Supplementary Information for 'Vaccination and Reduced Cohort Duration Can Drive Virulence Evolution: Marek's Disease Virus and Industrialized Agriculture'

Katherine E Atkins, Andrew F Read, Nicholas J Savill, Katrin G Renz, AFM Fakhrul Islam, Stephen W Walkden-Brown and Mark EJ Woolhouse

Contents

1 Dust shedding quantification

1.1 Data

In an experiment conducted and described by Islam and Walkden-Brown (2007), groups of broiler chickens were raised from an age of one day in isolators. All the dust from each isolator and its exhaust was retrieved every 24 or 48 hours. The total mass of dust shed per day per bird was found for weeks 1-8, giving a total of 8 data points per isolator.

1.2 Methods

To estimate the dust produced per bird of age t per day, $d(t)$, we fit a function of the form $d(t) = \eta_1 exp(-\eta_2/t^{\eta_3}) + \eta_4$.

where $\eta_i, \forall i \in [1, 4]$ were parameters to be estimated. Matlab (2011) (`lsqcurvefit` function) was used to estimate the best-fit parameters.

We have multiple dust profile curves and need to estimate an average dust shedding function. Minimizing the sum (over all observations) of the distances from each data point to the model point is equivalent to minimizing the distance from the arithmetic mean of multiple data points to the model point.

1.3 Results

A least squares fitting gives the quantity of dust (in mg per day) as

$$d(t) = 368\exp(-326/t^{1.64}) + 10.8 \tag{S1}$$

with the graph displayed in Figure S1.

A value for $P_{45} = 326$ was estimated because the time for maturation of the experimental birds to the required finishing weight was 45 days. This value was then used to calculate different growth curves for different cohort times. For example $P_{45} = 326$ gives $d(45, 45) = 206$, so P_{70} (the growth parameter for birds who are slaughtered after 70 days), for example, can be estimated by solving $206 = 368\exp(-P_{70}/70^{1.64}) + 10.8$. This then gives the new $d(t, T_c)$ which can be used for estimating the quantity of dust produced by a broiler on day t when in a cohort of duration 70 days.

2 MDV transmission

2.1 Introduction

We quantified the daily transmission rates to susceptible birds, who were either unvaccinated or HVT-vaccinated (Islam et al. *In Prep.*). Vaccinated hosts are still able to become infected with MDV. Here we elucidate the relationship between vaccination and susceptibility to infection.

2.2 Methods

Daily transmission probabilities for unvaccinated and HVT vaccinated birds were calculated independently and directly from the data via maximum likelihood. Each pen was analyzed separately for unvaccinated and HVT vaccinated birds.

Let X be defined as the random variable, the number of sampled individuals who are infected. At each sample time, $i = 1, 2, 3, 4$ (corresponding to 5, 10, 15, 20 days post exposure, respectively) birds are sampled without replacement. Therefore the number of sampled infected individuals follows a hypergeometric distribution (Kalbfleisch, 1985).

$$P(X = k_i) = \frac{\binom{M_i}{k_i}\binom{N_i - M_i}{n_i - k_i}}{\binom{N_i}{n_i}}$$

where,

M_i	total number of infecteds in the population at time i before sampling
N_i	total population at time i before sampling
n_i	sample size at time i

Now the likelihood can be defined such that

$$\mathcal{L}(k_1, k_2, k_3, k_4 | M_1, M_2, M_3, M_4) = \prod_{i=1}^{4} P(X = k_i | M_i)$$

where k_i is the observed number of infected individuals in each sample at time i and M_i are the parameters to estimate.

$$\max\{\mathcal{L}\} = \max_{M_j \leq N_j} \prod_{i=1}^{4} P(X = k_i | M_i) \tag{S2}$$

Since k_i, n_i and N_i are known, $M_i = \hat{M}_i$ can be calculated directly for the maximum likelihood estimate. Therefore there will be a set of $\hat{M}_i$ for each pen, for each vaccination group. The newly infected individuals between each time point, $m_i = \hat{m}_i$ can be calculated trivially, giving us the total number of infected individuals in each group in each pen between each sample time.

Assuming the number of newly infected individuals between each time point, L_i, follows a binomial distribution, with $\mathbb{E}(L_i) = m_i$, $L_i \sim \text{Bin}(N_i - \sum_{j<i} m_j + \sum_{j<i} k_j, q_i)$. The first parameter is the effective population size available to be infected at time step i, which is the number of un-sampled individuals at time i before sampling ($N(i)$), minus the number of infected individuals who have not yet been sampled ($\sum_{j<i} m_j - \sum_{j<i} k_j$). The second parameter is the probability

of infection between sample time $i-1$ and i. The maximum likelihood estimate of the expected probability of transmission within time period i is therefore

$$\hat{q}_i = \frac{\hat{m_i}}{N_i - \sum_{j<i} m_j + \sum_{j<i} k_j} \tag{S3}$$

Assuming there is an equal chance of infection on any of the 5 days between sampling, the daily infection per bird, p_i, is given by,

$$\hat{q}_i = 1 - (1 - \hat{p}_i)^5 \tag{S4}$$

$$\Rightarrow \hat{p}_i = 1 - (1 - \hat{q}_i)^{1/5} \tag{S5}$$

2.3 Results

The maximum likelihood estimates for the number of newly infected individuals, $\hat{m}_i$, and the daily probabilities of becoming infected, $\hat{p}_i$, were calculated for each replicate (pen) (Table S1 (unvaccinated) and Table S2 (HVT-vaccinated)).

The amount of virus in each pen is known at certain days, and linear interpolation estimates the average amount of virus (measured in VCN/m^3) between days 0-5, 5-10, 10-15 and 15-20. The average amount of virus in each pen and the probability of infection within that period is shown in Figure S2. The associated probabilities per day are shown in Figure S3.

Because the virus shed for one bird is much smaller than the quantities of virus examined in this experiment only the first datapoint (5 days post exposure) from each replicate is used in the to fit a linear regression between virus concentration

	Pen Number									
	1	2	3	4	5	6	7	8	9	10
k_1	0	0	0	1	1	0	0	0	1	1
k_2	3	3	2	3	2	2	1	3	3	5
k_3	4	5	4	5	4	5	5	5	4	5
k_4	5	5	5	5	5	5	5	5	5	5
m_1	0	0	0	4	4	0	0	0	4	4
m_2	9	9	6	6	3	6	3	9	6	12
m_3	2	4	4	4	4	6	8	4	2	0
m_4	1	0	1	0	1	0	0	0	1	0
M_1	0	0	0	4	4	0	0	0	4	4
M_2	9	9	6	9	6	6	3	9	9	15
M_3	8	10	8	10	8	10	10	10	8	10
M_4	5	5	5	5	5	5	5	5	5	5
q_1	0	0	0	0.2	0.2	0	0	0	0.2	0.2
q_2	0.6	0.6	0.4	0.5	0.25	0.4	0.2	0.6	0.5	1
q_3	0.5	1	0.67	1	0.67	1	1	1	0.5	-
q_4	1	-	1	-	1	-	-	-	1	-
p_1	0	0	0	0.044	0.044	0	0	0	0.044	0.044
p_2	0.17	0.17	0.10	0.13	0.056	0.10	0.044	0.17	0.13	1
p_3	0.13	1	0.20	1	0.20	1	1	1	0.13	-
p_4	1	-	1	-	1	-	-	1	-	-

Table S1: Transmission to unvaccinated birds: Maximum likelihood estimates for quantities from the hypergeometric distribution. For timestep i: k_i is the observed number of infected individuals (Islam et al. *In Prep.*); m_i is the maximum likelihood estimate of the number of newly infected individuals; M_i is the maximum likelihood estimate of the cumulative number of infected individuals; q_i is the estimated probability per timestep of infection per bird, and p_i is the estimated daily probability of infection per bird. The sampling was conducted in an unvaccinated population of birds where the number of newly infected individuals within a timestep is assumed to be binomially distributed.

	Pen Number									
	1	2	3	4	5	6	7	8	9	10
k_1	1	0	0	0	0	0	0	2	0	0
k_2	1	2	0	0	0	0	0	0	1	0
k_3	0	1	0	1	2	0	2	0	1	0
k_4	0	1	0	2	1	0	1	0	1	1
m_1	2	0	0	0	0	0	0	2	0	0
m_2	0	4	0	0	0	0	0	0	3	0
m_3	0	0	0	2	3	0	3	0	0	0
m_4	0	0	0	1	0	0	0	0	0	1
M_1	2	0	0	0	0	0	0	2	0	0
M_2	1	4	0	0	0	0	0	0	3	0
M_3	0	2	0	2	3	0	3	0	2	0
M_4	0	1	0	2	1	0	1	0	1	1
q_1	0.1	0	0	0	0	0	0	0.1	0	0
q_2	0	0.29	0	0	0	0	0	0	0.2	0
q_3	0	0	0	0.2	0.3	0	0.3	0	0	0
q_4	0	0	0	0.25	0	0	0	0	0	0.2
p_1	0.021	0	0	0	0	0	0	0.021	0	0
p_2	0	0.065	0	0	0	0	0	0	0.044	0
p_3	0	0	0	0.044	0.069	0	0.069	0	0	0
p_4	0	0	0	0.056	0	0	0	0	0	0.044

Table S2: Transmission to HVT-vaccinated birds: Maximum likelihood estimates for quantities from the hypergeometric distribution. For timestep i: k_i is the observed number of infected individuals; m_i is the maximum likelihood estimate of the number of newly infected individuals; M_i is the maximum likelihood estimate of the cumulative number of infected individuals; q_i is the estimated probability per timestep of infection per bird, and p_i is the estimated daily probability of infection per bird. The sampling was conducted in an unvaccinated population of birds where the number of newly infected individuals within a timestep is assumed to be binomially distributed.

and probability of infection per bird per day. Since the value of the intercept was not significantly different to zero and it makes biological sense to fit the line through the origin, the gradient was calculated as $\alpha(sham)$=8.97e-09 (p=0.07) for the unvaccinated birds and $\alpha(hvt)$=1.47e-09 (p=0.42) for the vaccinated birds. This relationship is shown in Figure S4.

3 Viral Shedding Regression

We performed statistical analyses describing the relationship between primary/secondary latent periods/shedding rates and virulence score/vaccine treatment (Atkins et al., 2011). Here we use these results to make more parsimonious statistical models where some associations were non-significant.

3.1 Primary Latent Period

Primary latent period was shown not to vary significantly between individuals in different vaccination treatment groups or between individuals infected with viruses with different virulence scores (Atkins et al., 2011). Therefore, we use the estimated primary latent period averaged over all individuals (4.7 days).

3.2 Secondary Latent Period

We removed the non-significant coefficients (at 15%) in turn to produce the most parsimonious statistical model (Adjusted R-squared=0.589, p-value= 0.000499):

	Estimate	Std. Error	t value	Pr($>\|t\|$)
Intercept	9.917	0.641	15.472	1.25e-10
vaccBiv	19.525	4.147	4.708	0.00028
VirulenceScore:vaccBiv	-26.464	6.633	-3.990	0.00118

3.3 Primary Shedding Rate

The full multiplicative model was used (Adjusted R-squared=0.0381, p-value=0.394)

	Estimate	Std. Error	t value	Pr(> |t|)
Intercept	86518	34657	2.496	0.0281
VirulenceScore	-119282	56107	-2.126	0.0549
vaccHVT	-91918	49012	-1.875	0.0853
vaccBiv	-86341	49012	-1.762	0.1036
VirulenceScore:vaccHVT	154640	79348	1.949	0.0751
VirulenceScore:vaccBiv	131342	79348	1.655	0.1238

3.4 Secondary Shedding Rate

We removed the non-significant coefficients (at 15%) in turn to produce the most parsimonious statistical model (Adjusted R-squared=0.761, p-value=3.29e-05):

	Estimate	Std. Error	t value	Pr(> |t|)
Intercept	-23914614	7461561	-3.205	0.006356
VirulenceScore	84525922	12657226	6.678	1.05e-05
VirulenceScore:vaccHVT	-10268274	6546052	-1.569	0.139056
VirulenceScore:vaccBiv	-30587445	6546052	-4.673	0.000359

4 R_0 is Independent of Number of Individuals

$R_0$is independent of the number of individuals within the cohort.

$$
\begin{aligned}
R_0(T_c, v, j, s_d) &= S_0 \sum_{t=T_s+1}^{T_c} p\left(\tfrac{M_e(t,T_c,v,j)}{V(S_0,s_d)}\right) L(t,v,j) && \text{(S6)}\\
&&& \text{(S7)}\\
&= S_0 \sum_{t=T_s+1}^{T_c} \alpha(j) \frac{M_e(t,T_c,v,j)}{V(S_0,s_d)} L(t,v,j) \qquad \text{(see Figure S4)} && \text{(S8)}\\
&&& \text{(S9)}\\
&= \frac{S_0\alpha(j)}{V(S_0,s_d)} \sum_{t=T_s+1}^{T_c} M_e(t,T_c,v,j) L(t,v,j) && \text{(S10)}\\
&&& \text{(S11)}\\
&= \frac{S_0\alpha(j)}{hwS_0/s_d} \sum_{t=T_s+1}^{T_c} M_e(t,T_c,v,j) L(t,v,j) && \text{(S12)}\\
&&& \text{(S13)}\\
&= \frac{\alpha(j)}{hw/s_d} \sum_{t=T_s+1}^{T_c} \gamma(t,T_c,s_d)[M_e(t-1,T_c,v,j) + m(t,T_c,v,j)] L(t,v,j) && \text{(S14)}
\end{aligned}
$$

Now since the reduction, $\gamma(t, T_c, s_d)$, can be further be broken down:

$$\gamma(t, T_c, s_d) = \min\Big[\frac{EhS_0/(s_d/w)}{\min[\sum_{s=1}^{t-1} S_0 d(s, T_c), EhS_0/(s_d/w)] + S_0 d(t, T_c)}, 1\Big] \tag{S15}$$

$$\tag{S16}$$

$$= \min\Big[\frac{EwhS_0/s_d}{\min[\sum_{s=1}^{t-1} S_0 d(s, T_c), EwhS_0/s_d] + S_0 d(t, T_c)}, 1\Big] \tag{S17}$$

$$\tag{S18}$$

$$= \min\Big[\frac{Ewh/s_d}{\min[\sum_{s=1}^{t-1} d(s, T_c), Ehw/s_d] + d(t, T_c)}, 1\Big] \tag{S19}$$

$$\tag{S20}$$

it is clear that R_0 is independent of S_0, but not s_d, the stocking density. Supposing the equilibrium value of dust has been reached, then since γ is a function of s_d, increasing the stocking density will reduce the fraction of dust remaining. Mathematically, this is true since the numerator is reduced by increasing s_d more than the denominator in the above formulation. This makes intuitive sense since the higher the stocking density, the more birds per unit volume and the more dust per unit volume which implies that more dust must be taken out if the equilibrium is to be maintained.

5 R_0 sensitivity to dust levels

We varied the range of maximum inhalable dust levels (optional exposure limits - OEL) in the barn atmosphere over the range of dust concentrations greater that those seen in European countries. Over this range, there is no change in optimal virulence score (Figure S5).

References

Atkins, K. E., A. F. Read, N. J. Savill, K. G. Renz, S. W. Walken-Brown, and M. E. Woolhouse (2011). Modelling Marek's Disease Virus (MDV) Infection: Parameter Estimates For Mortality Rate And Infectiousness. *BMC Vet Res* *7*(1), 70.

Islam, A. and S. W. Walkden-Brown (2007). Quantitative profiling of the shedding rate of the three Marek's disease virus (MDV) serotypes reveals that challenge with virulent MDV markedly increases shedding of vaccinal viruses. *J Gen Virol 88*, 2121–2128.

Kalbfleisch, J. G. (1985). *Probability and Statistical Inference, Volume 2*. Springer.

Matlab (2011). R2011a. The MathWorks Inc., Natick, MA.

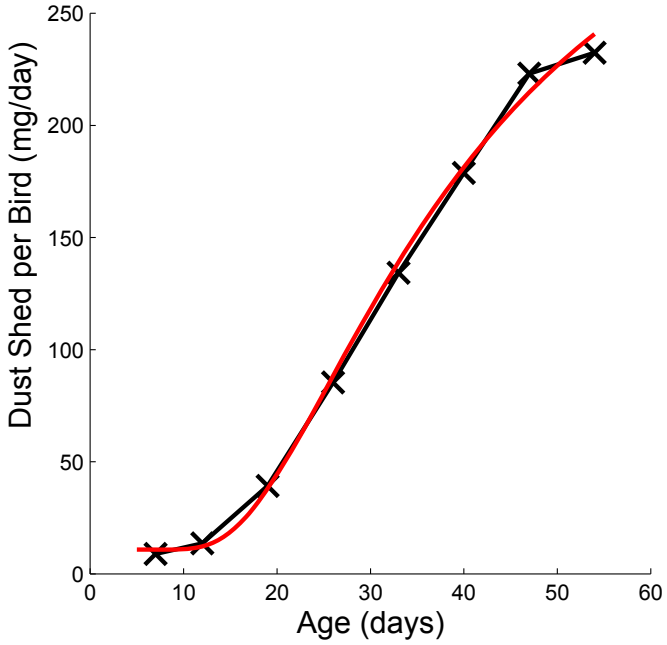
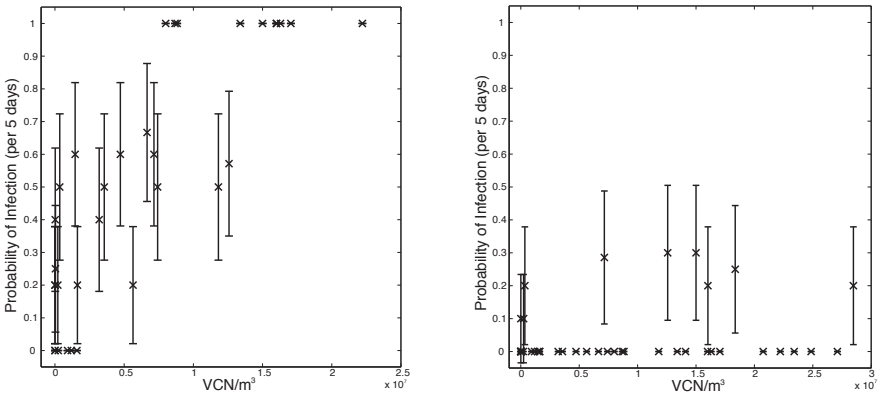


Figure S1

Dust Shedding: The amount of dust shed over time by a broiler chicken (black line) and the fitted function, $d(t)$ (red line).



(a) Sham (b) HVT

Figure S2

Transmission: 5 day probabilities for infection for different atmospheric virus concentrations (measured in VCN per m^3). Error bars are twice the standard error of the estimate (a) unvaccinated (b) HVT vaccinated.

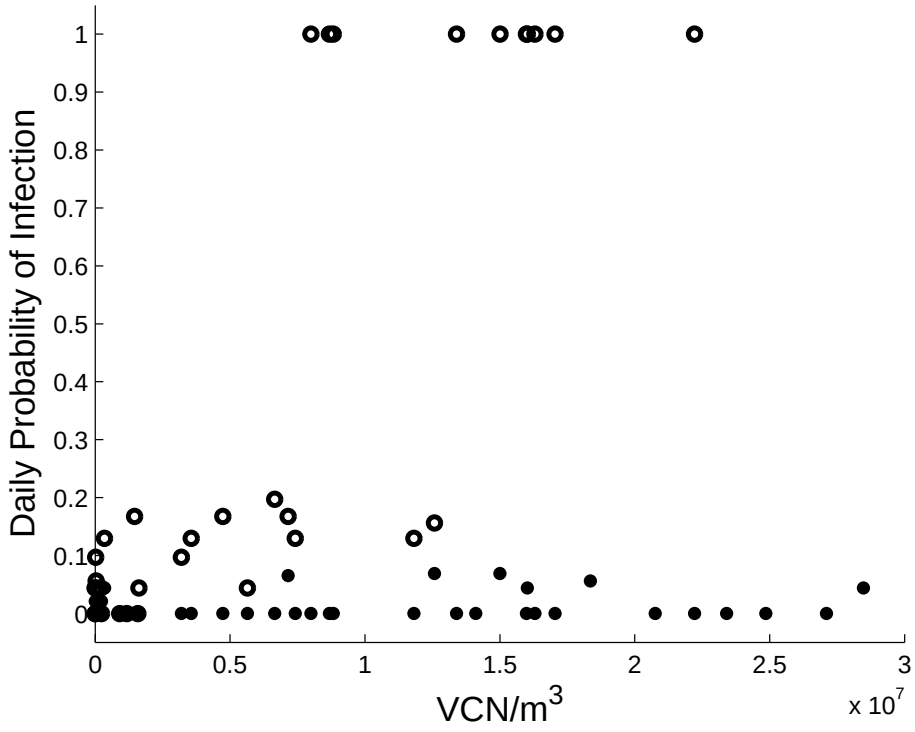


Figure S3

Transmission: Daily probability of infection per bird with different vaccination treatments. The circles are the maximum likelihood point estimates for p given for the different timesteps and pens (open for unvaccinated, filled for HVT vaccinated).

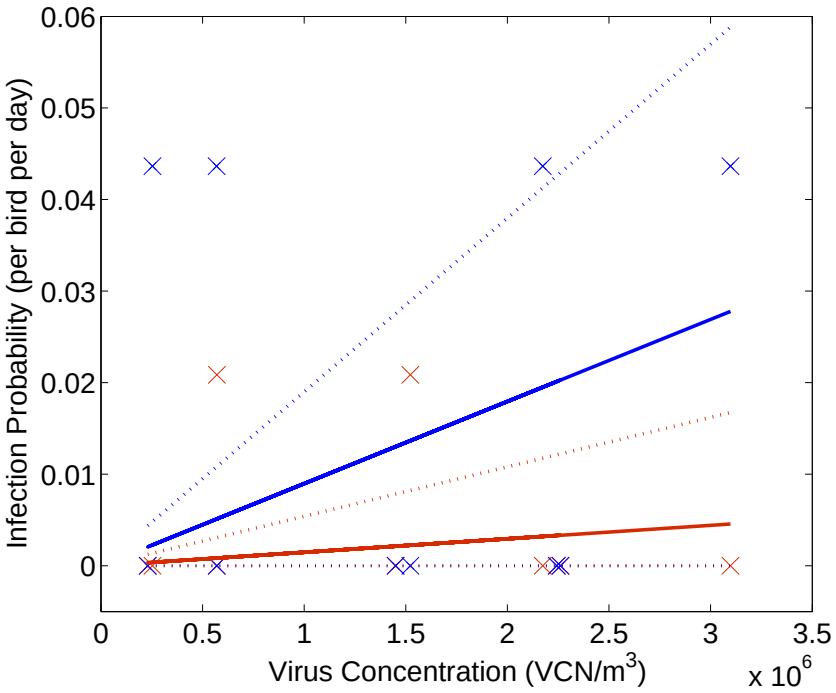


Figure S4

Transmission for Small Virus Concentrations: the probability of infection per day per bird, with the average estimated quantities of virus concentration in the atmosphere. The blue and red crosses are the unvaccinated and HVT vaccinated birds respectively. The blue and red lines give the least squares estimate of the line of best-fit to the unvaccinated and HVT vaccinated birds respectively. The dotted lines give the 95% confidence intervals on the regression line. Note that the dotted line at $y = 0$ is the limit for the lower confidence interval for both lines.

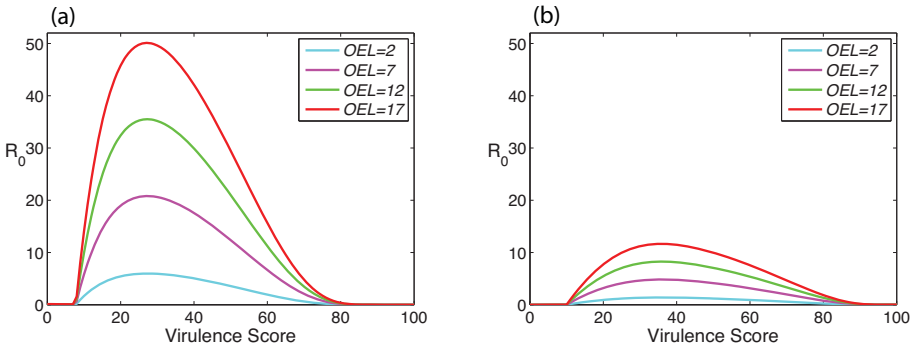


Figure S5

Effect of changing maximum dust levels (mg/m^3) on the reproductive ratio of MDV strains of different virulence scores. (a) Unvaccinated hosts (b) HVT-vaccinated hosts. Mortality rate is set at 0.0005 per bird per day. Cohort duration is set to 50 days.