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

Katelyn M Gostic1, Jennifer Whitaker2, Tristan Clark3, Lars Østergaard4, Winnie Tong5, Cecile Viboud6, Michael Worobey7, and James O Lloyd-Smith1,6\*

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

2Division of Infectious Diseases, Mayo Clinic, Rochester, MN, USA

3Southampton General Hospital UK

4Department of Infectious Diseases, Skejby Hospital, Århus, Denmark

5St Vincent’s Centre for Applied Medical Research, Sydney, Australia

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

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

\* jlloydsmith@ucla.edu

# Abstract

\cmt{KG: Rough outline. Would appreciate your input on the core message.}

\begin{itemize}

\item Two subtypes of influenza A virus currently circulate in humans.

\item Subtype H1N1 has disproportionate impacts in children and young people, whereas subtype H3N2 is more widely distributed across age groups.

\item Subtype H3N2 causes the majority of cases in persons over age 65, and the majority of influenza related deaths.

\item Differences in the age distribution and mortality impact of H1N1 and H3N2 may be linked to childhood imprinting, as older cohorts would have been exposed to H1N1 viruses in childhood, whereas younger cohorts are more likely to have had childhood exposures to H3N2. But few existing studies have analyzed differences in subtype-specific age distribution with respect to single year of age, and to date, no study has analyzed the association between single year of birth (a proxy for childhood imprinting history), and subtype-specific risk.

\item Here, we analyzed seasonal surveillance data collected between ### and ### across ## countries.

\item The data confirmed previous observations that H1N1 primarily impacts younger people, while H3N2 disproportionately impacts older people.

\item We fitted several competing models to the data, and model comparison supported the hypothesis that childhood imprinting to a particular HA or NA subtype shapes birth year-specific risk from H1N1 and H3N2. Model comparison showed little support for heterosubtypic protection at the HA group level.

\end{itemize}

# Introduction

Individuals develop particularly strong, lifelong immunity against influenza viruses antigenically similar to those encountered in childhood, a phenomenon alternatively referred to as original antigenic sin \cite{Francis1953}, antigenic seniority \cite{Lessler2012a}, or antigenic imprinting \cite{Gostic2016b, Ma2011a}. Thus, cohorts born during different antigenic eras can show dramatic differences in pre-existing immunity against modern influenza strains. These birth year-specific differences in immune memory, or cohort effects [CITE], are the suspected cause of numerous anomalies in observed age distributions of influenza incidence, morbidity or mortality.

%% Examples of cohort effects

For example, never in recorded history has pandemic influenza morbidity and mortality been uniformly elevated across the entire human population. Rather, certain cohorts have always shown ostensible immune protection and remained at low risk, while other cohorts have suffered disproportionately. Cohort effects are usually the suspected cause of differences in birth year-specific pandemic impact, and protective childhood imprinting has been associated with reduced risk of mortality or severe disease during every influenza pandemic in the modern epidemiological record: %risk of mortality or severe disease has been dramatically elevated in some birth cohorts, while other cohorts have been spared. In

1918 H1N1\cite{Worobey2014d}, 1957 H2N2\cite{Ma2011a}, 1968 H3N2\cite{Simonsen2004a}, 1977 H1N1 \cite{Kilbourne2006a} and 2009 H1N1\cite{Chowell2009}. Cohort effects are less well studied in the context of seasonal influenza, but may explain relatively low incidence \cite{Khiabanian2009a} and mortality \cite{Thompson2003b, Dushoff2006b} from seasonal H1N1 in elderly, H1N1-imprinted cohorts. Cohort effects are also the dominant driver of birth year-specific risk from emerging, avian subtypes H5N1 and H7N9 \cite{Gostic2016b}.

Recently, we showed that a simple model, based solely on demographic age distribution and reconstructed childhood imprinting patterns, yielded surprisingly accurate retrodictions of observed H5N1 and H7N9 case age distributions. In theory, similar chidhood imprinting reconstructions could be used to predict birth year-specific risk in future influenza pandemics \cite{Gostic2016b}, or seasonal epidemics. Such predictions could inform the distribution of limited vaccine and antiviral doses across age groups, and might eventually pave the way toward personalized, cohort-specific vaccination programs. But useful predictions of birth year-specific influenza risk from reconstructed childhood imprinting patterns depend on an improved understanding of which influenza antigens most strongly drive population immunity in different epidemic contexts.

%childhood imprinting is a strong driver of birth year-specific risk from emerging, avian subtypes H5N1 and H7N9 \cite{Gostic2016b}. Cohort effects are also a suspected driver of birth year-specific differences in risk from seasonal influenza [CITE], although their impacts on established seasonal subtypes has been less well studied than their more dramatic impacts on novel, avian or pandemic strains.

%The emergence of a novel influenza virus can magnify cohort-specific differences in pre-existing immunity, and influenza pandemics have shown some of the clearest evidence for cohort effects. For example, cohorts born before 1957 (who would have imprinted to H1N1 in childhood) were disproportionately protected during the 1977 re-emergence of H1N1 \cite{Kilbourne2006a}, and again during the 2009 H1N1 pandemic \cite{Chowell2009}. Similar associations between protective childhood imprinting and reduced risk of mortality or severe disease have been noted during

%Avian H5N1 and H7N9 influenza viruses have also shown odd, mirror-image age distributions as they spill over into humans. We recently showed that these odd age distributions are surprisingly predictable based solely on population-level reconstructions of childhood imprinting, and demographic age distribution.

%The ability to use reconstructed imprinting patterns to predict age distributions of influenza risk in other contexts would provided obvious advantages for influenza planning, preparedness and response. However, numerous remaining uncertainties about the exact drivers and epidemiological impacts of childhood imprinting must first be resolved.

%But despite growing scientific consensus that cohort effects shape birth year-specific influenza risk, it is difficult to establish definitive proof for cohort effects, and in many cases, their exact impacts remain speculative or unclear.

%One key unknown is which influenza antigens will be the strongest drivers of population-level cohort effects in a given epidemic context.

Influenza has numerous antigenic sites, which drift at different rates, and have shifted or reassorted at different historical time points. Should epidemiological predictions incorporate reconstructed patterns of imprinting specific to hemagglutinin (HA), to neuraminidase (NA), or to some combination of HA, NA and other internal immunogens? For imprinting to HA and NA, another open question is whether imprinting typically shows the strongest epidemiological signals of narrow subtype-level (homosubtypic) immunity or of broad, cross-subtype (heterosubtypic) immunity. As a rule of thumb, homosubtypic immunity typically plays a strong role in protection against familiar, seasonal influenza variants, whereas heterosubtypic immunity is more common against novel influenza variants [CITE]. However, these rules of thumb are based on a highly simplified view of the immune response, and may be fallible.

Given the complexity of antigenic space, and potential for antigenic recycling, the specific antigenic drivers of population immunity against novel, emerging influenza viruses may be especially inconsistent and difficult to predict. For example, population immunity against novel, avian subtypes H5N1 and H7N9 is strongly associated with HA group-level imprinting \cite{Gostic2016b}, as the immune system tends to target conserved HA epitopes upon exposure to a novel influenza subtype [CITE CITE]. However, group-level imprinting is not always the key driver of population immunity against emerging, pandemic strains, as pandemic strains are not always entirely novel to humans. For example, about a third of individuals born before 1957 had pre-existing titers against the 2009 pandemic strain \cite{Hancock2009a}, and these individuals tended to mount narrow, homosubtypic antibody responses against the 2009 pandemic strain (typical of antibody responses against seasonal influenza). Meanwhile, younger individuals who lacked pre-existing titers disproportionately mounted broadly protective antibody responses against conserved HA epitopes \cite{Andrews2015e} (typical of responses against novel, avian strains). In other words, heterosubtypic, HA group-level responses were strong drivers of immunity in some cohorts, while narrower, HA subtype-level responses were more common in other cohorts \cite{Andrews2015e, Gagnon2018}. Given these complexities, it would be difficult to model the relationship between childhood imprinting and protection against the 2009 pandemic strain, even retrospectively. True forward prediction of birth year-specific risk from the 2009 pandemic strain would have been nearly impossible without detailed genetic and antigenic data on the pandemic strain, well in advance of its emergence.

%Older cohorts with pre-existing titers disproportionately mounted antibody responses against variable immunodominant HA epitopes, whereas younger cohorts targeted more conserved HA epitopes.

For seasonal influenza viruses, the specific antigenic drivers of population immunity may be more consistent, and thus, it may be more feasible to predict birth year-specific risk in the seasonal context. The impacts of cohorts effects on seasonal influenza A epidemiology have not been systematically studied, but associations between the age-specific impacts of H1N1 and H3N2, and childhood immune history have been noted. For example, %seasonal subtype H1N1 typically has a low burden in older cohorts, while seasonal subtype H3N2 causes a disproportionate number of cases and deaths in the same cohorts [CITE]. These patterns are broadly consistent with childhood imprinting:

older cohorts would have imprinted to H1N1 as children, and may now be preferentially protected against modern H1N1 variants, but lack protection against H3N2 \cite{Khiabanian2009a}. A number of studies have speculated that childhood exposures to H1N1 may protect elderly cohorts against modern, seasonal H1N1 viruses. Others have noted that, relative to H1N1, H3N2 causes many times the number of influenza related deaths \cite{Thompson2003b, Dushoff2006b}. A lack of protection from childhood imprinting among high-risk elderly cohorts may at least partially explain why the mortality impact of H3N2 is much greater than that of H1N1. %To date, no systematic analysis has established whether the distributions of seasonal H1N1 and H3N2 incidence show consistent differences with respect to birth year. Furthermore, specific hypotheses on the antigenic drivers of ostensible imprinting protection have not been tested, and so it remains unclear whether childhood imprinting to HA, to NA, or to some broader combination of influenza antigens is the strongest driver of cohort-specific seasonal influenza risk.

If protection from childhood imprinting indeed currently attenuates the mortality impact of H1N1 in the elderly, then H1N1 may become more deadly in the future as cohorts with different imprinting histories eventually enter high-risk elderly age groups. The ability to anticipate such generational shifts in subtype-specific influenza risk would provide obvious advantages for seasonal influenza risk assessment.

Here, we developed methods to test whether observed differences in the age-specific impact of H1N1 and H3N2 are quantitatively associated with birth year-specific differences in imprinting, and whether these associations persist across time and space. We analyzed differences in observed age distributions of H1N1 and H3N2 incidence using two large epidemiological data sets, which collectively represented influenza A circulation across 23 years and 16 countries.

Preliminary analyses showed that birth year-specific differences in observed distributions of H1N1 and H3N2 cases are qualitatively consistent with patterns of childhood imprinting: H3N2 consistently has greater impacts in older cohorts than H1N1, whereas H1N1 consistently shows disproportionate impacts in younger cohorts.

Next, we developed a suite of antigen-specific hypotheses for the distribution of protection across birth years, and estimated probabilities of imprinting protection for each single year of birth from 1918 to 2017 under each hypothesis. Specifically, we tested whether birth year-specific imprinting at the hemagglutinin (HA) subtype level (driven by immune memory of variable HA epitopes), at the HA group level (driven by immune memory of conserved HA epitopes), or at the neuraminidase (NA) subtype level, could best explain observed birth year-specific differences in risk from H1N1 and H3N2. We used likelihood-based model comparison (AIC) to assess competing antigenic hypotheses. In reality, the immune response against influenza is incredibly complex, and driven by memory of internal proteins, in addition to epitopes on HA and NA. However, it would not be tractable to reconstruct patterns of population-level imprinting to every known antigenic site, or to model each site's respective contribution to overall immune protection. Rather, our approach aimed to test whether a model that reconstructed patterns of childhood imprinting specifically to conserved or variable regions of HA, or to NA was sufficient to capture the strongest signals of imprinting in epidemiological data. Model comparison (AIC) showed the strongest support for imprinting effects at the NA subtype or HA subtype level.

%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% RESULTS

%%%%%%%%%%%%%%%%%%%%%%%%%%%%

\section\*{Results}

\begin{comment}

\subsection\*{Review of observed imprinting patterns}

%% -----------------------

%% Table 1

%% -----------------------

\begin{table}[!ht]

\begin{adjustwidth}{-1.5in}{0in} % comment out/remove adjustwidth environment if table fits in text column.

\centering

\caption{{\bf Example Table.} This table demonstrates how to use the adjustwidth environment if you need that tad bit of extra width for your figures or tables.}

\setlength\tabcolsep{1pt}

\begin{tabular}{c|c|c|c}

\hline

Challenge & Year & Cohorts protected & Possible antigenic explanation \\

\hline

H5N1 avian influenza & 1997-2018 & Born before 1968 pandemic & Imprinted to HA group 1 (heterotypic immunity)\\

H7N9 avian influenza & 2013-2018 & Born after 1968 pandemic & Imprinted to HA group 2 (heterotypic immunity) \\

H2N2 pandemic & 1957 & Born before 1918 pandemic & Possible imprinting to a virus similar to H2N2 (speculative) \\

H1N1 pseudo-pandemic & 1977 & Born before 1957 pandemic & Imprinted to H1N1 \\

H1N1 pandemic & 2009 & Born before 1957 pandemic & Imprinted to H1N1 \\

H3N2 seasonal &c.1990 - present & Older cohorts - high risk & Subject of this study \\

H1N1 seasonal &c.1990-present & Older cohorts - low risk & Subject of this study \\

\hline

\end{tabular}

\end{adjustwidth}

\end{table}

\end{comment}

%% -----------------------

%% Subsection: Reconstructed imprinting patterns

%% -----------------------

\subsection\*{Reconstructed imprinting patterns}

%% -----------------------

%% Figure 1

\begin{figure}[ht] %s state preferences regarding figure placement here

\includegraphics[width=\textwidth]{Fig1\_hypotheses.pdf}

\caption{\color{Gray} \textbf{Hypotheses and modeling approach}. \textbf{A}, Reconstructed imprinting patterns. Plot shows a representative example for Thailand during the Northern Hemisphere 2016-2017 influenza season. Imprinting probabilities are specific to birth year, so age-specific reconstructions shift over time as specific birth years get older. \textbf{B} Qualitative predictions for imprinting at the HA group level, HA subtype level and NA subtype level. \textbf{C}, Cartoon of hypothetical age-specific risk. This hypothetical plot is not driven by data, but each model tested below (Figs. \ref{fig4}-\ref{fig5}) fit a similar step function to data. \textbf{D-F} Quantitative predictions from HA group-level, HA subtype-level and NA subtype-level imprinting reconstructions. Points show probability of imprinting protection against a particular subtype. \textbf{G-I} Hypothetical predictions from models that combine age-specific risk and imprinting risk. Plots assume the relative risk of infection, given imprinting protection, is 0.75.}

\label{fig1} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

We reconstructed birth year-specific childhood imprinting patterns, as described previously \cite{Gostic2016b}. Figure \ref{fig1}A shows one representative reconstruction for Thailand during the 2016-2017 Northern Hemisphere influenza season. Imprinting patterns mirror the timeline of influenza circulation in humans, in which pandemic years mark the transitions between eras of subtype-specific circulation. Older cohorts born between pandemics in 1918 and 1957 all imprinted to H1N1 as children. Middle-aged cohorts born between pandemics in 1957 and 1968 imprinted to H2N2. Ever since its emergence in 1968, H3N2 has dominated seasonal circulation in humans, and caused the majority of imprinting in younger cohorts. However, H1N1 has also caused some seasonal circulation since 1977, and thus a fraction of post-1977 cohorts are imprinted to H1N1 (Fig. \ref{fig1}A).

%% -----------------------

%% Expectations given broad vs. narrow imprinting

%% -----------------------

\subsection\*{Expectations under different models of imprinting}

We considered three models for the breadth of imprinting: HA group-level protection, HA subtype-level protection and NA subtype-level protection. All tested models assumed imprinting to past variants of H1N1 would provide lifelong protection against modern, seasonal H1N1. Similarly, all models assumed imprinting to past variants of H3N2 would provide lifelong protection against modern, seasonal H3N2. Only the imprinting protection status of H2N2-imprinted cohorts differed between models (Fig. 1B).

The HA group-level imprinting model assumed, H2N2 imprinted cohorts would be protected against modern H1N1 viruses, as H2 an H1 are in the same HA group. The HA subtype-level model assumed H2N2 imprinted cohorts would not be protected against either modern seasonal subtype. Finally, the NA group-level imprinting model assumed H2N2 imprinted cohorts would be protected against modern H3N2, due to the shared N2 subytype.

Collinearities between the predictions of different imprinting models were inevitable, given the limited diversity of influenza circulation in humans over the past century (Fig.1 D-I). However, differences in the shape of predicted risk in middle-aged, H2N2 imprinted cohorts provide some leverage to differentiate between hypotheses.

%% -----------------------

%% Subsection: Modeling approach

%% -----------------------

\subsection\*{Modeling approach}

% Age-specific factors shape seasonal influenza risk alongside birth year-specific effects.

%The intrinsic correlation between age and birth year is a key challenge for any analysis of imprinting effects on seasonal influenza epidemiology, as the respective impacts of age and birth year are easily confounded. %In the distant future, this correlation may be broken as large epidemiological data sets grow to span decades, and eventually track individual birth cohorts across an entire human lifetime.

The intrinsic correlation between age and birth year is a key challenge for any analysis of imprinting effects on seasonal influenza epidemiology. To tease apart age-specific risk factors from birth year-specific imprinting effects, we noted that age-specific risk factors are largely subtype-independent. Specifically, age-specific risk, or probabilities of case ascertainment 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. But 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 between H1N1 and H3N2 as a function of birth year (i.e. as a function of imprinting status). We used model comparison (AIC) to assess which, if any, of three tested imprinting models best explained observed subtype-specific differences in risk (Fig. \ref{fig1}).

%All three imprinting hypotheses predicted preferential protection against H1N1 in older, H1N1-imprinted cohorts, and preferential protection against H3N2 in younger, largely H3N2-imprinted cohorts. However, each model predicted a different pattern of protection in middle-aged, H2N2-imprinted cohorts. Respectively, the HA group-level, NA subtype-level and HA subtype-level models predicted H2N2 imprinted cohorts would have preferential protection against H1N1, against H3N2 or against neither subtype (Fig. \ref{fig1} B, D-F).

Figure \ref{fig1} (C-I) shows a cartoon of the overall modeling approach, in which a linear combination of age-specific risk and birth year-specific probabilities of imprinting protection defined predicted case age distributions for each subtype. Note that the age-specific risk curve shown in panel C is purely hypothetical, but probabilities of imprinting protection (D-F) reflect quantitative estimates from reconstructed patterns (e.g. Fig. \ref{fig1}A).

We used a multinomial likelihood to fit model-predicted risk distributions to observed age distributions of infection. Within each model, one free parameter described the relative risk of infection, given imprinting protection, and additional free parameters defined the age-specific risk function. When fitting to data from the INSIGHT study, additional patient details were available, which allowed us to fit parameters quantifying relative risk given influenza vaccination in the previous year, given antiviral treatment, given the presence of underlying conditions. Parameters were estimated simultaneously using maximum likelihood. Technical details and a link to all relevant code are provided in the Methods.

%% -----------------------

%% Subsection: Data

%% -----------------------

\subsection\*{Data}

We analyzed two, large epidemiological data sets. First, the Arizona State Department of Health Services (AZDHS) provided a data set of 18,812 influenza cases, each confirmed to subtype. Cases were observed from the 1993-1994 influenza season to the 2014-2015 season. Confirmed influenza cases were reported to a statewide surveillance database by the Arizona State Public Health Laboratory, and by participating commercial labs. Data from Arizona included only patient birth year, subtype and season in which each confirmed influenza case was observed. Denominator data on the age distribution of all tested cases were not available.

The second data set was provided by the INSIGHT outpatient study, and included 11,575 tested cases (1,998 PCR positive for H1N1 and 2,080 PCR positive for H3N2). Adult patients (18+) were enrolled after reporting to a participating medical provider with influenza like illness (ILI). Patients were enrolled from 17 countries, beginning in the 2009-2010 Northern Hemisphere (NH) influenza season, and ending with the 2016-2017 NH season. The INSIGHT data included some details that were not available in the Arizona surveillance data set, including relevant medical history, (influenza vaccination in the previous year, antiviral treatment used, underlying conditions present), and denominator data on the age distribution of cases that tested negative for influenza. We incorporated these additional data into our modeling analysis, and as a result, more factors were tested in model comparison when fitting to the INSIGHT data than when fitting to the Arizona data.

%% -----------------------

%% Subsection: Observed age distributions

%% -----------------------

\subsection\*{Observed age distributions}

%% -----------------------

%% Figure 2

\begin{figure}[ht] %s state preferences regarding figure placement here

\includegraphics[width=\textwidth]{age-dists-by-season-ARIZONA-ages18-90.pdf}

\caption{\color{Gray} \textbf{Observed age distributions, Arizona, 1993-2015.} Points show fraction of confirmed H1N1 or H3N2 cases observed in each single year of age (18-90). Lines show a smoothing spline fit to observed distributions. \textbf{A}, Age distributions of all observed cases. \textbf{B-G}, Season-specific age distributions from seasons in which H1N1 and H3N2 both circulated (i.e. seasons in which at least 50 cases of H1N1 and of H3N2 were confirmed).}

\label{fig2} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure 3

\begin{figure}[ht] %s state preferences regarding figure placement here

\includegraphics[width=\textwidth]{age-dists-by-country-and-season-INSIGHT.pdf}

\caption{\color{Gray} \textbf{Observed age distributions, INSIGHT outpatient data, 2009-2017.} Points show fraction of confirmed H1N1 or H3N2 cases observed in each single year of age (18-90). Lines show a smoothing spline fit to observed distributions. \textbf{A}, Age distributions of all observed cases. \textbf{B-G}, Season-specific age distributions. To facilitate comparison of H1N1 and H3N2 age distributions, only seasons in which H1N1 and H3N2 cocirculated are shown (i.e. seasons in which at least 50 cases of H1N1 and of H3N2 were confirmed). Case age distributions from all seasons, and plots showing alternate smoothing parameters shown in Figs. \ref{figS1}, \ref{figS2}.}

\label{fig3} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

To compare observed H1N1 and H3N2-specific age distributions of infection, we first plotted observed distributions from within each data set overall, and for individual countries and seasons in which H1N1 and H3N2 both circulated (Fig. \ref{fig2}-\ref{fig3}). Data from the INSIGHT study only included adults, (ages 18+), so we excluded children under age 18 from the plots in Fig. \ref{fig2} to facilitate comparison between data sets. Plots specific to countries or seasons in which few cases were observed, or in which only one subtype circulated, plots that include children age 0-17, and plots showing alternate smoothing parameters are shown in (Figs. \ref{figS1}-\ref{figS7}).

Age distributions of infection in the Arizona data were qualitatively consistent with the cohort effects presented in Fig. \ref{fig1}. In particular, H3N2 consistently caused a disproportionate number of cases in the oldest cohorts, while H1N1 caused a disproportionate number of cases in younger cohorts (Fig. 2).

Patterns observed in the INSIGHT data were also broadly consistent with cohort effects, in that H3N2 often showed disproportionate incidence in older adults, while H1N1 often showed disproportionately impact young adults. However, these signals were weaker and more variable in the INSIGHT data than in the Arizona data. The magnitude of subtype-specific differences in age distribution was generally smaller in the INSIGHT data than in the Arizona data, and the shape of subtype-specific differences in age distribution also fluctuated more from season to season. Clear differences in the age distribution of H1N1 and H3N2 cases were observed in some INSIGHT countries and seasons, while differences were barely perceptible in others (e.g. Argentina, Peru, NH 2013-2014). But despite a weaker overall signal of ostensible cohort effects in the INSIGHT data, observed patterns never contradicted expected imprinting patterns. In other words, whenever subtype-specific differences in age distribution were observed in the INSIGHT data, H3N2 always showed disproportionate impacts in the oldest cohorts, while disproportionate impacts of H1N1 were typically seen in younger cohorts.

\subsection\*{Model comparison}

%% -----------------------

%% Figure 4

\begin{figure}[ht] %s state preferences regarding figure placement here

\includegraphics[width=\textwidth]{AZ\_model\_fits.pdf}

\caption{\color{Gray} \textbf{Model fitting and comparison, AZ data.} Fitted effects from the best model (AN) of age \textbf{A} and imprinting \textbf{B}, as well as maximum likelihood parameter estimates, and profile confidence intervals \textbf{C}. Black points indicate effects are subtype-independent (i.e. effects impact risk from both seasonal subtypes. Blue indicates an effect on H1N1 risk and red indicates and effect on H3N2 risk. Dashed lines show relative risk of 1; values below 1 indicate reduced risk, while values above 1 indicate elevated risk, relative to baseline. \textbf{D}, Model comparison results, with shaded rectangles indicating which factors were included in each model. \textbf{E-F}, model fits to observed data.}

\label{fig4} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure 5

\begin{figure}[ht] %s state preferences regarding figure placement here

\includegraphics[width=\textwidth]{INSIGHT\_model\_fits.pdf}

\caption{\color{Gray} \textbf{Model fitting and comparison, INSIGHT data.} Fitted effects from the best model (ATVN) of \textbf{A} age, \textbf{B} antiviral treatment, \textbf{C} vaccination, and \textbf{D} imprinting, as well as \textbf{E} maximum likelihood parameter estimates, and profile confidence intervals. See Fig. \ref{fig4} legend for plotting conventions. \textbf{F}, Model comparison results, with shaded rectangles indicating which factors were included in each model. Models with $\Delta AIC\leq4$ are show in in color, while less preferred models are shown in gray. \textbf{G-H}, model fits to observed data. Only the best four models, and the null model (A) are shown.}

\label{fig5} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

Although visual assessment of the data shows somewhat consistent differences in birth year-specific risk from H1N1 and H3N2, a more formal analysis was necessary to test for quantitative associations between the data and each of three antigen-specific imprinting hypotheses. We used model comparison to test which, if any, of the three proposed imprinting hypotheses, (HA subtype level protection, HA group level protection or NA subtype level protection) was most consistent with observed differences in age distribution of H1N1 and H3N2 cases.

\subsubsection\*{Interpretation of age-specific risk parameters}

\cmt{KG: THIS FEELS LIKE DISCUSSION MATERIAL, BUT I ALSO THINK IT'S IMPORTANT FOR PROPER INTERPRETATION OF RESULTS... THOUGHTS?}

All models fit to INSIGHT data used denominator data on the age distribution of all tested cases as baseline expected age distribution in the corresponding country and season. Effectively, this denominator data described the age-specific risk of developing ILI and seeking medical treatment, and thus allowed us to model a large component of the age-specific risk and observation probabilities explicitly. Then, the age curve fitted to INSIGHT data only represented residual, age-specific differences in the rate at which subjects tested positive for influenza A. Age-specific differences in the test-positive rate might reflect true differences in the proportion of ILI caused by influenza viruses in different age groups. Differences in the test positive rate might also reflect that different age groups (e.g., students, working adults, retirees and the elderly), have different severity thresholds for seeking medical care. Ultimately, all fitted parameters that characterized age-specific relative risk in the INSIGHT data took values very close to 1, and had confidence intervals that overlapped 1, suggesting that test-positive rates did not differ significantly across age groups (Fig. \ref{fig5} A,E).

On the other hand, denominator data on the age distribution of tested cases were not available in the Arizona data set, so the age-specific risk curve fitted to Arizona data represented all facets of the infection and observation process, including (i) the age-specific risk of developing ILI and seeking medical treatment, and (ii) the age-specific rate of testing positive. The curve fitted to Arizona data thus shows more dramatic differences in age-specific risk, with relative risk being highest in the fixed reference group (children ages 0-4), and decreasing with age.

\subsubsection\*{Fits to Arizona data}

We fit a suite of four models to the Arizona data set. The simplest model contained only age-specific risk (A), and more complex models added effects from imprinting at the HA subtype level (AS), at the HA group level (AG), or at the NA subtype level (AN). 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+}. Models containing imprinting effects each added two additional free parameters, which described the relative risk of infection given birth year-specific imprinting protection against H1N1, or against H3N2, respectively.

Fitted age-specific risk curves took similar forms in all models, with risk being highest in children, decreasing until about age 40, and then decreasing much more slowly until the end of life. Parameter estimates consistently indicated that childhood imprinting protection against H1N1 was stronger than protection against H3N2. Figure \ref{fig4} A-C shows fitted effects, and parameter estimates with 95\% profile confidence intervals from the best model, AN. Estimates from all models fitted to Arizona Data are shown in Table \ref{tabS3}.

The best model included age-specific risk and imprinting at the NA subtype level (model AN). The second best model (AS) included HA subtype-level imprinting and had $\Delta AIC=21.65$. Model comparison showed effectively no statistical support for the null model (A), or the model that included group-level imprinting (AG), as each had $\Delta AIC > 400$ (Fig. \ref{fig4} D).

Difference in overall model fits to observed data are shown in Figure \ref{fig4} E-F. Because models were fit to a very large data set (N = 18,812 confirmed cases), small differences in model fit produced very large differences in model likelihood, and in AIC. Visual assessment confirms the two best models (AN and AS) provided better fits to data than models AG and A. The model containing no imprinting effects, A, did not allow for subtype-specific differences in predicted age distribution, and so its fit essentially split the difference between observed H1N1 and H3N2 age distributions. Thus, in older cohorts, model A to overestimated the number of expected H1N1 cases and underestimated the number of H3N2 cases. Model A systematically showed the opposite residual errors in younger cohorts. All three models that include imprinting effects improved upon the fit of model A by allowing some subtype-specific differences in predicted risk, but the model including HA group-level imprinting, (AG) provided the worst fits to H3N2 and H1N1 data within cohorts born before 1968. The two best models, AN and AS, provided the best fit to data in all cohorts, and their fits only differed substantially in the predicted distribution of H3N2 cases within H2N2-imprinted birth years (Fig. \ref{fig4} (as expected given the patterns shown in Fig. \ref{fig1}).

\subsubsection\*{Fits to INSIGHT data}

When fitting to the INSIGHT data, which contained additional medical details, we introduced three additional risk factors: vaccination (V), antiviral treatment (T), and presence of underlying conditions (U). Factors T and U each added one free parameter, which respectively characterized the relative risk of any influenza infection, given antiviral treatment or given the presence of underlying conditions. We assumed vaccination might have slightly different effects on H1N1 risk and H3N2 risk, so factor V added two free parameters which characterized the relative risk of H1N1 or of H3N2 infection, given vaccination against influenza. We tested all possible combinations of V, T, and U, in addition to age-specific risk and each of the three imprinting hypotheses (S, G, N).

The estimated relative risk associated with antiviral treatment consistently took values greater than one, which may reflects the fact that antiviral treatment is most often prescribed after an influenza infection is confirmed. The relative risk associated with vaccination consistently took values less than one, indicating some level of vaccine-induced protection, although confidence intervals around these parameter estimates often overlapped the null value of one, indicating that estimated reductions in risk from vaccination were not always significant. Parameter estimates from the best model suggested that vaccination provided slightly stronger protection against H1N1 than against H3N2. The presence of underlying conditions was typically associated with a negligible reduction in risk, possibly reflecting a slightly higher rate of medical testing in general within this risk group. Overall, risk parameters fitted to the INSIGHT data took values closer to 1 and had wider confidence intervals than risk parameters fitted to the Arizona data, which is consistent with the fact that the INSIGHT data contained fewer confirmed cases and showed weaker are-specific or subtype-specific differences in risk than the Arizona data. Parameter estimates and confidence intervals from the bets model fit to INSIGHT data are shown in Fig. \ref{fig5} all estimated from models are shown in Table \ref{tabS4}-\ref{tabS5}.

The model containing age-specific risk, antiviral treatment, vaccination and NA imprinting (ATVN) performed best in terms of AIC, however four other models, ATVS, ATUVN, ATV and ATUVS received almost equally strong statistical support with $\Delta AIC$ of 0.45, 1.95, 2.09 and 2.4, respectively (Fig. \ref{fig5}). We considered INSIGHT models with $\Delta AIC \leq 4$ to be top models. Two of the top models contained NA imprinting, two contained HA subtype-level imprinting, and one contained no imprinting effects. None of the top models contained impacts from imprinting at the HA group level, and overall, model selection showed the strongest support for imprinting effects at the HA or NA subtype level.

Akaike weights can be interpreted as the proportional support for a given model, out of all models tested. The total Akaike weight for models including imprinting at the NA subtype level was 0.44, and the total Akaike weight for models including imprinting at the HA subtype-level was 0.36. Although no single model fit to INSIGHT data was definitively preferred in terms of AIC, models containing imprinting at the NA or HA subtype level had a combined Akaike weight of 0.8, and thus received much more statistical support than models containing imprinting at the HA group level, or containing no imprinting effects.

\section\*{Discussion}

%%%%%%%

%% DRAFT SUMMARY

%%%%%%%

We analyzed two large, epidemiological data sets on influenza A to assess whether suspected impacts from childhood imprinting consistently shape birth year specific risk from H1N1 and H3N2 influenza. First, we reconstructed birth year-specific imprinting histories for all birth years from 1918 to the present, and then developed three antigen-specific hypotheses to describe expected patterns if childhood imprinting to conserved HA epitopes, variable HA epitopes, or to NA was the dominant driver of birth year-specific risk (Fig. \ref{fig1}). Then, we compared differences in the age distribution of risk from H1N1 and H3N2 across influenza seasons from 2007 to 2016, and found that H3N2 consistently caused more cases than H1N1 in older cohorts, while H1N1 consistently had the highest impacts in younger cohorts. In general, these differences between subtype-specific age distributions were qualitatively consistent with population-level imprinting protection (Fig. \ref{fig1}), although the magnitude and shape of differences between H1N1 and H3N2 distributions varied somewhat across time and space (\ref{fig2}-\ref{fig3}). Finally, we fit a suite of models to data using maximum likelihood, and performed model comparison to test which candidate imprinting hypothesis, if any, was most consistent with observed differences in risk from H1N1 or H3N2. Model comparison on both INSIGHT and Arizona data sets showed the greatest support for imprinting effects at the HA subtype level, or NA imprinting.

%%%%%%%%%%%%%%%%%%%%%

%% Differences between AZ, INSIGHT

%% \* age-specific sampling

%% \* recommendations

%%%%%%%%%%%%%%%%%%%%%

Overall, subtype-specific differences in age distribution were much less pronounced in the INSIGHT data than in the Arizona data. Given that differences in vaccination rate [CITE], population density [CITE], latitude [Cite], and connectivity [CITE] are known to modulate the intensity and seasonality of influenza A, observed differences between INSIGHT study data and Arizona data may reflect true differences in the distribution of subtype-specific impacts between Arizona, and the various countries in which INSIGHT data was collected. However, study-specific differences may also be due, at least partially, to study-specific or geographical differences in case ascertainment. The INSIGHT study did not enroll children, and enrolled adult populations contained relatively few subjects under the age of 25, or over the age of 70 (Fig. S\ref{figS8}). The relative dearth of cases at the extremes of age in the INSIGHT study may have dampened the signal of subtype-specific differences in risk; patterns of imprinting protection are much more distinct among the youngest and oldest cohorts (Fig. \ref{fig1}A,D-F), which were poorly sampled in the INSIGHT study, than among well-sampled middle aged adults. The age distribution of tested cases was not known in the Arizona data, but given the large numbers of confirmed cases in children, teens and the elderly (Fig \ref{figS8}), it is obvious that the extremes of age were comparatively well-sampled.

cmt{I'm considering adding a supplementary figure here, where we down-sample the Arizona data using sampled age distributions from the INSIGHT study. I'm curious to see whether a subset of cases from Arizona that largely excludes children and older adults erodes the signal of subtype-specific differences in age distribution. But I also already have tons of supplementary figures. Is this worth pursuing or should I leave it?}

%% Ideal data

The potential for age-specific sampling biases to erode or magnify the signal of imprinting effects in a given data set highlights some limitations of existing epidemiological data for analysis of birth year-specific imprinting patterns. The ideal comparison of H1N1 and H3N2's age distributions across multiple countries and multiple influenza seasons would be informed by an extremely large data set, in which all ages were sampled evenly, or proportionally to their demographic representation. Furthermore, a sufficient number of cases would be observed in each sampled season and location to precisely characterize a distribution across 90-100 single years of age, or birth years. Ideally, several hundred or more cases of each subtype would be observed in each sampled country and season, and data would be available continuously for several decades of influenza circulation. Long-term data collection is necessary to track subtype-specific circulation over time, as H1N1 and H3N2 do not circulate strongly each and every year. Long-term data collection also provides the ability to track individual birth years over time, thus eroding the correlation between birth year and age. In reality, sampling for influenza is usually somewhat opportunistic, uneven across age groups, and the total size or surveillance data sets changes over time. Although large surveillance data sets often span several decades, the total number of recorded cases has increased dramatically over the past decade, and sample sizes from the 2009 pandemic are often prohibitively small. Pandemic years or seasons in which influenza circulation has a particularly high burden tend to be relatively heavily sampled.

%% Recommended changes to data sharing

The INSIGHT data set analyzed here is remarkable in its size, geographic coverage, and in its inclusion of medical history and denominator data. However, as a focused clinical study, even the INSIGHT data set is smaller than ideal for population-level immuno-epidemiological inference. Seasonal surveillance data are of a more appropriate scale, and already collected regularly by most state or national public health agencies, but existing reporting and data sharing practices limit the utility of seasonal surveillance data for immuno-epidemiological inference. One major challenge is that when influenza surveillance data are shared publicly, age is often reported in 5-year bins, or broad categories, whereas single year of age, or birth year, is necessary to study cohort effects. Furthermore, the absence of denominator data on the age distribution of all tested subjects in many surveillance data sets poses a serious challenge for statistical inference. The expense and difficulty of maintaining large, public databases should not be taken for granted, and those responsible for collecting and curating high-quality data deserve more professional credit for their work. But if influenza surveillance data were more systematically shared with single year of age, and age-specific denominators, it could represent a turning point in the ability to leverage, massive, existing, seasonal surveillance data sets for basic science research linking influenza's immunology and virology with its epidemiology. As we enter the era of big data, a widespread data sharing initiative for influenza surveillance data would also lay the groundwork for a decades-long repository, in which specific birth cohorts could be tracked over time as they grew older, and in which the epidemiological impacts of relatively rare events (e.g. antigenic cluster transitions) could be systematically studied.

cmt{I PROBABLY NEED TO TONE THIS DOWN, BUT WANTED TO SEE WHAT YOU GUYS THINK.}

%% Interpretation of INSIGHT fits

cmt{I NEED TO WORK ON TRANSITION}

Comparison of models fit to INSIGHT data showed nearly equal statistical support for five tested models that all had $\Delta AIC < 4$. The spread of $\Delta AIC$ values among all tested models was also much narrower among models fit to INSIGHT data than among models fit to Arizona data. This is unsurprising given that the signal of imprinting effects was weaker overall in the INSIGHT data than in the Arizona data (Fig. \ref{fig2}-\ref{fig3}). The simplest model fit to INSIGHT data (A), whose fit was based solely on the observed age distribution of all tested cases, and on the fitted age-specific risk curve, provided a fairly good fit to observed data, and additional effects from patient medical history or imprinting improved on this simplest fit only marginally (Fig. \ref{fig5} G-H). Because the fitted data set was large (4,078 confirmed influenza A cases), these small improvements in fit translated to statistically meaningful differences in model likelihood, and in AIC.

%% Interpretation of AZ fits

Among models fit to the Arizona data, the spread of $\Delta AIC$ values was much wider, and visually, competing models showed much more dramatic differences in goodness of fit (Fig. \ref{fig4}). The best model fit to Arizona data, AN, would normally be interpreted as having definitive statistical support. Based on widely used rules of thumb for the interpretation of $\Delta AIC$ values, the second best model (AS, $\Delta AIC=21.65$), would normally be ruled out as inferior, and we should conclude that childhood imprinting to NA is a much better predictor of subtype-specific seasonal influenza risk than HA subtype-level imprinting. However, the only qualitative difference between the predictions of model AN (best) and model AS (second best), is the protection status of H2N2 imprinted cohorts against seasonal H3N2 (fits in Fig. \ref{fig4} E-F, consistent with differences between hypotheses in Fig. \ref{fig1} B). Visually, neither model AN nor AS is a precise fit to the observed data in the H2N2-imprinted region where their predictions differ (Fig. \ref{fig4} E-F), and so we hesitate to assert that NA-specific childhood imprinting is a much better predictor of seasonal influenza risk than HA subtype-specific imprinting. Rather, fits to the AZ data set show that models including either NA-level or HA subtype-level childhood imprinting provide better fits to data than models assuming HA group-level imprinting (AG), or no imprinting effects. However, none of the tested models provided a perfect fit to the Arizona data in all birth cohorts, indicating that models considering only the impacts of childhood imprinting to one antigen, HA or NA, are probably insufficient to precisely capture observed epidemiological patterns. Some combination of effects from HA imprinting, NA imprinting, and from imprinting to internal proteins most likely drives birth year-specific risk.

%% Complexity/colinnearity issues

We considered testing more complex models that would have incorporated a mixture of imprinting effects, but ultimately decided that, given extensive collinearities between the predictions of even the simple models tested here, the data were unlikely to support a more complex analysis. Deeper insights into the relative contributions of HA and NA to seasonal imprinting patterns are needed, but given the limited diversity of influenza circulation in humans over the past century, epidemilogical data can support a limited scope of inference. Deeper insights will most likely need to come from immunological data. The recent NIH initiative to fund large cohort studies [CITE], and technological advances in the sequencing of B and T cell repertoires [CITE].

%% HA and NA subtype-level protection is best

Model comparison on both data sets independently provided the strongest support for effects from childhood imprinting to NA, or subtype-level protection from imprinting to HA. These results are consistent with decades of research on seasonal influenza immunity, in which titers against HA and NA have always been considered the strongest correlates of protection, and in which narrow, homosubtypic immunity has been considered the norm against seasonal influenza.

%% Potential immunological explanation

Still, it is somewhat surprising given the known influence of antigenic drift, that signatures of homosubtypic imprinting protection persist across an entire human lifetime. On average, antigenic clusters persist for only 3.3 years [SMITH 2004]. Circulating H1N1 and H3N2 viruses drift by 0.62 and 1.01 antigenic units per year, respectively [cite Bedford eLife paper], which roughly correponds to a two-fold drop in titer for every 1.61, or 0.99 years of antigenic evolution between viruses. Cross-protective titers between viruses sampled 15 or more years apart are consistently too low to measure [Bedford, elife], indicating that cross-protection should be weak or absent. Thus, it is somewhat puzzling that influenza immunity primed in childhood provides any meaningful protection after adolescence, let alone throughout adulthood and into old age.

One possible explanation for the evident longevity of protection from childhood imprinting is that imprinting to a particular HA or NA subtype builds strong memory to epitopes that are relatively conserved at the subtype level, and drift more slowly than the epitopes that react in hemagglutination inhibition assays.

Another possible explanation is that the exact B cell clones developed during the first childhood influenza exposure are not necessarily identical to the clones deployed against strains encountered later in life. Memory B cells maintain some phenotypic plasticity, and through somatic hypermutation, have some capacity to tailor their binding affinity to match newly encountered antigens. Thus, childhood imprinting may provide preferential, lifelong protection against a particular HA or NA subtype by filling a child's B cell repertoire with clones that will serve in the future, not as final products, but as raw prototypes that can be rapidly and effectively tailored to recognize drifted influenza strains of the same subtype, but less effectively tailored to recognize more distantly related, heterosubtypic strains. A third possibility is that childhood imprinting depends strongly on a combination of B and T cell memory. The protective effects of T cells is not measured in the HI or neutralization assays used to build antigenic maps and characterize antigenic drift, and so these studies may systematically underestimate the functional strength of cross-immunity between drifted strains of the same subtype.

%% Group-level

Neither the INSIGHT nor the Arizona data set supported a strong role of HA group-level imprinting. Previously, we showed that childhood imprinting at the HA group level (presumably driven by memory of conserved HA epitopes) is a surprisingly strong predictor of birth year-specific risk from two emerging, avian influenza A subtypes, H5N1 and H7N9. Conserved HA epitopes are known to play a particularly strong role in the immune responses against novel influenza subtypes like avian H5 and H7 [CITE]. On the other hand, more variable HA epitopes are known to be immunodominant upon exposure to a more familiar, seasonal HA subtype [CITE]. Thus, it is not surprising that HA group-level imprinting appears to play a strong role against novel, emerging influenza viruses, whereas subtype-level imprinting plays a dominant role against established, seasonal variants.

%% Epidemiological interpretation

The vast majority of influenza-related deaths occur in adults over age ##, and H3N2 causes an estimated ## times as many deaths in these cohorts than H1N1. This study suggests that the low number of H1N1-related deaths in elderly cohorts is at least partially explained by the fact that childhood imprinting appears to reduce the incidence of clinically attended H1N1 disease in high-risk elderly age groups. H2N2 imprinted cohorts (born c. 1950-1968) will eventually replace the H1N1-imprinted cohorts currently at the highest risk of influenza-related mortality. Results from this study suggest that in the future, elderly H2N2 imprinted cohorts will not be preferentially protected against H1N1 from their childhood imprinting, and so epidemiologists may expect the mortality burden of H1N1 to increase in the future. H2N2-imprinted cohorts may instead enjoy some protection against H3N2 given childhood imprinting to the same NA subtype, or based on HA subtype-level imprinting, may not be protected against either circulating seasonal subtype. In the latter case, the overall mortality impact of influenza A may increase in the future.

NEED SOME SORT OF PUNCHY CONCLUSION...TO WRITE AFTER WE DISCUSS AND SETTLE ON A CORE MESSAGE.

\section\*{Materials and Methods}

\subsection\*{Data and exclusion criteria}

\subsubsection\*{Arizona Data}

The Arizona data set contained the variables birth year, subtype and season for 18,813 cases confirmed in the state of Arizona. AZDHS staff removed identifying details from the data, and only variables of interest to our study were shared. Birth year was extracted from each patient's reported date of birth, except in the 1993-1994, and 1994-1995 seasons, where birth year was approximated as (year of observation)-(age at time of observation). Birth year was necessary to estimate probabilities of imprinting protection for each patient. We calculated patient age as (year observed - birth year), where year observed was set to the second year of the influenza season associated with any case (e.g. 2014 for a case observed in the 2013-2014 season). We excluded one case whose birth year was not available, and ultimately analyzed 18,812 cases.

AZDHS staff classified cases into seasons using the standard definition, with the Northern Hemisphere (NH) season beginning in week 40, and ending in week 39 of the following calendar year (roughly Oct-Sept). Note that in the Arizona data, most H1N1 cases from the 2009 H1N1 pandemic were confirmed and reported during the 2009-2010 season, although some cases from early pandemic waves were reported during the 2008-2009 season. The relatively low number of pandemic H1N1 cases reported during the 2008-2009 season was due to lower case incidence early in the pandemic, and to limited testing capacity during the first pandemic wave. %% See pdf in raw\_data folder

\subsubsection\*{INSIGHT Data}

The INSIGHT outpatient data contained 11,679 total cases, and the variables: age, date of enrollment, antiviral treatment used (Y/N), presence of any underlying conditions (Y/N), influenza subtype (H1N1, H3N2, B or negative), and country. The data also contained variables for influenza vaccination in the past 6 months, or influenza vaccination in the past 12 months. We combined these data into a single binary variable, which classified subjects as vaccinated if they reported recent influenza vaccination at either timescale.

We used patient age and date of study enrollment to impute year of birth. Birth year can be approximated as (observation year)-(age), but this approximation is slightly biased, and biased in different ways for cases observed during the NH and SH influenza seasons. Cases observed earlier in the year (e.g. in January) are more likely to have a birthday that has not yet passed in the current calendar year, whereas cases observed later in the year (e.g. in December) are most likely to have already had a birthday pass in the current calendar year. Thus, the time of year at which a case is observed impacts how birth year should be estimated. Using logic laid out in Fig \ref{figS9}, we used the following three formulas to obtain three possible birth years for cases observed in the NH influenza season: (current year)-(age)-1, (current year)-(age)-, and (current year)-(age)-+1. We then took a weighted average of the three relevant, birth year-specific imprinting protection probabilities, using weights 0.0625, 0.875 and 0.0625, respectively. Meanwhile, for cases observed in a SH season, only two birth years were possible. Here, we used the formulas: (current year)-(age)-1, and (current year)-(age) to calculate both possible birth years, with each receiving a probabilistic weight of 0.5 (Fig \ref{figS9}). Table \ref{tabS1} shows

We excluded 7 subjects whose ages did not fall in the focal age range (18-90). We excluded 94 subjects for whom season, age, vaccination status, antiviral treatment use or presence of underlying conditions was unknown. We also excluded 3 cases of co-infection with multiple subtypes. After excluding 104 cases in total, we ultimately analyzed a data set of 11,575 tested cases, including 1,998 confirmed H1N1 cases and 2,080 confirmed H3N2 cases. Cases were confirmed in 16 countries (Argentina, Australia, Austria, Belgium, Chile, Denmark, Estonia, Germany, Greece, Japan, Peru, Poland, Spain, Thailand, UK, and USA), and our study analyzed cases observed during 16 influenza seasons, starting with the 2009-2010 Northern Hemisphere season (which included the second wave of the 2009 H1N1 pandemic), and ending with the 2016-2017 Northern Hemisphere season. To facilitate comparison between data sets, we defined INSIGHT cases enrolled in Oct-May as part of the Northern Hemisphere (NH) influenza season, and cases enrolled in June-Sept as part of the Southern Hemisphere (SH) season. October 1 roughly aligns with the week 40 NH season start date used in the Arizona data set. Table \ref{tabS2} shows confirmed case counts from each country and season.

\subsection\*{Splines}

In Figures \ref{fig2}-\ref{fig3}, smoothing splines were fit to aid visual interpretation of noisy data. We fit splines using the command \textit{smooth.spline(x = AGE, y = FRACTIONS, spar = 0.8)} in R version 3.5.0. Variables \textit{AGE}, and \textit{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. \ref{figS3}, \ref{figS6} \& \ref{figS7}. Observed differences in H1N1 and H3N2 age distribution were insensitive to our choice of smoothing parameter; although the overall shape of fitted splines changed with different smoothing parameter choices, differences between the splines fit to H1N1 data and the splines fit to H3N2 data usually remained qualitatively consistent within a given season or country.

\subsection\*{Reconstruction of imprinting patterns}

Using methods described previously [CITE], we estimated the probability that an individual born in a given year would have first been exposed in childhood to an H1N1, H2N2 or H3N2 influenza A virus. For birth years of ages 0-12 in the year of case observation, reconstructions included the probability that an individual remained naive to influenza A, and had not yet imprinted. We repeated reconstructions for all birth years, and for all countries and years of case observation relevant to the data. Code to perform reconstructions is available at in the Supplementary Materials.

We calculated probabilities of birth year-specific protection for each of the three tested imprinting hypotheses using the rules defined in Table \ref{tab1}, where $p\_{HxNx}$ gives the estimated probability of imprinting to subtype \textit{HxNx} for a given birth year.

\begin{table}[!ht]

\begin{adjustwidth}{0in}{0in} % comment out/remove adjustwidth environment if table fits in text column.

\centering

\caption{{\bf Estimated probabilities of imprinting protection}, where $p\_{HxNy}$ gives the probability of having imprinted to subtype $HxNy$ in childhood.}

\begin{tabular}{l|l|l}

\hline

\textbf{Imprinting hypothesis} & \textbf{Challenge} & \textbf{p(protected)} \\

\hline

HA subtype-level protection & H1N1 & $p\_{H1N1}$ \\

& H3N2 & $p\_{H3N2}$ \\

\hline

NA subtype-level protection & H1N1 & $p\_{H1N1}$ \\

& H3N2 & $p\_{H3N2}+p\_{H2N2}$$ \\

\hline

HA group-level protection & H1N1 & $p\_{H1N1}+p\_{H2N2}$ \\

& H3N2 & $p\_{H3N2}$ \\

\end{tabular}

\label{tab1}

\end{adjustwidth}

\end{table}

\subsection\*{Model formulation}

Information on individual case medical histories was available in the INSIGHT data set, and so we were able to model age-specific impacts from vaccination prior to the influenza season of interest (V), of underlying symptoms (U) and of antiviral treatment (T) explicitly when fitting to these data. However, when fitting to the Arizona dataset, these patient details were not reported, and so tested models only included age-specific risk (A), and birth year-specific risk from imprinting, which could provide protection at the HA subtype (S), HA group (G), or NA subtype (N) level.

We assumed underlying symptoms (U) and antiviral treatment (T) would have similar impacts on H1N1 and H3N2’s age distributions and therefore fit a single curve to describe these impacts on both H1N1 and H3N2, however, we fit distinct impacts of vaccination (V), and of imprinting protection (I) for each subtype. Below, the abbreviation I will be used to represent any of the three possible imprinting hypotheses, G, S or N.

For each unique country and season in which cases were observed, define $p$ as a vector whose entries, $p\_a$ represent the expected case age distribution for a given subtype, i.e., the expected probability that a randomly drawn H1N1 or H3N2 case was observed in an individual of age a. Then, use a linear model to define $p$, as follows:

\begin{equation}

p = null\*A\*T\*U\*\mathbbm{1}\_{H1N1}(V\_{H1N1}\*I\_{H1N1})\*\mathbbm{1}\_{H3N2}(V\_{H3N2}\*I\_{H3N2})

\end{equation}

Note $\mathbbm{1}\_{H1N1}$ is an indicator function that takes value 1 if $p$ describes the expected age distribution of H1N1 cases, and 0 otherwise. %Thus, $\mathbbm{1}\_{H1N1}$ and $\mathbbm{1}\_{H3N2}$ ensure that subtype-specific risk factors are only included in relevant subtype-specific predictions, whereas subtype-independent factors are always included, regardless of the focal subtype.

\subsubsection\*{Null}

When fitting to INSIGHT data, the null expected age distribution was proportional to the total number of subjects enrolled in the study, regardless of their influenza test outcome. Crucially, this null prediction ensured that the predicted number of cases was only greater than 0 in countries, seasons and ages where one or more subjects was enrolled. This null distribution was particularly important when fitting to the INSIGHT data, where the number of confirmed cases was often low, and where the age distribution of cases in a given country and season was not always fully populated (Table \ref{tabS2}).

AZDHS only collects surveillance data on confirmed influenza cases, and not on the age distribution of ILI cases that tested negative for influenza. Thus, we were unable to include a null distribution in models fitted to data from Arizona. Crucially, this technical difference in model formulation leads to a different interpretation of free parameters fitted to the Arizona and INSIGHT data sets, as noted above.

%When fitting to Arizona data, where no denominator data was available, the fitted age-specific risk curve describes a combination of the age-specific probability of developing ILI and seeking medical treatment, and of testing positive (i.e. the age distribution of all positive tests). On the other hand, when fitting to the INSIGHT data, the null distribution controls for the age-specific probability of developing ILI and seeking medical treatment, while and all other factors describe only the relative risk that an administered test is positive (i.e. the age-specific test positive rate). The age-specific rate of positive tests (INSIGHT fits) is much less variable across age groups than the age distribution of all positive tests (AZ fits), which is why the age-specific risk curve fitted to the AZ data (Fig. \ref{fig4} A), shows a much wider range of relative risk estimates than the age-specific risk curve fitted to the INSIGHT data (Fig. \ref{fig5} A).

\subsubsection\*{Age-specific risk}

Define age-specific risk, ($A$), as a step function, with each step representing the expected fraction of cases of a given subtype observed in each of $Z$ age bins. Fit $Z-1$ free parameters, $r\_z$ to describe relative risk in each age bin, except one arbitrarily chosen reference bin, in which relative risk is fixed to 1. Below, $\mathbbm{1}\_z$ are indicator functions that take value 1 if a given vector entry belongs in age bin \textit{z}, and 0 otherwise. To obtain the predicted fraction of cases observed in each single year of age, we normalized risk distribution given by equation 2 so that predicted risk across all age groups summed to 1.

\begin{equation}

A = norm(\mathbbm{1}\_{1}+\mathbbm{1}\_{2}r\_2...\mathbbm{1}\_{z-1}r\_{z-1}+\mathbbm{1}\_{z}r\_z)

\end{equation}

\subsubsection\*{Antiviral treatment}

Within a given country and season, $f\_T$ is a vector whose entries describe the fraction of tested cases of a given age that received antiviral treatment. Free parameter $r\_T$ defines the relative risk of a positive test, given antiviral treatment. Then, define risk factor \textit{T} as:

\begin{equation}

T = f\_Tr\_T+(1-f\_T)

\end{equation}

\subsubsection\*{Underlying conditions}

The underlying conditions risk factor takes the same form as factor \textit{T}:

\begin{equation}

U = f\_Ur\_U+(1-f\_U)

\end{equation}

\subsubsection\*{Vaccination and Imprinting}

Factors describing risk from vaccination (\textit{V}) and imprinting (\textit{I}) took forms similar to risk factors \textit{T} and \textit{U} (equations 3 and 4), but with subtype-specific impacts. An indicator function defined whether a given prediction vector described risk of confirmed H1N1 or H3N2 infection:

\begin{equation}

V\_{HxNy} = \mathbbm{1}\_{HxNy}(f\_V\*r\_V+(1-f\_V))

\end{equation}

Below, $f\_I$ describes the probability that an individual of a given age, and observed during a particular influenza season, is protected from the subtype of interest by their childhood imprinting.

\begin{equation}

I\_{HxNy} = \mathbbm{1}\_{HxNy}(f\_I\*r\_I+(1-f\_I))

\end{equation}

\subsection\*{Model fitting and model comparison}

We simultaneously estimated all free parameter values using maximum likelihood estimation (MLE). We calculated likelihood profiles for each fitted free parameter, and defined profile confidence intervals based on the range of fixed parameter values for which the profile likelihood exceeded a minimum threshold. The threshold was defined using the method of likelihood ratios, in which Wilks's Theorem states that $-2\ln(\frac{L\_{profile}}{L\_{full}}) \sim \chi ^2$. Then, the threshold is given by the value of $L\_{profile}$ for which the left-hand side of the above inequality is equal to the 95th quartile of a $\chi ^2$ distribution with one degree of freedom. All models fit to a given data set were compared using AIC.

\subsection\*{Code and data availability}

Code to perform all analyses and construct all plots is available \#\#HERE\#\#. Data from AZ AVAILABILITY? Data from the INSIGHT study are available by application, pending approval from the study's scientific review committee (LINK). Because we are not free to share the INSIGHT data, the posted code contains an INSIGHT data file with scrambled column entries.

\newpage

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% Supporting information

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

\section\*{Supporting Information}

%These commands reset the figure counter and add "S" to the figure caption (e.g. "Figure S1"). This is in case you want to add actual figures and not just captions.

\setcounter{figure}{0}

\setcounter{table}{0}

\renewcommand{\thefigure}{S\arabic{figure}}

\renewcommand{\thetable}{S\arabic{table}}

\subsection\*{Figures}

%% -----------------------

%% Figure S1

\begin{figure}[ht]

\includegraphics[width=\textwidth]{age-dists-by-season-ARIZONA-ages0-90.pdf}

\caption{\color{Gray} \textbf{AZ age distributions, including children ages 0-17.} Alternate version of Fig. \ref{fig2} showing observed age distributions with children ages 0-17 included. Children ages 0-17 were not enrolled in the INSIGHT study, and so children ages 0-17 were excluded from main text Fig. \ref{fig2}, in order to facilitate comparison between the Arizona and INSIGHT data sets.}

\label{figS1} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure S2

\begin{figure}[ht]

\includegraphics[width=\textwidth]{age-dists-ARIZONA-all-seasons.pdf}

\caption{\color{Gray} \textbf{AZ age distributions, all seasons.} Alternate versions of Fig. \ref{fig2} showing observed age distributions from all influenza seasons. Observed case fractions (points) were only plotted if 10 or more cases of a given subtype were confirmed, to avoid extreme stretching of the y axis. Smoothing splines were only plotted if 50 or more cases of a given subtype were observed, as fits to fewer data points would not have been meaningful.}

\label{figS2} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure S3

\begin{figure}[ht]

\includegraphics[width=\textwidth]{age-dists-by-season-ARIZONA-hilowsmoothpar.pdf}

\caption{\color{Gray} \textbf{Alternate smoothing parameters, AZ age distributions.} Alternate versions of Fig. \ref{fig2} B-G, with smoothing parameters chosen to fit smoothing splines that are less (\textbf{A-F}), or more (\textbf{G-L}) smooth than the splines shown in the main text.}

\label{figS3} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure S4

\begin{figure}[ht]

\includegraphics[width=\textwidth]{age-dists-INSIGHT-all-countries.pdf}

\caption{\color{Gray} \textbf{INSIGHT age distributions, all countries.} Observed age distributions from all countries in which one or more confirmed cases of either subtype was observed, (as opposed to Fig. \ref{fig3}, in which only countries reporting $>=$ 50 cases of both subtypes were shown). Smoothing splines were only plotted if 50 or more cases of a give subtype were observed, as fits to fewer data points would not have been meaningful.}

\label{figS4} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure S5

\begin{figure}[ht]

\includegraphics[width=\textwidth]{age-dists-INSIGHT-all-seasons.pdf}

\caption{\color{Gray} \textbf{INSIGHT age distributions, all seasons.} Alternate versions of Fig. \ref{fig3} showing observed age distributions from all seasons in which one or more confirmed cases were observed. Smoothing splines were only plotted if 50 or more cases of a give subtype were observed.}

\label{figS5} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure S6

\begin{figure}[ht]

\includegraphics[width=\textwidth]{age-dists-INSIGHT-hiSmoothPar.pdf}

\caption{\color{Gray} \textbf{INSIGHT age distributions, increased smoothness.} Alternate versions of Fig. \ref{fig3}, except with a higher smoothing parameter value (greater smoothness) used to fit smoothing splines.}

\label{figS6} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure S7

\begin{figure}[ht]

\includegraphics[width=\textwidth]{age-dists-INSIGHT-lowSmoothPar.pdf}

\caption{\color{Gray} \textbf{INSIGHT age distributions, increased smoothness.} Alternate versions of Fig. \ref{fig3}, except with a lower smoothing parameter value (less smoothness) used to fit smoothing splines.}

\label{figS7} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure S8

\begin{figure}[ht]

\includegraphics[width=\textwidth]{all-tested-cases.pdf}

\caption{\color{Gray} \textbf{Age distributions of all tested and confirmed cases.} \textbf{A}, Counts and \textbf{B}, frequencies from the Arizona and INSIGHT data. Note that data on the age distribution of all tested cases was not available in the Arizona data. The Arizona data contains comparatively more confirmed cases at the extremes of age.}

\label{figS8} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

%% -----------------------

%% Figure S9

\begin{figure}[ht]

\includegraphics[width=\textwidth]{FigS8-birth-yr-est-table.pdf}

\caption{\color{Gray} \textbf{Rationale for birth year estimation, INSIGHT data.} \textbf{A}, This worked example illustrates that the month of case observation determines potential birth dates, and cases observed earlier in the year have a higher probability of a birthday falling in the previous year. Because NH season cases are observed late in year y0, or early in year y1, an individual of age 0, observed during the NH 2000-2001 season would most likely have been born in 2000, with low probabilities of birth in 1999 or in 2001. However, an individual of age 0 observed in the SH season has an equal probability of birth in 2000 or 2001. \textbf{B}, the patterns shown in \textbf{A}, were generalized to develop a scheme to calculate the weighted average birth year for individuals of any age, whose cases were observed in a NH season spanning years y0-y1, or in a SH season during year y1.}

\label{figS9} % \label works only AFTER \caption within figure environment

\end{figure}

%% -----------------------

\newpage

\subsection\*{Tables}

%% -----------------------

%% Table S1

\begin{table}[!ht]

\begin{adjustwidth}{0in}{0in} % comment out/remove adjustwidth environment if table fits in text column.

\centering

\caption{{\bf Case counts from Arizona data}.}

\begin{tabular}{l l l}

\textbf{Season} & \textbf{H1N1 Count} & \textbf{H3N2 Count} \\

\hline

1993-1994 & 101 & 0 \\

1994-1995 & 38 & 12 \\

2002-2003 & 8 & 71 \\

2003-2004 & 71 & 0 \\

2004-2005 & 131 & 0 \\

2005-2006 & 321 & 1 \\

2006-2007 & 28 &212 \\

2007-2008 & 244 &196 \\

2008-2009 & 382 &2717 \\

2009-2010 & 1 &6261 \\

2010-2011 & 1204 & 472 \\

2011-2012 & 348 &595 \\

2012-2013 & 1578 & 80\\

2013-2014 & 151 &1475\\

2014-20915& 2109 & 5\\

\end{tabular}

\label{tabS1}

\end{adjustwidth}

\end{table}

% --------------------------

%% -----------------------

%% Table S2

\vspace\*{-2.5in}\begin{sidewaystable}[!ht]

\small

\centering

\caption{{\bf Case counts from INSIGHT data}.}

\begin{tabular}{lll|lll|lll|lll|lll}

Country & H1N1 & H3N2 & Country & H1N1 & H3N2 & Country & H1N1 & H3N2 & Country & H1N1 & H3N2 & Country & H1N1 & H3N2 \\

\hline

NH '09-'10 & & & NH '12-'13 & & & NH '15-'16 & & & SH '11 & & & SH '14 & & \\

\hline

Austria & 9 & 0 & Argentina & 1 & 0 & Belgium & 91 & 3 & Argentina & 6 & 238 & Argentina & 1 & 303 \\

Belgium & 15 & 0 & Belgium & 88 & 6 & Denmark & 7 & 2 & Australia & 2 & 3 & Australia & 0 & 1 \\

Denmark & 17 & 0 & Denmark & 0 & 5 & Estonia & 17 & 1 & Chile & 17 & 2 & Belgium & 0 & 1 \\

Estonia & 7 & 0 & Estonia & 8 & 7 & Germany & 7 & 0 & & & & Peru & 3 & 21 \\

Germany & 23 & 0 & Germany & 6 & 6 & Greece & 52 & 0 & & & & Poland & 2 & 0 \\

Greece & 91 & 0 & Peru & 0 & 21 & Peru & 6 & 1 & & & & Thailand & 30 & 23 \\

Poland & 16 & 0 & Poland & 3 & 1 & Poland & 34 & 0 & & & & & & \\

Spain & 14 & 0 & Spain & 2 & 1 & Spain & 10 & 2 & & & & & & \\

Thailand & 4 & 0 & Thailand & 1 & 1 & Thailand & 17 & 69 & & & & & & \\

USA & 121 & 0 & UK & 1 & 1 & UK & 6 & 1 & & & & & & \\

& & & USA & 5 & 56 & USA & 24 & 1 & & & & & & \\

\hline

NH '10-'11 & & & NH '13-'14 & & & NH '16-'17 & & & SH '12 & & & SH '15 & & \\

\hline

Australia & 1 & 0 & Belgium & 26 & 43 & Belgium & 0 & 7 & Argentina & 20 & 34 & Argentina & 23 & 267 \\

Belgium & 5 & 0 & Denmark & 9 & 1 & Denmark & 0 & 2 & Australia & 0 & 6 & Australia & 0 & 2 \\

Denmark & 5 & 0 & Estonia & 8 & 1 & Estonia & 0 & 4 & Peru & 3 & 7 & Peru & 0 & 21 \\

Estonia & 29 & 0 & Germany & 3 & 4 & Peru & 0 & 1 & Thailand & 1 & 0 & Thailand & 13 & 95 \\

Germany & 12 & 0 & Japan & 2 & 0 & Poland & 0 & 6 & USA & 0 & 2 & UK & 0 & 1 \\

Greece & 14 & 0 & Poland & 1 & 26 & Thailand & 13 & 32 & & & & & & \\

Poland & 10 & 0 & Spain & 10 & 2 & USA & 0 & 8 & & & & & & \\

Thailand & 11 & 10 & Thailand & 111 & 61 & & & & & & & & & \\

UK & 1 & 0 & USA & 48 & 2 & & & & & & & & & \\

USA & 5 & 3 & & & & & & & & & & & & \\

& & & & & & & & & & & & & & \\

& & & & & & & & & & & & & & \\

\hline

NH '11-'12 & & & NH '14-'15 & & & SH '10 & & & SH '13 & & & SH '16 & & \\

\hline

Belgium & 0 & 103 & Belgium & 22 & 116 & Australia & 21 & 1 & Argentina & 207 & 52 & Argentina & 400 & 0 \\

Denmark & 0 & 2 & Denmark & 6 & 11 & Thailand & 101 & 26 & Australia & 2 & 0 & Australia & 1 & 3 \\

Estonia & 0 & 6 & Estonia & 1 & 9 & & & & Chile & 1 & 2 & Belgium & 1 & 0 \\

Germany & 0 & 1 & Germany & 1 & 15 & & & & Peru & 36 & 2 & Peru & 3 & 2 \\

Peru & 0 & 3 & Greece & 0 & 6 & & & & Thailand & 2 & 39 & Spain & 1 & 0 \\

Spain & 0 & 4 & Japan & 0 & 10 & & & & & & & Thailand & 25 & 85 \\

USA & 1 & 4 & Peru & 1 & 0 & & & & & & & USA & 1 & 0 \\

& & & Poland & 7 & 17 & & & & & & & & & \\

& & & Spain & 0 & 3 & & & & & & & & & \\

& & & Thailand & 9 & 79 & & & & & & & & & \\

& & & UK & 0 & 3 & & & & & & & & & \\

& & & USA & 0 & 51 & & & & & & & &. &. \\

\end{tabular}\vspace{1in}

\label{tabS2}

\end{sidewaystable}

%% -----------------------

%% Table S3

\begin{table}[!ht]

\small

\centering

\begin{adjustwidth}{-.5in}{0in}

\caption{{\bf Paramter estimates from fits to Arizona Data}. Values represent maximum likelihood estimates of relative risk, given membership in each risk group. 95\% profile CIs are given in parentheses.}

\begin{tabular}{l|l|l|l|l}

\hline

\hline

& \textbf{A} & \textbf{AN} & \textbf{AS} & \textbf{AG} \\

\hline

\hline

$\Delta AIC$ & 0 & 21.65 & 449.7 & 1131.95\\

\hline

H1N1 imp. prot. & & 0.28 (0.24-0.33) & 0.25 (0.22-0.3) & 0.63 (0.57-0.7) \\

H3N2 imp. prot. & & 0.59 (0.53-0.67) & 0.62 (0.56-0.71) & 0.44 (0.39-0.5) \\

Ages 5-10 & 0.87 (0.82-0.92) & 0.81 (0.77-0.86) & 0.8 (0.76-0.86) & 0.88 (0.84-0.94) \\

Ages 11-17 & 0.59 (0.56-0.63) & 0.56 (0.53-0.6) & 0.55 (0.52-0.6) & 0.62 (0.58-0.67) \\

Ages 18-24 & 0.49 (0.46-0.53) & 0.48 (0.46-0.52) & 0.47 (0.44-0.52) & 0.53 (0.5-0.57) \\

Ages 25-31 & 0.42 (0.39-0.45) & 0.42 (0.4-0.46) & 0.41 (0.39-0.45) & 0.44 (0.42-0.48) \\

Ages 32-38 & 0.28 (0.26-0.31) & 0.25 (0.24-0.28) & 0.25 (0.23-0.28) & 0.29 (0.26-0.32) \\

Ages 39-45 & 0.25 (0.24-0.28) & 0.21 (0.2-0.24) & 0.20 (0.19-0.23) & 0.27 (0.25-0.3) \\

Ages 46-52 & 0.27 (0.24-0.3) & 0.23 (0.21-0.26) & 0.20 (0.18-0.22) & 0.30 (0.28-0.33) \\

Ages 53-59 & 0.22 (0.2-0.24) & 0.24 (0.22-0.28) & 0.22 (0.2-0.24) & 0.24 (0.22-0.26) \\

Ages 60-66 & 0.16 (0.15-0.18) & 0.23 (0.2-0.26) & 0.23 (0.21-0.26) & 0.17 (0.16-0.2) \\

Ages 67-73 & 0.14 (0.13-0.16) & 0.20 (0.18-0.24) & 0.21 (0.19-0.24) & 0.15 (0.14-0.18) \\

Ages 74-80 & 0.13 (0.12-0.16) & 0.19 (0.18-0.22) & 0.20 (0.18-0.24) & 0.14 (0.13-0.17) \\

Ages 81+ & 0.09 (0.08-0.1) & 0.12 (0.11-0.14) & 0.13 (0.12-0.15) & 0.09 (0.08-0.1)

\end{tabular}

\label{tabS3}

\end{adjustwidth}

\end{table}

%% -----------------------

%% Table S4

\begin{sidewaystable}[!ht]

\begin{adjustwidth}{-.3in}{}

\scriptsize

\centering

\caption{{\bf Paramter estimates from fits to INSIGHT Data (1/2)}. Values represent maximum likelihood estimates of relative risk, given membership in each risk group. 95\% profile CIs are given in parentheses.}

\begin{tabular}{l|l| l| l| l| l| l| l| l| l}

\hline

\hline

Model & \textbf{ATVN} & \textbf{ATVS} & \textbf{ATUVN} & \textbf{ATV} & \textbf{ATUVS} & \textbf{ATUV} & \textbf{ATVG} & \textbf{ATUVG} & \textbf{AVN} \\

\hline

\hline

$\Delta AIC$ & 0 & 0.45 & 1.95 & 2.09 & 2.4 & 4.01 & 4.65 & 6.57 & 8.99 \\

\hline

Ant. Vir. Trt. & 1.36 (1.14-1.62) & 1.36 (1.14-1.62) & 1.36 (1.14-1.62) & 1.36 (1.14-1.62) & 1.36 (1.14-1.62) & 1.36 (1.14-1.62) & 1.36 (1.14-1.62) & 1.36 (1.14-1.62) & \\

Und. Conds. & & & 0.98 (0.85-1.14) & & 0.98 (0.85-1.14) & 0.98 (0.85-1.14) & & 0.98 (0.85-1.14) & \\

Vaccin. H1N1 & 0.64 (0.48-0.83) & 0.64 (0.48-0.84) & 0.64 (0.48-0.83) & 0.62 (0.47-0.81) & 0.64 (0.48-0.84) & 0.62 (0.47-0.81) & 0.63 (0.48-0.82) & 0.63 (0.48-0.83) & 0.65 (0.49-0.84) \\

Vaccin. H3N2 & 0.85 (0.67-1.06) & 0.84 (0.67-1.06) & 0.85 (0.68-1.07) & 0.86 (0.69-1.08) & 0.85 (0.67-1.06) & 0.87 (0.69-1.08) & 0.85 (0.67-1.06) & 0.85 (0.67-1.07) & 0.85 (0.68-1.07) \\

Impr. H1N1 & 0.84 (0.62-1.12) & 0.78 (0.61-1.00) & 0.84 (0.62-1.12) & & 0.78 (0.61-1.00) & & 0.98 (0.79-1.24) & 0.98 (0.79-1.23) & 0.83 (0.62-1.12) \\

Impr. H3N2 & 0.90 (0.69-1.19) & 0.96 (0.81-1.16) & 0.89 (0.69-1.19) & & 0.96 (0.81-1.16) & & 0.92 (0.77-1.16) & 0.92 (0.77-1.16) & 0.90 (0.70-1.20) \\

Ages 18-24 & 0.97 (0.86-1.10) & 0.97 (0.86-1.09) & 0.97 (0.86-1.10) & 0.97 (0.86-1.09) & 0.97 (0.86-1.09) & 0.97 (0.86-1.09) & 0.97 (0.86-1.10) & 0.97 (0.86-1.10) & 0.97 (0.86-1.09) \\

Ages 25-31 & 1.04 (0.94-1.16) & 1.05 (0.94-1.17) & 1.04 (0.94-1.16) & 1.04 (0.94-1.16) & 1.04 (0.94-1.17) & 1.04 (0.94-1.16) & 1.04 (0.94-1.16) & 1.03 (0.94-1.16) & 1.04 (0.94-1.17) \\

Ages 39-45 & 0.99 (0.87-1.13) & 0.97 (0.86-1.09) & 0.99 (0.87-1.13) & 0.99 (0.89-1.11) & 0.97 (0.86-1.09) & 0.99 (0.89-1.11) & 1.00 (0.89-1.13) & 1.00 (0.89-1.13) & 1.00 (0.88-1.14) \\

Ages 46-52 & 0.95 (0.83-1.09) & 0.92 (0.81-1.04) & 0.95 (0.83-1.10) & 0.95 (0.85-1.07) & 0.92 (0.81-1.04) & 0.95 (0.85-1.08) & 0.94 (0.83-1.08) & 0.94 (0.83-1.09) & 0.95 (0.83-1.09) \\

Ages 53-59 & 0.95 (0.83-1.08) & 0.92 (0.81-1.06) & 0.95 (0.83-1.09) & 0.95 (0.83-1.08) & 0.93 (0.81-1.07) & 0.95 (0.84-1.08) & 0.93 (0.80-1.11) & 0.93 (0.80-1.12) & 0.95 (0.84-1.09) \\

Ages 60-66 & 0.96 (0.79-1.18) & 0.99 (0.82-1.19) & 0.96 (0.79-1.18) & 0.95 (0.82-1.11) & 0.99 (0.82-1.19) & 0.96 (0.82-1.12) & 0.93 (0.78-1.14) & 0.93 (0.78-1.15) & 0.97 (0.80-1.19) \\

Ages 67-73 & 1.01 (0.78-1.31) & 1.06 (0.84-1.33) & 1.02 (0.78-1.32) & 1.01 (0.82-1.23) & 1.07 (0.84-1.34) & 1.02 (0.83-1.25) & 0.98 (0.79-1.25) & 0.99 (0.79-1.26) & 1.02 (0.79-1.32) \\

Ages 74-80 & 1.00 (0.71-1.41) & 1.05 (0.76-1.44) & 1.01 (0.71-1.43) & 1.00 (0.74-1.33) & 1.06 (0.77-1.45) & 1.01 (0.74-1.35) & 0.97 (0.70-1.34) & 0.98 (0.71-1.36) & 1.04 (0.73-1.47) \\

Ages 81-90 & 0.90 (0.55-1.45) & 0.95 (0.58-1.48) & 0.91 (0.55-1.46) & 0.90 (0.56-1.38) & 0.96 (0.58-1.50) & 0.91 (0.56-1.40) & 0.88 (0.54-1.38) & 0.88 (0.54-1.39) & 0.96 (0.58-1.53) \\

\hline

\hline

Model & \textbf{AVS} & \textbf{ATN} & \textbf{ATS} & \textbf{AV} & \textbf{AUVN} & \textbf{AUVS} & \textbf{ATUN} & \textbf{ATUS} & \textbf{AUV} \\

\hline \hline

$\Delta AIC$ & 9.37 & 10.35 & 10.55 & 10.84 & 10.98 & 11.37 & 12.06 & 12.27 & 12.82 \\

\hline

Ant. Vir. Trt. & & 1.34 (1.12-1.60) & 1.34 (1.12-1.60) & & & & 1.35 (1.12-1.61) & 1.35 (1.12-1.60) & \\

Und. Cond. & & & & & 0.99 (0.86-1.15) & 0.99 (0.86-1.15) & 0.96 (0.83-1.11) & 0.96 (0.83-1.11) & 0.99 (0.86-1.15) \\

Vaccin. H1N1 & 0.65 (0.49-0.85) & & & 0.63 (0.47-0.81) & 0.65 (0.49-0.84) & 0.65 (0.49-0.85) & & & 0.63 (0.47-0.82) \\

Vaccin. H3N2 & 0.85 (0.67-1.06) & & & 0.87 (0.69-1.08) & 0.85 (0.68-1.07) & 0.85 (0.67-1.06) & & & 0.87 (0.69-1.09) \\

Impr. H1N1 & 0.78 (0.61-1.00) & 0.81 (0.59-1.09) & 0.76 (0.59-0.97) & & 0.83 (0.62-1.12) & 0.78 (0.61-1.00) & 0.81 (0.59-1.09) & 0.76 (0.59-0.97) & \\

Impr. H3N2 & 0.96 (0.81-1.16) & 0.89 (0.69-1.19) & 0.94 (0.79-1.13) & & 0.90 (0.70-1.20) & 0.96 (0.81-1.16) & 0.89 (0.69-1.19) & 0.94 (0.79-1.13) & \\

Ages 18-24 & 0.96 (0.86-1.09) & 1.00 (0.89-1.13) & 1.00 (0.89-1.12) & 0.97 (0.86-1.09) & 0.97 (0.86-1.09) & 0.96 (0.86-1.09) & 0.99 (0.89-1.12) & 0.99 (0.89-1.12) & 0.96 (0.86-1.09) \\

Ages 25-31 & 1.05 (0.94-1.17) & 1.06 (0.95-1.18) & 1.06 (0.96-1.19) & 1.04 (0.94-1.16) & 1.04 (0.94-1.17) & 1.05 (0.94-1.17) & 1.06 (0.95-1.18) & 1.06 (0.96-1.18) & 1.04 (0.94-1.16) \\

Ages 39-45 & 0.98 (0.88-1.10) & 0.98 (0.86-1.11) & 0.96 (0.86-1.09) & 1.00 (0.90-1.12) & 1.00 (0.88-1.14) & 0.98 (0.88-1.11) & 0.98 (0.86-1.12) & 0.97 (0.86-1.09) & 1.00 (0.90-1.12) \\

Ages 46-52 & 0.92 (0.81-1.04) & 0.93 (0.82-1.08) & 0.90 (0.80-1.02) & 0.95 (0.85-1.07) & 0.95 (0.83-1.09) & 0.92 (0.81-1.04) & 0.94 (0.82-1.08) & 0.91 (0.81-1.03) & 0.95 (0.85-1.08) \\

Ages 53-59 & 0.93 (0.81-1.07) & 0.93 (0.82-1.06) & 0.90 (0.79-1.04) & 0.95 (0.84-1.08) & 0.95 (0.84-1.09) & 0.93 (0.81-1.08) & 0.94 (0.82-1.07) & 0.91 (0.79-1.05) & 0.96 (0.84-1.09) \\

Ages 60-66 & 1.00 (0.83-1.20) & 0.93 (0.77-1.14) & 0.95 (0.80-1.15) & 0.96 (0.82-1.12) & 0.97 (0.79-1.19) & 1.00 (0.83-1.20) & 0.94 (0.77-1.16) & 0.96 (0.80-1.16) & 0.96 (0.82-1.13) \\

Ages 67-73 & 1.07 (0.85-1.34) & 0.97 (0.75-1.26) & 1.01 (0.81-1.27) & 1.02 (0.83-1.24) & 1.02 (0.79-1.33) & 1.07 (0.85-1.35) & 0.98 (0.76-1.28) & 1.02 (0.81-1.29) & 1.02 (0.83-1.25) \\

Ages 74-80 & 1.09 (0.79-1.49) & 0.97 (0.69-1.36) & 1.01 (0.73-1.38) & 1.04 (0.77-1.38) & 1.05 (0.73-1.48) & 1.09 (0.79-1.50) & 0.98 (0.69-1.39) & 1.02 (0.74-1.41) & 1.04 (0.77-1.40) \\

Ages 81-90 & 1.01 (0.61-1.56) & 0.87 (0.52-1.38) & 0.90 (0.55-1.40) & 0.95 (0.59-1.46) & 0.96 (0.58-1.54) & 1.01 (0.61-1.57) & 0.88 (0.53-1.41) & 0.91 (0.56-1.43) & 0.95 (0.59-1.47) \\

\end{tabular}

\label{tabS4}

\end{adjustwidth}

\end{sidewaystable}

\begin{sidewaystable}[!ht]

\begin{adjustwidth}{-.3in}{}

\scriptsize

\centering

\caption{{\bf Paramter estimates from fits to INSIGHT Data (2/2)}. Values represent maximum likelihood estimates of relative risk, given membership in each risk group. 95\% profile CIs are given in parentheses.}

\begin{tabular}{l|l| l| l| l| l| l| l| l| l}

\hline \hline

Model & AVG & AT & AUVG & ATG & ATU & ATUG & AN & AS & AUN \\

\hline \hline

$\Delta AIC$ & 13.55 & 14.19 & 15.54 & 15.57 & 15.81 & 17.2 & 18.48 & 18.66 & 20.33 \\

\hline

Ant. Vir. Trtmt & & 1.34 (1.12-1.59) & & 1.35 (1.12-1.60) & 1.34 (1.12-1.60) & 1.35 (1.12-1.61) & & & \\

Und. Conds. & & & 0.99 (0.86-1.15) & & 0.96 (0.83-1.11) & 0.96 (0.83-1.11) & & & 0.97 (0.84-1.12) \\

Vaccin. H1N1 & 0.64 (0.48-0.83) & & 0.64 (0.48-0.83) & & & & & & \\

Vaccin. H3N2 & 0.85 (0.68-1.07) & & 0.85 (0.68-1.07) & & & & & & \\

Impr. H1N1 & 1.00 (0.80-1.25) & & 1.00 (0.80-1.25) & 0.95 (0.77-1.20) & & 0.95 (0.77-1.20) & 0.80 (0.59-1.08) & 0.76 (0.59-0.97) & 0.81 (0.59-1.09) \\

Impr. H3N2 & 0.91 (0.77-1.15) & & 0.91 (0.77-1.15) & 0.91 (0.75-1.14) & & 0.91 (0.75-1.15) & 0.90 (0.70-1.20) & 0.95 (0.80-1.14) & 0.90 (0.70-1.20) \\

Ages 18-24 & 0.97 (0.86-1.09) & 1.00 (0.89-1.12) & 0.97 (0.86-1.09) & 1.00 (0.90-1.13) & 0.99 (0.89-1.12) & 0.99 (0.89-1.12) & 1.00 (0.89-1.12) & 0.99 (0.89-1.12) & 0.99 (0.89-1.12) \\

Ages 25-31 & 1.04 (0.94-1.16) & 1.05 (0.95-1.18) & 1.03 (0.94-1.16) & 1.05 (0.95-1.18) & 1.05 (0.95-1.17) & 1.05 (0.95-1.17) & 1.06 (0.95-1.18) & 1.06 (0.96-1.19) & 1.06 (0.95-1.18) \\

Ages 39-45 & 1.01 (0.91-1.14) & 0.98 (0.89-1.10) & 1.01 (0.91-1.14) & 0.99 (0.89-1.12) & 0.99 (0.89-1.11) & 1.00 (0.89-1.12) & 0.99 (0.87-1.12) & 0.97 (0.87-1.10) & 0.99 (0.87-1.13) \\

Ages 46-52 & 0.94 (0.83-1.08) & 0.94 (0.84-1.06) & 0.94 (0.83-1.08) & 0.93 (0.82-1.08) & 0.95 (0.84-1.07) & 0.94 (0.82-1.08) & 0.93 (0.82-1.08) & 0.90 (0.81-1.02) & 0.94 (0.82-1.08) \\

Ages 53-59 & 0.92 (0.81-1.11) & 0.93 (0.82-1.06) & 0.92 (0.80-1.11) & 0.91 (0.78-1.10) & 0.94 (0.82-1.07) & 0.92 (0.78-1.11) & 0.94 (0.82-1.07) & 0.91 (0.80-1.05) & 0.94 (0.82-1.08) \\

Ages 60-66 & 0.93 (0.79-1.14) & 0.92 (0.79-1.07) & 0.93 (0.79-1.15) & 0.90 (0.74-1.10) & 0.93 (0.80-1.09) & 0.91 (0.75-1.12) & 0.94 (0.77-1.16) & 0.96 (0.81-1.16) & 0.95 (0.78-1.17) \\

Ages 67-73 & 0.98 (0.80-1.25) & 0.96 (0.79-1.17) & 0.98 (0.79-1.26) & 0.94 (0.74-1.20) & 0.97 (0.80-1.19) & 0.95 (0.75-1.22) & 0.98 (0.76-1.27) & 1.02 (0.81-1.28) & 0.99 (0.77-1.29) \\

Ages 74-80 & 1.00 (0.73-1.38) & 0.96 (0.70-1.27) & 1.00 (0.73-1.40) & 0.94 (0.68-1.29) & 0.97 (0.71-1.30) & 0.95 (0.68-1.32) & 1.01 (0.71-1.42) & 1.05 (0.76-1.43) & 1.02 (0.71-1.44) \\

Ages 81-90 & 0.92 (0.56-1.44) & 0.85 (0.53-1.29) & 0.92 (0.56-1.45) & 0.83 (0.51-1.30) & 0.86 (0.53-1.32) & 0.84 (0.51-1.33) & 0.92 (0.55-1.46) & 0.95 (0.58-1.48) & 0.93 (0.56-1.48) \\

\hline \hline

Model & AUS & A & AG & AU & AUG & & & & \\

\hline \hline

$\Delta AIC$ & 20.51 & 21.99 & 23.69 & 23.77 & 25.48 & & & & \\

\hline

Ant. Vir. Trtmt & & & & & & & & & \\

Und. Conds. & 0.97 (0.84-1.12) & & & 0.97 (0.84-1.12) & 0.97 (0.84-1.12) & & & & \\

Vaccin. H1N1 & & & & & & & & & \\

Vaccin. H3N2 & & & & & & & & & \\

Impr. H1N1 & 0.76 (0.59-0.97) & & 0.97 (0.78-1.23) & & 0.97 (0.78-1.22) & & & & \\

Impr. H3N2 & 0.95 (0.80-1.14) & & 0.90 (0.75-1.14) & & 0.90 (0.75-1.14) & & & & \\

Ages 18-24 & 0.99 (0.88-1.12) & 0.99 (0.89-1.12) & 1.00 (0.89-1.12) & 0.99 (0.89-1.12) & 0.99 (0.89-1.12) & & & & \\

Ages 25-31 & 1.06 (0.96-1.18) & 1.05 (0.95-1.18) & 1.05 (0.95-1.18) & 1.05 (0.95-1.17) & 1.05 (0.95-1.17) & & & & \\

Ages 39-45 & 0.97 (0.87-1.10) & 0.99 (0.90-1.11) & 1.01 (0.90-1.14) & 1.00 (0.90-1.12) & 1.01 (0.90-1.14) & & & & \\

Ages 46-52 & 0.91 (0.81-1.03) & 0.94 (0.84-1.06) & 0.93 (0.82-1.07) & 0.95 (0.84-1.07) & 0.94 (0.82-1.08) & & & & \\

Ages 53-59 & 0.91 (0.80-1.06) & 0.94 (0.82-1.06) & 0.91 (0.79-1.10) & 0.94 (0.83-1.07) & 0.92 (0.79-1.11) & & & & \\

Ages 60-66 & 0.97 (0.81-1.17) & 0.93 (0.80-1.08) & 0.90 (0.75-1.11) & 0.93 (0.80-1.10) & 0.91 (0.76-1.12) & & & & \\

Ages 67-73 & 1.03 (0.81-1.30) & 0.97 (0.80-1.18) & 0.94 (0.75-1.20) & 0.98 (0.80-1.20) & 0.95 (0.76-1.22) & & & & \\

Ages 74-80 & 1.06 (0.76-1.45) & 0.99 (0.73-1.32) & 0.96 (0.69-1.32) & 1.00 (0.73-1.34) & 0.97 (0.70-1.35) & & & & \\

Ages 81-90 & 0.96 (0.58-1.50) & 0.89 (0.56-1.36) & 0.87 (0.53-1.36) & 0.91 (0.56-1.39) & 0.88 (0.54-1.39) & & & &

\end{tabular}\vspace\*{2in}

\label{tabS5}

\end{adjustwidth}

\end{sidewaystable}

\newpage

\section\*{Acknowledgments}

We thank just about everybody.

\nolinenumbers

%This is where your bibliography is generated. Make sure that your .bib file is actually called library.bib

\bibliography{2019Seasonal\_flu\_manuscript}

%This defines the bibliographies style. Search online for a list of available styles.

\bibliographystyle{abbrv}

\begin{comment}

Influenza epidemics and pandemics can show anomalous age-specific impacts, with disease burden episodically elevated in particular birth cohorts [CITE EXAMPLES]. These patterns are usually attributed to birth year-specific differences in immune history. In a phenomenon alternatively called original antigenic sin, antigenic seniority or immunological imprinting, individuals develop particularly strong, lifelong immune memory against influenza viruses antigenically similar to those encountered in childhood [CITE]. As a result, major antigenic transitions can cause generational shifts in immune protection, with cohorts imprinted in different antigenic eras showing dramatic differences in pre-existing immunity against a given influenza strain [CITE EXAMPLES].

The ability to predict which birth years will experience the highest burden of disease in future influenza epidemics or pandemics would provide myriad advantages for epidemic planning, preparedness and response, and could even pave the way toward individualized, or cohort-specific vaccination programs. In principle, birth year-specific risk should be predictable, as long as the evolutionary history of relevant antigenic targets is known. As a proof of concept, we recently generated surprisingly accurate predictions for the distributions of avian H5N1 and H7N9 cases across human birth years, based solely on population age structure, and reconstructed, birth year-specific childhood imprinting patterns [CITE].

However, in practice it is rarely obvious \textit{a priori} which antigenic sites will be the dominant immune targets in a particular epidemic context, and thus, it is difficult to choose which aspects of imprinting history are relevant for prediction of birth year-specific risk. For example, it was not previously clear whether childhood imprinting to neuraminidase (NA) or to conserved hemagglutinin (HA) epitopes would be the strongest driver of birth year-specific avian influenza risk, although formal statistical analysis now definitively supports HA-specific imprinting as the dominant effect.

%Another open question is whether population-level immune memory reliably targets the same few antigenic sites, or whether instead, population-level immunity arises as a mixture of diverse, individual-level memory responses. Epidemic patterns are more likely to be predictable based on imprinting history if the site-specificity of immune memory is somewhat synchronized at the population-level.

An even greater number of antigenic sites may play strong roles in protection against seasonal influenza subtypes H1N1 and H3N2, which have circulated in humans for years and are highly familiar to our immune systems. Specifically, immune memory of conserved and variable parts of HA, NA, and of epitopes on influenza's internal proteins may all play a role in the overall immune response. These various antigenic sites drift at different rates, and have shifted or reassorted at different historical timepoints. It is probably not tractable to model how imprinting to each of influenza's various eptiopes contributes to overall, birth year-specific protection against seasonal influenza. But if population-level immunity consistently targets the same antigens, then a simpler, tractable model may be sufficient. Here, we aimed to test whether any single kind of childhood imprinting remains a strong predictor of seasonal influenza risk in the long run.

% Furthermore, the relative strength of immunity against each antigenic site may vary from season to season, with immunodominant sites on HA most likely playing the strongest role upon exposure to familiar, weakly drifted seasonal influenza variant, but with various other, more conserved antigenic sites potentially playing a stronger role in the immune response against highly drifted, or newly glycosylated strains. Together, these complexities make it difficult to predict which antigenic facets of childhood immune history are relevant for prediction of birth year-specific influenza risk in any single influenza season, or whether any specific facet of childhood imprinting remains a strong predictor in the long run. , so histories of childhood imprinting to specific influenza antigens are out of sync.

Numerous studies have associated childhood imprinting with observed differences in birth year-specific risk from seasonal influenza A/H1N1 and A/H3N2. For example, seasonal H1N1 often has relatively mild impacts in older cohorts, who would have imprinted to historical H1N1 variants in childhood [CITE CITE CITE], and may now enjoy preferential protection against modern influenza variants of the same subtype. These same older cohorts generally experience relatively high incidence and mortality from H3N2 influenza [CITE], and greater H3N2 impacts in these older cohorts may be due to a lack of imprinting protection against H3N2.

While many studies have hypothesized that imprinting may shape cohort-specific risk from H1N1 and H3N2, existing studies have not yet proposed or tested any specific antigenic explanation for ostensible cohort effects observed in epidemiological data. Furthermore, many existing studies that note potential cohort effects on H1N1 have focused on the 2009 pandemic. Meanwhile, many studies of non-pandemic seasons have analyzed data from a limited time period, or from a limited geographic range.

Here, we performed a systematic analysis of two large, epidemiological data sets that each tracked age distributions of confirmed, seasonal influenza A cases across 21 years, and 9 years, respectively. Within these data, we compared age distributions of confirmed H1N1 and H3N2 cases to search for consistent birth year-specific differences in risk from each seasonal subtype. This analysis aimed to determine whether cohort effects are largely ephemeral, with birth year-specific risk periodically magnified by season-specific antigenic escape events, or instead whether cohort effects persist over time as predictable drivers of birth year-specific risk from seasonal H1N1 and H3N2.

Previously noted cohort effects are consistent across time and space, in that the relative impact of H3N2 tends to be greater than the impact of H1N1 in older cohorts. However, the magnitude of observed birth year-specific differences in risk fluctuated over time, differed across countries, and between data sets. We tested whether birth year-specific imprinting at the HA subtype level (driven by immune memory of variable HA epitopes), at the HA group level (driven by immune memory of conserved HA epitopes), or at the NA subtype level, could best explain observed birth year-specific differences in risk from H1N1 and H3N2. Model comparison (AIC) showed more consistent support for effects of imprintign at the NA subtype or HA subtype level than at the HA group level.

\end{comment}

\end{document}