Birth year, immune history and differences in risk from seasonal influenza H1N1 and H3N2

Katelyn M Gostic1, Jennifer A. Whitaker2, Tristan W. Clark3, Lars Østergaard4, Rebecca Bridge6, Shane Brady6, Cecile Viboud7, Michael Worobey8, and James O Lloyd-Smith1,7\*

1Dept. of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA

2 Molecular Virology and Microbiology; Section of Infectious Diseases, Baylor College of Medicine, Houston, TX

3 School of Clinical and Experimental Sciences, University of Southampton, Southampton, UK

4Dept. of Infectious Diseases, Aarhus University Hospital, 8200 Aarhus N, Denmark

6Arizona Department of Health Services, Phoenix Arizona, USA

6Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

7Dept. of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA

\* [jlloydsmith@ucla.edu](mailto:jlloydsmith@ucla.edu)

# Abstract

Across decades of co-circulation in humans, influenza A subtypes H1N1 and H3N2 have caused seasonal epidemics characterized by different age distributions of infection and mortality. H3N2 causes the majority of cases in high-risk elderly cohorts, and the majority of overall deaths, whereas H1N1 causes incidence shifted towards young and middle-aged adults, and fewer deaths. These contrasting age profiles may result from differences in childhood exposure to H1N1 and H3N2 or differences in evolutionary rate between subtypes. Here we analyze two detailed epidemiological data sets to test whether childhood immune imprinting shapes seasonal flu epidemiology, and if so, whether it acts primarily via narrow immune memory to a particular antigenic subtype of influenza or via broader immune memory that protects across subtypes. We also test the impact of evolutionary differences between influenza subtypes on the age distributions of cases they cause. Likelihood-based model comparison showed that narrow, within-subtype imprinting is the strongest driver of seasonal influenza risk. The data did not support a strong effect of evolutionary rate, or of broadly protective imprinting. Our findings emphasize that childhood exposures can imprint a lifelong immunological bias toward particular influenza subtypes, and that these cohort-specific biases shape epidemic age distributions. As a result, less “senior” antibody responses acquired later in life do not provide the same strength of protection as responses imprinted in childhood. Finally, we project that H1N1’s mortality burden (currently low) may increase in the coming decades, as cohorts that lack H1N1-specific imprinting eventually reach old age.

# Introduction

Childhood influenza exposures leave an immunological imprint, which has reverberating, lifelong impacts on immune memory. Foundational work on original antigenic sin (1) and antigenic seniority (2) showed, using serological data, that individuals maintain the highest antibody titers against influenza strains encountered in childhood. But how these serological patterns map to functional immune protection, and shape birth year-specific risk during outbreaks, remains an active area of inquiry. One open question is the breadth of cross-protection provided by immune memory imprinted in childhood, and the extent to which immune breadth depends on the antigenic familiarity of the challenge strain.

Many epidemiological examples highlight benefits from imprinting protection; every modern influenza pandemic has spared certain birth cohorts, presumably due to cross-protective memory primed in childhood (3–9). Recently, we showed that childhood imprinting also protects against novel, emerging avian influenza viruses (8,10). Childhood imprinting may additionally shape birth year-specific risk from seasonal influenza (11–13), but it is unclear what breadth of imprinting protection acts dominantly in this context.

Recent studies have highlighted childhood imprinting’s ability to shape multiple layers of influenza immune memory, both broad and narrow. Until recently, narrow influenza immunity, with cross-protection acting only among closely related variants of the same hemagglutinin (HA) subtype, has been considered the norm. Narrow, within-subtype protection against HA is primarily driven by memory of variable epitopes on the HA head. But protection may also be driven by memory of other antigens (e.g. neuraminidase, NA) (14–16), or by more broadly protective responses that target conserved epitopes, particularly on the HA stalk (10,17–19). Antibodies that target conserved HA epitopes can provide broad protection across HA subtypes in the same phylogenetic group (17,19,20), where HA group 1 contains seasonal subtypes H1 and H2, while group 2 contains seasonal H3 (10,18,21).

Narrow within-subtype responses are known to influence seasonal influenza’s epidemiology and evolution. But because narrow, within-subtype immune memory decays rapidly in the face of antigenic drift, narrow responses imprinted in childhood may not continue to shape cohort-specific protection across an entire human lifetime. Alternatively, broad, HA group-level immune memory can play a strong role in immune defense against strains whose variable HA epitopes are unfamiliar (17,19,20,22). Thus, group-level responses may provide a second line of defense against drifted seasonal strains whose conserved epitopes provide the only recognizable immune targets.

Since 1977, two distinct subtypes of influenza A, H1N1 and H3N2, have circulated seasonally in humans, with striking but poorly understood differences in their age-specific impact (8,11–13,23). These differences could be associated with childhood imprinting: older cohorts were almost certainly exposed to H1N1 in childhood (since it circulated from 1918-1957), and now seem to be preferentially protected against modern seasonal H1N1 variants (8,11–13). Likewise, younger adults have the highest probabilities of childhood imprinting to H3N2, which is consistent with relatively low incidence of seasonal H3N2 in these cohorts (***Fig. 1A***). Alternatively, differences in the evolutionary dynamics of H1N1 and H3N2 could explain the observed age profiles. Subtype H3N2 exhibits slightly faster drift in its antigenic phenotype than H1N1, and as a result, H3N2 may be more able to cause infections in older, immunologically experienced cohorts, whereas H1N1 may be relatively restricted to infecting immunologically naïve children (24).

Using two large data sets on seasonal influenza incidence, which together represent 13,063 confirmed influenza A cases across nearly two decades and 15 countries, we tested whether cohort effects from childhood imprinting primarily act against variable epitopes, only providing cross-protection against closely related HA or NA variants of the same subtype, or against more conserved epitopes, providing broad cross-protection across subtypes in the same HA group (***Fig. 1A-B***). Additionally, we considered the hypothesis that differences in evolutionary rate of H1N1 and H3N2, rather than imprinting effects, shape differences in age distribution. Our results have implications for long-term projections of seasonal influenza risk in elderly cohorts (12), who suffer the heaviest burdens of influenza-related morbidity and mortality, and whose imprinting status will shift through time as cohorts born during different inter-pandemic eras grow older.

# The Data

We analyzed two large epidemiological data sets. First, the Arizona Department of Health Services (ADHS) provided a dataset containing 9,451 seasonal H1N1 and H3N2 cases from their statewide surveillance system. Cases of all ages were confirmed to subtype by PCR and/or culture, primarily from virologic testing at the Arizona State Public Health Laboratory, with a smaller number of positive influenza tests obtained through mandatory reporting by other clinical labs (25). Cases were observed across 22 years of influenza surveillance, from the 1993-1994 influenza season through the 2014-2015 season, although sample sizes increased dramatically after the 2009 pandemic (Table 1).

A second data set provided by the INSIGHT influenza outpatient study (<http://insight.ccbr.umn.edu/>) contained 3,612 PCR-confirmed H1N1 and H3N2 cases, observed across 16 countries between 2009 and 2016 (***Table 2***). The study enrolled adults ages 18 and over who sought health care for influenza-like illness (ILI) as outpatients at participating clinics. The INSIGHT data sampled a greater geographical range, and contained information not available in the ADHS data, including the age distribution of patients with ILI who consented to the study who later tested negative for influenza, and targeted medical history.

We excluded cases observed during the 2009 H1N1 pandemic from both data sets (***Methods***), as pandemic age distributions and underlying drivers of immune memory differed from the normal, seasonal circulation patterns of interest to this study (13,17,20). The ***Methods*** provide additional definitions and exclusion criteria specific to each data set, but here we emphasize a few important points. First, note that age-specific sampling differed between data sets. The INSIGHT study excluded children under age 18, and enrolled relatively few elderly subjects, whereas the ADHS data contained large numbers of cases at the extremes of age, including in children (***Fig. S1,S2,S4-S5***). Second, the INSIGHT outpatient study did not enroll patients who were already hospitalized for influenza (hospitalized cases were the focus of a separate INSIGHT study, not analyzed here), whereas ADHS data captured positive test results from hospitals, long-term care facilities, correctional facilities, and outpatient clinics.

# The Model

## Reconstructed imprinting patterns

Reconstructed, birth year-specific probabilities of childhood imprinting are based on patterns of first childhood exposure to influenza A and reflect historical circulation (Fig. 1A). Most individuals born between pandemics in 1918 and 1957 experienced a first influenza A virus (IAV) infection by H1N1, and middle-aged cohorts born between pandemics in 1957 and 1968 almost all were first infected by H2N2 (note that because the first influenza exposure may occur after the first year of life, individuals born in the years leading up to a pandemic have some probability of first infection by the new pandemic subtype). Ever since its emergence in 1968, H3N2 has dominated seasonal circulation in humans, and caused the majority of first infections in younger cohorts. However, H1N1 has also caused some seasonal circulation since 1977, and has imprinted a fraction of all cohorts born since the mid-1970s (Fig. 1A).

We reconstructed birth year-specific probabilities of first infection by H1N1, H2N2 and H3N2 using methods described previously (10). Whenever supporting data was available, reconstructions incorporated country-specific differences in the relative dominance of H1N1 and H3N2 in recent influenza seasons. Reconstructions assumed children age 0-12 in the year of case observation might not yet have been exposed to any influenza virus. Interactions between imprinting and vaccination of naïve infants are possible, but poorly understood (10,26). We did not consider childhood vaccination effects here; none of the individuals in the INSIGHT data and few individuals in the ADHS data were born at a time when healthy infants were routinely vaccinated against influenza. We did however consider the effect of vaccination in the current influenza season in the INSIGHT data.

## Expected age distributions under alternate imprinting models

If narrow, HA subtype-level imprinting protection shapes seasonal influenza risk, primary exposure to H1 or H3 in childhood should provide lifelong protection against modern variants of the same HA subtype. If imprinting protection acts primarily against specific NA subtypes, lifelong protection will be specific to N1 or to N2 (Fig. ***1***). Alternatively, if broad HA group-level imprinting shapes seasonal influenza risk, then cohorts imprinted to H1 or H2 (both group 1) should be protected against modern, seasonal H1N1, while only cohorts imprinted to H3 (group 2) would be protected against modern, seasonal H3N2 (Fig. 1B). Collinearities between the predictions of different imprinting models (Fig. 1G-I) were inevitable, given the limited diversity of influenza circulation in humans over the past century. Note that middle-aged cohorts, which were first infected by H2N2, are crucial, because they provide the only leverage to differentiate between imprinting at the HA subtype, NA subtype or HA group-level level (Fig. 1B).

To tease apart age-specific risk factors from birth year-specific imprinting effects, we noted that age-specific risk factors for influenza infection are largely subtype-independent. Specifically, age-specific risk could be influenced by medical factors like age-specific vaccine coverage, age-specific risk of severe disease, and immunosenescence, or by behavioral factors like age-assorted social mixing, and age-specific healthcare seeking behavior. All these factors should have similar impacts on any influenza subtype.

Thus, we fit a single step function to characterize the shape of age-specific risk of any confirmed influenza infection. Then, we modeled residual, subtype-specific differences in risk as a function of birth year, which enabled us to focus on the possible role childhood imprinting status. Each tested model used a linear combination of age-specific risk (Fig. 1C) and birth year-specific risk (Fig. 1D-F) to generate an expected distribution of H1N1 or H3N2 incidence (Fig. 1G-I). Note that for a given birth cohort, age-specific risk changed across progressive years of case observation, whereas birth year-specific risk was constant over time.

To test quantitatively whether observed, subtype-specific differences in age distribution were most consistent with imprinting at the HA subtype, NA subtype or HA group level, or with no contribution of imprinting, we fitted a suite of models to each data set using a multinomial likelihood and then performed model selection using the Akaike information criterion (AIC). AIC is used to compare the relative strength of statistical support for a set of candidate models, each fitted to the same data, and favors parsimonious models that fit the data well. Akaike weights are derived from AIC can be interpreted as the fraction of statistical support allocated to any given model, out of all models tested (27,28). Technical details are provided in the Methods.

### ADHS Models

We fit a set of four models to the ADHS data set. The simplest model contained only age-specific risk (A), and more complex models added effects from imprinting at the HA subtype level (S), at the HA group level (G), or at the NA subtype level (N): AS, AG, and AN, respectively. The age-specific risk curve took the form of a step function, in which relative risk was fixed to 1 in age bin 0-4, and one free parameter was fit to relative risk in each of the following 12 age bins: {5-10, 11-17, 18-24, 25-31, 32-38, 39-45, 46-52, 53-59, 60-66, 67-73, 74-80, 81+}. Within models that contained imprinting effects, two additional free parameters estimated the relative risk of confirmed H1N1 or H3N2 infection given imprinting protection.

### INSIGHT Models

When fitting to the INSIGHT data, which contained additional medical details, the suite of tested models included three additional risk factors: vaccination prior to the current influenza season (V), antiviral treatment (T), and presence of underlying conditions (U). We expected vaccination to reduce the risk of confirmed infection with either subtype. We assumed the presence of any underlying condition might be associated with increased healthcare seeking behavior, and in turn, with greater probabilities of influenza testing and case ascertainment. Finally, although antiviral treatment is usually prescribed in response to a confirmed influenza infection, treatment may be obtained from personal stockpiles (29,30). We included risk factor T in case antiviral treatment prior to testing reduced viral loads and the probability of case detection. Factors T and U each added one free parameter, which characterized the relative risk of any influenza infection. Factor V added two free parameters which characterized the relative risk of H1N1 or of H3N2 infection, given recent influenza vaccination. Differences in vaccine effectiveness across study years would not have been identifiable, and so we did not include them in the model. We tested all possible combinations of V, T, and U, in and each of the three imprinting hypotheses (S, N and G), for a total of 32 tested INSIGHT models. All tested models contained the age-specific risk step function (A), and used the following 10 age bins: {18-24, 25-31, 32-38, 39-45, 46-52, 53-59, 60-66, 67-73, 74-80, 81-90}.

### Interpretation of the age-specific risk curve

When fitting to INSIGHT data, we input the age distribution of all tested cases as the null, expected age distribution in each country and season. Thus, age-specific risk curves fitted to INSIGHT data only represented age-specific differences in the rate of testing positive for influenza, residual to observed denominators. Corresponding denominator data were not available in the ADHS dataset. The ADHS age-specific risk curves must be interpreted differently, as they captured all aspects of the infection and case observation process, and ultimately showed stronger age-specific differences in risk than INSIGHT risk curves (***Fig. 2-3***).

**Effect of evolutionary rate on age profiles**

We used publicly available data from *Nextstrain* (31,32), and from one previously published study (33), to calculate annual antigenic advance, which we defined as the antigenic distance between strains of a given lineage (pre-2009 H1N1, post-2009 H1N1 or H3N2) that circulated in consecutive seasons (Methods). The “antigenic distance” between two strains is used as a proxy for immune cross-protection and can be obtained using a variety of methods that map serological or genetic data into Euclidian space (33,34), or onto phylogenetic trees (32).

To assess the impact of antigenic evolutionary rate on the epidemic age distribution, we tested whether the proportion of cases in children increased in seasons associated with large antigenic changes. If the rate of antigenic drift is a strong driver of age-specific influenza risk, then the fraction of influenza cases observed in children should be negatively related to annual antigenic advance (24). In other words, strains that have not changed much antigenically since the previous season should be unable to escape pre-existing immunity in immunologically experienced adults, and more restricted to causing cases in immunologically naïve children; strains that have changed substantially (i.e. those exhibiting greater antigenic advance) will be less restricted to children.

# Results

### Subtype-specific differences in age distribution

In both ADHS and INSIGHT data, H3N2 consistently caused more cases in older cohorts, while H1N1 caused more cases in younger cohorts (Figs. 2-3, S2-S7). These patterns are qualitatively consistent with the predicted effects of cohort-specific imprinting (Fig 1), and with previously reported differences in age distribution (11–13,23). Overall, differences between H1N1 and H3N2’s age distributions were more pronounced in the Arizona data than in the INSIGHT data, but whenever subtype-specific differences in age distribution were apparent, H3N2 always caused greater impacts in the oldest cohorts, while H1N1 caused greater impacts in young and middle-aged adults (see Figs. 2-3 for seasons where H1N1 and H3N2 co-circulated in substantial numbers, and Figures S2-S7 for the entire dataset and alternate smoothing parameters).

## Imprinting model selection

Whether we fit to INSIGHT or to ADHS data, models containing NA subtype-level imprinting received the most statistical support, and models containing HA subtype-level imprinting were the second most preferred in terms of AIC (Fig. 4, Table 3). The ADHS data showed a strong preference for NA subtype-level imprinting over HA subtype-level imprinting (ΔAIC=23.42), and effectively no statistical support for broad, HA group-level imprinting (ΔAIC=245.18), or for an absence of imprinting effects (ΔAIC=380.47). Visual assessment of model fits (Fig. 4F,G) confirmed that models containing imprinting effects at the narrow, NA or HA subtype levels provided the best fits to data. The lack of support for the no-imprinting model shows imprinting from the first exposure shapes lifelong seasonal influenza risk, just as it does avian-origin influenza (10, 12).

As expected (see Fig. 1G-I), predictions from the two best models were highly collinear, except in their risk predictions among middle-aged, H2N2-imprinted cohorts (birth years 1957-1968), and some other minor differences arising from normalization across birth-years.

Model selection on ADHS data showed much stronger effects of all tested factors, and clearer preference for particular models, than model selection on INSIGHT data (Fig. 4, Table 3). No single model fit to INSIGHT data was definitively preferred (six had ΔAIC<4, and differences between fits were negligible (Fig. 4, Table 3)), but results of model selection aligned qualitatively between INSIGHT and ADHS data. None of the six best models for INSIGHT contained HA group-level imprinting effects (Table 3). Overall, NA subtype-level imprinting or HA subtype-level imprinting received the most statistical support, in terms of AIC and Akaike weight (***Fig. 4E***).

## Fitted risk patterns

When fitted to ADHS data, age-specific risk curves took similar forms in all models, with risk decreasing rapidly from birth through adolescence, and then decreasing much more slowly until the end of life (Fig. 4A). When fitted to INSIGHT data, age-specific risk effects always took values close to 1, indicating that the age distribution of confirmed influenza cases was roughly proportional to denominator data on the age distribution of all tested cases (Fig. 4C, Fig. S1).

Tables S1-S3 show parameter estimates and 95% profile confidence intervals from all models fitted to ADHS And INSIGHT data. Fits to ADHS data estimated moderate reductions in risk due to imprinting protection, and fits to INSIGHT data estimated weak reductions in risk (Fig. 4B,D, Table S2-S3). Overall, risk parameters fitted to the INSIGHT data took values closer to 1 and had wider confidence intervals than risk parameters fitted to the ADHS data (Fig. 4, Tables S1-S3). As fitted to INSIGHT data, the estimated relative risk of infection given antiviral treatment was usually greater than one, which may reflect that antiviral treatment is often prescribed in response to a positive influenza test. Vaccination was consistently associated with small reductions in risk, although confidence intervals often overlapped the null value of one. The presence of underlying conditions did not strongly impact relative risk (Table 3, Tables S1-S3).

## Effect of evolutionary rate

Next, we searched for putative increases in the proportion of cases among children in seasons associated with antigenic novelty. The ADHS data showed a slight negative but not significant association between annual antigenic advance and the fraction of H3N2 cases observed in children (Fig. 5A). The data contained too few influenza seasons with >100 confirmed H1N1 cases to support meaningful Pearson correlation coefficients specific to pre-2009 or post-2009 H1N1 lineages.

If evolutionary rate is the dominant driver of epidemic age distribution, then outbreaks caused by H1N1 and H3N2 should converge in age distribution when annual antigenic advance is similar. However, the data showed that differences in H1N1 and H3N2’s age-specific impacts persisted, even when lineages showed similar annual advance (Fig. 5A). When comparing the fraction of cases observed in specific age classes, H1N1 data consistently clustered separately from H3N2, with H1N1 consistently causing fewer cases in children (0-10), and elderly adults (71-85), and more cases in adults than H3N2 strains with similar rates of antigenic advance. Smoothed density plots showed no clear relationship between annual antigenic advance and age distribution (Fig. 5B). Overall, the data did not show a strong signal that epidemic age distribution varies with the magnitude of annual antigenic drift.

# Discussion

Our analyses of two large datasets of influenza cases confirmed a difference in age-specific impacts of seasonal H1 and H3, which was consistent across multiple countries and seasons. We analyzed several possible drivers of these differences, and found greatest support for the hypothesis that immunological imprinting leads to lasting protection against the NA or HA subtype of the first influenza strain encountered in childhood (11,12). The data did not support strong effects from broader HA group-level imprinting, as recently detected for novel zoonotic or pandemic viruses (8,10), or from differences in rates of antigenic evolution (24). Our results show that the first childhood infection leaves a lifelong imprint of susceptibility to seasonal influenza, and that this imprint is not erased even after decades of exposure to or vaccination against seasonal influenza subtypes mismatched to the first infecting strain.

As additional evidence that birth year, rather than age, drives subtype-specific differences in seasonal influenza risk, we note that H3N2’s impacts have not always been focused in elderly cohorts. When H3N2 first emerged in 1968, it caused little or no excess mortality in the elderly, putatively because those who were elderly in 1968 had been exposed, as children or young adults, to an H3 virus that had circulated in the late 1800s (6,8). Meanwhile, H1N1-imprinted cohorts (those ~10-50 years old at the time of the H3N2 pandemic), experienced considerable excess mortality at the time of H3N2’s emergence (6), and continue to experience excess H3N2 morbidity and mortality today as elderly adults ((11–13,23), ***Fig. 2-3***). In short, across decades of circulation in humans, H3N2’s impacts have remained consistent with respect to birth year, but have shifted with respect to age.

Model comparison on both data sets independently provided the strongest support for effects from childhood imprinting to NA. Although NA is not as intensively studied as HA, these results emphasize the increasingly recognized importance of both antigens as drivers of protection against seasonal influenza (14–16). The models containing NA and HA subtype-level imprinting produced very similar, fits to data, and independently emerged as the top two models in terms of AIC. Realistically, some combination of effects from both HA and NA subtype-level imprinting probably shape seasonal influenza risk. Unfortunately, collinearities between predictions of the simple, single-antigen models considered here arose inevitably from influenza’s limited history of circulation in humans. Thus, we did not test more complicated models of combined effects from imprinting to HA and NA, or to other antigens such as internal proteins. Because analysis of population-level data can only support a limited scope of inference, deeper insights into the respective roles of HA, NA and other influenza antigens as drivers of cohort effects will most likely need to come from focused immunological cohort studies in which individual histories of influenza infection are known, such as those recently funded by the National Institutes of Health (35). Alternatively, the development of immunological biomarkers for diagnosis of imprinting status in individual patients could substantially increase the power of epidemiological inference, which (as in this study) currently relies instead on probabilistic reconstructions of imprinting histories according to birth year (10).

Our failure to detect a strong signal of impact from evolutionary rate on age distributions of H1N1 and H3N2 cases was surprising, but is consistent with growing recognition that existing methods to map HA antigenic distance between strains are based on hemagglutination inhibition (HI) data collected from ferrets, and do not always capture realized patterns of cross-reactivity in the human population. First, the evolutionary patterns considered in our analysis were purely based on HA, while our epidemiologic data supports an equal if not stronger role of NA. More work is required to map NA antigenic changes and understand the possible interactions between HA and NA evolution. Further, glycosylation of HA can cause antigenic escape in large subsets of the human population, which may be perceived as neutral in existing antigenic maps (36,37). One epidemiologically impactful example of incognito (unmapped) antigenic escape was observed during the 2013-2014 H1N1 epidemic (36). We speculate that a clearer relationship between epidemic age distribution and antigenic drift would emerge if antigenic distance measures were modified to incorporate cohort-specific variation in immune history, and impacts from multiple antigens.

Our finding that within-subtype imprinting has much stronger impacts than broader, HA group-level imprinting is consistent with the clear impact of narrow, within-subtype immunity on seasonal influenza’s evolutionary dynamics (24,38). Still, given that within-subtype immunity decays rapidly in the face of antigenic drift, it is striking that signatures of narrow, within-subtype imprinting protection persist across an entire human lifetime, and remain evident, even in the oldest cohorts. On average, H1N1 and H3N2 viruses drift by 0.62 and 1.01 antigenic units per year, respectively (33), which roughly corresponds to a two-fold drop in hemagglutination inhibition titer for every 1.61, or 0.99 years of antigenic evolution between strains. Strains that circulated more than 14 years apart do not usually show measurable cross-protective hemagglutination inhibition titers (33). In this context, it is not obvious that narrow, within-subtype influenza immunity primed in childhood should provide any meaningful protection after adolescence, let alone decades later in old age. However, we note that the serological assays used to map antigenic cross-reactivity measure serum antibodies are imperfect correlates of *in vivo* protection, and do not capture effects from cellular immunity (especially from CD4+ T cells), or from other mechanisms of protection (19,26,39,40).

Another potential explanation for the longevity of within-subtype childhood imprinting protection is that imprinting to a particular HA or NA subtype builds strong memory of epitopes conserved among variants of the same subtype, but not across subtypes. Immunological studies show B cell memory shifts over time to focus on conserved influenza epitopes, as a lifetime of exposures to drifted H1N1 or H3N2 variants repeatedly back-boosts memory of unchanged epitopes (41,42). Another explanation supported by recent immunological data (43), is that the memory B cell clones developed during the first childhood influenza exposure later adapt via somatic hypermutation to follow homosubtypic antigenic targets as they drift over time. Thus, childhood imprinting may provide preferential, lifelong protection against a particular HA or NA subtype by filling a child's memory B cell repertoire with clones that will serve in the future, not as final products but as prototypes that can be rapidly and effectively tailored to recognize drifted influenza strains of the same subtype.

A third possibility is that signals of imprinting protection are anomalously strong in the current cohort of elderly adults. For nearly four decades, from 1918-1957, H1N1 persisted as the only strain circulating in humans. The oldest subjects in our data were born slightly after its emergence in 1918, and would not have encountered an influenza virus of any subtype but H1N1 until after age 30. Decades of early-life exposures to H1N1 variants may have reinforced and expanded the breadth of H1N1-specific immune memory in these oldest cohorts. But this strong protection against H1N1 seems to come at a cost; even after decades of seasonal H3N2 exposure, and vaccination, older cohorts have evidently failed to develop equally strong protection against H3N2. Antigenic similarity between H1N1 strains that circulated earlier in the 20th century (which caused imprinting in older cohorts), and modern H1N1 lineages that emerged in 1977 and in 2009, may also have amplified the strength and longevity of H1N1 protection in these cohorts (4,44). One additional consideration in this context is that HA group 1 antigens appear to induce narrower immune responses than structurally distinct HA group 2 antigens, which may be better able to induce cross-group responses (21). Perhaps elderly cohorts imprinted to group 1 antigens have been trapped in narrower responses that offer exceptional protection against strains similar to that of first exposure but relatively poor adaptability to other subtypes.

Given that cohorts born before 1977 have had much more varied early life exposures to both H1N1 and H3N2, it is unclear whether equally strong, subtype-specific biases in protection will persist when pre-1977 birth cohorts eventually become elderly. Determining whether strong, lifelong biases in immune protection are a general feature of imprinting, or specific to the current cohorts of elderly adults, is necessary to project long-term shifts in influenza-related mortality. The vast majority of influenza-related deaths occur in adults over age 65, and H3N2 currently causes many times the number of fatalities as H1N1 (23,45). These patterns may reflect intrinsic differences in virulence, but we speculate that imprinting protection, which currently limits the incidence of clinically-attended H1N1 infection in the elderly, may also dramatically reduce H1N1-related mortality at present. In the future, cohorts imprinted to H2N2 (born c. 1950-1968) will become elderly, and their imprinting protection will act instead against seasonal H3N2 (via NA subtype-level imprinting), or against no seasonal strains (via HA subtype-level imprinting). If future elderly cohorts continue to show strong subtype-specific biases from imprinting, our results would corroborate the idea that mortality from H1N1 may increase in the future (8,12) as protection in the elderly shifts toward other subtypes. On the other hand, future generations of elderly adults may show a greater ability to act as immunological generalists, with effective defenses against multiple influenza subtypes.

Overall, differences between age distributions of infection caused by H1N1 or H3N2 were much more pronounced in the ADHS data than in the INSIGHT data. Differences between the datasets may arise due to geographic variation in influenza’s epidemiology; the INSIGHT data were collected across five continents, whereas all the ADHS data came from a single US state. Climatic or demographic characteristics, or high rates of influenza vaccination (46,47), may magnify subtype-specific differences in age distribution within the United States. It is noteworthy that within the INSIGHT data, the subset of cases observed in the United States showed more dramatic differences in age distribution than data collected in many other countries (Fig. 3). Similarly, apparent differences between H1N1 and H3N2’s age distributions were greater in the United States than in Europe in one previous study (13).

Differences in age-specific sampling probably also contributed to differences between the datasets. The INSIGHT study did not enroll children, and did not capture severe, hospitalized cases, which occur disproportionately in older adults. A dearth of enrolled cases in the youngest and oldest age groups may have dampened the signal of subtype-specific differences in risk in the INSIGHT data (Fig. S1). To illustrate the effects of age-specific sampling, we subsampled the ADHS data to match the sample size and age distribution of all confirmed influenza A cases from the INSIGHT study. Filtering the ADHS data in this way made differences in age-specific risk from H1N1 and H3N2 appear much smaller (Fig. S1D).

The potential for age-specific sampling biases to erode or magnify the signal of imprinting effects highlights some limitations of existing epidemiological surveillance data, which in turn limited this study’s scope of inference. The largest, long-term epidemiological data sets on influenza come from massive, global surveillance efforts. But due to practical and economic constraints, these data are often collected opportunistically, with uneven sampling efforts over time and across age groups, and rarely shared denominator data. Furthermore, while some aspects of surveillance data are already shared publicly by WHO (48), and by the US CDC (49), data on patient ages is obscured by aggregation into broad age categories. Focused clinical studies like INSIGHT can yield more extensive information on individual medical histories and on sampling denominators, but such clinical data sets typically contain far fewer cases than state or country-wide surveillance data, and may not sample all age groups. Arguably, only epidemiological surveillance provides enough confirmed cases to characterize epidemic age distributions with precision across multiple countries and influenza seasons. We emphatically echo earlier calls (50) for more systematic sharing of single year-of-age in influenza surveillance data, standardization of sampling effort, and reporting of age-specific denominators, which could substantially boost the scientific community’s ability to link influenza's genetic and antigenic properties with epidemiological outcomes.

Altogether, this analysis confirms that observed differences in the birth year-specific impacts of H1N1 and H3N2 are indeed shaped by cohort-specific differences in childhood imprinting (8,11,12,51). The finding that such imprinting acts at the narrow, HA or NA subtype-level against seasonal influenza facilitates prediction of the future epidemiological impact of specific seasonal subtypes in high-risk elderly cohorts. The lack of support for broader, HA group-level imprinting effects highlights the difficulty of deploying broadly protective memory B cell responses against familiar, seasonal strains. Overall, these findings further our understanding of how serological antigenic seniority translates to functional immune protection, and shapes cohort-specific risk during epidemics. The fact that elderly cohorts show relatively weak immune protection against H3N2, even after living through decades of seasonal exposure to or vaccination against H3N2, suggests that antibody responses acquired in adulthood do not provide the same strength of immune protection as responses primed in childhood. These findings suggest that antibody responses acquired in adulthood from back-boosting of existing memory B cells do not provide the same strength of protection as true, *de novo* responses. It remains to be seen whether the strong, lifelong immunological biases observed in elderly cohorts within this study are an intrinsic feature of childhood imprinting, or whether these biases have become unusually entrenched in the current cohort of elderly adults, due to the prolonged dominance of H1N1 circulation across the first half of the 20th century.

# Methods

## Data inclusion criteria

Following CDC standards, ADHS defines the influenza season as epidemiological week 40 (around early October) through week 39 of the following year (53). To facilitate comparison between data sets, we defined INSIGHT cases enrolled from Oct-May as part of the Northern Hemisphere influenza season, and cases enrolled from June-Sept as part of the Southern Hemisphere season.

In the ADHS data, the 2008-2009 and 2009-2010 influenza seasons were excluded from our analysis because these seasons spanned the first and second wave, respectively, of the 2009 H1N1 pandemic. We excluded INSIGHT data from the 2009-2010 Northern Hemisphere season, and the 2010 Southern hemisphere influenza season, which spanned the second pandemic wave in each hemisphere.

Additionally, from the ADHS data, we excluded 58 cases with birth years before 1918 (whose imprinting status could not be inferred unambiguously), and one case whose year of birth was recorded in error. From the INSIGHT data, we excluded 94 cases with missing information in any of five key categories: age, date of enrollment, vaccination, underlying conditions and antiviral treatment. We excluded an additional 7 cases that fell outside the focal age range of 18-90, and 3 cases whose infecting subtype could not be unambiguously identified due to coinfection.

## Estimation of age from birth year in ADHS data

The ADHS data contained three variables, influenza season, birth year and confirmed subtype. For most cases, birth year was extracted directly from the reported date of birth in patient medical records, but age was not known. To fit age-specific risk curves to the ADHS data, we estimated patient age at the time case observation using the formula [year of observation]-[birth year]. To ensure that the minimum estimated age was 0, the second year in the influenza season of case observation was considered the calendar year of observation (e.g. 2013 for the 2012-2013 season).

## Estimation of birth year from age in the INSIGHT data

The INSIGHT data contained patient age, and date of case enrollment, but not birth year. We estimated birth year using a method that took advantage of precise dates of case enrollment. The simplest calculation of birth year is (observation year)-(age), but this approximation is slightly biased, as cases observed earlier in the year (e.g. in January) are less likely to have passed a birthday in the current calendar year. As shown in Fig. S8, we determined the possible birth dates of an individual observed during each month of the year. We then determined that on average, cases observed during months in the Northern Hemisphere season (Oct.-May) had the following probabilities of birth in each of three possible years, relative to the year of case observation:

**1**

Cases observed during months of the Southern Hemisphere had different probabilities:

**2**

Using these probabilities, we took a weighted average of birth year-specific imprinting probabilities for cases observed in the Northern Hemisphere or Southern Hemisphere influenza seasons.

## Splines

In Figures 2-3, smoothing splines were fit to aid visual interpretation of noisy data. We fit splines using the command *smooth.spline(x = AGE, y = FRACTIONS, spar = 0.8)* in R version 3.5.0. Variables *AGE* and *FRACTIONS* were vectors whose entries represented single years of age, and the fraction of cases observed in the corresponding age group. The smoothing parameter 0.8 was chosen to provide a visually smooth fit. Alternative smoothing parameter choices (0.6 & 1.0) are shown in Figs. S3, S6-S7. Although the choice of smoothing parameter changed the shape of each fitted spline, qualitative differences between splines fitted to H1N1 or H3N2 were robust.

## Model formulation

For each unique country and season in which cases were observed, define *p* as a vector whose entriesrepresent the expected probability that a randomly drawn H1N1 or a randomly drawn H3N2 case was observed in an individual of age *a*. Each model defined, *p* as a linear combination of age-specific risk, birth year-specific risk (i.e. imprinting effects), and other medical history variables, and *p* took slightly different shapes for expected H1N1 and H3N2 case age distributions. All tested models were nested within the equation:

**3**

To include risk factors that only modulated risk from one subtype, we included indicator functions ***1H1N1*** and ***1H3N2***, which took value 1 if *p* described the expected age distribution of H1N1 or H3N2 cases, respectively, and 0 otherwise.

### Denominator data (D)

When fitting to INSIGHT data, *D* was a vector whose entries were proportional to the age distribution of all tested cases within a given country and year. As noted above, corresponding denominator data were not available in the ADHS dataset, and so factor D was not included in models fit to ADHS data.

### Age-specific risk (A)

Age-specific risk was defined as a step function, in which relative risk was fixed to value 1 in an arbitrarily chosen age bin, and then *z*-1 free parameters, denoted *r*2 to *rz*, were fit to describe relative risk in all other age bins. Below, ***1i*** are indicator functions specifying whether each vector entry is a member of age bin *i.* To obtain the predicted fraction of cases observed in each single year of age, we normalized the risk distribution so that predicted risk across all age groups summed to 1.

**4**

### Antiviral treatment (T) and underlying conditions (U)

Within each country and season, *fT* defined a vector whose entries describe the fraction of tested cases of a given age that had received antiviral treatment. Free parameter *rT* defines the relative risk of any confirmed influenza infection, given antiviral treatment. Then, risk factor *T* was defined as:

**5**

And risk factor *U* was described similarly:

**6**

### Imprinting (I) and vaccination (V)

Factors describing risk from vaccination and imprinting took forms similar to risk factors *T* and *U*, but with subtype-specific impacts. An indicator function defined whether a given prediction vector described risk of confirmed H1N1 or H3N2 infection. Let *fV* and *fIHxNy* be vectors describing the fraction of cases of each age that were recently vaccinated against influenza, or that were protected against strain *HxNy* by their childhood imprinting. Note that we used the general definition “recent influenza vaccination” because some INSIGHT study sites recorded whether patients had been vaccinated in the last 6 months, while other recorded vaccination in the last 12 months. We defined *rvHxNy* and *rIHxNy* as free parameters describing the risk of confirmed *HxNy* infection, given vaccination, or given imprinting protection. Finally, risk factors describing the effect of vaccination (V) and imprinting (I) were defined as:

**7**

**8**

## Likelihood

We used equations 3-8 to generate predicted case age distributions (*p*) for each influenza season (s) and country (c) in which cases were observed in the data. Then, the likelihood was obtained as a product of multinomial densities across all countries and seasons observed in the data. If *ncs* represents the total number of cases observed in a given country and season, *x0cs,…xmcs*each represent the number of cases observed in each single year of age/single year of birth, and if *p0cs…pmcs* each represent entries in the model’s predicted age/birth year-distribution of cases, then the likelihood is given by:

## Model fitting and model comparison

We fit models containing all possible combinations of the above factors to influenza data from each unique country and season in the ADHS data, or in the INSIGHT data. We simultaneously estimated all free parameter values using the optim() function in R. We calculated likelihood profiles and 95% profile confidence intervals for each free parameter. Confidence intervals were defined using the method of likelihood ratios (27).

## Antigenic advance

We obtained antigenic distance estimates from *Nextstrain (nextstrain.org)*(31,54)*,* and from source data associated with Figure 3 in Bedford et al. (33). *Nextstrain* calculates antigenic distance using genetic data from GISAID (55), and using methods described by Neher et al. (32). We analyzed “CTiter” estimates from *Nextstrain*, which correspond to Neher et al.’s “tree model” method. Datasets from *Nextstrain* and Bedford et al. both contained redundant antigenic distance estimates for the H3N2 lineage, but only Bedford et al. analyzed the pre-2009 H1N1 lineage, and only *Nextstrain* data analyzed the post-2009 H1N1 lineage. The antigenic distance estimates reported by Bedford et al. were roughly proportional to those reported on *Nextstrain*, but greater in absolute magnitude (32). To enable direct comparison among all three lineages on the same plot axes, we rescaled pre-2009 H1N1 estimates from Bedford et al. using the formula *dNextstrain* = 0.47*dBedford*. The scaling factor was chosen so that directly-comparable H3N2 distance estimates obtained using each method were well-aligned (Fig. S9). The *Nextstrain* data files used in this analysis are archived within our analysis code, as described below.

## Code and data availability

Code to perform all reported analyses and construct all plots is available \#\#HERE\#\#. ADHS data is available as a supplementary data file. Due to patient privacy considerations, data from the INSIGHT study are available by application, pending approval from the study's scientific review committee (<http://insight.ccbr.umn.edu/index.php>). Because we are not free to share the INSIGHT data, the posted code contains a dummy INSIGHT data file with scrambled column entries. In other words, the data files are formatted properly, and the code will run, but the actual data entries have no biological meaning.

# References

1. Francis T. On the Doctrine of Original Antigenic Sin. Proc Am Philos Soc. 1960;104(6):572–8.

2. Lessler J, Riley S, Read JM, Wang S, Zhu H, Smith GJD, et al. Evidence for Antigenic Seniority in Influenza A (H3N2) Antibody Responses in Southern China. PLOS Pathog. 2012 Jul 19;8(7):e1002802.

3. Xu R, Ekiert DC, Krause JC, Hai R, Crowe JE, Wilson IA. Structural Basis of Preexisting Immunity to the 2009 H1N1 Pandemic Influenza Virus. Science. 2010 Apr 16;328(5976):357–60.

4. Hancock K, Veguilla V, Lu X, Zhong W, Butler EN, Sun H, et al. Cross-Reactive Antibody Responses to the 2009 Pandemic H1N1 Influenza Virus. N Engl J Med Boston. 2009 Nov 12;361(20):1945–52.

5. Simonsen L, Spreeuwenberg P, Lustig R, Taylor RJ, Fleming DM, Kroneman M, et al. Global Mortality Estimates for the 2009 Influenza Pandemic from the GLaMOR Project: A Modeling Study. PLOS Med. 2013 Nov 26;10(11):e1001558.

6. Simonsen L, Reichert TA, Miller MA. The virtues of antigenic sin: consequences of pandemic recycling on influenza-associated mortality. Int Congr Ser. 2004 Jun 1;1263:791–4.

7. Ma J, Dushoff J, Earn DJD. Age-specific mortality risk from pandemic influenza. J Theor Biol. 2011 Nov 7;288:29–34.

8. Worobey M, Han G-Z, Rambaut A. Genesis and pathogenesis of the 1918 pandemic H1N1 influenza A virus. Proc Natl Acad Sci. 2014 Jun 3;111(22):8107–12.

9. Gagnon A, Miller MS, Hallman SA, Bourbeau R, Herring DA, Earn DJD, et al. Age-Specific Mortality During the 1918 Influenza Pandemic: Unravelling the Mystery of High Young Adult Mortality. PLoS ONE [Internet]. 2013 Aug 5 [cited 2019 Apr 4];8(8). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3734171/

10. Gostic KM, Ambrose M, Worobey M, Lloyd-Smith JO. Potent protection against H5N1 and H7N9 influenza via childhood hemagglutinin imprinting. Science. 2016 Nov 11;354(6313):722–6.

11. Khiabanian H, Farrell GM, George KS, Rabadan R. Differences in Patient Age Distribution between Influenza A Subtypes. PLOS ONE. 2009 Aug 31;4(8):e6832.

12. Budd AP, Beacham L, Smith CB, Garten RJ, Reed C, Kniss K, et al. Birth Cohort Effects in Influenza Surveillance Data: Evidence that First Influenza Infection Affects Later Influenza-Associated Illness. J Infect Dis [Internet]. [cited 2019 May 20]; Available from: https://academic.oup.com/jid/advance-article/doi/10.1093/infdis/jiz201/5485579

13. Lemaitre M, Carrat F. Comparative age distribution of influenza morbidity and mortality during seasonal influenza epidemics and the 2009 H1N1 pandemic. BMC Infect Dis. 2010 Jun 9;10(1):162.

14. Huang QS, Bandaranayake D, Wood T, Newbern EC, Seeds R, Ralston J, et al. Risk Factors and Attack Rates of Seasonal Influenza Infection: Results of the Southern Hemisphere Influenza and Vaccine Effectiveness Research and Surveillance (SHIVERS) Seroepidemiologic Cohort Study. J Infect Dis. 2019 Jan 9;219(3):347–57.

15. Cowling BJ, Sullivan SG. The Value of Neuraminidase Inhibition Antibody Titers in Influenza Seroepidemiology. J Infect Dis. 2019 Jan 9;219(3):341–3.

16. Memoli MJ, Shaw PA, Han A, Czajkowski L, Reed S, Athota R, et al. Evaluation of Antihemagglutinin and Antineuraminidase Antibodies as Correlates of Protection in an Influenza A/H1N1 Virus Healthy Human Challenge Model. mBio [Internet]. 2016 Apr 19 [cited 2019 May 31];7(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4959521/

17. Wrammert J, Koutsonanos D, Li G-M, Edupuganti S, Sui J, Morrissey M, et al. Broadly cross-reactive antibodies dominate the human B cell response against 2009 pandemic H1N1 influenza virus infection. J Exp Med. 2011 Jan 17;208(1):181–93.

18. Pica N, Hai R, Krammer F, Wang TT, Maamary J, Eggink D, et al. Hemagglutinin stalk antibodies elicited by the 2009 pandemic influenza virus as a mechanism for the extinction of seasonal H1N1 viruses. Proc Natl Acad Sci U S A. 2012;109(7):2573–8.

19. Krammer F. Novel universal influenza virus vaccine approaches. Curr Opin Virol. 2016 Apr;17:95–103.

20. Andrews SF, Huang Y, Kaur K, Popova LI, Ho IY, Pauli NT, et al. Immune history profoundly affects broadly protective B cell responses to influenza. Sci Transl Med. 2015 Dec 2;7(316):316ra192-316ra192.

21. Zost SJ, Wu NC, Hensley SE, Wilson IA. Immunodominance and Antigenic Variation of Influenza Virus Hemagglutinin: Implications for Design of Universal Vaccine Immunogens. J Infect Dis. 2019 Apr 8;219(Supplement\_1):S38–45.

22. Miller MS, Gardner TJ, Krammer F, Aguado LC, Tortorella D, Basler CF, et al. Neutralizing Antibodies Against Previously Encountered Influenza Virus Strains Increase over Time: A Longitudinal Analysis. Sci Transl Med. 2013 Aug 14;5(198):198ra107-198ra107.

23. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, et al. Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA. 2003 Jan 8;289(2):179–86.

24. Bedford T, Riley S, Barr IG, Broor S, Chadha M, Cox NJ, et al. Global circulation patterns of seasonal influenza viruses vary with antigenic drift. Nature. 2015 Jul;523(7559):217–20.

25. Arizona Department of Health Services. 2015–2016 Influenza Summary [Internet]. [cited 2019 May 23]. Available from: https://www.azdhs.gov/documents/preparedness/epidemiology-disease-control/flu/surveillance/2015-2016-influenza-summary.pdf

26. Erbelding EJ, Post DJ, Stemmy EJ, Roberts PC, Augustine AD, Ferguson S, et al. A Universal Influenza Vaccine: The Strategic Plan for the National Institute of Allergy and Infectious Diseases. J Infect Dis. 2018 Jul 2;218(3):347–54.

27. Bolker BM. Ecological Models and Data in R. Princeton University Press; 2008. 409 p.

28. Burnham KP, Anderson DR. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach [Internet]. 2nd ed. New York: Springer-Verlag; 2002 [cited 2019 Apr 16]. Available from: https://www.springer.com/us/book/9780387953649

29. Kramarz P, Monnet D, Nicoll A, Yilmaz C, Ciancio B. Use of oseltamivir in 12 European countries between 2002 and 2007 – lack of association with the appearance of oseltamivir-resistant influenza A(H1N1) viruses. Eurosurveillance. 2009 Feb 5;14(5):19112.

30. Gauld NJ, Jennings LC, Frampton C, Huang QS. Five years of non-prescription oseltamivir: effects on resistance, immunization and stockpiling. J Antimicrob Chemother. 2012 Dec 1;67(12):2949–56.

31. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics. 2018 Dec 1;34(23):4121–3.

32. Neher RA, Bedford T, Daniels RS, Russell CA, Shraiman BI. Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. Proc Natl Acad Sci. 2016 Mar 22;113(12):E1701–9.

33. Bedford T, Suchard MA, Lemey P, Dudas G, Gregory V, Hay AJ, et al. Integrating influenza antigenic dynamics with molecular evolution. Losick R, editor. eLife. 2014 Feb 4;3:e01914.

34. Smith DJ, Lapedes AS, Jong JC de, Bestebroer TM, Rimmelzwaan GF, Osterhaus ADME, et al. Mapping the Antigenic and Genetic Evolution of Influenza Virus. Science. 2004 Jul 16;305(5682):371–6.

35. RFA-AI-18-010: Impact of Initial Influenza Exposure on Immunity in Infants (U01 Clinical Trial Not Allowed) [Internet]. [cited 2019 Apr 15]. Available from: https://grants.nih.gov/grants/guide/rfa-files/RFA-AI-18-010.html

36. Linderman SL, Chambers BS, Zost SJ, Parkhouse K, Li Y, Herrmann C, et al. Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013–2014 influenza season. Proc Natl Acad Sci. 2014 Nov 4;111(44):15798–803.

37. Cobey S, Hensley SE. Immune history and influenza virus susceptibility. Curr Opin Virol. 2017 Feb 1;22:105–11.

38. Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, et al. Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. Science. 2004 Jan 16;303(5656):327–32.

39. DiLillo DJ, Palese P, Wilson PC, Ravetch JV. Broadly neutralizing anti-influenza antibodies require Fc receptor engagement for in vivo protection. J Clin Invest. 2016 Feb 1;126(2):605–10.

40. Henry Dunand CJ, Leon PE, Huang M, Choi A, Chromikova V, Ho IY, et al. Both Neutralizing and Non-Neutralizing Human H7N9 Influenza Vaccine-Induced Monoclonal Antibodies Confer Protection. Cell Host Microbe. 2016 Jun 8;19(6):800–13.

41. Henry C, Zheng N-Y, Huang M, Cabanov A, Rojas KT, Kaur K, et al. Influenza Virus Vaccination Elicits Poorly Adapted B Cell Responses in Elderly Individuals. Cell Host Microbe. 2019 Mar;25(3):357-366.e6.

42. Age-specific differences in the dynamics of protective immunity to influenza | Nature Communications [Internet]. [cited 2019 May 6]. Available from: https://www.nature.com/articles/s41467-019-09652-6

43. Tesini BL, Kanagaiah P, Wang J, Hahn M, Halliley JL, Chaves FA, et al. Broad Hemagglutinin-Specific Memory B Cell Expansion by Seasonal Influenza Virus Infection Reflects Early-Life Imprinting and Adaptation to the Infecting Virus. J Virol. 2019 Apr 15;93(8):e00169-19.

44. Rozo M, Gronvall GK. The Reemergent 1977 H1N1 Strain and the Gain-of-Function Debate. mBio. 2015 Sep 1;6(4):e01013-15.

45. Dushoff J, Plotkin JB, Viboud C, Earn DJD, Simonsen L. Mortality due to Influenza in the United States—An Annualized Regression Approach Using Multiple-Cause Mortality Data. Am J Epidemiol. 2006 Jan 15;163(2):181–7.

46. Palache A, Oriol-Mathieu V, Fino M, Xydia-Charmanta M. Seasonal influenza vaccine dose distribution in 195 countries (2004–2013): Little progress in estimated global vaccination coverage. Vaccine. 2015 Oct 13;33(42):5598–605.

47. Vanessen G, Palache A, Forleo E, Fedson D. Influenza vaccination in 2000: recommendations and vaccine use in 50 developed and rapidly developing countries. Vaccine. 2003 May 1;21(16):1780–5.

48. WHO | FluNet [Internet]. WHO. [cited 2019 Apr 15]. Available from: http://www.who.int/influenza/gisrs\_laboratory/flunet/en/

49. FluView Interactive | CDC [Internet]. 2018 [cited 2019 Apr 15]. Available from: https://www.cdc.gov/flu/weekly/fluviewinteractive.htm

50. Gagnon A, Acosta E, Miller MS. Reporting and evaluating influenza virus surveillance data: An argument for incidence by single year of age. Vaccine. 2018 Oct 8;36(42):6249–52.

51. Glezen WP, Keitel WA, Taber LH, Piedra PA, Clover RD, Couch RB. Age Distribution of Patients with Medically-Attended Illnesses Caused by Sequential Variants of Influenza A/H1N1: Comparison to Age-Specific Infection Rates, 1978–1989. Am J Epidemiol. 1991 Feb 1;133(3):296–304.

52. Monto AS, Malosh RE, Petrie JG, Martin ET. The Doctrine of Original Antigenic Sin: Separating Good From Evil. J Infect Dis. 2017 Jun 15;215(12):1782–8.

53. National Notifiable Diseases Surveillance System, Division of Health Informatics and Surveillance, National Center for Surveillance, Epidemiology and Laboratory Services. MMWR Week Fact Sheet [Internet]. [cited 2019 May 23]. Available from: https://wwwn.cdc.gov/nndss/document/MMWR\_Week\_overview.pdf

54. Sagulenko P, Puller V, Neher RA. TreeTime: Maximum-likelihood phylodynamic analysis. Virus Evol [Internet]. 2018 Jan 8 [cited 2019 Apr 12];4(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5758920/

55. Bogner P, Capua I, Lipman DJ, Cox NJ. A global initiative on sharing avian flu data. Nature. 2006 Aug;442(7106):981.

# Acknowledgements

We are grateful to Deborah Wentworth, Ken Komatsu and Kristen Herrick for their assistance with data access, and to Trevor Bedford for assistance accessing and interpreting antigenic distance data from *Nextstrain*. We thank Lone Simonsen, and the Cobey lab, especially Phil Arevalo for helpful discussions. KG was supported by the National Institutes of Health (F31AI134017, T32-GM008185). JLS was supported by NSF grants OCE-1335657 and DEB-1557022, SERDP RC-2635, and DARPA PREEMPT D18AC00031. The content of the information does not necessarily reflect the position or the policy of the U.S. government, and no official endorsement should be inferred.

# Competing interests

The authors declare no competing interests.

# Author contributions

MW, KG and JLS conceived of the questions and modeling analysis. CV and MW provided crucial input on data access and study design. JW, TC, LO, SB and RB supervised data collection. KG wrote the code, performed analyses and wrote the manuscript. All authors provided and input on analysis and interpretation of the results, and critical feedback on the manuscript text.