Qeios

Peer Review

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

**William Gallaher**[1]

1. LSU Health New Orleans, New Orleans, United States

This is a well written and thoroughly explained paper on an important topic. Given the heavy global impact of COVID, examination of the premises on which decisions are made for control of future outbreaks could not be more critical. An unbiased eye on the dynamics of the pandemic in the human population, with respect to viral mutations and immune selection, using well developed methodology, makes this paper a very interesting one.

In particular, the lengthy introduction made very clear the hypotheses to be examined and the alternatives to be tested. The authors developed a very clear roadmap.

Their consideration of 125 precisely defined “waves” or outbreaks in several different countries was a particularly attractive and detailed approach to the question addressed. I agree with several of their conclusions. Critical variants did not develop at peaks of the outbreaks. In several cases, outbreaks did not coincide with the arrival of new variant, but were clearly independent of variant development. The overall picture, that the global outbreak was an averaging of many quite different country-by-country outbreaks throughout the years of 2020 through 2022 is a correct one. There are well chosen examples illustrating this major point. Also, that the emergence of variants occurred in parallel with the pandemic and was in many cases uncoordinated with the case incidence. For the reasons they expound, as well as others, very strong selection was the driving force behind the emergence of variants during the pandemic.

I am especially fond of the concept behind Figure 6, directly comparing the course of the pandemic in 11 different countries where 16 different variants are said to have occurred. It does need some work, and I will be very specific in recommending changes below to this very important figure in the final version of this publication.



doi.org/10.32388/CJoACJ

Despite this broad agreement, however, in the old review system where I was the founding Deputy Editor of Virology Journal, I would have sent the paper back to the authors for significant revision in some places and technical or minor corrections in others, with an eye to finding it acceptable for publication. My comments will proceed page by page.

On page 3, in discussing the application of the SIR model, the compartments are described as "susceptible, infected, and immune or removed individuals". The authors contrast this with "immunity is a continuous variable", the measure of which is the titer of antiviral antibody, as shown later in Figure 1 on page 6. I believe this is a significant misstatement of the nature of immunity and purpose of immunization. What confers longer lasting immunity is not antibody titer but B and T memory cells, capable of producing a rapid and even higher level of immune mediators in the anamnestic response upon challenge. Without challenge, true immune status cannot be determined. This is a very important distinction that needs to be made. We do know that immunity to coronaviruses is transient. Individuals have contracted COVID up to three times in 30 months, especially nurses and respiratory therapists with high exposure. We do not know why. Antibody titers are measured because they are easy to do. There is, however, no assurance that they significantly illuminate what is going on in the more critical arenas of B and T memory cells. An alternative hypothesis is that reinfection only follows significant antigenic change in B or T cell epitopes of the S protein over time., permitting infection with high doses of inoculum that override even full immunization

Later in page 3, under A., the authors state that mutations are more likely to occur when viral population is high. In fact, each round of replication has an equal probability of producing a mutation. Several critical mutations occurred when the virus was still in China (the transposition producing a second ACGAAC splice acceptor site in N) or very early, by March of 2020 (the D614G in S), and April of 2020 (N501Y) when the total number of rounds of replication was very limited relative to later.

As important background information, mutation in SARS-CoV-2 poses serious conundrums for virologists. What we see is that mutations are heavily focused in the S gene, and even to just two regions of the S protein, the outer antigenic supersite in S1, and the areas flanking the receptor binding domain. Outside these very small regions of the genome, mutations are rare. ( https://virological.org/t/mutations-arising-in-sars-cov-2-spike-on-sustained-human-to-human-transmission-and-human-to-animal-passage/578/12 )

Another phenomenon unique to SARS-CoV2, and very unlike influenza, and even other closely related sarbecoviruses like RaTG13, is an incredibly low frequency of synonymous wobble base mutations

relative to nonsynonymous changes. Molecular evolution of SARS-CoV-2 in humans is literally unlike anything we have ever seen anywhere else. The overall picture suggests the virus is under extreme selective pressure focused on the S protein and S gene. We are seeing a highly selected subclass of the mutations that occur but do not emerge. This is consistent with the authors' conclusion about selection being the driving force, but directly observed in the viral sequence rather than inferred from the pattern of waves vs. emergence. I encourage them to note that their conclusion is consistent with this sequence data.

On page 4, under hypothesis 3, the authors introduce the concept of "partial immune evasion" as a general feature of infected individuals. This is graphically depicted in Figure 1 on page 5. The ordinate is antibody titer, and I refer back to my earlier comment that titer does not equal immunity. A window is defined for selective advantage. The time scale on the abscissa is not given, but presumably in months. In the period of 6.5 to 8.5 months after maximum antibody titer, itself at least 30 days after infection, in nearly all individuals the virus will obviously be gone or in such miniscule amounts that there is no virus to select. For selection to occur, one would need reinfection, which will surely boost antibody levels to greater than the original maximum within a few days. So, as described, this window does not exist in reality for normal individuals. I would eliminate Figure 1 and relevant discussion, and instead substitute from the following paragraph, documented as a true and not theoretical type of infected person.

There is one type of individual – not described by the authors – who not only exists but has been proposed, and in one case at least, identified, as responsible for the selection of variants. Especially variants with not one but several mutations that arise simultaneously in that variant. That person is the chronically infected host, with impaired immune responses, who is unable to eliminate the virus while at the same time producing some antibody that selects for mutations during extended replication. A Figure 1 describing this type of individual would be in order.

(https://www.frontiersin.org/articles/10.3389/fviro.2022.942555/full;
https://wwwnc.cdc.gov/eid/article/28/9/22-0875 article)

Variants have been few, in spite of truly apocalyptic levels of global viral replication, because these individuals are a very small subset of infected humans. This scenario, for which partial evidence has been published, validates the authors' alternate hypothesis of selection being the driving force, but, again, in a different way than they suggest.

Mutations occur randomly in such individuals, and emerge randomly, when a sufficient amount of virus is produced from the chronically infected, immunologically challenged individual, enough to infect

others.

In p. 9, the origin of the variant is defined as the country where the variant was first detected. Given the vagaries of testing and sequencing, there is a gap between the occurrence of the mutation and its detection. It must exceed about 10% of the virus population in an area, and the area needs to be adequately sampled, to permit such detection. This may mean that these designations are not in fact correct, and the virus may have been imported from elsewhere. Phylogenetic analysis gives an independent time estimate for origin, in terms of molecular diversion from other clades, and in all cases it significantly precedes the time of detection, in some cases like Omicron, by many months.

In Figure 3, p. 10, no units are given for either axis in either panel. The abscissa can be inferred as days, but this should be designated. The ordinates are quantitatively undeciperable. Earlier, Figure 2 was also without specification on either axis, but as a theoretical model it was acceptable there. This is a result, so the designations are necessary.

The sentence at the bottom of p. 10 begins with lower case “o”. Presumably, the preceding first letter was to be a capital “T”, and it is frankly alarming that this simple typo has gone uncorrected since the paper was posted on May 22.

This is all the more important because this paragraph is the core statement of the results of the paper. I substantively agree both with the findings and the conclusions from their analysis of these 125 waves of infection in many countries. This and the accompanying Figure 6 are what are most important in this publication.

So now we come to Figure 6, and several fixes that are needed to factually and graphically improve it.

1. Trivial but critical, the panels do not align vertically on the time scale, either on the left or right set of panels. The grid lines indicate that each panel is at a different magnification. This is unacceptable, but also an easy fix. Currently, one has to closely examine the figure to determine which waves the countries share – all have Omicron at virtually the same time – and which they do not, or if they share the same wave but slightly displaced in time, such as Delta first in India and then a bit later in Colombia, and then later in the US – in each case, setting off a major wave of infection in each country. The panels have widely different values on the ordinate, for case incidence. Just as they must be coordinate on the tme axis, they must be normalized somehow on the incidence axis, perhaps by making the Omicron peak in Dec/Jan 21/22 the same height in each, since it is a wave they all share.

2. The analysis makes all variants equal, but this is very clearly not the case. There are but five major variants corresponding to major new waves of infection, each renewing the pandemic in a major way. These are, in importance of impact, Delta, Omicron, Alpha, Gamma and Beta. All else pales, and some variants did not have a significant effect on overall morbidity or mortality in their country of origin, let alone globally. This is a critical error requiring correction, as will be noted below.

Of the 16 variants noted, 6 are of Omicron, and listed as originating in different places very close in time. This badly skews the data and conclusions. Clearly BA1, BA2 and BA3 originated in an unsampled environment in South Africa, and are recombinants of one another and an as yet unsampled precursor virus with a pre-Alpha sequence. This cluster of variants developed over many months, with an origin in early 2020, but remained hidden in coastal South Africa. (https://virological.org/t/omicron-is-a-multiply-recombinant-set-of-variants-that-have-evolved-over-many-months/775) With a large array of unique mutations, this cluster spread rapidly to many countries, overwhelming what began as a winter wave of Delta in the northern hemisphere but spreading globally. BA4,5 did develop later, but was only detected after the major and disastrous Omicron wave had subsided and anti-SARS-CoV-2 immunity greatly reinforced.

1. The dates and countries of origin are at variance with those of the WHO. D614G was detected in Italy, but Alpha was first detected in Kent, England, with its principal defining mutation N501Y, that increased transmission by 30% and mortality by 60%. Delta was first isolated in Mumbai, India, and initiated a disastrous epidemic that spread globally. Even if first detected in Slovenia, it had no impact there. But it did heavily impact India, which had been relatively unaffected until its arrival. This may be an inconvenient truth that upends in part your conclusion, but this is one major example of the arrival of a variant being the *sine qua non* for a very serious wave. Beta and Gamma and Omicron BA1 are correctly attributed.
2. I would rework the analysis using only the 5 major variants and still keep the 11 countries as an illuminating contrast in themselves. You will still find that not all waves corresponded to variant emergence in either the first country or later countries. But you will find that the 5 major changes in the virus were followed, rather immediately, by major waves. The arrival of variants is the major story emerging from the pandemic. Without the variants, the pandemic would have been over shortly after the vaccine was introduced. Immune escape, in a very limited percentage of the viral S gene, is what kept the pandemic going. The pandemic is still going. Johns Hopkins stopped collecting data in early March 2023, while in the US alone, there were nearly a million cases and

nearly 2000 deaths in the previous 4 weeks. For influenza, that would be a major epidemic inducing a significant public health response, but for a world tired of COVID, we define it as endemic.

When you focus on the major 5 variants, this upends Figure 7. Beta (5/20) was followed by Alpha (9/20) five months later, then Delta (10/20) one month later, then Gamma (11/20) one month later, and finally the Omicron cluster (11/21) 14 months after that. Most of the major variants appeared over a 7 month period in 2020. Omicron was a total shock and surprise, and especially the multiply recombinant nature of it, but phylogeny suggests it originated also in early 2020, but emerged after a long period of dormancy and recombination. (https://virological.org/t/omicron-is-a-multiply-recombinant-set-of-variants-that-have-evolved-over-many-months/775 ) All of the substantial replication in 2021 yielded no major variants at all, despite many waves in many countries.

After the experience of Omicron variants largely by recombination, we have been retrospectively sensitized to the role of recombination, rather than mutation, as a driving force in generating variants all along. Early on, there were two few mutations in the genome to make this clear. However, with the much larger number of mutations in Omicron, we now recognize the recombination has been covertly going on continuously in the evolution of the virus from the outset.

With respect to the Discussion, in view of the probable source of variants in a very small human population of chronically infected and immunologically compromised hosts, the discussion of decline in immunity is inappropriate as no role for it is demonstrated. This is especially so since antibody titer is not a good measure of immunity and also because, as I have indicated, Figure 1 does not describe a probable course of infection. Basically the discussion needs to be substantially rewritten.

The dynamics of the evolution of influenza is among the most heavily studied subjects in virology, and has, by informing the revision of vaccines through a global network of data collection, saved untold numbers of lives, even rapidly suppressing the influenza H1N1 2009 pandemic. It does not need revisiting. We know that antigenic shift follows reassortment of gene segments with novel mutations largely derived from the domestic animals that share influenza infection with humans. Genetic drift by random mutation in the human population could not be better documented.

The human immune system is maintained at high cost, deliberately. It protects us from virulent and avirulent agents alike. We have learned that virulence is less a property of the virus and more in the eye of the beholder. Comorbidities matter, and by a lot. Whether the host will respond with a cytokine storm is a function of the individual, not of the virus, and is deadly. In COVID, it is now recognized that serious

morbidity and mortality are low, the latter about 1%. Because of very serious undertesting and underdetection, we underestimated the number of persons exposed to SARS-CoV-2 by a factor of 3 to 10. Now with home testing, the underreporting of even clinically apparent cases is even greater than before. To properly assess virulence and mortality, one must have a good number for the denominator, the true incidence. The statistics on COVID in all but a few countries are junk data. The pandemic was vastly underestimated, and in many cases deliberately or by sheer negligence. We may never know its true dimensions.

In the United States, political and cultural factors had the most effect on generation of epidemic waves. There were only two clear influences on COVID in the US. As with all respiratory diseases throughout the Northern hemisphere, there is a peak in the winter months, beginning in early to late November, depending on latitude. In the US, there was a 50% difference in morbidity and mortality between regions, even in some cases between adjacent urban and suburban counties, based on one criterion alone – whether the geographical entity had voted a large majority for President Biden in the 2020 election (better medical outcome) or whether they had voted a large majority for President Trump (poorer medical outcome). The former complied with and enforced anti-pandemic measures, the latter opted for freedom – in this case the freedom of their elderly to get sick and die. The political/cultural divide in adhering to public health guidelines on gatherings, vaccine and masks determined one's risk of infection and even death. The per capita deaths in Texas and Florida, hotbeds of resistance, exceeded by 10s of thousands that in California, even though the population of California was larger than the combined population of the other two. This is an outrageous phenomenon that clearly demonstrates that populational behaviors were a major determinant in the periodicity and magnitude of waves of infection. In marked contrast, dense populations in South Korea and Taiwan, who were far more compliant, did far better in limiting the outcome of the pandemic in their countries.

Countermeasures such as shutdowns, limits on gatherings, masks, handwashing may have not worked as designed, but several results are clear. The initial isolation of Wuhan province did prevent the rapid spread of COVID to other large population centers of China, buying time and saving lives. The shutdowns did quell the initial surge of COVID in early 2020 in a then entirely naïve susceptible global population, or else the initial wave would have had the magnitude of infection not seen until alpha and delta appeared prior to the vaccine. Masks, especially common surgical masks, did not stop the virus directly, but did work as a psychological reminder to the wearer that they were in danger and to limit close transmission-effective contact with others. In the Far East they are commonly and voluntarily worn in the face of any

respiratory outbreak, with that same effect. We have learned that frequent handwashing, masks and keeping one's distance from others all work.

These considerations lie beyond the scope of your study, but clearly can be stated as "other factors" in generation of waves that did not coincide with arrival of the major variants.

Finally, one fact proves that the COVID vaccine, however faulted by excessive reactogenicity reducing acceptance, or outdated with respect to the variants it was designed against, saved an enormous number of lives. During the winter of 2021-22 and the emergence of Omicron, in the US a large majority, 75%, of hospitalized patients were unvaccinated. An even higher percentage, 90%, of those who died of COVID were unvaccinated. ( https://www.npr.org/sections/health-shots/2021/12/05/1059828993/data-vaccine-misinformation-trump-counties-covid-death-rate )

The vaccine was safe, with no documented deaths, even though toxicity was excessive, and effective. This continues a clear trend that is without exception from the 1940s, that vaccines are among the most cost effective, safest and most effective pharmaceuticals produced and distributed in mass quantities and have saved countless lives. Undermining confidence in vaccines kills people, as the fate of the unvaccinated continues to prove literally every day.

In the end, this paper deals with an important subject with a fresh perspective breaking the pandemic down into its component parts by country and each variant emergence. The review is lengthy in a desire to improve it, and the authors should not take it as a negative impression of the paper. Reviewers do not spend this much time making suggestions and critique of a paper they do not like. Rather, the intent is to help the authors circumvent and resolve issues that can indeed be readily be dealt with to bring the manuscript to its full potential.

**William R. Gallaher, Ph.D.**

*Professor of Microbiology, Immunology, and Parasitology, Emeritus*

*LSU Health New Orleans, LA*

## Declarations

**Potential competing interests:** No potential competing interests to declare.