Review of: "Evolution of new variants of SARS-CoV-2 during the pandemic: mutation-limited or selection-limited?"

Dominik Sziklai[1]

1 Semmelweis University

Potential competing interests: No potential competing interests to declare.

Brief summary:

The topic of the paper on a global scope is viral evolution, on a local scope it focuses on the recent SARS-CoV-2 pandemic. The paper proposes multiple models which could explain the dynamics of the ascension of new variants and how these new variants can initiate new waves. The three possible hypothesis is about whether the emergence of a new variant is mutation- or selection-limited. The authors give us a brief introduction on the currently trending mathematical models of pandemics like SIR and other models that incorporate immunity as a continuous variable. The authors collected data from reliable sources and run their own simulations on the available data. They tested the correlation of a newly appearing variant with the ascension of a new wave (measured in incidence). The text revealed interesting insights about the emergence of the detected new variants. Data suggests that the emergence of new variants are not correlated with the wave induced by other variants. The possible biases and limitations due to the data acquisition and limited data availability and the impact of these biases are described. In the last paragraph, the authors discuss their findings and draw their conclusions. As they suggest we can discard the mutation-limited theorem and that the dynamics of immunity must play a huge role in the formation of waves, however our knowledge of this matter is rather limited. Immunity itself is the main component of the selection limitation. There can be an evolutionary explanation about how human immunity evolved differently against different viruses and ultimately, since the mutation rates of different viruses are similar, the infectivity and transmissibility of different viruses might be the result of the differentiated immune responses. At the end, multiple practical epidemiological questions are raised, these are yet to answer, but they are relevant and important to ask.

Opinion:

- It is delightful that the authors focus on the dynamics of the pandemic rather than doing just time-series forecasting. Each time series of viral infection is the manifestation of millions of small-scale biophysical interactions. Identifying each of

these interactions is out of our reach, and the article emphasizes the limitations which we face when we attempt to do evaluations on the data.
- The article is thoughtful and informative, my issues is mainly with the stylistic side and some minor ambiguities.
- The article introduces briefly the mutation and selection limited way of variant ascension. However, this could be done with a bit more explanation and more emphasis on the differences between the two phenomena.
- The definitions of different time-series components are understandable and essential to understand the implemented model.
- In many paragraphs of the paper, we can read listings of different parameters or equations (especially when discussing the SIR model!), it would be nice to see these mathematical formulas and other possible listings of parameters in a properly structured format.
- The paper suggests that practically we should revise the theory of pandemics and the different mitigation processes. It would be interesting to see some possible suggestions with which humanity could counter the selection-limited viral pandemics.
- The paper ruled out Hypothesis 1. as a possible correct explanation for the dynamics, and later suggested that Hypothesis 3. might be valid. However, there was a Hypothesis 2. which was just to explain the possible biophysical/biological randomness. Hypothesis 2. was dropped under Fig. 6, but it was not explicitly written, that at this point Hypothesis 2. was dropped, which can be misleading.

Suggestions:

- Structured mathematical formulation of the SIR model. (Latex format maybe…)
- To make the figures more understandable for the audience, please label the the curves with arrows and text especially on Fig. 4. (this figure was the most ambiguous and hard to comprehend for me)
- x and y axis labels should be present on Fig. 5.
- Please explicitly write out when you drop Hypothesis 2.
- At least at the discussion section, compare the evidences pro/against the 3 hypotheses in a tabular and more follow-able format.
- Correct typos in Fig. 1.

Questions:

- Dropping Hypothesis 2. would mean that you drop the randomness factor out, however, as stated in the article, the data statistics is not robust enough to allow us

to draw such conclusions. The same goes for Hypothesis 1. Mathematically, wouldn't it be better to restate your conclusion in a way that all three Hypothesis might play a role, at least we should weight them differently according to specific conditions (population density, geolocation etc…)? What is your opinion about a mixed model?

- The article defines the Data Segment used as a two-week period. Is this resolution enough to draw reliable statistical conclusions?

- In your article, you state that there is a probability that the main paradigm (mutation-limited variants) of epidemic mitigation is flawed. If this is true, could you suggest any kind of practices that WHO policies could adapt to fight against epidemics according to the selection-limited paradigm?