Science Supporting Online Material

Public Health Risk from the Avian H5N1 Influenza Epidemic

Neil M. Ferguson, Christophe Fraser, Christl A. Donnelly,
Azra C. Ghani, Roy M. Anderson

doi: 10.1126/science. 1096898

Estimating R_0

We assume that in the absence of reassortment, all human infections with avian H5N1 virus have a reproduction number, R_0, below unity, and hence, although any such case may generate a small cluster of human-to-human transmitted H5N1 cases, such transmission is below the $R_0 = 1$ threshold required for self-sustaining transmission.

The dynamics of such "subcritical" outbreaks have been well studied in past work [reviewed in (*S1*)], and the probability distribution for the number of cases generated in an infectious disease outbreak with $R_0 < 1$ has been derived (*S1*, *S2*). If one assumes the host population is sufficiently large that depletion of susceptibles can be ignored for small outbreaks, the probability that an outbreak seeded by a single case will end after a total of n cases (including the index case) is

$$P(n \mid R_0) = \frac{(2n-2)!}{n!(n-1)!} \frac{{R_0}^{n-1}}{(R_0+1)^{2n-1}} \qquad \text{(A)}$$

if the infectious period is exponentially distributed, and

$$P(n \mid R_0) = \frac{(nR_0)^{n-1} \exp(-nR_0)}{n!} \qquad \text{(B)}$$

if the infectious period is of fixed duration.

Both distributions have mean $1/(1 - R_0)$, but the variances differ: they are $R_0 (R_0 + 1)/(1- R_0)^3$ and $R_0/(1 - R_0)^3$, respectively.

Here, we assume exponentially distributed infectious periods. Given data on the observed distribution of outbreak sizes (each outbreak being the number of H5N1 cases caused by each avian-to-human transmission event, including the avian-to-human case itself), one can then estimate R_0 using maximum-likelihood methods by directly fitting to distribution (A) above using a multinomial likelihood. Model goodness of fit can also be evaluated, as the saturated likelihood is well defined, although it should be noted that this goodness-of-fit statistic is not χ^2 distributed for R_0 close to 1, as distribution (A) is highly skewed, which leads to low expected numbers of cases in the last few occupied cells of the multinomial. For a χ^2-distributed goodness-of-fit statistic, it is necessary to group outbreak sizes so as to ensure all cells in the multinomial likelihood have sufficient cases (typically over 5).

This approach makes best use of outbreak size distribution data, but clearly requires complete data to be available. We therefore also explored methods requiring less information. The multinomial log-likelihood defined below uses the following 3 items of data: the number of avian-to-human transmitted cases, N; the number of those cases generating human-to-human case clusters of size > 1, n_C; the size of the largest cluster seen to date, $C_{\max}$; and the number of clusters of size $C_{\max}$ that have been seen, $n_{\max}$ [note that for large $(N - n_C)$, $n_{\max}$ is almost certain to be 1];

$$l(N, n_C, C_{\max}, n_{\max} \mid R_0) = \text{const} + (N - n_C) \log(P(1 \mid R_0)) + (n_C - n_{\max}) \sum_{i=2}^{C_{\max}-1} \log(P(i \mid R_0)) + n_{\max} \log(P(C_{\max} \mid R_0))$$

if $C_{\max} > 1$.

The following goodness-of-fit statistic can then be defined:

$$X^2\left(N, n_C, C_{\max}, n_{\max} \mid R_0\right) = 2\Big\{\left(N - n_C\right)\left[\log\left(1 - n_C / N\right) - \log\left(P(1 \mid R_0)\right)\right] + \left(n_C - n_{\max}\right)\left[\log\left(\left(n_C - n_{\max}\right)/N\right) - \sum_{i=2}^{C_{\max}-1} \log\left(P(i \mid R_0)\right)\right] + n_{\max}\left[\log\left(n_{\max}/N\right) - \log\left(P(C_{\max} \mid R_0)\right)\right]\Big\} \quad \text{(C)}$$

For large samples, this statistic will be approximately χ^2 distributed. However, the expected very low occupancy of the cell corresponding to the largest observed cluster in the multinomial likelihood suggests that this approximation may not hold (except in the region R_0 ~ 0.5). It is therefore necessary to evaluate the exact distribution of X^2 (and hence p values) by simulation. The relative performances of this estimator and that based on the full distribution are compared in Fig. S1.

Right-hand censoring will affect all estimates of R_0, because the number of reported human-to-human cases will increase over time. Statistics dependent only on the first few generations of infection within a cluster will be less sensitive to censoring, such as, for example, the overall proportion of avian-to-human cases that result in a cluster. The effect of censoring is to lower estimates of R_0 for early time points where few human-to-human clusters have been fully observed—as can be clearly seen in Fig. S1. The effect is more pronounced for the higher values of R_0 scenario (Fig. S1b) than for the low values (Fig. S1a). The extent of bias depends critically on the temporal frequency of avian-to-human transmission events and on the time course of natural infection: The longer the duration of infection (time to symptoms and generation time) and the higher the frequency of avian events, the more censoring becomes a problem. In the simulation, we have assumed a constant rate of 0.5 avian events per day and parameters consistent with the well-documented time course for epidemic H1N1 influenza (see legend to Fig. S1). Repeated simulations to ascertain the impact of censoring will be needed if the frequency of avian-to-human events changes as the avian epidemic develops, and detailed clinical records must be analyzed to determine the time course of infection with novel influenza strains.

Detecting anomalous clusters

The model likelihood given by equation (C) assumes that all outbreaks have the same R_0, and thus, we would expect that model goodness of fit would worsen if an anomalous cluster arose with a larger R_0, because such a cluster would tend to be significantly larger than even the largest expected with a smaller R_0. This argument suggests the use of model goodness of fit as an indicator of the emergence of a cluster with an anomalous R_0.

To test this hypothesis, we simulated 75 avian-to-human transmission events occurring randomly over 150 days (assuming a latent period of 1 day, an asymptomatic infectious period of 1 day, and a symptomatic infectious period of 4 days) and the resulting human-to-human clusters arising from these, assuming $R_0 = 0.2$ or $R_0 = 0.8$, together with a single anomalous cluster arising on day 100 with $R_0 = 1.5$. We generated 100 realizations of this simulation and recorded if and when the goodness-of-fit p value [as evaluated by simulation from the expected outbreak size distribution (A)] fell below 0.05.

For $R_0 = 0.2$, before the start of the anomalous cluster on day 100, 5% of simulations gave $p < 0.05$ (as expected). In the 41 realizations where the anomalous $R_0 = 1.5$ cluster grew exponentially, the $p < 0.05$ threshold was always crossed, after an average of 5.4 cases had been generated in the cluster. In the 59 realizations where the anomalous cluster went extinct by random chance, the $p < 0.05$ threshold was crossed in the 8 of these for which the size of the cluster exceeded seven cases and a further 14 instances where the cluster was smaller still.

For $R_0 = 0.5$, before the start of the anomalous cluster on day 100, 3% of simulations gave $p < 0.05$ (consistent with the 5% expected). Sensitivity to detect anomalous clusters was reduced because of the smaller difference between the value of R_0 in normal outbreaks and the anomalous value. Although again all anomalous clusters not going extinct were eventually identified, detection only occurred after an average of 30 cases. For $R_0 = 0.8$, sensitivity is reduced substantially, with an anomalous cluster only being detected after generating an average of 200 cases (though it should be noted that human clusters of 100 cases can occur by random chance for $R_0 = 0.8$). Before the start of the anomalous cluster on day 100, 3% of simulations gave $p < 0.05$. We conclude that goodness-of-fit–based detection is highly sensitive in contexts where the transmissibility of avian virus between humans is low (i.e., when $R_0 < 0.8$), as appears to be the case in the current epidemic.

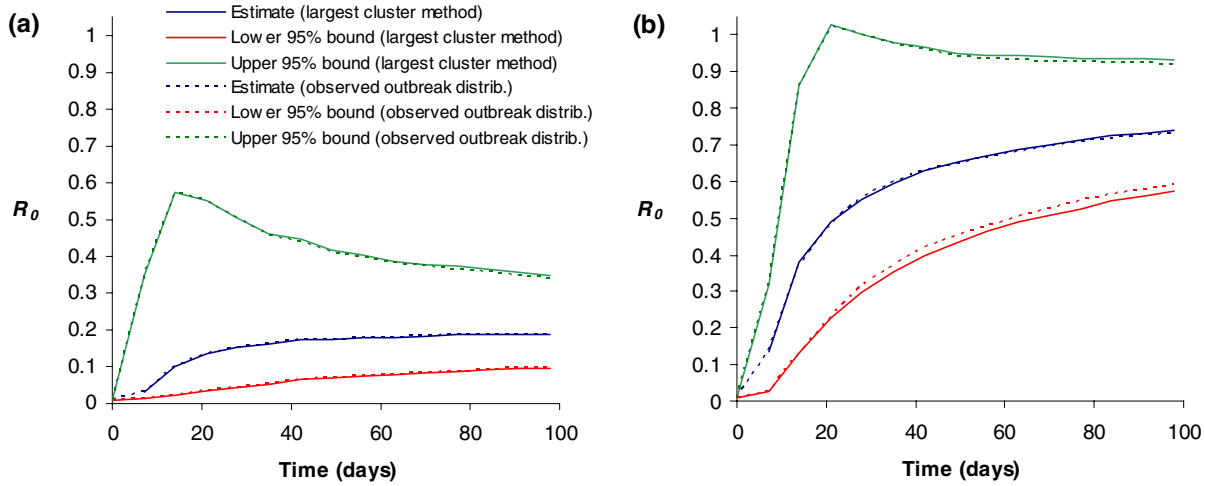

Fig. S1. Estimates and 95% confidence bounds for R_0 estimated from data collected during simulations of 50 avian-to-human transmission events occurring randomly over 100 days (assuming a latent period of 1 day, an asymptomatic infectious period of 1 day, and a symptomatic infectious period of 4 days) and the resulting human-to-human clusters arising from these, assuming (a) $R_0 = 0.2$, and (b) $R_0 = 0.8$. Curves shown are averages of estimates and confidence bounds calculated every day for each of 100 simulated datasets. Solid curves show the results obtained from using the goodness-of-fit statistic given in equation (C); dashed curves show results obtained by fitting to the complete empirical outbreak size distribution. Differences in the estimates produced by the two methods are negligible (both converge to the true value of R_0), and using the full distribution of outbreak sizes only results in marginally narrower confidence intervals. The underestimates of R_0 during the start of the outbreak reflect the impact of right-hand censoring in case reports.

It is interesting to note that using this method, the current case data give a marginal goodness of fit ($p \approx 0.05$), which indicates that under the assumption that two of the cases in the single cluster seen to date were human-to-human transmission, the size of that cluster is only marginally consistent with the observation of 33 cases in total and no other clusters.

In the text, the lower portion of the figure shows the threshold maximum cluster size required for detection as a function of the number of avian-to-human cases seen and the proportion of those cases generating human-to-human case clusters. Because we could not rely on X^2 being χ^2 distributed and because calculation of the distribution of X^2 conditional on particular values of N and n_C is highly computationally intensive, we calculated the maximum values of n_{max} shown in this figure as the 95th percentile values of $n_{\max}$ for the best-fit value of R_0 calculated, knowing N and n_C alone [namely, $R_0^* = n_C/(N - n_C)$]; i.e., the value of $n_{\max}$ for which $P\left(n_{\max} \mid R_0^*\right)^N \geq 0.95$. For $N = 75$ and R_0^* values of 0.2, 0.5, and 0.8, we verified that the threshold values of $n_{\max}$ calculated using this method closely match values obtained by determining the $p = 0.05$ goodness-of-fit threshold directly via simulation.

Case underascertainment

The analysis above assumes completely ascertainment of disease cases. Because no surveillance system is perfect, it is important to examine the impact of underascertainment, which is the result of cases that are not identified (or, equivalently, a proportion of cases that are relatively asymptomatic).

Simple random underascertainment (i.e., each case having a fixed probability of being reported) leads to underestimation of R_0, because the observed sizes of all clusters are reduced (*S1*). The magnitude of this effect is shown in Fig. S2. Note that the relative magnitude of bias is greatest for low values of R_0; so that if $R_0 = 0.8$, just 50% case ascertainment still gives an estimate value of 0.7. Furthermore, irrespective of the level of underascertainment, the true value of R_0 always falls within the confidence-bound estimates.

Underascertainment need not always lead to underestimation of R_0, however. Fig. S3 shows results for a scenario where random ascertainment affects initial case reports in the same manner as for Fig. S2, but then contact tracing is assumed to identify, retrospectively, all previously unidentified cases in clusters where at least one case was initially identified. This form of underascertainment leads to cluster sizes that are accurately observed when clusters are identified, but means that small clusters are completely missed with a disproportionately higher probability than larger clusters. This results in overestimation of R_0, although again confidence bounds on the estimate increase as ascertainment worsens.

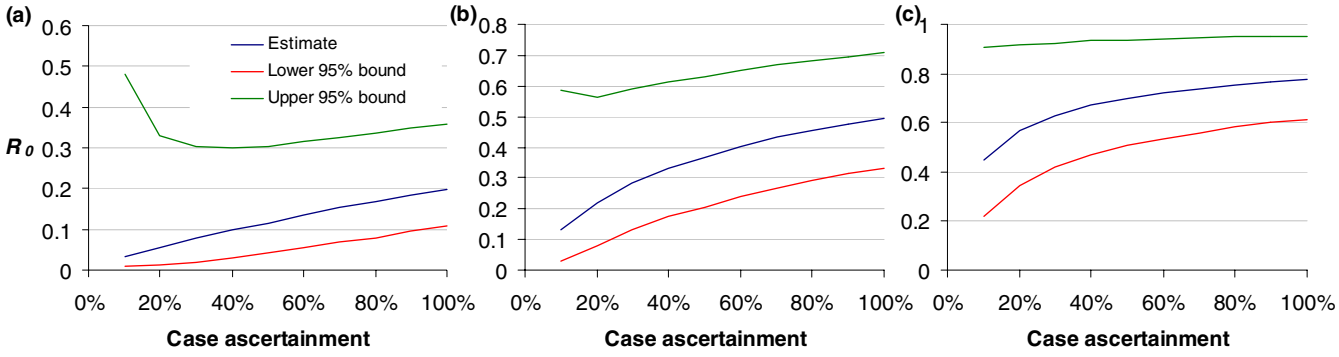

Fig. S2. Estimates and 95% confidence bounds for R_0, estimated from 50 simulated outbreaks with sizes drawn from distribution (A) above (to avoid the censoring issues illustrated in Fig. S1), as a function of the assumed level of case ascertainment. Ascertainment is assumed to be a random Bernoulli process, and the first ascertained case in any case cluster is attributed to avian-to-human transmission; clusters with no ascertained cases are not detected. Results for three different values of R_0 are shown: (a) $R_0 = 0.2$, (b) $R_0 = 0.5$, and (c) $R_0 = 0.8$. In all cases, random case underascertainment leads to underestimation of R_0, with the extent of bias increasing as ascertainment worsens.

A third possibility is that reporting is variable: The extreme would be that some regions and/or hospitals have perfect case ascertainment and contact tracing, whereas others report no cases. This is equivalent to reporting of a fixed proportion of all avian-to-human cases and only those human-to-human cases that are linked to reported avian-to-human cases. It is straightforward to see that this type of underascertainment does not bias R_0 estimates, or the method outlined above to detect anomalous clusters—although the loss of data widens confidence bounds and reduces detection power, respectively.

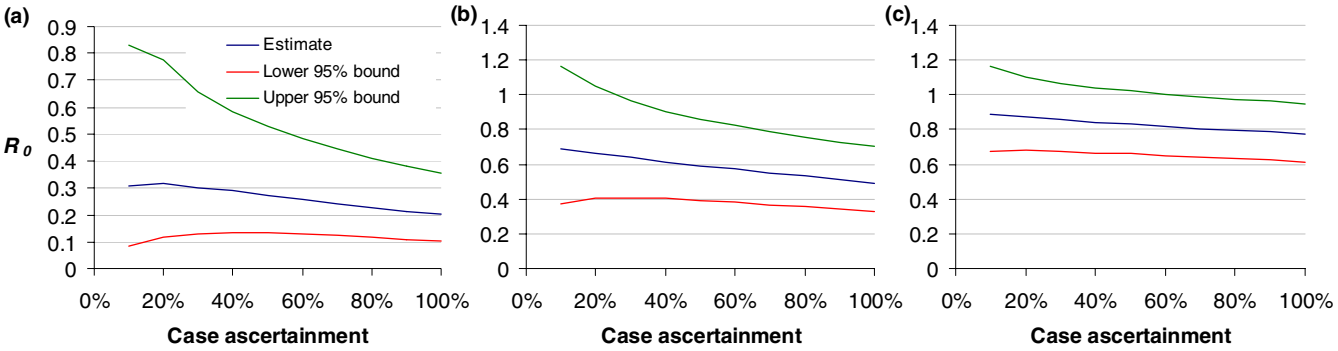

Fig. S3. As for Fig. S2, except that contact tracing is assumed to identify, retrospectively, all cases in any cluster where at least one case is identified through routine surveillance. Hence, here the only cases to be missed are those in case clusters where underascertainment led to a situation in which no cases were initially identified. This model of underascertainment leads to missing small clusters disproportionately, which results in overestimation of R_0.

References

S1. C. P. Farrington, M. N. Kanaan, N. J. Gay, *Biostatistics* **4**, 279–295 (2003).
S2. V. A. A. Jansen *et al.*, *Science* **301**, 804–804 (2003).