Research

The evolutionary dynamics of receptor binding avidity in influenza A: a mathematical model for a new antigenic drift hypothesis

Hsiang-Yu Yuan[1,2] and Katia Koelle[2,3]

[1]Program in Computational Biology and Bioinformatics, Duke University, Durham, NC 27708, USA
[2]Department of Biology, Duke University, Box 90338, Durham, NC 27708, USA
[3]Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

The most salient feature of influenza evolution in humans is its antigenic drift. This process is characterized by structural changes in the virus's B-cell epitopes and ultimately results in the ability of the virus to evade immune recognition and thereby reinfect previously infected hosts. Until recently, amino acid substitutions in epitope regions of the viral haemagglutinin were thought to be positively selected for their ability to reduce antibody binding and therefore were thought to be responsible for driving antigenic drift. However, a recent hypothesis put forward by Hensley and co-workers posits that cellular receptor binding avidity is the dominant phenotype under selection, with antigenic drift being a side effect of these binding avidity changes. Here, we present a mathematical formulation of this new antigenic drift model and use it to show how rates of antigenic drift depend on epidemiological parameters. We further use the model to evaluate how two different vaccination strategies can impact antigenic drift rates and ultimately disease incidence levels. Finally, we discuss the assumptions present in the model formulation, predictions of the model, and future work that needs to be done to determine the consistency of this hypothesis with known patterns of influenza's genetic and antigenic evolution.

1. Introduction

Human influenza viruses have long been known to evolve by antigenic drift, the process by which the gradual accumulation of mutations in the viruses' haemagglutinin (HA) surface glycoprotein results in evasion of host immunity [1]. Epidemiological models have therefore focused on understanding how this process affects influenza's ecological and evolutionary dynamics. The earliest models simply incorporated antigenic drift phenomenologically, through the consideration of susceptible–infected–recovered–(re)susceptible (SIRS) dynamics [2]. In recent analyses, these SIRS-type models have provided valuable insights into the drivers of influenza seasonality and the sources of influenza's interannual variability [3–5]. A more mechanistic approach for incorporating antigenic drift into epidemiological influenza models has been through the formulation of multi-strain models [6–13]. Although a subset of these models have looked at the consequences of antigenic drift on influenza's ecological dynamics, most of them have instead addressed how limited genetic and antigenic diversity can be maintained in the presence of high mutation rates. Whether phenomenological or mechanistic in structure, however, mathematical models of influenza, reflecting current understanding of influenza evolution, have commonly assumed that antigenic drift is a consequence of immune selection: viral strains harbouring mutations that reduce the ability of circulating antibodies to bind to the viral HA have higher fitness than strains without these mutations, leading to continual viral turnover [14].

In light of an increasing body of experimental studies [15–17], Hensley *et al.* [18] have recently questioned this model of antigenic drift, noting that viral escape from polyclonal antibodies by this mechanism would be exceptionally difficult. This is because escape mutants, having all the necessary epitope changes to allow for polyclonal immune escape, are extremely unlikely to arise within single hosts given current mutation rate estimates. In place of this model, they suggest that the evolutionary dynamics of influenza's HA are predominantly driven by cellular receptor binding avidity changes and that antigenic drift is a side effect of these mutational changes. Evolution can act on receptor binding avidity because this phenotype affects the rate at which viruses enter host cells, and thereby their ability to escape neutralization by circulating polyclonal antibodies. The authors support this new model of antigenic drift with passage experiments in mice: when passaged through immune mice, influenza A strains accumulate HA mutations that increase receptor binding avidity, with a subset of these mutations located in previously identified HA epitopes; when passaged through naïve mice, influenza A strains instead accumulate HA mutations that decrease receptor binding avidity, with a subset of these mutations again lying in known HA epitopes. Being appreciably different from the current model of antigenic drift, this new model may change our understanding of influenza's ecological and evolutionary dynamics. It may also affect the design of control strategies that aim to reduce disease incidence.

Although some of the dynamical consequences of this new antigenic drift model could presumably be intuited, others may be more difficult to predict. This is because the model, as verbally described, has nonlinear feedbacks between viral changes in receptor binding avidity, rates of antigenic drift and host immunity at the population level. Furthermore, it would be difficult for the verbal model to lead to quantitative predictions; this is particularly limiting when it comes to choosing between alternative disease-control strategies. Here, we therefore develop a mathematical model for the receptor binding avidity hypothesis outlined by Hensley and co-authors, with the assumption that selection acts solely on cellular receptor binding avidity. Through numerical simulation of the model, we show how epidemiological parameters, such as contact rates and host lifespans, affect receptor binding avidity levels and rates of antigenic drift. Finally, we use the model to quantitatively explore the consequences of alternative vaccination strategies on the rates of antigenic drift and ultimately on rates of disease incidence.

2. A mathematical formulation for the new model of antigenic drift

We formulate Hensley and co-authors' new model of antigenic drift mathematically by specifying an SIRS model, with hosts classified into discrete classes of susceptible hosts (S), infected hosts (I) and recently recovered and therefore temporarily immune hosts (R). We further subdivide each of these classes of hosts according to the cumulative number of times a host has been infected (including any current infections). The epidemiological dynamics are given by:

$$\frac{dS_0}{dt} = \mu N - \beta(0, V) \frac{S_0}{N} I_{tot} - \mu S_0,$$

and, for $i \geq 1$,

$$
\left.
\begin{aligned}
\frac{dI_i}{dt} &= \beta(i-1, V) \frac{S_{i-1}}{N} I_{tot} - (\gamma + \mu) I_i \\
\frac{dR_i}{dt} &= \gamma I_i - (\omega + \mu) R_i \\
\text{and} \quad \frac{dS_i}{dt} &= \omega R_i - \beta(i, V) \frac{S_i}{N} I_{tot} - \mu S_i,
\end{aligned}
\right\}
\tag{2.1}
$$

where μ is the birth/death rate, β is the transmission rate, γ is the recovery rate, ω is the rate of waning immunity, N is the total population size and I_{tot} is the total number of infected hosts, given by $\sum_{i=1}^{\infty} I_i$. The transmission rate $\beta(k,V)$ between an infected individual and a susceptible host depends on the number of times the susceptible host has previously been infected (k) as well as on the binding avidity of the virus (V) with which the susceptible host is being challenged. We let the transmission rate depend on the susceptible host's k because we use k as a proxy for immune status: individuals with a higher number of previous infections are assumed to have higher levels of circulating antibodies with which to counter a challenging infection, and thereby lower susceptibility to infection. A schematic of this model is shown in the electronic supplementary material, figure S1.

The model of antigenic drift described by Hensley and co-authors considers the evolutionary dynamics of two components of viral phenotype: the cellular receptor binding avidity of the HA protein and, as a side effect, viral antigenicity. Here, we first describe how we mathematically model receptor binding avidity dynamics given the SIRS model shown in equation (2.1). We then describe how we include antigenic drift within this framework and analyse the resulting evolutionary dynamics of this full model.

(a) Modelling the dynamics of cellular receptor binding avidity

Instead of modelling changes in receptor binding avidity within each infected individual, we consider how the mean receptor binding avidity changes at the population level using a quantitative genetics approach [19,20], an approach which assumes that the trait under selection has a unimodal distribution with narrow variance in the population. Further assuming that receptor binding avidity is the sole phenotype under selection, we can write that the rate of change of V, the mean receptor binding avidity, is proportional to its fitness gradient: $dV/dt \propto d(\text{fitness})/dV$. Fitness in ecological scenarios is commonly defined as the per capita growth rate [21], such that viral fitness in this case is simply given by the per capita growth rate of infected individuals, $(dI_{tot}/dt)/I_{tot}$. The rate of change of the mean receptor binding avidity is therefore given by

$$\frac{dV}{dt} = k_V \frac{d((dI_{tot}/dt)/I_{tot})}{dV}, \tag{2.2}$$

where the constant k_V quantifies the amount of genetic variance in receptor binding avidity. To simplify this expression, we first write out the total growth rate of infected individuals as the sum of the growth rates of all the infected classes:

$$\frac{dI_{tot}}{dt} = \sum_{i=1}^{\infty} \frac{dI_i}{dt} = \sum_{i=1}^{\infty} \left[\beta(i-1, V) \frac{S_{i-1}}{N} I_{tot} \right] - (\gamma + \mu) I_{tot}. \tag{2.3}$$
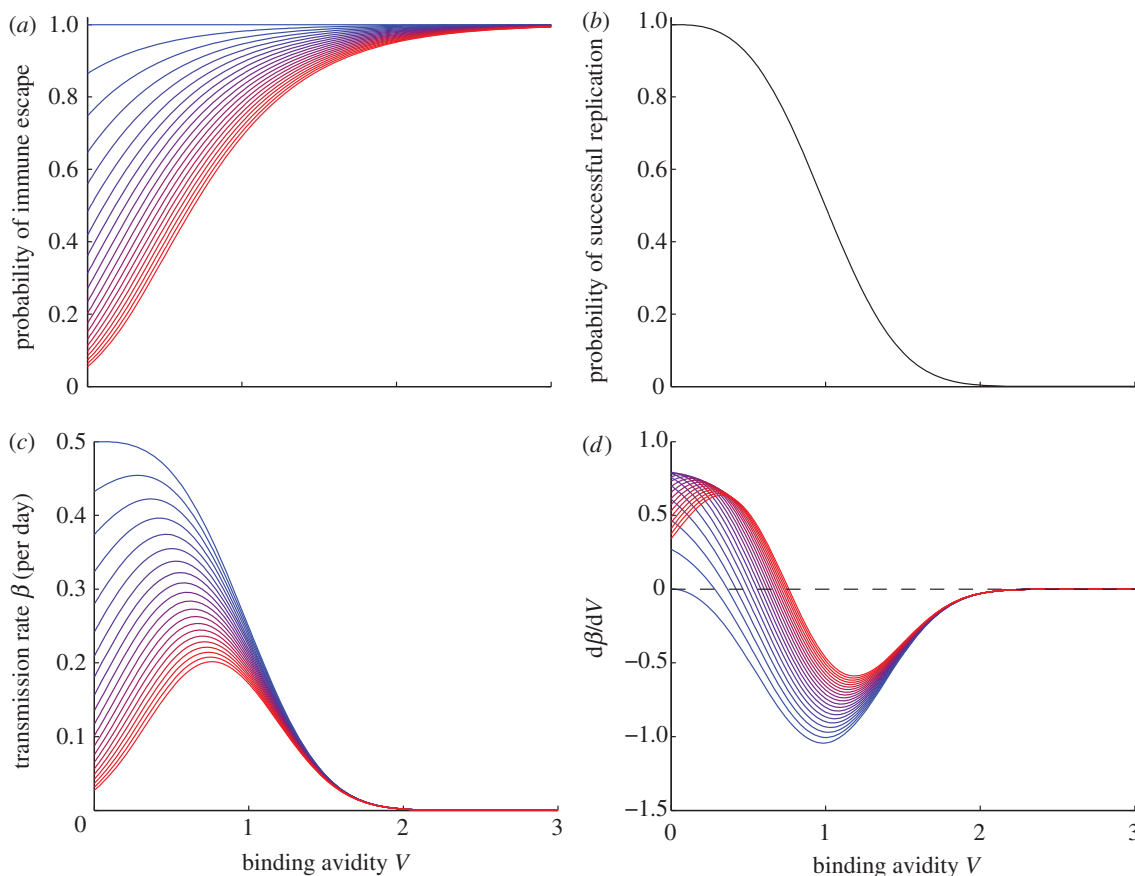
Figure 1. The effect of receptor binding avidity and previous exposure history on the transmission rate and its components. (a) The modelled relationship between binding avidity V and the probability that an infecting virus escapes neutralization by host antibodies, $f(k,V)$. Curves show $f(k,V)$, with k ranging from 0 (dark blue line) to 20 (red line). We let $f(k,V)$ be given by $[1 - e^{-p(V+1)}]^{rk}$, with positive constants r and p, where r controls the degree to which previous exposures reduce the probability of immune escape and p controls the degree to which binding avidity changes affect this probability. Here, we let $r = 1$ and $p = 2$. (b) The modelled relationship between binding avidity and the probability that a virus sufficiently replicates to cause a productive infection, $g(V)$. We let $g(V)$ be given by $g(V) = e^{-aV^b}$, with positive constants a and b, where a controls the rate at which this probability decreases with binding avidity and b controls the shape of this probability decrease. Here, we let $a = 0.7$ and $b = 3$. (c) The transmission rate $\beta(k,V)$, given by the product $cf(k,V)\,g(V)$. Curves are colour-coded as in (a). Contact rate $c = 0.5$ per day. (d) The derivative of the transmission rate with respect to binding avidity, $d\beta(k,V)/dV$. Curves are colour-coded as in (a).

The per capita infected growth rate is then:

$$\frac{(dI_{tot}/dt)}{I_{tot}} = \sum_{k=0}^{\infty} \left[\beta(k, V) \frac{S_k}{N} \right] - (\gamma + \mu), \tag{2.4}$$

where the subscript index has been changed to simplify notation. Taking the derivative with respect to V and substituting into equation (2.2) yields

$$\frac{dV}{dt} = k_V \sum_{k=0}^{\infty} \left[\frac{d\beta(k, V)}{dV} \frac{S_k}{N} \right]. \tag{2.5}$$

To describe the evolutionary dynamics of mean binding avidity V, we therefore need an expression for how the transmission rate β changes with V for any given number of previous infections k that a challenged susceptible host has experienced.

To derive this expression, we return to the passage studies performed by Hensley and co-authors, which indicate that increased binding avidity has benefits as well as costs. The benefit to increased binding avidity arises from an increased probability that circulating polyclonal antibodies in a host will fail to neutralize the virus. We therefore assume that, for a given individual having been exposed k previous times, the higher the binding avidity V of the virus, the higher the probability that the virus evades immune recognition. We also assume that, for a given binding avidity V, viruses will be less capable of evading immune recognition in individuals with a higher number of previous exposures k. This assumption reflects the observation that secondary immune responses result in more rapid and higher levels of antibody responses. Consistent with these patterns, figure 1a shows the probability that a virus evades neutralization by antibodies as a function of its own binding avidity and the number of previous infections the challenged host has experienced, $f(k,V)$.

The cost to increased binding avidity arises from a reduction in the ability of a virus to effectively replicate in a host (perhaps due to its slower release from infected host cells) and does not depend on the number of times a challenged susceptible host has previously been infected. We let the probability that a virus effectively replicates in a host be given by $g(V)$ (figure 1b).

Together, the product $f(k,V)g(V)$ yields the probability that a susceptible individual, having been infected k previous times, will become infected, given contact with a virus having binding avidity V. The overall transmission rate from an infected individual to a susceptible host with k previous infections is therefore $\beta(k,V) = c[f(k,V)g(V)]$ (figure 1c), where c is a positive constant parametrizing the contact rate per unit time.

Given this formulation, the basic reproduction number, defined as the expected number of secondary infections a

single infected individual will generate in a population of entirely susceptible, naïve hosts, is given by $R_0 = \beta(0,V)/(\gamma + \mu)$. Assuming that an infecting virus's mean binding avidity can quickly adapt to its optimal value in the susceptible host population, $R_0 = \beta(0,V=0)/(\gamma + \mu)$. This is because, as formulated, $V = 0$ is the optimal binding avidity for a virus infecting a naïve host.

The assumptions present in how the transmission rate and its components change with receptor binding avidity can be considered in light of the experiments described by Hensley and co-authors. In particular, viral mutants with higher receptor binding avidities have a greater ability to escape from polyclonal antibodies, as assessed by haemagglutination inhibition titre levels. This is consistent with our formulation of $f(k,V)$, shown in figure 1a: for a given level of circulating polyclonal antibodies (i.e. for a given number of previous exposures k), higher receptor binding avidity corresponds to a higher probability of immune escape. Additional experiments by Hensley and co-authors showed that viruses adapted to naïve hosts would, when passaged through immune mice, increase in their binding avidity. By contrast, viruses adapted to immune hosts would, when passaged through naïve mice, decrease in their binding avidity. Consistent with these patterns, figure 1c shows that viruses adapted to naïve (low k) hosts would increase their fitness in immune hosts by increasing their binding avidities, thereby allowing for more effective escape from polyclonal antibodies. This figure also shows that viruses adapted to immune (high k) hosts would increase their fitness in naïve hosts by evolving lower binding avidities (to increase their replication probabilities; figure 1b), thereby lowering their ability to escape from polyclonal antibodies (figure 1a).

We can now calculate the derivative of the transmission rate with respect to binding avidity V using our expression for $\beta(k,V)$ shown in figure 1c. We substitute this derivative, $d\beta(k,V)/dV$, shown in figure 1d, into equation (2.5).

The dynamics of the model including the evolutionary dynamics of mean binding avidity are therefore specified by equation (2.1) for the epidemiological dynamics, together with equation (2.5) for the mean binding avidity dynamics.

(b) Modelling the dynamics of antigenic drift

Until this point, we have considered only the evolutionary dynamics of influenza's cellular receptor binding avidity. Here, we further incorporate the evolutionary dynamics of viral antigenicity into our mathematical model. To do so, we turn to Hensley et al.'s [18] verbal model and the schematic of antigenic drift provided in their supplemental figure S7. According to these, the rate of antigenic drift (given by the rate at which escape from polyclonal antibodies occurs) should increase with an increase in the transmission frequency between naïve and immune individuals and should be higher the larger the change in binding avidity within an infected individual. Transmission frequency plays a role because the more transmissions occur per unit time, the more frequently the virus finds itself in a new host environment in which its binding avidity is suboptimal and in which it therefore evolves towards higher fitness by changing its cellular receptor binding avidity.

Because the epidemiological model we consider here does not only consider naïve ($k = 0$) and immune (high k) individuals, but a spectrum between them, we provide in the

electronic supplementary material, figure S2 a modified version of Hensley and co-authors' supplemental schematic. Our modification includes transmission of the virus between individuals with various previous exposure histories while retaining the salient features of the original schematic. Figure S2a in the electronic supplementary material shows how the receptor binding avidity of a virus transmitted from infected hosts to susceptible hosts may change over consecutive transmissions. A subset of the mutations contributing to changes in binding avidity also leads to changes in viral antigenicity (see the electronic supplementary material, figure S2b). To incorporate these changes in viral antigenicity into our epidemiological model (equation (2.1)), we interpret the rate of antigenic drift in terms of the parameter ω, the rate at which recovered (and immune) hosts become resusceptible: higher rates of antigenic drift correspond to higher rates of immunity loss. Consistent with Hensley and co-authors' schematic, the rate of immunity loss is no longer an independent parameter, but depends on the virus's epidemiological and evolutionary dynamics. In particular, it depends on the frequency of transmission events and the amount of antigenic change that occurs within infected hosts (as a side effect of evolutionary changes in binding avidity; electronic supplementary material, figure S2).

More formally, the rate of immunity loss can be written as $\omega = F\hat{A}$, where $F = \sum_{k=0}^{\infty} \beta(k,V)S_k/N$ is the frequency of transmission events (equivalent to $(\gamma + \mu)$ at endemic equilibrium) and $\hat{A}$ is the average amount of antigenic change that occurs within an infected individual. $\hat{A}$ can be further written as $\hat{A} = \sum_{l=0}^{\infty} \alpha_l \hat{A}_l$, where α_l is the proportion of transmission events that result in the infection of an individual who has previously been infected l times and $\hat{A}_l$ is the average amount of antigenic change that occurs within an infected individual who has previously been infected l times. α_l is given by $\beta(l,V)(S_l/N)/F$. Substituting into $\omega = F\hat{A}$ yields

$$\omega = \sum_{l=0}^{\infty} \left[\left(\beta(l,V)\frac{S_l}{N} \right)\hat{A}_l \right]. \tag{2.6}$$

With a fraction of receptor binding avidity mutations lying in HA epitopes, we let the average amount of antigenic change that accrues within an individual having previously been infected l times be proportional to the change in binding avidity within the individual:

$$\hat{A}_l = k_s|\Delta v_l|, \tag{2.7}$$

where Δv_l is the average change in receptor binding avidity that occurs in this class of infected individuals and k_s is a proportionality constant we use to map how changes in cellular receptor binding avidity alter viral antigenicity. Δv_l can in turn be written as $\Delta v_l = v_l(T) - v_l(0)$, where $v_l(t)$ is the receptor binding avidity at time t of the infection and $T = 1/F$ is the average amount of time an individual remains infected before transmitting the infection. On average, $v_l(0) = V$ for all l. $v_l(T)$ can be written as $v_l(T) = v_l(0) + \int_{t=0}^{T} (dv_l/dt)dt$. As before, we take a quantitative genetics approach, letting dv_l/dt be proportional to the receptor binding avidity fitness gradient within the host. In this case, a reasonable form for the fitness gradient is $d[f(l,v_l)g(v_l)]/dv_l$, where $[f(l,v_l)g(v_l)]$ is the probability that a virus with binding avidity v_l, residing in an infected individual who has been infected l previous times, yields a productive infection. Substituting, the

change in binding avidity Δv_l becomes

$$\Delta v_l = k_c \int_{t=0}^{T} \frac{\mathrm{d}[f(l,v_l)g(v_l)]}{\mathrm{d}v_l}\,\mathrm{d}t, \qquad (2.8)$$

where the constant k_c quantifies the amount of genetic variance in receptor binding avidity within a single host. With $v_l(0) = V$, we numerically solve Δv_l for all classes of infected individuals l. We then use these values to first calculate the average amount of antigenic change in a given class of infected individuals $\hat{A}_l$ (equation (2.7)) and then the rate of antigenic drift ω (equation (2.6)).

The full verbal model outlined by Hensley and co-workers can therefore be mathematically described by equation (2.1), where ω is given by equation (2.6), along with equation (2.5) describing the evolutionary dynamics of receptor binding avidity.

(c) The effect of epidemiological parameters on the rate of antigenic drift

To better understand the dynamics of the mathematical model developed above, we here determine, through numerical simulation, how some of the parameters of the epidemiological component of the model (equation (2.1)) affect the rate of antigenic drift. We specifically focus on the effects of three parameters: host lifespan ($1/\mu$), the contact rate (c) and the duration of infection ($1/\gamma$). We chose these parameters because they are the ones that are most likely to vary between species or can be affected by control policies such as quarantine measures and other non-pharmaceutical public health interventions. For each of these three sets of simulations, we keep constant all of the epidemiological parameters except for the one whose effect we explore. Along with showing how these parameters affect the rate of antigenic drift, figure 2 plots the equilibrium distribution of susceptible classes S_k/N and the equilibrium mean binding avidity V over the parameter ranges considered. We plot these two quantities because they are the main contributors to the rate of antigenic drift, interpreted as the rate of immunity loss ω (equation (2.6)).

Figure $2a-c$ shows the effect that host lifespan has on the equilibrium distribution of susceptible hosts S_k/N, on the equilibrium mean binding avidity V, and on the rate of antigenic drift ω, respectively. With longer host lifespans (and lower birth/death rates), the proportion of susceptible individuals in low-k classes decreases (figure $2a$). This is because longer-living individuals are able to accrue a higher number of previous infections. With the proportion of susceptible individuals in higher-k classes increasing with increasing host lifespans, the relative fitness of strains with higher receptor binding avidities increases. This in turn leads to higher mean binding avidities (equation (2.5) and figure $2b$). Figure $2c$ shows that the rate of antigenic drift ω increases (or, equivalently, that the duration of immunity decreases) with longer host lifespans. This can be interpreted in terms of the results shown in figure $2a,b$. Specifically, with short host lifespans, the S_k/N distribution is clustered around low-k classes, and the mean binding avidity V is low. The average change in viral antigenicity within an infected individual is therefore small: most individuals becoming infected will have had only a small number of previous infections and V has adapted to this situation. By equation (2.6), the rate of antigenic drift ω is therefore slow. In contrast, with long host lifespans, the S_k/N

distribution is broader (figure $2a$). Although V has also adapted to the average number of previous infections being higher, this broadening of the S_k/N distribution results in large amounts of antigenic change occurring in very low-k and very high-k infected individuals. This leads to higher rates of antigenic drift (figure $2c$). Although differences in the rates of antigenic drift (figure $2c$) will feed back to affect the S_k/N distribution (equation (2.1)), this feedback has negligible consequences on the equilibrium mean binding avidity (see the electronic supplementary material, figure S3). It is interesting to note here that this model therefore predicts that hosts with shorter lifespans would circulate viruses with lower rates of antigenic drift, all else being equal. Although there are arguably many more differences between swine hosts and human hosts than just their lifespans, this model would predict that immunity lasts longer in swine hosts (with lifespans of up to 5 years) than human hosts (with lifespans of approx. 70 years) if these other differences were ignored. This prediction is supported by recent work analysing the antigenic and genetic evolution of swine influenza A (H3N2) viruses: while the rates of genetic evolution in human and swine H3N2 viruses were similar, the rate of antigenic evolution in swine viruses was approximately six times slower than the rate in human viruses [22].

Figure $2d-f$ shows the effect that host contact rates have on the S_k/N distribution, on the mean binding avidity and on the rate of antigenic drift. Similar to the effect of longer host lifespans, the distribution of S_k/N changes from being clustered in low-k classes to being more evenly spread through these classes at higher host contact rates. This is because at higher host contact rates, susceptible individuals will become infected more rapidly, leading to a smaller proportion of susceptible hosts in low-k classes (figure $2d$). This again results in higher mean binding avidities (figure $2e$). For the same reasons as above, the rate of antigenic drift is therefore higher at higher contact rates (figure $2f$). As before, although differences in the rate of antigenic drift (figure $2f$) will feed back to affect the S_k/N distribution (equation (2.1)), this feedback has negligible consequences on equilibrium mean binding avidities (see the electronic supplementary material, figure S3). Interestingly, the results shown in figure $2d-f$ have implications for the control of influenza: public health measures that reduce contact rates (e.g. quarantine and improved hygiene measures) would not only reduce the number of infected individuals directly, but would have the indirect effect of slowing the rate of antigenic drift, thereby further reducing rates of disease incidence and extending the effectiveness of vaccines.

Finally, figure $2g-i$ shows the effect that the duration of infection has on the S_k/N distribution, on the mean binding avidity and on the rate of antigenic drift. Unlike in the previous cases, a longer duration of infection $1/\gamma$ leads first to higher and then lower mean binding avidities (figure $2h$). The increase in binding avidity at short durations of infection is because the basic reproduction number R_0 increases with an increase in $1/\gamma$ (we assume the transmission rate remains constant). This R_0 increase leads to an increase in the number of infected hosts, and therefore an increase in the force of infection. This leads to a shift in the S_k/N distribution to higher k classes (figure $2g$), with an increase in the average number of times susceptible hosts have previously been infected, and therefore selection for higher mean binding avidity (figure $2h$). As the duration of infection continues to
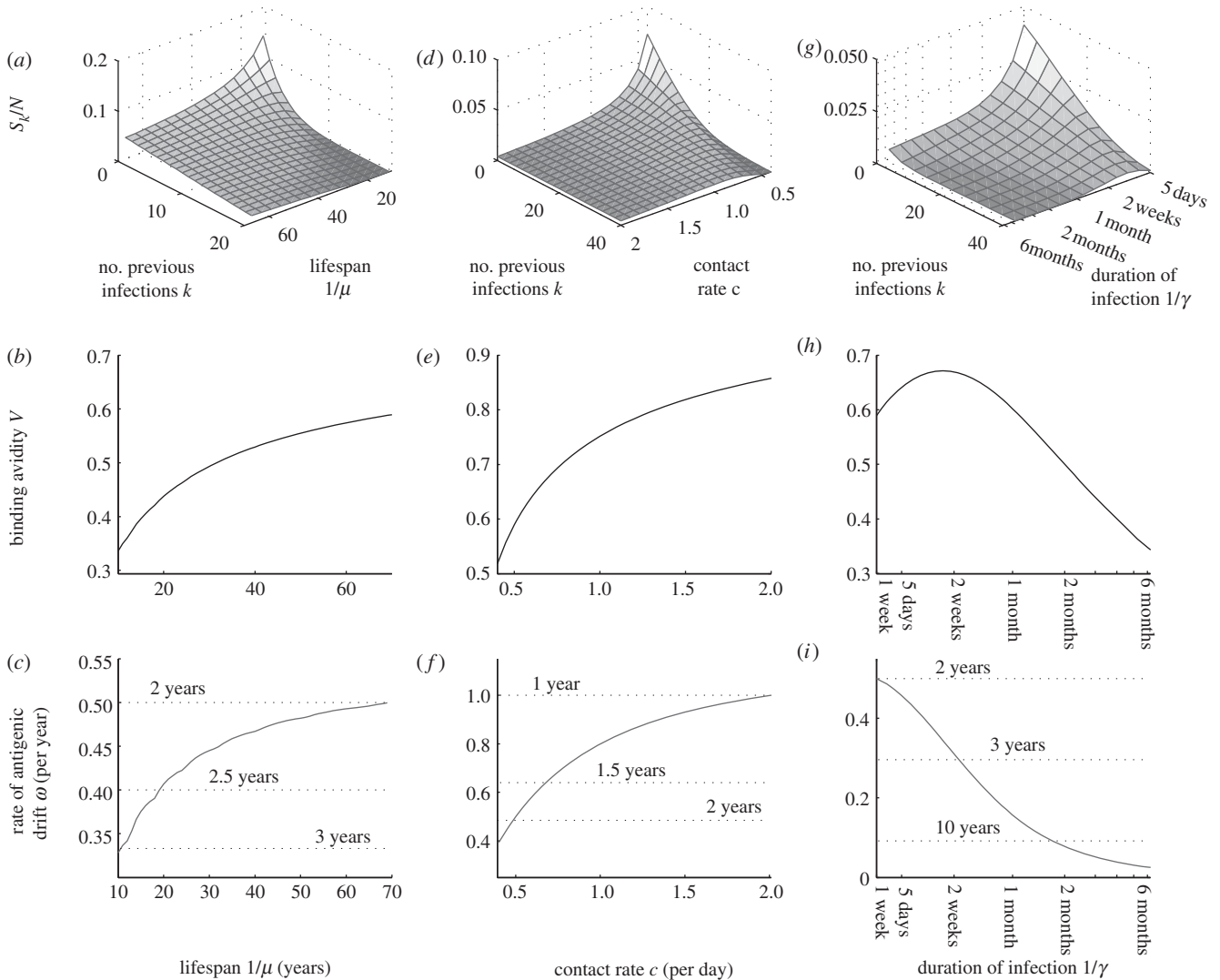
Figure 2. The effect of epidemiological parameters on the rate of antigenic drift. (a) The effect of host lifespan ($1/\mu$), on the equilibrium distribution of susceptible hosts, S_k/N. (b) The effect of host lifespan on the equilibrium mean receptor binding avidity V. (c) The effect of host lifespan on the rate of antigenic drift ω. (d) The effect of host contact rate on S_k/N. (e) The effect of host contact rate on binding avidity. (f) The effect of host contact rate on the rate of antigenic drift. (g) The effect of the duration of infection ($1/\gamma$) on S_k/N. See electronic supplementary material, figure S4 for the effect that the duration of infection has on S_k/S_{tot}, where $S_{tot} = \sum_{j=0}^{\infty} S_j$. (h) The effect of the duration of infection on binding avidity. (i) The effect of the duration of infection on the rate of antigenic drift. Epidemiological parameters are total population size $N = 100$ thousand, host lifespan $1/\mu = 70$ years (except in subplots $a-c$, where it is varied), host contact rate $c = 0.5$ contacts per day (except in subplots $d-f$, where it is varied), and duration of infection $1/\gamma = 5$ days (except in subplots $g-i$, where it is varied). Parameters p, r, a and b, used to parametrize the transmission rate β are as in figure 1. This parametrization leads to a basic reproduction number R_0 of 2.5. Parameters used in determining the amount of antigenic change occurring within infected individuals are $k_c = 0.09$ and $k_s = 0.0769$. The value of k_c was chosen such that viral binding avidity within infected hosts changed appreciably over the duration of infection, yet such that the optimal binding avidity within a host would not be reached. Given this value of k_c, the value of k_s was chosen such that the rate of antigenic drift ω became 0.5 year^{-1} when lifespan $1/\mu = 70$ years, the duration of infection $1/\gamma = 5$ days, and the contact rate $c = 0.5$ per day. The drift rate of 0.5 year^{-1} is a reasonable value for influenza A/H3N2. The parameter $k_V = 0.3$, but the equilibrium results do not depend on the value of this parameter. Horizontal lines with years marked in subplots (c), (f) and (i) correspond to $1/\omega$, the average duration of immunity.

increase, though, the proportion of susceptible hosts that belong to low-k classes starts to increase again (see the electronic supplementary material, figure S4). This non-intuitive effect is a result of two factors. First, it is because rates of antigenic drift slow down dramatically with longer infectious periods (figure 2i), such that non-infected hosts who have previously been infected linger in the recovered classes, rather than in the high-k susceptible classes. Second, it is due to infected individuals getting 'stuck' in the infectious class for a longer period of time. The return of predominantly low-k susceptible hosts results in selection for lower mean binding avidity. The importance of these two distinct factors can be seen by considering how mean binding avidity

changes with the duration of infection when the rate of antigenic drift is kept constant (see the electronic supplementary material, figure S3e,f); in this case, there is still a non-monotonic relationship between the duration of infection and mean binding avidity, although much longer durations of infection are needed to select for decreases in mean binding avidity. This is because only the second factor is at play in this case, resulting in a return in the predominance of low-k susceptible hosts only at very long infectious periods (see the electronic supplementary material, figure S5).

Finally, to understand why rates of antigenic drift decrease at longer infectious periods (figure 2i), it is easiest to return to the expression for the rate of antigenic drift

given by $\omega = F\hat{A}$. This expression shows that the rate of antigenic drift depends on the frequency of transmission events as well as on the average amount of antigenic change occurring within infected individuals. At equilibrium, the frequency of transmission events F is given by $(\mu + \gamma)$. Longer durations of infection $(1/\gamma)$ therefore lead to lower transmission frequencies. Assuming that the amount of antigenic change within an infected host does not increase considerably in the extra amount of time the host is infected, we would therefore expect the rate of antigenic drift to be slower with longer infectious periods. Our analysis of the duration of infection on the rate of antigenic drift illustrates the importance of formulating Hensley and co-authors' model of antigenic drift mathematically: without these simulations, it would have been difficult to predict that longer infectious periods would result in slower rates of antigenic drift. This prediction is also particularly interesting as it differs from the one arrived at under the commonly assumed model of antigenic drift in which several mutations in epitope regions of the viral HA have to occur to evade recognition by circulating polyclonal antibodies. In particular, this traditional model of antigenic drift would predict that the drift rate would be higher with longer infectious periods because individuals have to be infected for a long period of time for the virus to have the chance to generate an immune escape mutant. This juxtaposition in the direction of drift rate changes between the binding avidity model for antigenic drift and the traditional model of antigenic drift hints at an intriguing possibility: might there be some threshold for infectious duration, below which selection occurs on binding avidity changes and above which selection occurs directly on antigenic phenotypes? Although this might be the case, we note that a cellular receptor binding avidity model that did not have a cost associated with increased binding avidity, or one that allowed this cost to be overcome with compensatory evolution (perhaps occurring more frequently in individuals who were infected for a longer period of time), might also predict higher rates of antigenic drift with longer durations of infection.

(d) An application of the model to evaluate alternative vaccination policies

In §2c, we have seen that epidemiological factors will affect rates of antigenic drift. One of these factors, the host contact rate, was a particularly interesting parameter to consider because there is some, albeit limited, evidence showing that non-pharmaceutical disease-control strategies for influenza are able to affect this parameter [23,24]. We now consider instead the effect that vaccination—a pharmaceutical intervention—would have on rates of antigenic drift and ultimately disease incidence. In particular, we evaluate the effectiveness of two alternative vaccination strategies: a strategy that preferentially vaccinates children and a strategy that vaccinates individuals at random. We chose these two strategies in order to address a prediction made by Hensley *et al.* [18], namely that paediatric influenza virus vaccination would decrease the size of the naïve population and thereby slow rates of antigenic drift. As a consequence, the effectiveness of vaccines would be temporally extended [18].

To determine the dynamical consequences of these two control strategies on the rate of antigenic drift and disease incidence,

we extend our epidemiological model to incorporate vaccination:

$$\frac{dS_0}{dt} = \mu N - \beta(0, V)\frac{S_0}{N}I_{tot} - (\mu + p_0)S_0,$$

and, for $i \geq 1$,

$$\left.\begin{array}{l} \dfrac{dI_i}{dt} = \beta(i-1, V)\dfrac{S_{i-1}}{N}I_{tot} - (\gamma + \mu + p_i)I_i, \\[2mm] \dfrac{dR_i}{dt} = \gamma I_i - (\omega + \mu + p_i)R_i + p_{i-1}(S_{i-1} + I_{i-1} + R_{i-1}) \\[2mm] \text{and} \quad \dfrac{dS_i}{dt} = \omega R_i - \beta(i, V)\dfrac{S_i}{N}I_{tot} - (\mu + p_i)S_i, \end{array}\right\} \quad (2.9)$$

where p_k is the rate of vaccinating individuals with k previous infections. This formulation assumes that vaccination of susceptible individuals acts similarly to a natural infection: it results in temporary immunity and contributes an additional exposure. The formulation further assumes that vaccination of currently immune individuals also acts to contribute an additional exposure. Finally, it assumes that vaccination of currently infected individuals results in immediate recovery and an additional exposure. (This latter assumption greatly simplifies the fair comparison of vaccination strategies, as described below. Because there are few infected individuals at any point in time and because the parameters in our simulations are set such that $p_k \ll (\gamma + \mu)$ for all k, the results of our simulations, presented below, are not sensitive to this last assumption.)

We first consider the paediatric vaccination strategy. Being younger, children are likely to have had only a few previous infections. We therefore parametrize this vaccination strategy by setting vaccination rates p_k higher for lower k (figure 3 legend), such that overall, vaccination occurs at an annual rate of approximately 4 per cent. As anticipated by Hensley and co-authors, simulations of this model result in lower rates of antigenic drift compared with simulations without vaccination (figure 3a). Furthermore, this vaccination strategy results in an appreciable reduction in influenza's annual attack rate (figure 3b). The reason for the observed decrease in the rate of antigenic drift can be understood by considering the equilibrium S_k/N distribution (figure 3c) along with the equilibrium mean binding avidity V (figure 3d). As a consequence of the preferential vaccination of low-k individuals, the S_k/N distribution becomes narrower (figure 3c), whereas the mean binding avidity does not change appreciably (figure 3d; although it does increase, as expected, at higher vaccination rates; electronic supplementary material, figure S6). The significantly lower rate of antigenic drift that results from the paediatric vaccination strategy (figure 3a) is therefore due to the decrease in the number of low-k infections, which are the ones that contribute large antigenic changes.

We can now compare this paediatric vaccination strategy with one that vaccinates individuals at random. A random vaccination strategy has a vaccination rate that is independent of the number of times an individual has been previously infected, such that p_k is the same for all k. To fairly compare this vaccination strategy with the one that preferentially vaccinates children, we control for the total number of vaccines delivered (figure 3 legend). Simulations of this random vaccination scenario do not give rise to lower rates of antigenic drift (figure 3a). This is because neither the equilibrium S_k/N distribution (figure 3c) nor the equilibrium mean binding avidity (figure 3d) changes appreciably with this strategy. However, vaccinating at random does still decrease disease incidence moderately (figure 3b), although not to the same extent as the paediatric
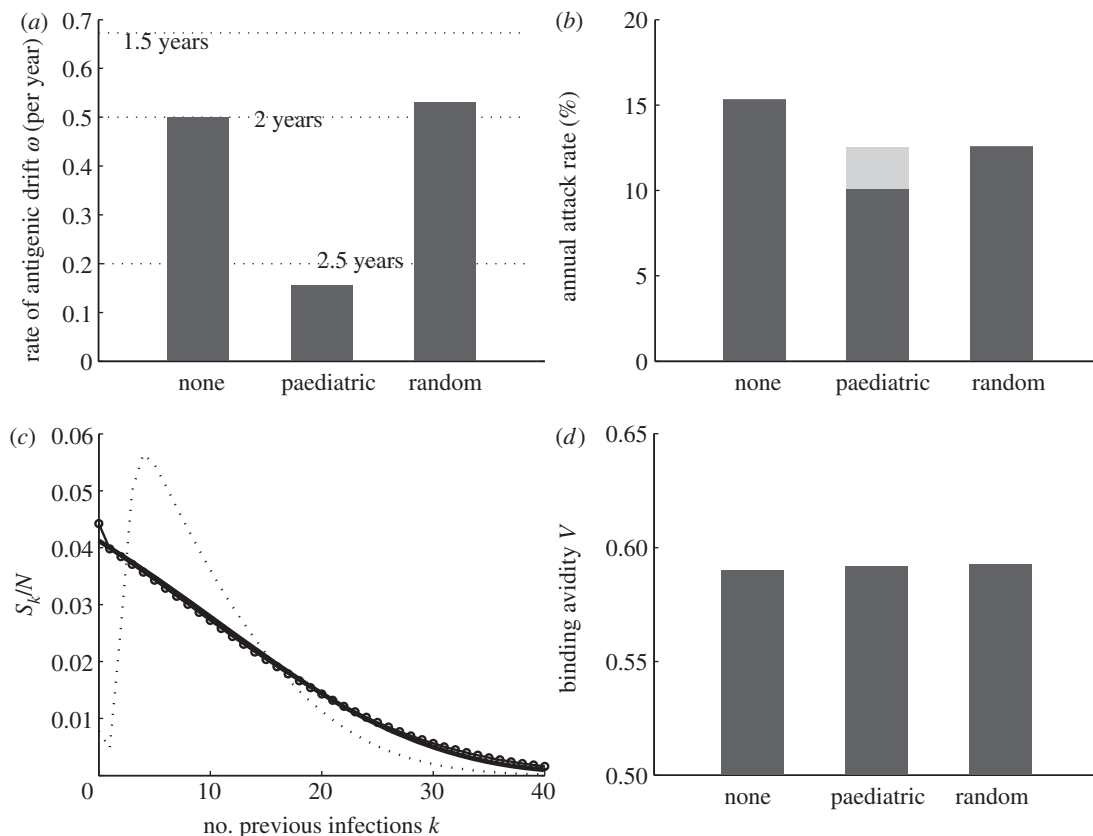
Figure 3. The effect of vaccination strategies on the rate of antigenic drift and disease incidence. (*a*) The equilibrium rate of antigenic drift under no vaccination, paediatric vaccination and random vaccination strategies. Horizontal lines with years marked correspond to $1/\omega$, the average duration of immunity. (*b*) The annual attack rate for influenza under these three strategies. The grey bar shows the annual attack rate under the paediatric vaccination strategy when the rate of antigenic drift ω is fixed at 0.5 year^{-1}. (*c*) The equilibrium S_k/N distribution under these three strategies. The solid bold line shows the distribution in the absence of a vaccination policy, the dotted line shows the distribution under the paediatric vaccination policy, and the line with circle marks shows the distribution under the random vaccination policy. (*d*) The equilibrium mean binding avidity under these three strategies. Parameter values used are as in figure 2. Vaccination rates, in units of year^{-1}, under the paediatric vaccination strategy are given by $p_k = m(1 - s)^k s$, with $m = 2$ and $s = 0.8$, resulting in vaccination rates that decrease with an increase in the number of previous exposures k. The vaccination rate, also in units of year^{-1}, under the random vaccination strategy is $p = 0.037$. These vaccination rates lead to the same total number of vaccinations delivered per unit time, with approximately 4% of the population vaccinated annually. The total number of vaccinations delivered was kept constant across the two vaccination strategies by first calculating the total number of individuals vaccinated per unit time for the paediatric vaccination strategy. This is given by $p_0 S_0 + \sum_{i=1}^{\infty} p_i(I_i + R_i + S_i)$. In the random vaccination scenario, the total number of individuals vaccinated per unit time is pN, where p is the vaccination rate that is independent of the number of previous exposures. In our simulations, we therefore chose p by dividing the total number of individuals vaccinated per unit time in the paediatric vaccination scenario by the population size N.

vaccination strategy. This is because vaccination still has the direct effect of protecting individuals from infection as well as the indirect effect of generating herd immunity.

The results of these simulations support Hensley and co-authors' predictions: vaccinating children would lead to lower rates of antigenic drift (figure 3*a*). Lower rates of antigenic drift bring two benefits. First, a lower rate of antigenic drift would temporally extend the effectiveness of vaccines [18]. Second, a lower rate of antigenic drift would reduce the annual attack rate, as it would take a longer amount of time for recovered individuals to become susceptible again, a prerequisite for becoming infectious. Indeed, we see in figure 3*b* that the annual attack rate is appreciably smaller in the case of paediatric vaccination compared with the case of random vaccination. However, it could be argued that the difference in annual attack rates between these vaccination strategies arises from the most susceptible individuals (children) being vaccinated in the first scenario, whereas vaccinations are 'wasted' on individuals who are less susceptible in the second scenario. That is, vaccinating a low-k susceptible individual should be more effective from a public health perspective than vaccinating a high-k susceptible

individual, because low-k susceptible individuals are always more susceptible than high-k ones (figure 1*c*). However, when we simulated the model under the paediatric vaccination strategy but did not allow the rate of antigenic drift to differ from the case without vaccination, the resulting annual attack rates were indistinguishable from those of the random vaccination strategy (figure 3*b*). This shows that it is the lower rate of antigenic drift under the paediatric vaccination policy, compared with the random vaccination strategy, that is responsible for the greater reduction in disease incidence.

It is interesting to compare these influenza vaccination strategies to those recommended by the US Centers for Disease Control and Prevention (CDC) Advisory Committee on Immunization Practices. In 2006, this committee recommended annual vaccination of children between the ages of six months to 5 years and persons over 50 years of age [25]. These recommendations were set to reduce the chance that individuals in these high-risk age groups get infected and develop serious influenza-related complications. In 2010, the committee modified its recommendation, urging that all persons older than six months be vaccinated annually [26].

If vaccinations were limited in number, the model presented here would argue that the vaccination policy recommended in 2006 might, if followed, lead to better health outcomes than the current policy because it would lower rates of antigenic drift and thereby result in a larger decrease in influenza incidence. However, if vaccinations are not limiting, the current policy would, if followed, increase the total number of vaccinations given annually and thereby increase herd immunity. This in turn would lower influenza incidence, perhaps below the level attainable by the previous policy.

3. Discussion

Here, we developed a mathematical model for the antigenic drift hypothesis recently presented by Hensley et al. [18]. Supported by passage studies in mice, this hypothesis argues that the antigenic drift of influenza A is predominantly a side effect of cellular receptor binding avidity changes. These changes are thought to accrue as the virus passes between individuals of differing immunity levels. We showed with this mathematical model that rates of antigenic drift are expected to be higher in populations with longer host lifespans, in populations with higher contact rates and in populations with shorter durations of infection. We have further used the model to evaluate the effects of two different influenza vaccination strategies: a strategy that focuses on vaccinating children, and a strategy that vaccinates individuals at random. In support of Hensley and co-authors' prediction, the paediatric vaccination strategy led to a lower rate of antigenic drift, whereas the random vaccination strategy did not appreciably affect the drift rate. These results stand in contrast to those made by previous influenza antigenic drift models [27], which predict that vaccination would increase the rate of antigenic drift. Correctly understanding the mechanism(s) driving antigenic drift will therefore have important consequences for accurately predicting the indirect effects brought about by control policies.

The mathematical model we have presented here contains several assumptions that still need to be more fully evaluated. One assumption is that increased cellular receptor binding avidity exacts a cost. Although passage experiments through naïve mice have shown that high avidity mutants evolve to lower avidity [18], thereby supporting the presence of a fitness cost, the mechanistic basis for such a cost is at this time unclear. One possibility is that the progeny of viruses with higher receptor binding avidities are not easily released from infected host cells, thereby resulting in low virion yields [28]. Efficient cleavage from the sialic acid receptor could, however, occur in adsorptive (high avidity) mutants that also had higher neuraminidase (NA) activity [29]. This opens up the possibility of compensatory NA mutations that could restore viral fitness of high avidity mutants. It also opens up questions about how this new antigenic drift hypothesis relates to the observation that a functional match between the HA and NA proteins needs to exist for a productive viral infection to occur [30,31]. Determining whether there is a cost to high binding avidity and whether/how it can be overcome is therefore important to understand in the context of this drift hypothesis.

A second assumption in our model is that host susceptibility depends on the number of times an individual has previously been infected. This assumption stands in contrast to the implicit assumption in SIRS models for influenza. By having only one susceptible class, these models implicitly assume that the number of times an individual has been infected has no bearing on the individual's susceptibility. Our assumption also differs from assumptions commonly made in multi-strain models for influenza. These often assume that the susceptibility of a host depends on the antigenic similarity between the challenging virus and the repertoire of previous strains the host has been infected with [6–8,10]. Finally, this assumption neglects the notions of original antigenic sin and antigenic seniority, which posit that a host's susceptibility depends on the antigenic differences between a challenging virus and the first virus, or first few viruses, the individual was ever exposed to [32,33]. We chose our model formulation specifically to be able to accurately accommodate the hypothesis put forward by Hensley and co-authors; because their hypothesis posits that antigenic drift is predominantly a side effect of binding avidity changes, we wanted, in a first model, to ignore the effects of antigenic strain variation altogether. However, as discussed below, considering the interaction between antigenic variation and binding avidity heterogeneity in the viral population might be important to understand the evolutionary dynamics of influenza A in humans and other hosts that exert immune pressure on the virus. We leave this possibility for future work.

A third assumption present in our model is that the amount of antigenic change occurring within a host is proportional to the amount of cellular receptor binding avidity changes within the host. Because the globular domain of the HA protein contains amino acid sites that affect cellular receptor binding as well as B-cell epitopes, with some overlap, this assumption is the simplest one to initially incorporate into a model. However, there is evidence that antibody binding is affected by certain amino acid residues located in the B-cell epitopes more than others [34]. Variation in the contribution of different amino acid residues to cellular receptor binding avidity is also likely. As of yet, the exact functional regions for antigenic sites and binding avidity sites, and their overlap, still need to be better identified to more fully understand the relationship between receptor binding avidity changes and antigenic changes. Statistical approaches that identify positively selected residues, such as the one recently developed by Tusche et al. [35], will be key to identifying which sites are likely to be relevant and may thereby provide good candidates for further study.

The mathematical formulation of Hensley and co-authors' antigenic drift model that we developed here can be used to make predictions that can be tested using publicly available influenza sequence data. For example, human hosts are generally long-lived compared with other influenza host species, and the duration of a typical influenza infection in humans is generally shorter or comparable to that of other host species. Given our results shown in figure 2, we would therefore anticipate that the cellular receptor binding avidities of influenza viruses circulating in humans should be higher than those circulating in other hosts. Following a pandemic emergence from either avian hosts or swine hosts, we therefore expect viral receptor binding avidities to increase over time in the human population. Indeed, Arinaminpathy & Grenfell [36] recently performed a retrospective analysis of influenza sequences, looking at how glycoprotein charge, which affects cellular receptor binding avidity in the case of the HA

protein, changes over time for several influenza A subtypes. In support of the above prediction, they show that the total charge of influenza A H3N2's HA increases from the time of its introduction in 1968 until 1986, when the total charge plateaus. However, this pattern is not present in influenza A subtypes H1N1 or H2N2. Furthermore, there are two confounders in interpreting this increase in H3N2 receptor binding avidity in humans according to the model presented here and a transition of the virus from a short-lived host to a long-lived host. First, the increase in H3N2 binding avidity in humans could be due to evolution of receptor specificity. Second, increases in receptor binding avidity in H3N2 may have evolved to compensate for the accumulation of glycosylation sites on the viral HA (which decrease receptor binding avidity) [37,38], rather than as a result of their stand-alone fitness advantage.

A second prediction made by the model, and the original hypothesis, is that currently infected hosts who have been infected more numerous times should, on average, harbour influenza viruses with higher binding avidities compared with currently infected individuals with fewer previous infections. This is because the model assumes that antigenic drift, occurring rather rapidly, is a side effect of frequent within-host changes in binding avidity (see the electronic supplementary material, figure S2), with higher binding avidity selected for in higher k individuals and lower binding avidity selected for in lower k individuals. Because an individual's age has been used as a proxy for the number of previous infections k, we therefore predict that the binding avidities of viruses isolated from patients (at least at the end of their infectious periods) should be an increasing function of patient age. This prediction is testable using age information available in the Influenza Virus Resource Database [39] and computational methods to predict binding avidity from viral sequence data. Here, we simply note that the prediction is one put forward by Hensley and co-authors' hypothesis, and that the mathematical formulation of their model presented here could perhaps be used to quantitatively predict the relationship between age and binding avidity, or, alternatively, that the observed relationship between age and binding avidity could be used to parametrize our mathematical model.

Finally, the model presented here is similar in structure to influenza models of the SIRS variety discussed in §1, albeit more complicated. In its current structure, it could therefore not address whether the new antigenic drift hypothesis is consistent with the ladder-like phylogeny of influenza's HA [40] and the emergence–replacement dynamics of influenza's antigenic clusters [41]. To determine whether this new antigenic drift model can reproduce these genetic and antigenic evolutionary dynamics, another formulation of the model would need to be developed. This formulation would be similar in structure to existing multi-strain influenza models [6–8,10–13], with the key difference being that strains would differ in their binding avidities along with their antigenic characteristics. For now, the new antigenic drift hypothesis presents an interesting and starkly different alternative to previous hypotheses of antigenic drift. A demonstration of its consistency with the evolutionary dynamics of influenza will surely make our commitment to it more 'binding'.

References

1. Nelson MI, Holmes EC. 2007 The evolution of epidemic influenza. *Nat. Rev. Genet.* **8**, 196–205. (doi:10.1038/nrg2053)

2. Pease CM. 1987 An evolutionary epidemiological mechanism, with applications to type A influenza. *Theor. Popul. Biol.* **31**, 422–452. (doi:10.1016/0040-5809(87)90014-1)

3. Dushoff J, Plotkin JB, Levin SA, Earn DJD. 2004 Dynamical resonance can account for seasonality of influenza epidemics. *Proc. Natl Acad. Sci. USA* **101**, 16 915–16 916. (doi:10.1073/pnas.0407293101)

4. Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M. 2010 Absolute humidity and the seasonal onset of influenza in the continental US. *PLoS Biol.* **8**, e1000316. (doi:10.1371/journal.pbio.1000316)

5. Casagrandi R, Bolzoni L, Levin SA, Andreasen V. 2006 The SIRC model and influenza A. *Math. Biosci.* **200**, 152–169. (doi:10.1016/j.mbs.2005.12.029)

6. Andreasen V, Lin J, Levin SA. 1997 The dynamics of co-circulating influenza strains conferring partial cross-immunity. *J. Math. Biol.* **35**, 825–842. (doi:10.1007/s002850050079)

7. Gog JR, Grenfell BT. 2002 Dynamics and selection of many-strain pathogens. *Proc. Natl Acad. Sci. USA* **99**, 17 209–17 214. (doi:10.1073/pnas.252512799)

8. Ferguson NM, Galvani AP, Bush RM. 2003 Ecological and immunological determinants of influenza evolution. *Nature* **422**, 428–433. (doi:10.1038/nature01509)

9. Andreasen V, Sasaki A. 2006 Shaping the phylogenetic tree of influenza by cross-immunity. *Theor. Popul. Biol.* **70**, 164–173. (doi:10.1016/j.tpb.2006.04.004)

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

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

12. Koelle K, Khatri P, Kamradt M, Kepler TB. 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)

13. Bedford T, Rambaut A, Pascual M. 2012 Canalization of the evolutionary trajectory of the human influenza virus. *BMC Biol.* **10**, 2–12. (doi:10.1186/1741-7007-10-2)

14. Earn DJD, Dushoff J, Levin SA. 2002 Ecology and evolution of the flu. *Trends Ecol. Evol.* **17**, 334–340. (doi:10.1016/S0169-5347(02)02502-8)

15. Yewdell JW, Webster RG, Gerhard WU. 1979 Antigenic variation in three distinct determinants of an influenza type A haemagglutinin molecule. *Nature* **279**, 246–248. (doi:10.1038/279246a0)

16. Yewdell JW, Caton AJ, Gerhard WU. 1986 Selection of Influenza A virus adsorptive mutants by growth in the presence of a mixture of monoclonal antihemagglutinin antibodies. *J. Virol.* **57**, 623–628.

17. Wang M-L, Skehel JJ, Wiley DC. 1986 Comparative analyses of the specificities of anti-influenza hemagglutinin antibodies in human sera. *J. Virol.* **57**, 124–128.

18. Hensley SE *et al*. 2009 Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. *Science* **326**, 734–736. (doi:10.1126/science.1178258)

19. Lande R. 1976 Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334. (doi:10.2307/2407703)

20. Abrams P. 2001 Modelling the adaptive dynamics of traits involved in inter- and intraspecific interactions: an assessment of three methods. *Ecol. Lett.* **4**, 166–175. (doi:10.1046/j.1461-0248.2001.00199.x)

21. Metz JAJ, Nisbet RM, Geritz SAH. 1992 How should we define 'fitness' for general ecological scenarios? *Trends Ecol. Evol.* **7**, 198–202. (doi:10.1016/0169-5347(92)90073-K)

22. de Jong JC et al. 2007 Antigenic and genetic evolution of swine influenza A (H3N2) viruses in Europe. *J. Virol.* **81**, 4315–4322. (doi:10.1128/JVI.02458-06)

23. Aledort JE, Lurie N, Wasserman J, Bozzette S. 2007 Non-pharmaceutical public health interventions for pandemic influenza: an evaluation of the evidence base. *BMC Public Health* **7**, 1–9. (doi:10.1186/1471-2458-7-208)

24. Bootsma MCJ, Ferguson NM. 2007 The effect of public health measures on the 1918 influenza pandemic in US cities. *Proc. Natl Acad. Sci. USA* **104**, 7588–7593. (doi:10.1073/pnas.0611071104)

25. Smith NM, Bresee JS, Shay DK, Uyeki TM, Cox NJ, Strikas RA. 2006 Prevention and control of influenza with vaccines: recommendations of the advisory committee on immunization practices (ACIP). *MMWR Recomm. Rep.* **55**, 1–42.

26. Fiore AE et al. 2010 Prevention and control of influenza with vaccines: recommendations of the advisory committee on immunization practices (ACIP). *MMWR Recomm. Rep.* **59**, 1–62.

27. Boni MF. 2008 Vaccination and antigenic drift in influenza. *Vaccine* **26**, C8–C14. (doi:10.1016/j.vaccine.2008.04.011)

28. Rudneva IA, Sklyanskaya EI, Barulina OS, Yamnikova SS, Kovaleva VP, Tsvetkova IV, Kaverin NV. 1996 Phenotypic expression of HA–NA combinations in human-avian influenza A virus reassortants. *Arch. Virol.* **141**, 1091–1099. (doi:10.1007/BF01718612)

29. Hensley SE, Das SR, Gibbs JS, Bailey AL, Schmidt LM, Bennink JR, Yewdell JW, de la Torre JC. 2011 Influenza A virus hemagglutinin antibody escape promotes neuraminidase antigenic variation and drug resistance. *PLoS ONE* **6**, e15190. (doi:10.1371/journal.pone.0015190)

30. Kaverin NV et al. 1998 Postreassortment changes in influenza A virus hemagglutinin restoring HA–NA functional match. *Virology* **244**, 315–321. (doi:10.1006/viro.1998.9119)

31. Wagner R, Matrosovich M, Klenk H. 2002 Functional balance between hemagglutinin and neuraminidase in influenza virus infections. *Rev. Med. Virol.* **12**, 159–166. (doi:10.1002/rmv.352)

32. Fazekas de St. Groth S, Webster RG. 1996 Disquisitions on original antigenic sin. I. Evidence in man. *J. Exp. Med.* **124**, 331–345. (doi:10.1084/jem.124.3.331)

33. Lessler J, Riley S, Read JM, Wang S, Zhu H, Smith GJD, Guan Y, Jiang CQ, Cummings DAT. 2012 Evidence for antigenic seniority in influenza A (H3N2) antibody responses in Southern China. *PLoS Pathog.* **8**, e1002802. (doi:10.1371/journal.ppat.1002802)

34. Liu Q, Hoi SC, Su CT, Li Z, Kwoh CK, Wong L, Li J. 2011 Structural analysis of the hot spots in the binding between H1N1 HA and the 2D1 antibody: do mutations of H1N1 from 1918 to 2009 affect much on this binding? *Bioinformatics* **27**, 2529–2536.

35. Tusche C, Steinbrück L, McHardy AC. 2008 Detecting patches of protein sites of influenza A viruses under positive selection. *Mol. Biol. Evol.* **24**, 1–9.

36. Arinaminpathy N, Grenfell B. 2010 Dynamics of glycoprotein charge in the evolutionary history of human influenza. *PLoS ONE* **5**, e15674. (doi:10.1371/journal.pone.0015674)

37. Das SR, Hensley SE, David A, Schmidt L, Gibbs JS, Puigbo P, Ince WL, Bennink JR, Yewdell JW. 2011 Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy. *Proc. Natl Acad. Sci. USA* **108**, E1417–E1422. (doi:10.1073/pnas.1108754108)

38. Kobayashi Y, Suzuki Y. 2012 Compensatory evolution of net-charge in influenza A virus hemagglutinin. *PLoS ONE* **7**, e40422. (doi:10.1371/journal.pone.0040422)

39. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, Ostell J, Lipman D. 2008 The influenza virus resource at the National Center for Biotechnology Information. *J. Virol.* **82**, 596–601. (doi:10.1128/JVI.02005-07)

40. Fitch WM, Bush WM, Bender CA, Cox NJ. 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)

41. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus ADME, Fouchier RAM. 2004 Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 371–376. (doi:10.1126/science.1097211)