Evolution of resistance to COVID-19 vaccination with dynamic social distancing

Evolution of resistance to COVID-19 vaccination with dynamic lockdown

Gabriela Lobinska*[1], Ady Pauzner[2], Arne Traulsen[3], Yitzhak Pilpel*[1], Martin A Nowak*[4]

1 Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel

2 Berglas School of Economics, Tel Aviv University, Tel Aviv 69978, Israel

3 Department of Evolutionary Theory, Max-Planck-Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Ploen, Germany

4 Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge MA 02138, USA

The COVID-19 pandemic has led to an unprecedented global response in terms of social lockdown in order to slow the spread of the virus [1], [2]**. Currently the greatest hope is based on world-wide vaccination**[3], [4]**. The expectation is that social and economic activities can gradually resume as more and more people become vaccinated. Yet, a relaxation of social distancing that allows increased transmissibility, coupled with selection pressure due to vaccination, will likely lead to the emergence of vaccine resistance** [5]**. Here we analyze the evolutionary dynamics of COVID-19 in the presence of dynamic lockdown and in response to vaccination. We use infection and vaccination data of 6 different countries (Brazil, France, Germany, Israel, UK, US)** [2]**. For slow vaccination rates, resistant mutants will appear even if strict lockdown is maintained. For fast vaccination rates (such as those used in Israel) the emergence of mutants can be prevented if strict lockdown is maintained during vaccination. We analyze multiple human factors that affect the evolutionary potential of the virus, including the extent of dynamic lockdown, vaccination campaigns, boosters and vaccine hesitancy. The emergence of vaccine resistance can also be prevented by developing two-gene vaccines. Our results provide guidelines for policies that aim to minimize the probability of emergence of vaccine resistance variants.**

The COVID-19 pandemic has had a devastating effect on global health and economy. Since the identification of the first SARS-COV-2 case in December 2019, 178.71 million infections have been recorded and at least 3.86 million people have died as a result of the infection (as of June 2021)[2]. The increased mortality and complication rates of SARS-COV-2[1] compared to the mild diseases caused by seasonal coronaviruses, such as HCoV-229E[6], have led to unparalleled governmental and individual-level responses in order to reduce the number of SARS-COV-2 infections.

Since the beginning of the pandemic, it has become clear that many of the non-pharmaceutical interventions (NPI), such as lockdowns, are economically and socially unsustainable in the long run. Periodical loosening and tightening of social distancing measures, which present an attempt at balancing economical and sanitary considerations, have led to waves of increase and decrease in the number of SARS-COV-2 infections per day[2] (see **Figure 1A**). Therefore, much hope has been placed on vaccine development, which would allow the immunization of a large fraction of the population, thereby substantially reducing mortality and potentially achieving herd immunity, which could in principle eradicate SARS-COV-2 altogether.

Mass vaccination campaigns have been launched in many countries (see **Figure 1B**), most notably Israel and the UK (both more than 60% of fully vaccinated population) and the US and Germany (both more than 50% of fully vaccinated population). Currently, four companies are producing vaccines that have been approved for emergency use either by the Food and Drug Administration (FDA)[3] or by the European Medicines Agency (EMA)[4]: Pfizer-Biontech, Moderna, AstraZeneca and Johnson & Johnson/Janssen Pharmaceuticals. Several other vaccines are also used outside of the European Union and the USA: Gamaleya (Sputnik V), Sinopharm Beijing, Sinovac, Sinopharm-Wuhan and Bharat-Biotech (Covaxin)[7].

However, the identification of new SARS-COV-2 variants has cast a shadow over the expectation of a swift end of the pandemic[8], [9]. The so-called "British" variant (B.1.1.7), now termed α, and "South African" variant (501.V2), now termed β, have been shown to be neutralized to a lesser extent by convalescent and vaccinee sera[10], although experiments on non-human primates have shown that this decrease might not necessarily cause a decrease in immunity[11]. Structural studies have mapped and predicted mutations that lead to antibody escape[12]–[14]. As vaccination around the world progresses, the continued evolution of SARS-COV-2 could eventually give rise to a fully vaccine resistant variant. Such a variant could quickly spread due to its ability to infect vaccinated and recovered in addition to fully susceptible individuals. The question of emergence of vaccine resistance has already been the subject of many research papers[15]–[18].

What policy could be exercised that would minimize the chance of emergence of vaccine resistant strains? Policymakers can vary the extent of social distancing imposed and regimes of vaccine administration. The critical biological parameters on the other hand include the infectivity of the various strains and the rate of mutation of the virus that may ultimately lead to emergence of a resistant strain. Here we introduce a mathematical model that examines various combinations of these parameters. Our model helps to design optimal policies that would minimize the chance of emergence of resistant strains or maximize the time until their occurrence.

Our paper is an addition to the extensive body of work that has been performed in the past year in order to understand the spread and evolution of SARS-COV-2[19]–[25]. SARS-COV-2 research has drawn on a very long

history of epidemiological research[26], [27], [36]–[38], [28]–[35]. Due to the global and urgent nature of the pandemic, many studies that could inform policy-making have been conducted[5], [39], [48], [40]–[47].

In order to understand the evolutionary potential of the virus in response to a vaccination program we study a stochastic model for infection dynamics and virus evolution in the presence of varying degrees of social lockdown and different vaccination rates. We distinguish between a wild-type virus (WT) and a vaccine resistant mutant virus (MT). The vaccine is effective against the WT strain, while the MT strain evades immunity induced by the vaccine either partially or completely. We build upon the mathematical framework of the Susceptible-Infected-Removed (SIR) model from epidemiology [32], [49], albeit with considerable adjustments necessitated by the specific problem at hand . Our model keeps track of people who are susceptible, infected by WT or MT, recovered from WT or MT, vaccinated or unvaccinated (**Figure 2**).

Crucially, we assume there is a dynamic lockdown guided by the number of new infections per day. As that number exceeds a threshold, governmental rules and individual responses reduce social activity. If the number of new infections falls below this threshold, the lockdown is somewhat relaxed and some people stop following the rules, thereby allowing higher transmission of the virus. We simulate these dynamics as a stochastic process. In consequence, we obtain fluctuating numbers of new infections per day. We introduce mass vaccination at alternative fixed rates. Then we compute the probability and timing of the wave of infection caused by the spontaneous emergence of a vaccine resistant virus.

In our approach, the mutation rate μ denotes the probability that a WT-infected individual will infect a susceptible individual with the MT strain. The exact value of this probability is currently unknown and complex to obtain empirically. For the simulations and calculations reported in this paper we therefore consider a wide range of mutation rates. From our model, we also derive an upper bound for the mutation rate using the fact that no wave of a vaccine-resistant variant has occurred up until now. Note that this rate can be very different from the per-base mutation rate of SARS-COV-2, which is about 10^{-5}.

DYNAMICS OF VIRAL INFECTION AND EVOLUTION

Our model keeps track of eight different variable states: individuals who are susceptible (x), infected with WT (y_1), non-vaccinated and infected with MT (y_{2A}), vaccinated and infected with MT (y_{2B}), recovered from WT (z_1), recovered from MT (z_2), vaccinated but susceptible to MT (w_1), vaccinated and recovered from MT (w_2); see **Figure 2**.

The WT strain can infect susceptible individuals (x), converting them to individuals infected with WT (y_1) at rate β_1. A mutation can occur with probability μ. In this case, a WT infected individual infects a susceptible individual (x) with a mutated version of the virus, in a mutation that have taken place in the infecting individual, thus converting the susceptible to a MT infected individual (y_{2A}). WT infected individuals either recover with rate a and become immune to future WT infection (z_1) or die at rate d. Susceptible individuals (x) and individuals recovered from WT (z_1) can become vaccinated individuals (w_1). The parameter c denotes the number of individuals vaccinated per day. Hence, the rates of vaccination of x, z_1 and z_2 are respectively $cx/(x + z_1 + z_2)$, $cz_1/(x + z_1 + z_2)$ and $cz_2/(x + z_1 + z_2)$. For simplicity we assume single-dose vaccination; for a double dose vaccine our model would describe the application of the second dose ignoring partial immunity caused by the first dose; extension of our model to a full two dose vaccination protocol is straightforward. Although several countries display a logistic-shaped vaccine distribution curve, we show that the above linear assumption does not significantly affect the probability of emergence of vaccine resistance (see **Supplementary Figures S1** and **S2**).

At rate β_2, the MT strain infects susceptible individuals (x), WT recovered individuals (z_1) and vaccinated individuals who are not immune to MT (w_1). MT infected individuals either recover with rate a and become immune to future MT and WT infection (z_2) or die at same death rate d as with the WT strain, i.e. assuming no difference in lethality between the two strains. While we consider death due to infection, the focus of this paper is not to model the number of deaths, nor public health metrics such as ICU usage or available respirators. We assume one-way cross-immunity induced by the viral strains: the MT strain can infect individuals that have recovered from WT or that have been vaccinated against WT, but the WT strain cannot infect individuals that have recovered from the MT. This assumption is reasonable because the MT evolves in the presence of the WT, but not vice versa. We note that our MT strain escapes both from the immunity that is induced by natural infection with WT and the immunity induced by vaccination against WT. We assume that the partial immunity to the vaccine resistant variant induced by infection with wild type is equal to that induced by vaccination. A recent study suggests that individuals that are both recovered and vaccinated benefit from higher immunity than only vaccinated individuals. In the Supplementary Materials (**Figures S3** and **S4**) we study this effect.

We need to distinguish between MT infected individuals that are or are not vaccinated: y_{2B} and y_{2A}, respectively. Upon recovery the former will not be vaccinated (again), while the latter will be vaccinated.

We also study partial immunity to the MT strain which can be acquired by recovery from WT infection or by vaccination. For partial immunity, the corresponding infection rates are multiplied by a parameter q, which is between 0 and 1. If $q = 1$ then WT infection or vaccination confers no immunity to MT at all; the mutant escapes completely. For $0 < q < 1$, the MT is a partial escape mutant. For $q = 0$, the MT does not escape at all.

Lockdown measures are implemented by multiplying the infectivity coefficients of each strain by a social activity parameter s which ranges in $[0,1]$. Unconstrained social interaction means $s = 1$, while $s = 0$ would be complete lockdown. The population tolerates a certain number of new infections, L, per day. Each day, if the number of new infections exceeds L, then s is decreased by a random, uniformly distributed number between 0 and 0.1, noted s_0. If the number of new infections is less than L, then s is increased by a random, uniformly distributed number between 0 and 0.1. In any case, s cannot decrease below 0.05 or increase above 1. In our simulations, we adjust the parameter s every day. We then obtain a near constant number of cases per day. In Supplementary Materials, we explore less frequent adjustments of the parameter (such as every week) and conclude that our results are robust (see **Figure S5** and **Figure S6**.) Our social activity parameter s can be interpreted to include other factors that could potentially affect transmissibility, such as seasonality. In Supplementary Materials (see **Figure S7** and **Figure S6**), we explore the effect of seasonality.

As an example, the rate of infection of the recovered from WT z_1 by the MT strain infected individuals y_2 is multiplied both by the lockdown coefficient s and the partial immunity coefficient q – hence this rate is given by $q\beta_2 s w_1 y_2$.

All model parameters are summarized in **Table 1.**

COMPUTATIONAL IMPLEMENTATION AND DATA

A Gillespie algorithm is commonly used to simulate stochastic systems with high variation in waiting times between consecutive events [50]–[53]. In our model the population is represented as a vector of length eight, corresponding to the eight categories. The rates of all possible events (infection, recovery, death, mutation and vaccination) are calculated. The time of the next event in the model is drawn from an exponential distribution, with parameter dependent on the sum of all event rates and an event is chosen, with probability proportional to its rate. The population is updated according to the event that occurred. The simulation is stopped when there are no more infected individuals in the population. The algorithm is presented in pseudocode in the **Appendix**, along with a table of the possible events of the model and their default rates.

In order to achieve feasible computation time and resources, we simulated populations of size up to $N = 10^6$. The results of those simulations can be scaled to larger population sizes by considering a population of for example $N = 10^7$ as $m = 10$ "batches" of 10^6 individuals, and computing the results for $N = 10^7$ as $1 - (1 - p)^m$, where p is the proportion of runs where the MT strain took over. **Figure S8** shows the strong agreement between simulated results and the results scaled from simulations with smaller population sizes.

For all our simulations, we have endeavored to use real world data for all model parameters. In particular, infection and vaccination data has been obtained from the database *Our World in Data* [2] (OWID) and downloaded on June 19th, 2021.

In our simulation, since the number of new infections each day is constant, the number of susceptible individuals decreases linearly with slope $-L/a$. Vaccination of both susceptible and recovered individuals proceeds at rate c. The social activity parameter, s, increases as more and more individuals become immunized either by infection or by vaccination. The WT reproductive rate, R_{WT} , is maintained at 1 as long as the MT has not appeared. The MT reproductive rate, R_{MT} , increases with the social activity parameter s until the MT strain takes over. After MT takeover, the MT reproductive rate R_{MT} is buffered at 1 by the dynamic lockdown (see **Figure 3** and **Methods**).

We performed 1000 runs of the stochastic simulation for each combination of parameters reflecting realistic values of the two model parameters determined by governmental policy: the tolerated number of infections per day, L, and the vaccination rate per day, c. Each square of the color map shown in **Figure 4** reflects the average value of these 1000 runs, which were performed for a population of $N = 10^6$ and then scaled to $N = 10^7$ and $N = 10^8$. At each combination of L and c the color maps denote the predicted probability of a mutant take over. We perform computations using $q = 1$ and $q = 0.4$ for complete and partial immune evasion by the mutant.

Allowing a large amount of infection cases and slow vaccination results in almost certain takeover of the MT strain. On the other hand, very fast vaccination coupled with a low number of tolerated new infections per day can prevent emergence of the MT. Partial immune evasion $(q = 0.4)$ of the mutant slightly reduces the probability of its takeover. Note that the shape of the parameter space where we observe takeover is similar for $q = 1$ and $q = 0.4$.

Although estimating COVID-19 mortality is not the focus of this paper, we have also recorded the number of deaths in the first 365 days of the simulation. The results are shown in **Figure S9** of the Supplementary Materials.

REPRODUCTIVE RATIO OF THE MUTANT AND PROBABILITY OF TAKEOVER

In **Figure 5** we show detailed data from 6 countries together with the estimated reproductive ratio, R_{MT}, of a vaccine resistant mutant and the probability of generating a wave of resistant virus. Data for the number of susceptible individuals $x(t)$, vaccinated individuals $w(t)$, recovered individuals $z(t)$, newly infected individuals $L(t)$, and an estimate for the reproductive rate R_{WT} of the WT can be obtained from *OWID* [2]. The reproductive rate R_{MT} of the escape mutant can be calculated according to:

$$R_{MT}(t) = R_{WT}(t)[x(t) + qw(t) + qz(t)]/x(t) \quad (1)$$

The probability of not producing an escape mutant in a given day is $(1-\mu)^{L(t)}$. The probability of not producing a surviving escape mutant is $(1-\rho(t)\mu)^{L(t)}$, where $\rho(t)$ is the survival probability of a mutant generated on that day. If $R_{MT}(t) < 1$ then $\rho(t) = 0$. If $R_{MT}(t) > 1$ we assume $\rho(t) = 1 - 1/R_{MT}(t)$. The probability that no surviving mutant is generated between time 0 and time t is given by

$$P(t) = \prod_{\tau=0}^{t} [1-\mu\rho(\tau)]^{L(\tau)} \quad (2)$$

In **Figure 5** we show the reproductive rate of the mutant $R_{MT}(t)$ and the probability $P(t)$ of generating a surviving escape mutant as a function of time. Prior to vaccination the reproductive rate of a potential escape mutant tracks closely that of the WT. As people become vaccinated in large numbers, R_{MT} starts to increase significantly above R_{WT}. Nevertheless, it is possible to keep R_{MT} below one by maintaining some measures of lockdown. (This is the case for Israel and UK). Overall, the probability that Israel generated a vaccine escape mutant (before June 2021) is of order of 1-2 percent (assuming $\mu = 10^{-7}$). For the same mutation rate the corresponding probability for the United States is 32 percent; the United States have a much larger total population size but also many more infections per million people. The corresponding probabilities for Brazil, France, Germany, and UK are 17, 8, 4 and 6 percent (see **Table 2**).

ESTIMATING THE MUTATION RATE μ

We suggest a method to estimate an upper bound of the mutation rate μ from WT to MT based on the observation that despite a large number of infections since the beginning of the pandemic and including recent vaccination campaigns, no immune evasive mutant has yet taken over. Our method for calculating an upper bound of the mutation rate is potentially applicable for estimating any mutation rate between two phenotypes in an evolving population. The upper bound is computed at any given time point, and can be updated and become tighter if in future an evasive strain still does not appear.

Let us assume that the real value of the mutation rate equals μ^*. For each country, we have an estimate of the reproductive coefficient of the MT strain and the number of new infection cases for each day (see **Figure 5**) for each day from the beginning of the pandemic until August 15th, 2021. Hence, we can calculate the probability that the MT strain would have emerged by August 15th, 2021 assuming a mutation rate μ^*. If this probability is higher than 0.5, then the probability that the MT strain would have been observed on August 15th, 2021 is higher than not, given that the mutation rate equals μ^*. Since as of August 15th, 2021, an MT strain has not been observed, then our estimate of the mutation rate must be lower than μ^*. We define the upper bound for the

mutation rate on a given day as the value μ for which the probability that the MT strain would have taken over by that day is 0.5.

Hence, for each day, we compute the probability that the MT strain would have taken over until that day given a mutation rate using Eq. (2). We assume $q = 1$, which means the MT strain is fully immune evasive. The resulting function for probability of mutant take over (for a given time point) versus mutation rate has a sigmoidal shape, with its midpoint corresponding to the mutation rate for which it is equally probable that the MT strain would have taken over or not, i.e. our definition for the upper bound of the mutation rate. In **Figure 6A**, we show the probability of emergence of the MT strain on July 30th, 2020 given data from the six considered countries. The upper bound estimate for the United States is the x-axis value for the point indicated by the red arrow, i.e. the midpoint of the sigmoidal function. The probability of emergence of the MT strain until August 15th, 2021 for a given mutation rate is higher than it was on July 30th, 2020 since many infections have occurred since then. Hence, the upper bound estimate for the United States has decreased, i.e. shifted to the left on the x-axis (see position of the red arrow) (see **Figure 6B**).

Therefore, the estimated upper bound on the mutation rate decreases over time as long as more infections do not give rise to a MT strain (see **Figure 6C**).

Since the probability of MT takeover (Eq. (2)) is strongly dependent on the number of infections, significant decreases in the estimated values correspond to periods with high infection rates in which, nonetheless, a mutant did not appear. The estimate for the upper bound of the mutation rate is expected to plateau as vaccination campaigns lead to a decrease in the number of infection cases. The estimate of 10^{-6} will decrease further if and when large countries such as the US will advance in the vaccination campaign with no mutant takeover. Using the world infection and vaccination data, we obtain $\mu = 10^{-7}$ as the order of magnitude for the upper bound for the rate at which immune evasive mutants appear. But estimates based on individual countries may be more informative since the world data reflects the average over an extremely heterogeneous population subject to very different policies.

Can we estimate a lower bound for the mutation rate? Our method could be potentially applicable for estimating the lower mutation rate, assuming that a vaccine resistant has emerged. Then, we could compute the value of μ given the date of emergence and assuming that the probability of emergence on that day has become higher than 0.5. Let us assume that an MT strain did takeover on August 15th, 2021. Then, numerically solving Eq. (2) plugging in data from the United States yields $\mu = 2.2 \cdot 10^{-7}$ (which is also the upper bound, assuming that the MT strain has not taken over).

A SIMPLE FORMULA FOR THE ESCAPE PROBABILITY

The dynamic lockdown captured by the social activity parameter $s(t)$ maintains the number of new infections per day fluctuating around a fixed value and thereby buffers the reproductive ratio of the wildtype R_{WT} around 1. The number of active infections is roughly constant and given by L/a, where a is the recovery rate (see **Figure 3**). If vaccination is slow, $c \ll L$, then the change in the number of susceptible, $x(t)$, and recovered individuals, $z_1(t)$, over time can be described by linear functions with slopes proportional to L (see **Methods and Figure S10A**).

Alternatively, for fast vaccination, $c \gg L$, the change in the number of susceptible $x(t)$ and vaccinated individuals $w_1(t)$ can be described by linear functions with slopes proportional to c before MT takeover, and with slopes proportional to L after MT takeover (see **Methods and Figure 3E**). Neglecting vaccination of recovered individuals (which is a reasonable approximation for $c \gg L$) we can write $(t) = N - Lt - ct$, $z(t) = Lt$ and $w(t) = ct$. The time when herd immunity against the WT is reached is given by

$$T_H = \frac{N}{c+L}\left(1 - \frac{1}{R_0}\right) \tag{3}$$

During vaccination the reproductive rate of the mutant increases as (see **Methods**)

$$R_{MT}(t) = \frac{N}{N-(L+c)t} \tag{4}$$

The reproductive rate of the MT is initially 1 and increases to R_0 as people recover from WT infection or are vaccinated (see **Figure 3D**). Once a mutant has been generated, the probability of its survival depends on the value of the reproductive rate, $R_{MT}(t)$. The probability that no surviving mutant has appeared before time t, where $t \leq T_H$, can be calculated to be (see **Methods**):

$$P(t) = \exp\left[-\left(\frac{\mu N}{2}\right)\left(\frac{L}{N}\right)\left(\frac{c+L}{N}\right)t^2\right] \tag{5}$$

The probability that no surviving mutant has appeared until herd immunity is

$$P(T_H) = \exp\left[-\left(\frac{\mu N}{2}\right)\left(\frac{L}{c+L}\right)\left(1 - \frac{1}{R_0}\right)^2\right] \tag{6}$$

Here $R_0 = \beta N/a$ is the basic reproductive ratio of the WT. The corresponding formulas for partial immune escape mutants are given in the **Methods**. Eq. (6) is in good agreement with the results of exact stochastic simulations (**Figure S11**). In addition, we derive $P(T_H)$ for more infectious mutants, i.e. $\beta_2 > \beta_1$ (see **Methods,** and **Figure S12**). We also explore simulation results for mutants that are less infectious than the wild type in the Supplementary Materials (see **Figure S13**). We notice that the probability $P(T_H)$ does not depend on parameters

β, a, s_{thres} and d. In order to confirm that the value of these parameters does not affect the probability of emergence of a vaccine resistant variant, we have included a sensitivity analysis in the Supplementary Materials (see **Figure S14**).

In **Table 3**, we show how the probability and timing of resistance depends on the vaccination rate and the number of new infections per day. We first consider a large country of $N = 10^8$ inhabitants and a mutation rate of $\mu = 10^{-7}$. If 10,000 new infections occur per day and 1 million people are vaccinated per day, then herd immunity is reached in 66 days and the probability of generating a vaccine resistant mutant is about 2 percent. For the same vaccination rate, if 50,000 new infections are tolerated each day, then the probability of generating an escape mutant increases to 10 percent. If 10,000 new infections occur per day but only 100,000 people are vaccinated every day, then the probability of generating vaccine resistance increases to 18 percent.

As the proportion of vaccinated individuals grows, social distancing measures relax, and the probability of emergence of resistance increases. Hence, higher vaccination rates are associated with higher probabilities of resistance after 50, 100 and 200 days (see **Table 2**). However, faster vaccination leads to earlier herd immunity. When herd immunity is reached, there are no more new infections and the cumulative probability of resistance plateaus. Therefore, we observe an interesting counterintuitive effect: the probability of resistance until a fixed time t increases with the vaccination rate c, but the probability of resistance until time T_H when herd immunity is achieved decreases with the vaccination rate c. (see **Table 3 and Figure S15**).

We can derive estimates for the emergence of vaccine resistant strains using current vaccination and infection rates from around the world. If the whole world ($N = 8 \cdot 10^9$) vaccinated as fast as the US ($c = 5000$ per day per million) and had slightly lower infection rates than Germany ($L = 100$ per day per million) then herd immunity would be achieved in $T_H = 131$ days; the probability that a resistant virus was generated and survived by that time would be 0.97 (for $\mu = 10^{-7}$) and 0.29 (for $\mu = 10^{-8}$). If the whole world vaccinated as fast as Brazil ($c = 3000$ per day per million) and had infection rates like the US ($L = 250$ per day per million) then herd immunity would be achieved in $T_H = 205$ days; the probability that a resistant virus was generated and survived by that time would be 0.999 (for $\mu = 10^{-7}$) and 0.75 (for $\mu = 10^{-8}$). Our results underline the importance of maintaining lockdown measures while herd immunity is not achieved and timely distribution of vaccines around the world.

PREVENTING EMERGENCE OF VACCINE RESISTANCE

Improved vaccine design

The probability of emergence of vaccine resistance can be reduced by increasing the number of vaccinated people per day (c) and reducing the number of allowed infections per day (L) until herd immunity is reached.

However, policymakers could also affect the mutation rate μ by determining what type of vaccine is used. We have extended our basic model to consider vaccine resistance achieved when two independent mutations are present (see **Figure S16** and **Figure S17**). Each mutation is neutral by itself. We find that for realistic mutation rates, that is mutation rates below our estimated upper bound, vaccine resistance never emerges when two-gene vaccines are distributed (see **Figure S18**). This extension can also be interpreted as the case where two independent mutations are needed to confer resistance to a vaccine based on a single gene. But using two mRNAs in a single vaccine could double the number of mutations needed to achieve immune escape.

Reducing vaccine hesitancy

In most developed countries, vaccine hesitancy has caused the number of vaccinated individuals to plateau. Vaccine hesitancy can also be reduced by policy decisions. In **Figure S19** and **S20**, we extend our model to account for the proportion of population that will not be vaccinated. We find that although high vaccine hesitancy substantially increases the probability of emergence of vaccine resistance, low vaccine hesitancy can actually have a negligible effect. For $R_0 = 3$, the levels of vaccine hesitancy necessary to significantly increase the probability of emergence of vaccine resistance are much higher than the vaccination hesitancy rates in many of the considered countries except for Brazil and the United States. However, more infectious mutants have emerged in the past months.

In order to achieve herd immunity, a certain fraction of the population must become immunized to the infecting agent. Immunization can either occur by recovery from infection or by vaccination. The minimum fraction of immunized required for herd immunity increases with the reproductive coefficient R. If a certain proportion of the population will not be vaccinated, then an excess of infections will occur in order to achieve the proportion of immunized required for herd immunity. This excess of infections increases the probability of emergence of vaccine resistance. For $R_0 = 3$, the fraction of immunized required for herd immunity is 2/3. Hence, 1/3 of the population can remain susceptible without affecting the probability of emergence of a vaccine resistance variant. However, with more infectious variants, this tolerated proportion of non-vaccinated is lower. For

example, for R = 5, the proportion of non-vaccinated such that the probability of vaccine emergence is unchanged is only 20%. As vaccination is expected to be extended to more age groups, the proportion of unvaccinated individuals could further decrease until vaccine hesitancy ceases to be a concern for emergence of vaccine resistance.

In most countries, vaccine hesitancy seems to be a function of age. Hence, we can estimate the expected vaccine hesitancy given the age structure of each of the six considered countries. Many countries have still not experienced a plateau in the total number of vaccinated for their younger population. The exception could be Israel. In Israel, 80.2% of the 20-29 year olds are vaccinated. In France, only about 50% of that age group is vaccinated, and the number of new vaccinations per day has not yet reached a plateau. We expect that many of the 20-29 year olds in France will still become vaccinated, thus increasing the total number of vaccinated people in the population.

Using boosters to counteract waning of immunity

The waning of immunity has recently become a concern. We have extended our model to consider waning of immunity after 180 days after vaccination or recovery (see **Figure S21** and **Figure S22**). When no booster vaccination is administered, the probability of emergence of vaccine resistance does increase substantially, especially for high vaccination rates. However, a booster vaccination campaign conducted after 180 days can reduce the probability of emergence of vaccine resistance back to basic model levels (see **Figure S23**).

SUMMARY

We have studied evolution of resistance to COVID-19 vaccination in the presence of dynamic lockdown. We use real world data to simulate the spread of the SARS-COV-2 virus. We have performed stochastic simulations and obtained analytical results. In particular, we have derived a simple intuitive formula for the probability of emergence of a vaccine resistant strain over time (Eqs. (5) and (6)).

Our basic model makes a series of simplifying assumptions, including: (i) no seasonality of the infection patterns; (ii) vaccination of the whole population; (iii) no waning of immunity; (iv) linear distribution of vaccine doses and (v) rapid social response to rising infection numbers. We have studied model extensions that remove those simplifications (see **Supplementary Materials Figures S1-S7, S12-S13, S16-S23**). We have assumed a stochastic model – this is because the appearance of a vaccine resistant strain, and in particular its non-extinction due to random drift, is by nature stochastic.

Our model most closely corresponds to the assumption that immune evasion could be due to a single point mutation. Nevertheless, our estimates of the mutation rate between the wild-type and immune evasive strains could signify that a combination of mutations is needed to achieve immune evasion. Therefore, we have explored lower effective mutation rates than the current estimation for the per-base mutation rate of the SARS-COV-2 virus.

The probability of takeover of an immune evasive strain is mostly dependent on the number of total infection cases that occur during the pandemic. Social distancing measures, such as lockdowns, can delay or event prevent the emergence of the MT strain. Each natural infection is an opportunity for the MT strain to appear and possibly take over. Hence, the main policy goal should be to maximize the proportion of the population which will be immunized to the virus through vaccination as opposed to natural infection.

In terms of policy implications, our result supports the maintenance of social distancing measures until the daily number of infections decreases substantially. Allowing a large number of infections can only be counterbalanced by very high vaccination rates, which ensure that herd immunity is reached before the MT strain can appear and takeover. Furthermore, our result underlines the importance of a worldwide effort to quickly vaccinate as many individuals as possible, especially in highly populated countries with low access to vaccines. Slow, or no vaccination, results in a large number of total cases in these areas and hence the emergence of an MT strain which could then spread over the whole world.

METHODS

Derivation of mathematical results

All of the stochastic simulations presented in **Figure 3** and **Figure 4** were run with the full model presented in **Figure 2**. However, in the following derivations, we make certain approximations in order to obtain analytical results. Among others, we neglect death and in some cases and we neglect vaccination of recovered individuals.

1. No Vaccination

First we consider the case without vaccination. We denote by x the number of susceptible individuals; by y the number of individuals infected with wildtype (WT); by z the number of individuals recovered from WT. The infection rate is β; the recovery rate α ; the mutation rate μ; the population size N. The social activity parameter

$s(t)$ captures the extent of imposed lockdown that varies over time. For simplicity, we neglect the number of individuals who die, hence the population size N is assumed to be constant.

Deterministic WT infection dynamics are given by the system of differential equations:

$$\dot{x} = -\beta s x y$$

$$\dot{y} = \beta s x y - a y \tag{7}$$

$$\dot{z} = a y$$

Initially, all of the population is susceptible to the WT strain, and no individuals are infected with or recovered from the WT strain. Therefore, we have: $x(0) = N$, $y(0) = 0$ and $z(0) = 0$. Social activity, $s(t)$, is adjusted such that $y(t) = L/a$ is constant (see **Figure S2C**). L is the number of new infections per day.

Without lockdown, $s = 1$, the basic reproductive ratio of the WT is given by $R_0 = \beta N/a$. If $R_0 > 1$, the number of infected individuals grows initially. With lockdown, $s < 1$, the reproductive ratio is $R_{WT} = \beta N s/a$. Since the social distancing measures maintain $y(t) = L/a$ at a constant value, we have $R_{WT} = 1$ and $\beta s(t) x(t) = a$ (see **Figure S2**). The parameter s can vary between 0 and 1.

Note that although R_{WT} and R_{MT} are initially equal, they are not equal at any later time point of the simulation. This is because the reproductive coefficient of the WT strain and MT strain are dependent on the number of individuals that can be infected by WT and the number of individuals that can be infected by MT. Initially, both strains can infect the whole population. As the simulation progresses, the WT strain can infect less individuals due to recovery of WT infected individuals and due to vaccination. On the other hand, the MT strain can still potentially infect every individual in the population since WT-recovered individuals and vaccinated individuals are susceptible to MT.

Each day, L individuals become infected and L individuals recover. Therefore, because of dynamic lockdown, the following equation is equivalent to (7):

$$\dot{x} = -L$$

$$\dot{y} = 0 \tag{8}$$

$$\dot{z} = L$$

The solution to this system of differential equations is

$$x(t) = N - Lt \tag{9}$$

$$z(t) = Lt$$

Hence, the number of susceptible individuals decreases linearly with slope L while the number of recovered individuals increases linearly with slope L. (See **Figure S2** for agreement with the stochastic simulation).

When $x(t)$ has declined such that $R_{WT} < 1$ and $s = 1$, there are not enough susceptible individuals to sustain the infection. This herd immunity is achieved when $x(t) < a/\beta$. Thus, the time T_H until herd immunity is given by $\beta(N - Lt) = a$. We obtain:

$$T_H = \frac{N}{L}\left(1 - \frac{1}{R_0}\right) \tag{10}$$

1.1 Rate of generating mutants in absence of vaccination

Each day, L new individuals become infected. Each of these infections has probability μ to be a vaccine-resistant MT. Hence, the rate of producing a mutant is $L\mu$ per day. Let $P(t)$ denote the probability that no mutant has been produced until time t. We have $\dot{P(t)} = -L\mu P(t)$, which leads to $P(t) = e^{-L\mu t}$.

The MT strain can be generated only during infection. Hence, if the MT strain has not been generated until the time when there are no more WT infections – that is, approximately when herd immunity is reached – it will never be generated. We neglect here the time of exponential decrease in the number of WT infections between time T_H (when herd immunity is reached) and the time when the number of WT infections has reached zero. The probability that no mutant will appear before time T_H is $P(T_H) = e^{-L\mu T_H}$. Inserting from eq (4) we obtain

$$P(T_H) = \exp\left(-N\mu\left(1 - \frac{1}{R_0}\right)\right) \tag{11}$$

1.2 Rate of generating surviving mutants in absence of vaccination

In order to calculate the probability that the MT strain will be generated and survive, we need to multiply the rate of generation of the MT strain with the probability that it will not become extinct by random drift. If $\rho(t)$ is the survival probability of the MT, then the rate of producing a surviving mutant is $L\mu\rho(t)$ per day. We approximate $\rho(t) = 1 - 1/R_{MT}(t)$, where $R_{MT}(t)$ is the reproductive ratio of the mutant at time T.

We have

$$R_{MT}(t) = \beta s(t) N/a \tag{12}$$

Since $s(t) = a/\beta x(t)$ and using Eq. (9) we obtain

$$R_{MT}(t) = \frac{N}{N - Lt} \tag{13}$$

And therefore we have $\rho(t) = Lt/N$.

Let $P(t)$ denote the probability that not surviving mutant has been produced until time t. We have $\dot{P(t)} = -L\mu\rho(t)P(t) = -L^2\mu tP(t)/N$. We can solve this differential equation to obtain $P(t) = \exp(-\mu L^2 t^2/2N)$. The probability that no surviving mutant has been produced until herd immunity, which is reached at time T_H, is given by

$$P(T_H) = \exp(-\frac{\mu N}{2}\left(1 - \frac{1}{R_0}\right)^2) \tag{14}$$

2. With Vaccination

Let us now add vaccination. Denote by w the number of vaccinated people. If both recovered and susceptible individuals are vaccinated at a total rate of c per day then deterministic infection and vaccination dynamics are given by

$$\begin{aligned} \dot{x} &= -\beta sxy - \frac{cx}{x+z} \\ \dot{y} &= \beta sxy - ay \\ \dot{z} &= ay - \frac{cz}{x+z} \\ \dot{w} &= c \end{aligned} \tag{15}$$

The initial condition is $x(0) = N$, $y(0) = 0$, $z(0) = 0$, $w(0) = 0$, $s(0) = 1$ and $R_0 = \beta N/a$. As before, we adjust $s(t)$ such that $y(t) = L/a$ is constant (see **Figure 3**).

Each day, L susceptible individuals become infected and $cx(x+z)$ susceptible individuals become vaccinated. Also, L infected individuals recover, and $cx(x+z)$ of recovered individuals become vaccinated. We have:

$$\begin{aligned} \dot{x} &= -L - \frac{cx}{x+z} \\ \dot{y} &= 0 \end{aligned} \tag{16}$$

$$\dot{z} = L - \frac{cz}{x+z}$$

$$\dot{w} = c$$

For simplicity let us assume that we only vaccinate susceptible people. This assumption is a reasonable approximation if $c \gg L$. In this case, we can write

$$\begin{aligned} \dot{x} &= -L - c \\ \dot{y} &= 0 \\ \dot{z} &= L \\ \dot{w} &= c \end{aligned} \tag{17}$$

The solution to this system of differential equations is

$$\begin{aligned} x(t) &= N - Lt - ct \\ z(t) &= Lt \\ w(t) &= ct \end{aligned} \tag{18}$$

Hence, the number of susceptible individuals decreases linearly with slope $L + c$, while the number of recovered individuals increases linearly with slope L, and the number of vaccinated individuals increases linearly with slope c.

The time T_H until herd immunity is given by

$$T_H = \frac{N}{c+L}(1 - 1/R_0) \tag{19}$$

2.1 Rate of generating mutants during vaccination

The rate of producing a mutant is $L\mu$ per day. Let $P(t)$ denote the probability that no mutant has been produced until time t. We have $\dot{P(t)} = -L\mu P(t)$, which gives $P(t) = \exp(-L\mu t)$.

The MT strain can be generated only during infection. Hence, if the MT strain has not been generated until the time when there are no more WT infections – that is, when herd immunity is reached – it will never be generated. Again we neglect here the time of exponential decrease in the number of WT infections between the time T_H when herd immunity is reached and the time where the number of WT infections reaches 0. Hence, the probability that no mutant will appear is $P(T_H) = \exp(-L\mu T_H)$. Using Eq. (19), the probability that no mutant has appeared until herd immunity is:

$$P(T_H) = \exp(-N\mu\left(\frac{L}{c+L}\right)\left(1-\frac{1}{R_0}\right)) \tag{20}$$

2.2 Rate of generating surviving mutants during vaccination

In order to calculate the probability that surviving mutants are generated, we again consider the survival probability $\rho(t) = 1 - 1/R_{MT}(t)$, where $R_{MT}(t)$ is the reproductive ratio of the mutant at time t. The rate of producing a surviving mutant is then $L\mu\rho(t)$ per day. We have:

$$R_{MT}(t) = \frac{\beta s(t) N}{a} \tag{21}$$

As explained above, $s(t) = a/\beta x(t)$. Using Eq. (18) we obtain

$$R_{MT}(t) = \frac{N}{N-(L+c)t} \tag{22}$$

And therefore $\rho(t) = (L+c)t/N$.

Let $P(t)$ denote the probability that not surviving mutant has been produced until time t. We have $\dot{P(t)} = -L\mu\rho(t)P(t) = -L\mu(c+L)tP(t)/N$. Let $v = c/N$ and $l = L/N$. We can solve this differential equation to obtain:

$$P(t) = \exp(-\frac{\mu N}{2} l(v+l)t^2) \tag{23}$$

The probability that no surviving mutant has been produced until herd immunity, which is reached at time T_H, is:

$$P(T_H) = \exp(-\frac{\mu N}{2}(\frac{l}{v+l})\left(1-\frac{1}{R_0}\right)^2) \tag{24}$$

2.3 Rate of generating surviving mutants with partial immune escape during vaccination

We can also study the case where the infectivity of the mutants is reduced by a factor $q \in [0,1]$ when infecting recovered or vaccinated people. For $q = 1$ we obtain full escape, while $q = 0$ means that the mutant does not escape at all.

A similar derivation to the one above leads to the following result. The probability that no surviving mutant with partial escape q has appeared until herd immunity is given by:

$$P(T_H) = \exp\left(-\frac{\mu N}{2}\left(\frac{l}{v+l}\right)A\right)$$

$$\text{with} \qquad A = \frac{2q}{1-q}\left(-\frac{R_0-1}{R_0} + \frac{1}{1-q}\log\frac{R_0}{1+q(R_0-1)}\right) \tag{25}$$

For $q = 1$ we obtain $A = (1-(1/R_0))^2$ leading to Eq. (23) above.

Relationship between the product formula and the exponential formula

Each day, L new WT infections occur. Each new infection has a probability of μ to be the MT strain. The survival probability of the mutant is approximately $1 - 1/R_m(t)$ where $R_m(t)$ is the basic reproductive ratio of the MT appearing at time t.

Hence, the probability that none of the L new WT infections in a day will generate a surviving mutant is $(1-\mu(1-1/R_m(t)))^L$. Then, we can write the probability P that no surviving mutant will be produced between time $t = 0$ and the time T_H when herd immunity is reached as the product

$$P = \prod_{\tau=0}^{T_H}\left[1-\mu\left(1-\frac{1}{R_{MT}(\tau)}\right)\right]^L \tag{26}$$

497 We have $T_H = [N/(c+L)](1 - 1/R_0)$ and $R_{MT}(t) = N/[N - (c+L)t]$.

498 Since $\rho(t) = 1 - 1/R_m(t) = (c+L)t/N$ we can write:

499 $$P = \prod_{\tau=0}^{T_H}(1 - \frac{\mu(c+L)\tau}{N})^L$$

500 Let us use the abbreviation $u = \mu(c+L)/N$. Then

501 $$P = \prod_{\tau=0}^{T_H}(1 - u\tau)^L$$

502 $$= \exp[log \prod_{\tau=0}^{T_H}(1 - u\tau)^L]$$

503 $$= \exp[L\ log \prod_{\tau=0}^{T_H}(1 - u\tau)]$$

$$= \exp[L \sum_{\tau=0}^{T_H} \log(1 - u\tau)] \quad (27)$$

504 Note that Eq. (26) is exactly equivalent to Eq. (23). Assuming $uT_H \ll 1$ which is the same as $\mu(1 - (1/R_0)) \ll 1$
505 we obtain

506 $$P = \exp[-uL \sum_{\tau=0}^{T_H} \tau]$$

507 $$= \exp\left[-\frac{uLT_H(T_H+1)}{2}\right]$$

508 Assuming $T_H \gg 1$ which is $N\left(1 - \frac{1}{R_0}\right) \gg c + L$, we obtain

509 $$P = \exp\left[-\frac{uL{T_H}^2}{2}\right]$$

510 $$= \exp\left[-\frac{(\mu(c+L)/N)L{T_H}^2}{2}\right]$$

511 Finally, inserting $T_H = (N/(c+L))(1 - 1/R_0)$ we get:

$$P(T_H) = \exp(-\frac{\mu N}{2}(\frac{l}{v+l})\left(1 - \frac{1}{R_0}\right)^2) \quad (28)$$

512 Which is equivalent to Eq.(24) (see above).

513

Dynamics after appearance of the MT strain

1. No Vaccination

After the MT strain has taken over, social distancing measures will continue maintaining the number of daily infections at L, which implies that $(y_1 + y_2) = L/a$ (see **Figure S2**). In practice, the WT strain rapidly goes extinct upon emergence of the MT strain; so we can consider $y_2 = L/a$ .The mutant strain can infect susceptible individuals x, and recovered individuals, z_1. The mutant strain infects those individuals with probabilities proportional to their frequencies at the time $t^*$of mutant takeover. Hence, for times $t > t^*$ we have:

$$x(t) = x(t^*) - \frac{x(t^*)}{z_1(t^*) + x(t^*)} L(t^* - t)$$
$$z_1(t) = z_1(t^*) - \frac{z_1(t^*)}{z_1(t^*) + x(t^*)} L(t^* - t) \quad (29)$$

After mutant takeover, the social distancing measures need to be readjusted to the mutant strain. Since more individuals are susceptible to it, $s(t)$ has to decrease (see **Figure S2F**):

$$s(t) = \frac{aN}{\beta(x(t) + qz_1(t))} \quad (30)$$

Which implies that $R_{MT} = 1$.

2. With vaccination

As for the case without vaccination, if the mutant strain survives, it will quickly replace the wild-type strain such that $y_2 = L/a$ (see **Figure 3C**). The number of susceptible individuals $x(t^*)$ at time of mutant takeover can be neglected for large enough vaccination rates. The number of vaccinated individuals, susceptible to the mutant strain w_1 will hence decrease linearly with the number of tolerated cases per day L, and the number of vaccinated individuals, recovered from the mutant strain w_2 will increase complementarily linearly with L. If the mutant takes over at time t^*, we have for all times $t > t^*$:

$$w_1(t) = w_1(t^*) - L(t^* - t) \quad (31)$$

$$w_2(t) = L(t^* - t)$$

The social activity parameter s needs readjustment to consider the additional groups of individuals that are now susceptible to the infecting strain. We have:

$$s(t) = \frac{a}{\beta} \frac{x(t) + q(z_1(t) + w_1(t)}{x(t) + q(z_1(t) + w_1(t)) - w_2(t)} \tag{32}$$

Which ensures that $R_{MT} = 1$. Here the parameter q in $[0,1]$ denotes the extent of escape.

Estimating the evolutionary potential of the virus

If μ is the mutation rate as described above and $L(t)$ is the time series giving the number of new infections on day t, then the probability that no mutant has been produced between time 0 and time T_H is given by:

$$P(T_H) = \prod_{\tau=0}^{T_H} [1 - \mu]^{L(\tau)} \tag{33}$$

This probability will overestimate the evolutionary potential of the virus to escape from vaccination because many mutants do not survive the initial random drift. The probability that no surviving mutant has been produced between time 0 and time T_H can be written as:

$$P(T_H) = \prod_{\tau=0}^{T_H} [1 - \mu\rho(\tau)]^{L(\tau)} \tag{34}$$

Here $\rho(t)$ is the survival probability of an escape mutant produced at time t. This probability depends on the basic reproductive ratio of the mutant on the day it is being produced (and the next few days until random drift is negligible). Approximately we can write:

$$\rho(t) = \min\{0, 1 - \frac{1}{R_M(t)}\} \tag{35}$$

For the potential of the virus to generate mutants (irrespective of whether they survive) what matters most is the total number of infections, $\sum_\tau L(\tau)$. But for the potential of the virus to generate surviving mutants one must also consider the time periods when lockdown is relaxed such that R_{MT} is above 1.

Analytic approximation for more infectious, vaccine escape mutants

1. No Vaccination

Now we calculate the probability that mutants are being generated which do become extinct by random drift. We denote f the relative infectiousness of the mutant vs. the wild type strain. Hence, for more infectious mutants, we have $f > 1$.

The rate of producing a surviving mutant is $L\mu\rho(t)$ per day. Here $\rho(t)$ is the survival probability, given by $\rho(t) = 1 - 1/R_M(t)$. The basic reproductive ratio of the mutant at time t is: $R_M(t) = f\beta s(t)N/a$. Since we have $\beta s(t)N/a = 1/x(t)$. We have $R_M(t) = fN/(N - Lt)$ and therefore:

$$\rho(t) = 1 - \frac{1}{R_M(t)} = \frac{N(f-1) + Lt}{Nf} \tag{36}$$

Let $P(t)$ denote the probability that no surviving mutant has been produced until time t.

We have $\dot{P}(t) = -L\mu\rho(t)P(t)$. Thus, $\dot{P}(t) = (\kappa + \lambda t)P(t)$, with the solution $P(t) = \exp(\kappa t + \frac{\lambda}{2}t^2)$.

Which already leads to $P(0) = 1$ as desired. In our original notation, the solution becomes:

$$P(t) = \exp(-L\mu\frac{N(f-1)}{Nf}t - L\mu\frac{Lt}{2Nf}t^2) \tag{37}$$

2. With vaccination

The rate of producing a surviving mutant is $L\mu\rho(t)$ per day. Here $\rho(t)$ is the survival probability given by $\rho(t) = 1 - 1/R_m(t)$. The basic reproductive ratio of the mutant at time t is $R_M(t) = f\beta s(t)N/a$. Since $\beta s(t)N/a = 1/x(t)$. We have $R_M(t) = fN/[N - (c + L)t]$ and:

$$\rho(t) = 1 - \frac{1}{R_M(t)} = \frac{N(f-1) + (c+L)t}{Nf} \tag{38}$$

Let $P(t)$ denote the probability that no surviving mutant has been produced until time t. We have $\dot{P}(t) = -L\mu\rho(t)P(t)$. Thus:

$$\dot{P}(t) = -L\mu\frac{N(f-1)+(c+L)t}{Nf}P(t) \tag{39}$$

This solution of this differential equation is given by:

$$P(t) = \exp(-L\mu\frac{f-1}{f}t - L\mu\frac{c+L}{2Nf}t^2) \tag{40}$$

REFERENCES

[1] J. T. Wu *et al.*, "Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China," *Nat. Med.*, 2020, doi: 10.1038/s41591-020-0822-7.

[2] "COVID-19 Data Explorer - Our World in Data." https://ourworldindata.org/explorers/coronavirus-data-explorer?zoomToSelection=true&time=2020-03-01..latest&country=USA~GBR~CAN~DEU~ITA~IND&region=World&pickerMetric=location&pickerSort=asc&Metric=Confirmed+cases&Interval=7-day+rolling+average&Align+outbreaks=false&Relative+to+Population=true (accessed Mar. 19, 2021).

[3] "Different COVID-19 Vaccines | CDC." https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines.html (accessed Mar. 19, 2021).

[4] "Safe COVID-19 vaccines for Europeans | European Commission." https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/safe-covid-19-vaccines-europeans_en (accessed Mar. 19, 2021).

[5] L. Böttcher and J. Nagler, "Decisive Conditions for Strategic Vaccination against SARS-CoV-2," *medRxiv*, p. 2021.03.05.21252962, Mar. 2021, doi: 10.1101/2021.03.05.21252962.

[6] E. E. Walsh, J. H. Shin, and A. R. Falsey, "Clinical impact of human coronaviruses 229E and OC43 infection in diverse adult populations," *J. Infect. Dis.*, 2013, doi: 10.1093/infdis/jit393.

[7] "Vaccines – COVID19 Vaccine Tracker." https://covid19.trackvaccines.org/vaccines/ (accessed Mar. 19, 2021).

[8] H. Tegally *et al.*, "Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa," *medRxiv*, vol. 10. medRxiv, p. 2020.12.21.20248640, Dec. 22, 2020, doi: 10.1101/2020.12.21.20248640.

[9] A. Rambaut *et al.*, "Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations - SARS-CoV-2 coronavirus / nCoV-2019 Genomic Epidemiology - Virological," 2020. https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563 (accessed Apr. 16, 2021).

[10] P. Wang *et al.*, "Increased Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7 to Antibody Neutralization.," *bioRxiv Prepr. Serv. Biol.*, 2021, doi: 10.1101/2021.01.25.428137.

[11] L. L. Luchsinger and C. D. Hillyer, "Vaccine efficacy probable against COVID-19 variants," *Science (80-.).*, vol. 371, no. 6534, pp. 1116 LP – 1116, Mar. 2021, doi: 10.1126/science.abg9461.

[12] T. N. Starr *et al.*, "Prospective mapping of viral mutations that escape antibodies used to treat COVID-19," *Science (80-.).*, vol. 371, no. 6531, pp. 850–854, Feb. 2021, doi: 10.1126/science.abf9302.

[13] T. N. Starr *et al.*, "Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on

Folding and ACE2 Binding," *Cell*, vol. 182, no. 5, pp. 1295-1310.e20, Sep. 2020, doi: 10.1016/j.cell.2020.08.012.

[14] A. J. Greaney *et al.*, "Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition," *Cell Host Microbe*, vol. 29, no. 1, pp. 44-57.e9, Jan. 2021, doi: 10.1016/j.chom.2020.11.007.

[15] R. N. Thompson, E. M. Hill, and J. R. Gog, "SARS-CoV-2 incidence and vaccine escape," *Lancet Infect. Dis.*, Apr. 2021, doi: 10.1016/S1473-3099(21)00202-4.

[16] P. J. Gerrish, F. Saldaña, B. Galeota-Sprung, A. Colato, E. E. Rodriguez, and J. X. V. Hernández, "How unequal vaccine distribution promotes the evolution of vaccine escape," *medRxiv*, 2021.

[17] S. Cobey, D. B. Larremore, Y. H. Grad, and M. Lipsitch, "Concerns about SARS-CoV-2 evolution should not hold back efforts to expand vaccination," *Nature Reviews Immunology*. Nature Research, pp. 1–6, Apr. 01, 2021, doi: 10.1038/s41577-021-00544-9.

[18] F. Geoffroy, A. Traulsen, and H. Uecker, "Vaccination strategies when vaccines are scarce: On conflicts between reducing the burden and avoiding the evolution of escape mutants," *medRxiv*, p. 2021.05.04.21256623, May 2021, doi: 10.1101/2021.05.04.21256623.

[19] R. E. Baker, W. Yang, G. A. Vecchi, C. J. E. Metcalf, and B. T. Grenfell, "Assessing the influence of climate on wintertime SARS-CoV-2 outbreaks," *Nat. Commun.*, vol. 12, no. 1, pp. 1–7, Dec. 2021, doi: 10.1038/s41467-021-20991-1.

[20] J. L. Geoghegan *et al.*, "Genomic epidemiology reveals transmission patterns and dynamics of SARS-CoV-2 in Aotearoa New Zealand," *Nat. Commun.*, vol. 11, no. 1, pp. 1–7, Dec. 2020, doi: 10.1038/s41467-020-20235-8.

[21] A. B. Komissarov *et al.*, "Genomic epidemiology of the early stages of the SARS-CoV-2 outbreak in Russia," *Nat. Commun.*, vol. 12, no. 1, pp. 1–13, Dec. 2021, doi: 10.1038/s41467-020-20880-z.

[22] N. R. Faria *et al.*, "Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil," *Science (80-.).*, p. eabh2644, Apr. 2021, doi: 10.1126/science.abh2644.

[23] K. A. Lythgoe *et al.*, "SARS-CoV-2 within-host diversity and transmission," *Science (80-.).*, vol. 372, no. 6539, p. eabg0821, Mar. 2021, doi: 10.1126/science.abg0821.

[24] N. L. Komarova, L. M. Schang, and D. Wodarz, "Patterns of the COVID-19 pandemic spread around the world: Exponential versus power laws: Patterns of the COVID-19 pandemic spread around the world: Exponential versus power laws," *J. R. Soc. Interface*, vol. 17, no. 170, Sep. 2020, doi: 10.1098/rsif.2020.0518.

[25] C. M. Saad-Roy *et al.*, "Immune life history, vaccination, and the dynamics of SARS-CoV-2 over the next 5 years," *Science (80-.).*, vol. 370, no. 6518, pp. 811–818, Nov. 2020, doi: 10.1126/science.abd7343.

[26] BERNOULLI and D., "Essai d'une nouvelle analyse de la mortalite causee par la petite verole, et des avantages de l'inoculation pour la prevenir," *Hist. l'Acad., Roy. Sci. avec Mem*, pp. 1–45, 1760, Accessed: Mar. 30, 2021. [Online]. Available: https://ci.nii.ac.jp/naid/10013124549.

[27] W. 0 Kermack and A. G. Mckendrick, "A contribution to the mathematical theory of epidemics," *Proc. R. Soc. London. Ser. A, Contain. Pap. a Math. Phys. Character*, vol. 115, no. 772, pp. 700–721, Aug. 1927, doi: 10.1098/rspa.1927.0118.

[28] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, "On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations," *J. Math. Biol.*, vol. 28, no. 4, pp. 365–382, 1990, doi: 10.1007/BF00178324.

[29] K. Dietz, "The estimation of the basic reproduction number for infectious diseases," *Stat. Methods Med. Res.*, vol. 2, no. 1, pp. 23–41, Mar. 1993, doi: 10.1177/096228029300200103.

[30] M. A. Nowak and R. M. May, "Superinfection and the evolution of parasite virulence," *Proc. R. Soc. B Biol. Sci.*, vol. 255, no. 1342, pp. 81–89, 1994, doi: 10.1098/rspb.1994.0012.

[31] M. A. Nowak and C. R. M. Bangham, “Population dynamics of immune responses to persistent viruses,” *Science (80-.).*, vol. 272, no. 5258, pp. 74–79, 1996, doi: 10.1126/science.272.5258.74.

[32] H. W. Hethcote, “Mathematics of infectious diseases,” *SIAM Rev.*, vol. 42, no. 4, pp. 599–653, 2000, doi: 10.1137/S0036144500371907.

[33] F. Brauer and C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology*, vol. 40. New York, NY: Springer New York, 2012.

[34] W. Hamer, “The Milroy Lectures ON EPIDEMIC DISEASE IN ENGLAND-THE EVIDENCE OF VARIABILITY AND OF PERSISTENCY OF TYPE.,” *The Lancet*, vol. 167, no. 4305. Elsevier, pp. 569–574, Mar. 03, 1906, doi: 10.1016/S0140-6736(01)80187-2.

[35] N. Bailey, “The Mathematical Theory of Infectious Diseases and its applications,” Wiley-Blackwell, 1978. Accessed: Mar. 30, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1457188/.

[36] H. E. Tillett, “Infectious Diseases of Humans; Dynamics and Control,” *Epidemiol. Infect.*, vol. 108, no. 1, p. 211, 1992, Accessed: Mar. 30, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2272191/.

[37] M. Nowak and R. M. May, “Virus Dynamics: Mathematical Principles of Immunology And Virology.” https://www.researchgate.net/publication/48378467_Virus_Dynamics_Mathematical_Principles_of_Immunology_And_Virology (accessed Mar. 30, 2021).

[38] O. Diekmann and J. A. P. Heesterbeek, “Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation.” https://www.researchgate.net/publication/48376881_Mathematical_Epidemiology_of_Infectious_Diseases_Model_Building_Analysis_and_Interpretation (accessed Mar. 30, 2021).

[39] M. Ragonnet-Cronin *et al.*, “Genetic evidence for the association between COVID-19 epidemic severity and timing of non-pharmaceutical interventions,” *Nat. Commun.*, vol. 12, no. 1, p. 2188, Dec. 2021, doi: 10.1038/s41467-021-22366-y.

[40] H. Rossman *et al.*, “Hospital load and increased COVID-19 related mortality in Israel,” *Nat. Commun.*, vol. 12, no. 1, p. 1904, Dec. 2021, doi: 10.1038/s41467-021-22214-z.

[41] C. M. Saad-Roy *et al.*, “Epidemiological and evolutionary considerations of SARS-CoV-2 vaccine dosing regimes,” *Science (80-.).*, p. eabg8663, Mar. 2021, doi: 10.1126/science.abg8663.

[42] C. M. Saad-Roy *et al.*, “Immune life history, vaccination, and the dynamics of SARS-CoV-2 over the next 5 years,” *Science (80-.).*, vol. 370, no. 6518, pp. 811–818, Nov. 2020, doi: 10.1126/science.abd7343.

[43] P. Ashcroft, S. Lehtinen, D. C. Angst, N. Low, and S. Bonhoeffer, “Quantifying the impact of quarantine duration on covid-19 transmission,” *Elife*, vol. 10, pp. 1–33, Feb. 2021, doi: 10.7554/eLife.63704.

[44] N. L. Komarova, A. Azizi, and D. Wodarz, “Network models and the interpretation of prolonged infection plateaus in the COVID19 pandemic,” *Epidemics*, p. 100463, May 2021, doi: 10.1016/j.epidem.2021.100463.

[45] M. Castro, S. Ares, J. A. Cuesta, and S. Manrubia, “The turning point and end of an expanding epidemic cannot be precisely forecast,” *Proc. Natl. Acad. Sci.*, vol. 117, no. 42, pp. 26190–26196, Oct. 2020, doi: 10.1073/PNAS.2007868117.

[46] M. Stich, S. C. Manrubia, and E. Lázaro, “Variable mutation rates as an adaptive strategy in replicator populations,” *PLoS One*, 2010, doi: 10.1371/journal.pone.0011186.

[47] O. Yagan, A. Sridhar, R. Eletreby, S. A. Levin, J. B. Plotkin, and H. V. Poor, “Modeling and Analysis of the Spread of COVID-19 Under a Multiple-Strain Model with Mutations,” *Harvard Data Sci. Rev.*, 2021, doi: 10.1162/99608f92.a11bf693.

[48] S. Lehtinen, P. Ashcroft, and S. Bonhoeffer, “On the relationship between serial interval, infectiousness profile and generation time,” *J. R. Soc. Interface*, 2021, doi: 10.1098/rsif.2020.0756.

[49] H. W. Hethcote, “Three Basic Epidemiological Models,” 1989.

[50] D. T. Gillespie, “A general method for numerically simulating the stochastic time evolution of coupled chemical reactions,” *J. Comput. Phys.*, vol. 22, no. 4, pp. 403–434, Dec. 1976, doi: 10.1016/0021-9991(76)90041-3.

[51] D. T. Gillespie, “Exact stochastic simulation of coupled chemical reactions,” in *Journal of Physical Chemistry*, 1977, vol. 81, no. 25, pp. 2340–2361, doi: 10.1021/j100540a008.

[52] W. Cota and S. C. Ferreira, “Optimized Gillespie algorithms for the simulation of Markovian epidemic processes on large and heterogeneous networks,” *Comput. Phys. Commun.*, vol. 219, pp. 303–312, Oct. 2017, doi: 10.1016/j.cpc.2017.06.007.

[53] R. I. Mukhamadiarov *et al.*, “Social distancing and epidemic resurgence in agent-based susceptible-infectious-recovered models,” *Sci. Rep.*, vol. 11, no. 1, p. 130, Dec. 2021, doi: 10.1038/s41598-020-80162-y.

TABLES

Parameter	Symbol	Biologically Significant Range	References/Notes
Wildtype R_0	β_1	$[1.8 - 3.6]$	[50]
Mutant R_0	β_2	$[1.8 - 3.6]$ [3.6-7.2] (more infectious variants)	[50], [51]
Recovery Rate	a	$[0.38 - 0.08]$	[52]
Social Distancing Parameter	s	$[0.05, 1]$	-
Death Rate	d	$[0.005 - 0.02]$	[2]
Population Size	N	$> 10^6$	-
Partial Immunity Coefficient	q	$[0,1]$	-
Mutation Rate	μ	$< 10^{-6}$	This mutation rate is the composite mutation rate from a vaccine sensitive phenotype to a vaccine resistant phenotype. See section: "Estimating the mutation rate μ for estimate.
Number of Vaccinated per day per million	c	$[0 - 10{,}000]$	[2]
Number of Infected per day per million	L	$10 - 1000$	[2]

Table 1: Summary of model's parameters and their biologically significant ranges. The mutation rate refers to the probability that someone infected with the wild type strain gives rise to someone infected with a vaccine resistant strain. Hence, its value is different from the SARS-COV2 molecular mutation rate that occurs during one replication cycle.

Rate of infection l per day per person ($\times \mathbf{10^{-6}}$)	Rate of vaccination c per day per person ($\times \mathbf{10^{-6}}$)	Time to herd immunity, $\boldsymbol{T_H}$ (in days)	Probability of resistance			
			t = 50 days	t = 100 days	t = 200 days	t = T_H
100	1000	606	0.001	0.005	0.022	0.183
100	5000	131	0.006	0.025	-	0.043
100	10,000	66	0.013	-	-	0.022
200	1000	556	0.003	0.012	0.047	0.310
200	5000	128	0.013	0.051	-	0.082
200	10,000	65	0.025	-	-	0.043
500	1000	444	0.009	0.037	0.139	0.532
500	5000	121	0.034	0.128	-	0.183
500	10,000	63	0.064	-	-	0.100

$N = 10^8; \mu = 10^{-7}$

Rate of infection l per day per person ($\times \mathbf{10^{-6}}$)	Rate of vaccination c per day per person ($\times \mathbf{10^{-6}}$)	Time to herd immunity, $\boldsymbol{T_H}$ (in days)	Probability of resistance			
			t = 50 days	t = 100 days	t = 200 days	t = T_H
100	1000	606	0.014	0.054	0.197	0.867
100	5000	131	0.062	0.225	-	0.353
100	10,000	66	0.119	-	-	0.197
200	1000	556	0.030	0.113	0.381	0.975
200	5000	128	0.122	0.405	-	0.575
200	10,000	65	0.225	-	-	0.353
500	1000	444	0.089	0.313	0.777	0.999
500	5000	121	0.291	0.747	-	0.867
500	10,000	63	0.481	-	-	0.652

$N = 10^9; \mu = 10^{-7}$

Table 2: Calculated probability of vaccine resistance for a range of vaccination rates and infection rates. We observe of counterintuitive effect of higher probability of resistance along time for higher vaccination rates, but lower probability of resistance overall. See also Figure S4.

Country	Population size ($\times 10^6$)	Average rate of infection per day per person before vaccination ($\times 10^{-6}$)	Average rate of infection per day per person after vaccination ($\times 10^{-6}$)	Average rate of vaccination per day per person ($\times 10^{-6}$)	Probability of resistance ($\mu = 10^{-7}$)
Brazil	212.6	128	259	3546	0.208
France	68.1	141	267	5142	0.125
Germany	83.8	64	117	5067	0.043
Israel	8.7	160	273	5853	0.026
United Kingdom	67.9	109	265	5819	0.157
United States	331.0	200	240	4907	0.454

Table 3: Calculated probability of emergence of vaccine resistance using real-world data from six countries: Brazil, France, Germany, Israel, the United Kingdom and the United States. The probability of vaccine resistance was calculated using the product formula in Eq. 2 and the data presented in Figure 5 assuming $\mu = 10^{-7}$.

FIGURES

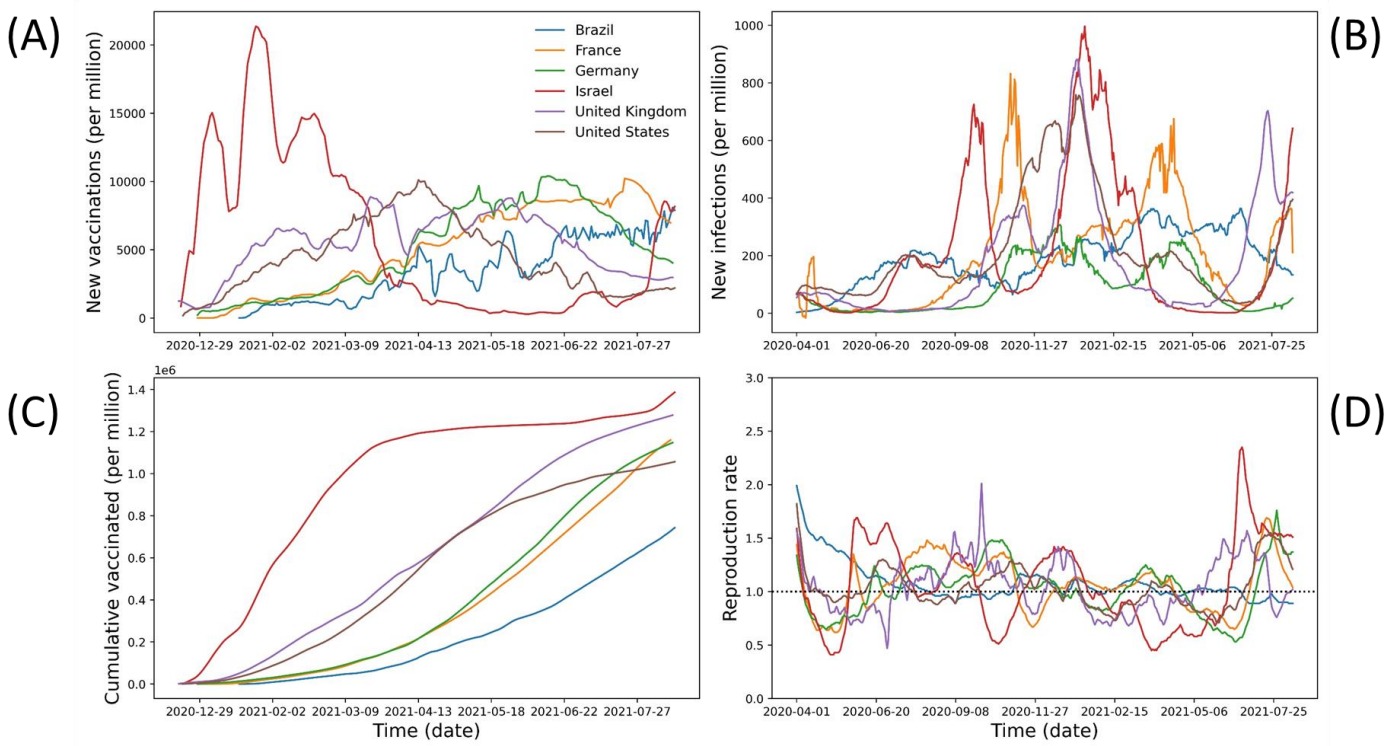

Figure 1: Vaccination, infection and basic reproductive ratio of SARS-CoV2 for Brazil, France, Germany, Israel, UK and USA. (A) Large scale vaccination programs commenced in December 2020. At the peak, Israel vaccinated more than 20,000 people per million (2%) per day. The vaccination rate decreased in April 2021 after most eligible individuals had been vaccinated. **(B)** In an attempt to balance economic and sanitary considerations, these six countries have gone through several cycles of loosening and tightening government-imposed restrictions, resulting in periodical increases and decreases in the number of new infections per day. The so-called "British variant", identified in November 2020, is most probably responsible for the increase in the number of infections in the UK and Israel at that time. The fourth wave in Israel (starting in July 2021) has been attributed to a combination of the emergence of the "Delta variant" and waning of immunity after vaccination. **(C)** Vaccination in most countries follows a logistic-shaped curve. At first, priority groups are vaccinated. Then, most of the population is vaccinated in a short timeframe (exponential growth). Lastly, the total number of vaccinated individuals plateaus due to vaccine hesitancy/ineligibility of the remaining non-vaccinated individuals. **(D)** The reproductive rate of SARS-COV2 infectious has hovered around 1 in the six considered countries due to the imposition of NPIs to curb the infection rates. Hence, the number of infections per day has fluctuated around an average for the duration of the pandemic.

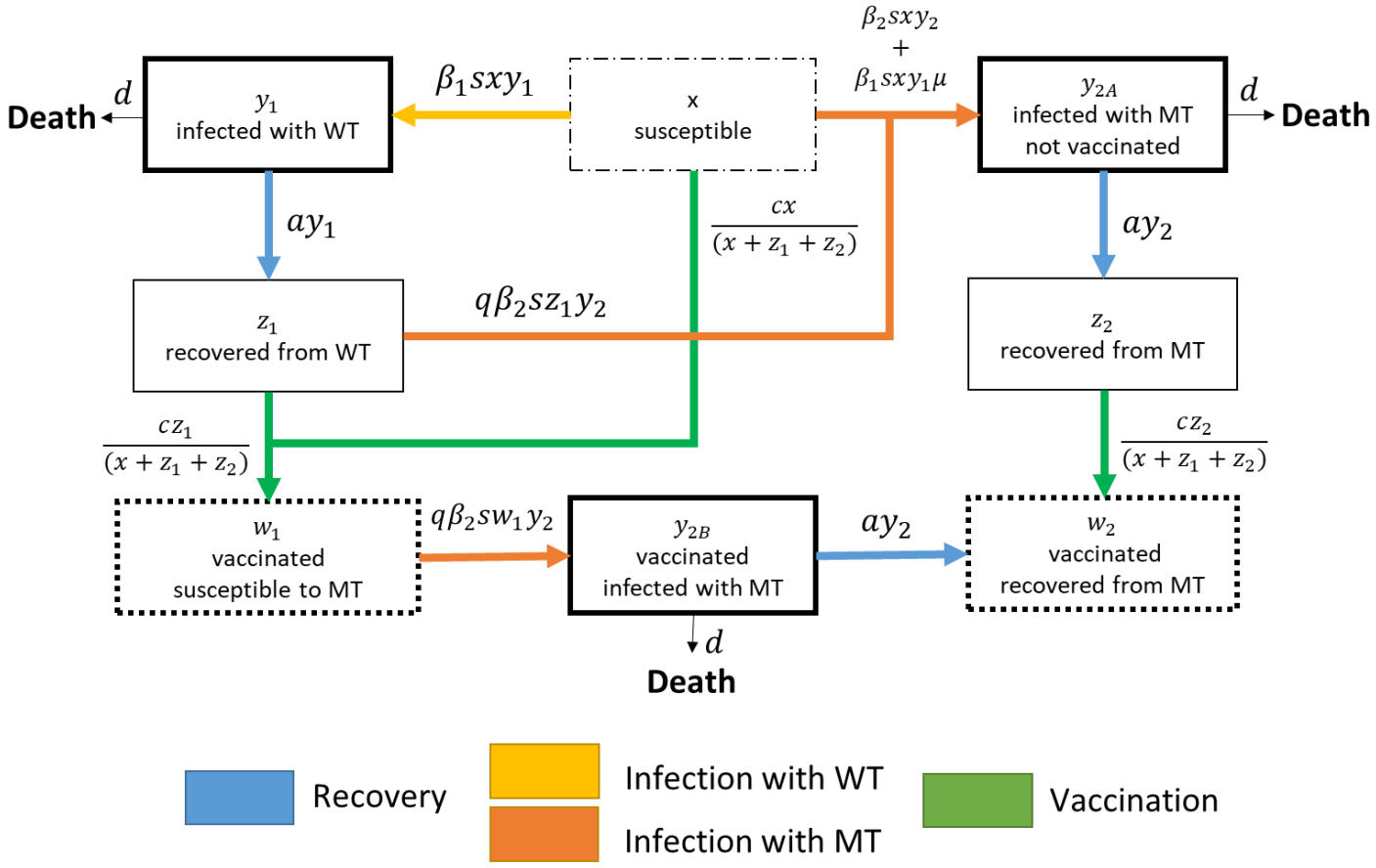

Figure 2: Infection dynamics, vaccination and resistance. Susceptible individuals (x) can be infected by wildtype (WT) or mutant (MT) virus. Infected people (y_1, y_{2A}, y_{2B}) die (at rate d) or recover (at rate a). People recovered from WT or vaccinated against WT can be infected by MT. People recovered from MT cannot be infected by WT. In our simplest model, we assume equal infectivity, recovery and death rates for both WT and MT. Vaccination occurs at rate c per day for all unvaccinated individuals (excluding those that are currently in active infection). Mutation happens (at rate μ) when exposure to a WT infected individual (y_1) results in the generation of a MT infected individual. Note that when exposure to a WT infected individual (y_1) results in the generation of a WT infected individual, the rate of infection should be multiplied by $1 - \mu$ in order to conserve the sum of mutation probabilities at 1. However, since μ is small, we neglect the term $1 - \mu$. The rates of these events are indicated on the arrows and are used in the Gillespie algorithm implementing stochastic dynamics.

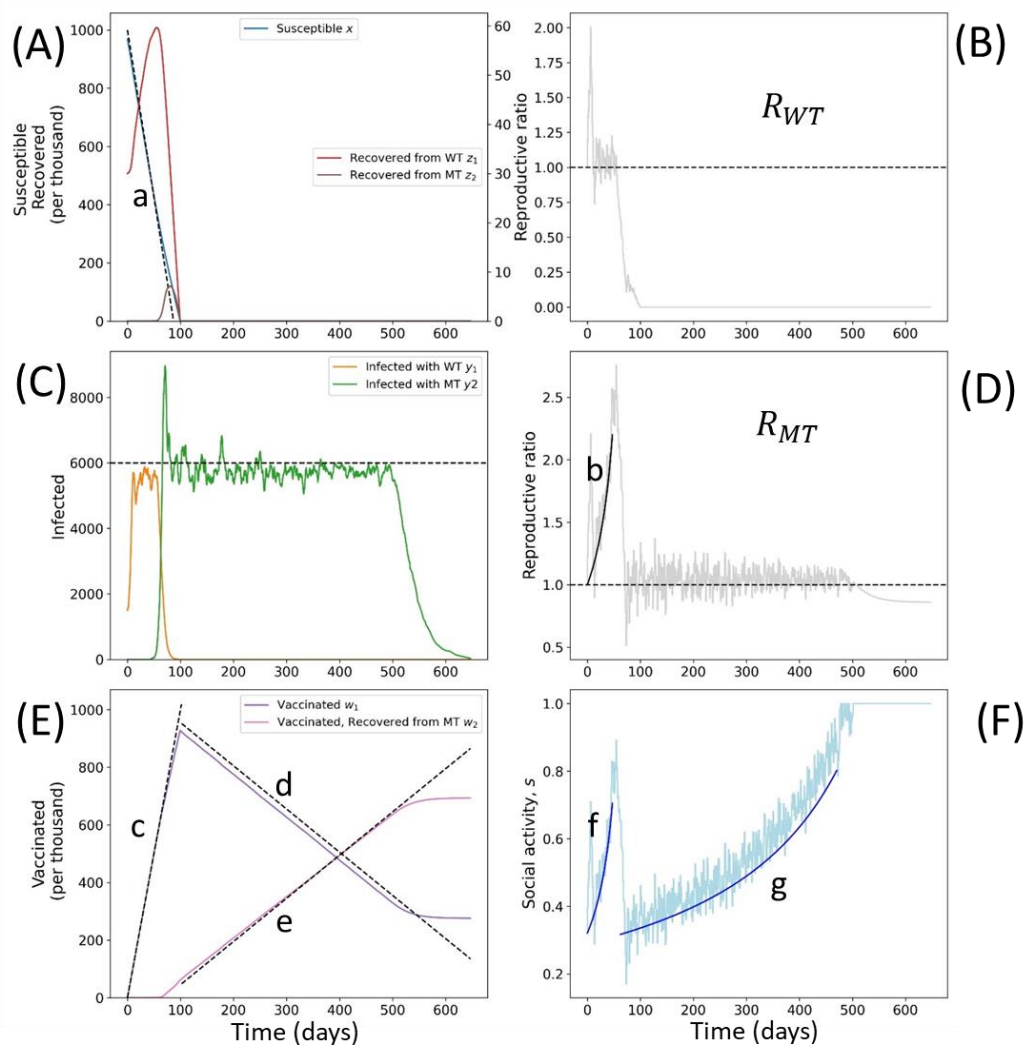

Figure 3: Evolution of resistance in presence of vaccination. (A) Before MT takeover, the decline in susceptible individuals (x) can be approximated by a linear function with slope equal to the vaccination rate c. Since vaccination is fast, individuals recovered from WT and non-vaccinated individuals recovered from MT are few. The equation of line (a) is $x(t) = x(0) - ct$ for $t < t^*$ where t^* is the time of takeover of the MT. (B) The reproductive rate R_{WT} is maintained at around 1 by dynamic lockdown. After mutant takeover, R_{WT} is less than 1, since the lockdown is now adjusted to the population susceptible to the MT strain. (C) The number of active WT infections before takeover and of active MT infections after takeover, is fluctuating around L/a until herd immunity to the MT is reached. (D) Before MT takeover, the reproductive rate of the MT grows as (b) $R_{MT} = \beta_2(x(t) + z_1(t) + w_1(t))/a$. After takeover, R_{MT} is maintained around 1. (E) The number of vaccinated individuals (w_1) first increases linearly with slope equal to the vaccination rate. After MT takeover, the number of individuals vaccinated to the WT and recovered from MT (w_2) increased linearly with slope L. The equations of the lines are given by (c) $w_1(t) = ct$ for $t < t^*$ (d) $w_1(t) = w_1(t^*) - L(t - t^*)$ for $t > t^*$ (e) $w_2(t) = L(t - t^*)$ for $> t^*$. (F) Before MT takeover, the dynamic lockdown is adjusted to the WT. As the number of individuals immune to WT grows, social activity increases. When the MT emerges, lockdown measures are reinstated. Subsequently, social activity increases as the population immune to the MT grows. The equations for the lines given by (f) $s(t) = a/\beta_1 x(t)$ for $t < t^*$; (g) $s(t) = a/\beta_2(x(t) + z_1(t) + w_1(t))$ for $> t^*$. Parameters: $N = 10^6$; $a = 0.25$; $d = 0.01$; $\mu = 10^{-6}$; $s_0 = 0.1$; $\beta_1 = \beta_2 = 7.5 \cdot 10^{-7}$; $c = 10{,}000$; $L = 1500$; $q = 1$.

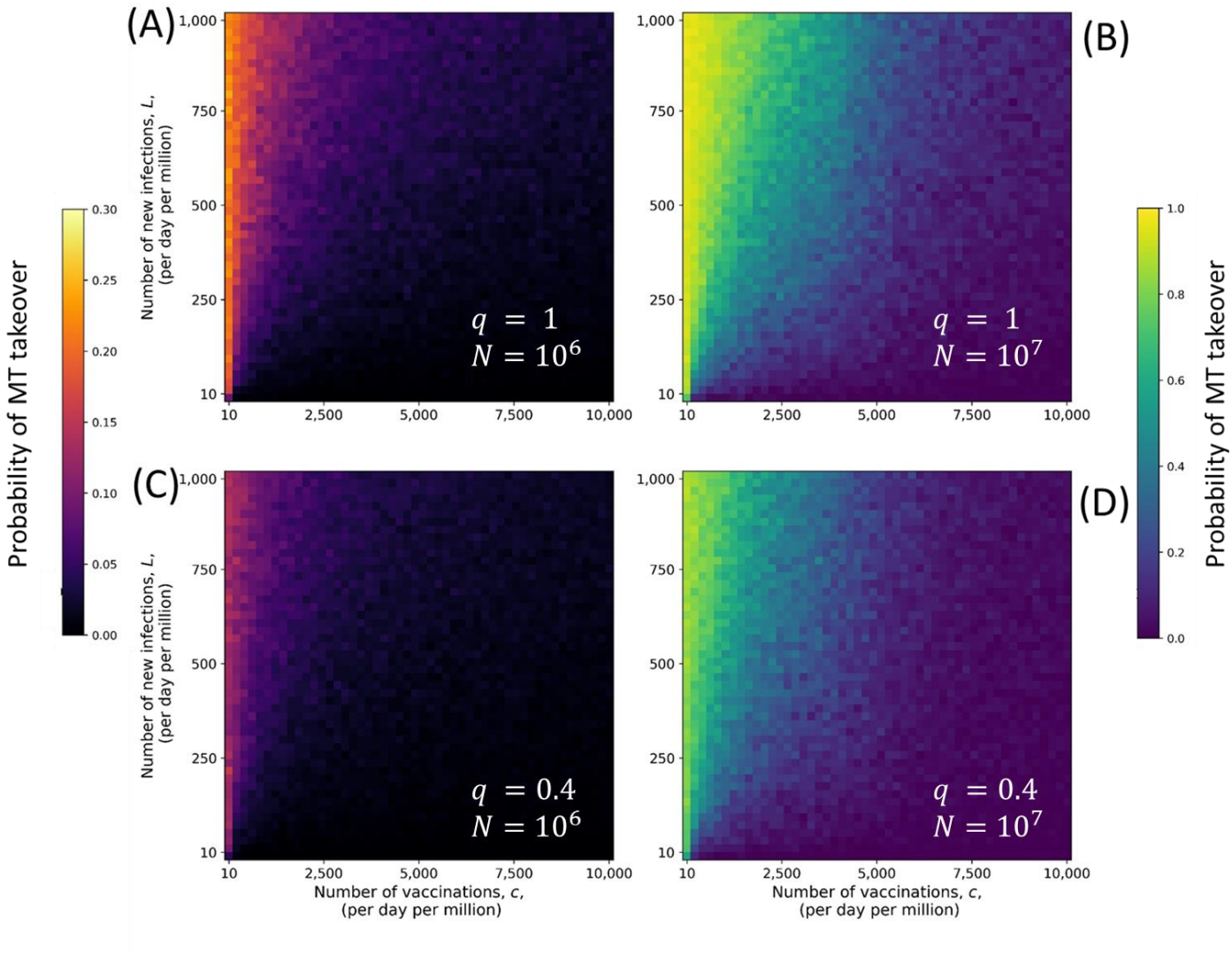

Figure 4: Probability of emergence of resistance. For each square of the color maps, the proportion of runs (out of 1000 runs) where the number of individuals infected with the MT strain exceeded the number of individuals infected with the WT strain is recorded. All simulations are run for a population size of $N = 10^6$, then scaled to obtain the results shown for $N = 10^7$. Results for color maps (B) and (D) were scaled according to $(1 - (1 - p)^{10})$, where *p* is the proportion of runs where the MT strain took over. We observe a triangular shape of (L, c) parameter sets for which the MT strain takes over, indicating that high vaccination rates can be safely associated with more lenient social distancing measures. On the other hand, very slow vaccination cannot be compensated by any strength of lockdown. Partial immunity to the MT strain - (C) and (D) - does not affect the shape of the parameter space where we observe MT takeover, but reduces its probability. Parameters: $a = 0.25; d = 0.01; \mu = 10^{-6}; s_0 = 0.1; \beta_1 = \beta_2 = 7.5 \cdot 10^{-7}$.

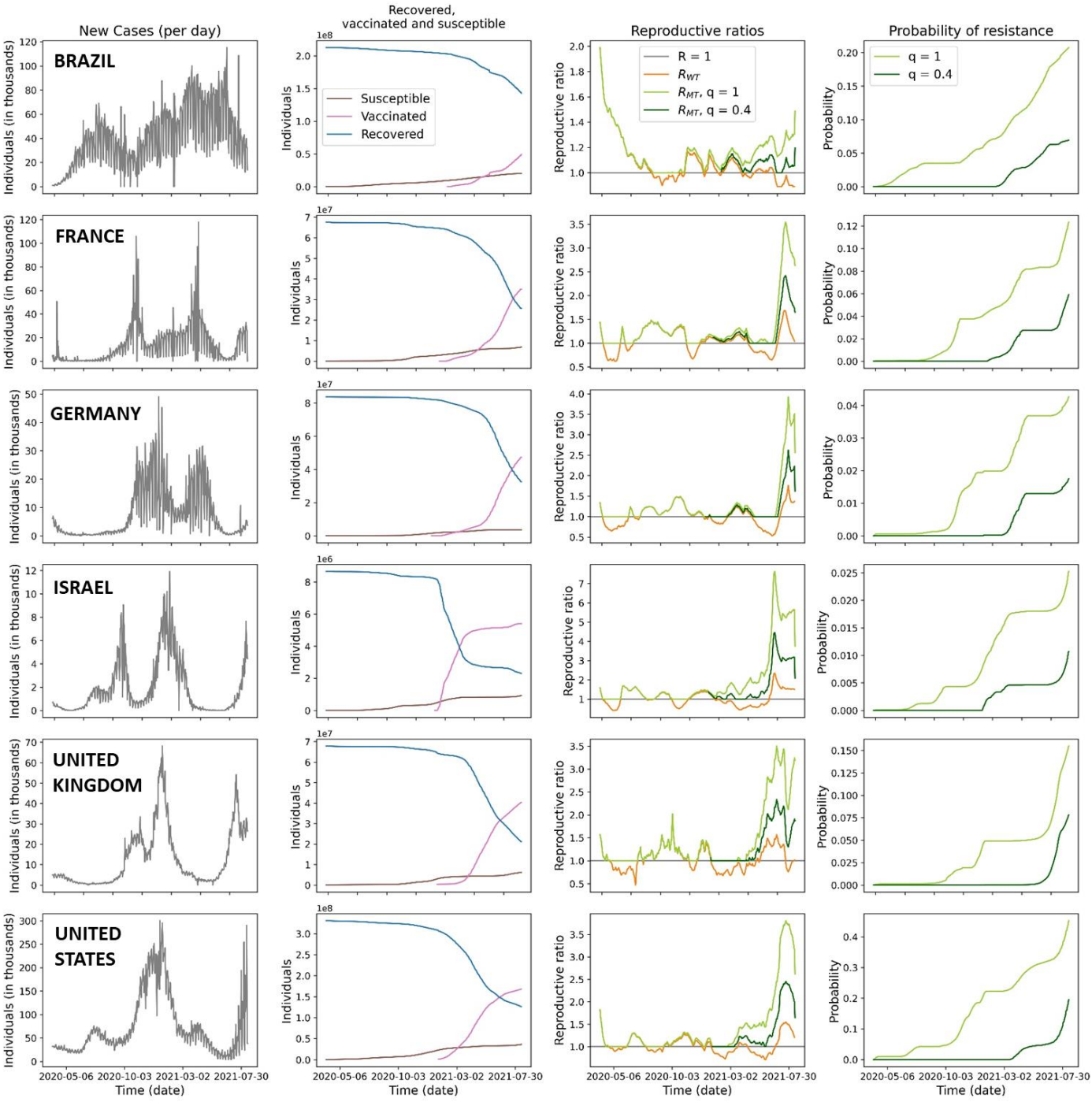

Figure 5: Infection, vaccination data, reproductive ratios and probability of resistance for Brazil, France, Germany, Israel, UK, and US. The total number of new cases per day, number of susceptible, recovered, and vaccinated individuals was downloaded from the *OWID* (Our World In Data) database. We used the *OWID* estimate for the WT reproductive coefficient R_{WT} to calculate the potential MT reproductive coefficient R_{MT} for a full escape mutant, $q = 1$ and a partial escape mutant, $q = 0.4$, (third column). We use Eq. 2 to estimate the probability that an escape mutant would have emerged by time t assuming $\mu = 10^{-7}$ (fourth column).

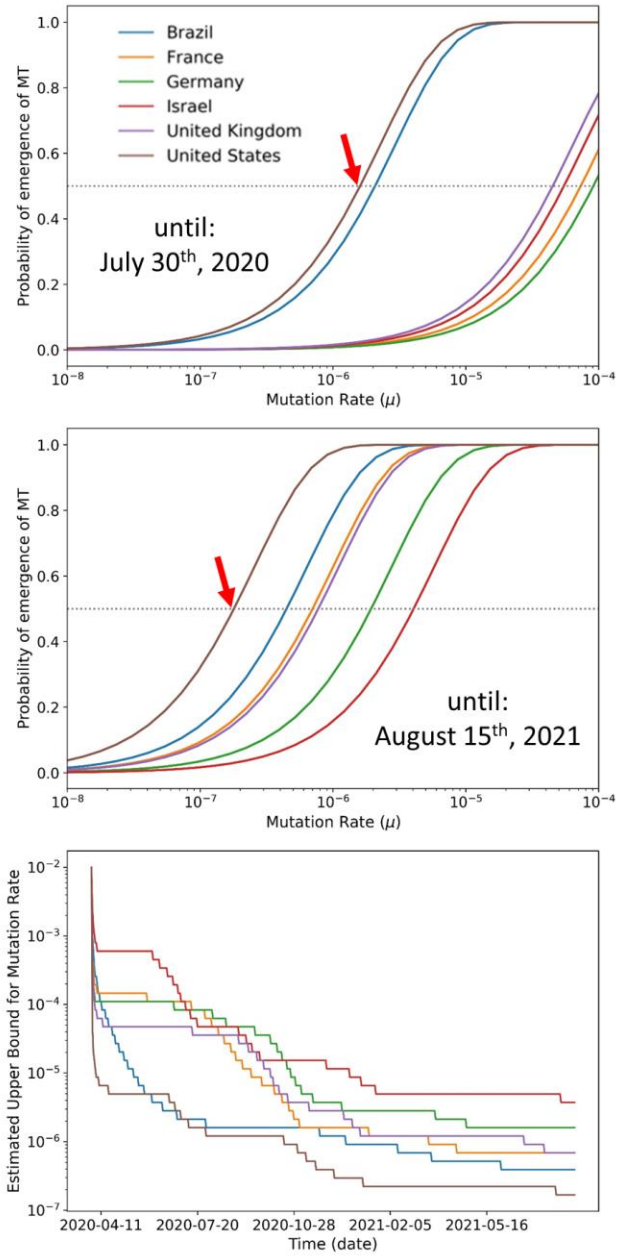

Figure 6: Estimating the mutation rate given that no vaccine resistant mutant has taken over. (A) Using Eq. 2, we calculate the probability that a MT strain would have taken over by July 30th, 2020. To this aim, we used the numbers of new infections and immunized individuals (needed to calculate the reproductive coefficient of the MT strain R_{MT} at each time point) from OWID [1]. The probability of MT strain takeover follows a sigmoidal function, where the midpoint is reached for the value of μ where MT strain takeover becomes more probable than not. We consider this value of μ the "upper bound" for the mutation rate. For the US, the estimated upper bound for the mutation rate on July 30th, 2020 would be about 10^{-6} (red arrow). **(B)** Using Eq. 2, we calculate the probability that a MT strain would have taken over by August 15th, 2021. We observe that the curves describing the probability of takeover of the MT strain along mutation rate have shifted left. This is because since July 30th, 2020, additional cases have occurred without a MT strain takeover. Therefore, the upper bound for the mutation rate decreases. For the US, the upper bound for the mutation rate would now be estimated at $2 \cdot 10^{-7}$ (red arrow). **(C)** The midpoint of the function (red arrows shown in (A) and (B)) describing the probability of MT strain takeover decreases in value as more time passes without takeover of an MT strain. We use this value as an upper bound of the mutation rate for our model.