

The importance of model assumptions in estimating the dynamics of the COVID-19 epidemic

Valery Forbes

▶ **To cite this version:**

Valery Forbes. The importance of model assumptions in estimating the dynamics of the COVID-19 epidemic. 2021, pp.100004. 10.24072/pci.mcb.100004 . hal-03300762

HAL Id: hal-03300762
https://hal.science/hal-03300762

Submitted on 27 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Peer Community In Mathematical & Computational Biology

The importance of model assumptions in estimating the dynamics of the COVID-19 epidemic

Valery Forbes *based on reviews by Bastien Boussau and 1 anonymous reviewer*

A recommendation of:

Estimating dates of origin and end of COVID-19 epidemics

Thomas Bénéteau, Baptiste Elie, Mircea T. Sofonea, Samuel Alizon *(2021), medRxiv, 2021.01.19.21250080, ver. 3 peer-reviewed and recommended by Peer Community in Mathematical and Computational Biology* ***10.1101/2021.01.19.21250080***

Open Access

Submitted: 23 February 2021, Recommended: 04 July 2021

Cite this recommendation as:
Valery Forbes (2021) The importance of model assumptions in estimating the dynamics of the COVID-19 epidemic. *Peer Community in Mathematical and Computational Biology, 100004.* ***10.24072/pci.mcb.100004***

Published: 27 July 2021

Copyright: This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nd/4.0/

Recommendation

In “Estimating dates of origin and end of COVID-19 epidemics”, Bénéteau et al. develop and apply a mathematical modeling approach to estimate the date of the origin of the SARS-CoV-2 epidemic in France. They also assess how long strict control measures need to last to ensure that the prevalence of the virus remains below key public health thresholds. This problem is challenging because the numbers of infected individuals in both tails of the epidemic are low, which can lead to errors when deterministic models are used. To achieve their goals, the authors developed a discrete stochastic model. The model is non-Markovian, meaning that individual infection histories influence the dynamics. The model also accounts for heterogeneity in the timing between infection and transmission and includes stochasticity as well as consideration of superspreader events. By comparing the outputs of their model with several alternative models, Bénéteau et al. were able to assess the importance of stochasticity, individual heterogeneity, and non-Markovian effects on the estimates of the dates of origin and end of the epidemic, using France as a test case. Some limitations of the study, which the authors acknowledge, are that the time from infection to death remains largely unknown, a lack of data on the heterogeneity of transmission among individuals, and the assumption that only a single infected individual caused the epidemic. Despite the acknowledged limitations of the work, the results suggest that cases may be detected long before the detection of an epidemic wave. Also, the approach may be helpful for informing public health decisions such as the necessary duration of strict lockdowns and for assessing the risks of epidemic rebound as restrictions are lifted. In particular, the authors found that estimates of the end of the epidemic following lockdowns are more sensitive to the assumptions of the models used than estimates of its beginning. In

summary, this model adds to a valuable suite of tools to support decision-making in response to disease epidemics.

References

Bénéteau T, Elie B, Sofonea MT, Alizon S (2021) Estimating dates of origin and end of COVID-19 epidemics. medRxiv, 2021.01.19.21250080, ver. 3 peer-reviewed and recommended by Peer Community in Mathematical and Computational Biology. https://doi.org/10.1101/2021.01.19.21250080

Reviews

Toggle reviews

Revision round #1

2021-04-22

Author's Reply

Download author's reply (PDF file)Download tracked changes file

Dear Dr Forbes,

We read the reviewer's comments and suggestions with great interest. We thank the reviewers for their suggestions, which helped to further improve our manuscript in the following ways.

First, we clarified the mathematical writing of the model, removing the unnecessary forms, and further defining the terms used. Second, we included a new analysis looking at the time until incidence reaches a threshold below which local measures would be sufficient to control the epidemic, following Dr Rousseau's suggestions. Third, we followed the reviewer's recommendations to correct for the writing mistakes, and in particular we made available the gitlab repository where all codes and raw simulation results are available.

In the response file, you will find our detailed comments to the issues raised by the reviewers. We also attach a .pdf file where the changes made are highlighted.

We herewith resubmit our manuscript and we hope that it is now acceptable for publication.

Awaiting your decision,

Thomas Bénéteau, Baptiste Elie, Mircea T. Sofonea, Samuel Alizon

Decision round #1

Dear Authors,

We have received two very thoughtful and detailed reviews of your manuscript. I would ask that you revise your preprint and indicate in a separate document how you have addressed each of the reviewers' comments. We look forward to receiving your revised preprint.

Sincerely,

Valery Forbes

*Preprint DOI: **10.1101/2021.01.19.21250080***

PCI Math & Comp Biol

Reviewed by anonymous reviewer, 2021-03-26 05:39

Download the review (PDF file)

Reviewed by Bastien Boussau, 2021-04-22 07:38

Bénéteau et al. investigate the estimations by several models of the dates of the beginning and the end of the SARS-CoV-2 epidemic in France. This is a difficult problem as the number of infected people on both tails of the epidemic is low, meaning that assumptions at the heart of commonly-used SIR-based deterministic models become inappropriate. They propose a new stochastic model, a version of which includes superspreaders, and compare the estimates of this model to a deterministic SIR-like model and to another published deterministic model that includes age stratification. They find that estimates of the end of the epidemic following lockdowns are more sensitive to the assumptions of the models used than estimates of its beginning.

General comments

The manuscript was most of the time clearly written and easy to follow. However, some figures were difficult to interpret, and in some cases the description of the results seemed to include mistakes (see specific comments). In spite of these mistakes, the results appeared convincing. I could not find links to the data or the implementation of the models to reproduce the results. Finally, I believe the discussion could be extended a bit as I explain below.

The reliance on several models allows for testing the influence of different factors, including superspreaders, age structure, and memory in the time from hospitalization to death. However, these models all rely on different implementations, and differ in several respects, making their comparisons difficult. It might have been cleaner to use one framework to implement all models and compare them by changing one parameter at a time; for instance, some Bayesian models that have been proposed in the literature on SARS-CoV-2 might be amenable to such an investigation. Nonetheless, the fact that the different models agree in a lot of their predictions suggests that the results would probably have been the same, and the reliance on several implementations also protects against implementation-specific bugs.

Among the results that stand out is the fact that several months of lock-down are necessary to reach extinction of the epidemic. This is not unexpected, but the relevance of it to public health is little discussed in the manuscript. In two places the authors mention "an audience not familiar with stochasticity"; if this means e.g. public health officials or the general public, then more discussion should be included. In particular, I believe that the relationship between the authors' result and the feasibility of the "zero-Covid" strategy should be discussed, as a cursory reading of the manuscript may be interpreted as an argument against the strategy.

Along similar lines, it seems a bit much to ask of a lock-down that it brings an epidemic to its extinction, especially when the epidemic is tackled a bit late. Would a different objective, i.e. that of reaching daily incidence levels that are compatible with a zero-covid-like strategy (control points, local lock-downs) also require several months of lock-down? Would the modeling approach proposed by the authors suffice to answer such a question, if the data are available?

Specific comments

p3: "Finally, we analyse a classical deterministic Markovian model, which is commonly used to analyse COVID-19 epidemics [?]." : missing reference

p4: "(see Figure S6)" : this is the first reference to a figure; it would probably make sense that this is Fig. S1, not S6.

p4: "a value much higher than the outbreak threshold above which a stochastic fade out is unlikely [10]": the number of daily deaths is not directly comparable to the outbreak threshold values provided in the reference cited. It would be convenient for the reader to detail the computations that ensure that the value chosen is much higher than the outbreak threshold.

Table S1: "Shape parameter (Gamma distribution)" : in this table, could the reader be reminded that the Gamma distribution is used to model heterogeneity in infectivity and/or infection duration?

Supp mat p3: "where η n measures the public health intervention impacts on the disease spread at day n,": for consistency with the stochastic model, perhaps it would be clearer to use t for the day?

Fig. S1: the legend to this figure should at least explain the meaning of the compartments, and possibly the parameters.

Supp mat p3-4: "We compared this model to the discrete time non-markovian model, and a SEAIRH4D model in which memory in the delay from hospitalization to death is implemented" : I find this description too short to really understand what was done, and the meaning of the acronym SEAIRH4D should be provided.

Supp mat p3: "The set of ODE shown in the previous paragraph is solved

using 'odeint' function from Numpy on Python 3.8.3.": Is the code for the deterministic models available? If so it could be stated here.

Supp mat p3: "We estimated the following parameters for the

SEAIRHD model using a maximum likelihood procedure" : could the authors provide the likelihood formula and specify what algorithm was used to maximize the likelihood?

Figure S4: "Generation time standard deviation impact on the starting date inference.": there is an inconsistency between the y axis that states "Serial interval standard deviation" and the legend.

Figure S5: I assume a serial interval of 2.3 was used? It would be useful to point it out.

Supp mat p7: "We can see that only the importation of

new infected individuals during the first days has an impact on the epidemic.": I do not understand how this conclusion is reached: is it by comparison of Figs. S4 and S5? I would need more details on the reasoning and possibly another figure to understand this.

p5: "with an estimated efficacy of 1 -\eta_{FR}= 76% [21]." : it would be good to define \eta_{FR} here rather than a few lines later.

p6: "finite lock-down extensions on the the probability": too many "the"s

p6: So \tau is defined per simulation, and p_0(t) is averaged over all simulations?

p6: "SEAIRHD" : This model does not include the possibility that asymptomatic individuals become recovered without ever becoming symptomatic, which is a big feature of Covid. Could the authors comment on the expected importance of the lack of such a feature?

p6: "Scripts for the SEAIRHD model can be found in the supplementary materials.": I have not found them.

p7: "the same as in our model" : the same as in our DS model

p7: "The likelihood of the deterministic SEAIRHD model was computed assuming a Poisson distribution of the daily mortality incidence data." : I think it would be good to explain how parameter inference was achieved with the non-Markovian deterministic model.

p7: "the time mortality incidence reaches" : I think it would help to remind the reader that this date is March 23.

p7: "67 days (equivalent to a first case on January 16 in France), with a 95% confidence interval (95% CI) between 62 and 79 days" : the numbers given in this section do not seem to match Fig. 1 "DS without heterogeneity". Was there an inversion in the names of the violin/boxplots between with and without heterogeneity?

p8: "However, consistently with earlier studies [21?]" : missing reference

p8: "the median delay for daily incidence to reach 100 deaths is decreased by 5 days when the serial interval standard deviation is decreased by one third (Fig. S4).": isn't it the opposite?

p8: "However, when assuming a more realistic scenario where all those cases are not imported on the same day, we find a much more limited impact on the delay" : I find it hard to be convinced, looking at the figures and trying to compare the two panels of Fig. S5. Could the authors provide trends or numbers, or maybe an additional supplementary figure, that would precisely convey this information?

p9 "Time to eradication": in this section a few comments about the results of the SEAIRHD model would be useful.

p10: "The results are shown in Figures 3 for the case without host heterogeneity and Fig. S8 with super-spreading events." : it is not clear to me why the authors chose to show the results of the superspreading model in supplementary material and the results of the model without superspreading in main? I would have expected the reverse.

p12: "as stressed by earlier studies [21?]." : missing reference

p13: "higher k parameter value that the one used here (0.30 versus 0.16 here)" : than instead of that