How long does it take to eliminate an epidemic without herd immunity?

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 global response to the SARS-Cov-2 pandemic has consisted of two main strategies both involving non-pharmaceutical interventions to control spread: mitigation, ultimately relying on herd immunity from vaccination, and elimination of infections locally. While simple theory for controlling an epidemic through herd immunity exist, there is no corresponding simple theory for the strategy of elimination with non-pharmaceutical interventions. Here we quantify an important aspect of the elimination strategy: the time to extinction without herd immunity, based solely on non-pharmaceutical interventions. Using a simple well-mixed stochastic SIR model, we find two new results: 1) using random walk theory we calculate a simple approximation of the mean extinction time and 2) using branching process theory the full distribution of times to extinction, which we show is given by the extreme value Gumbel distribution. We compare these results against complex spatially-resolved stochastic simulations to show very good quantitative agreement, demonstrating the validity of this simple approach. Overall, for SARS-Cov-2 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.

Introduction

The SIR model has remained a popular paradigm to understand the dynamics of epidemics (*1, 2*), despite its simplifications compared to real world epidemics, which have spatial structure (*3, 4*), heterogeneity (*5, 6*) in connection between regions and heterogeneity in contact rates between individuals, often giving rise to super-spreading events (*7*). The SIR model makes a simple prediction for a closed population (*2*): in the absence of any interventions and assuming individuals recovered from infection have permanent immunity, infections increase, eventually leading to the population developing herd immunity, at which point infections decline and after some time the epidemic goes extinct.

However, for infectious diseases, such as SARS-Cov-2 , we have seen such strategies, which carry a high burden of hospitalisation and mortality may not be socially and politically acceptable. This led to much of the globe adopting mitigation strategies, particularly in the West, which broadly, allowed infections to increase within the allowed capacity of health infrastructure, adopting non-pharmaceutical interventions (NPIs) when necessary, while a number of countries in the Asia-Pacific region adopted an elimination strategy that aimed to eliminate infections though NPIs (*8, 9*). Although, the latter strategy does not explicitly aim at eradication — since in the absence of concerted global cooperation, it is clear in a global pandemic an infectious disease cannot be eradicated, until it has been essentially eradicated everywhere — it does aim at a local eradication or elimination, such that infections are not routinely circulating in the population. Although, there has been some specific modelling regarding elimination, for example, in New Zealand (*9*), there is in general a basic lack of fundamental and simple theoretical results for this second scenario, which this paper addresses through calculation of the timescales of eradication/exctinction both at a national and global level.

In practice, the mitigation strategy is a stop-gap measure until population immunity can be achieved by vaccination and if the population is vaccinated to a sufficient fraction to achieve herd immunity, then the epidemic will decline. A key assumption that negates in practice the ultimate prediction of extinction of epidemics, is the lack of long-term immunity, particularly amongst the family of coronaviridae, which SARS-Cov-2 is a member, as well as the evolution of new escape variants. Currently, for SARS-Cov-2 the question of long-term immunity is not completely known, whether by natural immunity or vaccine induced; although there is evidence of waning antibody immunity on the timescale of a few months (*10*), the overall immunological response maybe more robust over the timescale of roughly a year (*11, 12*). Importantly, although vaccines may have high efficacy for reducing serious disease, the picture regarding a significant reduction in transmission is still not clear (*13–16*). In addition, as the recent emergence of new variants (*17, 18*) have shown, there is the possibility that vaccine escape mutants could evolve (*19*), reducing the efficacy or rendering redundant vaccines based on previously circulating antigen sequences. Although, much is still unknown, in a worst case scenario where immunity is short-lived and infections are endemic (*20*), there will be a continuing risk of vaccine escape, and so alternative strategies may be required; ultimately, even for those countries pursuing mitigation, elimination may be the only option available, other than naturally acquired immunity.

This paper examines the SIR model, but fully accounting for the discreteness of individuals that leads to stochasticity in the progress of an epidemic. This is critical to examine the question of extinction, since the continuous (deterministic) SIR model is unrealistic when only a few individuals are infected and gives the erroneous prediction that extinction only arises asymptotically at very long

times. There has been considerable work done on understanding stochastic aspects of epidemics (*21, 22*) from the role of critical community sizes in diseases such as measles (*23, 24*), stochastic phases in the establishment of epidemics (*25*), 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 (*26, 27*) or assuming a quasi-equilibrium has been reached through herd immunity in the population (*28*). However, the situation faced by many countries at the beginning stages of the SARS-Cov-2 pandemic and before vaccination, has not been strongly considered in previous modelling, that is where significant herd-immunity has not been achieved in the population, and there is the potential that NPIs alone can be used to reduce the reproductive number to less than 1, the critical threshold for growth and give rise to extinction of an epidemic.

It is with this scenario in mind, where a population still has many more susceptible (& recovered), compared to infected, where our main simple result is focussed, and we find the stochastic dynamics tractable within a simple birth-death branching process framework. Our main theoretical result is the distribution of the times to extinction of an epidemic, which surprisingly, we find is a Gumbel-type extreme value distribution. Although the extinction time distribution has been previously studied (*29*), a closed form solution for a Poisson offspring distribution was not obtained. Key to this result is a new threshold $I^\dagger = 1/(1-R_e)$, below which stochastic changes dominate and which we show arises from simple random walk theory. However, we then also extend the calculation using heuristic considerations that cover the whole range of $0 < R_e \leq 1$, accounting for the dynamics of R_e, when $R_e \lesssim 1$, where by necessity population immunity must play a role in the dynamics of the epidemic. As this result ignores spatial structure and heterogeneity of an epidemic, we then compare to simple and more complex spatial epidemic simulations, and find our theory captures the extinction time distribution very well, as long as R_e is appropriately rescaled to account for migration. We then use this theory to make broad predictions of extinction times within the UK, and globally to serve as a guide to more complex and detailed models. Our key message is that 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 and of order many years — on the other hand, extinction can be rapid with times much less than a year, or a few months, if restricted to $R_e < 0.5$.

Susceptible–Infected–Recovered (SIR) model of epidemiology

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). If we assume a rate β of an infected individual infecting a susceptible individual ($\mathrm{S} + \mathrm{I} \to 2\mathrm{I}$), and a rate γ that an infected person recovers from illness ($\mathrm{S} \to \mathrm{R}$), the ordinary differential equations describing the dynamics of this process are:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta I(S/N) \quad (1)$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta I(S/N) - \gamma I \quad (2)$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I. \quad (3)$$

A key aspect of this model is that it allows a simple characterisation of when number of infections will grow or decline: whatever the previous history of the epidemic, for growth we need $\frac{\mathrm{d}I}{\mathrm{d}t} > 0$ and this happens for the following condition on RHS of the 2nd equation above: $\beta S(t)/N - \gamma > 0$, or equivalently, $R_e = \frac{\beta S(t)/N}{\gamma} > 1$, where we have defined the dimensionless number R_e (also commonly called R_t) as the combination shown, 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 $\tau = 1/\gamma$. It is important to understand that this interpretation of R_e is within the context of a well-mixed model. In reality, locally there may be deviations from the global density of susceptible individuals ($S(t)/N$) and also differences in connectivity between different regions causing differing rates of infection locally.

It is worthwhile to briefly revisit how extinction arises through herd immunity in the standard SIR model, in contrast to the main mechanism we discuss in this paper, where there is extinction without herd-immunity and only using NPIs. Initially, it is assumed the whole population is susceptible, as the number of infected is 1 or very small, so $S(t=0) \approx N$ and the reproductive number in this case is $R_0 = \beta/\gamma$. The above SIR equations lead to a growing number of infected $I(t)$, and a decreasing susceptible pool $S(t)$, which leads to a decreasing R_e. The epidemic continues to grow until $R_e = 1$ (i.e. when the fraction of susceptibles has become sufficiently small), at which point $\mathrm{d}I/\mathrm{d}t = 0$ — this defines the classic herd-immunity threshold of the number of immune/recovered $1-1/R_e$. The classic herd immunity threshold only defines the point when the effective reproductive number is exactly 1, and in fact the number of susceptibles continues to decline beyond this point, causing $R_e < 1$, until it reaches a plateau S^∞ (*2*). The plateau corresponds to when infections have become sufficiently small that additional infections cause only a negligible change to the susceptible pool; once this plateau is reached we have a constant "ultimate" reproductive number $R_e^\infty = -W(-R_e e^{-R_0}) < 1$, where $W(z)$ is Lambert's W-function, which is defined by the solution to the transcendental equation $we^w = z$. In this limit, as discussed below, the number of infected then declines exponentially until extinction.

Assumption of constant R_e with small fraction of infected individuals in SIR model

As just discussed, in an idealised SIR epidemic, if β does not change due to behaviourial changes, the reproductive number is in general a constantly decreasing number, due to the susceptible pool of individuals diminishing — this eventually leads to herd immunity as the decreasing susceptible fraction brings $R_e < 1$. However, changes in social behaviour (non-pharmaceutical interventions or NPIs) can also bring $R_e < 1$ by controlling β, before any significant herd immunity is established, which was generally still the case in many countries with SARS-Cov-2 (before February 2021), when vaccination coverage was low. In this case, if we assume that the fraction of the population that are currently infected is small, compared to the number of susceptibles, we can assume that the effects of herd immunity are negligible; analogously to the description with herd immunity, this is manifested by an approximately unchanging susceptible pool and constant effective reproductive number R_e.

Using the United Kingdom as an example, just after the first lockdown, the UK Office for National Statistics (*30*) estimates from serological testing in England that 6.8% of the population had been infected to June 12th 2020 ($\approx 4.6 \times 10^6$ extrapolated to the UK population) and from random PCR

testing a current incidence of 0.055% ($\approx 3.7 \times 10^4$ were actively infected in UK) – in which case, extrapolating to a UK population size of $\approx 67 \times 10^6$, the total number of susceptibles ($\approx 62 \times 10^6$) is much greater than the number currently infected. More recently, with the emergence of more transmissible variants and varied application of NPIs, infections peaked at of order a million, and roughly 18.5% of the population estimated to have antibodies in February 2021 (*31*); this leads to a susceptible pool of approximately 55 million, which is still much greater than the number currently infected.

In this case, as long as R_e isn't very close to 1, it is reasonable 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. As we show in Supplementary Materials, this approximation is good as long as $R_e < R_e^*$, where $R_e^* = -W(-e^{-R_0(1-R(0)/N)})$, which 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, last summer 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 $R_e < 0.82$. Assuming that $I_0 \ll S_0$ and $R_e < R_e^*$, this model of constant R_e should also be very reasonable, when there is no immunity (SIS model), or in the presence of waning immunity, where immunity only lasts a finite time (SIRS), since increasing the susceptible pool again should have negligible effect, since we assume the total number of susceptibles is very large in comparison.

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 $I(t)$:

$$\frac{\mathrm{d}I}{\mathrm{d}t} = (\beta S_0/N - \gamma)I. \tag{4}$$

The last differential equation involving R can be ignored as it is really only there for book-keeping, as there is no direct effect of R on the dynamics of S and 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. Note that R_e is not a rate, it does not in absolute terms tell you anything about the time scales of change; however, ρ_e is a rate, and if it could be measured empirically, it would give information in the speed of spread of the infection, as well as having the same sign information for the direction of change ($\rho_e > 0$ the epidemic spreads, while $\rho_e < 0$ means the epidemic cannot spread). 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 $I(t) = 1$:

$$t^{\dagger} = \frac{1}{\rho_e} \ln (I_0), \tag{5}$$

Of course, we want to know the time to complete elimination $I(t) = 0$, but we cannot answer this question with a deterministic continuous approximation, since the answer would be ∞; the time it takes to go from 1 infected individual to 0 cannot be handled in a deterministic approach, since it ignores the discreteness of individuals and the stochasticity that lies therein. In fact, without understanding the stochasticity of the extinction process, it is difficult a priori to say anything about the goodness of this deterministic calculation, since in general we would expect stochasticity to be important far before there remains only a single infected individual. We answer below, using heuristic arguments, the 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; an infected person might typically take the tube to work, potentially infecting many people, whilst on another day decide to walk or take the car, reducing the chances of infecting others. When the epidemic is in full flow with large numbers of individuals infected, all the randomness of individual actions, effectively average out to give smooth almost deterministic behaviour. However, 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; in particular, we are primarily interested in calculating the distribution of the times to extinction.

We can initially confirm that the assumptions of a constant R_e due to a negligibly changing susceptible population of the previous section are accurate, by running multiple replicate stochastic continuous time simulations with Poisson distributed events (Gillespie or kinetic Monte Carlo simulations) (*32*) of the SIR model with $R_e = 0.7$, $\gamma = 1/7$ days^{-1}, $I_0 = 3.7 \times 10^4$ and an initial recovered population of $R(0) = 6 \times 10^6$, which for simplicity we take as 10% of population infected and recovered. Fig.1 plots the decline in number of infected over time $I(t)$. Each of the trajectories from the Gillespie simulations is a grey curve, whilst the deterministic prediction (Eqn.5) 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

We can see from Fig.1 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 (*35*) would define a stochastic threshold $I^{\dagger}$, below which stochastic forces are more important than deterministic, as indicated by the dashed black line in Fig.1; the time to

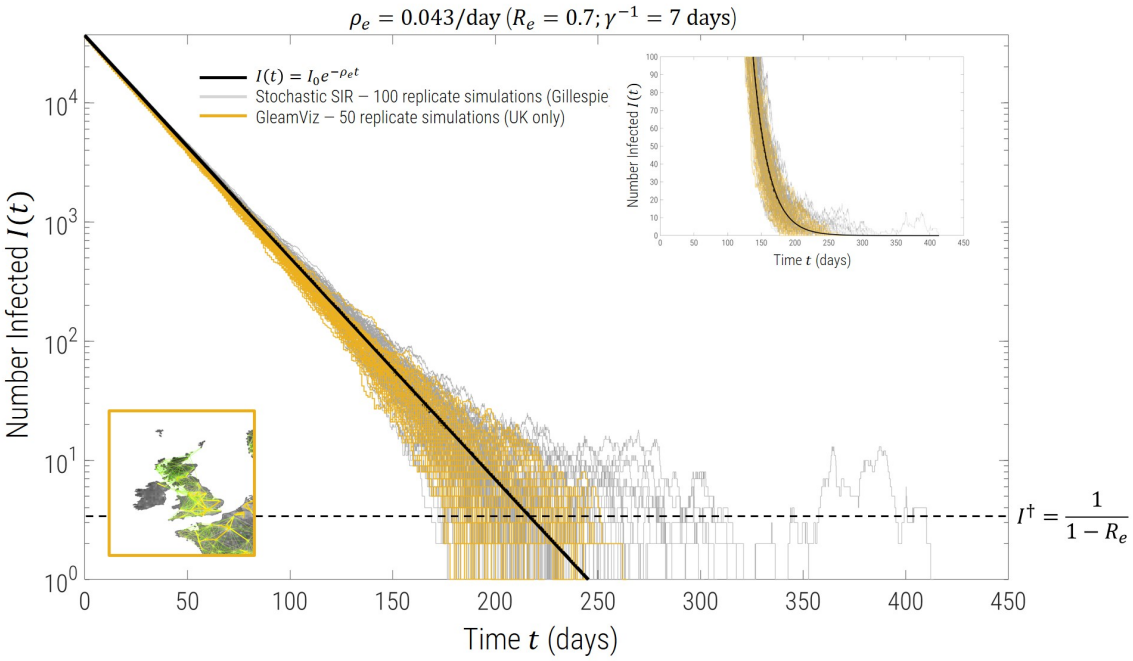

Figure 1: Simulation trajectories on log-linear scale (inset: linear-linear scale) for a decay rate of $\rho_e = 0.043$/day, corresponding to $R_e = 0.7$, $1/\gamma = 7$ days, $I_0 = 3.7 \times 10^4$ and an initial recovered population of $R(0) = 6 \times 10^6$. The solid black line is the deterministic prediction from Eqn.5, 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 (*33, 34*) 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.

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.1 for $R_e = 0.7$. Note that this threshold is closely related to Williams' threshold theorem (*36*), 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.8 below for a more exact expression of the mean).

Exact branching process analysis

The branching process framework used to calculate the distribution of extinction times is standard, but detailed, and so we will sketch the derivation here and leave details for the Supplementary Materials. The first step is to recognise that there are two independent stochastic events that give rise to the net change in the numbers of infected individuals, as depicted in Eqn. 4 for the continuum deterministic limit: 1) a susceptible individual is infected by an interaction with a infected individual, such that $I \to I+1$ and 2) an infected individual recovers spontaneously such that $I \to I-1$. This is a simple birth death branching process for which it possible to write down differential equations $(\mathrm{d}p_I(t)/\mathrm{d}t)$ for how the probability of I infected individuals changes with time in terms of the birth and death events just defined. It is possible to find after some calculation the probability generating function $G(z,t)$ of the birth-death process, from which the probability of having exactly $I = 0$ individuals as a function of time is given by:

$$p_0(t) = G(z=0,t) = \left(\frac{1 - e^{-\rho_e t}}{1 - R_e e^{-\rho_e t}}\right)^{I_0}. \tag{6}$$

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 Eqn.6, 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:

$$p^\dagger(t) = \frac{\mathrm{d}p_0(t)}{\mathrm{d}t} \approx \rho_e e^{-\rho_e(t-\tau^\dagger)} \exp(-e^{-\rho_e(t-\tau^\dagger)}) \tag{7}$$

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 $I^\dagger = \frac{1}{1-R_0}$; pleasingly, this is the same result as arrived by the simple heuristic analysis above. Fig.2 shows a histogram (grey bars) from Gillespie simulations of the SIR model with 5000 replicates of the number of infected individuals for $R_e = 0.7$, $\gamma = 1/7$ days, with initial number infected $I_0 = 3.7 \times 10^4$ and an initial recovered population of $R(0) = 6 \times 10^6$, corresponding to the situation in the UK in 12^{th} June 2020. The corresponding prediction from Eqn.7 is given by the solid black line — we see that there is an excellent correspondence. In addition, Fig.5, we see that for the range of $R_e < R_e^*$ the mean extinction time from simulation fits this prediction perfectly. Surprisingly, the extinction time

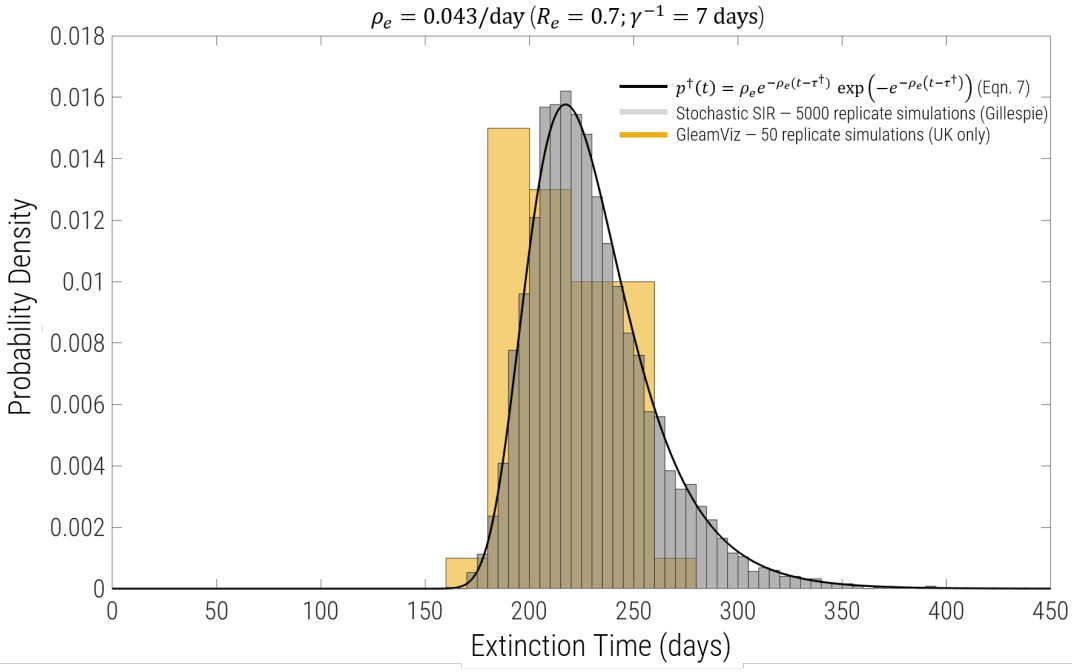

Figure 2: Probability density of extinction times for the same parameters as in Fig.1. Grey bars are a histogram of 5000 replicate simulations of Gillespie simulations normalised to give an estimate of the probability density, and the black curve is the prediction of the analytical calculation given in Eqn.7, 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.

distribution is a Gumbel-type extreme value distribution (*37, 38*); it is surprising as an extreme value distribution normally arises from the distribution of the maximum (or minimum) of some quantity, although here it is not clear how this relates to the extinction time.

There are number of standard results for the Gumbel distribution Eqn.7, so we can write down (or directly calculate) the mean and standard deviation of the extinction time:

$$\langle t \rangle = \frac{1}{\rho_e}\left(\Upsilon + \ln\left(\frac{I_0}{I^\dagger}\right)\right) \tag{8}$$

$$\sqrt{\langle\langle t^2\rangle\rangle} = \frac{\pi/\sqrt{6}}{\rho_e}, \tag{9}$$

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.6).

We can also calculate the cumulative distribution function

$$P^\dagger(t) = \int_0^t p^\dagger(t')\mathrm{d}t' = \exp(-e^{-\rho_e(t-\tau^\dagger)}), \tag{10}$$

from which the inverse cumulative distribution function $T^\dagger = (P^\dagger)^{-1}$ can be calculated:

$$T^\dagger(p) = \tau^\dagger - \frac{1}{\rho_e}\ln\left(-\ln\left(p\right)\right), \tag{11}$$

which enables direct generation of random numbers drawn from the extinction time distribution, by drawing uniform random u on the unit interval and calculating $T^\dagger(u)$. It also allows 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.7 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

National level (United Kingdom)

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) (*33, 34*), which includes a gravity model of migration, where rates of migrations between sub-populations are proportional to their population sizes (see Fig.1 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 12^{th} 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.1 – yellow lines) and histogram of extinction times (Fig.2 – 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, at a national level, we find the results of our simple model are accurate to within the width of the distributions of the extinction times.

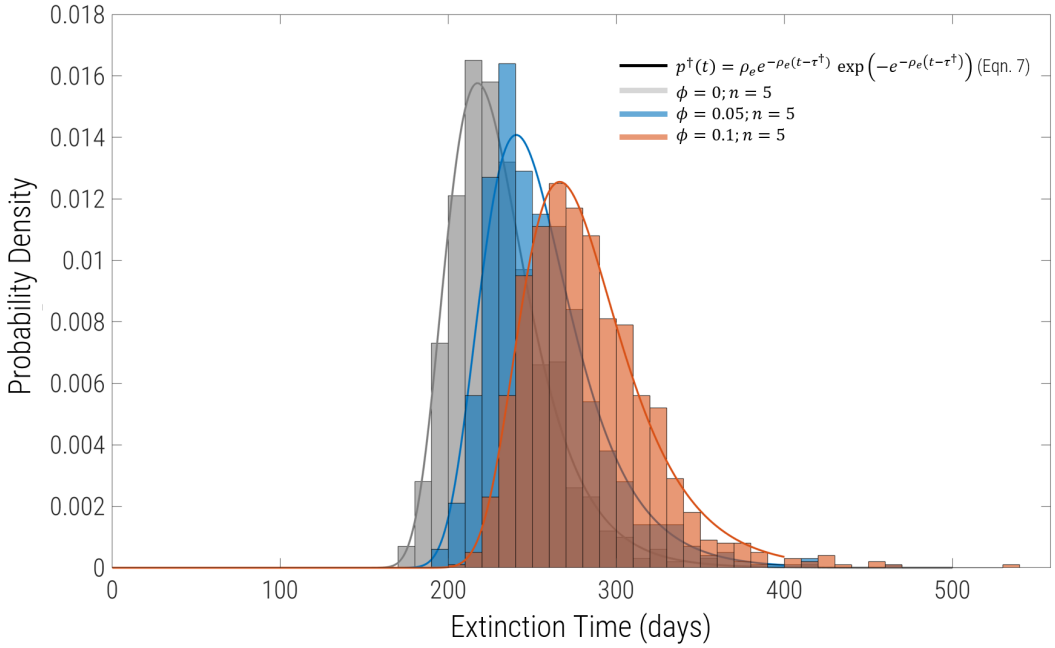

Figure 3: Probability density of extinction times for the same parameters as in Fig.1, but including migration and sub-division into equal sized populations. Each histogram comprises 1000 replicates for $n = 5$ regions connected by uniform migration with probability ϕ. Grey bars are $\phi = 0$ (complete isolation), blue correspond to $\phi = 0.05$ and $\phi = 0.1$ are the red bars. For $\phi = 0$ the solid line grey line is exactly the solid black line in Fig.2, showing that the extinction time distribution of identical to the single global well-mixed population of same aggregate size. The solid blue and red lines are fits to the histogram using Eqn.7 with a single free parameter R_e (with γ and I_0 constrained to the values used to run the simulations).

Global

It was not possible to repeat these simulations on a global scale as GleamViz does not record individual level changes in infections and deaths in its global output. Here instead we first consider the total extinction time distribution for a number of isolated regions (countries) with no migration between, but each with the same R_e. As we show in the Supplementary Materials, in fact, the extinction time distribution of the whole region (i.e. the distribution of the maximum time of all the groups) is exactly the same distribution as assuming a single unstructured/undivided population for the region. We verify this by Gillespie simulation of a simple birth-death model with growth rate γR_e and death rate γ for n isolated populations; the grey histogram in Fig.3 is the estimate of the extinction time distribution for isolated sub-populations and this matches the grey solid line, which is exactly the solid black line in Fig.2.

We now look at the effect of migration, where we examine the same Gillespie simulations of birth

and death, but with a probability of global migration per individual of ϕ. As we increase ϕ we see that the extinction time distribution shifts to longer times, yet still maintains the same form as given by Eqn.7 – fitting to this equation using only R_e as a free parameter, we find for $\phi = \{0.01, 0.05, 0.1\}$, $\hat{R}_e = \{0.709 \pm 0.001, 0.732 \pm 0.001, 0.760 \pm 0.001\}$, respectively (minimum R-sqd statistic of 0.975). These fits are shown as the blue and red solid lines in Fig.3 for $\phi = 0.05$ and $\phi = 0.1$, respectively, and we see that the fits follow the data very closely (the histogram and fits for $\phi = 0.01$ are not shown in Fig.3 for clarity, as they overlap closely with $\phi = 0$). We see that we can predict these estimated reproductive numbers $\hat{R}_e$ by simply rescaling the base R_e to $R_e \to (1+\phi)R_e = \{0.707, 0.735, 0.77\}$ for $\phi = \{0.01, 0.05, 0.1\}$, respectively. This finding is closely related to the literature on the group level reproductive number R_* (*3–5*), except here we are studying the decline and extinction of an epidemic/pandemic as opposed to its establishment, which has not been previously studied in this context. Overall, these results suggest that under the assumption that each national region has the same R_e, that the extinction time distribution is given by the stochastic SIR model (Eqn.7) but with a rescaled R_e to account for air traffic or migration between regions/countries.

Modification to theory for $R_e > R_e^*$

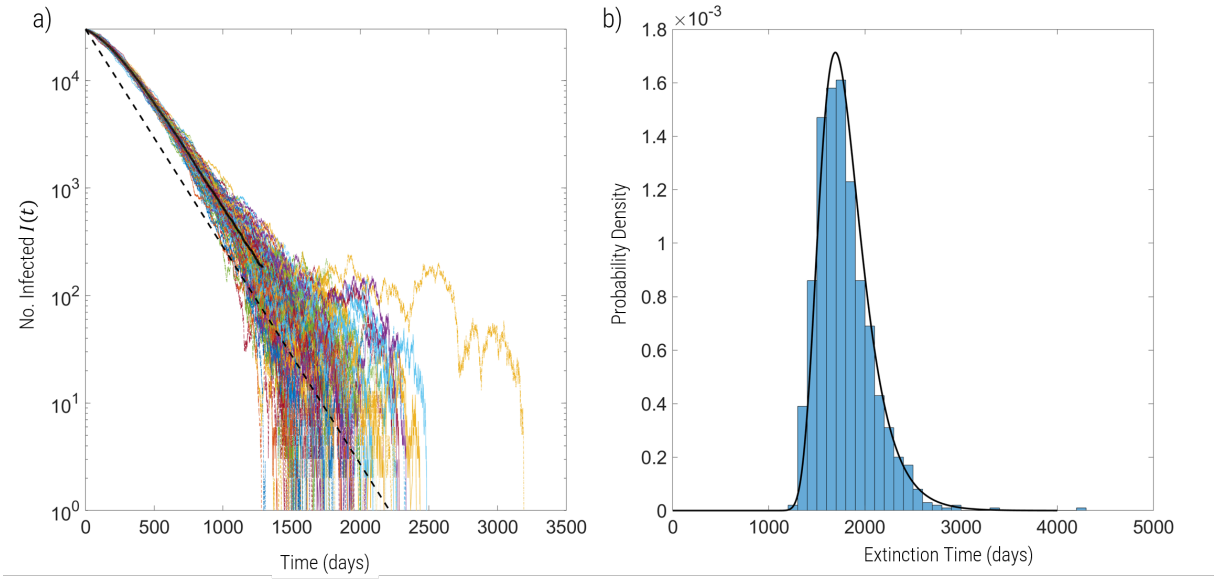

Figure 4: a) Trajectories of numbers infected $I(t)$ for simulations with $R_e = 0.99 > R_e^*$ for $I_0 = 3 \times 10^4$, $N = 67 \times 10^6$ and assuming 10% of population recovered at time $t = 0$, which gives $R_e^* \approx 0.97$; the solid line is the mean of the simulations, while the dashed line 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 Eqn7, 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 is close to 1, 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.4a, but at later times becomes exponential again, once R_e is sufficiently small that again changes in S become relatively negligible. As we detail in the Supplementary Materials, there is a critical value of R_e above which the constant R_e assumption is no longer an accurate approximation, which is given by

$$R_e^* = -W(-e^{-R_0(1-R(0)/N)}). \tag{12}$$

For $R_e > R_e^*$, we take a semi-heuristic approach and calculate the steady state value of number of susceptible S^∞, by integrating the SIR equations and then calculate

$$R_e^\infty = R_0 S^\infty/N = -W(-R_e e^{-R_0(1-R(0))}), \tag{13}$$

as the steady-state value of the reproductive number once infections have become sufficiently small. In Fig.4a, 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.7 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.4b, that this prediction for the extinction time distribution matches the histogram of times obtained by simulation very well. In addition, we see in Fig.5 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.5.

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.

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 current number of infected $I_0 = 7 \times 10^5$, assuming roughly 70% of the population are immune through a combination of infection and vaccination (*39*). In Fig.6a 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 (*40*).

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

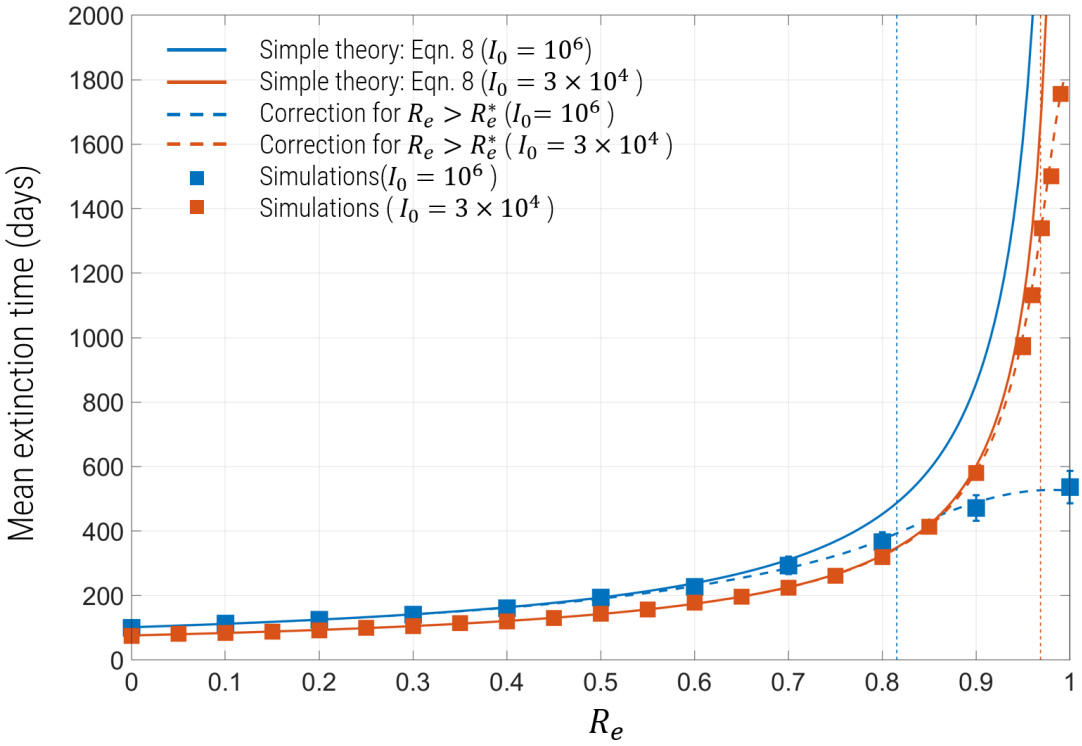

Figure 5: Mean extinction times for $I_0 = 3 \times 10^4$ (red) and $I_0 = 10^6$ (blue) for $N = 67 \times 10^6$ and assuming 10% of population has recovered. The simple theory (Eqn.8) that assumes a constant unchanging R_e are shown by solid lines, while the theory modified for values of $R_e > R_e^*$ 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.

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.8 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 more recent situation (mid-August

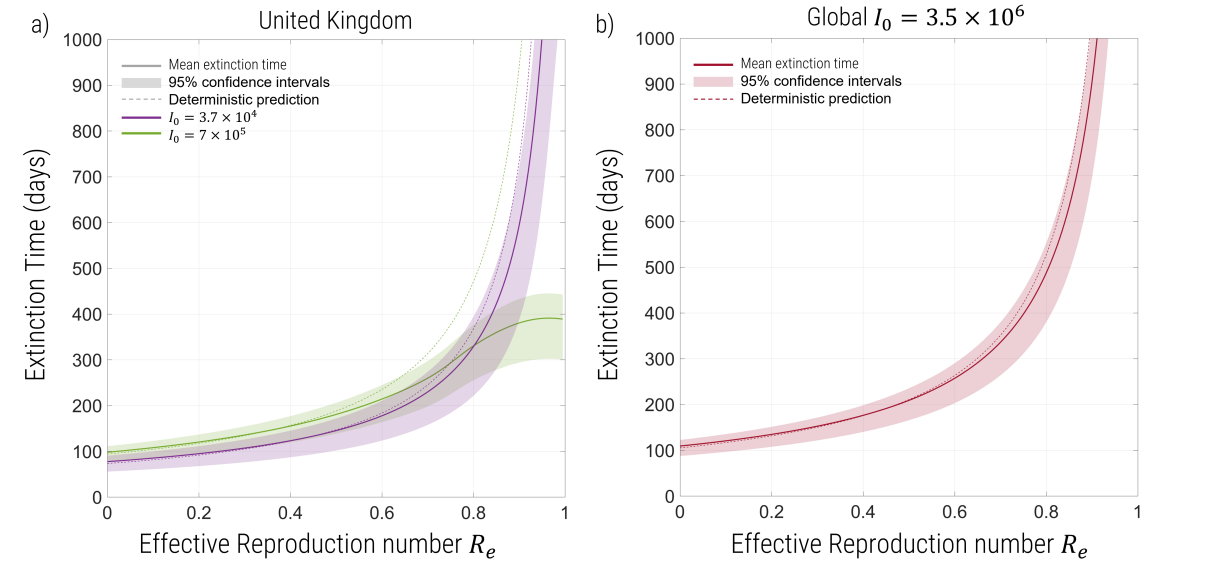

Figure 6: Prediction of the extinction times from analytical theory of the epidemic within the United Kingdom (a) and the pandemic globally (b) as a function of R_e. a) For the United Kingdom we use an initial infected population of $I_0 = 3.7 \times 10^4$ with 10% of the population recovered and immune (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.8), whilst each shaded region corresponds to the 95% confidence interval (Eqn.11) and the deterministic prediction (Eqn.5) are the dashed lines.

2021) where $I_0 \approx 7 \times 10^5$, we see extinction can occur 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. However, as argued above at a global level, we can have confidence that the predictions of the extinction time distribution Eqn.7 can be quantitatively correct, when the R_e value is effectively rescaled to account for migration/air-traffic between nations. 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. In June 2020, the global death rate was very approximately $5,000$ deaths per day, and so with an approximate infection fatality rate $\eta \approx 0.01$, (*41, 42*) the rate of change of deaths would be roughly $\mathrm{d}(deaths)/\mathrm{d}t = \eta\gamma I(t)$; inverting this gives a crude estimate of the current number of actively infected globally as $I_0 \approx 3.5 \times 10^6$, or roughly 3.5 million. More precise estimates would require understanding the age structure of the infection fatality rate, which is highly biased to older populations, as well as accounting for more recent improvements in treatments

of severe cases, but as extinction times are only logarithmically dependent on initial size (Eqn.8), these improvements would not significantly change these estimates. 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.6b 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

We have presented a new analysis of extinction in the stochastic SIR model in the context of populations with very little herd immunity and yet a significant number of infected individuals. Using simple random walk theory, we calculate the mean time to extinction and show that there is a critical threshold $I^\dagger = 1/(1 - R_e)$, below which random stochastic changes are more important, which suggests that for $R_e < 1$, but approaching 1, a simple deterministic prediction will be poor. With a more exact branching process analysis, we then calculate in closed form the extinction time distribution of an epidemic, which surprisingly, is an extreme value distribution of the Gumbel type, and is mainly dependent on the rate of decline of the epidemic $\rho_e = \gamma(1-R_e)$, with a weak logarithmic dependence on R_e explicitly. A key advantage of a closed form solution for the distribution is that we can discern broad trends very quickly without doing a large number of complex simulations over a large parameter space. Given the simplicity of SIR well-mixed model, we compared our results to a complex spatial epidemic simulator, GleamViz (*33, 34*) with explicit heterogeneity in connections between sub-populations, as well as simulations of uniform migration between sub-populations and find good overall agreement for the distribution of extinction times with our simple predictions. We have also extended the theory of extinction to the whole range of $0 < R_e < 1$, where for R_e close to unity, herd or population immunity must play a role in the dynamics of the epidemic.

We use these results to produce predictions of extinction times within the UK and globally of the SARS-Cov-2 epidemic/pandemic under various scenarios, as a function of R_e as a guide to the expected trends in more complex models. The results suggest that if the reproductive number was constrained to $R_e = 0.4$ (assuming $1/\gamma = 7$ days), within the UK from mid-June 2020, extinction would have occurred with 97.5% probability by the end of October 2020 (145 days). Alternatively, if the UK were to enact the same measures now with the help of immunity afforded by vaccination, the theory predicts that extinction would occur with 97.5% probability by the middle of February 2022 (177 days); note that in this scenario to achieve $R_e = 0.4$ with NPIs is on one hand much easier with vaccination, but also more difficult due to the higher transmissibility of the δ-variant. Globally, under a counterfactual scenario where all nations decided to adopt an elimination strategy in summer 2020 with $R_e = 0.4$, the theory predicts extinction would have occurred with 97.5% probability by the beginning of January 2021 (199 days).

However, unless R_e is strongly controlled, extinction times increase to times of order years for $R_e > 0.6$, as shown in Figs.6a&b and become increasingly unpredictable, which is an indication that the stochastic, random walk phase of the extinction process is dominant. On the other hand the same figures show that decreasing R_e much below $R_e = 0.4$ produces diminishing gains in reductions of extinction times; hence, given social consequences of lockdown measures this suggests an optimal $R_e \approx 0.4 \to 0.5$, which in the United Kingdom was achieved just after lockdown in regions such as London (*43*).

The basic calculation in this paper ignores spatial structure (*3, 4*) and heterogeneity of contacts between individuals (*5*). To assess the realism of our simple model, we performed simulations using a realistic spatial epidemic simulator, GleamViz, (*33, 34*) which gave a distribution of extinction times matching reasonably closely the Gumbel distribution in Eqn.7, with a slightly smaller mean time. On one hand we might expect spatial structure to give local deviations of the fraction of susceptibles from the global average, allowing herd immunity to arise locally in regions overall speeding up the decline of the epidemic, but on the other hand, migration between regions (as we see with the global migration simulations) act to slow down the decline. These results might suggest the well mixed model does reasonably well because of a cancellation of these two competing effects.

The simple SIR model also ignores age structure, as well as heterogeneity or individual variation in contact/spreading rates, which could give rise to super-spreading events. Although age structure has been shown to have a significant quantitative effect on the critical fraction of the population required to reach herd immunity (*44, 45*), here we expect for $R_e < R_e^*$ (Eqn.12) as the calculation assumes negligible herd immunity, we do not expect that different age transmission matrices will affect our results for the same mean rate of transmission β. On the other hand the role of superspreaders and a very leptokurtic off-spring distribution could have a significant quantitative effect, although we expect the qualitative behaviour to remain unchanged; analysis of contact patterns for the spreading of SARS-Cov-2 estimate a dispersion parameter k in the range of 0.1 to 0.6 (*46–48*), which suggests a small number of individuals cause a majority of infections. As previous work has shown strong super-spreading tends to enhance the probability of extinction, when starting from a single individual (*7*). If few individuals carry the majority of infections, this might suggest that the number of infected is effectively smaller giving rise to a shorter time to extinction, than for smaller individual variation; this however, may be a relatively weak logarithmic effect, as suggested by the mean time to extinction expression calculated in this paper (Eqn.8).

We have also shown how to calculate accurate extinction time distributions for the case when R_e is sufficiently close to 1 and when population immunity cannot be ignored, as shown in the predictions of Fig.6a, for the highly vaccinated UK population. When R_e is close to 1, changes in the susceptible pool due to new infections cause a large relative change in R_e itself and it cannot be considered constant; we find a simple expression for this threshold R_e^* and use a semi-heuristic method to calculate the extinction time distribution based on calculating the final or ultimate effective reproduction number R_e^∞, which is arrived at once again infections drop to a level that new infections cause a small relative change in the susceptible pool. In general, we find that as $R_e > R_e^*$ and approaches 1, the extinction times plateau, instead of diverging using the constant R_e theory. Interestingly, we find that in this regime higher initial infections lead to significantly shorter extinction times compared to lower initial infections, because the former leads to a lower R_e^∞. This new theoretical calculation also shows explicitly that for the SIR model, the epidemic declines for $R_e = 1$, culminating in an asymptotic $R_e = R_e^\infty < 1$ and ultimately extinction, whereas a naive expectation would be that the

epidemic does not grow or shrink when $R_e = 1$.

We have also made a very crude extrapolation of our results to extinction at the global level. The most problematic assumption, of course, is a single value of R_e globally, and so the predictions in Fig.6b should be viewed as a guide to what could be achieved globally if all countries acted roughly in the same way on average. As we show, sub-divided regions with uniform migration, simply lead to a upwards rescaling of R_e by factor $(1 + \phi)$ where ϕ is the migration probability, and so we can have confidence in these predictions, as long as R_e at a global level is suitably interpreted. Previous work on generalising the concept of the reproductive number to include spread between different regions, uses a different approach, by defining an analogous reproductive number R_* for regions, (*3–5*), i.e. how many sub-populations or regions have at least one infection, where migration plays an analogous role to individual contacts for the spread of infection in a single population. This gives rise to a condition for spread or decline of infections to multiple regions and eventually globally, if $R_* > 1$ or $R_* < 1$, respectively. In the context of a local reproductive number $R_e < 1$, as has been discussed previously (*3*), it is still theoretically possible that $R_* > 1$, particularly when infection is highly prevalent in a region (due to a previous time when $R_0 > 1$, as has occurred across many countries with SARS-Cov-2 in 2020 before lockdowns were imposed); this can happen if the global contact probability is sufficiently large, meaning that the infection can continue to spread between countries and regions. However, these long distance seedings of infection will not in themselves lead to regional or national outbreaks, as long as locally $R_e < 1$ (*49*); the simple upward rescaling of $R_e \to (1 + \phi)R_e$, which we observe for small ϕ captures this phenomenon. These results also suggest that at the national level, a simple rescaling of R_e should describe the reduction in the rate of decline of infections due to importation of infected cases, and should be accounted in more accurate estimates of extinction times.

This brings up an important question: whether a country should pursue elimination whilst other countries are not and there is the chance cases can be imported. As alluded to above, whether before or after zero infections have been achieved, if the country pursuing elimination keeps $R_e < 1$ through NPIs then there cannot be an outbreak; after zero infections have been achieved in practice there will clearly be a pressure to open up, but the counterfactual prediction above of global elimination by January 2021, essentially assumes this is not the case, that all countries keep $R_e < 1$ until zero cases worldwide. If this is not the case, then with any level of migration and non-zero cases abroad, extinction is not a stable state. So in which case, what do the extinction times mean practically at the national level? If there is a total rate M of migrations of infected cases into a country per day and $M \ll 1$ — for example, if there are strict control of immigration, and /or globally infections are rare — then extinction will be a meta-stable state with lifetime $\approx 1/M$ days. On the other hand, if $M \gtrsim 1$ due to open borders and/or highly prevalent infections abroad, then extinction cannot be achieved. As an illustration of these considerations, pre-pandemic levels of immigrations into UK were $\approx 400,000$ per day (*50*), and assuming a prevalence of $\approx 0.09\%$ globally (June 2020), that would mean $M \approx 360$ infected arriving per day; however, according to the Civil Aviation Authority (UK) passenger numbers for June 2020 had dropped significantly, to roughly 2.4% of their levels in June 2019 (*51*), which would then suggest $M \approx 6$ per day. If we further factor in quarantine rules for arrivals from specific countries, then it possible that in aggregate $M < 1$, or potentially even $M \ll 1$. In this case, extinction can persist for a period of time $1/M$, where a country would likely need to be prepared to act quickly to contain any potential outbreaks. This estimate itself is likely to be pessimistic, as each new infection must establish to the critical size $I^* \sim R_e/(R_e - 1)$,

which has probability $p^* = 1 - 1/R_e$ (Williams' threshold theorem (*36*)), before causing an outbreak and so if R_e is not too large many imported infections will initially stochastically die out, and as discussed if $R_e < 1$ the probability is zero. Overall, these considerations highlights that technically extinction can only occur nationally when all infections globally have been eliminated; nonetheless, a joined-up global strategy would entail each country aiming for a state of quasi-extinction, by keeping $R_e \ll 1$, where declining infections world-wide would lead to increasingly longer periods of zero infections nationally. Once infections have been eliminated locally/nationally, and while infections globally persist, this analysis would suggest only allowing R_e to be moderately greater than one, where the actual number would be a balance between acceptable border controls, relaxation of NPIs and how quickly NPIs can be re-introduced and affect a decline in infections, once new infections have been detected.

We see broadly the results in this paper demonstrate that in worst-case scenarios where vaccines cannot be developed for highly transmissible diseases, with relatively short recovery times ($1/\gamma$), the alternative strategy of NPIs can be used to eliminate the virus on practicable time-scales of many months. In the case of SARS-Cov-2 , at the time of writing it is still not very clear to what degree vaccines can cut transmission (*13–16*), even though they have been demonstrated to be highly effective at preventing serious illness. If we assume that current vaccines do not cut transmission significantly, then as the results in Fig.6a show it is possible, assisted by population immunity, to eliminate Sars-Cov2 within the UK within about 6 months given the levels of infection at the time of writing ($I \approx 7 \times 10^5$). However, it is also not known how long vaccine-induced immunity will last; if immunity wanes over the time-period of many months to years standard models, such as the SIRS model, show that our predictions are broadly correct — typically underestimating extinction times, since waning immunity leads to an increasing pool of susceptibles over time — except extinction is only possible if $R_0 < 1$, which could be achieved by NPIs. On the other hand opening up such that $R_0 > 1$, with waning immunity, leads to an endemic state.

Related to vaccine efficacy, this paper also does not consider the trade-off between mitigation and elimination in terms of the evolution of new variants of concern, particularly vaccine escape variants. However, very broadly the literature on population rescue and evolution of resistance (*52–54*) clearly predicts that resistance is more likely with increasing effective population size. In an epidemiological setting, there is a multi-scale aspect to the problem, involving within-host vs between-host evolution, and potential trade-offs between the two (*55–58*), however, if we treat each infected individual as a unit of evolution which produces mutants at a certain rate (determined by selection through an evolutionary substitution process within host), then this would predict that vaccine escape becomes more likely with increasing number of infected individuals. As the cumulative number of infections will be lower with an elimination strategy compared to mitigation, we would expect elimination to have the added benefit of being robust to the evolution of new variants.

Finally an aspect, which we have not considered, is the possibility of a non-human reservoir of SARS-Cov-2 , which could allow re-infections of human populations, such that, as with the case of migrations, extinction is not a permanent (stable) state. This could be accounted in a similar way as we account for migration in this theory; unlike human populations (and arguably even in human populations), we have much less control, or even understanding, of a potential non-human reservoir. However, if such reservoirs, whether bat or pangolin (*59, 60*), can be identified and surveyed (*61*), measures can be taken to control contact with human populations, such that effective global elimination in human populations is possible, even if the virus cannot be eliminated

from non-human populations.

Overall our results suggest an viable and relatively rapid ($\sim$ months) alternative strategy to eliminate infections without having to rely on herd immunity, either naturally acquired, or through vaccination; in a worst case scenario, for highly transmissible diseases where no effective vaccine is found, and/or immunity is short-lived, such alternative strategies may be the only option remaining.

References

1. W. O. Kermack, A. G. McKendrick, *Proceedings of the Royal Society,* **115A,**, 700721 (1927).
2. J. Murray, *Mathematical Biology I: An Introduction* (Springer-Verlag, 2002), chap. 10, pp. 319–326.
3. F. Ball, D. Mollison, G. Scalia-Tomba, *The Annals of Applied Probability* **7**, 46 (1997).
4. P. C. Cross, J. O. Lloyd-Smith, P. L. F. Johnson, W. M. Getz, *Ecology Letters* **8**, 587 (2005).
5. V. Colizza, A. Vespignani, *Physical Review Letters* **99**, 148701 (2008).
6. N. M. Ferguson, *et al.*, *Nature* **425**, 681 (2003).
7. J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, *Nature* **438**, 355 (2005).
8. W. R. Dowdle, *Bulletin of the World Health Organization* **76 Suppl 2**, 22 (1998).
9. S. Hendy, *et al.*, *Journal of the Royal Society of New Zealand* **51**, 1 (2021).
10. H. Ward, *et al.*, *medRxiv* (2020).
11. A. W. D. Edridge, *et al.*, *Nature Medicine* **26**, 1691 (2020).
12. J. M. Dan, *et al.*, *Science* **371**, eabf4063 (2021).
13. M. C. Shamier, *et al.*, *medrxiv* (2021).
14. I. Kroidl, *et al.*, *Eurosurveillance* **26**, 2100673 (2021).
15. N. E. Blachere, E. Hacisuleyman, R. B. Darnell, *New England Journal of Medicine* **385**, e7 (2021).
16. E. Hacisuleyman, *et al.*, *New England Journal of Medicine* **384**, 2212 (2021).
17. N. G. Davies, *et al.*, *Science* p. eabg3055 (2021).
18. K. Leung, M. H. Shum, G. M. Leung, T. T. Lam, J. T. Wu, *Eurosurveillance* **26** (2021).
19. S. A. Kemp, *et al.*, *Nature* pp. 1–10 (2021).
20. T. Crellen, *et al.*, *medRxiv* (2021).
21. T. Britton, *Mathematical biosciences* **225**, 24 (2010).

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

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

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

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

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

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

611 28. I. Nasell, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61**, 309
612 (1999).

613 29. C. P. Farrington, A. D. Grant, *Journal of Applied Probability* **36**, 771 (1999).

614 30. Office for National Statistics, Coronavirus (COVID-19) Infection Survey pilot: England, 12 June
615 2020.

616 31. Office for National Statistics, Coronavirus (COVID-19) Infection Survey, antibody data for the
617 uk: 16 February 2021.

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

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

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

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

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

623 37. R. A. Fisher, L. H. C. Tippett, *Mathematical Proceedings of the Cambridge Philosophical Society*
624 **24**, 180 (1928).

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

626 39. Office for National Statistics, Coronavirus (COVID-19) Infection Survey pilot: England, 19 Febru-
627 ary 2021.

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

629 41. R. Verity, *et al.*, *The Lancet. Infectious diseases* **20**, 669 (2020).

630 42. S. Ghisolfi, *et al.*, Working Paper 535: Predicted COVID-19 Fatality Rates Based on Age, Sex,
631 Comorbidities, and Health System Capacity (2020).

632 43. MRC Biostatistics Unit, University of Cambridge, COVID-19 Nowcast and forecast.
633 https://www.mrc-bsu.cam.ac.uk/now-casting/.

634 44. T. Britton, F. Ball, P. Trapman, *Science (New York, N.Y.)* (2020).

635 45. M. Chikina, W. Pegden, *PLOS ONE* **15**, e0236237 (2020).

636 46. A. Endo, C. for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group,
637 S. Abbott, A. J. Kucharski, S. Funk, *Wellcome Open Res* **5** (2020).

638 47. D. C. Adam, *et al.*, *Nature Medicine* **26**, 1714 (2020).

639 48. L. Wang, *et al.*, *Nature Communications* **11**, 5006 (2020).

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

642 50. UK Government: Home Office, National Statistics: How many people come to the UK each year
643 (including visitors)? https://www.gov.uk/government/publications/immigration-statistics-year-
644 ending-june-2019/how-many-people-come-to-the-uk-each-year-including-visitors.

645 51. Civil Aviation Authority, Airport data: https://www.caa.co.uk/data-and-analysis/uk-aviation-
646 market/airports/datasets/uk-airport-data/airport-data-2020-06/.

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

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

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

650 55. S. Bonhoeffer, M. A. Nowak, *Proceedings of the National Academy of Sciences* **91**, 8062 (1994).

651 56. E. C. Holmes, *Journal of Virology* **77**, 11296 (2003).

652 57. C. Fraser, *et al.*, *Science* **343**, 1243727 (2014).

653 58. K. A. Lythgoe, L. Pellis, C. Fraser, *Evolution* **67** (2013).

654 59. K. G. Andersen, A. Rambaut, W. I. Lipkin, E. C. Holmes, R. F. Garry, *Nature Medicine* **26**, 450
655 (2020).

656 60. A. Latinne, *et al.*, *bioRxiv* (2020).

657 61. M. Watsa, *Science* **369**, 145 (2020).

658 Acknowledgements

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

661 Competing interests

662 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.

Supplementary Materials for

How long does it take to eliminate an epidemic without herd immunity?

Bhavin S. Khatri*

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

This PDF file includes:

Supplementary Text
Figs. S1
References (S1)

Supplementary Text

Estimating extinction times from direct estimate of ρ_e

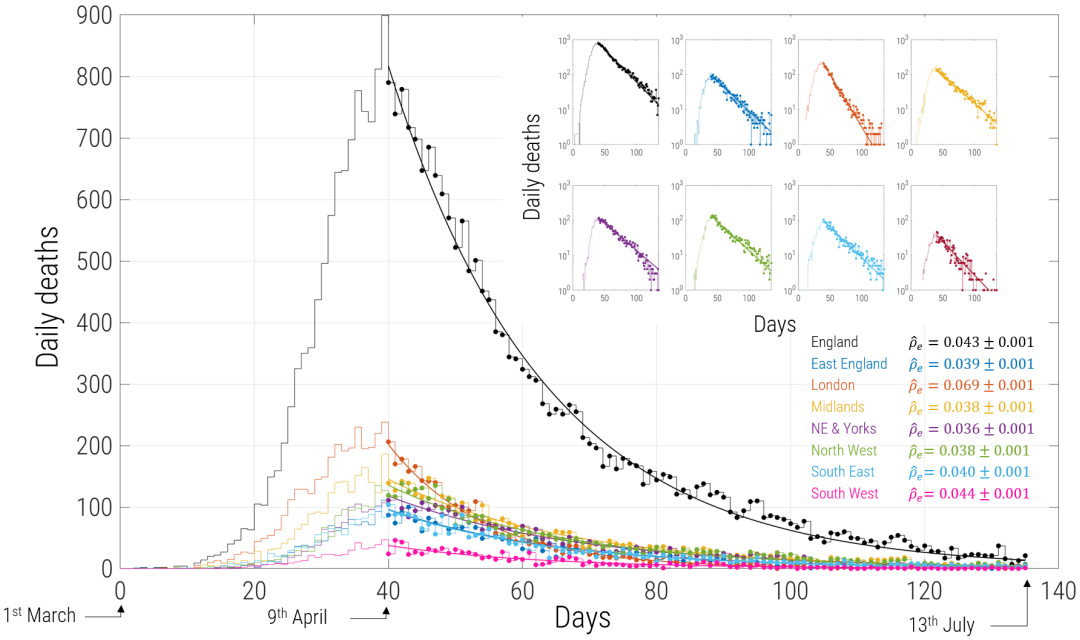

Figure S1: Data from NHS England (*S1*) on daily deaths within hospitals in England and different regions within England. Each region is fit using a simple decaying exponential from 40 days after the 1st March 2020, where each region first shows a clear decline in number of daily deaths.

Precise predictions of extinction times based on Eqn.7 in the main text require knowledge of both R_e and γ, which are generally quite difficult to estimate. However, Eqn.7 in the main text is mainly dependent on the rate of decline of the epidemic ρ_e, with weak dependence separately on R_e and γ. ρ_e can be determined more straightforwardly if we assume current daily deaths are proportional to the number infected; if infections are declining exponentially at a rate ρ_e then so will the number of daily deaths, so a curve fit will give an accurate measure of ρ_e even if we cannot determine the proportionality constant to translate deaths to infections. An alternative could be to look at time-series of number of daily infections per number of tests performed, to remove biases due to testing, however, here the aim is to illustrate why a direct estimate of ρ_e is useful, rather than to estimate this number very accurately.

As shown in Fig.S1, fitting decaying exponentials to daily number of deaths (date of death, not date of reporting) from the NHS UK (*S1*), from the moment of decline to 13th July 2020, shows

a very good fit (showing deaths are declining as a simple exponential and giving further weight to our simple model), giving an England wide estimate of $\rho_e = 0.043 \pm 0.001$ days^{-1}, with a range of $\rho_e = 0.036 \pm 0.001$ days^{-1} (slowest decline) for North Yorkshire and $\rho_e = 0.069 \pm 0.001$ days^{-1} (most rapid decline) in London. As an example, we estimate the mean time to extinction had the UK remained in lockdown from July 2013, when infections were $I_0 \approx 3.7 \times 10^4$; using the rate of decline of England, we estimate a mean extinction time in the UK as 231 ± 30 days (95% CI: $(187, 303)$ days), for $R_e = 0.7, 1/\gamma = 7$ days and 238 ± 30 days (95% CI:$(195, 310)$ days), for $R_e = 0.57, 1/\gamma = 10$ days; as we can see changing R_e and γ for a fixed ρ_e does not change the predictions significantly, and we suggest in general, determining ρ_e could be a more robust way to estimate extinction times.

Branching process analysis

A birth-death process with birth rate $b = \gamma R_e$ and death rate $d = \gamma$, corresponds to a pure exponential growth ($R_e > 1$) or decay ($R_e < 1$) phase of an epidemic, when the number of susceptible individuals $S(t)$ are far in excess of the total number infected $I(t)$. In this appendix, for presentational clarity, we will use $n = I$ to represent the number of infected individuals. To write down the rate of change of the probability of n infected individuals at time t, we need only consider the probability of having $n-1, n, n+1$ individuals and rates of transitions between them, since in the limit of infinitesimal (continous) changes in time, we consider only changes of single individuals. The rate of transition from $n-1 \to n$ happens with rate $\gamma R_e(n-1)$ and the rate of transition from $n+1 \to n$ happens with rate $\gamma(n+1)$, which both lead to an increase in the probability of n, while the rate of transition from $n \to n+1$ happens with rate $\gamma R_e n$ and the rate of transition from $n \to n-1$ happens with rate γn, which both decrease the probability of n. Using these facts we can write down the rate of change of the probability of n at time t:

$$\begin{aligned}\frac{\mathrm{d}p_n(t)}{\mathrm{d}t} = \gamma\big(&R_e(n-1)p(n-1,t) \\ &- (R_e+1)np(n,t) + (n+1)p(n+1,t)\big).\end{aligned} \tag{S1}$$

However, this description isn't complete, and we need to consider how the probability of the $n = 0$ state changes, since the above equation won't work for $n - 1$, since we cannot have a negative number of individuals:

$$\frac{\mathrm{d}p_{n=0}(t)}{\mathrm{d}t} = -\gamma\left(R_e np(n,t) + (n+1)p(n+1,t)\right). \tag{S2}$$

We can encompass both equations together in one by using the unit step function $U_n = 1$ for $n \geq 0$, while $U_n = 0$ for $n < 0$:

$$\begin{aligned}\frac{\mathrm{d}p(n,t)}{\mathrm{d}t} = \gamma\big(&U_{n-1}R_e(n-1)p(n-1,t) \\ &- U_n(R_e+1)np(n,t) + (n+1)p(n+1,t)\big).\end{aligned} \tag{S3}$$

For $n < 0$, as long as we have an initial condition, $p_{n<0}(t = 0) = 0$, the ODEs above guarantee that $p_{n<0}(t) \forall t$. Considering each value of $n : 0 \leq n < \infty$, we have an infinite set of coupled differential equations. The standard way to solve this is to use probability generating functions:

$$G(z,t) = \sum_{n=0}^{\infty} p_n(t) z^n, \tag{S4}$$

which is in general a complex function of a complex variable z. Using the fact that $z\partial G(z,t)/\partial t = \sum n p_n(t) z^n$, it is straightforward to show that the set of ODEs give the following first order partial differential equation for $G(z,t)$:

$$\frac{\partial G(z,t)}{\partial t} = \alpha(z) \frac{\partial G(z,t)}{\partial t}, \tag{S5}$$

where

$$\alpha(z) = \gamma(R_e z^2 - (R_e + 1)z + 1). \tag{S6}$$

This PDE can be solved by using the method of characteristics, which finds a parametric path $z(s), t(s)$ along which our original PDE is obeyed. The rate of change of $G(s)$ along this path in terms of our parameterisation is:

$$\frac{\mathrm{d}G(s)}{\mathrm{d}s} = \frac{\mathrm{d}t}{\mathrm{d}s}\frac{\partial G}{\partial t} + \frac{\mathrm{d}z}{\mathrm{d}s}\frac{\partial G}{\partial z}, \tag{S7}$$

and so with reference to the original PDE (Eqn.S5), we can identify that

$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1 \quad \frac{\mathrm{d}z}{\mathrm{d}s} = -\alpha(z). \tag{S8}$$

Integrating these pair of equations gives the characteristic paths for which $\mathrm{d}G(s)/\mathrm{d}s = 0$ is a constant:

$$\frac{z-1}{z-1/R_e} e^{\gamma(R_e-1)t} = C, \tag{S9}$$

where C is a constant that represents different possible initial conditions. Integrating $\mathrm{d}G(s)/\mathrm{d}s = 0$, gives

$$G(s) = \phi\left(\frac{z-1}{z-R_e^{-1}} e^{\gamma(R_e-1)t}\right), \tag{S10}$$

where ϕ is an arbitrary function to be determined by consideration of the initial conditions on $p_n(t)$. We can use the fact that at time $t = 0$ we assume we know the exact number of infected individuals is n_0 and hence, $p_n(t = 0) = \delta_{nn_0}$, where $\delta_{nn_0} = 0$ for $n \neq n_0$ and $\delta_{nn_0} = 1$ for $n = n_0$. Calculating the probability generating function for the initial delta function probability mass, we get $G(z, t = 0) = z^{n_0}$, and so we need to find a function ϕ satisfying:

$$G(z,0) = \phi\left(\frac{z-1}{z-R_e^{-1}}\right) = z^{n_0}. \tag{S11}$$

Substituting $x = (z-1)/(z-1/R_e)$, we can find $\phi(x)$, to give our solution:

$$G(z,t) = \left(\frac{1 + (z-1)e^{\gamma(R_e-1)} - zR_e}{1 + R_e(z-1)e^{\gamma(R_e-1)} - zR_e} \right)^{n_0}. \tag{S12}$$

Our probability mass function $p_n(t)$, should always be normalised $\sum_n p_n(t) = G(z=1,t) = 1$; substituting $z = 1$ we see this that the solution $G(z,t)$ behaves correctly. Finally, the reason this is all useful, is that we want to calculate the probability of zero individuals infected $p_0(t)$, which is simply given by $G(z=0,t)$, since $0^0 = 1$:

$$p_0(t) = G(0,t) = \left(\frac{1 - e^{\gamma(R_e-1)}}{1 - R_e e^{\gamma(R_e-1)}} \right)^{n_0}. \tag{S13}$$

Substituting $n_0 = I_0$ and $\rho_e = \gamma(1 - R_e)$ gives Eqn.6 in the main text.

Differentiating Eqn.S13 to obtain the extinction time distribution, we find

$$p^\dagger(t) = \frac{\mathrm{d}p_0(t)}{\mathrm{d}t} = (1-R_e)\rho_e n_0 \frac{(1-e^{-\rho_e t})^{n_0-1}}{(1-R_e e^{-\rho_e t})^{n_0+1}} e^{-\rho_e t}. \tag{S14}$$

This is an exact expression, which is valid for all values of I_0 and $I^\dagger$, as long as the original assumptions of the model that changes in susceptible numbers are negligible ($R_e < R_e^*$) is true, where R_e^* is given by Eqn.S19. If $I_0 \gg I^\dagger$ then we expect there to be a strong division between the deterministic phase and the stochastic phase, such that in Eqn.S14 the exponentials have sufficiently decayed such that $n_0 e^{-\rho_e t} \ll 1$, before any extinction is likely, then it is straightforward to show that the limiting form of Eqn.S14, is the Gumbel distribution, Eqn.7 in the main text, using the fact that $(1-e^{-\rho_e t})^{n_0} \approx (1 - n_0 e^{-\rho_e t}) \approx \exp(-n_0 e^{-\rho_e t})$.

Extinction for $R_e > R_e^*$

Infections decline when $R_e < 1$. $R_e(t) = R_0 \times S(t)/N$ is in general time-dependent, composed of two factors, $R_0 = \beta/\gamma$, which for simplicity we assume is time-independent and $S(t)/N$, which will tend to decrease in time as more susceptibles become infected, so the rate of decline $\rho_e = \gamma(1 - Re)$ is in general time-dependent and increasing over time. When $R_e \ll 1$, reductions in transmissions due to NPIs dominates the decrease in infections, compared to the fractional change in the susceptible pool and so $R_e \approx R_0 S_0/N$ is constant to a good approximation. However, when $R_e < 1$ but close to 1, this is no longer true, and the assumption that the number of susceptibles is roughly constant $S(t) \approx S_0$ with respect to changes in $I(t)$ and $R_e(t)$ is a poor one.

It is within this context that we would like to calculate the extinction time distribution. Although, an exact solution is not easily obtainable, we can make a semi-heuristic approximation that works very well. Initially R_e is time-dependent since the changing susceptible pool has significant affect on the decline in infections. However, once infections become sufficiently small the change in the susceptible pool, per unit time, once again becomes relatively small compared to its current value and $S(t) \to S^\infty$ attains its asymptotic value S^∞, at which point the constant R_e assumption becomes accurate again and infections decline at a constant rate. The asymptotic value S^∞ cannot be calculated via standard fixed point analysis of the SIR differential equations, since the only condition

for a fixed point is that $I = 0$, and this can happen for any value of S; the final asymptotic values depend on the initial conditions. Taking the SIR differential equations and calculating $\dot{S}/\dot{R}$ we have

$$\frac{\mathrm{d}S}{\mathrm{d}R} = -\frac{\beta S}{\gamma N} = \frac{R_0}{N} S. \tag{S15}$$

Integrating this equation, starting from an initial condition $S(0) = S_0$ and $R(0)$ to their final asymptotic values S^∞ and $R^\infty = N - I^\infty - S^\infty = N - S^\infty$, then we arrive at the following transcendental equation for S^∞:

$$S^\infty = S_0 e^{R_0(1-S^\infty/N - R(0)/N)}. \tag{S16}$$

The solution can however, be expressed using the Lambert W function:

$$S^\infty = -\frac{N}{R_0} W\left(-\frac{R_0 S_0}{N} e^{-R_0(1-R(0)/N)}\right), \tag{S17}$$

where $w = W(z)$ is the solution to the transcendental equation $we^w = z$. We are interested in finding the asymptotic effective reproductive number $R_e^\infty = R_0 S^\infty/N$ in terms of the initial effective reproductive number $R_e = R_0 S_0/N$, for which the above expression can be rearranged to give

$$R_e^\infty = -W\left(-R_e e^{-R_0(1-R(0)/N)}\right). \tag{S18}$$

We can replace $R_e \to R_e^\infty$ in Eqn.7 of the main text to calculate the distribution of extinction times to give a good approximation of the extinction times when $R_e \approx 1$ and where the above condition for constant R_e is not met. However, this gives a systematic underestimate of the time to extinction, since it effectively ignores the time it takes to attain these asymptotic values, which takes of order $1/\rho_e^\infty$ days, where $\rho_e^\infty = \gamma(1 - R_e^\infty)$. So finally an accurate and robust approximation to the extinction time distribution is obtained by the replacement $R_e \to R_e^\infty$ and $\tau^\dagger \to \tau^\dagger + 1/\rho_e^\infty$, as we can see in Fig.4b in the main text for simulations of $R_e = 0.99$ and $1/\gamma = 7$ days. Note that for sufficiently small R_e the correction to τ is not needed, as $R_e^\infty \approx R_e$ and the assumption of constant R_e is very accurate. We approximate this threshold value of R_e as the value of $R_e^\infty(R_e \to 1)$:

$$R_e^* = R_e^\infty(R_e = 1) = -W(-e^{-R_0(1-R(0)/N)}). \tag{S19}$$

which is roughly the plateau value of R^∞, which will robustly be close to the value of R_e that R_e^∞ begins to deviate from R_e. We can then also stitch together $\tau^\dagger$ for $R_e < R_e^*$ and $\tau + 1/\rho^\infty$ for $R_e > R^*$ using a standard $\tanh$ switching function centred on R_e^* and with width 0.05, which is used in Figs. 5&6 in the main text to provide the extinction time predictions across the whole range of $0 < R_e < 1$.

Invariance of extinction time distribution to population sub-division

If we imagine a single population to be divided into n equally sized sub-populations, each with a reproductive number R_e and zero-migration between, then the extinction time distribution of t_k in the k^{th} sub-population will be given by Eqn.7 in the main text, but with $I_0 \to I_0/n$. Now we want

to calculate the extinction time distribution of the whole population. Extinction will occur when all sub-populations have zero infected individuals. We can record the extinction times in each sub-population: $t_1, t_2, ..., t_k, ..., t_n$ and the extinction time of the whole population will be the maximum of this set: $\tilde{t} = \max\{t_1, t_2, ..., t_k, ..., t_n\}$. The cumulative distribution function of the maximum time $\tilde{t}$ will be the probability of the joint event that each sub-population k has an extinction time less than $\tilde{t}$:

$$\begin{aligned} P_n(\tilde{t}) &= P(t_1 < \tilde{t}, t_2 < \tilde{t}, ..., t_k < \tilde{t}, ..., t_n < \tilde{t}) \\ &= P(t_1 < \tilde{t})P(t_2 < \tilde{t})...P(t_k < \tilde{t})...P(t_n < \tilde{t}) \\ &= (P(\tilde{t}))^n \end{aligned} \tag{S20}$$

where $P(t)$ is the CDF for a single population given by Eqn.10 in the main text, but with $I_0 \to I_0/n$. Given the form of Eqn.10, these calculations can be performed exactly, whereas using extreme value theory it usually required that the tails of the distribution asymptotically obey some exponential form, which allows approximate calculation. Doing these calculations we find $(P(\tilde{t}))^n = (\exp(-e^{-\rho_e(\tilde{t}-\tau_n^\dagger)}))^n$, where $\tau_n^\dagger = \frac{1}{\rho_e}\ln(I_0/nI^\dagger)$. It is then simple to show that the n-dependence cancels in the final result to give

$$P_n(\tilde{t}) = P(\tilde{t}) = \exp(-e^{-\rho_e(\tilde{t}-\tau^\dagger)}). \tag{S21}$$

In other words, population sub-division into equal sized isolated populations does not affect the extinction time distribution of the whole global population. In fact, it is simple to extend these arguments to any population sub-division, where $I_0 = \sum_{k=1}^{n} I_k$, where I_k is the initial infected population in each, as long as the fraction of susceptible and R_e is the same in each sub-population. This is not surprising, as it is just a restatement of the mean-field/well-mixed approximation that infected individuals and sub-populations all experience the same probability of encountering a susceptible individual S_0/N which is set by the global number of susceptible individuals S_0.

References

S1. NHS England, NHS UK COVID-19 all announced deaths – 18th July 2020. https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-daily-deaths/.