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

Katelyn M Gostic1, Rebecca Bridge2, Shane Brady2, Cécile Viboud3, Michael Worobey4, and James O Lloyd-Smith1,3\*

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

2Arizona Department of Health Services, Phoenix Arizona, USA

3Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

4Dept. of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA

\* [kgostic@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 a large epidemiological surveillance dataset to test whether childhood immune imprinting shapes seasonal influenza epidemiology, and if so, whether it acts primarily via immune memory of a particular influenza subtype 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 infection. Likelihood-based model comparison shows that narrow, within-subtype imprinting is the strongest driver of seasonal influenza risk. The data do 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, newer, and 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 the relatively low mortality burden of H1N1 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 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 by influenza viruses encountered in childhood.

Many epidemiological studies 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 the importance of broadly protective immunity remains unclear 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, relatively narrow cross-protective immunity, which only protects against closely related antigenic variants of the same hemagglutinin (HA) subtype, has been considered the norm. Lymphocyte memory of variable epitopes on the HA head, (sites at which hemagglutinin antigens of different subtypes show limited homology), drive this narrow, within-subtype protection, which is the main mechanism of protection for the inactivated influenza vaccine. But protection may also be driven by memory of other influenza antigens (e.g. neuraminidase, NA) (14–16), or by immune response to conserved epitopes, particularly on the HA stalk (10,17–19). Antibodies that target conserved HA epitopes can provide broad protection across multiple HA subtypes in the same phylogenetic group (17,19,20), where HA group 1 contains hemagglutinin subtypes H1 and H2, while group 2 contains H3 (10,18,21). H1, H2 and H3 are the only HA subtypes that have circulated seasonally in humans since 1918.

Within-subtype cross-protection is known to shape seasonal influenza’s epidemiology and evolution (22). But because this type of narrow immunity decays rapidly in the face of antigenic drift, it would not be expected to shape cohort-specific protection across an entire human lifetime (23,24). Conversely, broad, HA group-level immune memory can play a strong role in immune defense against drifted influenza strains (or pandemic viruses) whose variable HA epitopes are unfamiliar, and would be expected to leave an epidemiological imprint in the risk of influenza infection throughout life (17,19,20,25). Thus, childhood immune imprinting may determine which birth cohorts are primed for effective defense against seasonal strains with conserved HA epitopes characteristic of group 1 or group 2. A similar line of reasoning may apply to immunity against the NA, although much less attention has been paid to this antigen.

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 (8,11–13,26). 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 escape pre-existing immunity and infect older, immunologically experienced adults, whereas H1N1 may be relatively restricted to infecting immunologically naïve children (27).

We analyzed a large data set on seasonal influenza incidence to test 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 HA subtypes in the same phylogenetic group (***Fig. 1A-B***). We fitted a suite of models to data using maximum likelihood and compared models using AIC. In a separate analysis, 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

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. A smaller number of positive influenza tests were obtained through reporting by other clinical labs, which has been mandatory in Arizona since 2004 (28). 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). The data included positive test results from hospitals, long-term care facilities, correctional facilities, and outpatient clinics, and thus captured a range of case severities.

Following CDC standards, ADHS defines the influenza season as epidemiological week 40 (around early October) through week 39 of the following year (29). The 2008-2009 and 2009-2010 influenza seasons spanned the first and second wave, respectively, of the 2009 H1N1 pandemic. We excluded cases observed during this time period, because age distributions of infection and molecular drivers of immune memory differed during the 2009 pandemic from the normal, drivers of seasonal influenza’s immuno-epidemiology of interest to this study (13,17,20). Additionally, 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.

# The Model

## Reconstructed imprinting patterns

Reconstructed, birth year-specific probabilities of childhood imprinting to H1N1, H2N2 or H3N2 were calculated using methods described previously (10). These probabilities 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, ***Fig. 1A***). 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).

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 plausible, but poorly understood (10,30). We did not consider childhood vaccination effects here; few individuals in the ADHS data were born at a time when healthy infants were routinely vaccinated against influenza.

## Expected age distributions under alternate imprinting models

If HA subtype-level imprinting protection shapes seasonal influenza risk, primary exposure to HA subtype 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. ***1B***). Alternatively, if broad HA group-level imprinting shapes seasonal influenza risk, then cohorts imprinted to HA subtype H1 or H2 (both group 1) should be protected against modern, seasonal H1N1 (also group 1), while only cohorts imprinted to H3 (group 2) would be protected against modern, seasonal H3N2 (also group 2) (Fig. 1B). Collinearities between the predictions of different imprinting models (Fig. 1G-I) were inevitable, given the limited diversity of influenza antigenic subtypes circulating in humans over the past century (reflected in ***Fig. 1A***). 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).

Our approach distinguishes between age-specific risk factors of influenza infection related to health and social behavior, from birth year-specific imprinting effects. 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. These factors should have similar impacts on any influenza subtype. In contrast, imprinting effects are subtype-specific. Thus, we fit a single step function to characterize the shape of age-specific risk of any confirmed influenza infection. Simultaneously, we modeled residual, subtype-specific differences in risk as a function of birth year, to focus on the possible role of childhood imprinting in H1N1 or H3N2 infections. 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 (as the cohort got older), whereas birth year-specific risk was constant over time.

To test quantitatively whether observed, subtype-specific differences in incidence 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 (31,32). Technical details are provided in the Methods.

## Tested models

We fit a set of four models to the ADHS data set. The simplest model contained only age-specific risk (abbreviated 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): abbreviated 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 represent 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 described the relative risk of confirmed H1N1 or H3N2 infection, given imprinting protection against that seasonal subtype.

**Effect of influenza evolutionary rate on age profiles**

We used publicly available data from *Nextstrain* (33,34), and from one previously published study(35), 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 influenza strains is used as a proxy for similarity in antigenic phenotype, and potential for immune cross-protection. A variety of methods have been developed to estimate antigenic distance using serological data, genetic data, or both (34–36).

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 (27). 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 will be less restricted to children.

# Results

### Subtype-specific differences in age distribution

Seasonal H3N2 epidemics consistently caused more cases in older cohorts, while H1N1 caused a greater proportion of cases in young and middle-aged adults (Figs. 2, S1-S2). These patterns were apparent whether we compared H3N2 epidemic age distributions with those caused by the pre-2009 seasonal H1N1 lineage, or with the post-2009 lineage. Observed patterns are consistent with the predicted effects of cohort-specific imprinting (Fig. 1), and with previously reported differences in age distribution of seasonal H1N1 and H3N2 incidence (11–13,26). See Fig. 2 for seasons where H1N1 and H3N2 co-circulated in substantial numbers, and Figs. S1-S2 for the entire dataset and alternate smoothing parameters.

## Imprinting model selection

The model containing NA subtype-level imprinting received the most statistical support, and the model containing HA subtype-level imprinting was the second most preferred in terms of AIC (Fig. 3, Table 2). 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. 3C-D) 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 supports the hypothesis that 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.

## Fitted risk patterns

Fitted age-specific risk curves took similar forms in all tested models, with risk decreasing rapidly from birth through adolescence, and then decreasing much more slowly until the end of life (Fig. 2A shows the fitted curve from the best model). Estimates of imprinting parameters were less than one, indicating some reduction in relative risk of infection (***Table 2***). Within the best model, estimated reductions in relative risk from childhood imprinting were stronger for H1N1 (0.34, 95% CI 0.29-0.42) than for H3N2 (0.71, 95% CI 0.62-0.82). Table 2 shows parameter estimates and 95% profile confidence intervals from all models fitted.

## Effect of evolutionary rate

Next, to test for effects of evolutionary rate on epidemic age distribution, we searched for increases in the proportion of cases among children in seasons associated with antigenic novelty. The data showed a slight negative but not significant association between annual antigenic advance and the fraction of H3N2 cases observed in children (Fig. 4A). 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 seasonal H1N1 lineages.

If evolutionary rate is the dominant driver of subtype-specific differences in epidemic age distribution, then when subtypes H1N1 and H3N2 show similar degrees of annual antigenic advance, their age distributions of infection should appear more similar. However, the data showed that differences in H1N1 and H3N2’s age-specific impacts did not converge, even when lineages showed similar annual advance (Fig. 4A). When comparing the fraction of cases observed in specific age classes, H1N1 data consistently clustered separately from H3N2, with H1N1 consistently causing fewer cases at the extremes of age (children 0-10 and elderly adults 71-85), but more cases in middle-aged 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. 4B). Overall, the data did not show a strong signal that epidemic age distribution varies systematically 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 H1N1 and H3N2, which was consistent across multiple 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 (27). Our results suggest 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 other influenza subtypes.

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 in the H3N2 pandemic (6), and continue to experience excess H3N2 morbidity and mortality today as elderly adults ((11–13,26), ***Fig. 2***). In short, comparing data from H3N2’s emergence in 1968 to its seasonal impacts today shows impacts that have remained consistent with respect to birth year, but that have shifted with respect to age.

In model comparison, the data showed 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). Realistically, some combination of effects from both HA and NA subtype-level imprinting probably shape seasonal influenza risk. The models containing NA and HA subtype-level imprinting produced very similar fits to data and emerged as the top two models in terms of AIC. Unfortunately, collinearities between predictions of the simple, single-antigen models considered here arose inevitably from influenza’s limited diversity of circulation in humans over the past century. These collinearities prevented us from testing 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 support only 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 (37). 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.

Our failure to detect a strong signal of impact of antigenic advance on age distributions of H1N1 and H3N2 cases was surprising. On the one hand, small sample sizes may have limited our power to detect a statistically significant relationship. The lack of signal is also consistent with growing recognition that existing methods to map antigenic distance between influenza strains rely on hemagglutination inhibition (HI) data collected from naïve ferrets, and do not always capture patterns of cross-reactivity in the human population that has been repeatedly exposed to influenza. Further, glycosylation of HA can cause antigenic escape in large subsets of the human population, yet such posttranslational modifications may be perceived as neutral in existing antigenic maps (38,39). One epidemiologically impactful example of incognito (unmapped) antigenic escape was observed during the 2013-2014 H1N1 epidemic (38). Moreover, existing metrics of evolutionary and antigenic advance are based on properties of HA (33–35), but our epidemiologic data support an equal if not stronger role of NA; more work is required to map NA antigenic changes and possible interactions between HA and NA changes. We speculate that a clearer relationship between epidemic age distribution and antigenic drift would emerge if antigenic distance measures were able to incorporate cohort-specific variation in immune history, and impacts from multiple antigens.

Traditionally, narrow, within-subtype influenza immunity is thought to decay rapidly in the face of antigenic drift. Signals of rapid drift are largely based on hemagglutination inhibition (HI) assay data, which measures antibody responses to just a handful of immunodominant, variable epitopes found near the receptor binding domain on the HA head. These epitopes experience fairly rapid drift, and so strains that circulated more than 14 years apart rarely show measurable cross-protective HI titers (35). The short timescale of immune protection from memory of variable HA head epitopes stands in contrast to patterns observed in our study and others (11–13), where within-subtype immune memory imprinted in childhood appears to persist for an entire human lifetime, remaining evident even in the oldest cohorts in the data. Thus, we speculate that within-subtype imprinting protection arises via different immune mechanisms than the well-studied antibody responses measured by the HI assay.

One possibility is that within-subtype imprinting protection is driven by antibody responses to more conserved epitopes, which might remain stable over time. We rule out a strong role from antibody responses against the best-studied conserved epitopes (e.g. those on the stalk), which tend to provide broad, cross-subtype protection. But responses to less-studied, intermediately conserved epitopes might provide the sort of long-lasting, within-subtype protection supported by the data. Immunological studies show that 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 epitopes held in common between past and current strains (23,24). Repeat boosting of conserved HA or NA antigens could explain the longevity of subtype-level imprinting protection.

Another potential explanation supported by recent immunological data (40), is that the memory B cell clones developed during the first childhood influenza exposure later adapt via somatic hypermutation to “follow” antigenic targets as they drift over time. Thus, the first influenza exposure in life may fill 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. The adaptability of the B cell repertoire would not be detectable in traditional HI panels, which are collected using sera from ferrets exposed to a single influenza variant, and do not reflect the development of the human B cell repertoire across repeated, seasonal influenza exposures. A final possiblility is that cellular immunity (e.g. CD4+ T cell memory), which would not be captured in serological assays, plays an underappreciated role in imprinting protection.

Signals of imprinting protection are anomalously strong in the current cohort of elderly adults, as reflected by higher estimates of imprinting protection for H1N1 than H3N2 in our data. 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 (4,41). 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 after 1968 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 post-1968 birth cohorts eventually become elderly. Determining the precise immune mechanism(s) responsible for subtype-level imprinting is necessary to project long-term shifts in influenza-related incidence, and possibly in mortality. The vast majority of influenza-related deaths occur in adults over age 65, and H3N2 has caused many times the number of fatalities in high-risk elderly cohorts as seasonal H1N1, even in the post-2009 pandemic period (12,26,42). These patterns may arise because H3N2 is intrinsically more virulent than H1N1, but we speculate that imprinting protection, which currently limits the incidence of clinically-attended H1N1 infection in the elderly may also explain these differences. In the future, cohorts imprinted to H2N2 (born c. 1950-1968) will become elderly, and would expect protection against H3N2 via NA subtype-level imprinting, while HA H2-level imprinting would not be of much use against currently circulating seasonal viruses. 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, especially those born after H1N1 and H3N2 began to co-circulate in 1977, may show a greater ability to act as immunological generalists, with effective defenses against multiple influenza subtypes.

One limitation of this study was that we could not model the impact of seasonal influenza vaccination explicitly, as the vaccination status of subjects in the ADHS data was unknown. We note that the influenza vaccine contains both an H1N1 and H3N2 strain, and so on average, influenza vaccination should protect individuals similarly against both subtypes. However, we also acknowledge that influenza vaccine effectiveness varies by season, age group, and subtype, in ways that are poorly understood and difficult to measure (43). These asynchronous and multi-dimensional shifts in vaccine effectiveness may contribute to variability in H1N1 and H3N2’s age distributions across influenza seasons.

Another limitation of this study was the low number of confirmed cases available in the pre-2009 era. To separate age-specific risk effects from birth year-specific cohort effects, the greatest power will come from large data sets collected continuously over decades, so that individual birth cohorts can be followed as they become considerably older. We emphatically echo earlier calls (44) 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 the burden of H1N1 and H3N2 is shaped by cohort-specific differences in childhood imprinting (8,11,12,45). The finding that such imprinting acts at the HA or NA subtype-level informs 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 antigenic seniority 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. Immunological experiments that consider multiple viral exposures, and cohort studies, in which individual histories of influenza infection are tracked from birth, promise to illuminate how B cell and T cell memory develops across a series of early life exposures. In particular, these studies may provide clearer insights than epidemiological surveillance data into which influenza antigens, epitopes and immune effectors play the greatest role in immune imprinting, and how quickly subtype-specific biases become entrenched across the first or the first few exposures.

# Methods

## 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. 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).

## Splines

In Figure 2, 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. S1-S2. 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:

**1**

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.

### 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.

**2**

### Imprinting (I)

An indicator function defined whether a given prediction vector described risk of confirmed H1N1 or H3N2 infection. Let *fIHxNy* be vectors describing the fraction of cases of each birth year that were protected against strain *HxNy* by their childhood imprinting. We defined *rIHxNy* as free parameters describing the risk of confirmed *HxNy* infection, given imprinting protection. Finally, the factor describing the effect of imprinting (I) was defined as:

**3**

## Likelihood

We used equations 1-3 to generate predicted case age distributions (*p*) for each influenza season (s) in which cases were observed in the data. Then, the likelihood was obtained as a product of multinomial densities across all seasons. If *ns* represents the total number of cases observed in a given 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:

4

## 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 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 (31).

## Antigenic advance

We obtained antigenic distance estimates from *Nextstrain (nextstrain.org)*(33,46)*,* and from source data from Figure 3 in Bedford et al. (35). *Nextstrain* calculates antigenic distance using genetic data from GISAID (47), and using methods described by Neher et al. (34). We analyzed “CTiter” estimates from *Nextstrain*, which correspond to Neher et al.’s “tree model” method. We repeated analyses using estimates from the similar “substitution model” method to verify that our choice of antigenic distance metric did not meaningfully impact our results. 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 (34). To enable visualization of 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 spanned the same range (Fig. S3). The *Nextstrain* data files used in this analysis are archived within our analysis code.

## 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.

# 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. 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.

23. 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.

24. 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

25. 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.

26. 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.

27. 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.

28. 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

29. 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

30. 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.

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

32. 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

33. 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.

34. 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.

35. 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.

36. 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.

37. 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

38. 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.

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

40. 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.

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

42. 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.

43. Lewnard JA, Cobey S. Immune History and Influenza Vaccine Effectiveness. Vaccines. 2018 Jun;6(2):28.

44. 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.

45. 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.

46. 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/

47. 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 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.

# Competing interests

The authors declare no competing interests.

# Disclaimer

This work does not necessarily represent the views of the US government or the NIH.

# Author contributions

MW, KG and JLS conceived of the questions and modeling analysis. CV and MW provided crucial assistance with data access and study design. SB and RB supervised data collection. KG wrote the code and performed analyses, with supervision from JLS, and drafted the manuscript. All authors provided and input on analysis and interpretation of the results, and helped revise and edit the manuscript text.