Robustness of the ferret model for influenza risk assessment studies: a cross-laboratory exercise

Jessica A Belser#*[1], Eric HY Lau#*[2], Wendy Barclay[3], Ian G. Barr[4], Hualan Chen[5], Ron AM Fouchier[6], Masato Hatta[7], Sander Herfst[6], Yoshihiro Kawaoka[7], Seema S Lakdawala[8], Leo Yi Yang Lee[4], Gabriele Neumann[7], Malik Peiris[2], Daniel R Perez[9], Charles Russell[10], Kanta Subbarao[4,11], Troy C Sutton[12], Richard J Webby[10], Huanliang Yang[5], Hui-Ling Yen*[2], Working group on the standardization of the ferret model for influenza risk assessment

[1]Influenza Division, Centers for Disease Control and Prevention, Atlanta, GA, USA; [2]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China; [3]Imperial College, London, UK; [4]WHO Collaborating Centre for Reference and Research on Influenza, VIDRL, Doherty Institute, Melbourne, Australia; [5]State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin, China; [6]Department of Viroscience, Erasmus MC, Rotterdam, The Netherlands; [7]University of Wisconsin, Madison, WI, USA; [8]University of Pittsburgh, Pittsburgh, PA, USA; [9]University of Georgia, Athens, GA, USA; [10]Department of Infectious Diseases, St. Jude Children's Research Hospital, Memphis, TN, USA; [11]Department of Microbiology and Immunology, University of Melbourne, Doherty Institute, Melbourne, Australia; [12]Pennsylvania State University, University Park, PA, USA

*Correspondence: JAB (jbelser@cdc.gov), EHYL (ehylau@hku.hk), HLY (hyen@hku.hk)
# These authors contributed equally to the study

Working group on the standardization of the ferret model for influenza risk assessment:
Theo Bestebroer[6], Amar Bhagwat[8], Elisabeth Blanchard[1], C Joaquin Caceres[9], Silvia Carnaccini[9], Shiho Chiba[7], Ka-Tim Choy[2], Jennifer L. DeBeauchamp[10], Dennis de Meulder[6], Lucas M. Ferreri[9], Rebecca Frise[3], Ginger Geiger[9], Meng Hu[10], Aeron C Hurt[4], Trushar Jeevan[10], Jennifer E Jones[8], Lisa A Kercher[10], Karen A Kormuth[8], Valarie Le Sage[8], Pascal Lexmond[6], Chengjun Li[5], Taronna R Maines[1], Devanshi R Patel[12], Kayla M Septer[12], Sin Fun Sia[2], Derek G Sim[12], Wen Su[2], Jie Zhou[3]

Abstract

Ferrets represent the preferred animal model for assessing the transmission potential of newly emerged zoonotic influenza viruses. However, heterogeneity among established experimental protocols and facilities across different laboratories may lead to variable results, complicating interpretation of transmission experimental data. Between 2018-2020, a global exercise was conducted by 11 participating laboratories to assess the range of variation in ferret transmission experiments using two common stock H1N1 influenza viruses that possess different transmission characteristics in ferrets. Inoculation route, dose, and volume were standardized, and all participating laboratories followed the same experimental conditions for respiratory droplet transmission, including a strict 1:1 donor:contact ratio. Additional host and environmental parameters likely to affect influenza transmission kinetics were monitored throughout. Overall transmission outcomes for both viruses across 11 laboratories were concordant, suggesting the robustness of the ferret model for zoonotic influenza risk assessment. To attain high confidence in identifying zoonotic influenza viruses with moderate-to-high or low transmissibility, our analyses support that as few as three but as many as five laboratories, respectively, would need to independently perform viral transmission experiments with concordant results. This exercise facilitates the development of a more homogenous protocol for ferret transmission experiments that are employed for the purposes of risk assessment.

Introduction

Pandemic influenza viruses may arise through interspecies transmission events of animal influenza viruses. Assessing the human-to-human transmission potential of animal influenza viruses that cause spillover infections in humans is essential for pandemic risk assessment. Ferrets have been used as a surrogate model for investigating the transmission potential of influenza viruses in humans, as they are naturally susceptible to infection with human and zoonotic influenza viruses, exhibit clinical signs during infection which closely resemble those of humans, and support influenza virus transmission via similar modes as humans. In particular, the respiratory droplet transmissibility of a specific influenza strain among ferrets often correlates with its transmission potential in humans (1). As such, ferrets are commonly used for assessing the pandemic potential of newly emerged zoonotic influenza viruses, and data from these experiments inform formal risk assessment rubrics established by the WHO and CDC (2, 3).

The transmission potential of influenza viruses is determined by multiple viral, host, and environmental parameters. As the ferret model becomes commonly employed in laboratories worldwide, there is an underappreciated heterogeneity among established experimental protocols and facility setups across different laboratories, which may lead to variable results between transmission experiments performed (4). Some of these variables, such as the dose, volume, and route of inoculation and animal age, have been confirmed to affect the kinetics of virus infection, replication, and transmission in the ferret model (5-7). However, the impact of other parameters, such as virus propagation procedures, caging designs, airflow directionality and number of air exchanges, and environmental conditions such as relative humidity, is largely unknown. Consequently, interpretation of results from ferret transmission experiments can represent a challenge when comparing data generated from multiple laboratories, even when the same virus

strain or subtype is being investigated (8). Considering the statistical limitations on small sample sizes in ferret experiments, and high potential for strain-specific variability, investigators often assess the pandemic potential of emerging virus subtypes as an aggregate of multiple viruses tested (9-11). As many public health efforts require cross-laboratory risk assessment studies for newly emerged zoonotic influenza viruses (12) and antiviral efficacy studies aiming to block influenza transmission between ferrets (13), a greater understanding of variability in transmission results obtained between independent groups is critical.

To assess the variability of ferret transmission results across laboratories under established protocols, we performed a global exercise using two common stock influenza viruses that possess different transmission characteristics in ferrets. Eleven independent laboratories inoculated ferrets with these stock viruses under uniform conditions; parameters known to affect influenza transmission kinetics were controlled in the experimental protocols while other potential parameters were carefully monitored and recorded, both prior to and during the transmission experiments. All aggregated data from these experiments were blinded and analyzed by an independent statistician. To inform future risk assessment activities, the confidence of drawing conclusions on virus transmissibility with concordant or discordant outcomes from multiple laboratories was also investigated. By assessing the range of variation present among ferret transmission experiments performed under established experimental protocols, this global exercise provides helpful guidance for data interpretation when cross-laboratory results are to be compared. The relatively concordant transmission results across 11 laboratories suggest that the ferret model is highly robust for influenza pandemic risk assessment studies under the semi-standardized conditions employed here. Furthermore, analyses investigating the role of host and environmental parameters as they contribute to virus transmission kinetics and outcomes is valuable for both

current risk assessment activities and for evaluation of countermeasures to block influenza transmission.

Results

Transmissibility of human A(H1N1)pdm09 virus. To evaluate potential heterogeneity in the transmission results between 11 laboratories, we first compared the transmissibility of a cell-grown isolate of the A(H1N1)pdm09 virus A/California/7/2009 (Cal/09), representative of early 2009 pandemic isolates and anticipated to exhibit moderate to high respiratory droplet transmissibility (14-17). Transmissibility was evaluated with 4 donor:contact pairs at a 1:1 ratio in each laboratory. Transmission to exposed respiratory droplet contact ferrets was defined by detection of infectious virus or seroconversion to the homologous virus in post-exposure sera. Following establishment of contact with donor ferrets 24 hours post-inoculation, detection of infectious virus and seroconversion in contacts was observed in 10/11 and 11/11 laboratories, respectively, with the reported transmission frequency ranging from 50-100% (Table 1). One out of 11 laboratories determined viral loads in nasal swabs and throat swabs (Group F, with throat swab viral loads used for subsequent analysis), while the other laboratories determined viral loads in nasal washes. Employing both virological and serological results, by Fisher's exact test of homogeneity, there was no significant difference in the transmission outcomes across labs with this virus ($p=0.797$). Collectively, infectious virus was detected from the nasal wash or throat swabs of 72.7% (32/44) exposed contacts, and seroconversion of contact ferrets to homologous virus was detected from 79.5% (35/44) of exposed contacts.

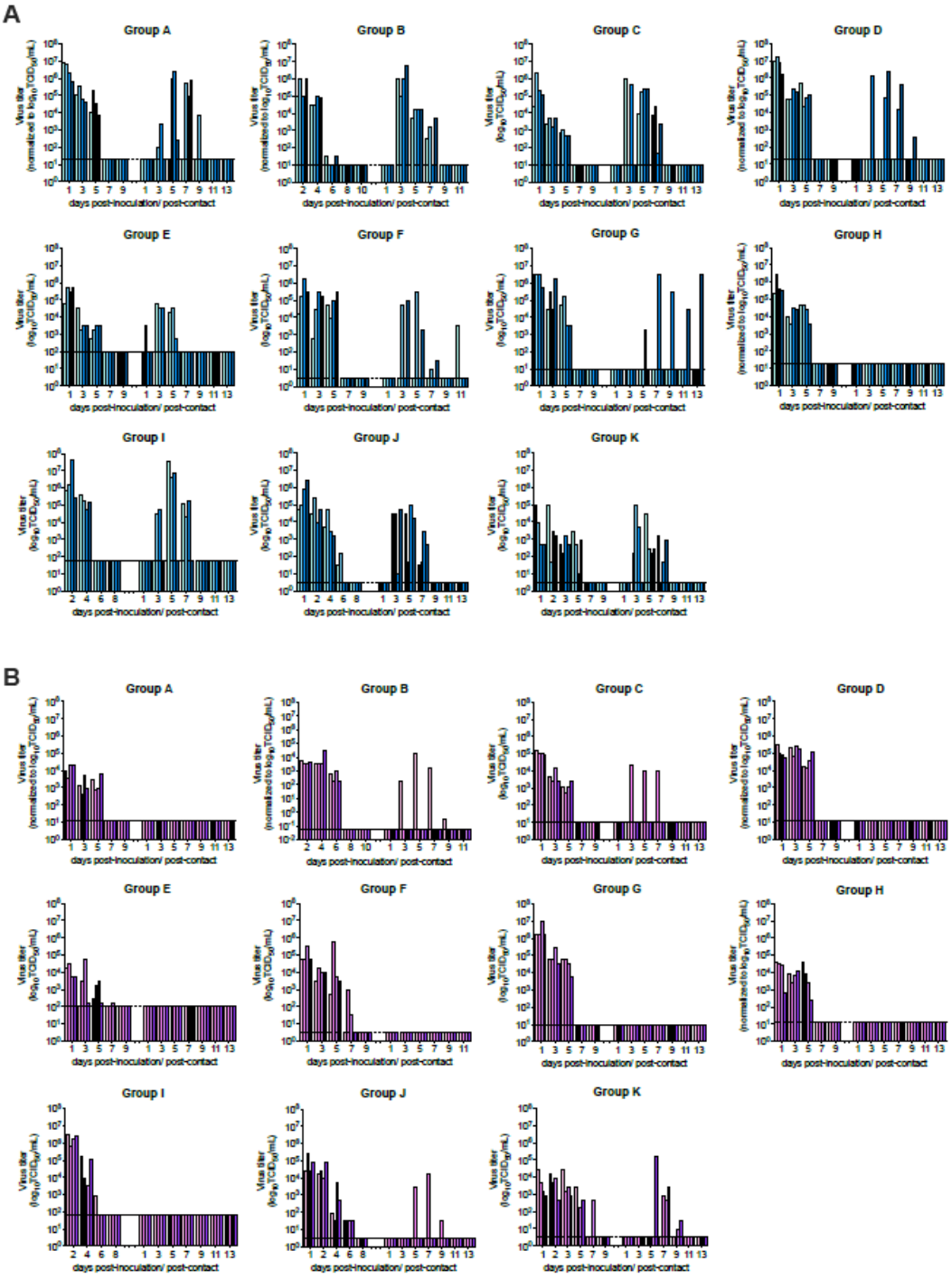
Table 1. Summary of virus transmissibility results from all laboratories.

Group	A(H1N1)pdm09 virus, A/California/7/2009			A(H1N1) avian influenza virus, A/ruddy turnstone/Delaware/300/20/2009		
	Viral load of inoculated donors (area under the curve)[a]	Transmission to aerosol contacts		Viral load of inoculated donors (area under the curve)[a]	Transmission to aerosol contacts	
		Virus detection[b]	Sero-conversion[c]		Virus detection[b]	Sero-conversion[c]
A	6.51±0.49	3/4	3/4	4.28±0.35	0/4	0/4
B	5.70±0.42	4/4	4/4	4.25±0.39	1/4	1/4
C	5.30±0.78	4/4	4/4	5.10±0.11	1/4	1/4
D	6.86±0.40	2/4	2/4	5.73±0.21	0/4	0/4
E	5.53±0.32	3/4	3/4	4.43±0.56	0/4	0/4
F	5.77±0.60	3/4	3/4	5.34±0.60	0/4	0/3
G	6.57±0.06	2/4	2/4	6.48±0.37	0/4	0/4
H	5.82±0.43	0/4	3/4	4.72±0.31	0/4	0/4
I	6.48±0.80	3/4	3/4	6.24±0.31	0/4	0/4
J	5.62±0.54	4/4	4/4	4.92±0.39	1/4	1/4
K	4.07±0.72	4/4	4/4	4.15±0.57	3/4	3/4

[a]Area under the curve (AUC) of viral titers from inoculated ferrets (following normalization of infectious units to $TCID_{50}$), expressed as log_{10} mean ± standard deviation. [b]Number of contact ferrets with infectious virus detected in respiratory specimens between 1-13 days post-contact (p.c.). [c]Number of contact ferrets that seroconverted to the exposed virus at the end of the study using hemagglutination inhibition assay.

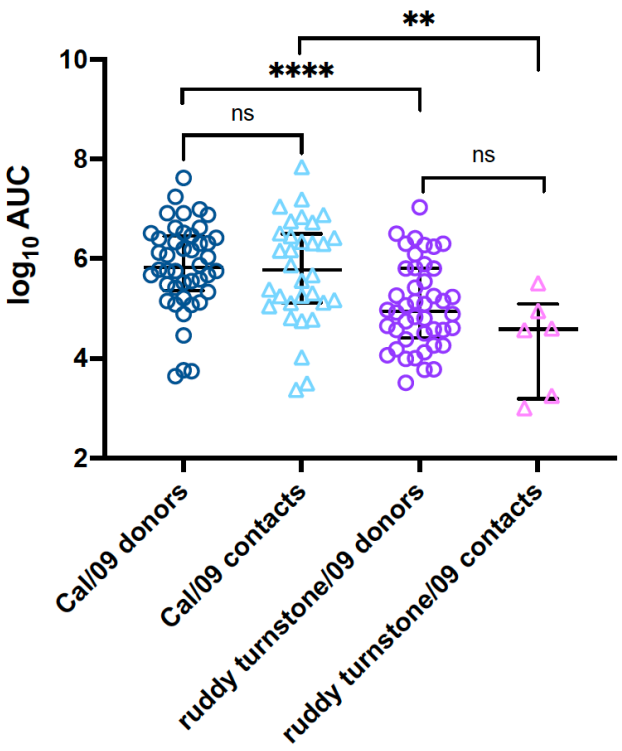
To allow comparison of the effect of viral load on transmissibility, viral titer units from nasal wash/throat swab samples (inclusive of $TCID_{50}$, PFU, and EID_{50} units, Supplemental Figures 1-2) were normalized to $TCID_{50}$ units (Figure 1), employing strain-specific conversions prior to analyses (Supplemental Table 1). From the inoculated donor ferrets, the peak viral titers detected in the nasal washes or throat swabs were at 5.72 ± 0.95, mean ± SD log_{10} $TCID_{50}$/mL after normalization, with the peak titers detected from 95.5% (42/44) of donors on the first sampling time point on 1 or 2 days post-inoculation (dpi), followed by a decline of infectious titer over time (Figure 1A). Area under the curve (AUC) after normalization was calculated to approximate total viral load shed by the Cal/09-inoculated donors, with a mean ± SD log_{10} AUC of 5.84 ± 0.89.

Figure 1. Transmission kinetics of A(H1N1) viruses in ferrets. A, normalized viral loads of donors (left bars) and aerosol contact ferrets (right bars) after inoculation or exposure to A(H1N1)pdm09 virus Cal/09. **B,** normalized viral loads of donors (left bars) and aerosol contact ferrets (right bars) after inoculation or exposure to avian H1N1 virus ruddy turnstone/09. Nasal washes (all groups except Group F) or throat swabs (Group F) were sampled to determine infectious viral loads which were normalized to $\log_{10}$ $TCID_{50}$/mL. Each bar represents individual ferrets. Limit of detection is indicated with a dashed line.



Next, to evaluate the transmission efficiency, the serial interval (first detection of viral shedding in contacts post-exposure from specimens collected every-other-day) was calculated for each infected contact ferret. The serial interval was 1 day for 3.1% (1/32) of the Cal/09 infected contact ferrets, followed by 3 days for 68.8% (22/32), 5 days for 21.9% (7/32), and 11 days for 6.3% (2/32), with a median serial interval of 3 days post-contact. Peak viral titers detected in the contact nasal washes or throat swabs were at 5.41 ± 1.06 mean ± SD $\log_{10}$ $TCID_{50}$/mL after normalization, with peak titers detected from 50% (16/32) and 34.4% (11/32) infected contacts on 3 dpi and 5 dpi, respectively. Altogether, the AUC for Cal/09 infected contact ferrets was 5.75 ± 1.05, comparable to the Cal/09 virus-inoculated donors (Mann-Whitney test, p=0.6547) (Figure 2).

Figure 2. Area under the curve of infectious viral loads detected from inoculated donors or infected contacts. Data points represent AUC values from individual ferrets from which infectious virus was detected. **, p< 0.01, ****, p< 0.0001, Mann-Whitney test.



Transmissibility of avian A(H1N1) influenza virus. We further evaluated the range of heterogeneity present in transmission results when using the A(H1N1) avian influenza virus A/ruddy turnstone/Delaware/300/2009 (ruddy turnstone/09) (18, 19), which has been reported to transmit in ferrets via respiratory droplets under the experimental setting of donor: direct contact: respiratory droplet contact at a 1:1:1 ratio, but not at a 1:1 donor:respiratory droplet contact ratio (R Fouchier, unpublished data) (18, 19). Here, the experimental setup and conditions were identical to those assessing Cal/09 virus transmissibility including a donor: respiratory droplet contact 1:1 ratio with no direct contact ferret. Transmission of an egg-derived isolate of ruddy turnstone/09 virus to exposed respiratory droplet contacts was only observed in 4 out of the 11 laboratories, with the transmission frequencies ranging from 25-75% across these four laboratories (Table 1). Compared to Cal/09 virus, there were greater differences in the ruddy turnstone/09 virus transmission outcomes across 11 laboratories, but the difference did not reach statistical significance by Fisher's exact test of homogeneity ($p=0.068$). Viral shedding and seroconversion to ruddy turnstone/09 virus were detected from 6/43 exposed contact ferrets across all laboratories, resulting in a transmission efficiency of 14.0%, which was significantly lower compared to that of Cal/09 virus (72.3%, paired t-test, $p < 0.001$).

From the inoculated donor ferrets, the peak viral titers detected in the nasal washes or throat swabs were at 4.85 ± 0.94, mean ± SD $\log_{10}$ $TCID_{50}$/mL after normalization, which was significantly lower than those detected in the Cal/09 inoculated donors (Mann-Whitney test, $p <0.0001$). Peak titers were detected from 88.6% (39/44) donors on the first sampling time point (1 or 2 dpi) followed by a decline of infectious titer over time (Figure 1B). Compared with Cal/09 virus inoculated donors, the mean ± SD $\log_{10}$ AUC of ruddy turnstone/09 virus-inoculated ferrets

was 5.06 ± 1.86, significantly lower than those inoculated with the Cal/09 virus (Mann-Whitney test, $p < 0.0001$) (Figure 2). Overall, ruddy turnstone/09 virus-inoculated donor ferrets shed lower titers of infectious virus than the Cal/09 virus-inoculated donors.

In contrast to the transmission efficiency of Cal/09 virus with a median serial interval of 3 days, for the ruddy turnstone/09 transmission experiments, the serial interval was 3 days, 5 days, or 7 days for 33.3% (2/6), 33.3% (2/6), and 33.3% (2/6) of the infected contact ferrets, respectively, with the median serial interval at 5 days. Peak viral loads (3.94 ± 0.94 mean ± SD $\log_{10}$ $TCID_{50}$/mL) detected from the six infected contact ferrets were lower when compared to the Cal/09 infected contact ferrets (Mann-Whitney test, $p=0.0022$). Peak titers were detected from 16.7% (1/6), 33.3% (2/6), and 50% (3/6) infected contacts on 3 dpi, 5 dpi, and 7 dpi, respectively. Furthermore, ruddy turnstone/09 virus-infected contact ferrets shed significantly less infectious virus (4.31 ± 0.98, mean ± SD $\log_{10}$ AUC) when compared to those animals directly inoculated with Cal/09 virus (Mann-Whitney test, $p=0.0033$) (Figure 2). Taken together, there was a longer serial interval and lower infectious virus shed by ruddy turnstone/09 virus-exposed contact ferrets when compared to those exposed to Cal/09 virus.

Contributing factors to ruddy turnstone/09 virus transmissibility. By standardizing the source stock virus, dose and volume of inoculation, and donor-to-contact ratio, we show that while infrequent discordant results were documented, the transmission outcomes of Cal/09 and ruddy turnstone/09 viruses independently performed by 11 laboratories were in general concordant, despite variabilities in the laboratory settings that were not standardized in the experiments (Supplemental Tables 2-4). As the transmission outcomes for the highly transmissible Cal/09 virus were more concordant than the less transmissible ruddy turnstone/09 virus, we attempted to

examine if any variable, including those not standardized between laboratories, may have been associated with differences in ruddy turnstone/09 virus transmissibility results.

Univariable logistic regression was performed to first evaluate if donor viral shedding kinetics were linked to ruddy turnstone/09 virus transmission efficiency. However, examination of several parameters, including AUC ($p=0.193$), peak viral titer ($p=0.197$), and days to peak titer ($p=0.473$), were not statistically associated with different transmission outcomes observed between laboratories (Supplemental Table 5), indicating that differences observed between laboratories were not attributable to virological measurements.

Numerous studies have indicated a role for environmental parameters in virus transmissibility (20, 21). Room temperature was generally consistent across all groups, with means of daily recordings within 3°C for all experiments performed (20.5-23.2°C, Supplemental Table 4). In contrast, the relative humidity reported varied widely, with regard to both the range of daily readings reported during 14-day individual experiments (varying 1-60% between low and high readings) and the mean recordings over the entirety of each experiment (32.7% to 77.0%). Despite this variability, there was no statistically significant association between transmission of ruddy turnstone/09 virus and temperature, relative humidity, or absolute humidity (all $p>0.3$, Supplemental Table 5).

Experimental cage setups varied widely between different groups, with extensive heterogeneity present with regard to cage dimensions, airflow directionality and air changes per hour, distance between cages, and other parameters (Supplemental Table 3). Groups employing caging with airflow directionality from inoculated to contact cages more frequently reported moderate to high transmissibility (≥50%) of both viruses compared with groups lacking this airflow directionality (6/6 vs 3/5 groups for Cal/09 virus, 3/6 vs 1/5 groups for ruddy turnstone/09

197 virus), however these findings did not reach statistical significance (both p>0.3, Supplemental
198 Table 5). Other specific features of cage setups, including distance between inoculated and contact
199 cages and air changes per hour (ACH) were also not statistically linked to the ruddy turnstone/09
200 transmission outcomes (both p>0.4, Supplemental Table 5). Taken together, despite substantial
201 heterogeneity in numerous non-standardized parameters in experimental setups employed between
202 groups, no one feature was identified as modulating transmission outcomes to a significant degree.
203
204 **Contributing factors to virus pathogenicity.** All ferrets inoculated with either Cal/09 or ruddy
205 turnstone/09 were productively infected, however measurements of morbidity varied between
206 groups for both viruses. Among Cal/09 virus-inoculated ferrets, mean maximum weight loss and
207 peak rise in body temperature between groups ranged from <1.0-15.6% and 0.6-2.1°C,
208 respectively (Supplemental Table 6, Supplemental Figure 3). Following ruddy turnstone/09 virus
209 inoculation, infected ferrets generally exhibited greater mean maximum weight loss (up to 19.6%)
210 and transient fevers (up to 3°C) (Supplemental Table 7, Supplemental Figure 4) compared to
211 ferrets with Cal/09 virus infections; ruddy turnstone/09-inoculated ferrets reached humane
212 experimental endpoints in 2/11 groups. The coefficient of variation between mean maximum
213 weight loss reported between groups was generally similar (56% and 52% for Cal/09 and ruddy
214 turnstone/09 viruses, respectively). No commonality with increased morbidity and ferret vendor,
215 gender, or pre-inoculation body weight was identified. Furthermore, no association was found
216 between morbidity and viral load (peak titer or AUC) or other environmental parameters, with the
217 exception of room temperature (with higher mean room temperatures associated with greater mean
218 weight loss) (Supplemental Table 8).
219

Confidence in virus transmission results generated from multiple laboratories. Collectively, the results from this exercise demonstrate a capacity for groups possessing differences in facilities designs and experimental protocols to report varying levels of relative transmissibility and pathogenicity following inoculation of ferrets with the same virus. To illustrate how confidence in risk assessments of virus transmissibility can increase as results from multiple groups are combined, we evaluated the hypothetical risk of a virus capable of moderate to high transmission (defined as p ≥50% transmission events per total pairs of ferrets as defined in Table 2) or non-transmissible (defined as p≤25% transmission events). In these analyses, concordant results are defined as multiple groups identifying a virus exhibiting the same transmission capacity, and discordant results are defined as multiple groups identifying a virus with different transmission capacities, as defined above. By assuming concordant results across laboratories which permits pooling of all transmission outcomes, as few as three groups (12 pairs of ferrets) will yield a probability of over 80% to conclude moderate to high transmissibility when transmission was observed in at least half of all experiments, and a probability of over 85% to conclude low transmissibility when at most one transmission event was observed over all experiments.

Table 2. Confidence in conclusions derived from pooled samples from multiple laboratories.

No. labs[a]	No. transmission pairs tested	No. pairs with transmission[b]	Probability of moderate to high transmissibility[c]	No. pairs without transmission	Probability of low transmissibility[d]
1	4	4	99%	0	76%
		3	88%	1	37%
2	8	≥7	>99%	0	92%
		6	99%	1	70%
		5	92%	2	40%
		4	68%		
3	12	≥8	>99%	0	98%
		7	95%	1	87%
		6	82%	2	67%
				3	42%
4	16	≥10	>99%	0	99%
		9	97%	1	95%
		8	90%	2	84%
		7	73%	3	65%
				4	43%

[a]Number of laboratories providing results (assumes 4 pairs of donor:contact at 1:1 ratio per laboratory; transmission in each pair is an independent event). [b]Number of transmission events among total number of ferrets from all groups specified (defined as detection of infectious virus and seroconversion to homologous virus in contact ferrets). [c]Moderate to high transmissibility is defined as p≥50% (eg. ≥2 infected out of 4 ferrets). [d]Low transmissibility is defined as p≤25% (eg. 0 or 1 infected out of 4 ferrets).

Alternatively, a voting system can be considered by first drawing a conclusion on transmissibility in each laboratory, with an overall conclusion drawn based on these 'votes' from multiple labs. When testing for moderate to high transmissibility, and assuming n=4 ferrets per laboratory, 3 laboratories are needed to conclude moderate to high transmissibility with confidence >90% if concordant results are obtained. In agreement with probabilities shown in Table 3, a greater number of laboratories contributing results are needed to demonstrate statistically significant results when testing for low transmissibility; to conclude low transmissibility with

252 >90% confidence, this would necessitate 5 contributing laboratories if concordant results are
253 obtained. In this scenario, a greater number of contributing laboratories (or a greater number of
254 donor:contact pairs per laboratory) would be required if the true transmission probability was
255 higher for confirming low transmissibility, or when the true transmission probability was lower
256 for confirming moderate to high transmissibility.

257

258 **Table 3. Confidence in conclusions derived from multiple laboratories considering a voting**
259 **system.**

No. labs[a]	No. labs with concordant result[b]	Probability of moderate to high transmissibility[c]	Probability of low transmissibility[d]
1	1	76%	56%
2	2	87%	72%
	1 (discordant)	31%	18%
3	3	92%	81%
	2 (discordant)	48%	32%
4	4	95%	87%
	3 (discordant)	62%	45%
	2 (discordant)	23%	12%
5	5	97%	95%
	4 (discordant)	72%	56%
	3 (discordant)	35%	21%
6	6	98%	94%
	5 (discordant)	80%	65%
	4 (discordant)	47%	30%

260 [a]Number of laboratories providing votes on the transmissibility of the tested virus. Each
261 laboratory will vote if the tested virus possess moderate to high transmissibility (p≥50%, eg. ≥2
262 infected out of 4 ferrets) or low transmissibility (p≤25%, eg. 0 or 1 infected out of 4 ferrets)
263 based on the experimental result. [b]Number of laboratories with concordant results on viral
264 transmissibility, or the number of laboratories reaching concordant results when discordant
265 results are included. [c]Moderate to high transmissibility is defined as p≥50% (eg. ≥2 infected out
266 of 4 ferrets). [d]Low transmissibility is defined as p≤25% (eg. 0 or 1 infected out of 4 ferrets).
267

268 Despite generally consistent results between all groups in this exercise, discordant results
269 are possible (Table 1), highlighting the need to better understand how to responsibly interpret and

account for these findings. As such, we also considered the scenario when discordant results between laboratories are recorded. To demonstrate moderate to high transmissibility, we found that 6 laboratories with 1 discordant result could still provide 80% confidence in the conclusion, while any discordant result significantly reduced confidence for concluding low transmissibility (Table 3). In both scenarios, if the results from different laboratories were more heterogeneous, the uncertainty around the conclusion from each lab increases and the overall confidence would decrease. This exercise is an illustration of the possible scenarios and confidence in drawing conclusions on transmissibility but would be affected by how moderate to high or low transmissibility were defined.

Discussion

The importance of the ferret model for influenza virus risk assessment studies cannot be understated (4, 22). Recent advances in molecular biology, aerobiology, genomics, and other areas highlight the ways the ferret model in general, and studies evaluating virus transmissibility by the airborne route specifically, continue to contribute towards our understanding of influenza viruses and the threat they pose to human health (23-25). However, as this model becomes more commonly employed in laboratories worldwide, there is a pressing need to capture the level of variability and heterogeneity intrinsic to this research. Cross-laboratory exercises have been employed in the past to evaluate the reproducibility of assays employed for influenza virus public health efforts (26), but no such exercise has been performed to date evaluating influenza virus transmissibility in the ferret. In this exercise, 11 laboratories across different continents independently evaluated the transmission potential of Cal/09 and ruddy turnstone/09 viruses with distinct transmission potential. With only a few experimental parameters (common virus stock,

293 standardized inoculation dose, route, volume, and the 1:1 donor:contact ratio) being controlled
294 across the participating laboratories, we observed homogenous transmission outcomes (that is,
295 outcomes did not differ statistically) across laboratories. Our results demonstrate the robustness of
296 the ferret model in influenza risk assessment studies.

297 Risk assessment rubrics have thoroughly evaluated a wide scope of influenza A viruses,
298 from viruses associated with poultry outbreaks in the absence of confirmed human infections, to
299 viruses such as A(H5N1) and A(H7N9) influenza viruses that have caused substantial human
300 disease and death (3, 27). As such, there is a need to evaluate heterogeneity of ferret transmission
301 models employing viruses possessing a similar scope of transmissibility phenotypes. While the
302 variability in transmission results for either the Cal/09 or ruddy turnstone/09 viruses tested in this
303 study were not statistically significant, the range of results obtained, especially with the ruddy
304 turnstone/09 virus, nonetheless illustrates a level of variability that can be present in transmission
305 readouts of viruses exhibiting both low to high transmission efficiency (Table 1). This variability
306 was present despite a high degree of standardization of virus stock, inoculation procedures, and
307 uniformity of donor:contact ratio.

308 As shown in the Supplemental Methods and Supplemental Tables 1-6, this exercise
309 captured the extensive heterogeneity in laboratory protocols and setups present between different
310 groups. Documented variation was present in every parameter examined, inclusive of ferrets, cage
311 setups, titration methods, and environmental conditions, among other features. Caging and airflow
312 considerations were especially variable (Supplemental Table 2). It is impossible to standardize all
313 contributing variables to these experiments, as institutional, animal welfare, and governmental
314 guidelines and requirements vary worldwide, as do cost implications. That said, this exercise
315 supports the capacity to harmonize results generated between disparate groups when a small

number of procedural parameters are fixed. Interestingly, the four groups that detected infectious virus in contact nasal wash specimens in ruddy turnstone/09 transmission experiments all found 4/4 virus transmission in the Cal/09 experiment; transmission percentages between the two viruses were highly correlated between laboratories (Spearman correlation = 0.86, $p < 0.001$). Furthermore, while directional airflow (OR=4) did not reach statistical significance, it is nonetheless of note that 3/4 laboratories for which ruddy turnstone/09 virus transmission was detected possessed directional airflow, versus 3/7 of the laboratories for which transmission with this virus was not detected; directional airflow from inoculated to contact animals was a feature in 6/11 laboratories in this exercise (Supplemental Table 3). While our results did not conclusively identify any one experimental parameter statistically associated with enhanced transmissibility outcomes, it is possible that a confluence of parameters is nonetheless capable of creating a more permissive environment for virus transmission to occur.

To improve interpretation of results from this standardization exercise, we concurrently investigated the hypothetical confidence in concluding low transmissibility (≤25% or ≤1 ferret infected out of 4 ferrets) or moderate to high transmissibility (≥50% or ≥2 ferrets infected out of 4 ferrets) from multiple contributing laboratories. These analyses assumed both a uniform prior distribution for the transmission probability for a novel pathogen, and independent transmission outcomes from the laboratories. We considered two scenarios: one scenario where strong homogeneity across laboratories could be assumed so the samples were pooled from multiple laboratories, and another scenario where each laboratory drew their own conclusion on transmissibility such that an overall conclusion was drawn as a voting system. As influenza viruses of notable public health importance are frequently assessed across multiple independent laboratories, these analyses provide a framework to rigorously interpret independently generated

339 findings, especially when discordant results between laboratories are reported. This is most critical
340 in the event of a novel virus believed to possess moderate-to-high transmissibility; our analyses
341 support that 4 independent laboratories with concordant results supporting an enhanced
342 transmissibility phenotype yields a 95% probability of this finding, with additional independent
343 groups or a greater number of total ferret donor:contact pairs necessary when discordant results
344 are present.

345 Collectively, the findings of this exercise support the potential benefit of increased
346 uniformity, or standardization, of some parameters when conducting risk assessment-specific
347 activities on the same viruses. Specifically, the donor:contact ratio represents such a parameter.
348 For a virus with moderate to high transmissibility, such as Cal/09 virus, modulation of this ratio
349 (e.g., conducting experiments with a 2:1 donor:contact ratio, as is the case when transmission
350 evaluations in a direct contact setting and via respiratory droplets employ a common donor) would
351 not substantially alter conclusions drawn. However, for a virus with reduced transmissibility at a
352 1:1 ratio, such as the ruddy turnstone/09 virus evaluated here, it is likely that an increased
353 donor:contact ratio (eg., 2:1) may enhance transmissibility by increasing virus-laden aerosols
354 exhaled from infected ferrets. Previous studies on ruddy turnstone/09 virus demonstrated airborne
355 transmission potential when employing a donor: direct contact: aerosol contact at 1:1:1 ratio;
356 efficient transmission by direct contact will subsequently affect the quantity and kinetics of virus-
357 laden aerosols that mediate transmission by air (18, 19). There is a need to better understand how
358 modulation of this ratio contributes to assessments of virus transmissibility. However, this does
359 underscore the potential complications posed by harmonizing data generated for risk assessment
360 purposes for which the donor:contact ratio diverges. With increased heterogeneity in results
361 between labs, uncertainty around the conclusions increases, and there is a corresponding decrease

362 in confidence in the results (Table 3), showing the utility in increasing homogeneity across findings
363 from different labs in order to reduce the total number of labs required to yield statistically
364 meaningful results in this sort of analysis.

365 The emergence of SARS-CoV-2 further corroborates the pandemic potential of viruses of
366 zoonotic origin. Early identification and risk assessments of novel viruses are essential for
367 preventing the next pandemic. Continued optimization and refinement of risk assessment protocols
368 will facilitate data interpretation in response to emerging pandemic threats. Collectively, a greater
369 appreciation of this heterogeneity, and understanding of the scope of variability present in risk
370 assessment settings, will permit more robust conclusions to be drawn from these efforts in the
371 future.

372

373 **Materials and methods**

374 **Viruses**. The A(H1N1)pdm09 virus A/California/07/2009 (Cal/09) was propagated in MDCK
375 cells (passage C3) at the US CDC as described previously (28). The low pathogenic avian influenza
376 A(H1N1) virus A/ruddy turnstone/Delaware/300/2009 (ruddy turnstone/09) was propagated in
377 eggs (passage E3) by St. Jude Children's Research Hospital as described previously (19). Stocks
378 were fully sequenced and tested for exclusivity to rule out the presence of other influenza virus
379 subtypes prior to distribution.

380 **Animal and experimental variability.** Groups obtained ferrets from multiple vendors and
381 independent breeders from North America, Europe and Asia, and animals varied in their age,
382 gender, health status, and other parameters (Supplemental Table 1). There was substantial
383 differences between laboratories in the specific caging employed for transmission experiments,
384 distance between cages, airflow directionality between cages, and air changes per hour

(Supplemental Table 2). Anesthesia protocols, sample collection methods, and decontamination procedures to prevent cross-contamination between contact and donor animals varied between groups and are reported in Supplemental Methods. All experiments were performed under country-specific legal guidelines and approved institutional-specific animal protocols as specified in the Supplemental Methods.

Standardized procedures. All laboratories received common stock viruses prepared by CDC and St. Jude Children's Research Hospital with the shipping temperature recorded. Stock viruses were diluted to 10^6 plaque forming units (PFU) in 500µl PBS based on predetermined viral titers, and donor ferrets were inoculated intranasally under in-house protocols for anesthesia (Supplemental Methods). On day 1 post-inoculation, one respiratory droplet contact ferret was introduced and exposed to each donor by housing in an adjacent cage, employing a strict 1:1 donor:contact ratio, with 4 transmission pairs tested for each virus. Ferret temperatures, weights, and nasal washes/swabs were collected every 24-48 hours. Daily room temperature and relative humidity readings were collected and are reported in Supplemental Table 3 employing pre-validated thermohygrometers with comparable readings (Testo Inc., 608-H1). Sera was collected at the end of each experiment for determination of seroconversion to homologous virus by hemagglutinin inhibition assay using established in-house serology protocols.

Sample titration and normalization. Infectious virus titers were determined by plaque assay, 50% tissue culture infectious dose ($TCID_{50}$) assay, or 50% egg infectious doses (EID_{50}) assay at each laboratory with varying limits of detection (Supplemental Table 4). To facilitate subsequent statistical assessments across laboratories, reported titers from each laboratory were normalized to $TCID_{50}$/mL for each virus based on PFU, $TCID_{50}$, and EID_{50} values pre-determined by a single laboratory to minimize titration methodology-specific variation.

Data blinding and analyses. Data blinding, aggregation and all statistical analyses were performed by an independent statistician. Transmission outcomes were compared across laboratories by each virus, using Fisher's exact test of homogeneity. Viral load between viruses were compared by testing difference in area under the curve (AUC) using t-test. Factors associated with transmissibility and morbidity were assessed by using logistic regression and linear regression models. We also investigated the confidence in concluding low transmissibility (≤25%, or ≤1 ferret infected out of 4 ferrets) or moderate to high transmissibility (≥50% or ≥2 ferrets infected out of 4 ferrets) from multiple contributing laboratories. We assumed a uniform prior distribution for the transmission probability for a novel pathogen was assumed, and independent transmission outcomes from the laboratories. The confidence of drawing conclusion on transmissibility with concordant or discordant outcomes from the laboratories is presented. We considered a scenario where strong homogeneity across laboratory can be assumed so the samples were pooled from multiple laboratories, and another scenario that each laboratory draw their own conclusion on transmissibility and the overall conclusion was drawn as voting system. All analyses were conducted in R version 4.0.4 (R Development Core Team).

Acknowledgements

This study was supported by Contract HHSN272201400006C from NIAID, NIH, USA. The findings and conclusions in this report are those of the authors and do not necessarily reflect the views of the Centers for Disease Control and Prevention/Agency for Toxic Substances and Disease Registry.

References

1. Buhnerkempe MG, Gostic K, Park M, Ahsan P, Belser JA, Lloyd-Smith JO. 2015. Mapping influenza transmission in the ferret model to transmission in humans. Elife 4.
2. Cox NJ, Trock SC, Burke SA. 2014. Pandemic preparedness and the Influenza Risk Assessment Tool (IRAT). Curr Top Microbiol Immunol 385:119-36.
3. WHO. 2016. Tool for Influenza Pandemic Risk Assessment (TIPRA).
4. Belser JA, Barclay W, Barr I, Fouchier RAM, Matsuyama R, Nishiura H, Peiris M, Russell CJ, Subbarao K, Zhu H, Yen HL. 2018. Ferrets as Models for Influenza Virus Transmission Studies and Pandemic Risk Assessments. Emerg Infect Dis 24:965-971.
5. Moore IN, Lamirande EW, Paskel M, Donahue D, Kenney H, Qin J, Subbarao K. 2014. Severity of clinical disease and pathology in ferrets experimentally infected with influenza viruses is influenced by inoculum volume. J Virol 88:13879-91.
6. Gustin KM, Belser JA, Wadford DA, Pearce MB, Katz JM, Tumpey TM, Maines TR. 2011. Influenza virus aerosol exposure and analytical system for ferrets. Proc Natl Acad Sci U S A 108:8432-7.
7. Bodewes R, Kreijtz JH, van Amerongen G, Fouchier RA, Osterhaus AD, Rimmelzwaan GF, Kuiken T. 2011. Pathogenesis of Influenza A/H5N1 virus infection in ferrets differs between intranasal and intratracheal routes of inoculation. Am J Pathol 179:30-6.
8. Belser JA, Eckert AM, Tumpey TM, Maines TR. 2016. Complexities in Ferret Influenza Virus Pathogenesis and Transmission Models. Microbiol Mol Biol Rev 80:733-44.
9. Belser JA, Maines TR, Katz JM, Tumpey TM. 2013. Considerations regarding appropriate sample size for conducting ferret transmission experiments. Future Microbiol 8:961-5.
10. Nishiura H, Yen HL, Cowling BJ. 2013. Sample size considerations for one-to-one animal transmission studies of the influenza A viruses. PLoS One 8:e55358.
11. Linster M, van Boheemen S, de Graaf M, Schrauwen EJA, Lexmond P, Manz B, Bestebroer TM, Baumann J, van Riel D, Rimmelzwaan GF, Osterhaus A, Matrosovich M, Fouchier RAM, Herfst S. 2014. Identification, characterization, and natural selection of mutations driving airborne transmission of A/H5N1 virus. Cell 157:329-339.
12. Martinez-Sobrido L, Blanco-Lobo P, Rodriguez L, Fitzgerald T, Zhang H, Nguyen P, Anderson CS, Holden-Wiltse J, Bandyopadhyay S, Nogales A, DeDiego ML, Wasik BR, Miller BL, Henry C, Wilson PC, Sangster MY, Treanor JJ, Topham DJ, Byrd-Leotis L, Steinhauer DA, Cummings RD, Luczo JM, Tompkins SM, Sakamoto K, Jones CA, Steel J, Lowen AC, Danzy S, Tao H, Fink AL, Klein SL, Wohlgemuth N, Fenstermacher KJ, El Najjar F, Pekosz A, Sauer L, Lewis MK, Shaw-Saliba K, Rothman RE, Liu ZY, Chen KF, Parrish CR, Voorhees IEH, Kawaoka Y, Neumann G, Chiba S, Fan S, Hatta M, Kong H, Zhong G, et al. 2020. Characterizing Emerging Canine H3 Influenza Viruses. PLoS Pathog 16:e1008409.

13. Lee LYY, Zhou J, Frise R, Goldhill DH, Koszalka P, Mifsud EJ, Baba K, Noda T, Ando Y, Sato K, Yuki AI, Shishido T, Uehara T, Wildum S, Zwanziger E, Collinson N, Kuhlbusch K, Clinch B, Hurt AC, Barclay WS. 2020. Baloxavir treatment of ferrets infected with influenza A(H1N1)pdm09 virus reduces onward transmission. PLoS Pathog 16:e1008395.
14. Pulit-Penaloza JA, Jones J, Sun X, Jang Y, Thor S, Belser JA, Zanders N, Creager HM, Ridenour C, Wang L, Stark TJ, Garten R, Chen LM, Barnes J, Tumpey TM, Wentworth DE, Maines TR, Davis CT. 2018. Antigenically Diverse Swine Origin H1N1 Variant Influenza Viruses Exhibit Differential Ferret Pathogenesis and Transmission Phenotypes. J Virol 92.
15. Baz M, Boonnak K, Paskel M, Santos C, Powell T, Townsend A, Subbarao K. 2015. Nonreplicating influenza A virus vaccines confer broad protection against lethal challenge. MBio 6:e01487-15.
16. Lakdawala SS, Lamirande EW, Suguitan AL, Jr., Wang W, Santos CP, Vogel L, Matsuoka Y, Lindsley WG, Jin H, Subbarao K. 2011. Eurasian-origin gene segments contribute to the transmissibility, aerosol release, and morphology of the 2009 pandemic H1N1 influenza virus. PLoS Pathog 7:e1002443.
17. Munster VJ, de Wit E, van den Brand JM, Herfst S, Schrauwen EJ, Bestebroer TM, van de Vijver D, Boucher CA, Koopmans M, Rimmelzwaan GF, Kuiken T, Osterhaus AD, Fouchier RA. 2009. Pathogenesis and transmission of swine-origin 2009 A(H1N1) influenza virus in ferrets. Science 325:481-3.
18. Kocer ZA, Krauss S, Stallknecht DE, Rehg JE, Webster RG. 2012. The potential of avian H1N1 influenza A viruses to replicate and cause disease in mammalian models. PLoS One 7:e41609.
19. Zanin M, Wong SS, Barman S, Kaewborisuth C, Vogel P, Rubrum A, Darnell D, Marinova-Petkova A, Krauss S, Webby RJ, Webster RG. 2017. Molecular basis of mammalian transmissibility of avian H1N1 influenza viruses and their pandemic potential. Proc Natl Acad Sci U S A 114:11217-11222.
20. Marr LC, Tang JW, Van Mullekom J, Lakdawala SS. 2019. Mechanistic insights into the effect of humidity on airborne influenza virus survival, transmission and incidence. J R Soc Interface 16:20180298.
21. Gustin KM, Belser JA, Veguilla V, Zeng H, Katz JM, Tumpey TM, Maines TR. 2015. Environmental Conditions Affect Exhalation of H3N2 Seasonal and Variant Influenza Viruses and Respiratory Droplet Transmission in Ferrets. PLoS One 10:e0125874.
22. Albrecht RA, Liu WC, Sant AJ, Tompkins SM, Pekosz A, Meliopoulos V, Cherry S, Thomas PG, Schultz-Cherry S. 2018. Moving Forward: Recent Developments for the Ferret Biomedical Research Model. mBio 9.
23. Karlsson EA, Meliopoulos VA, Savage C, Livingston B, Mehle A, Schultz-Cherry S. 2015. Visualizing real-time influenza virus infection, transmission and protection in ferrets. Nat Commun 6:6378.

24. Yen HL, Liang CH, Wu CY, Forrest HL, Ferguson A, Choy KT, Jones J, Wong DD, Cheung PP, Hsu CH, Li OT, Yuen KM, Chan RW, Poon LL, Chan MC, Nicholls JM, Krauss S, Wong CH, Guan Y, Webster RG, Webby RJ, Peiris M. 2011. Hemagglutinin-neuraminidase balance confers respiratory-droplet transmissibility of the pandemic H1N1 influenza virus in ferrets. Proc Natl Acad Sci U S A 108:14264-9.
25. Varble A, Albrecht RA, Backes S, Crumiller M, Bouvier NM, Sachs D, Garcia-Sastre A, tenOever BR. 2014. Influenza A virus transmission bottlenecks are defined by infection route and recipient host. Cell Host Microbe 16:691-700.
26. Stephenson I, Heath A, Major D, Newman RW, Hoschler K, Junzi W, Katz JM, Weir JP, Zambon MC, Wood JM. 2009. Reproducibility of serologic assays for influenza virus A (H5N1). Emerg Infect Dis 15:1252-9.
27. CDC. 2021. Summary of Influenza Risk Assessment Tool (IRAT) Results. https://www.cdc.gov/flu/pandemic-resources/monitoring/irat-virus-summaries.htm. Accessed Feb 10 2021.
28. Maines TR, Jayaraman A, Belser JA, Wadford DA, Pappas C, Zeng H, Gustin KM, Pearce MB, Viswanathan K, Shriver ZH, Raman R, Cox NJ, Sasisekharan R, Katz JM, Tumpey TM. 2009. Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. Science 325:484-7.