Influenza emergence in the face of evolutionary constraints

Adam Kucharski* and Julia R. Gog

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK

Different influenza subtypes can evolve at very different rates, but the causes are not well understood. In this paper, we explore whether differences in transmissibility between subtypes can play a role if there are fitness constraints on antigenic evolution. We investigate the problem using a mathematical model that separates the interaction of strains through cross-immunity from the process of emergence for new antigenic variants. Evolutionary constraints are also included with antigenic mutation incurring a fitness cost. We show that the transmissibility of a strain can become disproportionately important in dictating the rate of antigenic drift: strains that spread only slightly more easily can have a much higher rate of emergence. Further, we see that the effect continues when vaccination is considered; a small increase in the rate of transmission can make it much harder to control the frequency at which new strains emerge. Our results not only highlight the importance of considering both transmission and fitness constraints when modelling influenza evolution, but may also help in understanding the differences between the emergence of H1N1 and H3N2 subtypes.

Keywords: influenza; fitness costs; emergence; transmission; antigenic drift

1. BACKGROUND

The two influenza A subtypes common in humans, H1N1 and H3N2, frequently escape population immunity by changing their antigenic properties. Between 1983 and 2009, H3N2 fixed more amino acid changes in antigenic sites than H1N1 [1–4], and vaccines against H3N2 were updated more often than vaccines against H1N1 [5,6]. The higher rate of appearance and fixation of non-synonymous mutations in H3N2 could be owing to a combination of factors, presumably including effective virus population size, and selective pressures according to host prior immunity. However, the root cause of the epidemiological and evolutionary differences between subtypes of influenza in humans remains poorly understood [7,8].

Transmissibility, or more specifically the basic reproductive ratio R_0, the average number of secondary cases generated by the average infective individual in a naïve population, has been much studied for influenza [9,10]. The probability that a strain will become established is determined by its effective reproductive ratio, which depends on R_0 and the level of population immunity. Seasons in which H3N2 was the dominant subtype have been associated with a higher effective reproductive ratio [11,12]; there are also typically more deaths from pneumonia and influenza in seasons where H3N2 is the dominant strain [13]. Transmissibility must be therefore considered as a possible mechanism behind the difference in antigenic drift speeds between the two subtypes.

One strain of influenza may confer partial immunity to another [14], and multi-strain models can be used to study the evolutionary dynamics of the disease [15–17]. Here, we define a new strain to be a virus in which the surface proteins have undergone sufficient change, as a result of

mutation, to affect people who were previously immune to it. We will use such a model to examine the role that transmission plays in the emergence of new strains.

While multi-strain models are well established, few allow mutation to directly affect aspects of virus phenotype other than antigenicity. Although partially explored in individual-based simulations [1,18], a previous study [19] introduced a model to investigate the behaviour of influenza evolution when mutation affected the ability of a virus to transmit between hosts.

In studies *in vitro*, single nucleotide mutations in other RNA viruses have resulted, on average, in a fitness reduction [20–22]. If random mutation carries a fitness cost, there is a possibility that mutations associated with antigenic change also affect the phenotype in some other way, reducing viral fitness of antigenic escape mutants through antagonistic pleiotropy [23]. Alternatively, or additionally, 'deleterious hitchhikers'—mutations elsewhere on the genome that happen to be picked up in the process of emergence of the antigenic escape mutant—could have a negative effect (M. Lässig 2011, personal communication). The ill-fated side branches of influenza's evolutionary tree could be the result of costs incurred by either mechanism (M. Lässig 2011, personal communication; [24]). However, new influenza strains appear every few years [25]; this ongoing antigenic mutation and survival could be possible through subsequent compensatory mutations [26,27].

Gog [19] developed a simple model of these processes to show that a high-level fitness loss could in fact stop viral evolution ('strain lock'), and that vaccination could also nudge the system into this state. In this paper, a two-tier framework is used to separate the interaction of strains via population immunity, modelled deterministically, and the emergence of new strains, treated as a stochastic process. This improves the realism of the fitness loss and compensatory mutation processes, and makes

* Author for correspondence (ak640@cam.ac.uk).

the model computationally far more tractable, allowing the examination of a variety of assumptions.

We see that although the rate of emergence increases almost linearly with transmission in the absence of fitness constraints, when mutation incurs a cost, small changes in the R_0 of the fully fit virus can have a large impact on emergence. This is an example of phylodynamics [7], the relationship between epidemic and evolution dynamics.

Previous work has shown that often a threshold can exist where vaccination can slow or pause the emergence of new strains [19,28,29] and therefore we also look at how the use of vaccination in slowing antigenic evolution is affected by change in R_0. We show that vaccination can be far more successful in controlling evolution when fitness costs are present, highlighting the importance of considering such constraints when modelling influenza evolution.

2. METHODS
(a) *A two-tiered model*
A first approach to adding fitness costs to a multi-strain model, suggested by Gog [19], was to take the one-dimensional line of strains used in many previous models [16,30] and extend it to a two-dimensional space. Each variant had an intrinsic fitness, as well as an antigenic type (which changed in one direction). Mutation to a new antigenic type caused a reduction in fitness, and compensatory mutations were included so that strains can also regain fitness lost. A simple stochastic approximation was included; a test was imposed on new strains present in less than one individual to ensure that they did not emerge automatically. This meant that stochastic effects were present in a basic way, although the approximation had some drawbacks, which will be discussed here.

Suppose R_0 denotes the basic reproductive ratio of the fully fit virus, i.e. the top level of strains described by Gog [19]. This definition of R_0 will be used throughout this paper. A full fitness strain therefore has an effective reproductive ratio $R = R_0 S/N$, where S is the number of individuals susceptible to the strain, and N is the population size. Here, as in Gog [19], we assume that antigenic change incurs a fixed fitness cost, so a new strain has an effective reproductive ratio $R_1 = \chi R < R$ initially, where χ denotes the relative fitness of the mutant, reducing R by a factor $0 \leq \chi \leq 1$.

Let I_0 be the number of individuals infected with the new strain. If large numbers are infective with the dominant strain, even a low rate of mutation will lead immediately to $I_0 > 1$. Hence it circumvents the stochastic test, and means a possible 'overspill' effect, whereby a circulating strain can prop up its weaker mutant, allowing it to persist long enough to undergo a compensatory mutation, despite the fact that would not necessarily survive if both fitness loss and compensation were modelled as a branching process. Conversely, if $I_0 < 1$ always, as can happen when the dominant strain is endemic at low numbers, the stochastic test will not allow the weaker mutation to ever emerge if it has $R < 1$. However, it has been shown that if emergence is modelled stochastically, infections with $R < 1$ may still survive long enough to subsequently undergo a compensatory mutation [9]. A further issue, mentioned by Gog [19], is that, for $\chi = 1$, there are still two versions of each strain circulating, whereas

ideally the two-dimensional space should collapse to a one-dimensional line as χ approaches 1, and revert to a simple drift model similar to that of Gog & Grenfell [16].

As a simple model of strain emergence leads to the above issues, yet a fully stochastic process with mutation and compensation is computationally intensive, a balance must be sought. We shall therefore use a two-tiered model [31], focusing on the deterministic and stochastic processes separately.

A discrete-time deterministic model of infection at the population level forms the top tier; this is described in §2c. Mutation to a lower fitness strain occurs at a constant rate but the actual emergence of new strains, including compensatory mutation, is dictated by the bottom tier: a stochastic approximation, which has boiled down the details of emergence to a simpler process. This is outlined in §2b, and allows calculation of the probability that the new strain appears and causes an epidemic. This probability is then used in the top tier of the model as a test of the viability of a new strain appearing at each point in time.

(b) *Emergence process*
We consider two routes to emergence, as shown in figure 1. In particular, we wish to determine whether a strain with $R < 1$ can mutate before it goes extinct. Similar problems have been tackled for evolution in the face of selection pressure [32] and zoonoses [9,33].

First, we outline a useful result that will be used in the full problem: the probability that a strain with particular fitness fails to cause an epidemic. We assume that $I_0 \ll S/N$, and each infective individual can be considered independent. In the standard susceptible-infective-recovered model, where the effective reproductive ratio R is constant, the probability a single infective individual will fail to generate an epidemic is [34]:

$$\text{probability of extinction} = P = \begin{cases} \dfrac{1}{R} & \text{if } R > 1, \\ 1 & \text{else.} \end{cases}$$

As transmission by each infective individual can be considered independent, we have:

$$P = \min\left\{ 1, \left(\frac{1}{R}\right)^{I_0} \right\}. \tag{2.1}$$

Next, this result can be applied to the full evolutionary process, with compensatory mutation occurring at rate m. Let $I_0(t)$ be the number of individuals infected with the reduced fitness strain, which has reproductive ratio R_1, and $I_2(t)$ be the number of individuals infected with the compensated strain, with reproductive ratio $R_2 > R_1$. Although this method will work for more general fitness structures, in this paper we assume full fitness recovery, i.e. $R_2 = R_0 S_2/N$, where S_2 is the number of individuals susceptible to strain 2.

Following Keeling & Rohani [34], if $I_0 = I(0) = 1$, the probability the reduced fitness strain fails to cause an epidemic, P_1 can be written as a function of the recovery, transmission and compensatory mutation processes. P_2, the probability the full fitness strain fails to cause an epidemic, can be treated the same as P above. Rescaling time by the
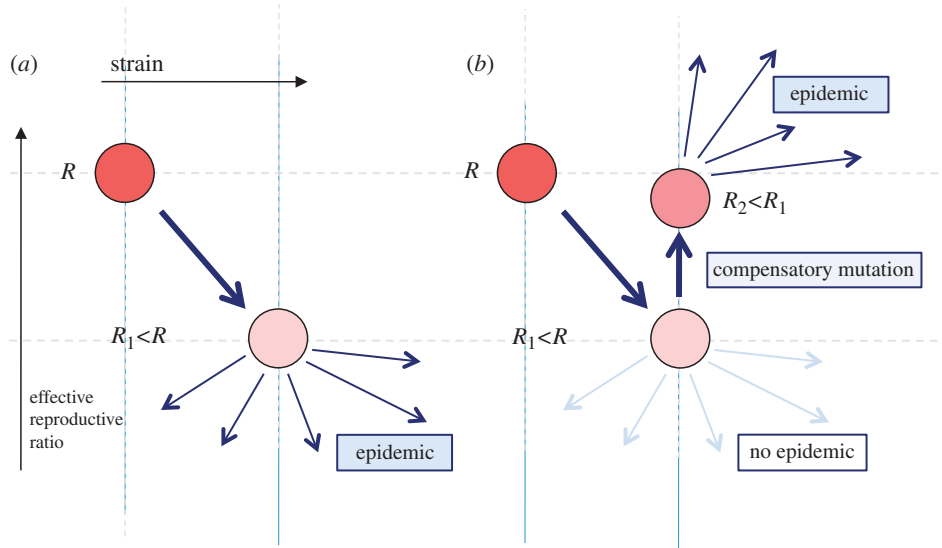
Figure 1. The two routes to emergence: (a) despite reduced fitness R_1, virus still causes epidemic; (b) no initial epidemic, but compensatory mutation occurs before extinction and virus with fitness R_2 subsequently causes outbreak.

infectious period $1/\gamma$, we have:

$$P_1 = \frac{1}{1 + R_1 + m_\gamma} + \frac{R_1}{1 + R_1 + m_\gamma}P_1^2$$
$$+ \frac{m_\gamma}{1 + R_1 + m_\gamma}P_2, \tag{2.2}$$

where $m_\gamma = m/\gamma$. Taking the smallest root in [0,1] of this equation gives us the probability that the reduced fitness strain fails to cause an epidemic [35],

$$P_1 = \min$$

$$\left\{ 1, 1 - \frac{(R_1 - 1 - m_\gamma) + \sqrt{(R_1 - 1 - m_\gamma)^2 + 4m_\gamma R_1(1 - P_2)}}{2R_1} \right\}, \tag{2.3}$$

where $P_2 = \min\{1, 1/R_2\}$.

If $I_0 \neq 1$ and $I_0 \ll S/N$ still, we can again use independence of interactions to obtain:

$$\mathbb{P}\left[\text{fails to emerge} \mid I(0) = I_0\right] = P_1^{I_0}. \tag{2.4}$$

Also note that if $m_\gamma = 0$ in equation (2.3), we recover the single strain case $P_1 = \min\{0, 1/R_1\}$, as expected. This is also true if $R_1 - 1 - m_\gamma \gg 4m_\gamma(1 - P_2)R_1$, implying that if there is little or no fitness cost, with $R_1 > 1$ and $m_\gamma \ll 1$, the probability of emergence behaves as if we had a single strain branching process. Therefore, the step-down and step-up do not have an unnecessary effect if R_1 is not reduced by a fitness cost.

We can use these results to approximate the probability a strain will emerge in our full model. Suppose strain a is our frontmost strain (i.e. the one that emerged most recently), at each point in time, we know the prevalence, I_a, and susceptibility to a mutant of this strain, S_{a+1}. If mutation occurs at a rate m, we would expect to see

$$I_0 = I_a(1 - e^{-m}) \tag{2.5}$$

new mutants appear in a single time step. Using the above results, we can therefore calculate $P_{a+1} = P_{a+1}(R_1, I_0)$, the probability that an infective individual fails to emerge with the new strain $a + 1$, with $R_1 = \chi S_{a+1}/N$, and I_0 as above.

Computationally, this only requires the following test to be applied at each time step:

$$I_{a+1} = \begin{cases} 0 & \text{if } P_{a+1} > \text{uniform random variable } X \in (0, 1), \\ 1 & \text{else.} \end{cases}$$

Although we have two routes to emergence (figure 1), and hence two possible values of R for I_{a+1}, it is reasonable to assume that if the virus does cause an epidemic, its numbers will end up sufficiently large for it to undergo the compensatory mutation during this epidemic [26,36]. I_{a+1} therefore represents individuals infective with the full fitness strain.

As in Gog [19], unrealistic strain survival is avoided by setting any strain that decreases in prevalence to $I < 1$, with an $R < 1$, equal to zero. This avoids the decay to infinitesimal levels (i.e. attostrains [37]), and potential re-emergence, which can occur in purely deterministic systems.

(c) *Epidemic process*
The second tier of the model keeps track of how immunity changes during epidemics at the population level. To do this, we use a status-based model with one level of fitness only [16], with S_a denoting the number of hosts susceptible to (at least) strain a; I_a the number of hosts infectious with strain a; and Λ_a the force of infection of strain a. We will use a discrete formulation, as it will make it easier to deal with the mutation step at each point in time (which is based on the random number tests in §2b). Our system is therefore:

$$S_a' = N(1 - e^{-\mu}) + S_a e^{-(\Lambda_a + \mu)} \tag{2.6}$$

and

$$I_a' = S_a(1 - e^{-\beta_a I_a}) + I_a e^{-\gamma}, \tag{2.7}$$

where

$$\Lambda_a = \sum_k \beta_k c_{ka} I_k.$$

Here S_a' and I_a' represent the values of S_a and I_a at the next time step. Each time step represents one day, and our parameters are scaled as such. We assume a fixed population size N, so μ acts as both the birth and death rate. β_a

Table 1. Parameters (sources: α [38,39]; γ [40]).

variable	description	value
N	population size	10 million
μ	host birth/death rate	$(80 \times 365$ days$)^{-1}$
α	coarseness of antigenic space	0.5
c_{ka}	cross-immunity between strain k and a	$e^{-\alpha|k-a|}$
$1/\gamma$	infectious period	5 days
R_0	basic reproductive ratio	varies
β	transmission rate	$R_0\gamma/N$
χ	relative fitness of mutant	varies
m	mutation rate	$(4$ years$)^{-1}$

represents the transmission rate for strain a. Recovery is at a fixed rate γ. Cross-immunity is given by the term c_{ka}. We set this to decay exponentially with antigenic distance, and we assume that strains give complete immunity to themselves, so c_{ka} must be such that $c_{aa} = 1$. Table 1 gives the values of the parameters used in the model. The infectious period in §2b was given by $1/\gamma$, so here γ includes both recovery from infection and natural death. However, for typical human influenza parameters, the death rate makes a negligible contribution to γ.

A moving strain space, as outlined by Gog [19], is also used to hasten computation. At each point in time all strains with $I > 0$ are identified. By the nature of mutation in the model, these cluster together and hence form a set of adjacent strains on our line. This set is active in the model, in that the above equations only use the variables corresponding to these prevalent strains. We update this active set at each time step. If the backmost (i.e. least recent) strain in the set falls below one infective, then we drop that variable, and if a new strain appears (see §2b), we add new variables. There is also a buffer in practice: if a new strain emerges, we create variables associated to the next strain on the line, in anticipation of its emergence.

We approximate the new level of susceptibility by multiplying the immunity with the previous strain by the reduction owing to the imperfection of cross-immunity:

$$S_{a+1} = (N - S_a)\,e^{-\alpha}.$$

This approximation underestimates immunity (in the form of reduced transmissibility) slightly [19]; this could affect accuracy under a high mutation rate, as more frequent appearance of mutants would mean that the strain space moves faster. However, it is less of an issue for rates similar to those considered in this paper.

Drift speed, used to measure the speed of evolution in our model, is defined as the rate of emergence: the mean number of strains that appear per year, over twenty 50-year simulations. If the system goes extinct, we only look at the behaviour up to the point of extinction. Over many runs, this rate approaches a fixed value for each R_0 and χ. We focus on $\chi \in [0.6,1]$, however, as below these values a tendency to lock or go extinct reduces the accuracy of the mean.

3. RESULTS
In a manner similar to that of Gog [19], our model exhibits three dynamical behaviours: *locked*, where only one strain circulates; *drifting*, with the continuing emergence
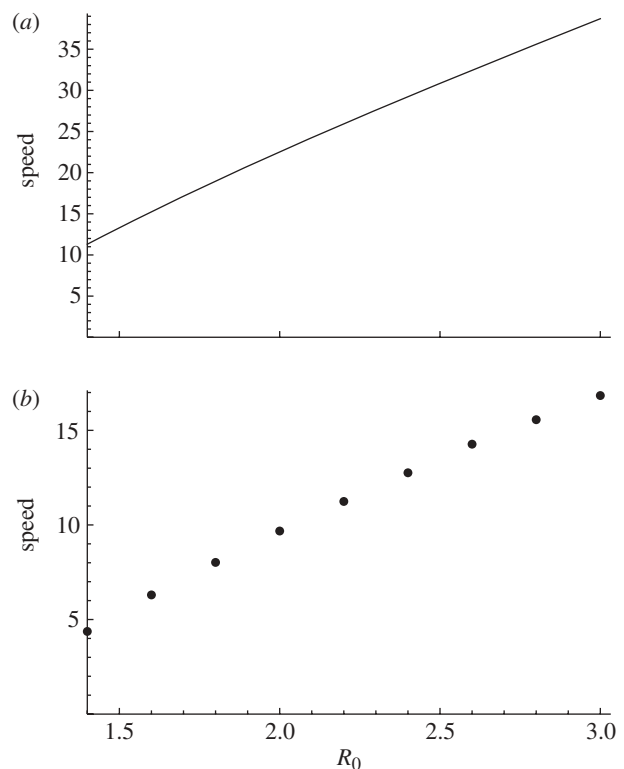
Figure 2. (a) Relationship between R_0 and emergence rate predicted by the deterministic model of Gog & Grenfell [16]; (b) relationship in our model in the absence of fitness costs.

of strains; and *extinction*, where no strains remain. Once drift is established, the system is more likely to continue drifting than switch to the other two, especially when χ is larger. As χ is reduced, a drifting system will have an increasing tendency to fall into a locked state, or go extinct.

First, to check the basic drift dynamics when there is no fitness loss ($\chi = 1$), this model is compared with a more basic deterministic model [16] using a continuous version of equations (2.6) and (2.7). In this simpler model, the authors could use an analytical method to calculate drift speed c, and it was shown that:

$$c = \min_{\lambda} \left\{ \frac{\beta - \gamma - \mu + 2m(\cosh\lambda - 1)}{\lambda} \right\}.$$

Such a result is not possible for a stochastic system of equations, but using $R_0 = \beta/\gamma$, we can compare the above theoretical result with the speed observed in our model (using the same values of γ, μ and m). Simulations were run for several values of R_0, with the model deliberately started in a drifting state, and calculated rate of emergence over twenty 50-year runs. This gives a measure of the speed of antigenic drift.

Figure 2 shows that the qualitative relationship between R_0 and speed of emergence is very similar, with the rate of emergence dependent on $R_0 \in [1.4,3]$ in a near-linear way. Although the model presented in this paper evolves at a slower rate, this is most probably owing to the stochastic step that tempers emergence.

However, when fitness costs are introduced we see something quite different. The left-hand side of figure 3 shows that when $\chi < 1$, the impact of R_0 can become
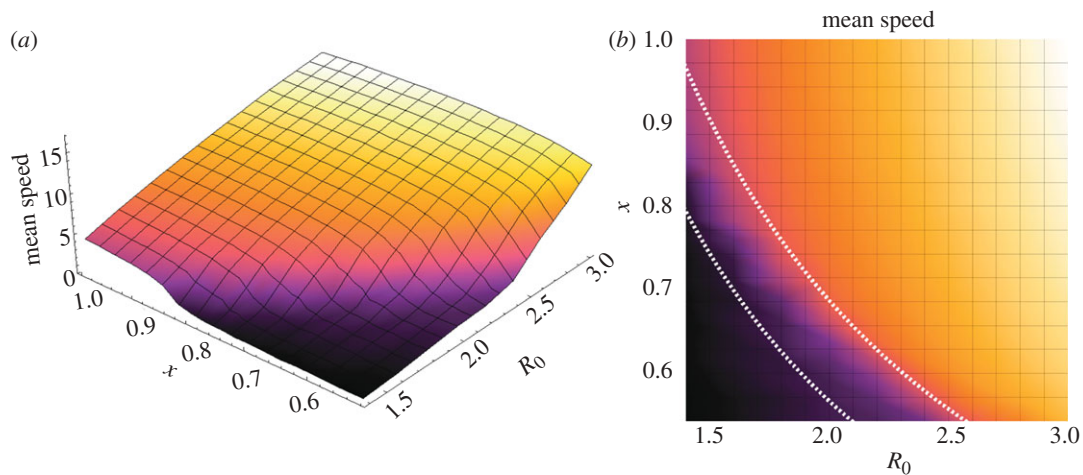
Figure 3. (a) Mean strains emerging per year, as a function of R_0 and fitness χ. Twenty 50-year runs were used to calculate speeds at 200 points in the R_0/χ plane. (b) A contour plot of this graph with the theoretical bounds of the drop-off overlaid as dotted white lines. Values of R_0 and χ leading to a high rate of emergence are in the top right and to a low rate in the bottom left. The lines are given by the value at which $R_1 = \chi R_0 S/N = 1$, for $S = 0.75$ (top) and $S = 0.9$ (bottom), as susceptibility to new strains falls within this range in model simulations.

nonlinear. A drop-off appears, along which small changes in transmission can have a disproportionate effect on the speed at which strains appear. Moreover, at the base of this 'drift cliff', the relationship is much flatter, and R_0 has even less effect on evolution than it did at when there were no fitness constraints.

It also shows that, if χ and R_0 do not vary much between strains, the appearance of a more fit strain with larger R_0 could immediately impact the rate of emergence, as any subsequent antigenic mutant will find it much easier to establish itself.

We can elucidate this result by considering the role of the stochastic step in the rate of emergence. Figure 4 shows the probability a strain emerges for a particular effective reproductive ratio, $R_1 = \chi R_0 S/N$, equal to $1 - P_1$ in equation (2.3). As R_1 decreases towards 1, it becomes less likely a new strain will emerge at that step. If S were fixed this would imply that, for each R_0, there is a critical value of χ below which the appearance of strains is relatively rare. Although a subsequent compensatory mutation and emergence event can still occur, as shown by the solid line in figure 4, this is rare for $R_1 < 1$ and the long wait for the appearance of a new strain slows down the rate of emergence.

However, in our model, S is not fixed because immunity to each strain varies, and depends heavily on other circulating infection. Therefore, we can only estimate this relationship between χ and R_0. Simulations show that susceptibility S to new strains falls between 0.75 and 0.9, so, by finding the region where $\chi R_0 S/N < 1$, this can be used to estimate the location of the drop-off in the rate of emergence. The right-hand side of figure 3 shows that this theoretical result agrees with the simulated behaviour, and hence provides an explanation for the location of the drop-off in emergence.

(a) Vaccination

Vaccination will be explored by observing how the effects of pulsed vaccination vary with fitness and transmission. For each χ and R_0, we let the system settle in a drifting state before vaccinating a fixed proportion of the
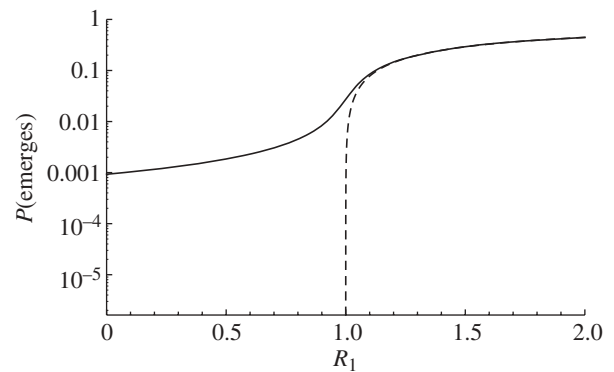

Figure 4. The probability a strain will emerge given $R_1 = \chi R_0 S/N$, as calculated from equation (2.3). Here $S/N = 0.9$ and $I_0 = 1$. The dashed line shows the probability that it will emerge in the absence of any compensatory mutations, as in equation (2.1). A strain with $R_1 < 1$ will find it much harder to cause an epidemic.

population chosen at random all at once, at time T, with a vaccine most similar to the most recently emerged strain, and cross-reactive with reduced factor $e^{-\alpha d}$ for a strain distance d away. Hence, if v is the vaccine strain, and we vaccinate a proportion p,

$$\text{new } S_a = (1 - p\,e^{-\alpha|a-v|})S_a.$$

The number of strains that emerge in the interval $[T, T + 365]$ is then counted and this number is compared with the expected rate of emergence for that χ and R_0, as calculated in figure 3, to obtain a relative speed between 0 and 1 as a result of vaccination of the population.

When 20 per cent of the population is vaccinated, the rate of emergence of the strain with $R_0 = 1.6$ is slowed much more than that of the one with $R_0 = 2$. Taking the mean from 200 runs, figure 5 shows that while R_0 makes little difference when $\chi = 1$, for $\chi < 1$, the infection with $R_0 = 2$ can be far harder to control.

This is because, for a strain with reduced fitness R_1 barely above 1, vaccination can lower susceptibility and put R_1
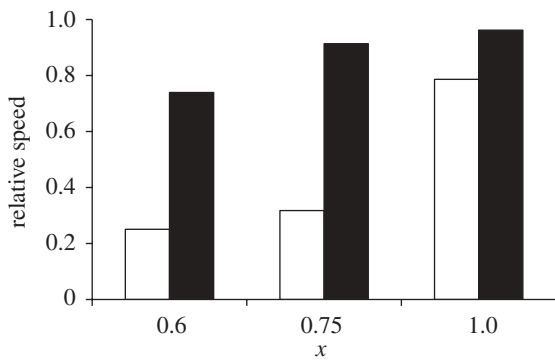
Figure 5. Relative speed of emergence as a result of vaccination of 20% of the population. Calculated as number of new strains in year post-vaccination divided by mean new strains per year if no vaccination. Two hundred simulations were performed for each value, and mean relative speed taken. Open bars, $R_0 = 1.6$; closed bars, $R_0 = 2$.

below 1, and on the slow side of the drop-off. If $R_0 = 1.6$, $R_1 = \chi R_0 S/N$ falls into this region. By contrast, with a small increase in transmissibility (in this case $R_0 = 2$), R_1 is far enough from 1 so as to be far less affected by vaccination. Even a 20 per cent reduction in the number of susceptibles is not enough to push the strain into the area of the drop-off, showing how R_0 has become disproportionately important.

4. CONCLUSIONS AND DISCUSSION

We have seen that if the fitness losses that hamper RNA viruses *in vitro* also affect an influenza virus escaping host immunity, drift speed can be affected. Not only does the relative fitness of the virus, χ, influence the rate of antigenic escape, there is a feedback between the evolutionary and epidemic processes, with susceptibility and transmission at the population level deciding the impact χ has, and vice versa. As seen in figure 3, under fitness constraints, R_0 can become disproportionately important in dictating dynamics. This implies that two strains with identical antigenic space and mutation rates could evolve at very different rates with only a marginal difference in R_0, unlike that required by the deterministic model of Gog & Grenfell [16]. This is illustrated by the drop-off we see between regions of high and low rates of emergence. We have shown that its location can be estimated from the probability that a strain emerges in a branching process; the theoretical prediction coincides well with the simulated results. It also explains the switch between the two stable patterns (locked and drifting) observed by Gog [19]. As χ decreases, the probability of switching from a drifting to a locked (or extinct) state becomes much larger, as emergence gets rarer. The drop-off in the overall rate of emergence reflects this reduction in tendency to drift. Further, the existence of such a 'drift cliff' between regions of high and low rates of emergence may affect the impact of antivirals, which have been known to reduce the infectious period (and hence the R_0) of influenza [41].

As influenza vaccines are selected around nine months in advance of each hemisphere's annual winter epidemic [42], it would be interesting to see whether there is any indication in historical data that a particularly transmissible season elsewhere in those nine months leads to a higher rate of emergence of new strains, and hence a vaccine update. This may be possible in retrospect, but it would be difficult to monitor the transmissibility of a strain in real-time, with the aim of providing an indication of the level of drift to be expected, as the accuracy of transmissibility estimates are reduced by reporting errors and gaps in surveillance [43]. In addition, calculating R_0 is difficult for seasons with milder epidemics [11]. However, even if the theoretical mechanisms suggested in this paper cannot immediately be applied to real-time prediction, they could be useful in understanding the differences between the emergence of H1N1 and H3N2 strains.

Second, we have seen that the disproportionate impact of transmission continues when the immediate period post-vaccination is considered; under fitness constraints, a small increase in R_0 can make controlling the rate at which new strains emerge much harder (figure 5), as well as controlling the rate at which they spread. However, it also implies that if strains do lose fitness when escaping immunity, vaccination could have a greater benefit than a simpler model would imply.

We have made several assumptions in our model. We have used status-based variables for tractability, simplifying the immune structure of the population. We have also imposed a one-dimensional strain structure on influenza evolution. Although this reflects the sequential appearance of strains implied by the ladder-like phylogenetic tree of influenza A [7,44], the causes of this phylogeny are likely to be complex, with several explanations having been proposed [1,17,18,45] for the observed evolution. By using a simple line of strains, this paper does not address the issue of strain dimensionality. Instead, it assumes that, in the long term, through some mechanism not included here, strain evolution follows a one-dimensional path through antigenic space, with cross-reactivity between strains decaying exponentially as the distance between them increases.

Like several status-based models, we have also assumed that cross-immunity acts to reduce transmission. We do not consider spatial structure, which may also affect how transmission rates impact on strain emergence [12], nor seasonality: this paper presents a single population with influenza constantly circulating. Our parameters, although chosen to be plausible, are also approximate. In particular, mutation rates can affect the steepness of the drop-off: an increase leads to easier emergence and a slightly flatter relationship, whereas a slower rate hinders the appearance of new strains and accentuates the drop (results not shown). Further, the mechanism of fitness loss and regain, via compensatory mutation, is necessarily approximate: a better understanding, ideally quantitative, of how antigenic escape might incur a fitness cost is needed if the effects of evolutionary constraints are to be modelled more accurately. If it was the case that the process of emergence involved several mutations, we would expect that the $R = 1$ boundary is still important in deciding the viability of a new strain. Although simplifications have been made, our framework could easily incorporate new developments in the understanding of influenza or RNA fitness constraints.

Taking a two-tiered approach, we focused on the emergence process and interaction of strains via epidemic dynamics separately, combining a stochastic approach

for the former with an ordinary differential equation model for the latter. This allowed us to not only draw attention to the role R_0 can have in dictating a strain's rate of antigenic escape, but also to show that inclusion of realistic evolutionary constraints, as well as careful consideration of the emergence process, can have implications on the use of vaccination in reducing the rate of such escape.

REFERENCES

1 Ferguson, N. M., Galvani, A. P. & Bush, R. M. 2003 Ecological and immunological determinants of influenza evolution. *Nature* **422**, 428–433. (doi:10.1038/nature01509)

2 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. (doi:10.1073/pnas.94.15.7712)

3 Suwannakarn, K., Chieochansin, T., Thongmee, C., Makkoch, J., Praianantathavorn, K., Theamboonlers, A., Sreevatsan, S. & Poovorawan, Y. 2010 Molecular evolution of human H1N1 and H3N2 influenza A virus in Thailand, 2006–2009. *PLoS ONE* **5**, e9717. (doi:10.1371/journal.pone.0009717)

4 Zaraket, H. *et al.* 2009 Molecular evolution of human influenza A viruses in a local area during eight influenza epidemics from 2000 to 2007. *Arch. Virol.* **154**, 285–295. (doi:10.1007/s00705-009-0309-9)

5 Hay, A., Gregory, V., Douglas, A. & Lin, Y. 2001 The evolution of human influenza viruses. *Phil. Trans. R. Soc. Lond. B* **356**, 1861–1870. (doi:10.1098/rstb.2001.0999)

6 Neuzil, K. & Bright, R. 2009 Influenza vaccine manufacture: keeping up with change. *J. Infect. Dis.* **200**, 835–837. (doi:10.1086/605507)

7 Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A. & Holmes, E. C. 2004 Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **303**, 327–332. (doi:10.1126/science.1090727)

8 Holmes, E. C. 2010 The comparative genomics of viral emergence. *Proc. Natl Acad. Sci. USA* **107**, 1742–1746. (doi:10.1073/pnas.0906193106)

9 Antia, R., Regoes, R., Koella, J. & Bergstrom, C. 2003 The role of evolution in the emergence of infectious diseases. *Nature* **426**, 658–661. (doi:10.1038/nature02104)

10 Grassly, N. & Fraser, C. 2008 Mathematical models of infectious disease transmission. *Nat. Rev. Microbiol.* **6**, 477–487. (doi:10.1038/hdy.1982.7)

11 Chowell, G., Miller, M. A. & Viboud, C. 2008 Seasonal influenza in the United States, France and Australia: transmission and prospects for control. *Epidemiol. Infect.* **136**, 852–864. (doi:10.1017/S0950268807009144)

12 Viboud, C., Bjornstad, O. N., Smith, D. L., Simonsen, L., Miller, M. A. & Grenfell, B. T. 2006 Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–451. (doi:10.1126/science.1125237)

13 Simonsen, L., Reichert, T., Viboud, C., Blackwelder, W., Taylor, R. & Miller, M. 2005 Impact of influenza vaccination on seasonal mortality in the US elderly population. *Arch. Int. Med.* **165**, 265–272. (doi:10.1001/archinte.165.3.265)

14 Potter, C. W. 1979 Determinants of immunity to influenza infection in man. *Br. Med. Bull.* **35**, 69–75.

15 Gupta, S., Ferguson, N. & Anderson, R. 1998 Chaos, persistence and evolution of strain structure in antigenically diverse infectious agents. *Science* **280**, 912–915. (doi:10.1126/science.280.5365.912)

16 Gog, J. R. & Grenfell, B. T. 2002 Dynamics and selection of many-strain pathogens. *Proc. Natl Acad. Sci. USA* **99**, 17 209–17 214. (doi:10.1073/pnas.252512799)

17 Koelle, K., Cobey, S., Grenfell, B. T. & Pascual, M. 2006 Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science* **314**, 1898–1903. (doi:10.1126/science.1132745)

18 Tria, F., Lässig, M., Peliti, L. & Franz, S. 2005 A minimal stochastic model for influenza evolution. *J. Stat. Mech.* P07008. (doi:10.1088/1742-5468/2005/07/P07008)

19 Gog, J. R. 2008 The impact of evolutionary constraints on influenza dynamics. *Vaccine* **26S**, C15–C24. (doi:10.1016/j.vaccine.2008.04.008)

20 Domingo-Calap, P., Cuevas, J. M. & Sanjuán, R. 2009 The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. *PLoS Genet.* **5**, e1000742. (doi:10.1371/journal.pgen.1000742)

21 Iglesia, F. & Elena, S. 2007 Fitness declines in tobacco etch virus upon serial bottleneck transfers. *J. Virol.* **81**, 4941–4947. (doi:10.1128/JVI.02528-06)

22 Sanjuán, R., Moya, A. & Elena, S. F. 2004 The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc. Natl Acad. Sci. USA* **101**, 8396–8401. (doi:10.1073/pnas.0400146101)

23 Hedrick, P. 1999 Antagonistic pleiotropy and genetic polymorphism: a perspective. *Heredity* **82**, 126–133. (doi:10.1038/hdy.1982.7)

24 Nelson, M. & Holmes, E. 2007 The evolution of epidemic influenza. *Nat. Rev. Genet.* **8**, 196–205. (doi:10.1038/nrg2053)

25 Cox, N. & Fukuda, K. 1998 Influenza. *Emerg. Infect. Dis.* **12**, 27–38.

26 Cowperthwaite, M. & Meyers, L. 2007 How mutational networks shape evolution: lessons from RNA models. *Annu. Rev. Ecol. Evol. Syst.* **38**, 203–230. (doi:10.1146/annurev.ecolsys.38.091206.095507)

27 Rimmelzwaan, G. F., Berkhoff, E. G. M., Nieuwkoop, N. J., Smith, D. J., Fouchier, R. A. M. & Osterhaus, A. D. M. E. 2005 Full restoration of viral fitness by multiple compensatory co-mutations in the nucleoprotein of influenza A virus cytotoxic T-lymphocyte escape mutants. *J. Gen. Virol.* **86**, 1801–1805. (doi:10.1099/vir.0.80867-0)

28 Gomes, M. G. M., Franco, A. O., Gomes, M. C. & Medley, G. F. 2004 The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy. *Proc. R. Soc. Lond. B* **271**, 617–623. (doi:10.1098/rspb.2003.2606)

29 Boni, M. F., Gog, J. R., Andreasen, V. & Feldman, M. W. 2006 Epidemic dynamics and antigenic evolution in a single season of influenza A. *Proc. R. Soc. B* **273**, 1307–1316. (doi:10.1098/rspb.2006.3466)

30 Andreasen, V., Lin, J. & Levin, S. A. 1997 The dynamics of cocirculating influenza strains conferring partial cross-immunity. *J. Math. Biol.* **35**, 825–842. (doi:10.1007/s002850050079)

31 Koelle, K., Khatri, P., Kamradt, M. & Kepler, T. B. 2010 A two-tiered model for simulating the ecological and evolutionary dynamics of rapidly evolving viruses, with an application to influenza. *J. R. Soc. Interface* **7**, 1257–1274. (doi:10.1098/rsif.2010.0007)

32 Iwasa, Y., Michor, F. & Nowak, M. 2004 Evolutionary dynamics of invasion and escape. *J. Theoret. Biol.* **226**, 205–214. (doi:10.1016/j.jtbi.2003.08.014)

33 Day, T., André, J.-B. & Park, A. 2006 The evolutionary emergence of pandemic influenza. *Proc. R. Soc. B* **273**, 2945–2953. (doi:10.1098/rspb.2006.3638)

34 Keeling, M. J. & Rohani, P. 2007 *Modelling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.

35 Athreya, K. B. & Ney, P. E. 1972 *Branching processes*. Heidelberg, Germany: Springer.

36 Burch, C. & Chao, L. 1999 Evolution by small steps and rugged landscapes in the RNA virus φ6. *Genetics* **151**, 921–927.

37 Mollison, D. 1991 Dependence of epidemic and population velocities on basic parameters. *Math. Biosci.* **107**, 255–287. (doi:10.1016/0025-5564(91)90009-8)

38 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.

39 Meiklejohn, G., Eickhoff, T., Graves, P. & Josephine, I. 1978 Antigenic drift and efficacy of influenza virus vaccines, 1976–1977. *J. Infect. Dis.* **138**, 618–624. (doi:10.1093/infdis/138.5.618)

40 Longini Jr, I. M., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cummings, D. A. T. & Elizabeth Halloran, M. 2005 Containing pandemic influenza at the source. *Science* **309**, 1083–1087. (doi:10.1126/science.1115717)

41 Longini, I., Halloran, M., Nizam, A. & Yang, Y. 2004 Containing pandemic influenza with antiviral agents. *Am. J. Epidemiol.* **159**, 623–633. (doi:10.1093/aje/kwh092)

42 World Health Organisation 2010 How recommendations are made on the composition of influenza vaccine. See http://www.who.int/entity/csr/disease/influenza/qanda_vaccinerechow.pdf.

43 Viboud, C., Alonso, W. J. & Simonsen, L. 2006 Influenza in tropical regions. *PLoS Med.* **3**, 468–471. (doi:10.1371/journal.pmed.0030089)

44 Bush, R. M., Bender, C. A., Subbarao, K., Cox, N. J. & Fitch, W. M. 1999 Predicting the evolution of human influenza A. *Science* **286**, 1921–1925. (doi:10.1126/science.286.5446.1921)

45 Recker, M., Pybus, O., Nee, S. & Gupta, S. 2007 The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proc. Natl Acad. Sci. USA* **104**, 7711–7716. (doi:10.1073/pnas.0702154104)