



OPEN

COVID-19 epidemic peaks distribution in the United-States of America, from epidemiological modeling to public health policies

Alexandre Vallée^{1✉}, Davide Faranda^{2,3,4} & Maxence Arutkin⁵

COVID-19 prediction models are characterized by uncertainties due to fluctuating parameters, such as changes in infection or recovery rates. While deterministic models often predict epidemic peaks too early, incorporating these fluctuations into the SIR model can provide a more accurate representation of peak timing. Predicting R_0 , the basic reproduction number, remains a major challenge with significant implications for government policy and strategy. In this study, we propose a tool for policy makers to show the effects of possible fluctuations in policy strategies on different R_0 levels. Results show that epidemic peaks in the United States occur at varying dates, up to 50, 87, and 82 days from the beginning of the second, third, and fourth waves. Our findings suggest that inaccurate predictions and public health policies may result from underestimating fluctuations in infection or recovery rates. Therefore, incorporating fluctuations into SIR models should be considered when predicting epidemic peak times to inform appropriate public health responses.

Understanding the predictions at the different stages of the evolution of key epidemic indicators remains a major goal for policy makers and health professionals. The dynamics of human-to-human transmission risk is associated with many factors, including response measures and other temporal factors that could affect the trajectory of the epidemic^{1,2}. It is essential to consider fluctuations of epidemiological parameters (e.g., variations in infection or recovery rates) in the modeling of the epidemic spread to better predict the epidemic peak date. It remains extremely challenging to provide an accurate epidemic scenario of an epidemic both because of the partial knowledge of the health status of the population and the variability of virus characteristics³. Previous studies investigated that COVID-19 prediction models are characterized by uncertainties resulting in fluctuating factors of the epidemic³⁻⁵. Daily fluctuations of recovery rates present a main role in peak epidemic timing based on Susceptible-Infected-Recovered (SIR) model dynamics. Using such a model with fluctuating control parameters, a previous work has shown that the infection counts follow a stochastic process with a log-normal distribution at an early stage, and that the epidemic peak is a random variable when considering that these parameters fluctuate. A form of the stochastic solution of the infection counts at an early stage of the epidemics was derived. With this framework, it has been shown that the deterministic models anticipate the epidemic peaks with respect to the stochastic model. In the latter, fluctuations of infection and recovery rates induce a realistic delay on the most probable and average date of the epidemic peak. Based on the data of the Italian regions, we previously explained in this country that the dispersion of the epidemic peak date can be modeled by including fluctuations in the control parameters in a stochastic SIR model³. This study aims to examine how daily fluctuations in infection and recovery rates affect the dynamics of a Susceptible-Infected-Recovered (SIR) model and the timing of the epidemic peak in the United States. The study will use data from three COVID-19 waves in the United States to analyze the impact of these fluctuations on the most probable and average peak time of the epidemic using a theoretical deterministic model. The basic reproduction number (R_0) will also be considered in the analysis. The results of this study will provide insight into the role of infection and recovery rate fluctuations

¹Department of Epidemiology-Data-Biostatistics, Delegation of Clinical Research and Innovation (DRCI), Foch Hospital, 92150 Suresnes, France. ²Laboratoire des Sciences du Climat et de l'Environnement, CEA Saclay l'Orme des Merisiers, UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay, IPSL, 91191 Gif-sur-Yvette, France. ³London Mathematical Laboratory, 8 Margravine Gardens, London W6 8RH, UK. ⁴Laboratoire de Météorologie Dynamique/IPSL, École Normale Supérieure, PSL Research University, Sorbonne Université, École Polytechnique, IP Paris, CNRS, 75005 Paris, France. ⁵School of Chemistry, The Center for Physics and Chemistry of Living Systems, Tel Aviv University, 6997801 Tel Aviv, Israel. ✉email: al.vallee@hopital-foch.com

on the spread and peak of an epidemic, particularly we will show that the deterministic SIR model anticipates the peak with respect to the most probable and average peak time of the stochastic model.

Methods

Theoretical modeling of the most probable epidemic peak time modeling. *Deterministic SIR model.* We build our analysis by showing that daily fluctuations in the infection and recovery rate are essential to improve prediction of the epidemic peak date and suggesting that it should be introduced in epidemiological models³. The epidemic peak date distribution allows for a new estimation of epidemic evolution using a Susceptible Infected Recovered (S.I.R) model with daily fluctuations on infection rates⁶.

The compartmental model divides the population into three groups, namely Susceptible (S), Infected (I), and Recovered (R) individuals, according to the following discrete-time evolution equations:

$$\frac{dS}{dt} = -\lambda \frac{I_t S_t}{N} \quad (1)$$

$$\frac{dI}{dt} = -\lambda \frac{I_t S_t}{N} - \beta I_t \quad (2)$$

$$\frac{dR}{dt} = \beta I_t \quad (3)$$

In the SIR model above, the parameters are the recovery rate (β), and the infection rate (λ), N is the total population.

At the beginning of the epidemic the number of susceptible people is considered constant ($S \sim N = \text{constant}$) and large with respect to the number of infected people, without parameter fluctuation, we recover the exponential growth at early stage:

$$I(t) \sim I_0 \exp((\lambda - \beta)t) \quad (4)$$

this solution shows that if $\lambda \leq \beta$ or $R_0 = \frac{\lambda}{\beta} \leq 1$ there is no epidemic outbreak, this is called the epidemic threshold and exhibit the importance of the R_0 to understand and control an epidemic dynamic⁷⁻¹⁰.

To consider time-dependent control factors, a stochastic approach is performed by which the control parameters $k \in \{\beta; \lambda\}$ are described through a stochastic process:

$$k_t = k_0(1 + \sigma_k \epsilon_{k,t}) \quad (5)$$

where ϵ is a reduced centered Gaussian random variable. $k_0 \in \{\beta_0; \lambda_0\}$ is set to the mean value of the parameter. We show that the infection counts follow a log-normal distribution and then, we can investigate the quantile of this solution¹¹. The log-normal distribution of the number of infected people implies sub-exponential divergence of the quantile of the solution from the average exponential growth behavior. Therefore, effectively managing an epidemic over a specific time frame and with a desired level of certainty should focus on managing a specific quantile of the solution Empirically we may consider the dynamic for the worst α (=95% for example) scenarios of the solution. The corresponding α -quantile q_α for the Brownian motion is:

$$q_\alpha(t) = \sqrt{2t} \operatorname{erf}^{-1}(2\alpha - 1) \quad (6)$$

Thus, the quantile of the number of infected people reads:

$$I_\alpha(t) = I_0 \exp\left(\left(m - \frac{\tilde{\sigma}^2}{2(1+m)^2}\right)t + \frac{\tilde{\sigma}}{(1+m)} q_\alpha(t)\right) \quad (7)$$

The log-normal distributions are positively skewed with long right tails due to low mean values and high variances in the random variables. This feature creates a balance between highly diffuse behavior at short time and drift domination at large time. A non-trivial time analogous to the time horizon appears canceling the exponent:

$$T^*(\alpha) = \frac{4 \frac{\tilde{\sigma}^2}{(1+m)^2} \operatorname{erf}^{-1}(2\alpha - 1)^2}{2m - \frac{\tilde{\sigma}^2}{(1+m)^2}} \quad (8)$$

with $m = \lambda_0 - \beta_0$ and $\tilde{\sigma} = \sqrt{\beta_0^2 \sigma_\beta^2 + \lambda_0^2 \sigma_\lambda^2}$.

Thus, numerical experiments are performed by discretizing a SIR model defined previously with an Euler scheme and a time step $\Delta t = 1$ day following the guideline defined by Faranda and Alberti⁴.

Most probable date of the epidemic peak. The incorporation of fluctuating parameters in the SIR model introduces a level of randomness to the solution, making the predicted epidemic peak date itself a random variable We will use tool borrowed from first passage time theory to compute the probability distribution of the epidemic peak date, the first passage modeling has shown to be ubiquitous in nature: diffusion-limited growth¹², neuron firing¹³, survival probability of a noble's man name (male descendent)¹⁴, or the triggering of stock options¹³. At the epidemic peak, the number of infected people has reached its maximum, as an approxi-

mation, we will use the deterministic peak level from the SIR model and compute the random time at which the number of infected people reaches that level.

Assuming that the number of infected people at $t=0$ is 1, for the deterministic SIR, an approximation of the number of infected people at the epidemic peak is^{10,15}:

$$I_{peak} \sim N\left(1 - \frac{1 + \log(R_0)}{R_0}\right) \quad (9)$$

using this approximation, the deterministic peak from the SIR model occurs at:

$$t_d = \frac{\theta}{m} \quad (10)$$

with $\theta = \log(I_{peak})$ and $m = \beta_0(R_0 - 1)$. Now let's assume that we introduce the fluctuation on the control parameters of the epidemic, the time at which the number of infected people will reach the level defined by Eq. (9) is a random variable defined as:

$$t_{peak} = \inf_t (at + bW_t = \theta) \quad (11)$$

with $a = (m - \frac{\sigma^2}{2(1+m)^2})$, $b = \frac{\sigma}{(1+m)}$, and W_t the Brownian motion. The probability distribution of this random variable is equivalent to the first passage time distribution of a lognormal process to a given threshold¹⁶:

$$f(t_{peak}) = \frac{\theta}{b\sqrt{2\pi}t^3} \exp\left(-\frac{(\theta - at)^2}{2b^2t}\right) \quad (12)$$

this time strongly depends on R_0 and is delayed with respect to the deterministic peak time due to the fluctuations. Another interesting time to study is the most probable date of the epidemic peak:

$$t_{mp} = \frac{3b^2}{2a^2} \left[\sqrt{1 + \frac{P_e^2}{9}} - 1 \right] \quad (13)$$

where $P_e = \frac{2a\theta}{b^2}$ is the Péclet number of the model. When fluctuation play a major role— $R_0 \approx 1$ —the most probable date for the epidemic peak scale like $t_{mp} = \frac{t_{mean} P_e}{6}$ with $P_e \rightarrow 0$, and could be much smaller than average peak date of the epidemic. It is true particularly for R_0 close to 1.

Thus, the epidemic peak time of the stochastic solution exhibits an inverse Gaussian probability distribution, that we will use to fit the spread of the epidemic peak times observed across the different regions/states³.

We recall that the probability distribution of the epidemic peak time is described by the following distribution:

$$f(t_{peak}) = \frac{\theta}{b\sqrt{2\pi}t^3} \exp\left(-\frac{(\theta - at)^2}{2b^2t}\right) \quad (14)$$

with $a = (m - \frac{\sigma^2}{2(1+m)^2})$, $b = \frac{\sigma}{(1+m)}$, $m = \beta_0(R_0 - 1)$, β_0 is the recovery rate and σ the amplitude of daily fluctuations of the control parameters (i.e., variations of infection rates or recovery rates). As shown above, this probability distribution one can easily get the most probable peak time, mean peak time and the confidence intervals of the epidemic peak time³.

Data analysis. Epidemic peak time, when the outbreak reaches its highest point, is crucial for controlling the spread of the disease. From a modeling perspective, the distribution of epidemic peak time can be derived analytically using the following approximations: we assume that the epidemic peak time is determined by a drifted lognormal distribution to the deterministic peak level, as seen above, for full probability distribution; the average epidemic peak time t_{mean} and the most probable epidemic peak time t_{mp} (for three waves in the United States of America) are derived analytically. The epidemic peak is delayed due to control parameters fluctuations with the SIR model (with conditions specified as above and $1.1 < R_0 < 2$) and from the analytical predictions. To compare our theoretical model to real data, we consider the U.S. states' infection counts. The data that support the findings of this study are openly available in <https://github.com/CSSEGISandData/COVID-19>. For each state, each epidemic wave started after the lowest number of infections between the previous wave to the next wave. The empirical distribution function of the epidemic peak time for each wave is fitted using maximum likelihood estimates of the theoretical epidemic peak time distribution defined above. Population of each US states came from: https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html#par_textimage. Analytic models were performed using MATLAB software.

Results

Figure 1 showed the different epidemic peak distributions according to the days from the beginning of a wave with different R_0 comprising between 1.1 and 1.7 based on the theoretical model. Figure 2 showed the median epidemic peak time for each wave and 90% confidence interval according to R_0 level based on analytical predictions and assuming control parameters daily fluctuations (as variations in infection rate or recovery rates) of 20%, with $R_0 \sim 1.7$ for the second wave and $R_0 \sim 1.4$ the third and fourth waves. The closer the R_0 fluctuations are to 1.0, the greater the epidemic peak distribution will be large, with a larger confidence interval, with 275 days for R_0 equal to 1.1 and with less than 50 days for $R_0 < 1.7$. We observed that the peak across US states, for the second

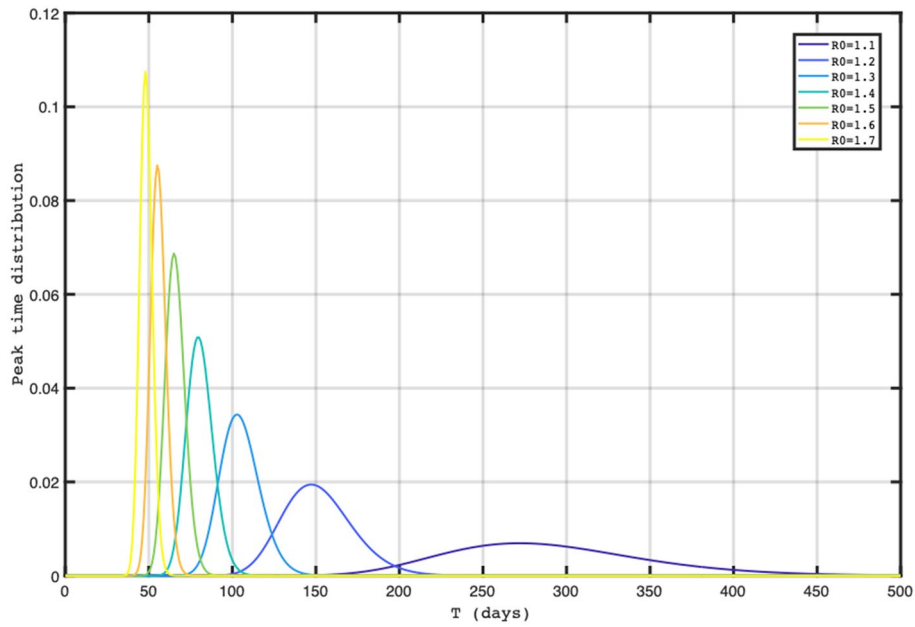


Figure 1. It illustrates the distribution of epidemic peaks at various intervals from the start of an outbreak, based on a theoretical model that considers R_0 values ranging from 1.1 to 1.7. The figure displays how the peak of the epidemic may vary depending on the number of days since the beginning of the outbreak and the corresponding R_0 value.

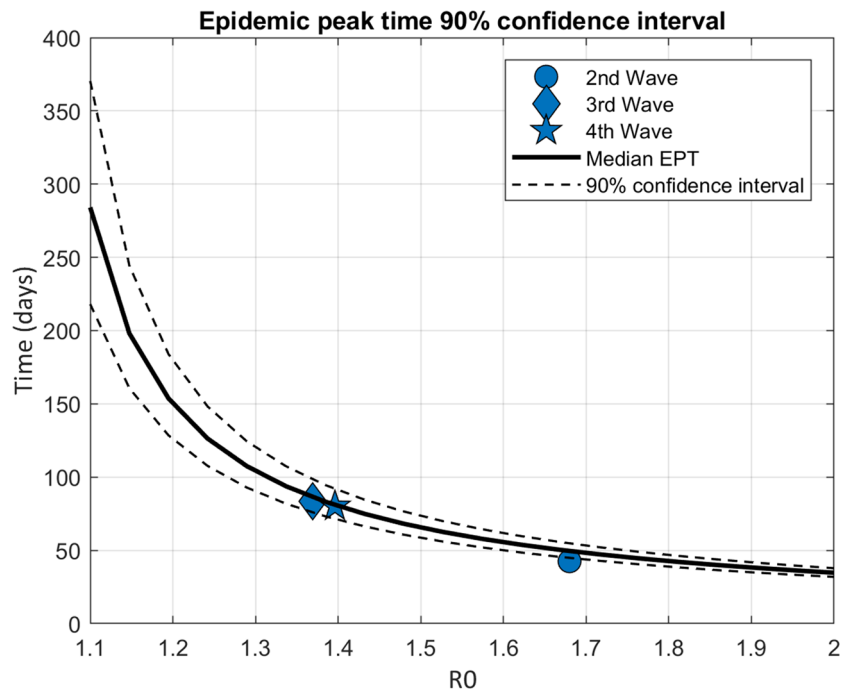


Figure 2. Median epidemic peak time for each wave and 90% confidence interval according to R_0 level assuming control parameters daily fluctuations of 20%. EPT: epidemic peak time.

wave, was distributed around 50 days (most probable peak date) (SD 24 days) (Fig. 3A), for the third wave around 87 days (SD 26 days) (Fig. 3B), and for the fourth wave around 82 days (SD 15 days) (Fig. 3C).

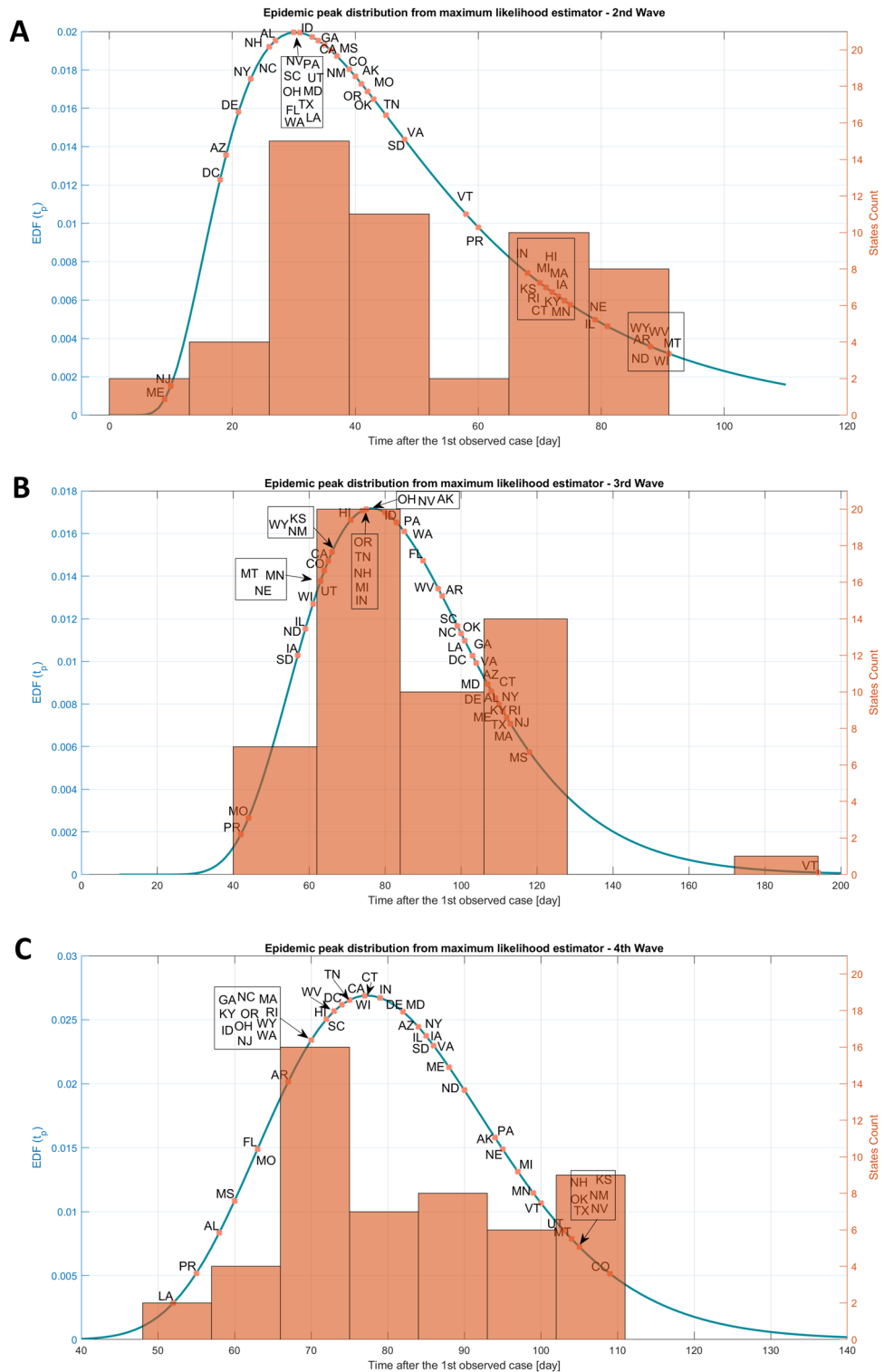


Figure 3. Epidemic peak distribution for the three last waves of the U.S. states, represented by alpha codes. Epidemic peak time is computed starting the time after the minimum of infection counts between two waves. The empirical distribution function of the epidemic peak is displayed in continuous lines, an histogram of the U.S. state counts date is shown in orange. (A) Epidemic peak distribution from maximum likelihood estimator for the second wave. (B) Epidemic peak distribution from maximum likelihood estimator for the third wave. (C) Epidemic peak distribution from maximum likelihood estimator for the fourth wave.

Discussion

For policymakers and health professionals, forecasting key pandemic indicators in the short term, such as the reproduction number (R_0) and the number of new cases, is a crucial goal. Accurate predictions can inform the implementation of effective response strategies and aid in the more efficient allocation of resources. The trajectory of the COVID-19 pandemic depends on a number of factors, including the attributes of the virus (such as its transmissibility), the characteristics of the location (such as population density and transportation patterns), individual behaviors in response to the pandemic, and government actions^{1,17–19}. Understanding how these factors influence the spread of the disease is essential for policymakers and health professionals as they work to develop effective strategies for managing the pandemic. By analyzing the interplay between these factors and the trajectory of the pandemic, policymakers and health professionals can better understand the drivers of the disease's spread and develop more effective response strategies.

These factors are correlated with a more linear growth of pandemics²⁰ but were still investigated in dynamic models of the COVID-19 transmission. Our study showed that epidemic peaks across the US states during the second, third and fourth waves were distributed around 50, 87 and 82 days from the beginning to the peak of the epidemic wave (Fig. 3). Thus, fluctuations in models should be considered in the epidemic modeling to predict the epidemic peak and plan appropriate public policies.

Epidemiological implications for public health policies. The epidemic peaks distributions are genuine features of the COVID-19 epidemic, and they originate from the combination of initial conditions (the health status of the population at the beginning of an epidemic wave) and the inherent fluctuations of the parameters that, in the SIR model, can be represented by stochastic fluctuations. Our previous study reported that epidemic peaks across the Italian region during the first wave were distributed around 55 days (most probable peak date), and around 130 days for the second wave³. Other models have shown a 6-week (~42 days) errors for cumulative death below 10%¹⁹, a median absolute percentage error at 10 weeks (~70 days) of forecasting COVID-19 resurgence for the Institute for Health Metrics and Evaluation (IHME) SEIR model²¹.

Many hypotheses have been mentioned to explain the divergence between the predicted and observed epidemic peaks, many inaccuracies and incompleteness of available information²², difficulties in confirming large numbers of cases by specific tests, presence of asymptomatic cases and possible delays in diagnosis, lack of testing, individual behavioral responses²³, seasonality, meteorological factor¹⁷, variant spread²⁴, worse the situation for modeling the different scenarios.

Uncertainties in predicting the peak of a pandemic can affect the efficacy of health policies and strategies. These uncertainties may be due to errors in long-term forecasts and to variations in the mechanisms of viral transmission. The expected exponential growth of transmission may not always be observed, which can be attributed to government interventions as well as individuals' reactions to the epidemic, such as self-isolation and practicing social distancing. These behavioral responses are indeed associated with a sub-exponential growth of epidemics^{20,23}. However, these observations remain difficult to implement in the dynamic modeling of SARS-CoV-2 transmission. Restrictions in activities, such as non-pharmaceutical interventions and non-physical distancing factors, may probably help to delay the epidemic peak by playing a part in mitigating potential spikes in cases, especially when physical distancing measures are relaxed²⁵. However, this work included limitations with large uncertainties for the estimate R_0 .

The validity of such claims depends on the evidence to support the hypotheses regarding the impact of a policy on transmission¹⁸. Different dynamics can interact with these models and impact the predicted gross scale of the epidemic. Indeed, the public health policies, including precautionary measures and quarantine, modulate the possible trajectories of outbreak²⁶. Uncertainty in peak and date sizes can be due to numerous factors, including stochasticity of early dynamics, heterogeneity of contact profiles, spatial variation, and dynamics of epidemiological parameters⁸. While strong control policies have been associated with inmate growth in cases where house-stay restrictions were unlikely to be the one-size-fits-all agreement, a gradual approach to restrictive measures could be of concern²⁷.

The prediction of R_0 remains a major epidemiological challenge with practical consequences due to it supports governments policies to develop rapid strategies to counteract the growth of the outbreak. In this study, we propose a policy-maker tool which show the consequences of possible fluctuations in policy strategies on different R_0 levels (Fig. 2)²⁸.

Limitations. Our data were provided from a public data source and thus, were limited to the accuracy of their report.

Conclusion

Our study suggests that the distribution of epidemic peaks across different regions of the United States during each wave is not solely determined by the mean infection and recovery rates, but also by the fluctuations in these rates. This is an important finding as it highlights the need to consider fluctuations in predictive models and public health policies in order to have a more accurate prediction of the epidemic peak time.

Inaccurate predictions of both epidemic scenarios and public health policies could be the consequence of an underestimation of these fluctuations. This means that without considering the fluctuations in the infection and recovery rates, the predictions of the epidemic peak time could be incorrect. This could lead to inadequate and ineffective public health policies, which in turn can lead to a failure in controlling the spread of the disease.

To address this issue, our study proposed a policy-maker tool that incorporates fluctuations in R_0 into predictive SIR models. This tool could be a valuable resource for policymakers in developing rapid strategies and implementing appropriate public health policies in response to outbreaks. By incorporating fluctuations in R_0 ,

the tool would provide policymakers with a better understanding of the potential impacts of different policy strategies on the spread of a disease. This could allow them to make more informed decisions and effectively allocate resources to prevent or mitigate outbreaks.

Overall, we emphasize the importance of considering the fluctuations in infection and recovery rates in order to have a more accurate prediction of the epidemic peak time, and thus to have a better epidemic control. This is crucial for policymakers and health professionals to develop effective response strategies and allocate resources in a more targeted way.

Data availability

The data that support the findings of this study are openly available in <https://github.com/CSSEGISandData/COVID-19>.

Received: 5 November 2022; Accepted: 14 February 2023

Published online: 27 March 2023

References

- Riou, J. & Althaus, C. L. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro Surveill.* **25**, 2000058 (2020).
- Vallée, A. Heterogeneity of the COVID-19 pandemic in the United States of America: A geo-epidemiological perspective. *Front. Public Health* **10**, 818989 (2022).
- Vallée, A. Underestimation of the number of COVID-19 cases, an epidemiological threat. *Epidemiol. Infect.* **150**, e191 (2022).
- Arutkin, M., Faranda, D., Alberti, T. & Vallée, A. Delayed epidemic peak caused by infection and recovery rate fluctuations. *Chaos* **31**, 101107 (2021).
- Faranda, D. & Alberti, T. Modeling the second wave of COVID-19 infections in France and Italy via a stochastic SEIR model. *Chaos* **30**, 111101 (2020).
- Faranda, D. *et al.* Asymptotic estimates of SARS-CoV-2 infection counts and their sensitivity to stochastic perturbation. *Chaos* **30**, 051107 (2020).
- Alberti, T. & Faranda, D. On the uncertainty of real-time predictions of epidemic growths: A COVID-19 case study for China and Italy. *Commun. Nonlinear Sci. Numer. Simul.* **90**, 105372 (2020).
- Dietz, K. The estimation of the basic reproduction number for infectious diseases. *Stat. Methods Med. Res.* **2**, 23–41 (1993).
- Anderson, R. M., Heesterbeek, H., Klinkenberg, D. & Hollingsworth, T. D. How will country-based mitigation measures influence the course of the COVID-19 epidemic?. *Lancet* **395**, 931–934 (2020).
- Obadia, T., Haneef, R. & Boëlle, P.-Y. The R0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC Med. Inform. Decis. Mak.* **12**, 147 (2012).
- Bailey, N. T. J. *The mathematical theory of infectious diseases and its applications*. 2nd edition. The mathematical theory of infectious diseases and its applications. 2nd edition. (1975).
- Bouchaud, J.-P. Elements for a theory of financial risks. *Physica A* **263**, 415–426 (1999).
- Krapivsky, P. L., Redner, S. & Ben-Naim, E. *A Kinetic View of Statistical Physics*. (Cambridge University Press, 2010). <https://doi.org/10.1017/CBO9780511780516>.
- Gerstein, G. L. & Mandelbrot, B. Random walk models for the spike activity of a single neuron. *Biophys. J.* **4**, 41–68 (1964).
- On the Probability of the Extinction of Families. | BibSonomy. <https://www.bibsonomy.org/bibtex/ce44ef06781c9e8b0ff642e736a1afa2>.
- Cadoni, M. How to reduce epidemic peaks keeping under control the time-span of the epidemic. *Chaos Solitons Fractals* **138**, 109940 (2020).
- Redner, S. *A Guide to First-Passage Processes*. (Cambridge University Press, 2001). <https://doi.org/10.1017/CBO9780511606014>.
- Ma, Y., Pei, S., Shaman, J., Dubrow, R. & Chen, K. Role of meteorological factors in the transmission of SARS-CoV-2 in the United States. *Nat. Commun.* **12**, 3602 (2021).
- Friedman, J. *et al.* Predictive performance of international COVID-19 mortality forecasting models. *Nat. Commun.* **12**, 2609 (2021).
- Zucman, D., Fourn, E. & Vallée, A. The COVID-19 vaccine health pass fraud in France. *Clin. Microbiol. Infect.* <https://doi.org/10.1016/j.cmi.2022.04.006> (2022).
- Jentsch, P. C., Anand, M. & Bauch, C. T. Prioritising COVID-19 vaccination in changing social and epidemiological landscapes: a mathematical modelling study. *Lancet Infect. Dis.* **21**, 1097–1106 (2021).
- IHME COVID-19 Forecasting Team. Modeling COVID-19 scenarios for the United States. *Nat. Med.* **27**, 94–105 (2021).
- Karatayev, V. A., Anand, M. & Bauch, C. T. Local lockdowns outperform global lockdown on the far side of the COVID-19 epidemic curve. *PNAS* **117**, 24575–24580 (2020).
- Hodcroft, E. B. *et al.* Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **595**, 707–712 (2021).
- Prem, K. *et al.* The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *Lancet Public Health* **5**, e261–e270 (2020).
- Nishiura, H. *et al.* Modelling potential responses to severe acute respiratory syndrome in Japan: the role of initial attack size, precaution, and quarantine. *J. Epidemiol. Community Health* **58**, 186–191 (2004).
- Studdert, D. M. & Hall, M. A. Disease control, civil liberties, and mass testing—Calibrating Restrictions during the Covid-19 pandemic. *N. Engl. J. Med.* **383**, 102–104 (2020).
- Tkachenko, A. V. *et al.* Time-dependent heterogeneity leads to transient suppression of the COVID-19 epidemic, not herd immunity. *Proc. Natl. Acad. Sci. USA* **118**, (2021).

Author contributions

Conceptualization, A.V., M.A., D.F.; methodology, A.V., M.A., D.F.; formal analysis, A.V., M.A., D.F.; writing—original draft preparation, A.V., M.A., D.F.; The authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023