A simple stochastic theory of extinction shows rapid elimination of a Sars-like pandemic

Bhavin S. Khatri*

Department of Life Sciences, Imperial College London,
Silwood Park, Ascot, SL5 7PY, United Kingdom
The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, United Kingdom

*To whom correspondence should be addressed; E-mail: bkhatri@imperial.ac.uk

The SARS-Cov-2 pandemic has seen the challenge of controlling novel zoonotic diseases that have high infection fatality rates, including a natural capacity for the evolution of variants that transmit more easily and evade immunity. In dealing with current and future similar pandemics, the question arises: what is the optimum strategy to control infections. Although a complex question, a key neglected component to appraise the elimination strategy is simple theory predicting the expected timescales of elimination. We use simple random walk and branching process theory to provide new insights on the process of elimination using non-pharmaceutical interventions. Our central achievement is a full theory of the distribution of extinction times — which we find is an extreme value Gumbel distribution — for any value of the reproductive number including some degree of population immunity. Overall, for the original SARS-Cov-2 variant our results predict rapid extinction — of order months — of an epidemic or pandemic if the reproductive number is kept to $R_e < 0.5$; in a counterfactual scenario with global adoption of an elimination strategy in June 2020, SARS-Cov-2 could have been eliminated world-wide by early January 2021. Looking to the future, our results would suggest that elimination using NPIs is a more optimal strategy to control — and potentially eradicate — a Sars-like pandemic, in its early stages before the evolution of variants with greater transmissibility.

Introduction

The SARS-Cov-2 pandemic has revealed a complexity of the interplay between the dynamics of infections, the evolution of the virus and the policies that governments have adopted to mitigate or eliminate the epidemic. The predictions of the SIR model (*1, 2*) would seem inadequate when compared to the history of infections seen during the SARS-Cov-2 pandemic, which would have allowed infections to rise until herd immunity brings infections under control; instead when faced with the high burden of hospitalisation and mortality, the majority of governments decided on alternative strategies of mitigation or elimination, which to different degrees use non-pharmaceutical interventions (NPIs) to control infections (*3–6*). However, in addition we have seen the evolution of new variants that are more transmissible and or gained the ability to evade immunity (*7–9*). Ultimately, for the SARS-Cov-2 pandemic, vaccination has provided a largely effective prophylaxis, making the question of mitigation vs elimination less relevant, particularly in light of SARS-Cov-2 having evolved variants of far higher transmissibility, making the latter strategy beyond what is generally considered practicable.

However, should the globe face a similar Sars-like pandemic in the future the question remains: what lessons can be learnt and what is the optimal strategy (*10*)? Looking back, the mitigation strategy with SARS-Cov-2 has encountered a number of difficulties, despite an unprecedented pace of vaccine development: natural and vaccine induced antibody immunity is short-lived, and generally much less than a year (*11–13*), vaccines do not sufficiently cut transmission (*14–20*) to achieve $R_e < 1$, vaccine roll-outs take significant time, vaccine hesitancy has been significant (*21, 22*) and finally persistent worldwide infections have led to the evolution of variants that are more transmissible and an ability to evade vaccine and natural antibody immunity (*23–25*). On the other hand, a number of countries in the Asia-Pacific region adopted an elimination strategy that aimed to eliminate local infections though NPIs (*26, 27*), which on many metrics has been seen to be successful (*28*), but was ultimately abandoned by some countries in light of the advent of vaccines that prevent serious disease and the difficulty in using NPIs to control more transmissible variants (*29*). However, despite this, the questions remains of whether elimination is a viable strategy: can infections be controlled at a small number, without requiring an "endless" lockdown, or can infections be completely eradicated.

Although, this is a complex question, using simple random walk theory and branching process analysis, we determine the full stochastic properties of the time to extinction. A stochastic approach is necessary as it accounts for the discreteness of individuals that will dominate the final stages of an epidemic, which cannot be captured by the deterministic SIR model. There has been considerable work done on understanding stochastic aspects of epidemics (*30, 31*) from the role of critical community sizes in diseases such as measles (*32, 33*), stochastic phases in the establishment of epidemics (*34*), to stochastic extinction. With regard stochastic extinction most results have been focussed on understanding the time to extinction either through the whole course of an epidemic (*35, 36*) or assuming a quasi-equilibrium has been reached through herd immunity in the population (*37*). In essence, the situation faced by many countries at the beginning stages of the SARS-Cov-2 pandemic and before vaccination, has not been considered in previous modelling; at the early stages significant herd-immunity had not been achieved in the population, and there was the potential that NPIs alone could be used to reduce the reproductive number to less than 1 and give rise to extinction of an epidemic. Our branching process analysis draws on standard results on sub-critical branching processes that show the distribution of extinction times is an extreme value

Gumbel distribution (*38, 39*); our key contribution here is to show the range of applicability of these results for effective reproductive numbers less than a critical value R_e^* and, using a heuristic approach, further extend their range to the full domain of $0 < R_e < 1$.

Using this new theory, our key practical finding is to show extinction/elimination can be achieved on the timescale of months, not years, and therefore there is a strong case that elimination should be the first defense against a new zoonosis like SARS-Cov-2 (*10, 28*), before it has evolved greater transmissibility. Specifically, for an infectious disease like SARS-Cov-2 , where infection durations are of order a week, reproductive numbers $R_e > 0.6$ give extinctions times which are long, of order many years, and dominated by the stochastic phase — on the other hand, extinction can be rapid with times much less than a year, or a few months, if restricted to $R_e < 0.5$. We show that if the globe adopted of an elimination strategy in June 2020, SARS-Cov-2 could have been eliminated world-wide by early January 2021. However, within the current context of a persistent epidemic that has evolved more transmissible variants and in the face of vaccines that only cut transmission moderately, this means that using NPIs to achieve a given R_e becomes proportionally more difficult. In light of this, we argue that elimination via NPIs is likely only optimal and feasible for early stages of a pandemic with moderate intrinsic transmissibility of the virus.

Results

SIR model with small fraction of infected individuals leads to non-changing R_e

The SIR model divides the population of N individuals in a region into 3 classes of individuals: susceptible S (not infected and not immune to virus), I infected and R recovered (and immune, so cannot be re-infected). The model assumes a rate β of an infected individual infecting a susceptible individual (S + I $\rightarrow$ 2I), and a rate γ that an infected person recovers from illness and is assumed permanently immune/recovered. The model allows a simple characterisation of when infections will grow or decline: $R_e = \frac{\beta S(t)/N}{\gamma} > 1$, where R_e is dimensionless and is called the reproductive number (also commonly called R_t) and will in general be time-dependent, as the number of susceptible individuals in a population change. R_e represents the average number of individuals an infected person infects through the duration of the infection $\tau = 1/\gamma$. The classic course of an epidemic culminates in the herd immunity threshold being reached when sufficient number of individuals have become immune and $R_e < 1$ leads to a decline of infections to eventual extinction.

The use of non-pharmaceutical interventions (NPIs) provides an alternative means to reduce R_e by reducing the contact rate β between individuals to allow $R_e < 1$, without ever requiring herd immunity; this is the basis behind elimination strategies to control infectious diseases. As we show in the Supplementary Materials, as long as R_e is less than a critical value R_e^* it is a good approximation to assume that the population of susceptible individuals $S(t) = S_0$ is roughly constant and the reproductive number unchanging $R_e = \frac{\beta S(t)/N}{\gamma} \approx \frac{\beta S_0/N}{\gamma}$; although each time an individual is infected, we loose exactly one susceptible the relative change of the susceptible pool is negligible, since the total number of susceptible individuals is very large. This critical value of the effective reproductive number is given by

$$R_e^* = -W(-e^{-R_0(1-R(0)/N)}). \tag{1}$$

where W is the Lambert W-function. This regime corresponds to an initial value of R_e such that the decline is sufficiently rapid that the error due to ignoring the change in the susceptible pool is negligible. Calculating R_e^* within the United Kingdom during summer 2020 infections were small $I_0 \approx 3 \times 10^4$ and assuming roughly 10% had recovered ($R(0) = 0.1N$), we find the constant R_e assumption to be good for $R_e < 0.98$, i.e. for all but R_e very close to 1, however, using numbers from January 2021, $I_0 \approx 10^6$, and 15% recovered this requires $R_e < 0.82$.

This means for the case where only a small fraction of the population are ever currently infected, the SIR dynamics results in a single differential equation for $I(t)$: $\frac{\mathrm{d}I}{\mathrm{d}t} = (\beta S_0/N - \gamma)I$. The solution to this is of course an exponential function: $I(t) = I_0 e^{(\beta S_0/N-\gamma)t} = I_0 e^{\rho_e t}$, where $\rho_e = \gamma(R_e - 1)$ is the effective growth rate for $R_e > 1$, and decay rate when $R_e < 1$, and $I_0 = I(0)$ the initial number of infected individuals. As we will see ρ_e more directly determines the dynamics of the extinction process than R_e or γ separately, and is in fact an easier quantity to determine (Supplementary Materials).

We are interested in understanding extinction of an epidemic and so from here on we define the rate $\rho_e = \gamma(1 - R_e)$ to be a positive quantity, making the assumption that $R_e < 1$. In this case we can make a simple deterministic prediction for the time to extinction, by calculating the time for the infected population to reach $I(t) = 1$:

$$t^\dagger = \frac{1}{\rho_e} \ln(I_0), \tag{2}$$

However, without understanding the stochasticity of the extinction process, it is difficult a priori to say anything about the goodness of this deterministic calculation. As we show below, using heuristic arguments, there is minimum number of infected individuals needed to overcome stochastic effects and confirm that this threshold also arises as a key determinant in the extinction time distribution in an exact branching process calculation.

Stochastic extinction of an epidemic

The above analysis assumes deterministic dynamics with no discreteness – it ignores any randomness in the events that lead to changes in number of infected individuals. At the beginning of the epidemic, or towards the end, there are very small numbers of individuals infected, so these random events can have a large relative effect in how the virus spreads and need a stochastic treatment to analyse. We are interested in analysing the stochasticity of how the number of infected decreases when $R_e < 1$ and eventually gives rise to extinction, i.e. when there is exactly $I = 0$ individuals and the distribution of this time.

We first run stochastic continuous time simulations with Poisson distributed events (Gillespie or kinetic Monte Carlo simulations) (*40*) of the SIR model to confirm that the assumptions of a constant R_e due to a negligibly changing susceptible population of the previous section are accurate in the regime $R_e < R_e^*$. Fig.1a plots the decline in number of infected over time $I(t)$, where each of the trajectories from the Gillespie simulations is a grey curve, whilst the deterministic prediction (Eqn.2) is shown as the solid black line. We see that for $I(t) \gg 1$ the stochastic trajectories are bisected by the deterministic prediction, indicating that the assumption of a constant R_e is a good one.

Simple random walk analysis

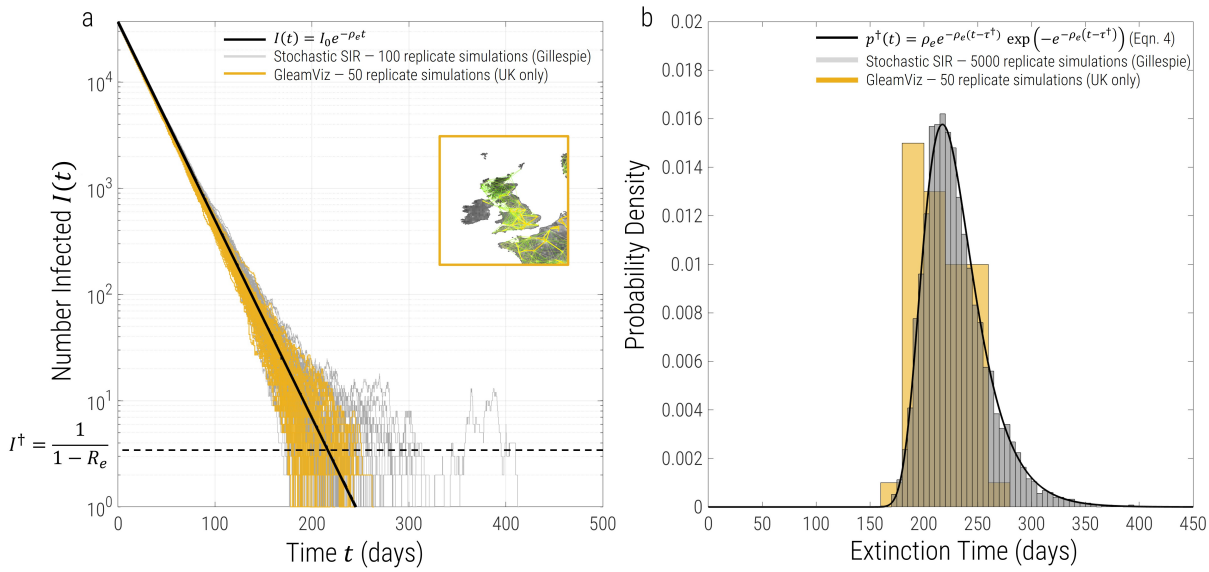

Figure 1: a) Simulation trajectories on log-linear scale for a decay rate of $\rho_e = 0.043$/day, corresponding to $R_e = 0.7$, $1/\gamma = 7$ days, $I_0 = 3.7 \times 10^4$ and an initial recovered population of $R(0) = 6 \times 10^6$. The solid black line is the deterministic prediction from Eqn.2, grey trajectories are 100 replicate Gillespie simulations of a standard SIR model, whilst the yellow trajectories are from 50 replicates using the spatial epidemic simulator GleamViz (*41, 42*) restricted to the United Kingdom with a gravity model between heterogeneous sub-populations as shown in the inset map of the UK. The dashed black line is the threshold number of infected individuals $I^\dagger$, below which changes in infected number of individuals is mostly stochastic. b)Probability density of extinction times for the same parameters as in a). Grey bars are a histogram of 5000 replicate simulations of Gillespie simulations normalised to give an estimate of the probability density, and the black curve is the prediction of the analytical calculation given in Eqn.3, which we see matches the simulations extremely well. The yellow bars are histograms from the GleamViz spatial epidemic simulator with 50 replicates, which we see gives similar results to the predictions of the stochastic SIR model.

We can see from Fig.1a that as $I(t)$ approaches extinction, as expected the trajectories become more and more varied as the number of infected becomes small. A simple heuristic treatment inspired from population genetics (*43*) defines a stochastic threshold $I^\dagger$, below which stochastic forces are more important than deterministic, as indicated by the dashed black line in Fig.1a; the time to extinction is then approximately the sum of the time it takes to go deterministically from I_0 to $I^\dagger$ ($\frac{1}{\rho_e}\ln(I_0/I^\dagger)$) and the time it takes to go from $I^\dagger$ to $I = 0$ by random chance.

Assuming such a threshold $I^\dagger$ exists, this latter stochastic time can be approximated as follows: if there are $n \leq I^\dagger$ infected individuals and changes are mainly random, then we are randomly drawing individuals from a pool of n infected individuals and $N - n$ non-infected individuals — a binomial

random walk — which when $n \ll N$ has standard deviation $\approx \sqrt{n}$ per random draw, which means we need $k = n$ random draws, such that the standard deviation over those k draws is $\sqrt{kn} \approx n$; a single random draw corresponds to one infection cycle of the virus, which is $\tau = 1/\gamma$ days, so the time to extinction starting with n individuals is approximately n/γ.

How do we estimate $I^\dagger$? It is given by the threshold size at which random stochastic changes, change the number of infected by the same amount as the deterministic decline. In one cycle or generation of infection, if there was no stochasticity, the number of infected would decline by $\approx \rho_e I^\dagger/\gamma$, so equating this to the expected standard deviation of purely random changes, $\sqrt{I^\dagger}$, we find $I^\dagger = 1/(1 - R_e)$, which is shown in Fig.1a for $R_e = 0.7$. Note that this threshold is closely related to Williams' threshold theorem (*44*), where the probability of establishment of an epidemic from a single infected individual is $p^* = 1 - 1/R_e$, in the case that $R_e > 1$, which then gives a critical number of infected $I^* \sim 1/p^* = R_e/(R_e - 1)$, below which infections changes as a random walk.

As discussed below, and in more detail in the Supplementary Materials, a more exact calculation of these considerations, using branching process theory, gives exactly the same expression for $I^\dagger$. This means the typical stochastic phase lasts $I^\dagger/\gamma = \frac{1}{\rho_e}$ days and so adding the deterministic and stochastic phases, the mean time to extinction $\approx \frac{1}{\rho_e}(1 + \ln(I_0/I^\dagger))$ (see Eqn.4 below for a more exact expression of the mean).

Exact branching process analysis

Given that in the regime $R_e < R_e^*$, the SIR dynamics can be accurately reduced down to tracking only the dynamics of I, our stochastic dynamics of infections is described by a standard birth death process, where we can use standard results from branching process theory (*45, 46*) (reproduced in Supplementary Materials for completeness) that determines the probability $p_0(t)$ that there are exactly $I = 0$ individuals as a function of time. If $p^\dagger(t)$ is the distribution of times to extinction (i.e. the probability of an extinction occuring between time t and $t + \mathrm{d}t$ is $p^\dagger(t)\mathrm{d}t$), then clearly the integral of this distribution, between time 0 and t is exactly $p_0(t)$, and hence the distribution of times to extinction is simply the derivative of $p_0(t)$ with respect to time. Doing this and also taking the limit that $I_0 \gg I^\dagger$, we find:

$$p^\dagger(t) = \frac{\mathrm{d}p_0(t)}{\mathrm{d}t} \approx \rho_e e^{-\rho_e(t-\tau^\dagger)} \exp(-e^{-\rho_e(t-\tau^\dagger)}) \tag{3}$$

where $\tau^\dagger = \frac{1}{\rho_e}\ln(I_0/I^\dagger)$, which is the time it takes for number infected to change from the initial number I_0 to the critical infection size, which this calculation shows is given by $I^\dagger = \frac{1}{1-R_e}$; pleasingly, this is the same result as arrived by the simple random walk analysis above. Fig.1b shows a histogram (grey bars) from Gillespie simulations of the SIR model with 5000 replicates of the number of infected individuals for $R_e < R_e^*$. The corresponding prediction from Eqn.3 is given by the solid black line — we see that there is an excellent correspondence. In addition, Fig.3, we see that for the range of $R_e < R_e^*$ the mean extinction time from simulation fits this prediction perfectly. This result has been calculated using other methods (*38, 39*), however, here we make clear the range of validity and the role of the stochastic threshold $I^\dagger$. The connection between the extinction time distribution and the extreme value Gumbel distributed (*47, 48*) arises since in a population that has I_0 initially infected individuals, where each recover independently with some distribution of times, the population

extinction time is the maximum extinction time of this population, which leads to extreme value statistics.

There are number of standard results for the Gumbel distribution Eqn.3, so we can write down (or directly calculate) the mean and standard deviation of the extinction time:

$$\langle t \rangle = \frac{1}{\rho_e}\left(\Upsilon + \ln\left(\frac{I_0}{I^\dagger}\right)\right) \tag{4}$$

$$\sqrt{\langle\langle t^2 \rangle\rangle} = \frac{\pi/\sqrt{6}}{\rho_e}, \tag{5}$$

where $\Upsilon \approx 0.577$ is the Euler-Mascheroni constant (conventionally assigned the symbol γ, but here γ is the recovery rate). We see that the heuristic calculation overpredicts the stochastic part of the extinction time by a factor of ≈ 2. Note that the standard deviation or dispersion of the distribution only depends on the inverse of the rate of decline ρ_e and as expected not on the initial number of infected individuals I_0; hence as ρ_e decreases (R_e gets closer to 1), we see that the distribution of extinction times broadens (as we see below in Fig.4).

The cumulative distribution function for the Gumbel distribution has the well-known double-exponential form: $P^\dagger(t) = \exp(-e^{-\rho_e(t-\tau^\dagger)})$, from which the inverse cumulative distribution function $T^\dagger = (P^\dagger)^{-1}$ is:

$$T^\dagger(p) = \tau^\dagger - \frac{1}{\rho_e}\ln\left(-\ln\left(p\right)\right), \tag{6}$$

which allows direct calculation of arbitrary confidence intervals, for example, the 95% confidence intervals, by calculating $T^\dagger(0.025)$ and $T^\dagger(0.975)$, as well as the median $T^\dagger(1/2) = \tau^\dagger - \frac{1}{\rho_e}\ln\left(\ln\left(2\right)\right)$.

Finally, it is important to stress that the distribution of extinction times Eqn.3 and the following results all assume that $I_0 \gg I^\dagger$, so that there is a clear separation of the deterministic and stochastic phases of the decline in infections. A more general and exact result for the distribution of extinction times is given in the Supplementary Materials.

Extinction time distribution with spatial structure and heterogeneity

A potentially valid criticism is that real populations have spatial structure and heterogeneity of contacts between regions. To make comparison to our simple predictions, we used a complex epidemic simulator GleamViz (v7.0) (*41, 42*), which includes a gravity model of migration, where rates of migrations between sub-populations are proportional to their population sizes (see Fig.1a inset map of UK), and each sub-population based on accurate census data within a grid of 25 km. We ran 50 replicate simulations for an SIR epidemic within the United Kingdom and with zero air travel to other countries, with the same parameters as the stochastic SIR simulations in the previous section (corresponding to 12th June 2020: $R_e = 0.7$, $\gamma = 1/7$ days, initial recovered population $R(0) = 6 \times 10^6$ – in addition, each definable sub-population in the UK was given a current infection incidence of 0.06% giving a total $I_0 \approx 3.7 \times 10^4$). We see the trajectories (Fig.1a – yellow lines) and histogram of extinction times (Fig.1b – yellow bars) compare very favourably to the predictions of the stochastic SIR model (black solid line and grey histogram bars); the mean and standard deviation including the gravity migration model is 211 ± 16 days, which is slightly smaller than the prediction of

the stochastic SIR model which has no migration or spatial structure (231 ± 30 days). This suggests that heterogeneity and migration might together have the net effect of reducing extinction times, as below we see increasing migration uniformly, has the opposite effect; nonetheless within the UK it would seem the overall effect of heterogeneity and migration is of second order to predictions of a well mixed model. Overall, we find the results of our simple model are accurate to within the width of the distributions of the extinction times.

Accounting for population immunity: modification to theory for $R_e > R_e^*$

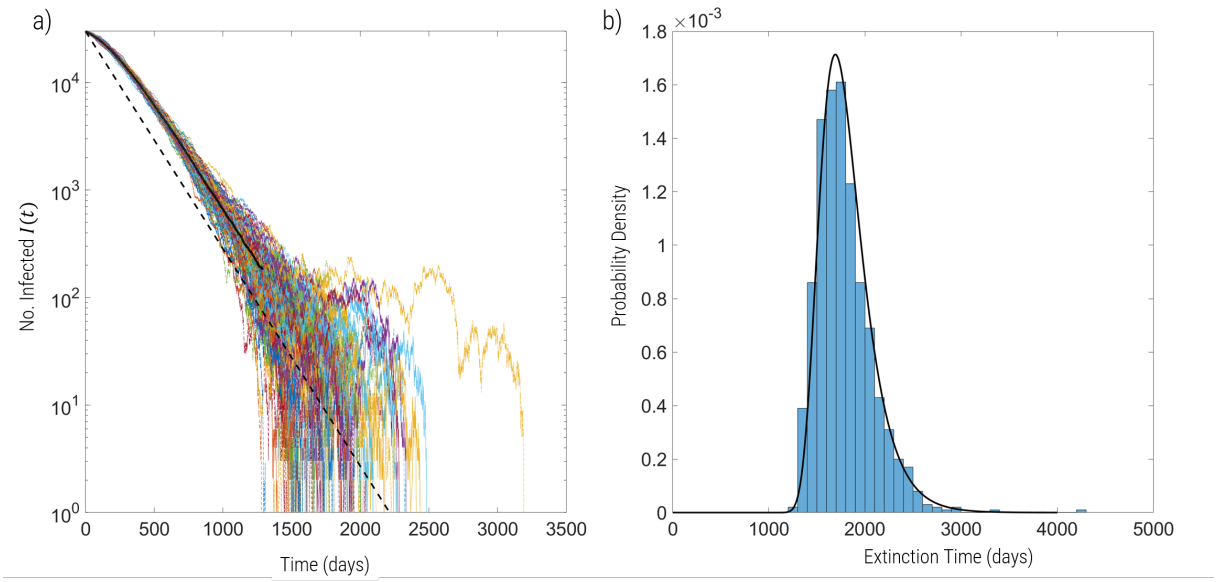

Figure 2: a) Trajectories of numbers infected $I(t)$ for simulations with $R_e = 0.99 > R_e^*$ for $I_0 = 3 \times 10^4$, $N = 67 \times 10^6$ and assuming 10% of population recovered at time $t = 0$, which gives $R_e^* \approx 0.97$; the solid line is the mean of the simulations, while the dashed line shows the decline of infections if $R_e = R_e^\infty$ from $t = 0$, where we see the slope is the same as the mean at longer times. b) Histogram of extinction time distributions of the same simulations with prediction given by Eqn3, but with $R_e \to R_e^\infty$ and $\tau^\dagger \to \tau^\dagger + 1/\rho_e^\infty$, as detailed in the text.

When $R_e > R_e^*$, we can no longer assume that changes in the number of susceptibles has a negligible effect on R_e itself. In this case the decline of infections is initially non-exponential, since R_e decreases over time, as we see in the trajectories of Fig.2a, but at later times becomes exponential again, once R_e is sufficiently small that again changes in S become relatively negligible. As we show in the Supplementary Materials, we can use a semi-heuristic approach to calculate the steady state value of number of susceptible S^∞, by integrating the SIR equations and then calculate

$$R_e^\infty = R_0 S^\infty / N = -W(-R_e e^{-R_0(1-R(0))}), \tag{7}$$

as the steady-state value of the reproductive number once infections are sufficiently small. In Fig.2a, we plot how infection would decline with $R_e = R_e^\infty$, which we see has the same slope on a log-linear plot as the asymptotic mean of the simulation trajectories at later times (solid line).

To calculate the extinction time distribution for $R_e > R_e^*$, we substitute for R_e^∞ for R_e in Eqn.3 and in addition, make the substitution $\tau^\dagger \to \tau^\dagger + 1/\rho_e^\infty$, where $\rho^\infty = \gamma(1 - R_e^\infty)$ to account for the time it takes to reach this steady state. We see in Fig.2b, that this prediction for the extinction time distribution matches the histogram of times obtained by simulation very well. In addition, we see in Fig.3 that the mean extinction time of simulations fits the predictions very well for $R_e > R_e^*$. Finally, we can simply "stitch" the solutions for $R_e \leq R_e^*$ and $R_e > R_e^*$, if needed, using a standard $\tanh$ switching function for $\tau^\dagger$, as detailed in the Supplementary Materials, and shown as the dashed line in Fig.3.

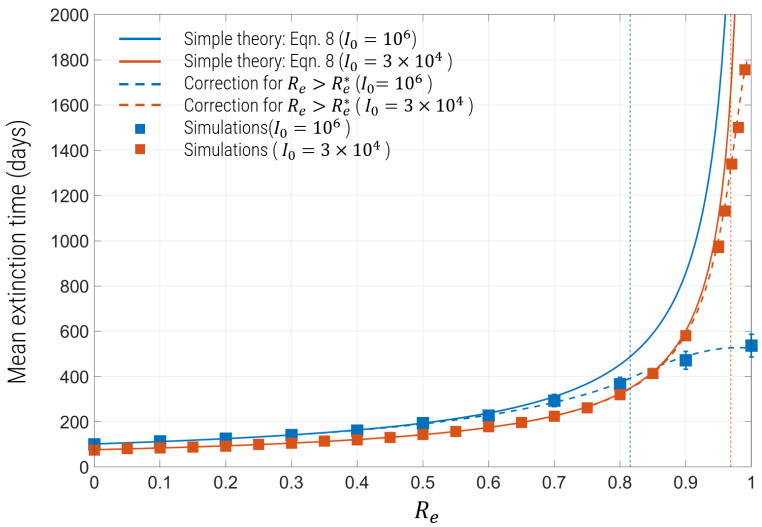

Figure 3: Mean extinction times for $I_0 = 3 \times 10^4$ (red) and $I_0 = 10^6$ (blue) for $N = 67 \times 10^6$ and assuming 10% of population has recovered. The simple theory (Eqn.4) that assumes a constant unchanging R_e are shown by solid lines, while the theory modified for values of $R_e > R_e^*$ are shown as dashed lines. The values of R_e^* are indicated by the vertical dashed lines, which for $I_0 = 3 \times 10^4$ is $R_e^* \approx 0.97$ and for $I_0 = 10^6$ is $R_e^* \approx 0.83$. The filled squares are the mean extinction times obtained using simulations for same sets of parameters.

It is interesting to note that for $R_e > R_e^*$, the extinction times are significantly lower for a higher number of initial infected because R_e^∞ is much lower; in essence the higher infection levels lead to a significant *relative* reduction in the susceptible pool causing R_e to drop more dramatically.

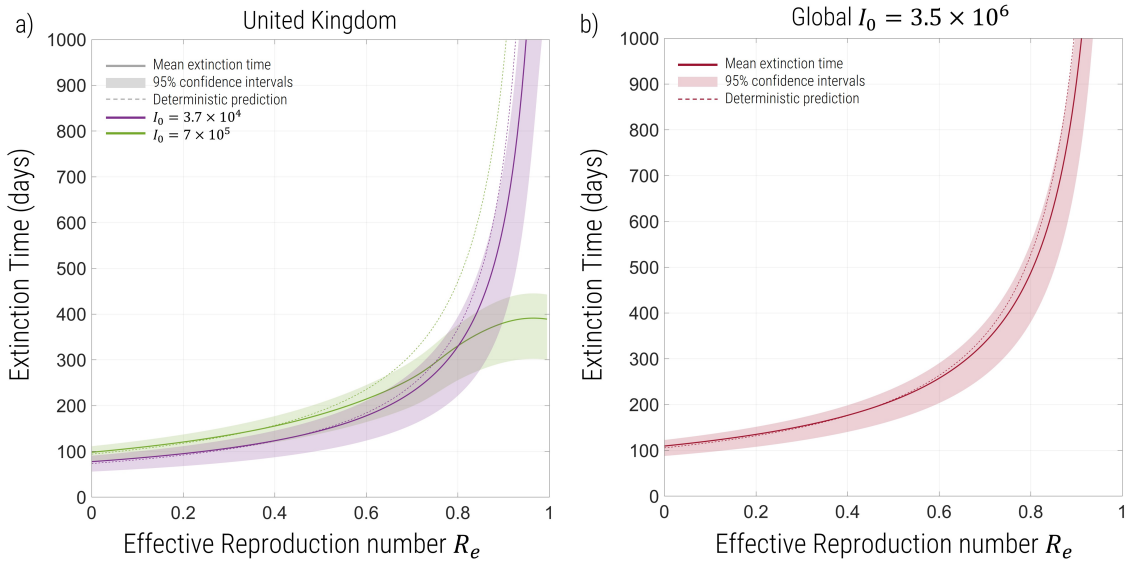

Figure 4: Prediction of the extinction times from analytical theory of the epidemic within the United Kingdom (a) and the pandemic globally (b) as a function of R_e. a) For the United Kingdom we use an initial infected population of $I_0 = 3.7 \times 10^4$ with 10% of the population recovered and immune (purple) and $I_0 = 7 \times 10^5$ with 70% of the population immune (green), both with $1/\gamma = 7$ days. b) For the global prediction, we use an initial infected of $I_0 = 3.5 \times 10^6$ and $1/\gamma = 7$ days. Solid lines are the predictions of the mean (Eqn.4), whilst each shaded region corresponds to the 95% confidence interval (Eqn.6) and the deterministic prediction (Eqn.2) are the dashed lines.

Extinction time predictions for SARS-Cov-2

United Kingdom

We first consider what this model predicts for the extinction of the SARS-Cov-2 epidemic within the United Kingdom, given an estimate of number infected of $I_0 \approx 3.7 \times 10^4$ with approximately 10% of the population immune for June 2020, when the epidemic was near it's lowest incidence, and for the number of infected $I_0 = 7 \times 10^5$ in mid-August 2021, assuming roughly 70% of the population are immune through a combination of infection and vaccination (*49*). In Fig.4a we have plotted the estimates of mean (solid line) together with 95% confidence intervals (shaded region), given an initial number of infected I_0 for various reproductive numbers R_e between $0 < R_e < 1$. In all these estimates we assume a typical duration of $1/\gamma = 7$ days for infections (*50*).

We see three broad trends: 1) for $R >\approx 0.6$ the extinction times are very long of order years, whilst at the same time the 95% confidence intervals becomes increasingly broad (of order years themselves) meaning increasing unpredictability; 2) also for $R_e >\approx 0.6$, the deterministic prediction increasingly and significantly overestimates the mean extinction time; 3) for $R_e < 0.5$ the extinction times plateau with diminishing returns for further decreases in the reproductive number. Regarding point 3), we see that there is minimum time to extinction, given by $R_e = 0$; from Eqn.4 in the

limit that $R_e \to 0$, the mean time to extinction $\langle t \rangle \to 1/\gamma(\ln(I_0) + \Upsilon)$, which shows the extinction process is ultimately limited by the rate of recovery γ, imposing a maximum speed limit on the rate of decline of infections.

In the case corresponding to summer 2020, we see that the simple stochastic SIR model predicted that extinction or elimination of the epidemic could have occurred within the United Kingdom within 4 to 5 months, which would be October/November 2020, if R_e can be kept to below about 0.5, and assuming no immigration of cases. More precisely, if $R_e = 0.4$, the mean time is 123 ± 15 days with 95% confidence intervals:$(87, 145)$ days. On the other hand for $R_e > 0.6$ we see extinction times increase very rapidly and are of order years. However, it should be noted that should $1/\gamma = 7$ days be an underestimate of the infectious period then we would expect extinction times to be approximately scaled upwards in proportion. On the other hand, taking the number of infected the following summer (mid-August 2021) where $I_0 \approx 7 \times 10^5$, we see extinction could have occurred within about 6 months if $R_e < 0.5$, or roughly by February 2022; more precisely the mean extinction time is predicted to be 157 ± 15 days $(120, 177)$ if $R_e = 0.4$.

Global extinction time predictions

We can also use our calculation to make approximate predictions of global extinction times of SARS-Cov-2 , with all the same broad caveats, given the simplifications of the model. We choose to examine the time to extinction from the period of summer 2020, as a hypothetical counterfactual scenario, where nations decided globally to adopt an elimination strategy. We assume in June 2020, current infections were $I_0 \approx 3.5 \times 10^6$, or roughly 3.5 million, based on a global death rate of $5,000$ per day. Given a global population of 7.8 billion this corresponds to a global incidence of $\approx 0.09\%$. We further assume that as in the UK at the time very roughly 10% of the population had acquired immunity naturally. In Fig.4b we plot how the predicted extinction time changes with effective reproductive numbers between $0 < R_e < 1$ for $1/\gamma = 7$ days. As would be expected, the results mirror the predictions for within the United Kingdom, except the plateau as $R_e \to 1$ occurs at much longer times; for sufficiently small values of R_e ($R_e < 0.5$), the theory predicts global extinction on a timescale of slightly greater than 200 days or 7 to 8 months, whilst $R_e > 0.6$ lead to very long extinction times ($\sim$years). More precisely, for $R_e = 0.4$ and assuming $1/\gamma = 7$ days, we find a mean extinction time of 177 ± 15 days (95% CI: $(140, 199)$ days). In this counterfactual scenario, global extinction of SARS-Cov-2 could have occurred by the beginning of January 2021.

Discussion & Conclusions

Quantifying the time to extinction for elimination strategies is of great importance for learning lessons from the SARS-Cov-2 pandemic and future pandemics with similar characteristics. In this paper, we have developed a complete theory of the extinction time distribution for well-mixed stochastic SIR model, which develops two important new theoretical insights: 1) there is a stochastic phase of declining infections when $I(t) \ll I^\dagger = 1/(1 - R_e)$, where infections change only randomly and 2) there is a critical effective reproductive number R_e^* — below which existing population immunity can be ignored, infections decline exponentially and standard sub-critical branching process can be applied — and above which decline is non-exponential due to the role of existing population immunity. In the latter case, we develop a heuristic approach to calculate a very accurate approximation of the

distribution of extinction times. Using these results we show that for $R_e = 0.4$, extinction can be rapid and of order months with specific counterfactual predictions that from mid-June 2020, extinction would have occurred within the UK with 97.5% probability by the end of October 2020 (145 days) or globally, with 97.5% probability by the beginning of January 2021 (199 days). It is worthwhile to note that these dates are before any variants of concern had been designated, but roughly coinciding with dates at which the initial variants of concern (alpha, beta, gamma, and delta) were first detected; although, this might suggest this calculation could be invalidated, or at least more complicated, in this counterfactual scenario, the probability of the evolution of such variants would be significantly reduced with fewer global infections (*54–57*) — see further discussion below.

It is clear that with global cooperation elimination of SARS-Cov-2 was feasible on a timescale of months for the original Wuhan strain – the politics, sociology and economics of achieving such global cooperation, and the many detailed factors that contribute to controlling infections (*4–6*) is outside the scope of this paper — however, with the evolution of far more transmissible strains such as the delta and omicron-variants it is arguable that elimination becomes impractical. If we use $R_e = 0.5$ as a target, then with the original Wuhan strain estimates of the reproductive number $R_0 \approx 2.5$, which means assuming a completely susceptible population, the contact rate β would have to be reduced, using NPIs, by a $\frac{1}{5}$, or crudely individuals would need to reduce their contacts to 20% of their pre-pandemic levels; note that in the first wave this was achieved in many countries (*6*). For the δ variant, we use a representative value $R_0 \approx 5$ (*51, 52*) for the reproductive number, which would suggest contacts would need to be reduced to 10% of pre-pandemic levels to achieve $R_e = 0.5$, which as was seen in New Zealand was difficult to achieve, even with some level of vaccination coverage. However, for many countries with higher infection levels of the delta-variant, immunity levels are likely be higher and so this would need to be revised accordingly. Nonetheless, without vaccines that significantly cut transmission these numbers demonstrate that elimination on the timescale of months is likely only practicable early on in a pandemic like SARS-Cov-2 .

Another important aspect of an elimination strategy is the role of the migration of infections between countries, which we ignore in the extinction time predictions, but the more complex simulations using GleamViz suggest play a relatively small role. The reason for this is likely the misconception that in order to enact an elimination strategy, all migration needs to be stopped. However, as long the rate of migration is not enormous, any country following an elimination strategy where R_e is sufficiently less than 1, there will be a stable fixed point with a small incidence of infections, due to a migration-extinction balance, and it would not be possible for a local outbreaks to occur (*53*). Hence, at a global level as long as individual countries maintain $R_e < 1$, infections will decline globally even in the face of migration. It is in this context we have estimated the global time to extinction for different values of R_e in Fig.4b and the counterfactual scenario where extinction would have occurred by January 2021, if global elimination had been pursued from June 2020.

Elimination also potentially has the advantage of preventing the evolution of new variants of concern; although effectively anecdotal, it appears none of the variants of concern have evolved in countries pursuing elimination. This is supported by the literature on population rescue and evolution of resistance (*54–56*) which clearly predicts that resistance is more likely with increasing effective population size. In particular, recent results on the evolution of multi-site resistance (*57*) suggest the evolution of variants with changes at a large number of multiple epitope sites can happen relatively rapidly in large populations with weakly deleterious standing genetic variation.

In summary, we have demonstrated a number of new results concerning the time to extinction

of an epidemic, including a complete theory for all values of the reproductive number: $0 < R_e < 1$. The key contribution of this paper is to show that elimination is a viable and relatively rapid strategy, on the time scale of $\sim$ months not years, to eliminate infections without having to rely on herd immunity, either naturally acquired, or through vaccination; this strategy could be critical in controlling infections early on in a pandemic like SARS-Cov-2 , before the development of vaccines and before the evolution of more transmissible variants, and in fact could help prevent their evolution.

References

1. W. O. Kermack, A. G. McKendrick, *Proceedings of the Royal Society,* **115A**, 700 (1927).
2. J. Murray, *Mathematical Biology I: An Introduction* (Springer-Verlag, 2002), chap. 10, pp. 319–326.
3. S. Chen, *et al.*, *International Journal of Environmental Research and Public Health* **18**, 8686 (2021).
4. N. Haug, *et al.*, *Nature Human Behaviour* **4**, 1303 (2020).
5. J. M. Brauner, *et al.*, *Science* **371**, eabd9338 (2021).
6. S. Flaxman, *et al.*, *Nature* **584**, 257 (2020).
7. K. Tao, *et al.*, *Nature Reviews. Genetics* **22**, 757 (2021).
8. M. S. Dhar, *et al.*, *Science* **374**, 995 (2021).
9. R. Viana, *et al.*, *Nature* **603**, 679 (2022).
10. D. Lewis, *Nature* **609**, 236 (2022).
11. H. Ward, *et al.*, *medRxiv* (2020).
12. A. W. D. Edridge, *et al.*, *Nature Medicine* **26**, 1691 (2020).
13. J. M. Dan, *et al.*, *Science* **371**, eabf4063 (2021).
14. M. C. Shamier, *et al.*, *medrxiv* (2021).
15. I. Kroidl, *et al.*, *Eurosurveillance* **26**, 2100673 (2021).
16. N. E. Blachere, E. Hacisuleyman, R. B. Darnell, *New England Journal of Medicine* **385**, e7 (2021).
17. E. Hacisuleyman, *et al.*, *New England Journal of Medicine* **384**, 2212 (2021).
18. A. S. Shah, *et al.*, *New England Journal of Medicine* **385**, 1718 (2021).
19. R. J. Harris, *et al.*, *New England Journal of Medicine* **385**, 759 (2021).

20. C. Stokel-Walker, *BMJ* **376**, o298 (2022).

21. S. Machingaidze, C. S. Wiysonge, *Nature Medicine* **27**, 1338 (2021).

22. M. Sallam, *Vaccines* **9**, 160 (2021).

23. N. G. Davies, *et al.*, *Science* p. eabg3055 (2021).

24. K. Leung, M. H. Shum, G. M. Leung, T. T. Lam, J. T. Wu, *Eurosurveillance* **26** (2021).

25. S. A. Kemp, *et al.*, *Nature* pp. 1–10 (2021).

26. W. R. Dowdle, *Bulletin of the World Health Organization* **76 Suppl 2**, 22 (1998).

27. S. Hendy, *et al.*, *Journal of the Royal Society of New Zealand* **51**, 1 (2021).

28. M. Oliu-Barton, *et al.*, *Lancet (London, England)* **397**, 2234 (2021).

29. M. Oliu-Barton, *et al.*, *The Lancet Global Health* **10**, e142 (2022).

30. T. Britton, *Mathematical biosciences* **225**, 24 (2010).

31. N. T. J. Bailey, *Biometrika* **50**, 235 (1963).

32. M. S. Bartlett, *Journal of the Royal Statistical Society. Series A (General)* **120**, 48 (1957).

33. M. S. Bartlett, *Applied Statistics* **13**, 2 (1964).

34. F. Ball, *Journal of Applied Probability* pp. 227–241 (1983).

35. A. D. Barbour, *Biometrika* **62**, 477 (1975).

36. P. Holme, *PloS one* **8**, e84429 (2013).

37. I. Nasell, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61**, 309 (1999).

38. A. G. Pakes, *Advances in Applied Probability* **21**, 243 (1989).

39. P. Jagers, F. C. Klebaner, S. Sagitov, *Proceedings of the National Academy of Sciences* **104**, 6107 (2007).

40. D. T. Gillespie, *Journal of computational physics* **22**, 403 (1976).

41. W. V. d. Broeck, *et al.*, *BMC infectious diseases* **11**, 37 (2011).

42. D. Balcan, *et al.*, *BMC medicine* **7**, 45 (2009).

43. M. M. Desai, D. S. Fisher, *Genetics* **176**, 1759 (2007).

44. T. Williams, *Advances in Applied Probability* **3**, 223 (1971).

45. M. Bartlett, *An Introduction To Stochastic Processes With A Special Reference To Methods And Applications* (ACM Transactions on Multimedia Computing Communications and Applications, 1960).

46. D. G. Kendall, *The Annals of Mathematical Statistics* **19**, 1 (1948).

47. R. A. Fisher, L. H. C. Tippett, *Mathematical Proceedings of the Cambridge Philosophical Society* **24**, 180 (1928).

48. E. J. Gumbel, *The Annals of Mathematical Statistics* **12**, 163 (1941).

49. Office for National Statistics, Coronavirus (COVID-19) Infection Survey pilot: England, 19 February 2021.

50. A. W. Byrne, *et al.*, *medRxiv* p. 2020.04.25.20079889 (2020).

51. F. Campbell, *et al.*, *Eurosurveillance* **26**, 2100509 (2021).

52. Y. Liu, J. Rocklöv, *Journal of Travel Medicine* **28**, taab124 (2021).

53. A. F. Siegenfeld, N. N. Taleb, Y. Bar-Yam, *Proceedings of the National Academy of Sciences* p. 202011542 (2020).

54. H. A. Orr, R. L. Unckless, *The American naturalist* **172**, 160 (2008).

55. H. A. Orr, R. L. Unckless, *PLoS genetics* **10**, e1004551 (2014).

56. J. Hermisson, P. S. Pennings, *Genetics* **169**, 2335 (2005).

57. B. S. Khatri, A. Burt, *Proceedings of the National Academy of Sciences* **119**, e2200567119 (2022).

Acknowledgements

I thank John McCauley (The Francis Crick Institute), Austin Burt, Vassiliki Koufopanou, Tin-Yu Hui (Imperial College) and Ace North (Oxford) for their insights and useful comments on the manuscript.

Competing interests

The author declares that they have no competing interests.

Data and materials availability

Code to plot extinction time predictions and distributions, as well as for performing the Gillespie simulations can be found at https://github.com/BhavKhatri/Stochastic-Extinction-Epidemic.