UC Berkeley
UC Berkeley Previously Published Works

Title

The effect of the definition of pandemic on quantitative assessments of infectious disease outbreak risk.

Permalink

Journal

Authors

Singer, Benjamin
Thompson, Robin
Bonsall, Michael

Publication Date

DOI

Copyright Information

scientific reports

OPEN

The effect of the definition of 'pandemic' on quantitative assessments of infectious disease outbreak risk

Benjamin J. Singer[1✉], Robin N. Thompson[2,3] & Michael B. Bonsall[1]

In the early stages of an outbreak, the term 'pandemic' can be used to communicate about infectious disease risk, particularly by those who wish to encourage a large-scale public health response. However, the term lacks a widely accepted quantitative definition. We show that, under alternate quantitative definitions of 'pandemic', an epidemiological metapopulation model produces different estimates of the probability of a pandemic. Critically, we show that using different definitions alters the projected effects of key parameters—such as inter-regional travel rates, degree of pre-existing immunity, and heterogeneity in transmission rates between regions—on the risk of a pandemic. Our analysis provides a foundation for understanding the scientific importance of precise language when discussing pandemic risk, illustrating how alternative definitions affect the conclusions of modelling studies. This serves to highlight that those working on pandemic preparedness must remain alert to the variability in the use of the term 'pandemic', and provide specific quantitative definitions when undertaking one of the types of analysis that we show to be sensitive to the pandemic definition.

In the early stages of an infectious disease outbreak, it is important to determine whether the pathogen responsible may go on to cause an epidemic or a pandemic[1–5]. There is extensive literature on determining the probability of a major epidemic given a small population of initial infected hosts[6–9]. This research has produced a natural mathematical definition of an epidemic, based on the bimodal distribution of outbreak sizes given by simple stochastic epidemiological models when R_0 is larger than but not close to one[10]. The term 'pandemic' has no corresponding theoretical definition, and there is no consensus mathematical approach to determining the probability of a pandemic. In this study, we examine how alternative definitions of 'pandemic' affect quantitative estimates of pandemic risk assessed early in an infectious disease outbreak.

The term 'pandemic' is used extensively, appearing in phrases such as 'pandemic preparedness'[11–13], 'pandemic influenza'[14–16], and 'pandemic potential'[17–19]. A Google Scholar search returns 25,800 results using the term 'pandemic' for 2019 alone.

The International Epidemiology Association's Dictionary of Epidemiology defines a pandemic as "an epidemic occurring worldwide, or over a very wide area, crossing international boundaries and usually affecting a large number of people"[20]. Notably this definition makes an explicit reference to national borders. Contrastingly, a World Health Organization (WHO) source makes reference to a pandemic as "the worldwide spread of a new disease"[21]. The use of the word 'new' here is ambiguous in the context of infectious diseases. HIV/AIDS is often referred to as a global pandemic, but is certainly not new on the timescale of, say, the emergence of influenza strains[22,23]. A study by Morens et al. in 2009 finds that there is little in common between all disease outbreaks that have been referred to as pandemics, except that they have a wide geographical extension[24].

These kinds of differences between pandemic definitions can often go unnoticed, but in certain circumstances they can cause confusion between different stakeholders (e.g. between scientists and governments, or governments and the public), who may not have a shared background understanding of the term. In 2009, the WHO declared a pandemic of H1N1 influenza, using criteria related to the incidence and spread of the virus in different WHO regions[25]. The criteria did not include reference to morbidity or mortality[26]. This fact led to some controversy over whether the declaration of a pandemic was appropriate, as the declaration prompted some governments to mount an intensive response to an outbreak that resulted in fewer yearly deaths than a typical strain of seasonal flu[27–30].

[1]Department of Zoology, University of Oxford, Oxford, UK. [2]Christ Church, University of Oxford, Oxford, UK. [3]Mathematical Institute, University of Oxford, Oxford, UK. ✉email: benjamin.singer@bnc.ox.ac.uk

International health organisations such as the WHO have not provided any formal definitions of the term 'pandemic', and the WHO no longer uses it as an official status of any outbreak[25,31]. It would, however, be hasty to dismiss the importance of the term on these grounds. Although the WHO no longer uses the term 'pandemic' officially, the WHO Director-General drew attention to their use of the term as recently as March 2020, to describe the status of the COVID-19 outbreak[32]. The Director-General cited "alarming levels of inaction" as one of the reasons to use the term, along with the caveat that "describing the situation as a pandemic does not change WHO's assessment of the threat posed by this virus". The WHO's use of the term was of interest to the public, receiving extensive press coverage[33–35]. The term 'pandemic' clearly continues to be important to indicate serious risk during disease outbreaks.

Regardless of the extent to which the pandemic definitions currently in use do or do not agree, they are all qualitative in nature, using descriptions such as "very wide area" and "large number of people". Perhaps as a result of this, many quantitative studies on pandemics do not make use of a quantitative definition of a pandemic, but instead focus on causally related concepts, such as sustained transmission[19], or emergence of novel viruses[36]. Others treat the spread of a pathogen at a pandemic level as a context in which to study transmission dynamics, without paying special attention to how those dynamics might lead to a pandemic as distinct from an epidemic or a more limited outbreak[37–39]. In this paper, we examine the effects of alternative pandemic definitions on the analysis of key epidemiological questions. The results provide a foundation for deciding the appropriate quantitative definition of 'pandemic' in a given context.

We use a metapopulation model to investigate the effects of pandemic definition on the results of a quantitative assessment of the probability of a pandemic. Metapopulation models are commonly applied to pathogens that spread between regions of the world, and so are appropriate for modelling pandemics[40–45]. We represent states of our metapopulation model as states of a Markov chain, allowing us to calculate the probability of a pandemic directly, as opposed to simulating many stochastic outbreaks and recording the proportion which result in pandemics. We explore two different kinds of pandemic definition, following Morens et al. 2009[24], specifically:

- the family of *transregional* definitions, where a pandemic is defined as an outbreak in which the number of regions experiencing epidemics meets or exceeds some threshold number *n*. We refer to specific transregional definitions as *n-region* transregional definitions, e.g. a 3-region transregional definition.
- the *interregional* definition, where a pandemic is defined as an outbreak in which two or more non-adjacent regions experience epidemics.

Note that these definitions require a specific sense of 'region'. These regions could be countries, or they could be larger or smaller than individual countries—from counties to health zones to WHO regions. Our metapopulation model (detailed in the Methods section below) can be used to model regions of any size. We have chosen not to include definitions with criteria relating to the number of people infected or killed, instead of, or in addition to, geographical extension. Extension is the only universal factor in pandemic definitions, and so is the focus of the current study[24].

Three questions that help form public health policy at the beginning of an outbreak are:

- Would interventions restricting travel reduce the risk of a pandemic?
- Does a portion of the population have pre-existing immunity, and does this affect the risk of a pandemic?
- How is the risk of a pandemic affected by regional differences in transmission?

Using our metapopulation model, we explore how changing the pandemic definition does or does not affect our answers to these questions. We show that the precise definition of a pandemic used in modelling studies can (but does not always) affect the inferred risk. The predicted effects of travel restrictions, the influence of pre-existing immunity, and the impact of regional differences in transmission can all vary when alternative definitions of 'pandemic' are used. This demonstrates clearly the need to consider carefully the pandemic definition used to assess the risk from an invading pathogen. This is necessary for clear communication of public health risk.

Results

Travel rates. One important question about pandemic risk is what effect inter-regional travel rates have on the probability of a pandemic occurring[16,17,46,47]. Here we model epidemics occurring in regions connected on a network in which the connections and their weighting can be set at fixed values representing the rates of travel between regions. We consider simple networks that can illustrate the effects of our different pandemic definitions—namely, the star network, in which one central region is connected to all others with equal weighting and the non-central regions lack any other connections, and the fully connected network, in which each region is connected to every other with equal weighting. Figure 1 illustrates that the connectivity of the full network is much higher than that of the star network. Using the star network allows us to make the distinction between adjacent and non-adjacent regions, thus allowing us to distinguish between transregional and interregional pandemic definitions.

Unless otherwise stated, all figures in the current study are generated with a transmission rate of $\beta = 0.28$ per day, a recovery rate of $\mu = 0.14$ per day, and an inter-regional travel rate of 2×10^{-4} per day. This corresponds to a within-region basic reproduction number (R_0) of 2. These values are within the plausible range for both seasonal and pandemic influenza, and as such they can be used to simulate a plausible pathogen of pandemic potential[38]. We further assume an initial population of 1000 susceptible individuals in each region, and that the outbreak is seeded by a single infectious individual in one region. In the full network, all regions are equivalent,
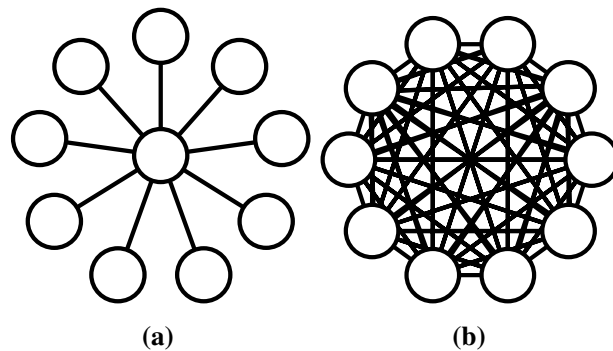
Figure 1. Illustrations of (**a**) a star network and (**b**) a full network, each with ten regions. Circles represent regions, and straight lines represent travel routes between regions.
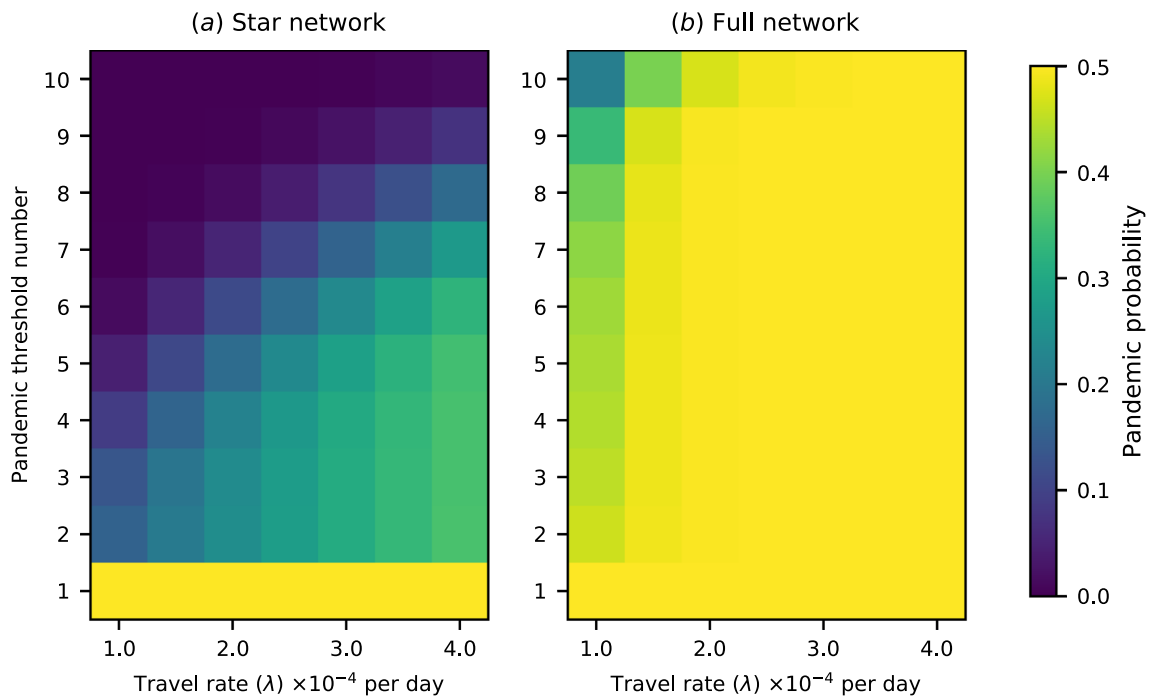

Figure 2. Pandemic probability for a range of between-region travel rates and a range of transregional pandemic definitions. The "pandemic threshold number" refers to the minimum number of regions that must experience epidemics before a pandemic is declared. The pandemic probability is, in general, sensitive to the pandemic definition used, but the degree of sensitivity depends on network structure and travel rates. (**a**) Pandemic probability for a star network. The pandemic probability is, in general, highly sensitive to the pandemic definition used. (**b**) Pandemic probability for a fully connected network. The sensitivity of the pandemic probability to the pandemic definition used is significantly reduced at high travel rates.

so we seed the outbreak in a single arbitrary region. In the star network, we take the average probability of a pandemic over outbreaks seeded in each region.

Using a model with ten regions allows us to test a range of different transregional definitions of a pandemic. The pandemic probability under a range of n-region transregional definitions for a 10-region network with a variety of travel rates is shown in Fig. 2. An n-region transregional definition effectively provides a threshold number n—if more than n regions experience epidemics, the outbreak is counted as a pandemic, and otherwise it is not. Thus we indicate the different possible n-region definitions through their threshold numbers in Figs. 2, 5, and 6.

The 1-region transregional definition merges the definitions of 'pandemic' and 'epidemic' in an implausible way, but it is included in these figures for comparison. The comparison between the pandemic probability according to the 2-region definition and according to the 10-region definition shows the difference between pandemic definitions that are satisfied by any transregional transmission and definitions that are satisfied only by truly global spread. For the star network, or for the fully connected network with low travel rates, there is a marked difference between the probability of either of these definitions being satisfied. However, for the fully connected network at medium or high rates of travel, if the pathogen invades the initial region successfully, then it will go
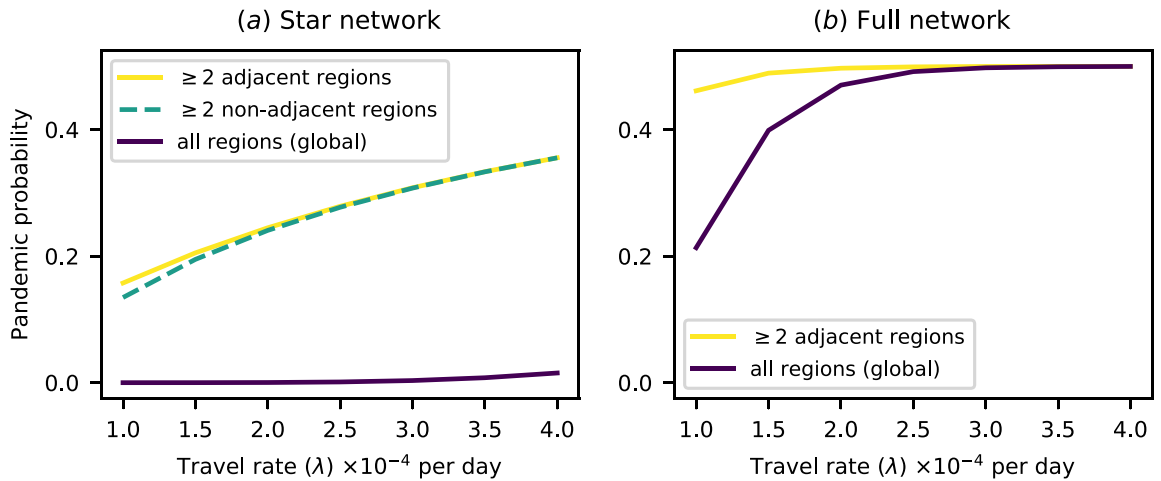
(a) Star network   (b) Full network

Figure 3. Plots of pandemic probability against between-region travel rate for a range of pandemic definitions. The difference in probability for different pandemic definitions changes as travel rates increase. (**a**) Plot of pandemic probability for a star network. (**b**) Plot of pandemic probability for a full network. For a fully connected network all regions are adjacent, so no line is shown for the interregional definition, which requires non-adjacent regions to experience epidemics.

on to spread globally. As such, the probability of a pandemic is nears the maximum of 0.5 (i.e. $1 - 1/R_0$) at all thresholds. For any definition, the probability of a pandemic increases with the connectivity of the network, and with travel rates across the network.

We can also explore the difference in pandemic probability between the transregional and interregional definitions, which make use of a distinction between adjacent and non-adjacent regions. This is shown for a 10-region star network in Fig. 3a, in which we consider the 2-region transregional and 2-region interregional definitions. We choose a star network as it is one of the simplest network types in which there are adjacent and non-adjacent regions. There is a difference between the 2-region interregional and transregional definitions, but the difference is much smaller than that between the 2-region interregional and 10-region (global) definition, and reduces as travel rates increase. In the case of a fully connected network, all regions are essentially adjacent to each other, so we compare only the 2-region transregional and global definitions. We find that the definitions are clearly distinct for low travel rates, but as the travel rate increases the difference between the likelihood of a pathogen causing an epidemic in one region and the likelihood of it causing epidemics in all regions disappears. This is due to the fact that the pathogen can be introduced into any population from any other.

In this section we have shown that, when a pandemic is defined in terms of which regions experience epidemics of a disease, different definitions can produce very different estimates of the pandemic probability at low connectivity or travel rates, but have a much smaller effect at high connectivity and travel rates. In the supplementary information, we illustrate that effects due to network structure are mostly due to the difference in motility between the full network and the star network, although topology still plays an important role.

Cross-immunity. Some pathogens with pandemic potential have a prior history of infecting humans, such as pandemic influenza. Newly emerged pathogens with no history of infecting humans are less likely than these established pathogens to encounter regions where susceptible individuals have partial immunity to infection. Established pathogens may encounter individuals with partial immunity acquired from infections with previously circulating strains—i.e. *cross-immunity*[48,49]. It can be important in responding to an outbreak to consider whether any individuals might have existing immunity. We can therefore investigate the interaction between immunity generated by prior exposure and pandemic definition by examining how cross-immunity affects our calculation of the pandemic probability on a network.

We modelled the spread of a pathogen over a ten-region network with no cross-immunity initially, where the initial infected individual could originate in any region. We only included cases where at least one region experienced an epidemic of this initial pathogen. To simulate the emergence of a strain with higher pandemic potential, we then introduced a second pathogen with a higher transmission rate of $\beta = 0.42$ (corresponding to a basic reproduction number of 3), to which infection with the initial pathogen conferred some degree of partial immunity to infection. The strength of this immunity is written as α. See the Methods section for details of how cross-immunity is incorporated into our modelling framework. We defined a pandemic as occurring when all ten regions experienced epidemics of the second pathogen, and repeated the model for two values of the level of cross-immunity at a variety of between-region travel rates. The results are presented in Fig. 4.

First, increasing cross-immunity decreases the probability of a pandemic. Second, the presence of cross-immunity changes how pandemic probability scales with travel rates. In general, the pandemic probability increases faster with travel when the level of cross-immunity is low, except when it reaches a point of saturation as in Fig. 4b.
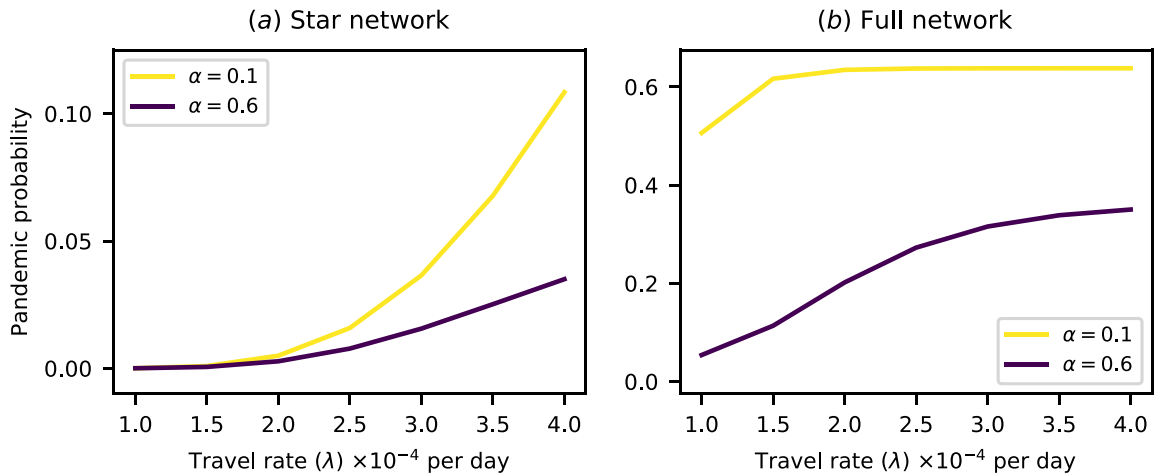
Figure 4. Plots of pandemic probability against travel rate for high and low levels of cross-immunity (α) on ten-region networks. A pandemic is defined here as all ten regions experiencing epidemics, i.e. the 10-region transregional definition. The plots show a large relative difference both in the probability of pandemics and in how that probability scales with travel rates for different levels of cross-immunity. The initial infected individual for each outbreak originates in a randomly chosen region. (**a**) Plot of pandemic probability for a star network. (**b**) Plot of pandemic probability for a full network.
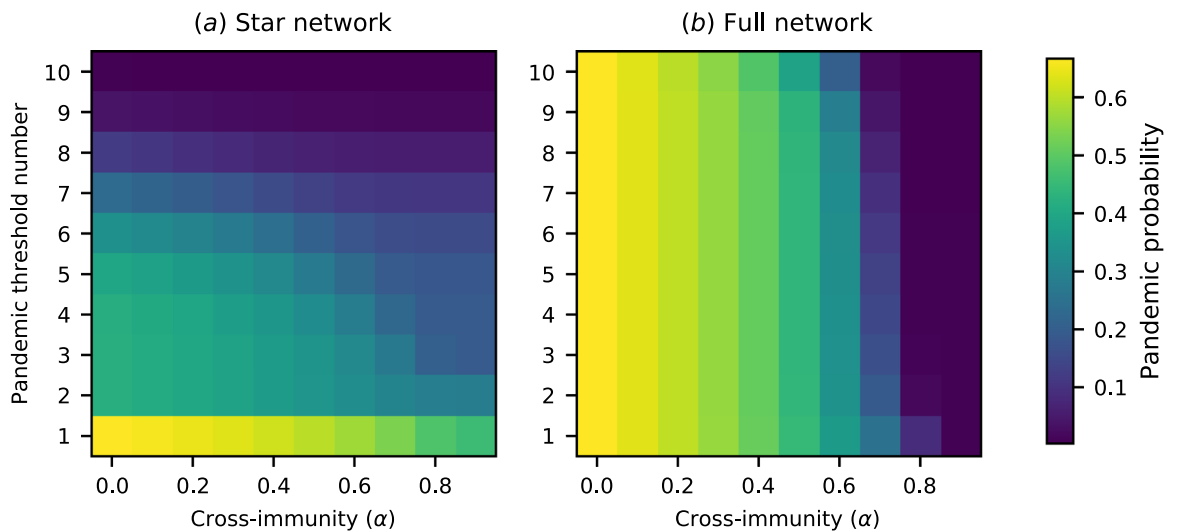

Figure 5. Pandemic probability for various levels of cross-immunity (α) and a range of transregional pandemic definitions, on a ten-region network. (**a**) Pandemic probability for a star network. (**b**) Pandemic probability for a fully connected network. Here the sensitivity of the pandemic probability to the pandemic definition used increases with cross-immunity, until the probability of any epidemic becomes very low.

Figure 5 shows the simultaneous effects of different n-region transregional pandemic definitions and the degree of cross-immunity in determining the pandemic probability. Here we fix the travel rate at 2.0×10^{-4} per day. In the full network there is a distinct transition from higher risk to lower risk, as cross-immunity approaches one. However, in the star network there is, on average, less circulation of the initial pathogen, so the effect of cross-immunity is less dramatic. Increased cross-immunity can also increase the difference in risk for different pandemic definitions—for the fully connected network, when cross-immunity exceeds $\alpha = 0.5$, differences in probability between different thresholds become visible that are much smaller at lower values. This suggests that the probability that an outbreak will develop into a pandemic may be more sensitive to the exact pandemic definition for outbreaks of pathogens that encounter pre-existing immunity than for pathogens which encounter only fully susceptible populations. However, this effect is not seen for the star network, in which the low connectivity of the network results in larger differences in probability between different thresholds even at low levels of cross-immunity.

Heterogeneous transmission. A topic of great concern during a pandemic is heterogeneity in risk between different countries or regions[50,51]. Cross-immunity can create one kind of heterogeneity, since it is common for previous exposure to a pathogen to differ between regions[52]. Another kind of heterogeneity is that due
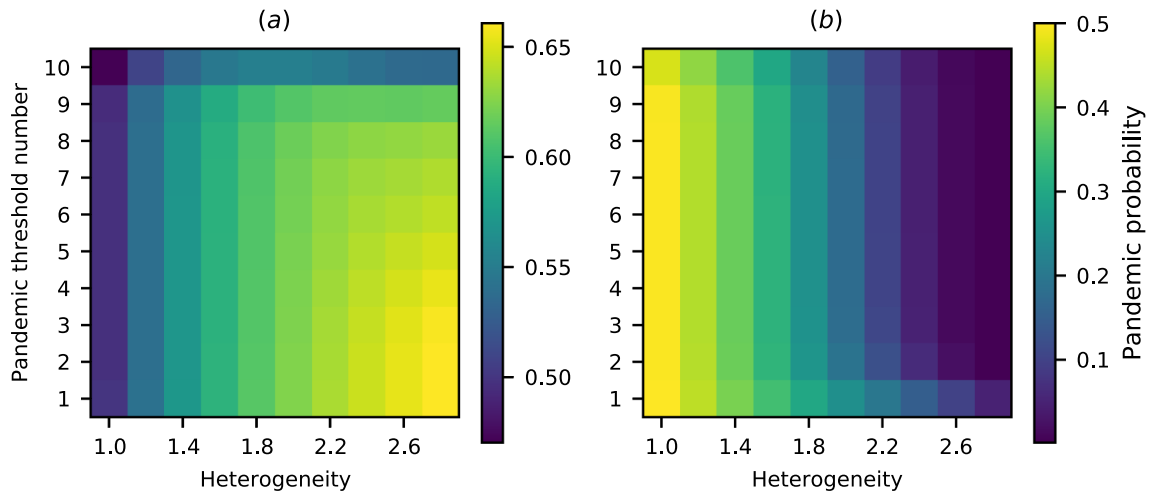
Figure 6. Pandemic probability for various degrees of heterogeneity of transmission rates and a range of transregional pandemic definitions, on a fully connected ten-region network where five regions are classed as higher-transmission and the other five regions are classed as lower-transmission. Note that the colour scales differ between the two plots, in order to make the variation in plot (**a**) clearer. (**a**) Pandemic probability for a pathogen emerging in a higher-transmission region. For low thresholds heterogeneity increases the pandemic probability, but at the 10-region threshold the pandemic probability grows and then decreases with increasing heterogeneity. (**b**) Pandemic probability for a pathogen emerging in a lower-transmission region. At all thresholds increasing heterogeneity decreases the pandemic probability.

to different public health interventions. Here we ignore cross-immunity and instead examine a heterogeneous fully connected network of ten regions, five of which have a higher rate of transmission of the pathogen than the other five. This can be thought of as an approximation to the difference between poor regions with a relative lack of public health interventions, and wealthy regions with well-funded public health organisations and increased access to healthcare.

The level of heterogeneity was defined as the ratio of the transmission rate in the higher-transmission regions to the transmission rate in the lower-transmission regions. The average transmission rate across all regions was kept fixed at $\bar{\beta} = 0.28$ per day, corresponding to a basic reproduction number of 2. The simultaneous effects of heterogeneity and the pandemic definition in determining the pandemic probability are illustrated in Fig. 6.

The row for the 1-region definition shows how the risk of any outbreak varies with the changing basic reproduction number of the pathogen in the region in which it emerges. More complex effects can be seen for higher n-region definitions, especially the 10-region definition, where, at high levels of heterogeneity, even pathogens emerging in higher-transmission regions are prevented from spreading globally due to the low chance of epidemics in lower-transmission regions. Thus the probability of a pandemic under a 10-region definition increases and then decreases with increasing heterogeneity. In the supplementary information, we show that this increasing-decreasing effect exists in networks of different sizes and structures. It appears at different thresholds in different networks. No corresponding effect exists for a pathogen emerging in a lower-transmission region, where increasing heterogeneity always decreases the chance of a pandemic, however it is defined.

Discussion

In this study, we have developed a theoretical framework to estimate the probability of a pandemic, as detailed in the Methods section below. We use a Markov chain technique based on SIR dynamics to model the spread of a pathogen. The results of this modelling framework reveal in which situations the definition of 'pandemic' has a strong effect on the calculated pandemic risk and in which situations it does not. The models also illustrate the effects of differing epidemiological parameters on the pandemic risk under different definitions, and how these effects interact with each other.

Returning to the three epidemiological questions introduced in the introduction, we can see that our results show how the answers can depend on our definition of a pandemic, and on key population and pathogen parameters. The first question was "Would interventions restricting travel reduce the risk of a pandemic?" In Fig. 2, we see that reductions in travel rates always reduce risk in a network with low connectivity, where travel occurs mainly through a central hub. However, in a highly connected network with high travel rates, travel would have to be extremely highly suppressed to change the probability of a pandemic substantially, under most definitions. This accords with previous findings regarding the effectiveness of restricting travel[53]. Additionally, in the highly connected network, changing the definition of a pandemic makes little difference to the pandemic probability, for high enough values of the travel rate.

Figure 3 further illustrates the effects of different definitions. Changing the pandemic definition can sometimes greatly alter the estimated probability of a pandemic, as seen in Fig. 3a between the yellow line, representing the 2-region transregional definition, and the purple line, representing the 10-region transregional definition. The effect on the pandemic risk of reducing travel rates also differs substantially between these two definitions. However, there are situations where changing the definition does not significantly change the pandemic

probability, as seen in the same figure between the yellow line and the dashed green line, representing the 2-region interregional definition. Both the estimated risk and the effect of reducing travel are very similar in these two cases. So, while some changes in definition do not cause a large change in quantitative analyses of the risk of a pandemic, others may significantly alter both our point estimates and the predicted effects of key parameters. Figure 3b shows that this may depend on the values of those key parameters themselves. For low travel rates, the pandemic probability is very different for the two illustrated definitions, but at high travel rates the pandemic probabilities for the two definitions converge.

The second question was "Does a portion of the population have pre-existing immunity, and does this affect the risk of a pandemic?" The presence of immunity can significantly alter the results discussed in the paragraphs above. In Fig. 5b, the leftmost column is equivalent to the column from Fig. 2b in which $\lambda = 2.0 \times 10^{-4}$ per day, but with a higher transmission rate of $\beta = 0.42$. However, as cross-immunity increases, a marked difference in the pandemic probability between different definitions becomes visible. This shows that the conclusion that precise pandemic definitions are of reduced importance in a highly connected network with high travel rates is context sensitive—if the population has high immunity, differences between definitions re-emerge.

The third question was "How is the risk of a pandemic affected by differences between regions?" In Fig. 6, we examined how heterogeneous transmission rates in different regions affect the pandemic probability. Many pathogens have higher transmission rates in lower income countries, and novel pathogens are more likely to emerge in low income countries[50,51,54]. Putting these two facts together, we see that pathogens are most likely to emerge in countries in which they have higher transmission rates. Motivated by this, we compared the scenarios of emergence in a higher-transmission and lower-transmission region, finding that pandemic definition makes a larger difference for diseases emerging in a higher-transmission region. In particular, when the pandemic definition requires many countries to experience epidemics to qualify an outbreak as a pandemic, including countries with lower transmission rates, we see striking non-linearity in the relationship between heterogeneity and the pandemic probability. For these definitions, as the difference in transmission rates between higher- and lower-transmission regions increases, the pandemic probability increases initially, before decreasing. This initial rise is due to the enhanced spread between high-transmission regions increasing the importation rate to low-transmission regions. This result implies that, when the mean value of the transmission rate is fixed, a small gap in the effectiveness of public health infrastructure between wealthy and poor regions puts all regions at greater risk, while a larger gap protects wealthier regions while the risk for poor regions continues to increase.

To illustrate this concept, consider the contrasting examples of Ebola and COVID-19. The 2014 outbreak of Ebola virus followed the pattern of high incidence in low income countries but low incidence in high income countries. The virus spread through several low-income African countries but was effectively contained when introduced to high-income countries[55–57]. In this case, high-income countries had the capacity to prevent a pandemic from taking hold, being able to quickly isolate and treat symptomatic individuals. This generated high heterogeneity in transmission, corresponding to the right side of Fig. 6a, with low-income countries at high risk and high-income countries at low risk. In contrast, high-income countries have not been able to escape the pandemic of COVID-19, in part due to asymptomatic and presymptomatic transmission of SARS-CoV-2 allowing it to evade surveillance and public health measures[58,59]. This has led to more similar transmission rates between different countries, corresponding to the left side of Fig. 6a, where risk is more uniform between regions and therefore between pandemic definitions.

In our analyses, we use a metapopulation modelling framework. Metapopulation models are widely used in pandemic modelling[40–45]. Our novel Markov chain approach allows us to calculate pandemic probabilities directly, without requiring large numbers of simulations to generate an approximation. We expect our overall conclusion, that the effects of key parameters on pandemic risk depend on the pandemic definition, to hold irrespective of the underlying modelling framework. Future studies could replicate our analyses using different models and modelling approaches, such as metapopulation models with additional epidemiological complexity[43,45,60,61] or the widely used global epidemic and mobility (GLEaM) model[62–64]. Exploring how our quantitative results vary for different modelling frameworks in the field of mathematical epidemiology[14,16,65–67] is a target for further investigation.

Other future work using our modelling framework could address the role of pandemic definitions in quantifying the effects of additional epidemiological parameters on pandemic risk, such as use of different types of travel (e.g. within-country transport or international flights)[45,68,69], the rate of nosocomial infections[70], or age structure[71]. Our metapopulation modelling framework is generally applicable, and this framework could be extended to represent outbreaks of many different specific pathogens emerging in various locations. An important factor for response planning is the timescale over which outbreaks develop into pandemics. The duration of the initial phase of outbreaks has been a subject of previous study[72], as has the overall duration of outbreaks[10,73–76]. In theory, Markov chain models could be used to assess the time for a local epidemic to develop into a pandemic, and we leave this as an avenue for further work.

In summary, we have developed a novel modelling framework for estimating the pandemic risk. We have applied this framework to assess the pandemic risk in a range of different scenarios, and have interpreted the results under a variety of pandemic definitions. We have found that certain relationships, such as the effect of heterogeneity in transmission between regions on the risk of a pandemic, are highly dependent on the definition of 'pandemic' used, while others, such as the effect of high travel rates on pandemic risk in a highly connected network, are not. This work provides a foundation for improved communication about pandemic risk, by highlighting the contexts in which pandemic definitions need to be provided in quantitative detail. In general, we contend that, when assessing the risk that an outbreak will develop into a pandemic, the precise pandemic definition used for a given analysis should be considered and stated clearly. Future work could investigate the effects of alternative definitions in more detailed epidemiological models, and extend this framework to investigate different dynamical features of pandemics.

Methods

We have combined standard epidemiological modelling techniques with a novel Markov chain treatment of metapopulation dynamics to produce a method for calculating the probabilities of epidemics and pandemics in a network of population regions. At each step of this chain, we resolve information about which regions may experience epidemics. The order in which the status of any given region is resolved does not necessarily match the order in which the given epidemics occur in calendar time. A benefit of our model is that we can calculate the probabilities of different final outcomes directly, without requiring large numbers of stochastic simulations to estimate these values. This comes at the cost that temporal information is not represented explicitly in our model: we focus on the pandemic probability, accounting for all possible ways that a pandemic could occur, rather than estimating the possible times at which epidemics could occur in different regions or the timescale over which an outbreak will develop into a pandemic (see Discussion).

We model the transmission of a pathogen through n regions labelled $P_1, P_2, P_3, \ldots, P_n$. Each region P_j has associated with it some intra-region pathogen transmissibility β_j, disease recovery rate μ_j, and population size N_j. From these quantities it is possible to calculate a region-specific basic reproduction number $R_{0,j}$. This can be fixed across all regions for a particular pathogen, or allowed to vary from region to region to reflect local epidemiological differences.

First let us consider the spread of the pathogen in a single region, using well-established results of stochastic Susceptible-Infected-Recovered (SIR) models. If a region P_j contains an initial number of infected individuals $I_j(0)$, then in the stochastic SIR model, the probability that these individuals do *not* cause an epidemic in P_j is $(1/R_{0,j})^{I_j(0)}$ when $R_{0,j} \geq 1$, and 1 otherwise[17]. We also define the final size of an epidemic $R_j(\infty)$ (not to be confused with $R_{0,j}$) as the number of recovered individuals in P_j at the end of the epidemic. This equals the total number of individuals in P_j who become infected at any time, and is given by the solution of the following equation[77].

$$R_j(\infty) = N_j - N_j \exp\left(\frac{-R_{0,j}R_j(\infty)}{N_j}\right). \tag{1}$$

Infected individuals are assumed to travel from region P_j to region P_m at a rate λ_{jm}. We seek the probability that infected individuals travelling from P_j will *not* cause an epidemic in P_m, in the case where initially infected individuals in P_m do not cause an epidemic in P_m (including the case where there are no initially infected individuals in P_m). This is equal to the probability that i infected individuals migrate from P_j to P_m, multiplied by the probability that this number of individuals fails to cause a major epidemic, summed over possible values of i. The minimum value of i is the case where no infected individuals migrate, and the maximum value is the case where all individuals in P_j that become infected at any point migrate. This gives us an expression for q_{jm}, the conditional probability that, if P_j experiences an epidemic and P_m does not experience an epidemic due to a source of infected individuals other than P_j, P_m does not experience an epidemic.

$$\begin{aligned}
q_{jm} &\approx \sum_{i=0}^{R_m(\infty)} \left(\frac{1}{R_{0,m}}\right)^i \binom{R_j(\infty)}{i} \left(\frac{\lambda_{jm}}{\lambda_{jm} + \mu_j}\right)^i \left(\frac{\mu_j}{\lambda_{jm} + \mu_j}\right)^{R_j(\infty)} \\
&= \left(\frac{\lambda_{jm}}{R_{0,m}(\lambda_{jm} + \mu_j)} + \frac{\mu_j}{\lambda_{jm} + \mu_j}\right)^{R_j(\infty)},
\end{aligned} \tag{2}$$

This approximation is valid when the number of infected individuals that travel between regions is much smaller than the size of the regions.

We assume that infected individuals travelling from a region P_j cannot cause an epidemic in a neighbouring region P_m if P_j does not itself experience an epidemic. Then computing the value of q_{jm} for every pair of populations P_j and P_m gives us sufficient information to determine the probability of any particular set of regions connected on a network experiencing epidemics so long as there are no interactions between different groups of migrants arriving in a region, and the total numbers of migrants in any region remains very small relative to the region's size. If these assumptions hold, we can imagine the regions on a network with weighted directed edges, where the weight of the edge directed from region P_j to region P_m is q_{jm}.

To determine how the final probabilities of epidemics depend on the pairwise probabilities q_{jm}, we use a Markov chain. The states of this Markov chain assign one of three states to each region—N (for neutral), where it is not yet determined whether the region will experience an epidemic, E (for epidemic), where it is determined that the region will experience an epidemic but it is not yet determined in which further regions it will cause epidemics, and T (for terminal), where it is determined that the region will experience an epidemic and in which further regions it will cause epidemics due to onward transmission. As our model does not explicitly represent dynamical processes occurring over time, these states should not be interpreted as actual states of infection and recovery within regions, but rather as bookkeeping devices for the role of various regions in determining the spread of the pathogen through the network.

Suppose we have a network connecting n regions. In the initial state, each region where the initially infected individuals have caused an epidemic is in state E, and all the other regions are in state N. The global state of the network is simply the product of the states of each system. We can then define a transition matrix $\mathbf{T}$ that acts on the global state. The elements of this matrix are denoted $t_{x_1 x_2 \ldots x_n \to y_1 y_2 \ldots y_n}$.

$$t_{x_1 x_2 \ldots x_n \rightarrow y_1 y_2 \ldots y_n}$$

$$= \left[\prod_{m=1}^{n} \left\{ \delta((x_m, y_m), (N, N)) + \delta((x_m, y_m), (N, E)) + \delta((x_m, y_m), (E, T)) + \delta((x_m, y_m), (T, T)) \right\} \right]$$

$$\cdot \left[\prod_m \left(1 - \prod_j \left(q_{jm} \right)^{\delta(x_j, E)} \right)^{\delta((x_m, y_m), (N, E))} \right] \tag{3}$$

$$\cdot \left[\prod_m \prod_j \left(q_{jm} \right)^{\delta(x_j, E) \delta((x_m, y_m), (N, N))} \right],$$

where

$$\delta((a, b, \ldots, z), (A, B, \ldots, Z)) = \begin{cases} 1 & \text{if } (a, b, \ldots, z) = (A, B, \ldots, Z) \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

x_j is the state (N, E, or T) of region P_j before the transition, and y_j is its state afterwards. The expression inside the first set of square brackets ensures that the only acceptable transitions for any given region are $N \rightarrow E$ and $E \rightarrow T$, and requires that all epidemic regions in the initial state must be terminated in the transition (this prevents double-counting of possible transmission paths). The expression inside the second set of square brackets gives the probability of each $N \rightarrow E$ transition, and the expression inside the final set of square brackets gives the probability of each $N \rightarrow N$ transition, given the regions that are in state E before the transition.

Note that these transitions do not represent a dynamical process—the order of transitions in this model does not necessarily correspond to the order in which regions experience epidemics. Instead, the transitions are simply stages along the exploration of different routes and outcomes from the disease spreading process.

The initial probability of each global state $z_1 z_2 \ldots z_n$ (where $z_i \in \{N, E, T\}$) is given by:

$$p_{z_1 z_2 \ldots z_n} = \left[\prod_j \delta(z_j, T) \right] \cdot \left[\prod_j \left(Q_j \right)^{\delta(z_j, N)} \right] \cdot \left[\prod_j \left(1 - Q_j \right)^{\delta(z_j, E)} \right], \tag{5}$$

where $Q_j = \min((1/R_{0,j})^{I_j(0)}, 1)$ is the probability that the initial population of infective individuals does not cause an epidemic in region P_j. Essentially, no region can begin in state T, and the probability of each initial global state is given by the product of the probabilities of each region being in the corresponding initial regional state.

In this system, all states in which no region is epidemic are absorbing, and in each transition at least one epidemic state must become terminal. This means that the system must reach an absorbing state in at most n transitions, since at least one region becomes terminal in each transition, and a fully terminal state is absorbing. So the final probability vector p_{final} is given by

$$p_{\text{final}} = \mathbf{T}^n p_{\text{initial}}, \tag{6}$$

with $\mathbf{T}$ as the transition matrix and p_{initial} as the vector whose elements defined by Eq. (5). This final vector gives the probabilities of each configuration of the metapopulation, with populations in state N never experiencing an epidemic, and regions in state T experiencing an epidemic at some point.

Cross-Immunity. The model described above can incorporate certain epidemiological details, such as heterogeneity of population parameters, but is restricted to treating quite simple disease dynamics. In this section we expand the model to treat pathogens that give those who overcome infection cross-protection against future strains of that pathogen. This is necessary to be able to investigate how pre-existing immunity changes how pandemic definitions affect the results of our model.

We first describe the spread of a pathogen strain X using the methods above, introducing a superscript X to the relevant parameters to mark the strain, e.g. R_0^X, $R^X(\infty)$, and p_{final}^X. We assume that infection with pathogen X confers cross-immunity α to a second strain of the pathogen, which we call Y. In each population P_j we can define an effective basic reproductive number for Y in the case that P_j has experienced an epidemic of X, which we call $R_{e,j}^Y$.

$$R_{e,j}^Y = \left(1 - \alpha \frac{R_j^X(\infty)}{N_j} \right) R_{0,j}^Y. \tag{7}$$

This expression simply multiplies the basic reproductive number by the effective number of susceptible individuals given the prevalence of cross-immunity in the population. It is through this expression that cross-immunity enters the model—the parameter α does not otherwise appear in what follows.

We can write down an equation for the expected total number of individuals in P_j infected in an epidemic of Y in analogy to Eq. (1). In the case where there has been no previous epidemic of X in P_j, the expected epidemic size is the solution $R_{j,\text{noX}}^Y(\infty)$ of

$$R_{j,\text{noX}}^Y(\infty) = N_j - N_j \exp \left(\frac{-R_{0,j}^Y R_{j,\text{noX}}^Y(\infty)}{N_j} \right). \tag{8}$$

In the case where there has been a previous epidemic of X in P_j, the expected epidemic size is the solution $R_{j,X}^Y(\infty)$ of

$$R_{j,X}^Y(\infty) = N_j - N_j \exp\left(\frac{-R_{e,j}^Y R_{j,X}^Y(\infty)}{N_j}\right). \tag{9}$$

We assume that individuals infected with Y travel at the same rate as individuals infected with X. We then define the pairwise probabilities of transmission of Y between populations in analogy to Eq. (2). That is,

$$q_{jm}^Y = \left(\frac{\lambda_{jm}}{R_{c,m}^Y(\lambda_{jm} + \mu_j^Y)} + \frac{\mu_j^Y}{\lambda_{jm} + \mu_j^Y}\right)^{R_{j,b}^Y(\infty)}, \tag{10}$$

where $R_{c,m}^Y = R_{0,m}^Y$ when P_m has not experienced a previous epidemic of X, $R_{c,m}^Y = R_{e,m}^Y$ when P_m has experienced a previous epidemic of X, $R_{j,b}^Y(\infty) = R_{j,\text{noX}}^Y(\infty)$ when P_j has not experienced a previous epidemic of X, and $R_{j,b}^Y(\infty) = R_{j,X}^Y(\infty)$ when P_j has experienced a previous epidemic of X.

These expressions for q_{jm}^Y can be substituted for q_{jm} in Eq. (3) to yield a transition matrix for modelling the spread of Y, which we will call $\mathbf{T}^Y(s_1 s_2 \ldots s_n)$, where s_j is the final state (either N or T) of the X outbreak in P_j. We find the initial probabilities of each state with regards to Y, p_{initial}^Y, in analogy to Eq. (5), given an initial number of individuals infected with Y in each population $I_j^Y(0)$.

$$p_{z_1 z_2 \ldots z_n}^Y(s_1 s_2 \ldots s_n) = \left[\prod_j \delta(z_j, T)\right] \cdot \left[\prod_j (Q_j^Y)^{\delta(z_j, N)}\right] \cdot \left[\prod_j (1 - Q_j^Y)^{\delta(z_j, E)}\right], \tag{11}$$

where $Q_j^Y = \min[(1/R_{0,j}^Y)^{I_j^Y(0)}, 1]$ when P_j has not experienced a previous epidemic of X (i.e. $s_j = N$), and $Q_j^Y = \min[(1/R_{e,j}^Y)^{I_j^Y(0)}, 1]$ when P_j has experienced a previous epidemic of X (i.e $s_j = T$). We can then write the final probabilities of each combination of possible epidemics of Y, for a given set of previous epidemics of X, as

$$p_{\text{final}}^Y(s_1 s_2 \ldots s_n) = \left(\mathbf{T}^Y(s_1 s_2 \ldots s_n)\right)^n p_{\text{initial}}^Y(s_1 s_2 \ldots s_n). \tag{12}$$

To find the overall probability of each combination of epidemics of Y in various populations given a prior probability of each combination of epidemics of X (given by $p_{\text{final}}^X(s_1 s_2 \ldots s_n)$ defined in Eq. (6)), we sum over the possible values of $(s_1 s_2 \ldots s_n)$, weighted by their probability.

$$p_{\text{total}}^Y = \sum_{(s_1 s_2 \ldots s_n) \in \{N,T\}^n} p_{\text{final}}^X(s_1 s_2 \ldots s_n) \cdot p_{\text{final}}^Y(s_1 s_2 \ldots s_n) \tag{13}$$

Code availability
Code is available on the Open Science Framework at https://osf.io/z52te/.

References
 1. Craft, M. E., Beyer, H. L. & Haydon, D. T. Estimating the probability of a major outbreak from the timing of early cases: an indeterminate problem?. *PLoS ONE* **8**, 1–7. https://doi.org/10.1371/journal.pone.0057878 (2013).
 2. Thompson, R. N., Gilligan, C. A. & Cunniffe, N. J. Detecting presymptomatic infection is necessary to forecast major epidemics in the earliest stages of infectious disease outbreaks. *PLoS Comput. Biol.* **12**, 1–18. https://doi.org/10.1371/journal.pcbi.1004836 (2016).
 3. Jamison, D. T. *et al.* (eds.) *Disease Control Priorities: Improving Health and Reducing Poverty*, vol. 9 (World Bank, Washington (DC), 2018), 3rd edn. https://doi.org/10.1016/S0140-6736(15)60097-6.
 4. Adalja, A. A., Watson, M., Toner, E. S., Cicero, A. & Inglesby, T. V. Characteristics of microbes most likely to cause pandemics and global catastrophes. *Curr. Top. Microbiol. Immunol.* https://doi.org/10.1007/82_2019_176 *(2019).*
 5. Thompson, R. N. Novel coronavirus outbreak in Wuhan, China, 2020: intense surveillance is vital for preventing sustained transmission in new locations. *J. Clin. Med.* **9**, 498. https://doi.org/10.3390/jcm9020498 (2020).
 6. Tarwater, P. M. & Martin, C. F. Effects of population density on the spread of disease. *Complexity* **6**, 29–36. https://doi.org/10.1002/cplx.10003 (2001).
 7. Miller, J. C. Spread of infectious disease through clustered populations. *J. R. Soc. Interface* **6**, 1121–34 (2009).
 8. House, T., Ross, J. V. & Sirl, D. How big is an outbreak likely to be? Methods for epidemic final-size calculation. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **469**, 20120436–20120436 (2012).
 9. Ball, F., Sirl, D. & Trapman, P. Threshold behaviour and final outcome of an epidemic on a random network with household structure. *Adv. Appl. Probability* **41**, 765–796 (2009).
10. Thompson, R., Gilligan, C. & Cunniffe, N. Will an outbreak exceed available resources for control? Estimating the risk from invading pathogens using practical definitions of a severe epidemic. *J. R. Soc. Interface* https://doi.org/10.1098/rsif.2020.0690 *(2020).*
11. Kyoon-achan, G. & Wright, L. Community-based pandemic preparedness: COVID-19 procedures of a Manitoba First Nation community. *J. Community Saf. Well-Being* **5**, 45–50 (2020).
12. Marston, H. D., Paules, C. I. & Fauci, A. S. The critical role of biomedical research in pandemic preparedness. *JAMA* **318**, 1757 (2017).

13. Monto, A. S., Comanor, L., Shay, D. K. & Thompson, W. W. Epidemiology of pandemic influenza: use of surveillance and modeling for pandemic preparedness. *J. Infect. Dis.* **194**, 92–97 (2006).

14. Lunelli, A., Pugliese, A. & Rizzo, C. Epidemic patch models applied to pandemic influenza: contact matrix, stochasticity, robustness of predictions. *Math. Biosci.* **220**, 24–33. https://doi.org/10.1016/j.mbs.2009.03.008 (2009).

15. Chowell, G., Nishiura, H. & Bettencourt, L. M. Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. *J. R. Soc. Interface* **4**, 154–166. https://doi.org/10.1098/rsif.2006.0161 (2007).

16. Colizza, V., Barrat, A., Barthelemy, M., Valleron, A. J. & Vespignani, A. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS Med.* **4**, 95–110. https://doi.org/10.1371/journal.pmed.0040013 (2007).

17. Thompson, R. N., Thompson, C. P., Pelerman, O., Gupta, S. & Obolski, U. Increased frequency of travel in the presence of cross-immunity may act to decrease the chance of a global pandemic. *Philos. Trans. R. Soc. B Biol. Sci.* https://doi.org/10.1098/rstb.2018.0274 *(2019)*.

18. Thompson, R. Pandemic potential of 2019-nCoV. *Lancet Infect. Dis.* **3099**, 30068. https://doi.org/10.1016/S1473-3099(20)30068-2 (2020).

19. Breban, R., Riou, J. & Fontanet, A. Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk. *The Lancet* **382**, 694–699. https://doi.org/10.1016/S0140-6736(13)61492-0 (2013).

20. Porta, M. *A Dictionary of Epidemiology* (Oxford University Press, USA, 2008). https://www.oxfordreference.com/view/10.1093/acref/9780195314496.001.0001/acref-9780195314496.

21. World Health Organization. What is a pandemic? (2010). https://www.who.int/csr/disease/swineflu/frequently_asked_questions/pandemic/en/.

22. Osman, A. S. HIV / AIDS in the last 10 years. *Eastern Mediterranean Health Journal* **14**, 90–96 (2007). http://www.emro.who.int/emhj-volume-14-2008/volume-14-supplement/hivaids-in-the-last-10-years.html.

23. Cohen, M. S. *et al.* The spread, treatment, and prevention of HIV-1: evolution of a global pandemic. *J. Clin. Investig.* **118**, 1244–1254. https://doi.org/10.1172/JCI34706.1244 (2008).

24. Morens, D., Folkers, G. & Fauci, A. S. What is a pandemic?. *JAMA J. Am. Med. Assoc.* **321**, 910. https://doi.org/10.1001/jama.2019.0700 (2009).

25. Doshi, P. The elusive definition of pandemic influenza. *Bull. World Health Organ.* **89**, 532–538. https://doi.org/10.2471/BLT.11.086173 (2011).

26. Coninx, J. K. *et al.* (eds.) *Pandemic influenza preparedness and response: a WHO guidance document* (WHO Press, 2009). https://www.who.int/influenza/resources/documents/pandemic_guidance_04_2009/en/.

27. Bonneux, L. & Van Damme, W. Health is more than influenza. *Bull. World Health Organ.* **89**, 539–540. https://doi.org/10.2471/BLT.11.089078 (2011).

28. Simonsen, L. *et al.* Global mortality estimates for the 2009 influenza pandemic from the GLaMOR project: a modeling study. *PLoS Med.* https://doi.org/10.1371/journal.pmed.1001558 *(2009)*.

29. Coronavirus Disease (COVID-19) Press Conference 17 February 2020 (2020). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/media-resources/press-briefings.

30. Thompson, R. N., Gilligan, C. A. & Cunniffe, N. J. Control fast or control smart: when should invading pathogens be controlled?. *PLoS Comput. Biol.* **14**, 1–21. https://doi.org/10.1371/journal.pcbi.1006014 (2018).

31. Nebehay, S. WHO says it no longer uses 'pandemic' category, but virus still emergency (2020). https://www.reuters.com/article/uk-china-health-who/who-says-it-no-longer-uses-pandemic-category-but-virus-still-emergency-idUKKCN20I0PD.

32. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020 (2020). https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.

33. Kopecki, D., Lovelace Jr., B., Feuer, W. & Higgins-Dunn, N. WHO declares the coronavirus outbreak a global pandemic (2020). https://www.cnbc.com/2020/03/11/who-declares-the-coronavirus-outbreak-a-global-pandemic.html.

34. Boseley, S. WHO declares coronavirus pandemic (2020). https://www.theguardian.com/world/2020/mar/11/who-declares-coronavirus-pandemic.

35. Wan, W. What is pandemic? Why did WHO just declare one? (2020). https://www.washingtonpost.com/health/2020/03/11/who-declares-pandemic-coronavirus-disease-covid-19/.

36. Webster, R. G. Predictions for future human influenza pandemics. *J. Infect. Dis.* **176**, S14–S19. https://doi.org/10.1086/514168 (1997).

37. Towers, S. & Feng, Z. Pandemic H1N1 influenza: predicting the course of a pandemic and assessing the efficacy of the planned vaccination programme in the United States. *Eurosurveillance* **14**, 1–3 (2009).

38. Coburn, B. J., Wagner, B. G. & Blower, S. Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1). *BMC medicine* **7** (2009). http://www.ncbi.nlm.nih.gov/pubmed/19545404.

39. Chowell, G., Bettencourt, L. M., Johnson, N., Alonso, W. J. & Viboud, C. The 1918–1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact. *Proc. R. Soc. B Biol. Sci.* **275**, 501–509. https://doi.org/10.1098/rspb.2007.1477 (2008).

40. Apolloni, A., Poletto, C., Ramasco, J. J., Jensen, P. & Colizza, V. Metapopulation epidemic models with heterogeneous mixing and travel behaviour. *Theor. Biol. Med. Model.* **11**, 1–26. https://doi.org/10.1186/1742-4682-11-3 (2014).

41. Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proc. Natl. Acad. Sci.* **101**, 15124–15129. https://doi.org/10.1073/pnas.0308344101 (2004).

42. Watts, D. J., Muhamad, R., Medina, D. C. & Dodds, P. S. Multiscale, resurgent epidemics in a hierarchical metapopulation model. *Proc. Natl. Acad. Sci.* **102**, 11157–11162. https://doi.org/10.1073/pnas.0501226102 (2005).

43. Colizza, V., Barrat, A., Barthelemy, M. & Vespignani, A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc. Natl. Acad. Sci.* **103**, 2015–2020. https://doi.org/10.1073/pnas.0510525103 (2006).

44. Ball, F. *et al.* Seven challenges for metapopulation models of epidemics, including households models. *Epidemics* **10**, 63–67 (2015).

45. Pei, S., Kandula, S., Yang, W. & Shaman, J. Forecasting the spatial transmission of influenza in the United States. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 2752–2757. https://doi.org/10.1073/pnas.1708856115 (2018).

46. Brownstein, J. S., Wolfe, C. J. & Mandl, K. D. Empirical evidence for the effect of airline travel on inter-regional influenza spread in the United States. *PLoS Med.* **3**, 1826–1835. https://doi.org/10.1371/journal.pmed.0030401 (2006).

47. Mateus, A. L., Otete, H. E., Beck, C. R., Dolan, G. P. & Nguyen-Van-Tam, J. S. Effectiveness of travel restrictions in the rapid containment of human influenza: a systematic review. *Bull. World Health Organ.* https://doi.org/10.2471/BLT.14.135590 *(2014)*.

48. Andreasen, V. Dynamics of annual influenza A epidemics with immuno-selection. *J. Math. Biol.* **46**, 504–536. https://doi.org/10.1007/s00285-002-0186-2 (2003).

49. Valkenburg, S. A. *et al.* Immunity to seasonal and pandemic influenza A viruses. *Microbes Infect.* **13**, 489–501. https://doi.org/10.1016/j.micinf.2011.01.007.Immunity (2013).

50. Walker, P. G. T. *et al.* The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* **422**, 413–422. https://doi.org/10.1126/science.abc0035 (2020).

51. Bonds, M. H., Keenan, D. C., Rohani, P. & Sachs, J. D. Poverty trap formed by the ecology of infectious diseases. *Proc. R. Soc. B Biol. Sci.* **277**, 1185–1192. https://doi.org/10.1098/rspb.2009.1778 (2010).

52. Penman, B. S., Gupta, S. & Shanks, G. D. Rapid mortality transition of Pacific Islands in the 19th century. *Epidemiol. Infect.* **145**, 1–11. https://doi.org/10.1017/S0950268816001989 (2017).
53. Hossain, M. P. *et al.* The effects of border control and quarantine measures on the spread of COVID-19. *Epidemics.* https://doi.org/10.1016/j.epidem.2020.100397 *(2020).*
54. Jones, K. E. *et al.* Global trends in emerging infectious diseases HHS Public Access. *Nature* **451**, 990–993. https://doi.org/10.1038/nature06536 (2008).
55. Merino, J. G. Response to Ebola in the US: misinformation, fear, and new opportunities. *BMJ (Online)* **349**, 1–2. https://doi.org/10.1136/bmj.g6712 (2014).
56. World Health Organization. Ebola virus disease - United Kingdom (2014). https://www.who.int/csr/don/30-december-2014-ebola/en/.
57. Gomes, M. F. C. *et al.* Assessing the international spreading risk associated with the 2014 West African ebola outbreak. *PLoS Curr. Outbreaks* https://doi.org/10.1371/currents.outbreaks.cd818f63d40e24aef769dda7df9e0da5 *(2014).*
58. Gandhi, M., Yokoe, D. S. & Havlir, D. V. Asymptomatic transmission, the achilles' heel of current strategies to control Covid-19. *N. Engl. J. Med.* **382**, 2158–2160. https://doi.org/10.1056/NEJMe2009758 (2020).
59. Lai, C. C., Shih, T. P., Ko, W. C., Tang, H. J. & Hsueh, P. R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges. *Int. J. Antimicrob. Agents* **55**, 105924. https://doi.org/10.1016/j.ijantimicag.2020.105924 (2020).
60. Meakin, S. R., Tildesley, M. J., Davis, E. & Keeling, M. J. A metapopulation model for the 2018 Ebola outbreak in Equateur province in the Democratic Republic of the Congo. *bioRxiv* 1–30 (2018). https://doi.org/10.1101/465062.
61. Baguelin, M. *et al.* Control of equine influenza: Scenario testing using a realistic metapopulation model of spread. *J. R. Soc. Interface* **7**, 67–79. https://doi.org/10.1098/rsif.2009.0030 (2009).
62. Balcan, D. *et al.* Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21484–21489. https://doi.org/10.1073/pnas.0906910106 (2009).
63. Bajardi, P. *et al.* Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PLoS ONE* https://doi.org/10.1371/journal.pone.0016591 *(2011).*
64. Gonçalves, B., Balcan, D. & Vespignani, A. Human mobility and the worldwide impact of intentional localized highly pathogenic virus release. *Sci. Rep.* **3**, 1–7. https://doi.org/10.1038/srep00810 (2013).
65. Bosch, F. V. D., Metz, J. A. J. & Zadoks, J. C. Pandemics of focal plant disease, a model. *Anal. Theor. Plant Pathol.* https://doi.org/10.1094/PHYTO.1999.89.6.495 *(1999).*
66. Chao, D. L., Halloran, M. E., Obenchain, V. J. & Longini, I. M. FluTE, a publicly available stochastic influenza epidemic simulation model. *PLoS Comput. Biol.* https://doi.org/10.1371/journal.pcbi.1000656 *(2010).*
67. Thompson, R. N. & Brooks-Pollock, E. *Detection, forecasting and control of infectious disease epidemics: modelling outbreaks in humans, animals and plants* (Trans. R. Soc. B Biol. Sci, Philos, 2019).
68. Khan, K. *et al.* Infectious disease surveillance and modelling across geographic frontiers and scientific specialties. *Lancet Infect. Dis.* **12**, 222–230. https://doi.org/10.1016/S1473-3099(11)70313-9 (2012).
69. Poletto, C., Tizzoni, M. & Colizza, V. Human mobility and time spent at destination: impact on spatial epidemic spreading. *J. Theor. Biol.* **338**, 41–58. https://doi.org/10.1016/j.jtbi.2013.08.032 (2013).
70. Barbarossa, M. V. *et al.* Transmission dynamics and final epidemic size of ebola virus disease outbreaks with varying interventions. *PLoS ONE* **10**, 1–21. https://doi.org/10.1371/journal.pone.0131398 (2015).
71. Nishiura, H., Chowell, G., Safan, M. & Castillo-Chavez, C. Pros and cons of estimating the reproduction number from early epidemic growth rate of influenza A (H1N1) 2009. *Theor. Biol. Med. Model.* **7**, 1–13. https://doi.org/10.1186/1742-4682-7-1 (2010).
72. Miller, J. C., Davoudi, B., Meza, R., Slim, A. C. & Pourbohloul, B. Epidemics with general generation interval distributions. *J. Theor. Biol.* **262**, 107–115. https://doi.org/10.1016/j.jtbi.2009.08.007 (2010).
73. van Herwaarden, O. A. & Grasman, J. Stochastic epidemics: major outbreaks and the duration of the endemic period. *J. Math. Biol.* **33**, 581–601. https://doi.org/10.1007/BF00298644 (1995).
74. Grasman, J. Stochastic epidemics: the expected duration of the endemic period in higher dimensional models. *Math. Biosci.* **152**, 13–27. https://doi.org/10.1016/S0025-5564(98)10020-2 (1998).
75. Barbour, A. D. The duration of the closed stochastic epidemic. *Biometrika* **62**, 477–482. https://doi.org/10.1093/biomet/62.2.477 (1975).
76. Allen, L. J. & Allen, E. J. A comparison of three different stochastic population models with regard to persistence time. *Theor. Popul. Biol.* **64**, 439–449. https://doi.org/10.1016/S0040-5809(03)00104-7 (2003).
77. Miller, J. C. A note on the derivation of epidemic final sizes. *Bull. Math. Biol.* **74**, 1. https://doi.org/10.1007/s11538-012-9749-6 (2012).

Acknowledgements

Author contributions

B.J.S., R.N.T., and M.B.B. conceived the study. B.J.S. carried out the analysis, wrote the manuscript, and prepared the figures. R.N.T. and M.B.B. supervised the research. All authors revised the manuscript and gave final approval for submission.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-81814-3.

Correspondence and requests for materials should be addressed to B.J.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.