Can we predict the limits of SARS-CoV-2 variants and their phenotypic consequences?

This is a refresh of the research analysis published on the 30th July 2021 (https://www.gov.uk/government/publications/long-term-evolution-of-sars-cov-2-26-july-2021/long-term-evolution-of-sars-cov-2-26-july-2021). In the original analysis we predicted that antigenic drift would lead to current vaccine failure (Scenario Two, point 5 – Likelihood: Almost certain. Impact: Medium). Within a few months of the report, the Omicron variant emerged, that is the farthest antigenic variant of SARS-CoV-2 from the 'Wuhan' vaccine so far. Whilst a boost of the original vaccine provided protection against severe disease with the Omicron variant, it did not prevent transmission. This should remain a long-term goal of future vaccine development and strategy. We caution that the milder symptoms in the human population, and in animal models, associated with the Omicron variant compared to previous variants, is likely a chance event. The next variant to achieve UK/global dominance is likely to have the same pathogenicity as previous variants. The loss of virulence as viruses evolve is a common misconception.

New variants do not necessarily emerge from previous globally dominant variants e.g. Omicron did not arise from Delta and new variants also do not necessarily arise from where vaccine selective pressure is greatest. Whilst focus is rightly placed on changes in the spike protein, our understanding and monitoring of changes at other sites in the genome should not be neglected as these proteins are also major determinants of transmission, fitness and disease outcome in coronaviruses. In the UK we have seen almost complete replacement of previous variants by Alpha, then Delta, then Omicron. Other parts of the world have less complete replacement, but the general trend is similar, with some co-circulation of different lineages in different parts. This is a feature that we also see in influenza, albeit with some co-circulation of different lineages.

We recommend further urgent analysis of the origin of different SARS-CoV-2 variants particularly focused on potential immune compromised hosts where longer term (persistent) infections can become established. Given the promiscuity of the virus potential zoonotic cross over events should be investigated. We have no knowledge of the long-term evolutionary trajectory of human coronaviruses and archived material/biobanks of appropriate material should be identified and analysed to help predict future scenarios. We remain concerned that a new serotype can emerge for which no current vaccine is effective. Generic anti-viral therapy and other measures should be considered to provide sufficient breathing space until a new vaccine can be manufactured and distributed.

As eradication of SARS-CoV-2 will be unlikely, we have high confidence in stating that there will always be variants. The number of variants will depend on control measures. We describe hypothetical scenarios by which SARS-CoV-2 could further evolve and acquire, through mutation, phenotypes of concern, which we assess according to possibility. For this purpose, we consider mutations in the 'body' of the virus (the viral genes that are expressed in infected cells and control replication and cell response), that might affect virus fitness and disease severity, separately from mutations in the spike glycoprotein that might affect virus transmission and antibody escape. We assess which scenarios are the most likely and what impact they might have and consider how these scenarios might be mitigated. We provide supporting information based on the evolution of SARS-CoV-2, human and animal coronaviruses as well as drawing parallels with other viruses.

Scenario One: A variant that causes severe disease in a greater proportion of the population than has occurred to date with potential vaccine escape. For example, with similar morbidity/mortality to other zoonotic coronaviruses such as SARS-CoV (~10% case fatality) or MERS-CoV (~35% case fatality). This could be caused by:

1. Point mutations or recombination with other host or viral genes. This might occur through a change in SARS-CoV-2 internal genes such as the polymerase proteins or accessory proteins. These genes determine the outcome of infection by affecting the way the virus is sensed by the cell, the speed at which the virus replicates and the anti-viral response of the cell to infection. There is precedent for Coronaviruses (CoVs) to acquire additional genes or sequences from the host, from themselves or from other viruses.

2. By recombination between two VOC or VUIs. One with high drift (change in the spike glycoprotein) from the current spike glycoprotein gene used in the vaccine and the other with a more efficient replication and transmission determined by internal genes, for example, a recombination between beta and alpha or delta variants respectively. Alternatively, recombination may occur between two different variants with two different strategies for overcoming innate immunity, combining to give an additive or synergistic change of phenotype resulting in higher replication of the virus – and potentially increased morbidity and mortality. We note that the Omicron variant has shown how a disease phenotype can alter in response to genotypic changes.

Likelihood of genotypic change in internal genes: Likely whilst the circulation of SARS-CoV-2 is high.

Likelihood of increased severity phenotype: Realistic possibility.

Impact: High. Unless there is significant drift in the spike glycoprotein gene sequence, then the current spike glycoprotein-based vaccines are highly likely to continue to provide protection against serious disease. However, an increase in morbidity and mortality would be expected even in the face of vaccination since vaccines do not provide absolute sterilising immunity, meaning they do not fully prevent infection in most individuals.

What can we do?

- Consider vaccine booster doses to maintain protection against severe disease.
- Reduce transmission of SARS-CoV-2 within the UK (to reduce risk of point mutations, recombination).
- Minimise introduction of new variants from other territories (to reduce risk of recombination between variants).
- Targeted surveillance for reverse zoonoses and consideration of strategies to prevent reverse zoonosis.
- Continue to monitor disease severity associated with variants (to identify changes in phenotype).
- Continue to develop improved prophylactic and therapeutic drugs for SARS-CoV-2 and disease symptoms.
- Consider stockpiling prophylactic and therapeutic drugs for SARS-CoV-2.

Scenario Two: A variant that evades current vaccines. This could be caused by:

Antigenic 'shift':

3. Natural recombination events that insert a different spike gene sequence (or partial sequence) from human CoVs MERS-CoV (highly unlikely due to the low frequency of MERS-CoV infections), or from currently circulating endemic human CoVs (more likely due to the prevalence of these viruses). This would recombine into the 'body' of SARS-CoV-2 that is capable of high replication in human cells. The consequence could be a virus that causes disease at a level similar to COVID-19 when it first emerged but against which our current battery of spike glycoprotein-based vaccines would not work.

   Likelihood: Realistic possibility.

   Impact: High for a completely new spike, medium/low if a spike from a seasonal CoV is introduced since we expect a proportion of the population to have antibodies to these endemic viruses.

   What could we do? In the case of introduction of a completely different spike glycoprotein, a similar vaccine platform could be rapidly employed as has been used successfully on the original Wuhan SARS-CoV-2 and subsequent variants. However, there would be a time lag for roll out whilst these vaccines were generated in sufficient quantities to control and mitigate the effects of infection.

Longer-term shift:

4. A longer-term version of shift whereby SARS-CoV-2 undergoes a reverse zoonotic event into an animal reservoir(s). This virus is then on a separate evolutionary trajectory because the virus animals is subject to different selection processes than in humans. The SARS-CoV-2 decedents then re-emerge into humans at a later time when vaccines that have been updated to keep pace with drift in humans sufficiently mismatched so as not able to provide immunologic cross protection.

   Likelihood: Realistic possibility.

   Impact: Medium.

   What could we do?: Maintain a capacity to make vaccines with updated/different spike protein variants and begin to develop broader CoV immunity in the human population to diverse coronaviruses. For example, begin to develop a universal coronavirus vaccine with strong cross protection to other CoVs potentially using other viral proteins rather than just the spike glycoprotein.

Antigenic drift:

5. A gradual or punctuated accumulation of antigenic variation that eventually leads to current vaccine failure. Worst case is that this drift combines with significant antigenic sin (vaccination resulting in an immune response that is dominated by antibodies to previously experienced viruses/vaccines) meaning that it becomes difficult to revaccinate to induce antibodies to the new strains. Genetic and antigenic drift are almost inevitable. Antigenic sin has not yet been reported for SARS-CoV-2 so we consider this possibility less likely.

   Likelihood: Almost certain. Impact: Medium.

   What could we do? Need to continue vaccinating vulnerable age groups at regular periods with updated vaccines to the dominant antigenic drift variants to increase an individual's immunological protection against SARS-CoV-2 variants.

   - Monitor antigenic variants and update candidate vaccines to cover antigenic escape variants
   - Conduct clinical trials of re-vaccination with antigenically distant vaccines
   - Consider clinical trials of multi-valent vaccines
   - Re-vaccinate vulnerable age groups at regular periods with updated vaccines to the dominant antigenic drift variants to increase an individual's immunological landscape to SARS-CoV-2 variants.
   - Reduce transmission of SARS-CoV-2 within the UK (to reduce risk of point mutations, recombination)
   - Minimise introduction of new variants from other territories (to reduce risk of recombination between variants)
   - Targeted surveillance for reverse zoonoses and consideration of strategies to prevent reverse zoonosis.
   - Continue to develop improved prophylactic and therapeutic drugs for SARS-CoV-2.
   - Stockpile prophylactic and therapeutic drugs for SARS-CoV-2.

Scenario Three: Emergence of a drug resistant variant after anti-viral strategies. This could be caused by:

6. Emergence of new variants following the administration of directly acting antiviral therapies. As we begin to use directly acting antiviral drugs it is highly likely the virus will evolve resistance to individual agents. For example, drugs that target the viral 3C protease, drugs that target the polymerase, monoclonal antibodies that target the spike glycoprotein. Resistance to some therapies has been identified in cell culture models. If the drugs are used as a mono therapy, then resistant variants have a high probability of emerging. This may render all drugs in that category unusable.

   Likelihood: Likely - unless the drugs are used correctly. Impact: medium unless a scenario arises where drugs are needed more widely.

   What can we do? Use the deep knowledge of deploying antivirals against RNA viruses such as HIV and influenza virus.

- Only use antiviral combination therapy, using ≥2 drugs with different targets or mechanism of action.
- Preserve antiviral use for an emergency situation in which a SARS-CoV-2 variant is more severe, and a matched vaccine is unavailable and takes time to develop.
- Use antivirals cautiously in immunocompromised people in whom long term evolution can happen – monitor for treatment failure and resistance, minimise risk of onwards transmission of resistant variants.

Scenario Four: SARS-CoV-2 follows an evolutionary trajectory with decreased virulence. This could be caused by:

7. Variants arising with increased transmissibility but decreased pathogenesis/virulence as the virus becomes fully adapted to the human host becoming an endemic infection. Coupled with the likelihood of eventual high populations immunity the infection produces less disease. In other words, this virus will become like other human CoV that causes common colds, but with much less severe disease predominantly in the old or clinically vulnerable.

   Likelihood: Unlikely in the short term, realistic possibility in the long term.

General considerations for further reducing the impacts of variants:

8. Whilst we feel that current vaccines are excellent for reducing the risk of hospital admission and disease, we propose that research be considered on vaccines that also induce high and durable levels of mucosal immunity in order to reduce infection of and transmission from in vaccinated individuals. This could also reduce the possibility of variant selection in vaccinated individuals that evade prior immunity.

9. The UK should continue to proactively support a strategy of worldwide effective vaccination in order to drive down global viral load reducing the likelihood of dangerous variants emerging in other parts of the world.

10. Implement a long-term strategy for national and global genomic surveillance of SARS-CoV-2 to monitor for variants and the rapid assessment of their affects.

11. Genomic surveillance alone is not enough as phenotypes cannot be unambiguously predicted. Therefore, we recommend ensuring sustainability for the rapid laboratory phenotypic evaluation of variants at scale to run alongside clinical observations to assess risk compared to contemporary variants.

12. Invest in laboratory-based studies that can be used to predict forward evolution of variants and further interpret genomic surveillance.

13. Given the current state of predictive capabilities of artificial intelligence (AI) in biological systems we are unlikely to be able to identify new variants of risk using these approaches alone at the moment. We recommend a long-term strategy where systems are in place in peace time for quick genomic surveillance and phenotyping but where the data can be integrated with AI methods to build their utility.

14. We recommend that the UK lead on processes and legislation for the rapid obtaining and sharing of viral sequence, clinical and biological materials, especially virus isolates, on the global stage. This will allow for the rapid study and evaluation of the threat posed by novel variants.

15. We recommend process and onshoring of capacity is undertaken to reduce the time to identify and manufacture vaccines and medical countermeasures to mitigate the impact of new variants that have a more significant clinical impact.

16. We recommend careful use of antivirals, as in the face of a threat posed by new variants, these therapeutic options are precious resources and should not be squandered.

Supporting information:

17. In addition to the oldest human endemic CoV (NL63), in the last 150 years there have been 5 known coronavirus incursions from animals into humans, of these three ended up as human endemic pathogens (229E, OC43 & HKU1) and two have caused limited but severe disease (SARS-CoV & MERS-CoV). Therefore, CoV zoonosis appears a common consequence of human proximity to animals.

18. Since the beginning of the SARS-CoV-2 outbreak in Wuhan in October/November 2019, an estimated 175 million people have been infected. The outbreak has seen multiple waves of infection, latterly with genetic variants of the virus, each with differing potential public health concerns.

19. SARS-CoV-2 uses RNA as a genetic material. RNA viruses are known to make errors as they replicate, leading to mutations (genetic changes in the genome) and therefore accumulation of genetic diversity over time.

20. The genome of SARS-CoV-2 in infected humans and animals can and has been sequenced to great effect. Sequencing the genetic material of a virus provides information on the spread and evolution of the virus and potentially allows prediction of how changes in its genome may influence future infection waves.

21. Genetic variation of a viral pathogen is a natural and expected process. The overall rate of accumulation of genetic changes, and therefore the risk of a new variant emerging with altered biological properties, is dependent on the virus mutation rate, the incidence and prevalence of infection, and the advantage a new variant has over other co-circulating variants.

22. When a human (or animal) is infected with SARS-CoV-2, multiple viruses with slightly different genetic sequences are generated as replication ensues which leads to a population of viral variation with an individual person.

23. Over time, the relative proportions of variants of SARS-CoV-2 will change within the infected individual, and specific variants may come to dominate if they confer the virus with a fitness advantage. This process happens both within single infected individuals, and across the population as variants are transmitted.

24. Over the course of the SARS-CoV-2 outbreak, several variants with transmission advantages have come to dominate in infection waves. These include the alpha variant and the delta variant that have dominated the second and third waves in the UK. These variants have come to prominence through founder effect and selection favouring increase transmission, allowing alpha and then delta to out-compete previous variants. In addition, the beta and gamma variants that have circulated in South Africa and Brazil respectively may have a fitness advantage because they are antigenically distant from first wave viruses and can reinfect people more efficiently.

25. Several different processes can lead to change or growth in mutation/variant frequency and not all observed changes in variant frequency will be due to the action of natural section. It is not possible to predict with any certainty what or when new variants will emerge, or their phenotypic significance. Assessing diverse streams of data, including genetic characterisation, epidemiological trends, and laboratory studies of measurable virus properties, are necessary to evaluate the significance of new variants. What can be said is that the more virus in circulation the greater number of variants.

26. The observation that a mutation or variant has increased in frequency (or even become fixed) is not, by itself, sufficient evidence to conclude that it confers a selective benefit to the virus. New variants can rise in frequency due to events that are unrelated to the functional effect of the mutations it carries, especially when numbers of infections are low (e.g. introduction into a new location, so called ‘founder’ effect). Further, an inconsequential mutation may increase in frequency because it happens to be present on the same genome as another mutation that is beneficial to the virus (‘genetic hitchhiking’). Therefore, it is essential to use all data available to determine if changes in absolute and relative frequency are a result of a variant being “luckier” or “more fit”.

27. Mutations that enable variants to infect new hosts more rapidly and effectively will be favoured by natural selection and may increase in frequency. Mutations that enhance onward transmission can be associated with many different aspects of virus biology, including changes to disease severity, the kinetics of infection, enhancement to virus shedding or cell binding, evasion of immune responses, etc.

28. The Omicron variant of concern has illustrated how genetic change may alter transmission and disease phenotypes. Data from the human population and animal models suggest that the Omicron variant causes a milder disease. However, we emphasise that it is a common misconception that ‘viruses mutate to cause less severe disease’. This belief is from a very specific example of myxomatosis in Australia where the virus caused a case fatality rate of nearly 100%, where a strong selection pressure was placed on rabbits to be resistant and less virulent strains to emerge.

How does genetic change occur in SARS-COV-2?

29. Genetic change in SARS-CoV-2 can occur in several ways: point mutations due to polymerase copying errors leading to single nucleotide polymorphisms and recombination allowing acquisition of new genetic material including viral and host.

    a) RNA is made up of building blocks known as nucleotides. These can be any one of four different types: A, G, C, or U. The RNA genome of SARS-CoV-2 is about 30,000 nucleotides in length. Every time the viral replication factory copies the genome, random genetic change can occur. For example, an A nucleotide at a particular position on the genome could change to a U nucleotide. This is known as a single nucleotide polymorphism, or SNP. The intrinsic error rate of coronavirus genome replication is in the order of 1×10^{-6} to 1×10^{-7} mutations per nucleotide per genome replication (1 mutation in 1-10 million nucleotides replicated). As the virus genome is about 30,000 nucleotides long, then 1 mutation is introduced about every 33-330 replications. In an infected person the peak number of virus genomes exceeds 100 million genomes; therefore, the virus has the potential to mutate every nucleotide of its genome hundreds of times per infected person, therefore variant generation is common. Many SNPs do not result in an amino acid change, many cause deleterious changes and some are compatible with a viable virus, and it is these that can result in changes to amino acids in viral proteins giving them new properties. An example of this is the D614G SNP, a defining feature of an amino acid change in the spike glycoprotein leading to a change from the B to B1 lineage early in the outbreak. Phenotypic assessment showed this change is associated with increased transmission because it enhances the ability of the Spike protein to bind to the ACE2 receptor.

    b) Recombination is the process by which viruses swap genetic material between genomes, producing new combinations of genetic sequences. This can have several effects: parts of the genome can be deleted, or new sequences can be incorporated into the genome as small insertions or as whole extra genes producing a chimeric virus. New sequence can be acquired from the same type of coronavirus or a different source of RNA altogether. For example, SARS-CoV-2 is thought to have arisen by a recombination event(s) between other animal coronaviruses and/or host RNA. The presence of a furin protease cleavage site at the S1/S2 junction in the spike glycoprotein gene of SARS-CoV-2 is a consequence of such recombination. This furin site is associated with increased disease and transmission in animal models. Importantly, certain regions of the coronavirus genome have higher frequency of recombination, so called ‘hotspots’, one of which occurs at the Spike gene, promoting recombination that could lead to new virus receptor usage and alter species or cell type tropism.

30. The accumulation of genetic change has consequences for the emergence of new variants. These changes can be neutral (no effect), deleterious (decreases viral fitness) or advantageous (increases viral fitness) depending on the selection pressure.

31. SARS-CoV-2 variants are constantly being generated during infection and these provide the raw material for the virus to respond to different selection pressures. Keeping national and global levels of SARS-CoV-2 low through vaccination and/or isolation/containment will reduce the number of possible future variants.

What are the effects of genetic change on viral genome biology and protein function?

32. A change in the genotype (genetic information) of a virus is not always associated with a change in its phenotype (its observable characteristics). However, on occasion, a change in the genetic sequence of a virus may be associated with a change in its transmissibility, severity, or other characteristic such as susceptibility to immune control, vaccines or drugs.

33. Examples of changes in the genetic sequence of SARS-CoV-2 that have been associated with changes in its phenotype have been seen in the spike glycoprotein region of all variants of concern (VOCs). For example, mutation at position P681H of spike of alpha variant has been shown to increase the efficiency of furin cleavage, enhancing entry of the virus into the cell. In addition the alpha variant contains other important mutations amongst the constellation of mutations that define the alpha lineage. A deletion in the spike protein ΔH69/ΔV70 (Δ = a deletion) is associated with higher infectivity, compared to wild type, and a further mutation N501Y in spike enhances the interaction with the ACE2 receptor.

34. Several viral RNA structures and proteins are critical for virus biology and infection and therefore mutations in these regions can be lethal to the virus (and therefore not passed on). However, other regions are non-essential to critical function. For example, there are several variants with deletions of viral proteins involved in subverting the innate immune response. One of these was identified in the ORF8 region of SARS-CoV in patients in Singapore. This deletion was associated with a milder disease profile. In contrast the B.1.1.7 variant has a mutation that led to changes in several genes including a deletion in nsp6, mutations in ORF8 protein and mutations that affect potentially lead to production of new versions of viral proteins, along with the mutations in the spike glycoprotein described in point 32 above. The alpha variant is associated with greater transmissibility than previous strains but a relationship with increased morbidity and mortality were harder to establish and this particular variant has since been outpaced by others. Therefore, crucially the whole virus genome and the complete set of virus proteins contribute to a complex phenotype for the virus.

What can we learn from animal coronaviruses and how does this apply to SARS-CoV-2?

35. Coronaviruses have been studied for many decades and what has happened in other coronaviruses infections of humans and animals may paint a picture of the evolutionary journey of SARS-CoV-2.

36. Variants can lead to escape from current vaccines. Infectious bronchitis virus (IBV) is a coronavirus that infects chickens causing mainly respiratory disease. Gene translocations, deletions and recombination between IBVs from different genetic lineages have contributed to a complex population of IBV. As seen for SARS-CoV-2 the S1 region of spike is sufficient to induce good protective immunity. However, only a few amino acid differences in the spike glycoprotein gene of IBV have a detrimental impact on immunologic protection, consequently a large number of antigenic escape mutants have evolved. The further away in sequence space from the vaccine strain the spike glycoprotein gene of a particular variant, the less cross protective immunity is conferred by a vaccine. Whilst vaccines can control disease as for SARS-CoV-2, there can be sustained asymptomatic transmission leading to further evolution. Current vaccine strategies for IBV employ a combination of initial immunisation with a live attenuated vaccines followed by boosting with inactivated vaccines in order to provide cross protection, but this option is not available for SARS-CoV-2.

37. Radical changes in pathogenesis can occur in an infected individual through genetic change in animal coronaviruses. In 10% of cats infected with feline enteric coronavirus (FeCV), mutations occur in the viral genome, many targeted to the furin cleavage site in the spike glycoprotein. This allows white blood cells becoming infected and causes a new disease, feline infectious peritonitis virus (FIPV) with increased morbidity and mortality. This occurs in persistently infected cats and perhaps provides a model for the genetic diversity observed in SARS-CoV-2 in persistently infected immune compromised hosts. Similar changes have been observed in IBV, where variants cause more severe disease by infecting the kidney, oviduct and testes. For SARS-CoV-2 mutation in the furin cleavage site has been associated with increases in virulence and transmission. Variants with divergent phenotype have now been observed for SARS-CoV-2 e.g. the Omicron variant.

38. New variants can arise through interactions between wild type and vaccine strains. Transmissible gastroenteritis virus (TGEV), a coronavirus infecting pigs, has close similarity to porcine epidemic diarrhoea virus (PEDV). Mice and cats can also act as reservoirs for the transmission of PEDV. Different genotypes (variants) of PEDV can be highly virulent leading to 80-100% morbidity and up to 100% mortality in pigs. Such viruses have arisen through recombination between low pathogenic attenuated vaccine strains of TGEV and pathogenic circulating PEDV strains. This is a common feature in coronavirus vaccines in the animal world and in the closely related virus species the arteriviruses. This strongly suggests that any move to live attenuated vaccines for SARS-CoV-2 should be resisted.

39. New human coronaviruses can originate from domesticated animals. HCoV-OC43 is thought to have originated from cattle in the 19th century, possibly after a recombination event that allowed it to acquire a new gene from an influenza virus. This strongly indicates the potential for domesticated animals to serve as reservoirs for new variants of SARS-CoV-2 that can remerge into humans.

40. Whilst different situations in animal coronaviruses can point towards what may be expected with SARS-CoV-2 one of the greatest gaps in our current knowledge is how seasonal human coronaviruses vary with time. We would urge the identification of appropriate biobanks with respiratory samples over the course of years. These should be utilised to identify and sequence human coronaviruses to measure the rate of evolution and genomic diversity.

What type of changes have we seen so far in SARS-CoV-2 and what might we expect?

41. Although genetic change accumulates at random, environmental factors known as selection pressures will influence whether that genetic change increases or decreases viral fitness and therefore rises or declines in the population. In the case of SARS-CoV-2, there are particular selection pressures that are more concerning because they may encourage the emergence of variants that may be more harmful or more difficult to control.

42. There is a remarkable degree of convergent evolution evident amongst the variants of SARS-CoV-2 that have thus far emerged. In different combinations and using slightly different coding changes, all variants have mutations in the spike protein that appear to enhance the direct binding to ACE2 receptor, the presentation of the receptor binding domain and often the efficiency of entry via enhanced furin cleavage that primes for fusion. Changes at amino acid positions 203 and 204 in the nucleoprotein are further examples of convergent evolution. Variants also all harbour mutations in other part of the genome likely enhance polymerase activity (P323L in nsp12) or affect the virus' ability to antagonize the host innate response (changes to ORF8, nsp6 deletions – whose function are not yet understood and changes in spectrum of viral gene expression). There are multiple insertions and deletions as well as SNPs, but so far, the virus has not acquired any whole new genes. Taken together, the constellations of mutations enhance the interaction of the virus with its new human host. As more information is uncovered about the way SARS-CoV-2 interacts with human cells, we might foresee some concerning possibilities for further evolution.

43. For example, the interferon system is a major determinant of disease outcome, individuals deficient in interferon fare badly. The interferon system works by sensing the presence of an invading virus and responding by production of hundreds of host proteins that have antiviral activity. Emerging data suggests the alpha variant has evolved to be less readily sensed by the infected cell. Transfer by recombination of the genetic determinants of this phenotype to other variants or independent acquisition by forward evolution is expected to increase transmission and disease severity by tipping a balance in favour of virus over host.

44. A second example is the possibility for the virus to acquire additional interferon antagonists by recombination. Emerging data finds that one particularly potent interferon stimulated gene that controls SARS-CoV-2 replication in infected cells is oligoadenylate synthetase 1 (OAS1). OAS1 expression reduces the amount of virus replication and controls infection. Other coronaviruses are also sensitive to OAS1 but have overcome this by acquiring an enzyme called PDE from the host. It is to be expected that if SARS-CoV-2 acquired PDE by recombination with another human coronavirus such as OC43 or from a host cell, the capacity to replicate in the face of an interferon response would be increased and this would lead to higher transmission and disease.

45. As vaccines against SARS-CoV-2 are deployed across populations, it is possible to create a selection pressure for variants that are able to escape the vaccine-acquired immune response. Over the past few months, several variants have emerged which show a reduced susceptibility to vaccine-acquired immunity, though none appears to escape entirely. These variants largely emerged before vaccination was widespread, thus selection pressure from vaccines is unlikely to have made a significant contribution to their emergence. However, as vaccines become more widespread, the transmission advantage gained by a virus that is able to evade vaccine-acquired immunity will increase.

46. The extent of vaccine escape is difficult to predict. The SARS-CoV-2 spike protein appears to be quite plastic accommodating a plethora of new mutations, similar to what occurs in animal coronaviruses. Deep mutational scanning has been used to try to predict which (single) mutations in spike will evade antibodies. So far, the approaches use artificial methods to generate and express the spike variants such as yeast, phage or expression on another virus, VSV. None of these enable the full spectrum of mutational possibilities generated by the coronavirus polymerase. Nonetheless, important mutations were predicted by these methodologies such as the E484K mutation the spike receptor binding domain that is the most potent of all single point escape mutations thus far seen.

47. The use of anti-viral therapies also presents a selection pressure for the evolution of drug resistance, as virus that is able to evade the treatment will replicate, and possibly transmit to others as readily happens with anti-influenza drugs. Indeed, all antiviral drugs used to date (against HIV, Herpesviruses, influenza virus) have generated drug resistant mutants that limited the effectiveness of the drugs. Where the anti-viral therapy consists of more than one antiviral drug (combination therapy), the emergence of escape variants is less likely. These combinations act to keep viral load very low and thus statistically there are less variants. If combination therapy was used for SARS-CoV-2, the virus would need to develop resistance to multiple drugs simultaneously in order to replicate, this is unlikely. Extreme care in deployment of anti-SARS-CoV-2 drugs is therefore critical, particularly if monotherapy is considered.

48. In immunocompromised hosts, viral replication and shedding may be prolonged. This increases the potential for emergence of escape variants, particularly under the pressure of monoclonal antibodies therapy targeting the spike protein or convalescent plasma treatment, again mainly targeting the spike glycoprotein. Case studies have documented the emergence of unusually high numbers of genetic changes in immunosuppressed individuals treated with convalescent plasma.

49. Evolution in immunocompromised hosts may be a way by which viruses traverses a fitness landscape through accumulating constellations of mutations that confer phenotypes and epistatic mutations that compensate for fitness costs. It is notable that all the current SARS-CoV-2 VOCs harbour constellations of mutations across the genome implying they either evolved during long term persistent infection, occurred in under circumstances of intense transmission with wide selection bottlenecks or occurred in environments were sustained but undocumented transmission occurred. On the other hand, it is noteworthy that single point mutants on top of VOCs such as alpha and delta have not (yet) emerged to predominate over the existing VOCs.
50. Variants can potentially change the transmission of the virus leading to different modes of infection within community or demographies associated with potential novel properties, for example a faecal oral transmission rather than respiratory. Examining other coronavirus in animals and humans show faecal oral transmission can occur as an efficient additional means of transmission as was the cases with SARS-CoV and transmission in the Amoy Garden Complex. There is not yet evidence for alternate routes of transmission for SARS-CoV-2 but delta variant has been associated with increased frequency of GI symptoms. We would note that in the UK and other countries with water treatment and clean drinking water transmission of (possible future) gastro-intestinal variants would be unlikely.

51. SARS-CoV-2 can infect a wide range of animals both in nature/farms (such as minks) laboratory animals (several species of non-human primates, mice, rats, ferrets and hamsters) and companion animals (cats and dogs). SARS-CoV-2 is thought to have originated from bats. Thus SARS-CoV-2 has a broad host range and is capable of continuous interchange between humans and an animal reservoir (e.g. mink farms) which could lead to the generation and selection of new variants.

52. Infections in mink farms have been observed throughout the world. The widespread presence of the virus in an animal population will render eradication even more unlikely. There is likely to be a different level of risk of exposure and potential for new variants between farmed animals (high density) and companion animals. In either case, reverse zoonosis may occur (already seen in Denmark from mink). Zoonotic reservoirs could lead to a large, expanded population of the virus with the potential for future dramatic variant change in the virus through recombination with another coronavirus already prevalent in that animal species, akin to antigenic shift in influenza virus in terms of conferring new virus properties.

53. Re-infection following SARS-CoV-2 infection is well described in case reports. The frequency of re-infection and the predisposing risk factors are less well understood. Evidence is emerging from PHE/UK based longitudinal cohorts and international publications that prior infection is generally protective against re-infection over at least a 9-12 month period in most people. Estimates from SIREN, a large national HCW cohort and from the PHE Care homes study indicate that re-infection occurs in <5% individuals and that the serological profile in cases of re-infection suggests that low levels of neutralising antibody confer susceptibility. However, most of these data relate to the alpha VOCs and these data may change when the delta VOC is considered. Reinfections are likely to drive further evolution.

54. The immune response to SARS-CoV-2 involves multiple mechanisms, including innate defences, antibodies, T cells, and B cells. While virus-neutralising antibodies are usually against specific sites exposed on the surface of virus proteins, T cells recognise peptide fragments from a wider range of viral proteins that may be conserved between viral variants, reducing the likelihood that immune escape will emerge. SARS-CoV-2 infection induces strong T cell responses, which may provide additional selection pressure for SARS-CoV-2 evolution. For example, there is some suggestion that, in addition to increased transmissibility and a level of antibody escape, the delta variant escapes T cell responses in one HLA type common in Asia. How many T cell epitope escape mutations are needed to decrease T cell immunity and if there is a hierarchy of crucial T cell epitopes is not known.

55. Non-pharmaceutical interventions such as mask-wearing and social distancing shape the environment in which a virus transmits and may act as selection pressures for increased viral transmission. For example, where social distancing exists, virus that is able to better transmit through air may have a fitness advantage, and therefore be better able to survive and replicate.

56. Should a genetic change that benefits the virus arise under any of these selection pressures, there is no guarantee that it will rise to dominance. If the change occurs in conjunction with others which are deleterious, then the variant in which it is present may die out, because there is no overall benefit to the combination. Even if the variant is fitter than others, in order to become dominant, it must also be able to spread between people and communities. If population mixing and transmission is low, new variants are less likely to spread.

Antivirals could provide a major selection pressure in the emergence of drug-resistant variants.

57. The replication of the SARS-CoV-2 leads to many of the clinical manifestations of COVID-19. Specific antiviral drugs, which work by inhibiting or suppressing viral replication and thereby minimising the impact of tissue damage caused by SARS-CoV-2, are therefore being investigated for the treatment of COVID-19.

58. For antiviral drugs, it is important to find viral targets which can be selectively inhibited, so as to minimise the impact on host cell replication. Typically, key points in viral replication cycle can be targeted, including:
    a. Viral entry (via the angiotensin-converting enzyme 2 [ACE2] receptor and transmembrane serine protease 2 [TMPRSS2]).
    b. Viral membrane fusion and endocytosis.
    c. Viral replication involving the RNA-dependent RNA polymerase complex. This can involve the application of mutagens that decrease viral fitness through error catastrophe. For example, Molnupiravir (Lagevrio) has been approved by the MHRA in the UK. We have no experience of the wide scale application of this type of compound which in certain circumstances could unwittingly drive evolution through generation of variants. We note that resistance has been identified in other viral infections treated with such anti-virals.

d. Viral protein or RNA synthesis involving the activity of viral enzymes, such as the SARS-CoV-2 3-chymotrypsin-like protease (3CLpro).

59. Neutralising mAbs have been given emergency use authorisation and recently shown to be effective in treatment of symptomatic disease. This class of molecule is an attractive and flexible approach to therapy for emerging viral infections. Monotherapy with single target mAbs is likely to generate pressure for the emergence of viral variants. Cocktails of mAbs will provide a higher barrier to the emergence of escape variants. Activity of mAbs can also be reduced by variants that escape convalescent or vaccine induced antibody immunity.

60. There is no historic precedent for the mass administration of antiviral medication in the community as prophylaxis, apart from the use of anti influenza Neuraminidase Inhibitors, which were used to a limited extent in this way in the early phases of Influenza Pandemic of 2009 in the UK. The safety and efficacy profile must be extremely well established for a mass administration strategy to work and poor compliance will likely rapidly lead to the selection of drug resistant variants, rendering such a strategy short lived.

61. Combination therapy is recognised as a way of minimising viral evasion, typically involving two or more drugs or drugs and/or mAbs combined targeting different viral functions, creating a higher barrier to the evolution of resistance. Preliminary animal studies in the ACE2-transgenic mouse model illustrate that two anti-viral compounds with different modalities can be used together and result in enhanced viral clearance. However, mass administration of combination therapy is an even harder strategy to achieve.

62. None of these points are reasons not to use anti-virals in clinically vulnerable or severely ill patients but we would caution that genomic and phenotypic capabilities should remain in place to rapidly monitor and flag the emergence of potential resistance.

63. The application of anti-virals should not be for patients who are already on the road to recovery – although this will be a clinical judgement.