Available online at www.sciencedirect.com
SCIENCE DIRECT®
Theoretical Population Biology 65 (2004) 179–191
Theoretical Population Biology
http://www.elsevier.com/locate/ytpbi

Influenza drift and epidemic size: the race between generating and escaping immunity ☆

Maciej F. Boni,[a,d,*] Julia R. Gog,[b] Viggo Andreasen,[c] and Freddy B. Christiansen[d]

[a] Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA
[b] Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK
[c] Department of Mathematics and Physics, University of Roskilde, DK-4000 Roskilde, Denmark
[d] Department of Biology, University of Aarhus Ny Munkegade, Bldg. 540, DK-8000 Aarhus C, Denmark

Received 22 May 2003

Abstract

Influenza in humans is characterised by strongly annual dynamics and antigenic evolution leading to partial escape from prior host immunity. The variability of new epidemic strains depends on the amount of virus currently circulating. In this paper, the amount of antigenic variation produced each year is dependent on the epidemic size. Our model reduces to a one-dimensional map and a full mathematical analysis is presented. This simple system suggests some basic principles which may be more generally applicable. In particular, for diseases with antigenic drift, vaccination may be doubly beneficial. Not only does it protect the population through classical herd immunity, but the overall case reduction reduces the chance of new variants being produced; hence, subsequent epidemics may be milder as a result of this positive feedback. Also, a disease with a high innate rate of antigenic variation will always be able to invade a susceptible population, whereas a disease with less potential for variation may require several introduction events to become endemic.

Keywords: Influenza; Antigenic drift; Cross-immunity; Herd immunity; Annual epidemics; Pathogen invasion; Evolutionary epidemiology

1. Introduction

Influenza A remains at high incidence levels in the human population through its ability to evade host immune surveillance by mutations that gradually changes the viral antigens (Cox and Subbarao, 2000; Fitch et al., 1991, 1997; Pease, 1987; Webster et al., 1992). The host immune response changes continuously in the sense that immunity to one strain of influenza confers a partial immunity against related strains (Gill and Murphy, 1976; de Jong et al., 2000). This cross-immunity wanes with decreasing relatedness between the immunising and the challenging strain (de Jong et al., 2000; Pease, 1987). Thus each epidemic must meet two requirements to allow for the possibility of a subsequent epidemic. First, the viral population must be large enough to produce sufficient genetic variation to allow for antigenically novel strains to appear; Pease (1987), de Jong et al. (2000), and Fitch et al. (1991) have all recognised the necessity of this to the evolution of influenza A. Second, the epidemic must avoid conferring too much immunity such that the resulting susceptible pool is too small to support subsequent epidemics of arising strains. The purpose of this paper is to formulate a model that includes the interplay between the amount of drift and the epidemic size, and to explore its behaviour.

Antibodies are formed to several antigenic sites; the most important of which are associated with a surface protein hemagglutinin (HA) which undergoes rapid positive immuno-selection (Bush et al., 1999; Fitch et al., 1997). Five key epitopes appear to allow a wide range of antigenic variation, without impairing viral functionality (Wilson and Cox, 1990). This accumulation of mutations over the years is known as influenza *drift*. In addition to drift dynamics, so-called *shifts* occur every few decades, which are more dramatic changes that are the result of reassortment events, in which new

☆ The authors contributed equally to this work.
*Corresponding author. Fax: +509-562-2296.
E-mail addresses: maciek@charles.stanford.edu (M.F. Boni), julia@zoo.cam.ac.uk (J.R. Gog), viggo@ruc.dk (V. Andreasen), freddy@birc.dk (F.B. Christiansen).

0040-5809/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.tpb.2003.10.002

sub-types can emerge to which there can be little immunity (Webby and Webster, 2001). However we shall focus on the dynamics of influenza evolution during periods of drift of a single subtype.

Pease considered the effect of drift on influenza's transmission dynamics using a model with a constant rate of drift with time. Andreasen et al., (1996) analysed a similar model that focused on the dynamics of influenza A in antigenic space. More recently, Andreasen (2003) utilised the seasonality of influenza epidemics (Cox and Subbarao, 2000), modelling the drift process by assuming that a new strain arises each year and causes an epidemic, the size of which depends on the available susceptible pool. The model presented here shares the annual structure of the Andreasen model, in that we assume a new epidemic strain at the beginning of every influenza season; in addition, we allow a two-way interaction between the epidemiological and evolutionary dynamics.

The key new element introduced in this paper is that the amount of season-to-season drift depends on the size of the previous season's epidemic. A large epidemic may cause enough genetic variation to give rise to a strain which almost completely escapes host immunity during the next season while a small epidemic will cause only minor drift in the antigenic type of next year's strain. With each infection, influenza generates immunity in the infected host; it may also mutate contributing to antigenic drift which will help it escape next year's herd immunity. This model allows us to study the balance between these two opposing forces.

There are three subsequent sections. Section 2 introduces the model, constructing the epidemic phase, the season-to-season phase, then finally interleaving these to form the full model. Section 3 presents an analysis of the model, including discussion of the implications. First, we consider steady equilibrium dynamics, where epidemics are the same size every year, then the whole map is considered and behaviour classified within parameter space; finally we consider an initial condition of an entirely susceptible population in order to mimic invasion dynamics, perhaps after a shift event. Section 4 presents a general discussion and conclusions.

2. The model

In this section, we construct the mathematical model. There are two distinct phases to the model: an epidemic phase, and a season-to-season phase. A range of different formulations can result in the same dynamical system, a point we return to below. For the description presented here, we use a history-based formulation, where partial cross-immunity acts to reduce how infectious a host is, if he ever gets infected again (Andreasen et al., 1997; Gupta et al., 1998).

The susceptibles in our host population are described by the variables q_i for $i = 0, 1, 2, 3$, etc. and q_*. These denote the proportion of the population that was last infected i years ago (q_i), or the proportion of individuals that have never been infected by influenza (q_*). The population class q_0 refers to hosts who were infected this year and are now recovered. Later, we will include basic population demographics (births and deaths).

For simplicity, we assume that there is one strain each season, but this strain changes from season to season. We assume that mutations accumulate: mutations back to previous strains are unlikely. Denote the antigenic distance mutated since the strain i years ago by d_i, so that $d_0 = 0$. Cross-immunity, τ, is a decaying function of this distance:

$$\tau(d_i) = \mathrm{e}^{-ad_i}, \quad \tau(d_*) = 0.$$

The exponential function is chosen as it has the property $\tau(a+b) = \tau(a)\tau(b)$ which will be used below. In terms of immunity, this property means that effective immunity decays at an exponential rate dependent only on antigenic distance to the last challenging strain.

The infectiousness of a host who was previously infected i years ago is reduced by a factor $(1 - \tau(d_i))$. Hence $\tau = 1$ means total immunity, and $\tau = 0$ means the host has no effective immune memory. Hosts infected during the current year, q_0, are totally immune and hosts never previously infected, q_*, are totally susceptible. We assume that individuals cannot get infected twice in one season; this assumption has no consequences since in a reduced infectivity model, an individual experiencing his second infection would not be able to infect anyone else.

2.1. The epidemic phase

In temperate regions of the northern hemisphere, influenza A is a winter disease (de Jong et al., 2000; Webster et al., 1992). During the months between November and April, an epidemic runs through susceptible host populations. To model this, we use the classical SIR model (Anderson and May, 1991), modified by cross-immunity from previous infections:

$$\dot{q}_i = -\beta q_i \left(p_* + \sum_{j=0}^{\infty} (1 - \tau(d_j)) p_j \right) \quad \text{for } i = *, 1, 2, \ldots,$$

$$\dot{p}_i = +\beta q_i \left(p_* + \sum_{j=0}^{\infty} (1 - \tau(d_j)) p_j \right) - \nu p_i \quad \text{for } i = *, 1, 2, \ldots,$$

$$\dot{q}_0 = \sum_{i=1}^{\infty} \nu p_i + \nu p_*,$$

where p_i is the proportion who are currently infected, but were previously infected i years ago. The dynamic variable p_* represents infected individuals who have never previously been infected, while the variable $p_0 = 0$ since individuals cannot be reinfected once they enter the q_0-class. These equations can be much simplified by using the variables I and S to denote *effective* infectious and susceptible proportions, where proportions are weighted by their infectiousness:

$$S = q_* + \sum_{j=0}^{\infty} (1 - \tau(d_j))q_j, \tag{1}$$

$$I = p_* + \sum_{j=0}^{\infty} (1 - \tau(d_j))p_j. \tag{2}$$

These sums could equally start at $j = 1$, since the p_0 class is empty and since $\tau(d_0) = 1$. Using these variables, the system reduces to the classical SIR system,

$$\begin{aligned}\dot{S} &= \dot{q}_* + \sum_{j=0}^{\infty} (1 - \tau(d_j))\dot{q}_j \\ &= -\beta\left(q_* + \sum_{j=0}^{\infty} (1 - \tau(d_j))q_j\right) \\ &\quad \times \left(p_* + \sum_{k=0}^{\infty} (1 - \tau(d_k))p_k\right) \\ &= -\beta SI,\end{aligned}$$

$$\begin{aligned}\dot{I} &= \dot{p}_* + \sum_{j=0}^{\infty} (1 - \tau(d_j))\dot{p}_j \\ &= +\beta\left(q_* + \sum_{j=0}^{\infty} (1 - \tau(d_j))q_j\right) \\ &\quad \times \left(p_* + \sum_{k=0}^{\infty} (1 - \tau(d_k))p_k\right) \\ &\quad - \nu\left(p_* + \sum_{j=0}^{\infty} (1 - \tau(d_j))p_j\right) \\ &= +\beta SI - \nu I,\end{aligned}$$

where differentiation of infinite sums is valid as long as the partial sums and their derivatives converge uniformly. These equations describe a simple epidemic, and the final values of the system may be easily found:

$$\frac{dI}{dS} = -1 + \frac{1}{rS},$$
$$I = -S + r^{-1}\log S + \text{constant},$$

where $r = \beta/\nu$ is the basic reproductive ratio of influenza: the average number of secondary cases a single case would generate in a totally susceptible population. If the epidemic starts and ends with a very small number of infections, we have

$$-S_{initial} + r^{-1}\log S_{initial} = -S_{final} + r^{-1}\log S_{final}.$$

Define y to be the severity of the epidemic—in that it describes the proportion of individuals who were not infected—as follows:

$$y = \frac{S_{final}}{S_{initial}}.$$

The smaller y is, the more severe the impact of the epidemic. We have

$$-S_{initial} + r^{-1}\log S_{initial} = -yS_{initial} + r^{-1}\log(yS_{initial}),$$

$$0 = rS_{initial}(1 - y) + \log y.$$

The trivial solution, $y = 1$, corresponds to there being no epidemic. We seek a non-trivial solution and conditions for its existence. Dividing by $(1 - y)$, we define $\phi(y)$ as follows:

$$rS_{initial} = \frac{-\log y}{1 - y} = \phi(y). \tag{3}$$

It is then possible to show

- $\phi'(y) < 0$,
- $\phi(1) = 1$,
- $\lim_{y\to 0}\phi(y) = +\infty$.

This can be demonstrated by application of the simple inequalities

$$\frac{y - 1}{y} \leqslant \log y \leqslant y - 1, \tag{4}$$

which hold for all $y > 0$, with each equality if and only if $y = 1$.

Then, every value $rS \geqslant 1$ corresponds to a unique $y \in (0, 1]$. Given a value for $S_{initial} \geqslant r^{-1}$, the implicit equation (3) gives a value for y which in turn gives S_{final}. See Fig. 1. For the epidemic to take off, we must have $rS_{initial} \geqslant 1$. If $S_{initial} < 1/r$ then there are not enough susceptibles, and we deem this to be a stopping point of the model.

We could also calculate final values for the q_i using y, but as we discover below, it is not necessary to do this explicitly.

2.2. Season-to-season phase

In this phase, we calculate the next season's initial values from the previous season's final values. The q_i are all shifted by one year, a new strain is identified, and host births and deaths are taken into account.

The new strain is some distance d away from the previous epidemic strain, which then generates all the new d_i. The distance mutated depends on the number and severity of infectious cases in the previous epidemic; we call this I_{total}. The higher the severity of the epidemic,

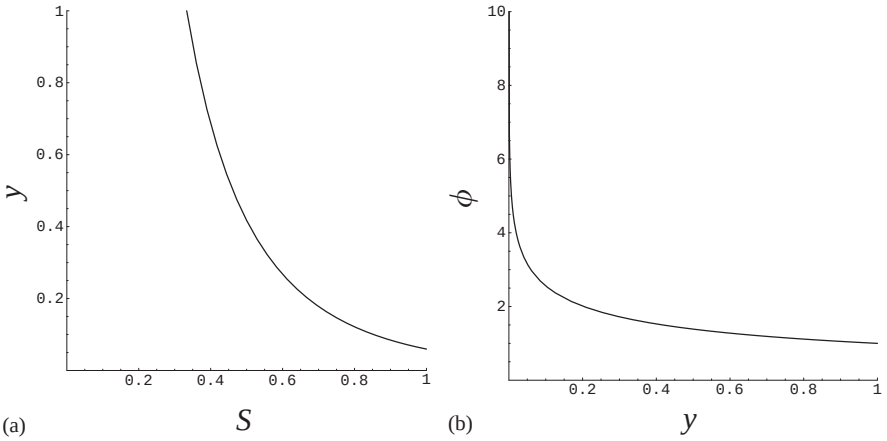

Fig. 1. Epidemic impact as a function of initial susceptibles and $\phi(y)$. Here, $r = 3$. In panel (a), when $S \geqslant 1/r$, y is given by the solution of (7). When $S < 1/r$, there is no epidemic. The larger S is, the smaller y is, and hence the more severe the epidemic. Panel (b) shows $\phi(y)$ which decreases monotonically from $+\infty$ to 1 on the interval $[0, 1]$.

the more likely a distant strain is to arise. This implicitly assumes that there is some selection process at work—a strain which has more potential to infect the population is more likely to be the new epidemic strain than one to which the population is largely immune. A wider choice of variants results in a higher expected amount of change. As a simple model, we take

$$d = \delta + bI_{total},$$

where b is a constant, and δ is a minimal amount of mutation. This minimal mutation can account for mutation in a chain of infection persisting over the summer months, for example in tropical or sub-tropical regions where influenza persists year round (Webster et al., 1992); note that δ can be zero. Later we will see that this parameter is not important on its own, as it becomes compounded with population births and deaths.

Using hats to denote values at the end of the previous season, we have

$$d_i = \hat{d}_{i-1} + d \quad \text{for } i = 1, 2, 3, \ldots, \quad (d_0 = 0).$$

Births and deaths are modelled between seasons as a proportion μ of the population dying each year, and the same number being born. Some of these might occur during the epidemic, but supposing this proportion is small and does not much affect the epidemic, then to a good approximation we can deal with births and deaths between seasons:

$$\begin{aligned} q_0 &= 0, \\ q_i &= \hat{q}_{i-1}(1-\mu) \quad \text{for } i = 1, 2, 3, \ldots, \\ q_* &= \hat{q}_*(1-\mu) + \mu. \end{aligned}$$

Considering the effective susceptible population:

$$\begin{aligned} S &= q_* + \sum_{j=0}^{\infty} (1 - \tau(d_j))q_j \\ &= \hat{q}_*(1-\mu) + \mu + \sum_{j=1}^{\infty} (1 - \tau(\hat{d}_{j-1} + d))\hat{q}_{j-1}(1-\mu) \\ &= \mu + (1-\mu)\left(\hat{q}_* + \sum_{j=0}^{\infty} (1 - \tau(\hat{d}_j)\tau(d))\hat{q}_j\right) \\ &= \mu + (1-\mu)\left(\hat{q}_* + \sum_{j=0}^{\infty} \hat{q}_j - \tau(d) \sum_{j=0}^{\infty} \tau(\hat{d}_j)\hat{q}_j\right), \end{aligned}$$

but at the end of the previous epidemic

$$\hat{q}_* + \sum_{j=0}^{\infty} \hat{q}_j = 1$$

and

$$\sum_{j=0}^{\infty} \tau(\hat{d}_j)\hat{q}_j = 1 - \hat{S} \quad \text{by (1)}.$$

Simplifying, we find S does not depend on any q_i:

$$\begin{aligned} S &= \mu + (1-\mu)(1 - \tau(d)(1 - \hat{S})) \\ &= 1 - (1-\mu)\tau(d)(1 - \hat{S}) \\ &= 1 - (1-\mu)\tau(\delta + bI_{total})(1 - \hat{S}) \end{aligned} \tag{5}$$

and we notice that the quantity $1 - S$, which represents something related to the immunity in our population, is reduced by a factor $(1-\mu)\tau(\delta + bI_{total})$ between seasons.

2.3. The full model

The final step is to combine the epidemic dynamics with the season-to-season dynamics to form a map $f(S)$ for the whole year giving the initial S value for the next

year. The function f will map the effective susceptible population S at the beginning of this season to S at the beginning of next season, yielding a one-dimensional map of the form $S_{n+1} = f(S_n)$. We already have S_{n+1} as a function of $\hat{S}$ from Eq. (5), and the epidemic section shows how $\hat{S}$ (in that section S_{final}) depends on S_n. The necessary relationships to construct f are:

$$\hat{S} = yS,$$
$$I_{total} = (1-y)S,$$

where y is the solution of

$$rS = \frac{-\log y}{1-y}.$$

Hence, combining these with (5):

$$\begin{aligned} f(S) &= 1-(1-\mu)\tau(\delta + b(1-y)S)(1-yS) \\ &= 1-(1-\mu)e^{-a\delta}e^{-ab(1-y)S}(1-yS) \\ &= 1-(1-\mu)e^{-a\delta}e^{+abr^{-1}\log y}(1-yS) \\ &= 1-\kappa y^{\lambda}(1-yS), \end{aligned}$$

where

$$\kappa = (1-\mu)e^{-a\delta},$$
$$\lambda = \frac{ab}{r}.$$

In summary, the whole year to year dynamics can be described by a one-dimensional map, with three parameters: r, κ, and λ:

$$f(S) = \begin{cases} 1-\kappa y^{\lambda}(1-yS), & S \geqslant 1/r, \\ \textit{model stops}, & S < 1/r, \end{cases} \tag{6}$$

where y solves

$$rS = \frac{-\log y}{1-y}. \tag{7}$$

The basic reproduction ratio of influenza in this population is r, and we assume $r > 1$. κ represents the loss of immune memory of the population, both through background antigenic drift and population turnover. We have $\kappa \in (0,1)$; the smaller κ is, the faster the population loses immunity.

λ is the hardest to interpret, but is the most interesting parameter to emerge as it concerns the trade-off between escaping and generating immunity. It is formed of three simpler parameters used in model formulation. b is related to how much a single viral replication would add to antigenic drift, and a is the rate at which that drift decays immunity, so it is natural these two only occur together. The product ab is mutability—how easily immune escape variants can arise. On the other hand, r is related to the number of secondary cases a single infection generates—the number of people a single case removes from the susceptible population. λ is the ratio of these numbers and thus parameterises the balance between finding new susceptibles (by immune escape) and removing them (through infection and conferred immunity). A high value of λ means that it is very easy for the virus to escape its own immunity; a low value of λ means that the virus generates too much immunity to be successful. Despite λ being most natural to work with in terms of algebra, mutability, $ab = \lambda r$, is more biologically intuitive and is used again later in this paper, denoted by m.

This derivation of the model followed the path of using a history-based formulation and reduced transmission. In fact, it is possible to derive an identical system by using a status-based model with cross-immunity acting to either reduce susceptibility to influenza, or transmissibility once infected (Gog and Swinton, 2002; Gog and Grenfell, 2002). The range of possible models for cross-immunity has been much studied (Dawes and Gog, 2002; Ferguson and Andreasen, 2001). As there is only one strain circulating at a given time, different choices of model formulation are unlikely to have any qualitative impact on the dynamics. However, the particular assumptions used here allow the model to be collapsed *exactly* to the one-dimensional system. The key steps required are: first, that the dynamics of the effective susceptibles and infecteds (S and I here) can be described without recourse to the more detailed class variables (q_i and p_i), and second, that immunity changes between seasons in such a way that, again, the detailed population classes are not explicitly needed. This second step is met by having immunity that decays exponentially with antigenic distance and an infinite number of population classes.

3. Endemic and invasion dynamics

3.1. Unique and stable equilibrium

In this section, we consider the basic properties of the map, and show that there is a unique fixed point. The map $f(S)$ is a continuous map from the interval $[1/r, 1]$ to $[0,1]$ which we interpret as a stopping point if $f(S) < 1/r$. First looking at the end points of the domain, it is easy to see that $S = 1/r$ corresponds to $y = 1$ and then

$$f(1/r) = 1-\kappa(1-1/r) > 1/r.$$

Also, $S = 1$ corresponds to a non-zero y, and so $f(1) \neq 1$, hence

$$f(1) < 1$$

and there is at least one fixed point in this region. To establish its uniqueness and stability, consider the

derivative of f:

$$f(S) = 1 - \kappa y^{\lambda}(1 - yS),$$

$$f'(S) = -\kappa\lambda y^{\lambda-1}(1 - yS)\frac{\partial y}{\partial S} + \kappa y^{\lambda} S\frac{\partial y}{\partial S} + \kappa y^{\lambda} y,$$

$$\frac{1}{\kappa y^{\lambda}} f'(S) = y + (S - \lambda y^{-1}(1 - Sy))\frac{\partial y}{\partial S}.$$

From (7) we have

$$\frac{\partial y}{\partial S} = \frac{-ry(1-y)}{1 - rSy}$$

and using (7) and the basic inequalities (4):

$$rSy \leqslant 1.$$

Then,

$$\frac{1}{\kappa y^{\lambda}} f'(S) = y - (S - \lambda y^{-1}(1 - Sy))\frac{ry(1-y)}{1 - rSy} \tag{8}$$

$$= \frac{y - rSy}{1 - rSy} + \lambda\frac{r(1 - Sy)(1 - y)}{1 - rSy} \tag{9}$$

$$\geqslant \frac{y - rSy}{1 - rSy} \quad \text{as } \lambda > 0 \text{ and } (1 - rSy) \geqslant 0 \tag{10}$$

$$= \frac{y(1-y) + y\log y}{1 - y + y\log y} \quad \text{by (7).} \tag{11}$$

Using the basic inequalities (4), one can then show that this function of y decreases from 0 to -1 on $y \in [0, 1]$; see Appendix A. Hence

$$\frac{1}{\kappa y^{\lambda}} f'(S) \geqslant -1,$$

$$f'(S) \geqslant -\kappa y^{\lambda} \geqslant -\kappa$$

and for all S we have

$$f'(S) > -1. \tag{12}$$

Hence the gradient can be bounded above minus one for any given parameter values. This is true for the whole range of S, and not just at the fixed point, and this will be useful again later.

Now, to find an upper bound for the gradient and establish stability, we will need to use that we are considering a fixed point. The calculations are slightly tedious and are relegated to Appendix B. For any S^* a fixed point of f, we have

$$f'(S^*) < 1.$$

This establishes uniqueness. The function starts the range above the diagonal, and ends below. The upper bound on its derivative gives that it cannot cross the diagonal from below, so there is a unique fixed point. The derivative at this fixed point is between minus and plus one, and hence it is stable.

It is straightforward to locate this equilibrium numerically. Fig. 2(a) shows S^* against ranges of r and m (mutability, λr). For high values of m, the equilibrium susceptible proportion is not monotonic in r: for low values of r few are infected, and for high values of r many new variants are generated, both extremes resulting in a large effectively susceptible population.

Fig. 2(b) shows I_{total}, the proportion of hosts affected by the epidemic, perhaps the most easily observable quantity. The epidemic size increases with both m and r. The interesting feature here is the rapid change from a small number of hosts affected by each epidemic, to the plateau where a substantial proportion are affected each year. Considering a fixed value of m and varying r, there is a critical range of r in which incidence increases rapidly. This range can be seen in the figure to be distinct from the usual $r = 1$ threshold. We interpret this as the point at which the virus has the capacity to escape

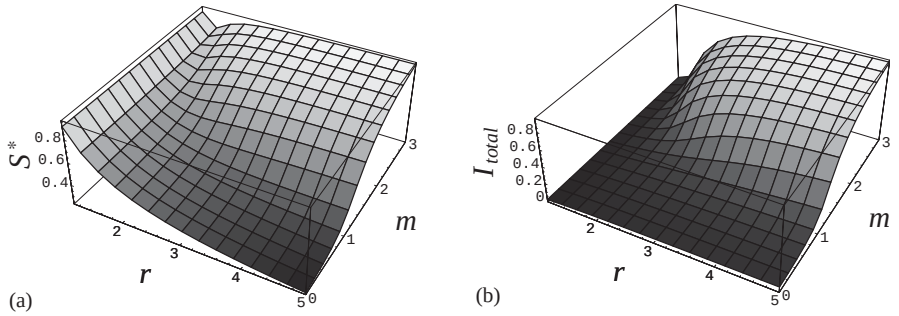

Fig. 2. Equilibrium values. (a) numerical evaluation of the equilibrium susceptible size (S^*). (b) shows the total proportion infected over the course of the epidemic (I_{total}). This is monotonic increasing in both r and m. The key feature is that there is a rapid change from a small number of hosts affected by the epidemic, to the plateau where a substantial proportion are affected each year. As mutation rate is increased, there is a critical range in which there is positive feedback. Large mutation means a larger drift rate, higher susceptibility next year, so more infections and variants produced, hence more drift. Where the graph ramps up steeply, influenza gains the ability to escape its own tail, and runaway drift occurs. This same effect is seen for increasing r, as is discussed in the text.

prior immunity: large enough transmission means many cases, this leads to higher production of antigenic variants, meaning a higher drift rate, more effective susceptibles next year and hence more cases. In this range of r, positive feedback enables runaway drift.

Conversely, in the same parameter range, vaccination can be disproportionately effective as it reduces the effective value of r. The interpretation is simple: effects of vaccination of an antigenically variable pathogen are enhanced beyond classical herd immunity—not only does vaccination protect the individual and any they would infect, it also reduces the chances of new variants appearing. Though perhaps an obvious point to emerge, we expect this principle may hold in more complex models. The magnitude of the effect, and the critical range of r will be sensitive to model assumptions.

3.2. Full dynamics

The following section develops the equilibrium results to describe the full dynamics of the map. Parameter space will be partitioned into two regions according to the sign of the initial gradient of the map. One region is then further split into four cases according to whether the map includes a region in which epidemics stop, and whether the map has a turning point. Overall, either the stable equilibrium is globally attracting, or some initial conditions will lead to epidemics of decreasing size that will eventually fail to take off, and all other initial conditions are attracted to the fixed point.

We divide parameter space into two regions according to whether the initial gradient of the map is positive or negative. We have from the previous section

$$\frac{1}{\kappa y^{\lambda}} f'(S) = y + (S - \lambda y^{-1}(1 - Sy)) \frac{\partial y}{\partial S}.$$

And since $S = 1/r$ corresponds to $y = 1$,

$$\frac{\partial y}{\partial S} = \frac{-ry(1-y)}{1 - rSy} = \frac{-ry(1-y)^2}{(1-y) + y\log y},$$

$$\lim_{S \to 1/r} \frac{\partial y}{\partial S} = -2r$$

by l'Hôpital's rule. Hence

$$\begin{aligned}\frac{1}{\kappa} f'(1/r) &= 1 + (1/r - \lambda(1 - 1/r))(-2r) \\ &= 2\lambda(r-1) - 1\end{aligned}$$

giving

$$\lambda \geqslant \frac{1}{2(r-1)} \;\Leftrightarrow\; m \geqslant \frac{1}{2 - 2r^{-1}} \;\Leftrightarrow\; f'(1/r) \geqslant 0. \qquad (13)$$

This is the condition that corresponds to whether the initial gradient is positive or negative. The same bound on λ is also important for the existence of turning points of the map. From (9),

$$\frac{1}{\kappa y^{\lambda}} f'(S) = \frac{y - rSy}{1 - rSy} + \lambda \frac{r(1 - Sy)(1 - y)}{1 - rSy},$$

so at a turning point $f'(S) = 0$, and we obtain

$$\lambda = \frac{-(y - rSy)}{(r - rSy)(1 - y)} = \frac{y(-(1-y) - \log y)}{(1-y)(r(1-y) + y\log y)} \qquad (14)$$

using (7) to eliminate S. For fixed r, this is $\lambda(y)$, giving a potential location of a turning point (but only potential, as the y value may not correspond to $S \leqslant 1$). See Fig. 3. At the edges of possible y:

$$\lambda(0) = 0, \quad \lambda(1) = \frac{1}{2(r-1)}$$

using l'Hôpital's rule for the latter. It is possible to show that $\lambda(y)$ is an increasing function, as outlined in Appendix C.

So for fixed r, for every $\lambda \leqslant 1/(2(r-1))$ there is a unique y which could give a turning point of the map. If this corresponds to a feasible S, this turning point exists and is unique. Recall from above that y is a decreasing function of S (see Fig. 1), decreasing from 1 down to some minimum value at $S = 1$. Hence there is a minimum λ that gives a valid turning point. Below this, $f(S)$ is monotonic decreasing. If $\lambda > 1/(2(r-1))$ then our map is monotonic increasing.

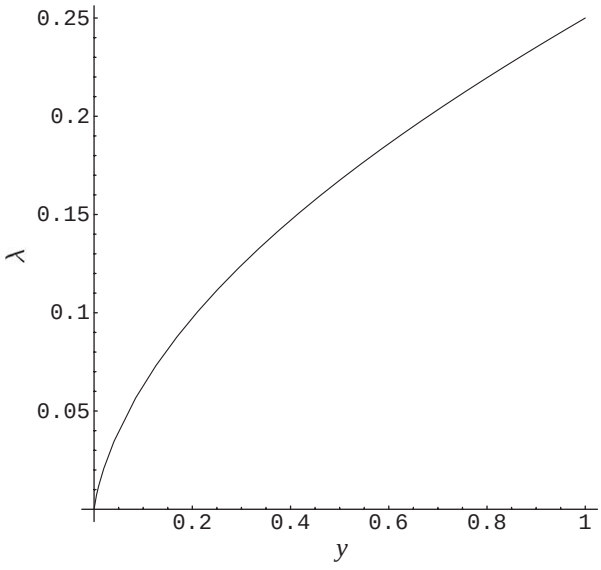

Fig. 3. Finding turning points of the map ($r = 3$). Setting the derivative of the map to be zero gives λ as a function of y for fixed r, see Eq. (14). This allows exploration of possible turning points. $\lambda(y)$ is monotonic, growing from $\lambda = 0$ to $1/(2(r-1))$. Only y larger than a certain value will correspond to $S \leqslant 1$ for each r. Here, $y < 0.06$ yields S-values larger than unity, hence, low enough λ-values will yield a map with no turning point.

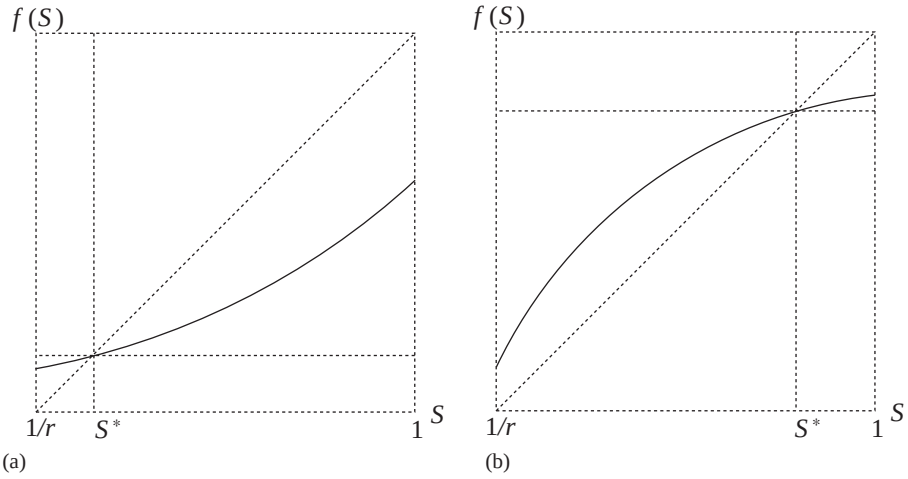

Fig. 4. High λ. This is a schematic diagram of the map. For $\lambda > 1/(2(r-1))$, the map is monotonic increasing, and the unique stable fixed point (S^*) is globally attracting. (a) λ just large enough to make f monotonically increasing. Equilibrium sits close to $1/r$ and corresponds to repeated epidemics but of a much smaller size than the population could allow for. (b) Runaway drift. λ large enough to make f concave as well as monotonically increasing. Equilibrium sits close to 1 and corresponds to large repeated epidemics. A high mutation rate allows the virus to escape prior immunity and infect a large number of susceptibles each season.

3.2.1. High λ

We have divided up parameter space into two regions, according to the size of λ. From the model formulation, high λ corresponds to a pathogen that can mutate easily enough to escape the immunity that it generates. Mathematically, for $\lambda > 1/(2(r-1))$ we have the unique fixed point S^*, and $f'(S) > 0$ so:

$$S < f(S) < S^* \quad \text{for } S < S^*,$$
$$S^* < f(S) < S \quad \text{for } S > S^*,$$

see Fig. 4. So all solutions increase or decrease, converging to S^*. The dynamics are very simple—the fixed point is globally attracting.

3.2.2. Low λ

In this region, the gradient of the map is initially negative. There may be a unique turning point; otherwise the map is monotonic decreasing. In this region, there is also the possibility that the map dips below the halting level. With these two properties, we divide this region up into four cases, see Fig. 5.

Case (a): $f(S)$ is monotonic decreasing but not so far that it reaches $f(S) = 1/r$. Combining with (12), we can bound the derivative throughout:

$$-1 < f'(S) < 0,$$

$$|f(S)| < 1$$

and hence the map is contracting. The unique fixed point is globally stable.

Case (b): Here, there is a unique turning point in f, and $f(S) \geqslant 1/r$ throughout. The equilibrium in this case is believed to be globally stable, as it appears there is always a region $[1/r, S_\varepsilon]$ containing S^* where $|f'| < 1$. S_ε is the point at which $f'(S_\varepsilon) = 1 - \varepsilon$.

Case (c): $f(S)$ is monotonic decreasing and does indeed reach $f(a_0) = 1/r$ at some a_0. With a_0, we split $[1/r, 1]$ into two regions, P_0 and H_0 (P will stand for persistence, and H for hole). See Fig. 6.

$$f(S) > 1/r \;\Leftrightarrow\; S \in [1/r, a_0] = P_0,$$
$$f(S) < 1/r \;\Leftrightarrow\; S \in (a_0, 1] = H_0.$$

For $S \in H_0$, the resulting number of susceptibles the next season is too low (too much immunity from the previous epidemic) and the model is at a stopping point. For $S \in P_0$, we see by the monotonic decreasing property that

$$f(P_0) = [f(a_0), f(1/r)] = [1/r, f(1/r)],$$

but

$$f'(S) > -1 \;\Rightarrow\; \frac{f(a_0) - f(1/r)}{a_0 - 1/r} > -1 \;\Rightarrow\; f(1/r) < a_0$$

by the mean value theorem. Hence,

$$f(P_0) \subset P_0.$$

In P_0 the dynamics are as in case (a), where $|f(S)| < 1$, and so all solutions that start in P_0 will remain in P_0 and are attracted to the unique fixed point S^*.

Case (d): This is the most interesting case. Here, there is a turning point, and also $f(S) < 1/r$ somewhere. If the map does not climb back above $1/r$ after the turning point, then it is exactly as case (c).

If however, there is a second point b_0 with $f(b_0) = 1/r$, then P_0 is defined as before: the leftmost, non-stopping region, which will contain the fixed point—$[1/r, a_0]$. But now, H_0, the first "hole," is (a_0, b_0) and does not extend to the right edge of the interval. Instead, there is a further interval, in which the map has positive derivative, lies above $1/r$, but lies

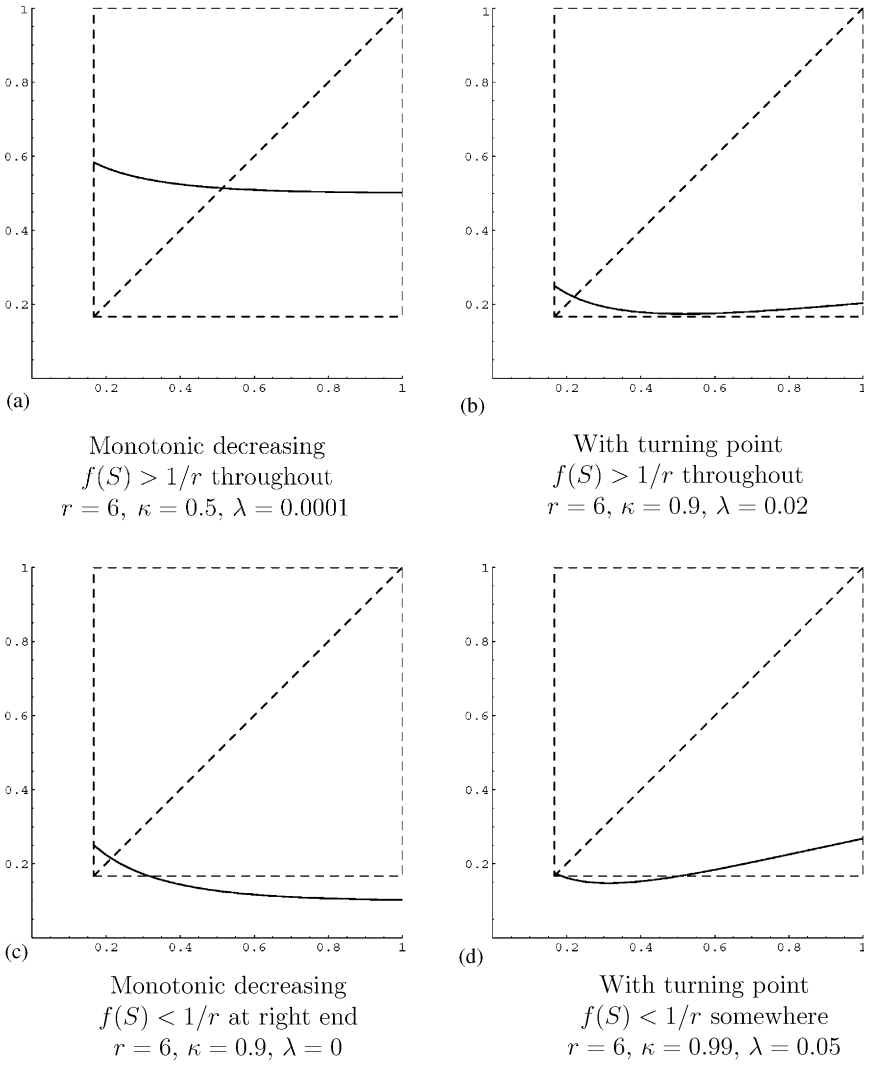

Fig. 5. For $\lambda < 1/(2(r-1))$ the gradient is initially negative. There are then four cases according to whether the map decreases enough to have a stopping point somewhere (bottom row), and whether the map has a turning point (right column).

below the diagonal. The first part of this new interval maps into P_0; call this part P_1. If $f(1) < a_0$, then P_1 extends the whole way to 1 and the whole domain is accounted for. If not, then the right edge of P_1 is a point a_1 that maps to a_0: $f(a_1) = a_0$. Then there is a further interval, H_1, which maps into H_0. Again, this might account for the whole interval, otherwise there may be more intervals in the order: P_2, H_2, P_3, H_3, and so on. P_n maps into P_{n-1}, and H_n maps into H_{n-1}. H_n is the set of starting S values for which disease will fail to take off in n seasons time. P_n is the set of values for which the nth iterate will be in P_0, where the map will settle to the fixed point.

We define the sequences a_n and b_n to be the preimages of a_0 and b_0 lying further and further to the right, as far as these exist. H_n lies between a_n and b_n, P_n lies between b_{n-1} and a_n. See Fig. 7.

If we write $H_1 \to H_0$ to mean that $f(H_1) \subset H_0$, then the dynamics can be summarised as follows:

$$H_n \to H_{n-1} \to \cdots \to H_1 \to H_0 \to \text{STOP},$$
$$P_m \to P_{m-1} \to \cdots \to P_1 \to P_0 \to \text{fixed point } S^*,$$

where $m = n$ or $m = n+1$ depending on whether the domain ends in a P or an H interval.

There are two different behaviours according to initial conditions. If the initial conditions are in a "hole"—i.e. an H_n-interval—then the epidemic will decrease in size each year, then finally fail to take off. All other initial conditions settle down to the fixed point.

3.3. Invasion

This section considers the specific case of the initial condition $S = 1$, corresponding to the introduction of influenza to a naïve population. This describes the scenario after a substantial antigenic *shift* event or the scenario where influenza is introduced to a population that was previously influenza-free.

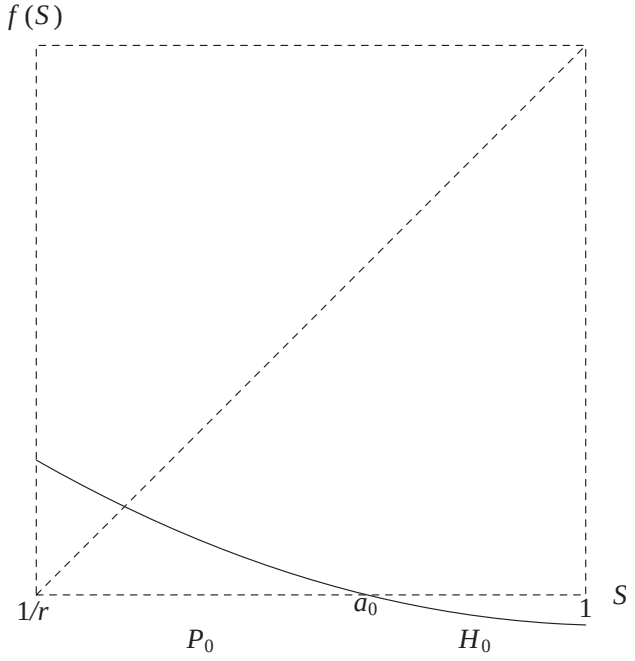

Fig. 6. Low λ, case (c). This is a schematic depiction of the location of the intervals P_0 and H_0. In this case, the map is monotonic decreasing, higher values of S leading to stopping points. $f(a_0) = 1/r$, and $[1/r, a_0] = P_0$, $(a_0, 1] = H_0$. Initial conditions in P_0 are attracted to the fixed point, and initial conditions in H_0 stop after one season.

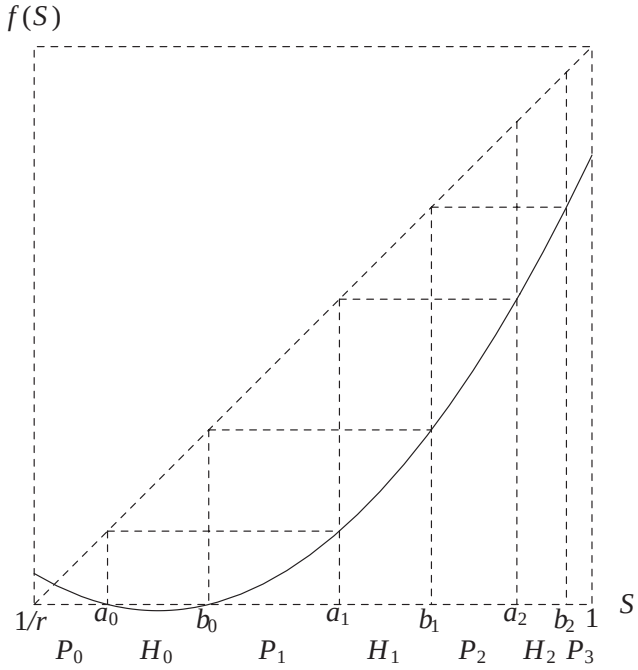

Fig. 7. Low λ, case (d). This is a schematic figure. In the first interval (P_0), the fixed point is attracting. Immediately to the right is the first "hole" (H_0) where influenza has generated too much immunity or not enough variation and does not reestablish itself next season. Further right, P_1 maps entirely to P_0 and so is part of the basin of attraction for the fixed point. H_1 maps into H_0 so starting in H_1 means two seasons of flu before it fails. Continuing to the right, the intervals alternate between points attracted to the fixed point, and those ultimately mapped into the hole.

The previous section characterised the dynamics of the whole domain of S for each case in parameter space. The identification of different intervals allows an easy description here: either $S = 1$ is in the basin of attraction of the fixed point (because the fixed point is globally stable, or $S = 1$ lies in a P interval) or $S = 1$ is in a hole H_n, where the epidemics will eventually cease.

Figs. 8(a) and (b) show the outcome of introduction of infection to a susceptible population for a range of λ or r, for two different values of κ. For high λ or high r, the epidemics persist indefinitely, as they settle towards the equilibrium. If both λ and r are low, then there is no second epidemic—the infection is both too feeble, and unable to mutate sufficiently. If κ is close to one, then there may be further regions, where epidemics happen for more than one year, but ultimately die out.

The boundaries between different regions in Fig. 8 have a hyperbolic-like shape in r and λ, so it is natural to consider how these different regions might appear in terms of m and r (Figs. 8(c) and (d)). Apart from very low values of r, the result of invasion in a susceptible population is much more sensitive to m than it is to r. A highly mutable virus will always persist, even for low values of r. Pathogens with large enough values of r will also persist even for low values of m—however, the required r-values may be too large to be realistic.

In the several *shift* events of the twentieth century where a new subtype appeared, that subtype initially caused annual epidemics, generally of decreasing size, and subsequently established itself in the human population until a new shift event occurred and replaced the prevailing subtype. In 1977, the H1N1 subtype was reintroduced into the population without replacing H3N2; these two subtypes have been co-circulating since (Earn et al., 2002; Webster et al., 1992). Mortality data for the US consolidated by Simonsen et al. (1998) and global epidemiological analyses (Cox et al., 1994; Cox and Subbarao, 2000) of the shift events and their corresponding pandemics help us understand the pattern that a new subtype follows after it is introduced. In the present model, the pattern of decreasing pandemics may correspond to each of these subtypes having what we call a low λ value: they have the ability to create one or several large epidemics but appear to settle to epidemics of a smaller size in the long run.

Note that we may be observing a low effective r value, but still a high m value; see conditions (13). This could make the gradient of our map initially negative and place us in a low-λ region. It is important to remember that a complete discussion of influenza invasion dynamics would include details about population structure which would have profound effects on the parameters r and b. Therefore, demographic effects—effects extrinsic to influenza's ability to elude host

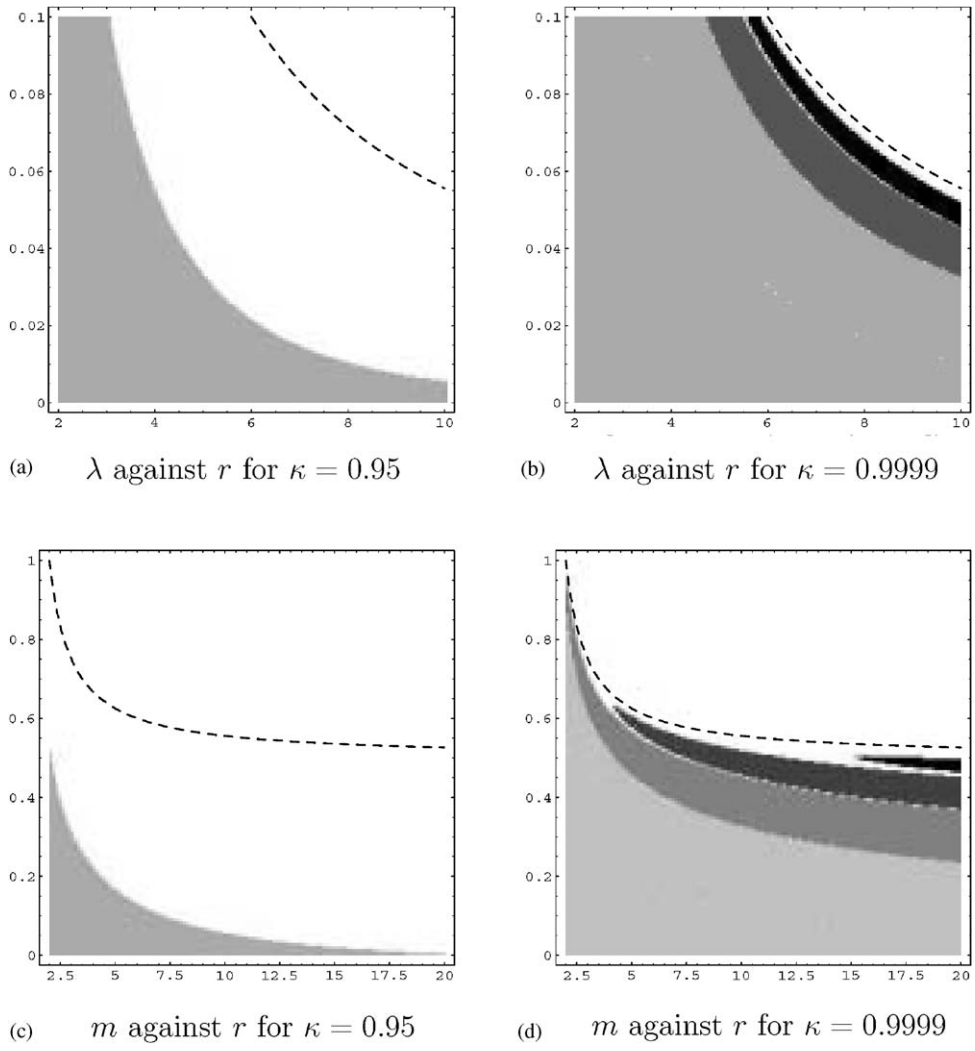

Fig. 8. Consequences of invasion. r is the horizontal axis in all panels while λ (top row) and m (bottom row) are the vertical axes. Calculated numerically by noting if P_0 was the whole range or not, and if not, iterating the map until it reached P_0 or H_0. Influenza can always cause an initial epidemic in a susceptible population. These figures denote subsequent behaviour against two different parameters, keeping κ fixed. White means the $S = 1$ initial condition results in a epidemics settling to the equilibrium, lightest grey means no second epidemic ($S = 1$ is in H_0), next lightest means only two epidemics (H_1), etc. The dotted lines correspond to $2\lambda(r-1) = 1$ which was the condition for zero gradient initially, above which all initial conditions are attracted to the fixed point for all κ.

immunity—may be able to push influenza epidemic dynamics into a region of runaway drift.

4. Discussion

We have constructed a model where the amount of drift each year depends on the size of the previous year's epidemic. Analysis of the derived map shows that there is always a unique stable equilibrium, and that it is globally stable except for solutions that ultimately lead to a situation in which the epidemic fails to take off one year.

En route, some basic qualitative ideas emerged. Specifically, the equilibrium has two fairly distinct regimes of high and low severity. The transition between them is fairly rapid in parameter space. This can be interpreted as the pathogen gaining the ability to escape its own tail, leading to a high number of susceptibles each year. Further, the implication of this is that a vaccine which reduces the effective transmission rate has the potential to reduce the production of new variants below this critical range, having a substantial impact on the continued drift evolution and severity.

A second idea emerges from considering the introduction of the pathogen into a susceptible population. If there is a high potential for mutation, then the pathogen will succeed in long-term establishment. A large initial epidemic will produce sufficiently much antigenic variation that the second and subsequent epidemics will have a sufficient number of effective susceptible hosts. Moreover, high enough mutation rates may cause repeated large annual epidemics and runaway drift. In contrast, a

pathogen with low mutation potential will cause a large initial epidemic, and leave itself without enough susceptibles; either immediately or perhaps in later years, it may fail and die out. Hence, if the introduction was a single event, the pathogen will not necessarily become established. A pathogen with low potential for antigenic variation would need to be introduced repeatedly for it to become established in the longterm.

In considering the possibilty of extinction, case (d) of the low-λ region presents a good example of the race between generating and escaping immunity. For low values of S, influenza does not generate too much immunity and is able to cause subsequent epidemics. For very high values of S, the resulting epidemic yields enough antigenic variation to select a novel variant for next season that escapes this season's immunity. However, for some intermediate values of S, the pathogen generates more immunity than its variants can escape, and it is not able to cause a subsequent epidemic.

Our assumptions were specifically chosen in order to make the model as simple as possible mathematically, so as to allow a full investigation of the system. It would be possible to extend the model in a number of ways, and use this model as a basis for comparison. For example, population structure could be made more realistic. Adding age structure would show which groups are infected for different disease characteristics (e.g. a disease with low mutation would largely infect children as effective immunity would remain for many years, whereas a disease like influenza infects all age classes), and thus which groups contribute to the epidemic and the evolution of the pathogen. Analysis of multiple lineages of cocirculating strains would be useful as recent data point to this phenomenon (Schweiger et al., 2002) and since our model's runaway drift scenario may yield an unexpected amount of diversification. The analysis in Plotkin et al. (2002) on swarms, clusters, and quasispecies also suggests that our model should be adapted to look more finely at multiple strains circulating and evolving during the time course of one epidemic; such a modified model would aid in vaccine recommendations and would allow a revisiting to the famous herald wave of Glezen et al. (1982). Galvani (2003) presents a review of these and other problems and challenges in evolutionary epidemiology.

Influenza dynamics is currently a highly active area of interest (Andreasen, 2003; Earn et al., 2002; Plotkin et al., 2002; Ferguson et al., 2003). Simple models can provide conceptual tools and language for more general and less manageable systems; they also serve to suggest general features and expected behaviours. As models of high complexity are produced, capturing influenza drift with increasing realism, the minimal model presented here offers a tractable analog when considering the interplay between influenza drift and epidemic size.

Acknowledgments

This work was supported in part by NIH grant 1 RO1 GM607929, NIH grant GM28016 to Marcus W. Feldman, and the Danish Natural Science Research Council grant 51-00-0392. JRG is supported by Queens' College, Cambridge, UK. Maciej F. Boni would like to express his sincere gratitude to the Danish-American Fulbright commission who helped fund and encourage this research.

Appendix A. Lower bound for f'

Define

$$g(y) = \frac{y(1-y)+y\log y}{1-y+y\log y} = 1 - \frac{-(1-y)^2}{1-y+y\log y}$$

and note that $g(0)=0$ and $g(1)=-1$. Then,

$$g'(y) = \frac{2(1+y\log y/(1-y))+\log y}{(1+y\log y/(1-y))^2}$$

and we can show that the numerator is negative by applying the inequality

$$2y\log y > y^2 - 1 \quad \text{for } y\in(0,1),$$

which is a direct result of (4).

Appendix B. Upper bound for f' at S^*

At equilibrium, we have

$$f(S^*) = 1-\kappa y^{\lambda}(1-yS^*) = S^*,$$
$$\kappa y^{\lambda} = \frac{1-S^*}{1-yS^*}.$$

Using this, we bound λ at the fixed point values by taking logs and using $\kappa<1$:

$$-\lambda\log y < \log\frac{1-yS^*}{1-S^*},$$
$$\lambda r S^*(1-y) < \log\frac{1-yS^*}{1-S^*} \quad \text{by (7)}.$$

From (9) and the above inequality we get

$$\frac{1}{\kappa y^{\lambda}}f'(S^*) = \frac{y-rS^*y}{1-rS^*y} + \frac{\lambda r(1-S^*y)(1-y)}{1-rS^*y},$$

$$\frac{1}{\kappa y^{\lambda}}(f'(S^*)-1) = \frac{y-rS^*y}{1-rS^*y} + \frac{\lambda r(1-S^*y)(1-y)}{1-rS^*y} - \frac{1-yS^*}{1-S^*}$$
$$< -\frac{(1-y)(1-rS^{*2}y)}{(1-S^*)(1-rS^*y)} + \frac{(1-S^*y)}{S^*(1-rS^*y)}\log\frac{1-yS^*}{1-S^*}$$

and placing more positive terms on the left-hand side, we obtain

$$\begin{aligned}(1-S^*)(1-rS^*y)\frac{1}{\kappa y^{\lambda}}(f'(S^*)-1)\\ <-(1-y)(1-rS^{*2}y)+\frac{(1-S^*)(1-S^*y)}{S^*}\log\frac{1-yS^*}{1-S^*}\\ =-(1-y)-S^*y\log y+\frac{(1-S^*)(1-S^*y)}{S^*}\log\frac{1-yS^*}{1-S^*}\\ =\mathrm{RHS}(S^*,y).\end{aligned}$$

Thinking of the function $\mathrm{RHS}(S^*,y)$ as a function of y for fixed S^*, it is straightforward to show:

- $\mathrm{RHS}(S^*,1)=0$,
- $(\partial/\partial y)\mathrm{RHS}(S^*,y)=0$ at $y=1$,
- $(\partial^2/\partial y^2)\mathrm{RHS}(S^*,y)<0$ for $y<1$.

Hence $\mathrm{RHS}(S^*,y)\leqslant 0$ and thus the desired upper bound on the derivative at a fixed point: $f'(S^*)<1$.

Appendix C. $\lambda(y)$ increasing

From (14), we differentiate to obtain

$$\frac{\partial\lambda}{\partial y}=\frac{-2r+y+4ry-2y^2-2ry^2+y^3-r\log y+ry^2\log y-y^2(\log y)^2}{(\ldots)^2}.$$

Considering the numerator, which we wish to show to be positive, direct calculation and use of simple inequalities (4) gives

- $(\partial^3/\partial y^3)$ numerator <0 for $y\in(0,1)$,
- $((\partial^2/\partial y^2)\ \text{numerator})|_{y=1}=0\ \Rightarrow$ $(\partial^2/\partial y^2)$ numerator $\geqslant 0$ for $y\leqslant 1$,
- $((\partial/\partial y)\ \text{numerator})|_{y=1}=0\ \Rightarrow$ $(\partial/\partial y)$ numerator $\leqslant 0$ for $y\leqslant 1$,
- $(\text{numerator})|_{y=1}=0\ \Rightarrow$ numerator $\geqslant 0$ for $y\leqslant 1$.

Hence $\partial\lambda/\partial y\geqslant 0$ for $y\leqslant 1$.

References

Anderson, R.M., May, R.M., 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford Science Publications, Oxford.

Andreasen, V., 2003. Dynamics of annual influenza A epidemics with immuno-selection. J. Math. Biol. 46, 504–536.

Andreasen, V., Levin, S., Lin, J., 1996. A model of influenza A drift evolution. Z. Angew. Math. Mech. 76 (S2), 421–424.

Andreasen, V., Lin, J., Levin, S.A., 1997. The dynamics of cocirculating influenza strains conferring partial cross-immunity. J. Math. Biol. 35, 825–842.

Bush, R.M., Fitch, W.M., Bender, C.A., Cox, N.J., 1999. Positive selection on the h3 hemagglutinin gene of human influenza virus A. Mol. Biol. Evol. 16 (11), 1457–1465.

Cox, N.J., Brammer, T.L., Regnery, H.L., 1994. Influenza: global surveillance for epidemic and pandemic variants. Eur. J. Epidemiol. 10, 467–470.

Cox, N.J., Subbarao, K., 2000. Global epidemiology of influenza: past and present. Annu. Rev. Med. 51, 407–421.

Dawes, J.H.P., Gog, J.R., 2002. The presence of oscillations in models of the dynamics of multiple virus strains. J. Math. Biol. 45, 471–510.

de Jong, J.C., Rimmelzwaan, G.F., Fouchier, R.A.M., Osterhaus, A.D.M.E., 2000. Influenza virus: a master of metamorphosis. J. Infect. 40, 218–228.

Earn, D.J.D., Dushoff, J., Levin, S.A., 2002. Ecology and evolution of the flu. Trends Ecol. Evol. 17 (7), 334–340.

Ferguson, N.M., Andreasen, V., 2001. The influence of different forms of cross-protective immunity of the population dynamics of antigenically diverse pathogens. In: Blower, S., Castillo-Chavez, C., Cooke, K.L., Kirschner, D., van der Driessche, P. (Eds.), Mathematical Approaches for Emerging and Reemerging Infections: Models, Methods and Theory, IMA Volumes in Mathematics and its Applications. Springer, New York, pp. 157–169.

Ferguson, N.M., Galvani, A.P., Bush, R.M., 2003. Ecological and immunological determinants of influenza evolution. Nature 422, 428–433.

Fitch, W.M., Leiter, J.M.E., Palese, P., 1991. Positive darwinian evolution in human influenza A viruses. Proc. Natl. Acad. Sci. USA 88, 4270–4274.

Fitch, W.M., Bush, R.M., Bender, C.A., Cox, N.J., 1997. Long term trends in the evolution of H(3) HA1 human influenza type A. Proc. Natl. Acad. Sci. USA 94, 7712–7718.

Galvani, A.P., 2003. Epidemiology meets evolutionary ecology. Trends Ecol. Evol. 18 (3), 132–139.

Gill, P.W., Murphy, A.M., 1976. Naturally acquired immunity to influenza type A: a clinical and laboratory study. Med. J. Aust. 2, 329–333.

Glezen, W.P., Couch, R.B., Six, H.R., 1982. The influenza herald wave. Am. J. Epidemiol. 116, 589–598.

Gog, J.R., Grenfell, B.T., 2002. Dynamics and selection of many-strain pathogens. Proc. Natl. Acad. Sci. USA 99, 17209–17214.

Gog, J.R., Swinton, J., 2002. A status-based approach to multiple strain dynamics. J. Math. Biol. 44, 169–184.

Gupta, S., Ferguson, N., Anderson, R.M., 1998. Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. Science 280, 912–915.

Pease, C.M., 1987. An evolutionary epidemiological mechanism, with applications to type A influenza. Theor. Popul. Biol. 31, 422–452.

Plotkin, J.B., Dushoff, J., Levin, S.A., 2002. Hemagglutinin sequence clusters and the antigenic evolution of influenza a virus. Proc. Natl. Acad. Sci. USA 99 (9), 6263–6268.

Schweiger, B., Zadow, I., Heckler, R., 2002. Antigenic drift and variability of influenza viruses. Med. Microbiol. Immunol. 191 (3–4), 133–138.

Simonsen, L., Clarke, M.J., Schonberger, L.B., Harden, N.H., Cox, N.J., Fukuda, K., 1998. Pandemic versus epidemic influenza mortality: A pattern of changing age distribution. J. Infect. Dis. 178, 53–60.

Webby, R.J., Webster, R.G., 2001. Emergence of influenza A viruses. Philos. Trans. R. Soc. Lond. B 356, 1817–1828.

Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., Kawaoka, Y., 1992. Evolution and ecology of influenza A viruses. Microbiol. Rev. 56, 152–179.

Wilson, I.A., Cox, N.J., 1990. Structural basis of immune recognition of influenza virus hemagglutinin. Annu. Rev. Immunol. 8, 737–771.