Outline of manuscript:

—> We often, but not always find a negative relationship between imprinting protection and risk.

—> But this relationship is not universal. E.g. positive relationship between imprinting protection and ICU risk.

—> Neither AIC nor cross-validation (AUC) points to inclusion of imprinting in the best model. While we can’t rule out some imprinting effects,

they don’t seem to be strong predictors of incidence or severity outcomes.

🡪 Vaccination does seem to have an effect.

🡪 NEED MORE SYSTEMATIC DATA COLLECTION. SEVERITY, SINGLE YEAR OF AGE.

🡪 NEED THIS DATA OVER A LONG TIME PERIOD.

🡪 NIH STUFF

TO DO:

🡪 Verify GAM methods

🡪 Add results of glycan analysis into model formulation and double-check.

🡪 Draft manuscript

**Methods**

**Data**

Our data included ## cases from FLU 002 and ## cases from FLU 003…

**Estimation of imprinting probabilities**

We estimated the probability that each case in the data imprinted to an H1N1, H2N2 or H3N2 virus during childhood. Imprinting probabilities were reconstructed using previously described methods (CITE), except as noted below. To summarize briefly, we first estimated a case’s birth year based on their age in the year their case was observed. Then, we estimated the probability that the cases’s first childhood influenza exposure occurred 0, 1, 2, etc. years after the birth year. Finally, we cross-referenced these estimates with historical data on the influenza subtypes circulating during each possible year of first, childhood influenza exposure. Thus, we estimated the overall probability of first exposure to H1N1, H2N2 or H3N2.

The only notable modification to previously described methods (CITE), is that previously we estimated birth year as *y-a*, where *y* represents the year of case observation and *a* represents age. We have now updated the methods to better address uncertainty inherent to this approach*.* For example, consider an infant of age 0 observed in May 2000. Using our old method, this case would have been assigned a birth year of *y-a*=2000. However, in reality, this individual could have been born any time between late May 1999 and early May, 2000. This implies 7.5 possible birth months in 1999 [*y-a-1*], and only 4.5 possible birth months in 2000 [*y-a*].

We have now generalized the above argument to estimate the probability that each case was born in year *y-a-1*, or in year *y-a* (Table ##). Then, we reconstructed imprinting probabilities for each of two possible birth years, and took a weighted average as appropriate to the month of case observation.

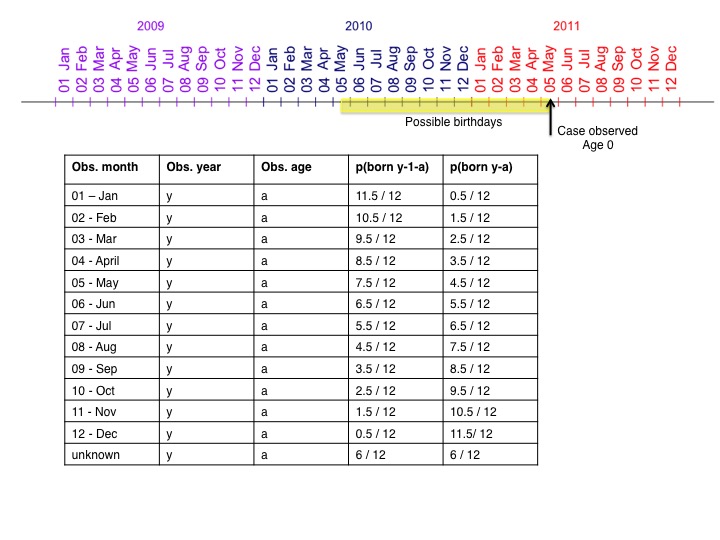


Figure xx or Table xx draft.

For input into each model, we ultimately wished to calculate the probability that each case was protected by their childhood imprinting against the influenza challenge that they most likely encountered, given the season and country in which their case was observed. To compile this single, imprinting input vector, we used the following protocol:

* If a case has confirmed H1N1 infection, we entered the probability ofimprinting protection against an H1N1 challenge.
* If a case had confirmed H3N2 infection, we entered the probability of imprinting protection against an H3N2 challenge.
* For cases who did not have any confirmed influenza infection, we used epidemiological surveillance data from WHO FluNet to determine the probability that the individual would have been challenged by H1N1 or by H3N2 in the season and country of interest. Here, we calculated the H1N1 fraction as the total number of confirmed H1N1 cases reported in the season and country of interest, divided by the total number of confirmed H1N1 and H3N2 cases reported in the season and country of interest. We defined the northern hemisphere season as spanning October 1 – March 31 (weeks 40-13), and the southern hemisphere season as spanning April 1 – Sept 30 (weeks 14-39). Finally, we assumed each case was challenged by H1N1 with a probability equal to the calculated H1N1 fraction (in which case we input the probability of imprinting protection against H1N1), or that the case was challenged by H3N2 with a probability equal to the calculated H3N2 fraction (in which case we input the probability of imprinting protection against H3N2).

**Methods-ish:**

One major challenge in this analysis is to detangle the effects of age and immune history. For example, H3N2 has a particularly strong impact in elderly cohorts. This pattern may arise either because elderly individuals are intrinsically at higher risk of severe influenza infection, or because childhood immune histories in elderly cohorts leave this group unprotected against H3N2.

In an attempt to separate the effects of childhood imprinting from age, we fit models to H1N1 and H3N2 data simultaneously. This approach allowed us to estimate a single age coefficient, which can be interpreted as describing the effects of age on the risk of infection with *any* influenza challenge. We also fit a single imprinting coefficient, which can be interpreted as the overall effect of protective childhood imprinting given exposure to any challenge within the same genetic group or subtype.

For example, factors like immunosenescence [CITE], underlying health status and social mixing patterns [CITE] all vary with age, and should influence the risk of influenza infection or severe outcome in similar ways, regardless of whether an individual is challenged by an H1N1 or H3N2 virus.

Meanwhile, we also estimated a single coefficient to describe the effect of protective imprinting status on infection risk. One limitation is that the strength of childhood imprinting protection against H1N1 may be different than the strength of protection against H3N2, whereas our approach is only able to estimate a single, overall protective coefficient.