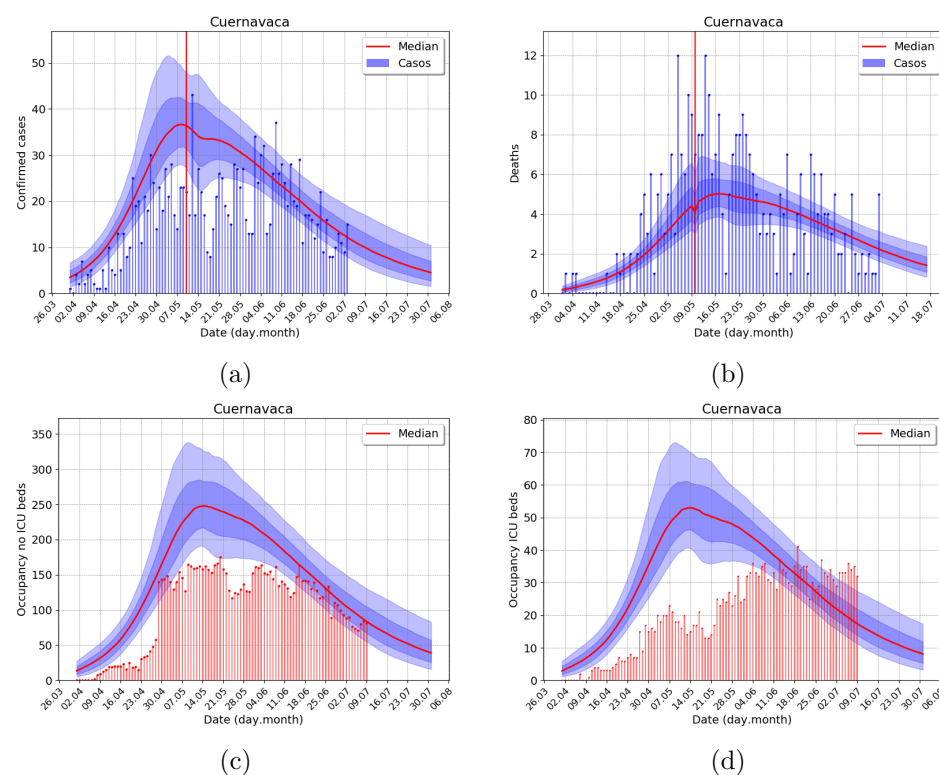


Fig S3. Outbreak analysis for Cuernavaca (the state of Morelos, Mexico central highlands) metropolitan area, using data from 7 July. (a) Incidence of confirmed cases, (b) Incidence of deaths (c) No ICU, and (d) ICU demand of hospital beds. Total population 1,059,521 inhabitants. In the city of Cuernavaca, the model captures the slow decline of the outbreak.

S1.4 Acapulco

363

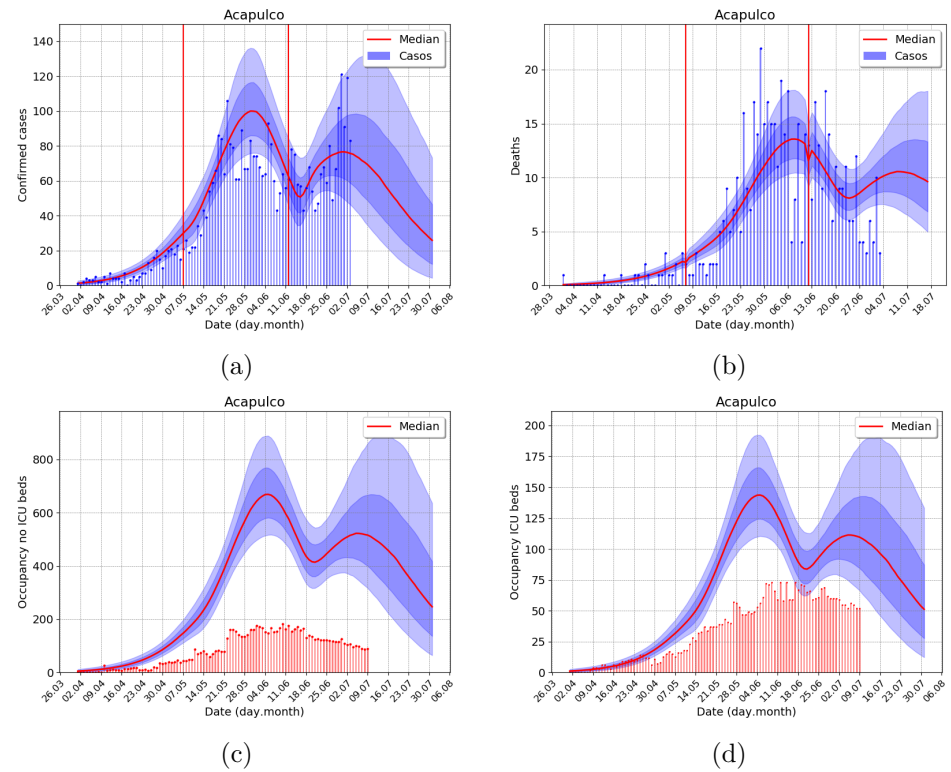


Fig S4. Outbreak analysis for Acapulco (the state of Guerrero, Mexico south pacific shore) metropolitan area, using data from 7 July. (a) Incidence of confirmed cases, (b) Incidence of deaths, (c) No ICU, and (d) ICU demand of hospital beds. Total population 1,059,521 inhabitants. The outbreak in the city of Acapulco is an example of a second outbreak of the same size.

S1.5 Culiacan

364

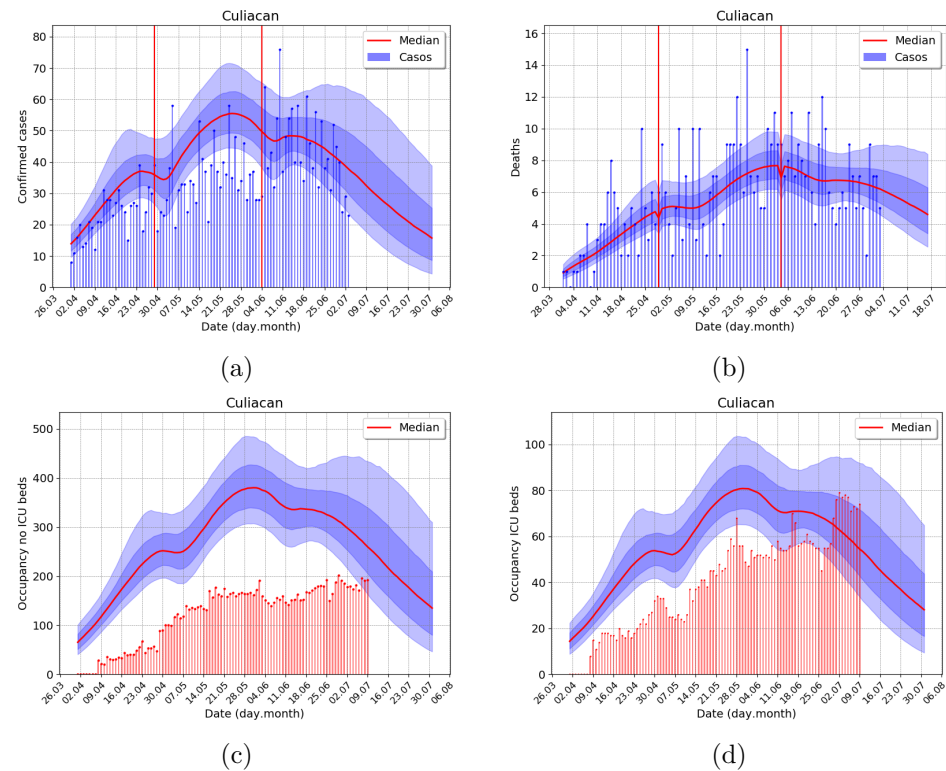


Fig S5. Outbreak analysis for Culiacan (state of Sinaloa, Mexican north pacific shore) metropolitan area, using data from 7 July. (a) Incidence of confirmed cases, (b) Incidence of deaths, (c) No ICU, and (d) ICU demand of hospital beds. Total population 962,871 inhabitants. Example of secondary outbreak with two lockdown-induced second waves.

S2 Model

365

We developed a dynamic transmission compartmental model to simulate the spread of the novel coronavirus SARS-CoV-2. A definition of the state variables is given in Table S1. Additionally, an “Erlang series” is included for most of these state variables to account for non-exponential residence times. The model may be described conceptually with the graph in Figure 1. Without showing the Erlang series for the

Table S1. Description of the state variables in the dynamic model

Variable	Description
S	Susceptibles
E	Latent individuals
I^A	Asymptomatic/mild-symptomatic individuals
I^S	Symptomatic individuals
I^C	Out-patients
H^1	Hospitalized patients, initial stage
H^2	Hospitalized patients (no ICU)
U^1	Hospitalized patients (ICU or respiratory support)
U^2	Hospitalized patients (ICU or respiratory support) critical day
H^3	Hospitalized patients recovering after ICU or respiratory support
R	Recovered
D	Deceased

sub-compartments the system of equations in the model is as follows:

$$\begin{aligned}
 \frac{dS}{dt} &= -\frac{(\beta_A I^A + \beta_S I^S)}{\omega N} S & \frac{dE}{dt} &= \frac{(\beta_A I^A + \beta_S I^S)}{\omega N} S - \sigma_1 E \\
 \frac{dI^A}{dt} &= (1-f)\sigma_1 E - \gamma_1 I^A & \frac{dI^S}{dt} &= f\sigma_1 E - \sigma_2 I^S \\
 \frac{dI^C}{dt} &= (1-g)\sigma_2 I^S - \gamma_2 I^C & \frac{dH^1}{dt} &= g\sigma_2 I^S - \sigma_3 H^1 \\
 \frac{dH^2}{dt} &= (1-h)\sigma_3 H^1 - \gamma_3 H^2 & \frac{dU^1}{dt} &= h\sigma_3 H^1 - \sigma_4 U^1 \\
 \frac{dU^2}{dt} &= \sigma_4 U^1 - \mu U^2 & \frac{dD}{dt} &= i\mu U^2 \\
 \frac{dH^3}{dt} &= (1-i)\mu U^2 - \gamma_4 H^3 & & \\
 \frac{dR}{dt} &= \gamma_1 I^A + \gamma_2 I^C + \gamma_3 H^2 + \gamma_4 H^3 & &
 \end{aligned}$$

In Table S2, we give a brief description of all the parameters in the model.

S2.1 Infection force and basic reproductive number \mathcal{R}_0

For the infection force (λ) we assume that individuals that spread the infection correspond to the mild-symptomatic/asymptomatic (I^A) and symptomatic individuals (I^S) before they get contact with the health care system or doctor, e.g.

$$\lambda = \frac{\beta_A I^A + \beta_S I^S}{N_{eff}}$$

We compute the basic reproductive number R_0 of the epidemic by the next generation matrix method [44] and obtain

$$R_0 = (1-f)\frac{\beta_A}{\gamma_1} + f\frac{\beta_S}{\sigma_2}.$$

S2.2 Values of model parameters

Since our QoI are related to the hospital pressure we choose all parameters conservatively. For each metropolitan area, we assume that N corresponds to its full

Table S2. Description of model parameters.

Parameter	Description	Units
N	Total number individuals in the population	–
$\omega(t)$	Percentage susceptible individuals in the population at time t	–
$\beta_A(t)$	Transmission rate of asymptomatic/mild-symptomatic individuals (asx/mild-sym) at time t	per day
$\beta_S(t)$	Transmission rate of symptomatic individuals at time t	per day
κ	Relative strength between the transmission the rate of asx/mild-sym and symptomatic	–
f	Proportion of infected persons with strong enough symptoms to visit a hospital	–
g	Proportion of infected persons that need hospitalization	–
h	Proportion of hospitalized patients requiring respiratory support or ICU care	–
i	Proportion of Respiratory-assisted or ICU patients deceased	–
$1/\sigma_1$	Average incubation time	day
$1/\sigma_2$	Average time from symptomatic onset to hospital visit	day
$1/\sigma_3$	Average time from hospital admission to respiratory support or ICU care	day
$1/\sigma_4$	Average time with respiratory support or ICU care	day
$1/\mu$	Average length of critical stage of respiratory-support/ICU between death and recovery	day
$1/\gamma_1$	Average time that asymptomatic/mild-symptomatic individuals remain infectious	day
$1/\gamma_2$	Average time of symptomatic individuals that recover without visiting a hospital	day
$1/\gamma_3$	Average time from hospital admission to hospital discharge	day
$1/\gamma_4$	Average time from respiratory-support/ICU care release to hospital discharge	day

population, as defined by Instituto Nacional de Estadística y Geografía (INEGI). The values of the transition probabilities are summarized in Table S3.

Table S3. Transition probabilities at bifurcations in the model

Parameter	Value	Reference
f	0.40	postulated
g	0.04375	[1], IMSS
h	0.25	[1], IMSS
i	0.5	[1], IMSS

S2.3 Erlang series and sub-compartments

To make the intrinsic generation-interval of the renewal equation in each compartment more realistic we divide each compartment of the model into m equal sub-compartments to generate an Erlang-distributed waiting time [14]. The Erlang distributions of each compartment is calibrated by two parameters: the rate λ_E and the shape m , a positive integer that corresponds to the number of sub-compartments on the model. In terms of these parameters the mean of the Erlang distribution is m/λ_E , this mean correspond to the average times in the dynamic model.

We use recent publications and information generously shared by the Instituto Mexicano de Seguridad Social (IMSS) to estimate the average time and the shape parameter of the Erlang series in each compartment. In Table S4, we give details of Erlang series lengths, residence times and imputed values.

Table S4. Average times and Erlang shape parameters for sub-compartments

Variable	Rates	Average time	Erlang shape m	Reference
S	β_S	Inferred	1	—
E	$1/\sigma_1$	5 days	4	[2]
I^A	$1/\gamma_1$	7 days	3	[45]
I^S	$1/\sigma_2$	4 days	3	[46]
I^C	$1/\gamma_1$	7 days	3	[47]
H^1	$1/\sigma_3$	2 days	10	[3]
H^2	$1/\gamma_3$	10 days	3	[3]
U^1	$1/\sigma_4$	10 days	3	IMSS
U^2	$1/\mu$	1 day	1	[47]
H^3	$1/\gamma_4$	4 days	5	[3]
R	None	—	—	—
D	None	—	—	—

S2.4 Relative strength between the transmission the rate of asymptomatic/mild-symptomatic and symptomatic

In our methodology, we aim to infer the force of the infection λ . This parameter is defined in terms of contact rate of asymptomatic/mild-symptomatic individuals β_A and contact rate of symptomatic individuals β_S . Due to the functional dependence of λ in these parameters, there is a lack of identifiability between β_A and β_S that can not be resolved without further assumptions. We assume that the relative strength between the transmission rate of asymptomatic/mild-symptomatic and symptomatic is modeled as a

fixed ratio κ . We model the value of κ directly as the ratio of the viral load of symptomatic and asymptomatic/mild-symptomatic patients [48, 49] and fixed it to $\kappa = 2$. Hence, the force of infection becomes $\lambda = \beta^S(I^S + \kappa I^A)/N_{eff}$.

S3 Data and observational model

For inference, we, therefore, consider daily confirmed cases c_i of patients arriving at H^1 and daily reported deaths d_i , for the metropolitan area or region being analyzed.

The first default model for count data is a Poisson distribution; however, epidemiological data tends to be over disperse. Thus, an over disperse generalized Poisson distribution may be needed to correctly, and safely, model these types of data. Following [16] (see main paper) the NB distribution is re parametrized in terms of its mean μ and “overdispersion” parameters θ and ω_{NB} , with $r = \frac{\mu}{\omega_{NB} - 1 + \theta\mu}$ and $p_{NB} = \frac{1}{\omega + \theta\mu}$, the number of failures before stopping and the success probability, respectively, in the usual NB parametrization. For data y_i we let $y_i \sim NB(p\mu(t_i), \omega, \theta)$, with fixed values for the overdispersion parameters ω_{NB}, θ and an additional reporting probability p . The index of dispersion is $\sigma^2/\mu = \omega_{NB} + \theta\mu$. Over dispersion with respect to the Poisson distribution is achieved when $\omega_{NB} > 1$ and the index of dispersion increases with size if $\theta \neq 0$; both desirable characteristics in outbreak data, adding variability as counts increase. In both cases we found good performance fixing $\omega_{NB} = 2$. To model daily deaths, we fixed $\theta = 0.5$ and for daily cases $\theta = 1$ implying higher variability for the later. The reporting probabilities are 0.95 for deaths and 0.85 for cases, with the assumption that the c_i ’s are confirmed sufficiently severe cases arriving at hospitals. As explained in the main paper, the theoretical expectations estimated in terms of the dynamical model are given by $\mu_D(t_i)$ and $\mu_c(t_i)$ for dead and cases, respectively.

S4 Modeling interventions and Bayesian Inference

We assume conditional independence in the data, and therefore from the NB model, we obtain a likelihood. Our parameters are the contact rate parameters β ’s, the ω ’s and crucially we also infer the initial conditions $E(0), I^A(0), I^S(0)$. Letting $S(0) = N - (E(0) + I^A(0) + I^S(0))$ and setting the rest of the parameters to zero, we have all initial conditions defined and the model may be solved numerically to obtain μ_D and μ_c to evaluate our likelihood. We use the *lsoda* solver available in the *scipy.integrate.odeint* Python function.

Moreover, as explained in the main paper, we also estimate ω_i with $N_{eff} = \omega_i N$, both before and after a change point (“relaxation day”; we show intervention days with black vertical lines and relaxation change points with red vertical lines in our plots, see next section). To make it a bit simple, we also force the model to have a new β parameter after a relaxation day. However, the process of numerically solving the system of ODE’s is slightly more complex. For a set of parameter values, including the , to evaluate the likelihood, one needs to apply the solver from time $t = 0$ to the first relaxation day, considering $N_{eff} = \omega_1 N$. Then for $\omega_2 > \omega_1$ we use the last values of all state variables as initial values for a second solve now with $N_{eff} = \omega_2 N$, and so forth.

To sample from the posterior we resort to MCMC using the “t-walk” generic sampler [42]. The MCMC runs semi-automatic, with a fairly consistent burn-in of 1,000 iterations (sampling initial values from the prior). We perform subsampling using the Integrated Autocorrelation Time, with pseudo-independent sample sizes of 1,000 to 1,500 with 400,000 iterations of the MCMC. This process takes roughly 60 min in a 2.2 GHz processor.

To illustrate the whole posterior distribution, for any state variable V (or $\mu_c(t_i)$), for each sampled initial conditions and β 's the model is solved at time t_1, t_2, \dots, t_k , including possibly future dates, obtaining a sample of $V(t_i)$ values for each t_i . The median and other desired quantiles are plotted vertically for each date considered, obtaining the plots as in Figure 3. Note that the traced median or other plotted quantiles do not necessarily correspond to any given model trajectory. It provides a far richer Uncertainty Quantification approach than the classical parameter estimates plug-in approach. Indeed, the sampled values for $V(t_i)$ do correspond to Monte Carlo samples of the posterior predictive distribution for $V(t_i)$.

S5 Adding age structure

Adding group ages is straightforward in these types of models. The number a of age groups is decided, and our model is repeated a times. Different residence times may be included [50] but we preferred to concentrate on the different transition probabilities g, h that vary nearly two orders of magnitude using age groups $[0, 25], (25, 50], (50, 65], (65, 100]$ [1]. The age structure is used to divide the initial infectious and susceptible population proportional to each age group size. We infer the same number of parameters, using a single β , with an optional weighting contact matrix [51] to model different contact rates among age groups in specific regions. Using a sufficiently flexible software design, progressing to an age-structure model is not complex; nonetheless, the MCMC may run substantially slower. We have experimented with our age-structured model using census data from Mexico and both uniform and non-uniform contact matrices. However, we do not report any of these results here, given that our non-age structure model has sufficiently enough predictive power, as already discussed.

S6 Confounding effect of $N_{eff} \times f$

To explain the confounding effect of $N_{eff} \times f$ we have two observations. First, if we let $f = \tilde{f}/\alpha$ for some $\alpha \in (0, \tilde{f})$ then differential equations for the variables $I^c, H^1, H^2, H^3, U^1, U^2$ and D remain invariant and the equations for I^s becomes

$$\frac{dI^s}{dt} = \frac{\tilde{f}}{\alpha} \sigma_1 E - \sigma_2 I^s.$$

By letting $\tilde{E} = E/\alpha$ the equation for I^s is also invariant with the substitution of E by \tilde{E} . Now, the equation for \tilde{E} is given by

$$\frac{d\tilde{E}}{dt} = \frac{(\beta_A I^A + \beta_S I^S)}{\alpha N_{eff}} S - \sigma_1 \tilde{E}.$$

By letting $\tilde{N}_{eff} = \alpha N_{eff}$ the latter equations becomes also invariant under the substitution of E by \tilde{E} . Therefore for the lower branch in the model (see Figure ??) the system of equations is invariant under the change of f and N_{eff} by \tilde{f} and \tilde{N}_{eff} provided $\tilde{N}_{eff} \times \tilde{f} = N_{eff} \times f$ holds. We need to adapt the equations for S, I^A , and R to get a consistent system of equations.

Second, to infer parameter β we inform the system with data at H^1 and D compartments. If $\tilde{N}_{eff} \times \tilde{f} = N_{eff} \times f$ holds, in view of our first observation, to fit these data the fluxes $f\sigma_1 E$ and $\tilde{f}\sigma_1 \tilde{E}$ in either case have to be the same. The solutions in the compartment I^s and after do not change in this case, but the individuals in the I^A compartment does change depending on which combination of N_{eff} and f or \tilde{N}_{eff}

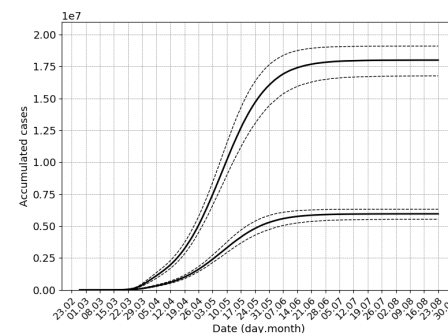


Fig S6. Outbreak analysis for Mexico city metropolitan area, data until 15 May. Total number of recovered R ; with $N_{eff} = N$, total population, and $f = 0.05$, $R(\infty)$ reaches approximately $17.5 \cdot 10^6$ while with $N_{eff} = N/3$, $f = 0.05 \cdot 3$, $R(\infty)$ only reaches roughly $5.5 \cdot 10^6$. However, the fit for cases and deaths and the predictive curves for hospital demand are identical (results not shown).

and \tilde{f} is considered. There is a range of validity for α where the inference of β does not change, but we do not explore this property further.

We also present numerical simulations to confirm this confounding effect (see Figure S6), but until the asymptomatic infection is fully described, it is not possible to resolve this issue.

S7 Data sources

We take metropolitan areas delimitation and population from [52] and [53], respectively. Official records of COVID-19 confirmed cases and deaths are reported in [54].

References

1. Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. London: Imperial College COVID-19 Response Team, March. 2020;16.
2. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of COVID-19 disease. medRxiv. 2020;.
3. Novel CPERE, et al. Zhonghua liu xing bing xue za zhi= Zhonghua liuxingbingxue zazhi, Number = 2, Pages = 145, Title = The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China, Volume = 41, Year = 2020;.
4. Zhou X, Li Y, Li T, Zhang W. Follow-up of asymptomatic patients with SARS-CoV-2 infection. Clinical Microbiology and Infection. 2020;.
5. Gandhi M, Yokoe DS, Havlir DV. Asymptomatic Transmission, the Achilles' Heel of Current Strategies to Control Covid-19. The New England Journal of Medicine. 2020;.
6. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond

- Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance*. 2020;25(10). doi:<https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180>.
7. Day M. Covid-19: identifying and isolating asymptomatic people helped eliminate virus in Italian village. *BMJ*. 2020;368. doi:10.1136/bmj.m1165.
8. Cordova-Villalobos JA, Macias AE, Hernandez-Avila M, Dominguez-Cherit G, Lopez-Gatell H, Alpuche-Aranda C, et al. The 2009 pandemic in Mexico: Experience and lessons regarding national preparedness policies for seasonal and epidemic influenza. *Gac Med Mex*. 2017;153(1):102–10.
9. Ha KM. A lesson learned from the MERS outbreak in South Korea in 2015. *Journal of Hospital Infection*. 2016;92(3):232–234.
10. Emanuel EJ. The lessons of SARS. *Annals of Internal Medicine*. 2003;139(7):589–591.
11. Hoffman SJ, Silverberg SL. Delays in global disease outbreak responses: lessons from H1N1, Ebola, and Zika. *American journal of public health*. 2018;108(3):329–333.
12. Gates B. The next epidemic—lessons from Ebola. *New England Journal of Medicine*. 2015;372(15):1381–1384.
13. Perra N, Gonçalves B. Modeling and predicting human infectious diseases. In: *Social phenomena*. Springer; 2015. p. 59–83.
14. Champredon D, Dushoff J, Earn DJ. Equivalence of the Erlang-distributed SEIR epidemic model and the renewal equation. *SIAM Journal on Applied Mathematics*. 2018;78(6):3258–3278.
15. Wearing HJ, Rohani P, Keeling MJ. Appropriate Models for the Management of Infectious Diseases. *PLOS Medicine*. 2005;2(7). doi:10.1371/journal.pmed.0020174.
16. Lindén A, Mäntyniemi S. Using the negative binomial distribution to model overdispersion in ecological count data. *Ecology*. 2011;92(7):1414–1421.
17. Zarebski AE, Dawson P, McCaw JM, Moss R. Model selection for seasonal influenza forecasting. *Infectious Disease Modelling*. 2017;2(1):56–70.
18. Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*. 2020;.
19. Park SW, Bolker BM, Champredon D, Earn DJ, Li M, Weitz JS, et al. Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: framework and applications to the novel coronavirus (SARS-CoV-2) outbreak. *medRxiv*. 2020;.
20. Eksin C, Paarporn K, Weitz JS. Systematic biases in disease forecasting—The role of behavior change. *Epidemics*. 2019;27:96–105.
21. Weitz JS, Park SW, Eksin C, Dushoff J. Moving Beyond a Peak Mentality: Plateaus, Shoulders, Oscillations and Other ‘Anomalous’ Behavior-Driven Shapes in COVID-19 Outbreaks. *medRxiv*. 2020;.
22. Liu Y, Eggo RM, Kucharski AJ. Secondary attack rate and superspreading events for SARS-CoV-2. *The Lancet*. 2020;395(10227):e47.

23. Yamana T, Pei S, Shaman J. Projection of COVID-19 Cases and Deaths in the US as Individual States Re-open May 4, 2020. medRxiv. 2020;.
24. Henderson M, *et al* . Covid Act Now;. Available from: https://docs.google.com/document/d/1cd_cEpNiI11TzUJBvw9sHLbrbUZ2qCcgN32IqVLa3Do/edit.
25. Bhatia S, *et al* . Short-term forecasts of COVID-19 deaths in multiple countries;. Available from: <https://mrc-ide.github.io/covid19-short-term-forecasts/index.html>.
26. COVID I, Murray CJ, *et al* . Forecasting the impact of the first wave of the COVID-19 pandemic on hospital demand and deaths for the USA and European Economic Area countries. medRxiv. 2020;.
27. Moghadas SM, Shoukat A, Fitzpatrick MC, Wells CR, Sah P, Pandey A, *et al* . Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proceedings of the National Academy of Sciences*. 2020;117(16):9122–9126.
28. Salje H, Kiem CT, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, *et al* . Estimating the burden of SARS-CoV-2 in France. *Science*. 2020;.
29. Karin O, Bar-On YM, Milo T, Katzir I, Mayo A, Korem Y, *et al* . Adaptive cyclic exit strategies from lockdown to suppress COVID-19 and allow economic activity. medRxiv. 2020;.
30. Di Domenico L, Pullano G, Sabbatini CE, Boëlle PY, Colizza V. Expected impact of reopening schools after lockdown on COVID-19 epidemic in Île-de-France. medRxiv. 2020;.
31. Weitz JS, Beckett SJ, Coenen AR, Demory D, Dominguez-Mirazo M, Dushoff J, *et al* . Modeling shield immunity to reduce COVID-19 epidemic spread. *Nature Medicine*. 2020; p. 1–6.
32. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, *et al* . Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*. 2020;.
33. Daunizeau J, Moran R, Brochard J, Mattout J, Frackowiak R, Friston K. Modelling lockdown-induced secondary COVID waves in France. medRxiv. 2020;doi:10.1101/2020.06.24.20139444.
34. Jewell NP, Lewnard JA, Jewell BL. Predictive Mathematical Models of the COVID-19 Pandemic: Underlying Principles and Value of Projections. *JAMA*. 2020;doi:10.1001/jama.2020.6585.
35. Oden JT, Moser R, Ghattas O. Computer predictions with quantified uncertainty, part I. *SIAM News*. 2010;43(9):1–3.
36. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, *et al* . Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases*. 2020;.
37. Silal S, Pulliam J, Meyer-Rath G, Nichols B, Jamieson L, Kimmie Z, *et al* . Estimating cases for COVID-19 in South Africa Update: 19 May 2020; 2020. Available from: https://www.nicd.ac.za/wp-content/uploads/2020/05/SACovidModellingReport_ProvincialLongTermProjections_Final-00000002.pdf.

38. Wang W. Epidemic models with nonlinear infection forces. *Mathematical Biosciences and Engineering*. 2006;3(1551-0018₂006₁₂67) : 267. doi : [http : //dx.doi.org/10.3934/mbe.2006.3.267](http://dx.doi.org/10.3934/mbe.2006.3.267).
39. Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. *American Journal of Epidemiology*. 2013;178(9):1505–1512. doi:10.1093/aje/kwt133.
40. Preliminary Estimate of Excess Mortality During the COVID-19 Outbreak — New York City, March 11–May 2, 2020;. Available from: <http://dx.doi.org/10.15585/mmwr.mm6919e5>.
41. Singh VP, Rajagopal A, Singh K. Derivation of some frequency distributions using the principle of maximum entropy (POME). *Advances in Water Resources*. 1986;9(2):91–106.
42. Christen JA, Fox C. A general purpose sampling algorithm for continuous distributions (the t -walk). *Bayesian Anal*. 2010;5(2):263–281. doi:10.1214/10-BA603.
43. Ahmed A. As Official Toll Ignores Reality. *New York Times*. 9 May 2020;A.
44. Van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*. 2002;180(1-2):29–48.
45. Team EE, et al. Updated rapid risk assessment from ECDC on the novel coronavirus disease 2019 (COVID-19) pandemic: increased transmission in the EU/EEA and the UK. *Eurosurveillance*. 2020;25(10).
46. Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *The Lancet Infectious Diseases*. 2020;.
47. Buchholz U, et al.. Modellierung von Beispielszenarien der SARS-CoV-2-Ausbreitung und Schwere in Deutschland; 2020.
48. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *New England Journal of Medicine*. 2020;382(12):1177–1179.
49. He X, Lau EH, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature medicine*. 2020; p. 1–4.
50. Moghadas SM, Shoukat A, Fitzpatrick MC, Wells CR, Sah P, Pandey A, et al. Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proceedings of the National Academy of Sciences*. 2020;117(16):9122–9126. doi:10.1073/pnas.2004064117.
51. Prem K, Cook A, Jit M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Computational Biology*. 2017;13:e1005697. doi:10.1371/journal.pcbi.1005697.
52. SEDATU-CONAPO-INEGI. Delimitación de las zonas metropolitanas de México 2015;. Available from: <https://www.gob.mx/conapo/documentos/delimitacion-de-las-zonas-metropolitanas-de-mexico-2015.html>.
53. CONAPO. Proyecciones de la Población de México y de las Entidades Federativas, 2016-2050;. Available from: <https://datos.gob.mx/busca/dataset/proyecciones-de-la-poblacion-de-mexico-y-de-las-entidades-federativas-2016-2050>

54. de Salud S. Datos Abiertos - Dirección General de Epidemiología;. Available from: <https://www.gob.mx/salud/documentos/datos-abiertos-152127>.