The effects of strain heterology on the epidemiology of equine influenza in a vaccinated population

A. W. Park^{1,2*}, J. L. N. Wood¹, J. M. Daly¹, J. R. Newton¹, K. Glass^{1,2,3}, W. Henley¹, J. A. Mumford¹ and B. T. Grenfell²

¹Animal Health Trust, Lanwades Park, Kentford, Newmarket CB8 7UU, UK

²Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

³National Centre for Epidemiology and Public Health, The Australian National University, Canberra, ACT 0200, Australia

We assess the effects of strain heterology (strains that are immunologically similar but not identical) on equine influenza in a vaccinated population. Using data relating to individual animals, for both homologous and heterologous vaccinees, we estimate distributions for the latent and infectious periods, quantify the risk of becoming infected in terms of the quantity of cross-reactive antibodies to a key surface protein of the virus (haemagglutinin) and estimate the probability of excreting virus (i.e. becoming infectious) given that infection has occurred. The data suggest that the infectious period, the risk of becoming infected (for a given vaccine-induced level of cross-reactive antibodies) and the probability of excreting virus are increased for heterologously vaccinated animals when compared with homologously vaccinated animals. The data are used to parameterize a modified susceptible, exposed, infectious and recovered/resistant (SEIR) model, which shows that these relatively small differences combine to have a large effect at the population level, where populations of heterologous vaccinees face a significantly increased risk of an epidemic occurring.

Keywords: equine influenza; stochastic model; epidemiology; heterology

1. INTRODUCTION

Influenza viruses are a cause of highly contagious respiratory diseases in mammals (Turkington & Ashby 1998). The hierarchy of the virus has been extensively studied. There are three types of influenza virus (A, B and C) differing in their internal protein ensemble (Webster *et al.* 1992). Type A is the dominant type, affecting several mammalian species and causing the majority of epidemics (Cliff *et al.* 1986). The high mutation and evolutionary rates of the virus (Ferguson *et al.* 2003) mean that constant surveillance of outbreaks is necessary to monitor the viral evolution (termed antigenic drift) and vaccines are updated periodically (Mumford 1999; Plotkin *et al.* 2002). However, there is always a possibility that circulating strains and vaccine strains will be heterologous (i.e. they are immunologically similar but not identical). Strain heterology can arise through failure to predict circulating strains correctly or through delays in updating vaccine composition. Therefore, it is important to relate strain heterology thoroughly to the risk of infection at both the individual and the population levels. This information is vital for contingency planning in the event of there being a problem in obtaining a virus that is a close match to the circulating strain(s) for inclusion in vaccines. Also, from the general perspective of the ecology of infectious diseases, the impact of strain heterology on epidemic development and viral evolution is an under-explored area (Grenfell *et al.* 2004).

Influenza vaccination is used widely in humans, particularly in the elderly, where infection is a significant cause of increased mortality (Sprenger *et al.* 1993). It has also been practised in horses since the 1960s and is compulsory for racehorses in the UK and several other countries. However, cases of equine influenza in vaccinated populations are not uncommon (Newton *et al.* 2000a; Morley *et al.* 2000). For almost 20 years, all outbreaks of equine influenza have been attributed to one subtype of influenza A (H3N8). An interesting feature of this subtype is that *ca.* 15 years ago its strains diverged into two separate sublineages, which continue to co-circulate. For influenza A subtypes in man, usually only one lineage dominates for any length of time (Fitch *et al.* 1997; Ferguson *et al.* 2003). The probable cause of the divergence of the equine subtype is simply a spatial one, the majority of viruses in one lineage being isolated on the North American continent, and in the other on the European continent. Compared with human populations, equine populations are less mixed. Typically, only animals involved in elite competition and breeding travel extensively.

It is currently recommended that licensed equine influenza vaccines should contain current strains from both sublineages (OIE 1996) because, although they are of the same subtype, there is significant antigenic and genetic variation between viruses of the two sublineages and there is not complete cross-protection (Yates & Mumford 2000). Experiments conducted at the Animal Health Trust have aimed to establish the degree of cross-protection conferred by a vaccine from one sublineage against a challenge strain from the other sublineage (Daly *et al.* 2004). We use these data from individual horses to parameterize a mathematical model to examine the effects of heterology between vaccine and circulating strains at the

* Author and address for correspondence: Department of Mathematics and Statistics, York University, N520 Ross Building, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada (awp@mathstat.yorku.ca).

population level. The population that we consider is a typical race-training-yard thoroughbred population. We use high-quality data on yard sizes, age structure and vaccination, combined with the experimental data, to estimate the risk of an influenza epidemic in a realistic vaccinated population and, further, to elucidate the factors that determine risk.

Because training-yard populations are small, typically 20–150 horses (Bell 2002), it is important to use stochastic modelling, which incorporates chance effects (Iofescu 1973). This applies to population demography (e.g. purchase of individual horses into the yard) and also to the infection dynamics, where a chance event, such as the failure of an infectious horse to transmit the virus, may become important (Mollison 1977; Grenfell & Dobson 1995). Accordingly, we use stochastic formulations throughout this paper.

Using individual-level data that relate the probability of being protected from infection to antibody levels conferred by vaccination, we develop a modified susceptible, exposed, infectious and recovered/resistant (SEIR) model. The standard model, which defines hosts according to their infection status (SEIR), is modified to incorporate the effects of imperfect influenza vaccines on epidemic development. We show that heterology of the vaccine and circulating strains affects many aspects of the epidemiology and that the distribution of infectious periods can impact significantly on the probability of a large outbreak. Finally, through model simulation, we demonstrate how a mismatch in the vaccine and circulating strains drastically increases the risk of an outbreak at the population level, despite only relatively small differences being apparent at the level of the individual host.

2. DATA

(a) *Seroconversion and virus excretion*

When a host is infected with the influenza virus, the immune system responds in many ways. One of the key defences is the generation of antibodies that recognize haemagglutinin, a surface protein on the virus that allows the virus access to host cells (Hobson *et al.* 1972; Askonas *et al.* 1982). The virus becomes inactivated following antibody binding to key antigenic sites (Wiley *et al.* 1981; Dimmock 1984). The amount of these antibodies circulating in the blood in a host can be established by using a technique called single radial haemolysis (SRH) (Schild *et al.* 1975; Mumford *et al.* 1983). If a population of hosts is sampled prior to exposure to the virus then each host will have an antibody level (an unvaccinated host that has not previously been infected will have an antibody level of zero, and a vaccinated host should have some positive antibody level). If the same population is sampled after exposure to the virus then those hosts that were infected will show a marked increase in their antibody levels, indicating that the immune system responded to the infection (these hosts are said to have ‘seroconverted’). We define seroconversion as a significant increase in antibody level between an acute sample (taken at the time of infection) and a convalescent sample (taken approximately two weeks after infection).

At the Animal Health Trust, three experiments have been conducted in groups of influenza-naïve ponies

(Yates & Mumford 2000; Daly *et al.* 2004). Because it has been ascertained that the animals have not previously been vaccinated or infected (as they are young animals from an isolated location and all tested negative for pre-existing antibodies to equine influenza), these data provide a unique opportunity to study the effects of strain heterology without the complicating factor of the presence of antibodies from previous vaccination and/or natural infection. Each experiment comprised three groups of horses: an unvaccinated control group, a group vaccinated with the same viral strain as the challenge strain (termed the ‘homologous group’) and a group vaccinated with a current viral strain from the alternative sublineage (termed the ‘heterologous group’). All animals were vaccinated twice (four weeks apart) with inactivated-virus vaccine containing 50 µg of haemagglutinin. The animals were challenged with a standardized dose of virus two weeks after the second vaccination. We amalgamated the data from these experiments resulting in 57 animals in the homologous vaccine group, 35 animals in the heterologous vaccine group and 32 animals in the control group.

All control animals seroconverted and excreted virus. However, for both the homologous and heterologous groups, seroconversion showed a marked dependence on pre-exposure antibody levels. To use antibody status in the modelling work and also to summarize the functional dependence of the probability of seroconverting on antibody level, logistic regression models were fitted to these data. Logistic regression models are natural models to use in situations where a binary response (e.g. either ‘did seroconvert’ or ‘did not seroconvert’) depends on a continuous explanatory variable (e.g. antibody level). In summary, the models enable the fitted probability of seroconverting to be high when antibody level is low and vice versa. They also assume a gradual change in the probability of seroconverting as antibody level changes. Explicitly, the models are of the form $\text{Pr}(\text{seroconverting}) = 1/(1 + \exp(a + bx))$, where x is the antibody level. The parameter a determines the probability of seroconverting when the antibody level is zero, and the parameter b adjusts how quickly the probability falls as antibody level increases. The parameters a and b are estimated using maximum-likelihood techniques (Eliason 1993). Fitting the logistic regression models suggests that the probabilities of seroconverting are not the same for the heterologous and homologous vaccine groups (figure 1).

Ponies in the heterologous group have an increased, though not statistically significantly so, risk of becoming infected ($p = 0.164$). Moreover, the discrepancy between the two groups is greatest for intermediate antibody levels, which are the levels most frequently attained by animals following vaccination (Park *et al.* 2003). Although it is not surprising that the largest discrepancies occur for intermediate antibody levels (as the logistic regression curves are bounded between probabilities of zero and one), the fact that these intermediate levels agree with independent estimates of the antibody levels attained by vaccinated animals suggests that typical homologously and heterologously vaccinated animals may have different probabilities of seroconverting. Illustratively, the largest discrepancy in the probabilities of seroconverting between the two groups occurs at an antibody level of 138 mm² (SRH) where the

Figure 1. Logistic regression model fitted to the data for homologous and heterologous groups, showing the dependence of the probability of seroconverting on the pre-exposure antibody level. The model fitted is $\text{Pr}(\text{seroconverting}) = 1/(1 + \exp(a + bx))$, where x is the antibody level. For the homologously vaccinated population (dotted line) $a = -6.743$ and $b = 0.05277$. For the heterologously vaccinated population (solid line) $a = -13.634$ and $b = 0.08779$.

probabilities for the homologous and heterologous groups are 0.38 and 0.83, respectively. Following the method of Casagrande *et al.* (1978), assuming a common sample size, n , for the two groups then the difference between these two probabilities is statistically significant (at the 95% confidence level, 80% power level) for values as low as $n = 22$. This indicates the feasibility of obtaining a statistically significant result to strengthen the existing suggestion from the data that heterologously vaccinated animals are more likely to seroconvert following exposure than homologously vaccinated animals. This is an effect that has been noted previously in equine influenza (Yates & Mumford 2000) but has not been quantified. Importantly, the available data have allowed the probabilities of seroconverting to be estimated for the two distinct groups.

It is well documented that imperfect equine influenza vaccines can reduce the symptoms in infected animals (Park *et al.* 2003). Indeed, these experiments have shown that seroconversion does not necessarily imply that an animal will excrete virus. All control animals seroconverted and excreted virus. However, some vaccinated animals seroconverted but did not excrete detectable virus. The probabilities of excreting detectable virus (given that an animal seroconverted) are slightly different for the two vaccine groups, with those animals that received the heterologous vaccine being more likely to excrete virus ($p = 0.8$ compared with $p = 0.73$ for the homologous-vaccine group; 95% confidence intervals are [0.61, 0.92] for the heterologous group and [0.54, 0.88] for the homologous group). No significant correlation between antibody level and this probability was observed.

(b) Latent and infectious periods

From the experimental-challenge studies it is possible to estimate the latent and infectious periods for both the homologous and heterologous groups. On each day, the

animals were sampled once using nasopharyngeal swabs to determine the presence or absence of detectable virus. As the time of challenge was known, the latent period was estimated as the time elapsed between the challenge and the first sign of virus excretion. Similarly, the number of inclusive days between the first and last signs of detectable virus was used as an estimate of the infectious period. Analysis of the latent and infectious periods for the vaccinated groups reveals some important features (figure 2). The distributions of the latent periods for both groups are similar, as are the distributions of the infectious periods for both groups. In fact, the latent periods are also quantitatively very similar and are characterized by the majority of animals having a 2 day latent period, with a few animals having latent periods of 1, 3 and 4 days. The distributions of the infectious periods are effectively bimodal, with one peak corresponding to a short infectious period of 1 day and another peak corresponding to a significantly longer infectious period, which is 4 days in the homologous group and 5 days in the heterologous group. Also, the maximum infectious periods recorded for the homologous and heterologous groups were 5 and 6 days, respectively. These results suggest that the reduced specificity of the antibodies causes an increase in the time taken to eliminate the infection. From a modelling aspect, another striking feature of these distributions is that they bear no resemblance to the exponential distributions commonly assumed for latent and infectious periods, particularly when using standard Monte Carlo techniques (Bartlett 1961), although there are some notable exceptions where alternative distributions have been considered in models (Anderson & Watson 1980; Keeling & Grenfell 1998; Lloyd 2001*a,b*). Indeed, the empirical distributions shown in figure 2 do not bear a close resemblance to any standard parametric distributions, and, consequently, we have avoided trying to describe the distributions for the latent and infectious periods in this way.

(c) Transmission rates

Previously, the transmission rate, β , for equine influenza in an unvaccinated population has been estimated to be 1.85 days^{-1} (Glass *et al.* 2002), which corresponds to a basic reproductive number (Diekmann *et al.* 1990) of $R_0 = 10.18$. The transmission rate was estimated by fitting average epidemic curves generated from a standard stochastic SEIR model to data from a published outbreak (Scholtens & Steele 1964). The transmission rate used in the stochastic model was varied to allow comparison of the model predictions with the data. The discrepancy between the model and the data was measured as the sum of the squares of the differences between the numbers of new cases predicted by the model and shown by the data on each day. The transmission rate that minimized this difference was then selected as the best estimate.

Although there are some data available for outbreaks in vaccinated populations, they are sporadic and often without details of antibody levels and the number of uninfected animals, both of which are necessary when trying to estimate the transmission rate in a vaccinated population. Here, we assume that the transmission rate is unaffected by vaccination and use the same value of $\beta = 1.85$. This results in effective reproductive numbers (Anderson & May 1991; Glass *et al.* 2002) of $R^* = 2.497$ for the

Figure 2. Distributions of (a,c) latent and (b,d) infectious periods for (a,b) homologously and (c,d) heterologously vaccinated animals.

homologously vaccinated population and $R^* = 4.948$ for the heterologously vaccinated population. These numbers represent a reduction when compared with unvaccinated animals but, significantly, are greater than 1.0, which is consistent with the fairly frequently observed outbreaks in vaccinated populations (Newton *et al.* 2000a; Morley *et al.* 2000; Wood *et al.* 2003). The effective reproductive numbers were calculated using the formula $R^* = \beta S/gN$, where S/N is the fraction of hosts that are susceptible (estimated as the proportion of animals that seroconverted upon challenge in each group), N is the population size given by $N = (S + E + I + R)$ (see figure 3) and $1/g$ is the mean infectious period for each group.

3. THE MODEL

(a) Population demography

In the UK, thoroughbred racehorses are vaccinated against influenza in accordance with Jockey Club rules. Effectively, it is recommended that young horses be vaccinated twice a year and older horses be vaccinated annually (Newton *et al.* 2000a). This means that the age structure of a yard population must be taken into consideration. From previous studies conducted at the Animal Health Trust, there are good-quality data concerning how antibody levels in horses of different ages change over time under the Jockey Club's vaccination policy (Newton *et al.* 2000a,b). Consequently, for any given population, it is relatively straightforward to ascertain the dynamics of the distribution of antibody levels over time.

The population we consider here is that of a typical flat-race-training yard. It comprises approximately 100 horses. Changes in the population over the course of a year reflect the management practices in such a yard (e.g. seasonal sale of older horses and purchase of younger horses). The

characteristics of the population are presented in more detail elsewhere (Park *et al.* 2003).

(b) Dynamics of infection

Previous modelling work on equine influenza has shown that standard SEIR models are not appropriate when considering imperfect vaccines (Park *et al.* 2003). We further develop the approach not only by taking into account the relationship between pre-exposure antibody level and seroconversion (figure 1) but also by considering the distributions of the latent and infectious periods as well as the probability of seroconverting but not excreting detectable virus.

To determine the initial status of the horses, their antibody levels are drawn from a previously parameterized distribution based on the current vaccination policy in the UK. Details of this distribution are presented elsewhere (Park *et al.* 2003). The distribution is based on data from over 600 animals (Newton *et al.* 2000a,b), allowing the distribution of typical antibody levels to be characterized for each week of the year. Because the normal distributions occasionally allow a negative antibody level to be sampled, antibody levels are sampled until a non-negative value is obtained. This occurs only occasionally and therefore the sample distributions are not significantly different from the underlying normal distributions.

Using the logistic regression curves derived (see figure 1), an antibody level is then converted to a probability of seroconverting upon challenging. A binomial trial for each horse (with the relevant probability) is performed to determine whether it should be counted as a 'seroconverter' or not. We interpret those animals that do not seroconvert as being protected at the onset of the epidemic, and accordingly we consider them resistant to infection and they are put in the R class of the SEIR system. Naturally,

Figure 3. Flow diagram showing all the possible transitions in the model along with their rates. A horse may be susceptible (S), exposed (E), infectious (I) or recovered/resistant (R). The population size is given by $N = (S + E + I + R)$ and the parameters are: β , transmission rate; $1/a$, latent period; $1/g$, infectious period; p , probability of excreting virus (given seroconversion).

horses that seroconverted are counted as initially susceptible and put in the S class.

Our data have shown that once an animal is exposed it will not necessarily become infectious and therefore an animal may follow an $S \rightarrow E \rightarrow R$ pathway as opposed to the more commonly assumed $S \rightarrow E \rightarrow I \rightarrow R$ pathway. The probability of becoming infectious, p , has been estimated for the homologous and heterologous groups (see § 2a).

The discrete latent and infectious periods (figure 2) were integrated into the continuous-time stochastic model in the following way. From the distributions, whenever a period, L (measured in days), was sampled it was recalculated as $L = L + (U - 0.5)$, where U is a variate drawn randomly from the uniform distribution in the interval between 0 and 1. This is to prevent any results being purely an artefact of the oversimplified approach of using only integer values for the periods.

Following de Jong & Kimman (1994), we assume a mass-action transmission term. This is based on the heuristic idea that training yards with large populations tend to be more spacious and, therefore, there is no reason to assume a strong crowding effect on the transmission of infection. As the populations we consider here are effectively constant over the year, alternative formulations of the transmission term (e.g. pseudo-mass action) will not have a marked effect on model results (de Jong & Kimman 1994; McCallum *et al.* 2001). A schematic diagram of all the possible transitions is given together with the rates at which they occur (figure 3).

4. RESULTS

(a) *Exponential versus empirical distributions for the latent and infectious periods*

We show how the probability of a large outbreak occurring depends on the distributions of the latent and infectious periods and on the initial number of infected animals. A population of 100 horses was used and in each replicate simulation the antibody level of each horse was drawn from a normal distribution (mean = 114.9, s.d. = 44.4) that was based on a sample of horses from a training yard.

The probability of a large outbreak occurring was estimated from 10 000 replicate simulations. The 10 000

Figure 4. Probability of a large outbreak as a function of the number of initially infected animals, calculated using exponential distributions (filled circles) and empirical distributions (open circles) for both the latent and infectious periods.

replicates give rise to a distribution of epidemic sizes that is bimodal (Anderson & Watson 1980; Näsell 1995). The lower mode represents simulations where the infection faded out of the population without causing a large epidemic (owing to a break in the chain of transmission). The higher mode represents those simulations where the epidemic took off. The probability of a large epidemic occurring was estimated by considering the relative sizes of the two modes.

Two different versions of the model (figure 3) were used: a standard Monte Carlo stochastic simulation and a data-driven alternative. The Monte Carlo stochastic simulation implicitly assumes that the latent period ($1/a$) and the infectious period ($1/g$) are exponentially distributed (Kermack & McKendrick 1927; Diekmann & Heesterbeek 2000). The alternative model differs only in the distributions of the latent and infectious periods: instead of using the exponential distributions, the empirical distributions for the homologous group, shown in figure 2, are used. Both models have the same mean values for the latent and infectious periods. Consequently, many features of the epidemiology are conserved (e.g. epidemic duration, size of population infected). However, the models differ significantly in the probability of a large outbreak. The traditional Monte Carlo stochastic simulation significantly underestimates the probability of a large outbreak occurring when compared with the stochastic model using latent and infectious periods based on data (figure 4). The discrepancy is most marked when the epidemic is initiated from one infected animal, which is a plausible scenario for equine influenza in training yards.

(b) *Risk of an outbreak for homologously and heterologously vaccinated populations*

The following results are estimations of the probabilities of epidemics of certain sizes occurring (at least 5% and at least 20% of the population). For each of the two threshold epidemic sizes, 20 different populations were simulated to capture the demographic variability of yard

Figure 5. Probability of an outbreak affecting (a) at least 5% and (b) at least 20% of a yard population as a function of the week of the year in which the initial case occurred. The risk for a homologously vaccinated population is shown by a dotted line and the risk for a heterologously vaccinated population is shown by a solid line.

populations. For each population, 20 different binomial trials were performed to incorporate the chance of different outcomes for initial conditions. Finally, 20 different epidemics were simulated for each initial condition (because of the chance effects in the stochastic process of the spread of infection). Therefore, a total of 8000 replicate simulations were used to estimate the probability of an epidemic of a certain size occurring. All epidemics were initiated by changing the status of one horse from susceptible to infectious. The model ran until the infection had died out of the population, after which the size of the epidemic was recorded.

Based on the results from § 4a, we use the version of the model that assumes latent and infectious periods drawn from the empirical distributions presented in figure 2. The key questions we address are as follows.

- (i) If equine influenza were recruited into the yard, what is the probability of an epidemic affecting $x\%$ of the population?
- (ii) How does this probability depend on the homology or heterology of the vaccine and circulating strains?
- (iii) How does this probability depend on the week of the year in which the epidemic is initiated?

The last question (dependence on time) is of interest because the vaccination schedule for thoroughbreds means that at some times of the year antibody levels are much lower than at other times of the year. It is possible to consider the risk of epidemics of any given size occurring. However, here we restrict our investigation to the probabilities of epidemics affecting 5% and 20% of the yard population. The reason we are interested in small outbreaks as well as large outbreaks is that small outbreaks have been shown to occur frequently (Glass *et al.* 2002) and are potentially important for maintaining influenza in the population.

The results show that the risk of an outbreak is much higher for a heterologously vaccinated population than for an equivalent homologously vaccinated population (figure 5). The risk of an epidemic in each population changes throughout the year as a result of the temporal dynamics of antibody levels. Increases in antibody levels following vaccination, especially around weeks 6 and 31 (Park *et al.* 2003), correspond to drops in the risk of an epidemic occurring. Conversely, antibody levels fall owing to antibody decay, after which the risk of an epidemic occurring increases.

The probability of an epidemic in a heterologously vaccinated population is less sensitive to the time of the year in which the epidemic is initiated. For example, for epidemics affecting at least 20% of the yard population, the minimum and maximum probabilities of an epidemic occurring at some point in the year are 7.5×10^{-4} and 0.76, respectively, for the homologously vaccinated population (covering a range of 0.76) and 0.42 and 0.89, respectively, for the heterologously vaccinated population (covering a range of 0.46).

5. DISCUSSION

Heterology of the vaccine and circulating strains of equine influenza affects the epidemiology of the virus in a variety of ways. Although further experiments and studies are required to test for statistical significance, the available data suggest that animals given a heterologous vaccine to the circulating strain have a higher probability of becoming infected at a given level of pre-exposure antibodies (figure 1). When antibody levels are very low then both types of vaccinees are at risk and there is no discernible difference in the probability of infection between the two groups (the probability is close to one in both cases). Also, when antibody levels are very high then both groups are protected (shown by the probability of becoming infected being close to zero for both groups). The available evidence suggests that if there is sufficient antibody conferred by a heterologous vaccine then it is apparently capable of producing the same degree of protection as the strain-specific antibody generated by a homologous vaccine. Further experiments should be done to clarify this effect. However, most vaccinees attain only intermediate antibody levels. It is in this region that the difference between the homologous and heterologous vaccines is most marked.

We have assumed that the transmission rate is equal to that previously estimated for unvaccinated animals (Glass *et al.* 2002). Future work will aim to establish the effect, if any, of vaccination on the transmission rate, as has been observed for a few other communicable diseases, notably Aujeszky's disease virus in pigs (de Jong & Kimmann 1994) and HIV (Longini *et al.* 1996). If the transmission rates for the homologous and heterologous groups were different, it is much more likely that the heterologous group would have the greater transmission rate, in which case the difference between the risk profiles of the two vaccine groups (figure 5) would further increase.

We have also assumed a true mass-action term for the transmission of infection. We have based this on the heuristic idea that yards with larger populations tend to be more spacious. Although this is a reasonable assumption

for thoroughbred populations, careful consideration would have to be given to how transmission may scale with population size when dealing with other populations, as it is well documented that the way in which transmission is modelled can impact significantly on the results of these models (de Jong & Kimmman 1994; McCallum *et al.* 2001).

The problems associated with estimating the transmission rate in a vaccinated population demonstrate the need for research to continue in this area. The difficulty in estimating the transmission rate would be helped by a rapid response to outbreaks, allowing data such as acute antibody level to be established. Often the focus is on the infected and infectious animals, but data on the hosts that escape infection are equally important. A possible experimental method to determine the transmission rate would be to try to establish the total amount of virus excreted throughout the infectious period. Current data of this type tend to involve infrequent sampling and can be sensibly used only to confirm virus excretion qualitatively. However, the problem remains of how to relate the amount of virus excreted to the transmission rate. There may well be a saturation effect, whereby further increasing the amount of virus excreted has no effect on the transmission rate. The resulting models are likely to be more complex. For example, they may incorporate the degree of infectivity, which would require a move away from the traditional SEIR compartment models.

The latent period does not appear to be affected by the type of vaccine administered. However, the infectious period has a tendency to increase when the vaccine strain is heterologous to the challenge strain (figure 2*b,d*). Generally, the infectious period has a bimodal distribution, with some horses having only a 1 day infectious period and the remainder having a longer infectious period. For those animals given a heterologous vaccine, both the position of the peak of the second mode (i.e. that corresponding to longer infectious periods) and the maximum recorded infectious period are increased when compared with the homologous group. Ideally, more data would be used to give better definitions to these distributions. However, the available data indicate that the heterologous vaccinees take longer to clear the infection. This may have important implications for the coexistence of strains (Gog & Grenfell 2002) and for the generation of escape mutants (Rocha *et al.* 1991) when within-host modification of strains is sufficient to give rise to a viral strain that hosts are not protected against. An increased infectious period would allow the virus more replication cycles and a greater chance to evolve by a combination of random mutations and (immune) selection pressure (Grenfell *et al.* 2004).

The distributions of the latent and infectious periods for the homologous (figure 2*a,b*) and heterologous (figure 2*c,d*) vaccinees do not bear any resemblance to exponential distributions (figure 2), whereas models often assume exponential distributions for the latent and infectious periods. The importance of this assumption is demonstrated by the comparative study undertaken in § 4*a*, which shows that using exponential distributions for the latent and infectious periods results in a substantial underestimate of the probability of large outbreaks occurring. This is because, when using the empirical distributions, there is a significant chance of an animal remaining

infectious for 4–6 days (figure 2), which reduces the chance of a break in the chain of transmission in the early phases of the epidemic. Mathematical models in epidemiology are often used to assess the risk of large outbreaks occurring (e.g. Anderson & May 1991). Therefore, as demonstrated here, it is vital to take into account the real distributions of the latent and infectious periods.

We have used data relating to individual animals (antibody levels, seroconversion status, latent period, infectious period, etc.) to parameterize a modified SEIR model designed to incorporate the effects of imperfect influenza vaccines. The model takes into account the fact that some vaccinated animals become infected (measured serologically) but not infectious (no detectable virus excreted). Data such as these demonstrate that the 'exposed' status used in the model is much more than simply a holding class that would only generate a delay in the infection dynamics.

The risk of an outbreak of influenza is highly dependent on whether the vaccine strain is homologous or heterologous to the challenge strain (figure 5). This is owing to a combination of three observed differences between the two groups: first, the probability of seroconverting as a function of antibody level (figure 1); second, the probability of seroconverting but not excreting detectable virus; and, finally, the distribution of infectious periods (figure 2). Although each factor taken in isolation gives rise to only small differences between the two vaccine types, the population-modelling work shows that the increased risk to a population of animals is considerable. The striking difference between the risk profiles for the two groups provides a stark warning of the need to continue surveillance of outbreaks and update vaccines regularly.

The probability of an outbreak in the heterologously vaccinated population (figure 5, solid line) is not strongly affected by the time of year that the outbreak begins (cf. the homologously vaccinated population, dotted line). Essentially this is because vaccinated horses typically maintain antibody levels of *ca.* 100–150 mm² (SRH) (Newton *et al.* 2000*a,b*; Park *et al.* 2003). In this region there is a larger range of probabilities of protective immunity in the homologously vaccinated population than in the heterologously vaccinated population (figure 1). Consequently, the heterologously vaccinated population is inherently less variable. This effect reduces the potential impact of a timing strategy in the heterologously vaccinated population, whereby the dosing schedule is designed to maximize protection at the population level during periods of high risk of recruiting the infection (e.g. the peak of the racing season, or the annual sales season).

The fact that the experimental populations were immunologically naive provided the opportunity to study the effects of strain heterology on the epidemiology of equine influenza without the complication of pre-existing antibodies. The challenge remains to develop the research to incorporate the effects of pre-existing antibodies (conferred through either prior infection or prior vaccination), as those results would apply to a broader class of populations, including humans.

In future work, we aim to quantify and qualify the heterology between vaccine and epidemic strains. In the first instance, quantification may be achieved by considering the number of amino acid differences between two strains

in established antigenically important regions (Wiley *et al.* 1981; Bizebard *et al.* 2001). Indeed, modelling work has begun to grapple with the 'shape space' in which influenza viruses evolve (Smith *et al.* 1997; Plotkin *et al.* 2002; Gog & Grenfell 2002; Ferguson *et al.* 2003), and the recent outbreak of equine influenza in a vaccinated population (Wood *et al.* 2003) highlights the need to continue this work.

This manuscript was much improved following the comments of two anonymous referees, whom the authors thank for their constructive criticism. The authors are very grateful to the Horserace Betting Levy Board, who have funded much of this work through their support of the Animal Health Trust's diagnostic and surveillance services. B.T.G. was supported by the Wellcome Trust and BBSRC. K.G. was supported by the Wellcome Trust.

REFERENCES

- Anderson, D. & Watson, R. 1980 On the spread of a disease with gamma distributed latent and infectious periods. *Biometrika* **67**, 191–198.
- Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans: dynamics and control*. Oxford University Press.
- Askonas, B. A., McMichael, A. J. & Webster, R. G. 1982 The immune response to influenza virus and the problem of protection against infection. In *Basic and applied influenza research* (ed. A. S. Beare), pp. 159–188. Boca Raton, FL: CRC Press.
- Bartlett, M. S. 1961 Monte Carlo studies in ecology and epidemiology. In *Proc. 4th Berkeley Symp. Math. Statist. Prob.*, vol. 4 (ed. J. Neyman), pp. 39–55. Berkeley, CA: University of California Press.
- Bell, L. (ed.) 2002 *Horses in training*. London: Raceform.
- Bizebard, T., Barbey-Martin, C., Fleury, D., Gigant, B., Barrere, B., Skehel, J. J. & Knossow, M. 2001 Structural studies on viral escape from antibody neutralization. *Curr. Topics Microbiol. Immunol.* **260**, 55–64.
- Casagrande, J. T., Pike, M. C. & Smith, P. G. 1978 An improved approximate formula for calculating sample sizes for comparing binomial distributions. *Biometrics* **34**, 483–486.
- Cliff, A. D., Haggett, P. & Ord, J. K. 1986 *Spatial aspects of influenza epidemics*. London: Pion Ltd.
- Daly, J. M., Yates, P. J., Newton, J. R., Park, A. W., Henley, W., Wood, L. J. N., Davis-Poynter, N. & Mumford, J. A. 2004 Evidence supporting the inclusion of strains from each of the two co-circulating lineages of H3N8 equine influenza virus in vaccines. *Vaccine*. (In the press.) (DOI 10.1016/j.vaccine.2004.02.048.)
- de Jong, M. C. M. & Kimman, J. G. 1994 Experimental quantification of vaccine-induced reduction in virus transmission. *Vaccine* **12**, 761–766.
- Diekmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. Chichester: Wiley.
- Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. 1990 On the definition and computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382.
- Dimmock, N. J. 1984 Mechanisms of neutralisation of animal viruses. *J. Gen. Virol.* **65**, 1015–1022.
- Eliason, S. R. 1993 *Maximum likelihood: logic and practice*. Beverley Hills, CA: Sage.
- Ferguson, N. M., Galvani, A. P. & Bush, R. M. 2003 Ecological and immunological determinants of influenza evolution. *Nature* **422**, 428–433.
- 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.
- Glass, K., Wood, J. L. N., Mumford, J. A., Jessett, D. & Grenfell, B. T. 2002 Modelling equine influenza. 1. A stochastic model of within-yard epidemics. *Epidemiol. Infect.* **128**, 491–502.
- Gog, J. R. & Grenfell, B. T. 2002 Dynamics and selection of many-strain pathogens. *Proc. Natl Acad. Sci. USA* **99**, 17 209–17 214.
- Grenfell, B. T. & Dobson, A. P. (eds) 1995 *Ecology of infectious diseases in natural populations*. Cambridge University Press.
- 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.
- Hobson, D. D., Curry, R. L., Beare, A. S. & Ward-Gardner, A. 1972 The role of serum haemagglutinin inhibiting antibody in protection against challenge infection with influenza A2 and B viruses. *J. Hygiene* **70**, 767–777.
- Iofescu, M. 1973 *Stochastic processes and applications in biology and medicine*. Berlin: Springer.
- Keeling, M. J. & Grenfell, B. T. 1998 Effect of variability in infection period on the persistence and spatial spread of infectious diseases. *Math. Biosci.* **147**, 207–226.
- Kermack, W. O. & McKendrick, A. G. 1927 A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* **115**, 700–721.
- Lloyd, A. L. 2001a Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. *Theor. Popul. Biol.* **60**, 59–71.
- Lloyd, A. L. 2001b Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. *Proc. R. Soc. Lond. B* **268**, 985–993. (DOI 10.1098/rspb.2001.1599.)
- Longini, I. M., Datta, S. & Halloran, M. E. 1996 Measuring vaccine efficacy for both susceptibility to infection and reduction in infectiousness for prophylactic HIV-1 vaccines. *J. AIDS* **13**, 440–447.
- McCallum, H., Barlow, N. & Hone, J. 2001 How should pathogen transmission be modelled? *Trends Ecol. Evol.* **16**, 295–300.
- Mollison, D. 1977 Spatial contact model for ecological and epidemic spread. *J. R. Statist. Soc. B* **39**, 283–326.
- Morley, P. S., Townsend, H. G. G., Bogdan, J. R. & Haines, D. M. 2000 Risk factors for disease associated with influenza virus infections during three epidemics in horses. *J. Am. Vet. Med. Assoc.* **216**, 545–550.
- Mumford, J. A. 1999 The equine influenza surveillance program. *Adv. Vet. Med.* **41**, 379–387.
- Mumford, J. A., Wood, J. M., Scott, A. M., Folkers, C. & Schild, G. C. 1983 Studies with inactivated equine influenza vaccine. 2. Protection against experimental infection with influenza A/equine/Newmarket/79 (H3N8). *J. Hygiene* **90**, 385–395.
- Nåsell, I. 1995 The threshold concept in stochastic epidemic and endemic models. In *Epidemic models: their structure and relation to data* (ed. D. Mollison), pp. 71–83. Cambridge University Press.
- Newton, J. R., Lakhani, K. R., Wood, J. L. N. & Baker, D. J. 2000a Risk factors for equine influenza serum antibody titres in young thoroughbred racehorses given an inactivated vaccine. *Prevent. Vet. Med.* **46**, 129–141.
- Newton, J. R., Townsend, H. G. G., Wood, J. L. N., Sinclair, R., Hannant, D. & Mumford, J. A. 2000b Immunity to equine influenza: relationship of vaccine induced antibody in young thoroughbred racehorses to protection against field infection with influenza A/equine-2 viruses (H3N8). *Equine Vet. J.* **32**, 65–74.

- OIE 1996 Conclusions and recommendations from the consultation meeting of OIE and WHO experts on equine influenza, Newmarket, United Kingdom, September 18–19, 1995. *Office Int. Epizootics Bull.* **108**, 482–484.
- Park, A. W., Wood, J. L. N., Newton, J. R., Daly, J., Mumford, J. A. & Grenfell, B. T. 2003 Optimising vaccination strategies in equine influenza. *Vaccine* **21**, 2862–2870.
- Plotkin, J. B., Dushoff, J. & Levin, S. A. 2002 Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *Proc. Natl Acad. Sci. USA* **99**, 6263–6268.
- Rocha, E., Cox, N. J., Black, R. A., Harmon, M. W., Harrison, C. J. & Kendal, A. P. 1991 Antigenic and genetic variation in influenza A (H1N1) virus isolates recovered from a persistently infected immunodeficient child. *J. Virol.* **65**, 2340–2350.
- Schild, G. C., Pereira, M. S. & Chakraverty, P. 1975 Single radial haemolysis: a new method for the assay of antibody to influenza haemagglutinin. *Bull. WHO* **52**, 43–50.
- Scholtens, R. G. & Steele, J. H. 1964 U.S. epizootic of equine influenza. *Public Hlth Rep.* **79**, 393–402.
- Smith, D. J., Forrest, S., Hightower, R. R. & Perelson, A. S. 1997 Deriving shape space parameters from immunological data. *J. Theor. Biol.* **189**, 141–150.
- Sprenger, M. J. W., Beyer, W. E. P., Kempen, B. M. & Mulder, P. G. H. 1993 Risk factors for influenza mortality? In *Options for the control of influenza II* (ed. C. Hannoun, A. P. Kendal, H. D. Klenk, A. McMichael, K. G. Nicholson & A. Oya), pp. 15–23. Amsterdam: Excerpta Medica.
- Turkington, C. & Ashby, B. 1998 *Encyclopaedia of infectious diseases*. New York: Facts on File.
- Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M. & Kawaoka, Y. 1992 Evolution and ecology of influenza A viruses. *Microbiol. Rev.* **56**, 152–179.
- Wiley, D. C., Wilson, I. A. & Skehel, J. J. 1981 Structural identification of the antibody binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* **289**, 373–378.
- Wood, J. L. N., Newton, J. R., Daly, J., Park, A. W. & Mumford, J. A. 2003 It's all in the mix: infection transmission in populations. *Equine Vet. J.* **35**, 526–528.
- Yates, P. & Mumford, J. A. 2000 Equine influenza vaccine efficacy: the significance of antigenic variation. *Vet. Microbiol.* **74**, 173–177.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.