

Reviewed Preprint

v1 • February 9, 2026

Not revised

✉ For correspondence:

bgarciacarreras@gmail.comderek.cummings@jhu.eduAuthor notes: See [page 13](#)

Reviewing editor: JT McCrone, Fred Hutch Cancer Center, United States

© 2026, García-Carreras et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Assessing evidence of immune imprinting in serological patterns of influenza immune response

Bernardo García-Carreras^{1,2}✉, Bingyi Yang³, James A Hay⁴, C Jessica E Metcalf⁵, Matt DT Hitchens², Claire P Smith^{6,7}, Siyu Chen⁵, Jonathan M Read⁸, Huachen Zhu⁹, Chaoqiang Jiang¹⁰, Kin On Kwok^{11,12}, Steven Riley¹³, Sarah Cobey¹⁴, Justin Lessler^{6,7,15}, Derek AT Cummings^{1,2,16}✉

¹Department of Biology, University of Florida, Gainesville, United States • ²Emerging Pathogens Institute, University of Florida, Gainesville, United States • ³WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China • ⁴Pandemic Sciences Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom • ⁵Department of Ecology and Evolutionary Biology, Princeton University, Princeton, United States • ⁶Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, United States • ⁷Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, United States • ⁸Lancaster Medical School, Lancaster University, Lancaster, United Kingdom • ⁹School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China • ¹⁰Guangzhou No. 13 Hospital, Guangzhou, China • ¹¹The Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China • ¹²Hong Kong Institute of Asia-Pacific Studies, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China • ¹³United Kingdom Health Security Agency, London, United Kingdom • ¹⁴Department of Ecology and Evolution, University of Chicago, Chicago, United States • ¹⁵Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, United States • ¹⁶Department of Biomedical Engineering, Johns Hopkins Whiting School of Engineering, Baltimore, United States

eLife Assessment

This **valuable** study highlights the challenge of identifying the role of immune imprinting in influenza immunity. The manuscript provides **solid** evidence that statistical support for imprinting depends heavily on model choice and can be found in the absence of imprinting due to age-related processes. However, the results are **incomplete** in that the impact of incorrectly modeling imprinting is not clear. The work will be of interest to researchers who study adaptive immunity in any system where imprinting may be observed.

<https://doi.org/10.7554/eLife.109291.1.sa2>

Abstract

The subtype of the first influenza infection shapes infection risk, illness and mortality associated with subsequent infection subtypes, as well as vaccine efficacy. This phenomenon, termed ‘immune imprinting’, may result from immune recognition rooted in structural similarity of the hemagglutinin protein within subtypes of influenza viruses. However, it is difficult to isolate from cohort effects resulting from the annual dominance of specific influenza subtypes. We used a long-term cohort study in southern China to examine whether hemagglutination inhibition responses show evidence of imprinting, and explore whether such patterns could emerge in the absence of imprinting using simulations. Our statistical analysis reveals patterns consistent with immune imprinting. However, simulations showed that similar patterns can emerge without an imprinting

mechanism, meaning these may partly be due to age-related variations in immune responses driven by other factors. This work lays the foundations for further research into immune imprinting while accounting for cohort effects.

Introduction

The burden of influenza on public health is a challenge compounded by the limited efficacy of influenza vaccines. Vaccine efficacies are low in part because of barriers posed by pre-existing immunity resulting from prior vaccination or infection. Pre-existing immunity can blunt or amplify responses based upon the correspondence of antigens in vaccines with prior exposures (Reber et al., 2015 [↗](#); Moritzky et al., 2023 [↗](#)). Annual cycles in different influenza subtypes mean that this can potentially also lead to cohort effects. Immune imprinting represents a specific example of a cohort effect, wherein exposure to specific influenza strains during childhood influence the immune system's memory, impacting future responses to strains based on their similarity to those early exposures. This phenomenon has been shown to drive age-specific patterns of illness and death, as evidenced by the fact that H3N2 underlies a greater proportion of reported cases in older individuals than H1N1 (Oidtman et al., 2021 [↗](#)), and by models fit to data (Gostic et al., 2016 [↗](#); Gagnon et al., 2018 [↗](#); Gostic et al., 2019 [↗](#); Arevalo et al., 2020b [↗](#)). Immune imprinting may pose a substantial challenge to vaccination strategies (Reber et al., 2015 [↗](#); Moritzky et al., 2023 [↗](#)), as the immune system's "memory" may either enhance or hinder the effectiveness of subsequent influenza vaccines, depending on the degree of similarity between the circulating strains and those encountered earlier in life. Understanding the intricacies and mechanisms of immune imprinting is crucial for refining vaccination approaches and addressing the complex dynamics of pre-existing immunity in ongoing efforts against influenza.

A suggested mechanism for imprinting is that the similarity of the hemagglutinin (HA) protein within groups of influenza viruses drives correlated serological responses across subtypes. HA subtypes are clustered into two structural groups, with H1, H2, and H5 among others, in one, and H3, H7, and H9, among others, in another (Gostic et al., 2016 [↗](#)). The imprinting hypothesis posits that early exposure to a virus from one group will enhance responses later in life to viruses from the same group relative to viruses of the other group.

One important and straightforward line of evidence for evaluating how imprinting shapes protection from infection from different subtypes would be to evaluate the cross-subtype serological profile of individuals with imprinting resulting from different subtypes. Antibody titers as measured by the hemagglutination inhibition (HI) assay have been shown to correlate with protection (Hobson et al., 1972 [↗](#)), and have previously been used in studies on antigenic seniority (Lessler et al., 2012 [↗](#)). To date, however, there has only been limited investigation of imprinting leveraging serology. Ranjeva et al. (2019 [↗](#)) fit an individual-level model to longitudinal serological data. They found limited evidence of imprinting, which in their study was defined as inferred protection from infection, as opposed to differences in immune response. Edler et al. (2024 [↗](#)) observed biases in immune responses to influenza B viruses, with titers tending to be higher against viruses from the same lineage as those encountered early in life. Other studies explored related questions, but did not explicitly test the immune imprinting hypothesis. Miller et al. (2013 [↗](#)), for instance, conducted a study following 40 participants over 20 years, and provided additional support for antigenic seniority, also reported by Lessler et al. (2012 [↗](#)) using a different cohort. However, all participants in the Miller *et al.* study were born prior to 1953, and thus likely exclusively exposed to one subtype (H1N1) early in life, so that they could not directly test the immune imprinting hypothesis.

Here, we examine immune profiles from a long-term cohort study in southern China to determine if patterns of hemagglutination inhibition (HI) responses show evidence of imprinting between subtypes. Specifically, we test the hypothesis that individuals imprinted by a virus from one group (e.g., H1N1 or H2N2) will lead to systematically higher immune responses to viruses from the same group, compared to those imprinted by a virus from another group (i.e., H3N2).

Results

HI titers were measured for 21 A(H3N2) and three A(H1N1) strains. We fit generalized additive models (GAMs) to log-transformed HI titers, aiming to identify any systematic variation consistent with the imprinting hypothesis, after adjusting for factors known to affect immune responses. The main challenge was to determine whether imprinting had an effect, independent of the known variation in titers with age (whether at the time of sampling, at the time of strain isolation, or both). We fit models that accounted for age in several ways: no age included, a spline on age at sampling, a spline on age at isolation, or both. The key “imprinting variable” represents the first influenza subtype exposure, with models fit both with and without it. Two approaches were used to assign the imprinting subtype: one “non-probabilistic”, based solely on birth year and which implicitly assumes the earliest exposures happen during the first year of life, and another “probabilistic”, allowing for later (beyond the first year) exposure to new subtypes (see Methods and Materials).

We also used simulations from a mechanistic model (described in [Yang et al. \(2022\)](#)) that simulates the evolution of immune responses to H3N2 viruses from 1968 to 2014 for 777 participants, as null models for testing the imprinting hypothesis. This model does not explicitly include any imprinting mechanism (or any other between-subtype interaction). As a positive control, we also simulated a hypothesised imprinting effect by boosting responses by 50% for participants born on or after 1968, relative to those born earlier (see Methods and Materials).

We find some support for the imprinting hypothesis

The best model by BIC included both age at sampling and at isolation, and the probabilistic imprinting variable, but in this model, only the coefficient for the H2N2 cohort was consistent with the imprinting hypothesis ([Fig 1A](#) and [Table S3](#) in Supplemental Information (SI)). However, the support for the imprinting hypothesis was sensitive to which age variables and imprinting variables were included ([Fig 1](#)). Hence, we consider our results to weakly support the imprinting hypothesis.

More generally, when using the non-probabilistic imprinting variable, coefficients aligned with the imprinting hypothesis when the model included age at sampling ([Fig 1A](#)). Conversely, when age was not included, the coefficients were not significant, and when only age at isolation was included, only the coefficient for the H2N2 cohort was significant. In contrast, for the probabilistic imprinting variable, only the models incorporating age at sampling yielded coefficients that were largely in line with the imprinting hypothesis. However, when age was not included, or when only age at isolation was included, coefficients for the H1N1 cohort (and to a lesser extent for the H2N2 cohort) did not significantly deviate from those for the H3N2 cohort. Models that included age at isolation were significantly better (by BIC) compared to models that only included age at sampling ([Fig 1B](#)). The inclusion of both age variables did not always consistently improve model performance relative to only having age at isolation. Only when age at isolation was included did models with the imprinting variables consistently outperform models without it ([Fig 1C](#)). However, results for models including both age splines should be interpreted with caution, given that in these models, the inclusion of an imprinting variable led to significant changes in the other splines and coefficients included in the model (see [Fig S15](#) in contrast to [Figs S12–S14](#) in SI).

When fitting to H1N1 titers, the best models included the probabilistic imprinting variable. In contrast to the results above, the best model by BIC included age at sampling ([Fig S16](#) and [Table S5](#) in SI). None of the coefficients were consistent with the imprinting hypothesis in these best models. More broadly, the results show that when age at isolation is included in the model, coefficients for the H1N1 cohort were higher than those for the H3N2 cohort, aligning with the imprinting hypothesis ([Fig S16](#) in SI). However, the coefficients for the H2N2 cohort did not clearly deviate from those for the H3N2 cohort.

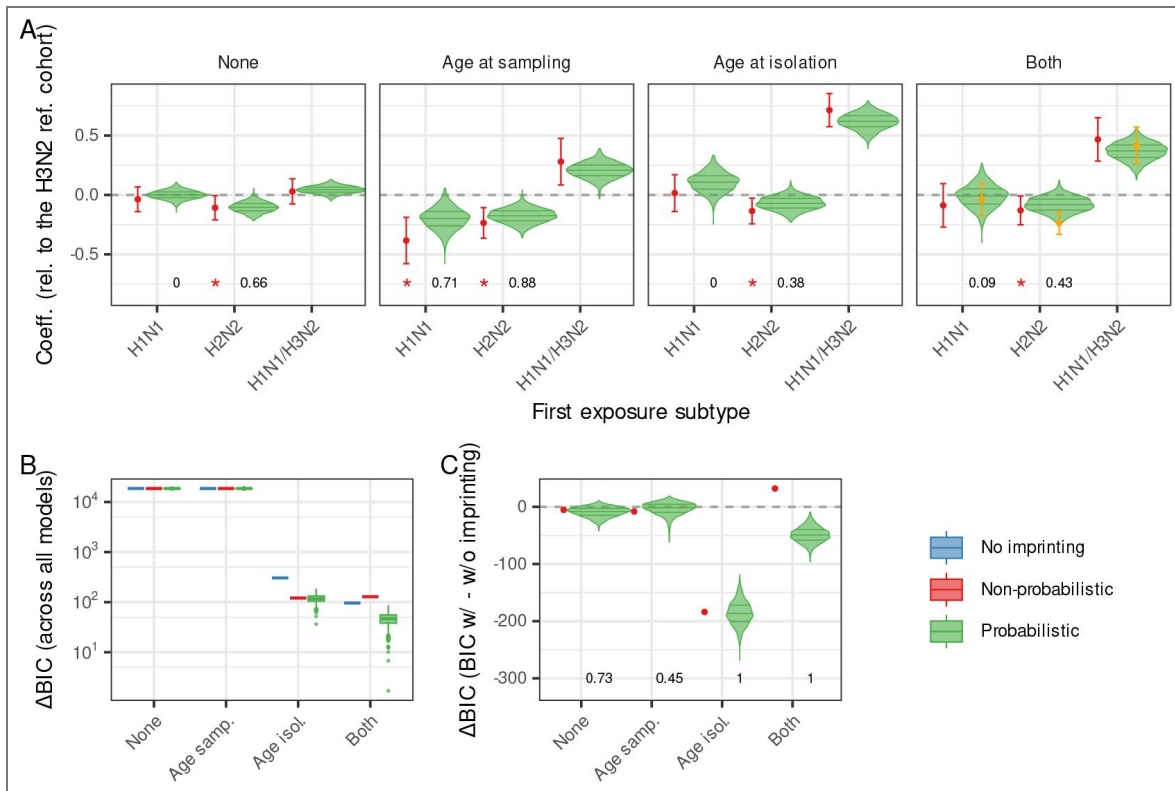


Figure 1. Coefficients for imprinting variables and model comparisons for GAMs fit to participants' H3N2 HI titers.

(A) The coefficients for the imprinting variables, relative to the reference H3N2 cohort. Each panel in (A) represents different adjustments for age, with “Both” indicating models incorporating both age at sampling and at isolation. Within each panel, the red points and lines depict the coefficients and uncertainty intervals for the non-probabilistic imprinting variable. Red stars highlight coefficients consistent with the imprinting hypothesis (i.e., coefficients for the H1N1 and H2N2 cohorts that are significantly lower than the H3N2 reference cohort). The violin plots show the distributions of 1000 coefficients for the probabilistic imprinting variable, with numbers below indicating the proportion of the 1000 coefficients that are both statistically significant and consistent with the imprinting hypothesis. The orange points and lines show the coefficients for the best model by BIC. See Table S3 in SI for the coefficients for the three best models by BIC. (B) BIC values relative to the model with the lowest BIC, across all models, both without and with (non-probabilistic and probabilistic) imprinting variables. (C) The differences in BICs between models with and without the imprinting variable; negative values suggest better performance of the model with the imprinting variable. The numbers below the violin plots in (C) show the proportions of the comparisons for which the difference in BIC is three or more in favour of the model with the imprinting variable. See Fig S11 in SI for analogous results using deviance explained and root mean square error (RMSE) rather than BIC.

Simulation models without imprinting can produce similar results

As in the models fit to the H3N2 titer data, the best model fit to the null simulation model (without an imprinting effect) by BIC included age at isolation and the probabilistic imprinting variable, although in contrast to the fits to the H3N2 titer data, the coefficients did not support the imprinting hypothesis.

More broadly, however, there are clear similarities in the patterns of the coefficients observed for the fits to the data, and those for the fits to the null model output, which lacks any imprinting mechanism (Fig 2A). Models that incorporate both age variables provide minimal support for the imprinting hypothesis, while models only including age at isolation suggest only the H2N2 cohort may have systematically lower titers than the H3N2 cohort. Nonetheless, inclusion of the imprinting variable in the models is still strongly favoured (by BIC; Fig 2C). Including age at sampling into a model already incorporating age at isolation does not lead to significant improvements in the model fit (by BIC; Fig 2B,C).

When the effects of an imprinting mechanism are included in the model (the positive control model), the imprinting variable successfully identifies the pattern and produces coefficients that clearly align with the imprinting hypothesis (Fig 2A). The signal is strongest when using the non-though in contrast to the fits to the H3N2 titer data, the coefficients did not support the imprinting hypothesis.

More broadly, however, there are clear similarities in the patterns of the coefficients observed for the fits to the data, and those for the fits to the null model output, which lacks any imprinting mechanism (Fig 2A). Models that incorporate both age variables provide minimal support for the imprinting hypothesis, while models only including age at isolation suggest only the H2N2 cohort may have systematically lower titers than the H3N2 cohort. Nonetheless, inclusion of the imprinting variable in the models is still strongly favoured (by BIC; Fig 2C). Including age at sampling into a model already incorporating age at isolation does not lead to significant improvements in the model fit (by BIC; Fig 2B,C).

When the effects of an imprinting mechanism are included in the model (the positive control model), the imprinting variable successfully identifies the pattern and produces coefficients that clearly align with the imprinting hypothesis (Fig 2A). The signal is strongest when using the non-probabilistic imprinting variable (as expected, given that the effects of the imprinting mechanism in the model are encoded in the same way). However, when both age variables are included in the model, the signal for the H1N1 cohort is diminished when the probabilistic imprinting variable is used, perhaps explaining why in our analyses of the data, the coefficients for the H2N2 cohort were more likely to be consistent with the imprinting hypothesis than the coefficients for the H1N1 cohort.

Discussion

While a wide range of data on disease and mortality indicates evidence for immune imprinting, our analysis demonstrates that patterns in serological data are somewhat equivocal. Statistical models of our cohorts HI titers showed evidence in support of immune imprinting, however, this support was dependent on the other variables included in the model and significance patterns were also sensitive to specific models considered. Evidence was also stronger for imprinting by H2N2 being associated with reduced H3N2 titers than other combinations (a pattern corroborated even in our positive control simulation). Statistical patterns consistent with immune imprinting (i.e., systematically lower HI titers against H3N2 strains in cohorts likely first exposed to and infected by H1N1 and H2N2) could be reproduced by a simulation model that lacked an explicit imprinting mechanism or subtype-subtype interactions, although the best models fit to the simulation output had coefficients that were not consistent with the imprinting hypothesis.

Nevertheless, these results do not represent evidence for the absence of immune imprinting. Indeed, there are many factors that may help explain why we might not detect a clear and strong imprinting effect in our analysis of serology, even if the mechanism were at play (Oidtmann et al.,

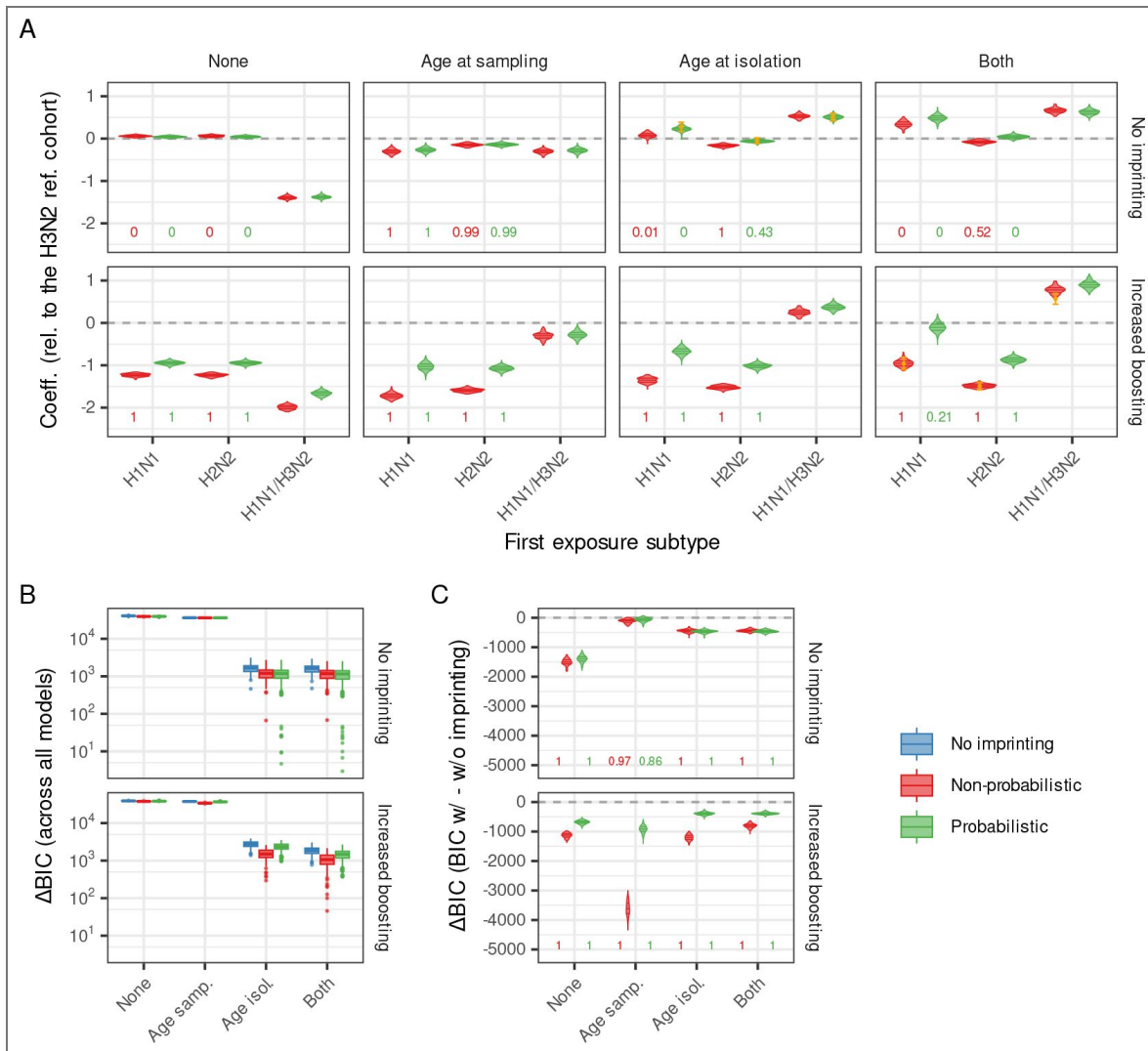


Figure 2. Coefficients for imprinting variables, and model performance comparisons for GAMs fit to simulation output.

The coefficients for the GAMs are for fits to both the null model (without an imprinting mechanism) and the positive control model (including imprinting via increased boosting; Methods and Materials). (A) The coefficients for the imprinting variables, relative to the reference H3N2 cohort. Each panel in (A) represents different adjustments for age, with “Both” indicating models incorporating both age at sampling and at isolation. Within each panel, the red violin plots show the distributions of the coefficients for the non-probabilistic imprinting variable (one for each of the 100 stochastic realisations of the model). The green violin plots show the distributions of 1000 coefficients for the probabilistic imprinting variable. The numbers below the violin plots indicate the proportions of the coefficients that are both statistically significant and consistent with the imprinting hypothesis (i.e., coefficients for the H1N1 and H2N2 cohorts that are significantly lower than those for the reference H3N2 cohort). Orange points and lines indicate the coefficients and 95% confidence intervals for the best models by BIC. (B) BIC values are relative to the model with the lowest BIC, across all models, both without and with (non-probabilistic and probabilistic) imprinting variables. (C) The differences in BICs between models with and without the imprinting variable; negative values suggest better performance of the model with the imprinting variable. The numbers below the violin plots in (C) show the proportions of the comparisons for which the difference in BIC is three or more in favour of the model with the imprinting variable.

2021 [↗](#)). HI assays quantify titers of antibodies that predominantly bind to sites present in the HA head (Miller et al., 2013 [↗](#)), while imprinting may be mainly driven by antibody responses to the HA stalk (Gostic et al., 2016 [↗](#); Knight et al., 2020 [↗](#); Arevalo et al., 2020a [↗](#)), by other components of immune memory (Nelson and Sant, 2019 [↗](#)), or by the quantity and specificity of the antibodies (Oidtman et al., 2021 [↗](#)). Stalk-specific antibodies produced in response to an imprinting subtype have been shown to weakly bind to virus strains of a different subtype (Arevalo et al., 2020a [↗](#)). This could lead to more infections, and as a result, a relative increase in the number of head-specific antibodies, and therefore, higher HI titers against the strain of the non-imprinting subtype. Furthermore, prior evidence for imprinting relied on reported severe and fatal cases, suggesting that imprinting may primarily help reduce disease severity. However, the relationship between HI titers and protection (however it is defined) and disease severity is complex, and likely a function of age, subtypes, and other factors (Fox et al., 2015 [↗](#); Ranjeva et al., 2019 [↗](#); Oidtman et al., 2021 [↗](#)). Our dataset may include participants who have either not been infected or were asymptomatic; in these cases, immune imprinting could be absent or its effects subtler. Immune imprinting might also not be solely mediated by antibody responses, e.g., other branches of the immune response (e.g., CD4 T cells; Nelson and Sant (2019)) may play a role, although previous studies indicating stratification in immune responses based on hemagglutinin type suggest that these effects would still be subtype specific (Gostic et al., 2016 [↗](#), 2019 [↗](#)).

In ascribing a subtype of likely first exposure to participants, we made assumptions on how early children are likely to be infected. The assumptions ranged between a first infection being certain within the first year, to possible within the first twelve (assumption made by Gostic et al. (2019 [↗](#))). The reality may well be between the two. Brouwer et al. (2022 [↗](#)) found that a significant proportion of children under one year of age have antibodies against influenza strains, but this fraction decreased as a function of the number of years since a new antigenic cluster was introduced. Furthermore, whether imprinting is driven exclusively by a first exposure, or whether more generally by the first few exposures, also remains an open question (Oidtman et al., 2021 [↗](#)) of particular significance in understanding the role immune imprinting may be playing post-1977, given the co-circulation of H1N1 and H3N2.

Given our focal question, accounting for participant infection history was a necessary element of all our analyses, achieved by including participant age (either at sampling and/or at isolation) in the model. However, any imprinting effect would also manifest as a birth cohort effect. To account for age effects, we used GAMs, a highly flexible non-parametric approach, and then included a categorical variable to detect systematic differences by birth cohorts aligning with the imprinting hypothesis. The implicit presumption is that systematic patterns caused by imprinting would be more parsimoniously captured by this categorical variable rather than in the age splines. To minimise the risk of the GAMs potentially masking patterns attributable to imprinting, we performed model selection with double penalisation within the GAMs. Despite these efforts, it is still possible that the splines and other model coefficients absorbed some of the imprinting signal (if there was one). For instance, coefficients for the H1N1/H3N2 cohort were often significantly positive, likely because this cohort is composed of younger participants who are more likely to have had a recent infection. These systematically higher titers were to some extent captured in the imprinting variable, as well as in the age splines. Nevertheless, results for models using both age splines should be interpreted with caution. In these models, most coefficients and splines changed significantly when the imprinting variable was added compared to when it was absent, indicating that any potential imprinting effect could be masked by an intricate interplay of compensating factors (see Section “Sensitivity analyses” in SI). Our approach did detect the imprinting mechanism when its effects were included in the simulation model. However, we simulated the imprinting effects by increasing boosting by 50% for participants imprinted by a virus of the same group as the infecting strain. Were the effects of imprinting subtler, or were they to operate through other mechanisms, the resulting coefficients might be harder to distinguish from those resulting from the model without an imprinting mechanism.

As a control for the effects of participant infection history in the context of possible cohort effects driven by years of different strain circulating, we fit the same statistical models to the outputs of a simulation model. This model did not include subtype-subtype interactions, and made some simplifying assumptions. For example, the model assumes that antigenic change between strains varies linearly as a function of time. The model was parameterised using fits to a subset of the cohort data analysed here, as well as data from a separate cohort in Vietnam. It is therefore possible that certain patterns, consistent with and potentially attributable to immune imprinting, may have been implicitly subsumed in the resulting parameterisation. This could possibly account for the observed similarity in patterns within the coefficients of the models fit to the simulation output and those of the models fit to the data.

To further explore what patterns we could expect, were imprinting to be at play, the simulation model could be further refined to introduce immune imprinting by explicitly including other subtypes and their interactions.

A greater understanding of immune responses and the intricate interplay between viral evolution and the life course of immunity to influenza, holds significant promise in shaping more effective vaccine policies and formulations. Determining the mechanisms underlying immune imprinting could lead to better predictions of how individuals may respond to potential antigenic drifts and shifts in circulating strains, to influenza vaccine candidates and to exposures to non-human hosted influenza viruses, based on their previous exposures. Our analysis demonstrates that imprinting-like patterns can emerge as a result of the confluence of age and temporal phenomena, even in the absence of any of the hypothesised immunological mechanisms of imprinting. The unique experiences of specific cohorts in relation to the timing of particular infections or the time since the introduction of a new antigenic cluster can result in apparent subtype-subtype interactions (Brouwer et al., 2022 [↗](#)). This highlights the complexity of interpreting patterns in the context of age and temporal dynamics. These findings may also apply to other pathogen systems characterised by complex interactions between rapidly evolving (e.g., SARS-CoV-2) or co-circulating and interacting (e.g., dengue) viruses and the host exposure history and immune response.

Methods and Materials

Ethical approval

The following institutional review boards approved the study protocols: Johns Hopkins Bloomberg School of Public Health (IRB 1716), University of Florida (IRB201601953), University of Liverpool, University of Hong Kong (UW 09-020), and Guangzhou No. 12 Hospital ('Research on human influenza virus immunity in Southern China'). Written informed consent was obtained from all participants over 12 years old; verbal assent was obtained from participants 12 years old or younger. Written permission from a legally authorized representative was obtained for all participants under 18 years old.

Data

The Fluscape cohort study is described by Jiang et al. (2016 [↗](#)). In summary, between 2009 and 2015, 4130 serum samples from 2008 individuals between 2 and 95 years of age were collected in Guangzhou, China, across 40 different locations and 920 households. From these, 61,614 HI titers against 21 A(H3N2) strains that circulated in the region over 47 years (1968–2014; with the exception of X31, a 1970 synthetic strain), and 5845 HI titers against three A(H1N1) strains (1977, 2005, and 2009), were measured. Samples were collected at roughly annual intervals, and on average, two samples were collected per participant. Serum samples were processed and titers estimated in three different batches (lab rounds). See Fig 3 [↗](#) for an overview.

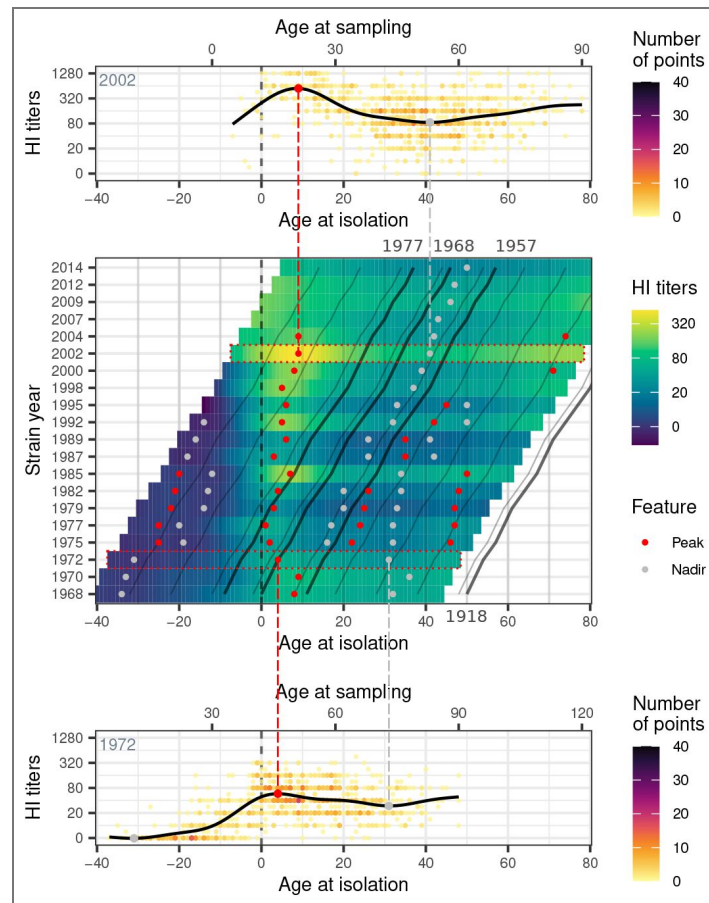


Figure 3. Overview of the Fluscape dataset, with emphasis on the patterns of HI titers with age.

The middle panel displays the HI titers predicted for each H3N2 strain as a function of age at isolation, for visit 4 (2014–2015) and lab round 3. The predictions were generated using GAMs fitted independently to each strain, including household ID as a random effect (these differ from the GAM models used in the main analysis, in which all strains were included in the models). Red and grey points indicate the locations of peaks and nadirs in the predicted HI titers for each strain. The diagonal contour lines follow birth cohorts, with thicker lines representing birth cohorts of specific influenza relevance: 1918, 1957, 1968, 1977. For explanation, we select two strains (1972 on the bottom panel and 2002 on the top panel), both highlighted in the heatmap with red dotted lines, and overlay the predicted titers from the GAM on the data used for the fits. The colour of the points in the top and bottom panels indicates the total number of overlaid points. Note the systematically higher titers in individuals who were young when each strain was circulating (as previously observed by Lessler et al. (2012) using a subset of the same data), and how individuals born shortly after the introduction of H2N2 in 1957 appear to have systematically low titers. In an analogous heatmap expressed as a function of age at sampling, the left and right edges of each row would vertically line up.

Statistical approach

We fit competing generalised additive models (GAMs) using the ‘bam’ function in the ‘mgcv’ package in R (R Development Core Team, 2023 [↗](#); Wood, 2011 [↗](#)). We transformed titers (expressed as the inverse of the two-fold serial dilutions) using $z = \log_2(x/5)$ where x are the titers, and we set $z = 0$ when titers were undetectable. The log titers (z) were the dependent variable in the models. These models assume normally-distributed residuals.

Immune responses are known to vary systematically with age (see Fig 3 [↗](#); Lessler et al. (2012 [↗](#)); Yang et al. (2020 [↗](#))); we control for age using two different variables: the age of the participant at the time of sampling (age at sampling), and the age the participant would have had when a given strain was circulating (age at isolation).

We fit a suite of models, each accounting for age in different ways: (i) a model without age, and then, models with a (thin plate regression) spline on (ii) age at sampling; (iii) age at isolation; and (iv) both, one spline on age at sampling and the other on age at isolation (additively). Furthermore, in all cases, we account for the time of sampling (the visit number as a categorical variable), the time since the start of the study (with a spline), and the participant gender. Finally, preliminary analyses showed that titers can vary systematically across lab rounds and as a function of the strain (e.g., see Fig 3 [↗](#)). For this reason, we include an interaction between the strain and lab round (and the main effect of each). To account for pseudoreplication (within-person correlations), we include a random effect on participant ID (encoded using `bs = "re"` in ‘mgcv’). To mitigate the potential for overfitting in the GAMs, we use double penalisation (via the `select = TRUE` argument in the ‘mgcv’ R package). We use restricted maximum likelihood to select the smoothing parameters.

Each of these models were fit with and without an “imprinting” variable, which is the variable of interest. It encodes the influenza subtype of likely first exposure based on year of birth. The coefficients of this variable (corresponding to each of the subtypes) indicate whether, when accounting for the aforementioned factors, participants imprinted with any of the subtypes tend to have systematically higher or lower H3N2 or H1N1 titers. The hypothesis being tested here, that imprinting should be detectable in serological data, would correspond to systematically lower H3N2 titers in participants imprinted with H1N1 or H2N2, when compared to participants imprinted with H3N2. Conversely, for the analyses of H1N1 titers, the expectation is reversed: participants imprinted with H3N2 should have systematically lower H1N1 titers compared to those imprinted with H1N1 or H2N2.

We use two approaches to determine the subtype of first exposure. In the first approach, we implicitly assume that participants are exposed to influenza within their first year of life, so that the subtype of first exposure is solely determined by the year of birth (see Fig 4 [↗](#)). For example, the subtype of first exposure would be H1N1 for anyone born before 1957, H2N2 for participants born between 1957 and 1967 (inclusive), and H3N2 for those born between 1968 and 1976. In 1977, H1N1 reemerged, and since then, both H1N1 and H3N2 have been co-circulating. We keep this as a separate mixed category (“H1N1/H3N2”), and explain the reason behind this choice in Section “Rationale for grouping the H1N1 and H3N2 cohorts for participants born after 1977” in SI. We refer to this approach as the “non-probabilistic” immune imprinting variable.

In the second approach, we assume that individuals may not be exposed to influenza within their first year of life, meaning that individuals born prior to the introduction of a new subtype may still have a non-zero probability of being imprinted with the new subtype. For example, a participant born in 1955 (corresponding to the period when H1N1 was circulating) may still be imprinted with H2N2 (which emerged in 1957). For this purpose, we use the probabilities of imprinting with a subtype as a function of year of birth estimated by Gostic et al. (2016 [↗](#)) for China (see Fig 4 [↗](#)). For each participant we sampled the subtype of first exposure using these probabilities. We then fit all models 1000 times, where the subtypes of first exposure for each participant were resampled each time. As in the first approach, we treat the period characterised by the reintroduction of H1N1 as its own mixed category of H1N1/H3N2. We refer to this approach based on sampling probabilities as the “probabilistic” immune imprinting variable.

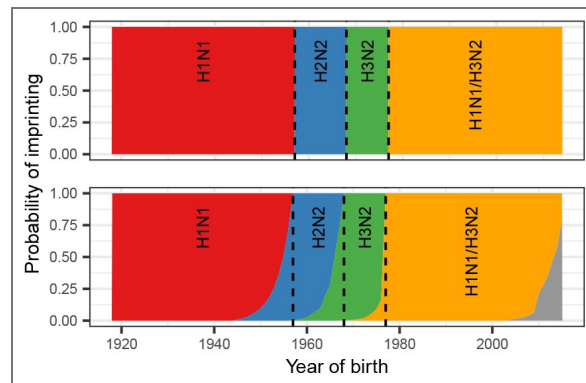


Figure 4. Proportions of each birth cohort with an initial exposure to each subtype.

These proportions correspond to 2014 in China, as obtained from Gostic et al. (2016). The top panel assumes that participants get infected in their first year of life, while the bottom panel relaxes this assumption, leading to non-zero probabilities of being infected and imprinted to more than one subtype in the years prior to the introduction of a new subtype (Gostic et al., 2016).

The model formula thus takes the following form:

$$z = \beta_0 + f_1(a_{\text{samp}}) + f_2(a_{\text{isol}}) + f_3(t) + \beta_1 \mathbf{l} + \beta_2 \mathbf{s} + \beta_3 \mathbf{l} \cdot \mathbf{s} + \beta_4 \mathbf{v} + \beta_5 \mathbf{p} + \epsilon, \quad (1)$$

where f_1, f_2 , and f_3 are smooth functions, $\beta_0, \beta_1, \dots, \beta_5$ are vectors of coefficients, a_{samp} and a_{isol} are the ages at sampling and isolation, respectively, \mathbf{l} is the lab round, \mathbf{s} is the strain, \mathbf{v} is the visit, \mathbf{p} the imprinting variable, and ϵ the error term. The age terms can both be excluded or included, or only included one at a time. Models are run including and excluding the imprinting variable. See Table S1 in SI for the rationale behind each parameter, and Table S2 in SI for a list of the set of models fit.

Simulation models

We use previously described simulations from mathematical models (Yang et al., 2022) as null models for testing the imprinting hypothesis. This mechanistic model primarily focuses on the dynamic interplay between antibody responses, incorporating within-subtype cross-reactivity, and antibody-level dependent infection risk, while no imprinting mechanism (or any other between-subtype interaction) is explicitly included. We refer to this model as the “null model”. In brief, the model simulates the evolution of immune responses to a A(H3N2) from 1968 to 2014, for a subset of 777 participants with a similar birth year distribution as the participants in our study. Prior to simulating each season, we used each participant’s immune profile to calculate the probability of infection, based on immunity-dependent protection (Hoa et al., 2022). Infection outcomes for each participant were randomly drawn annually, and the immune profile was updated according to the infection outcomes and cross-reactivity. We assumed that an A(H3N2) virus can cross-react to viruses that were isolated 10 years before or after the virus isolation, which is in line with previous estimates of the rate of evolution (Fonville et al., 2014; Kucharski et al., 2018). The overall population attack rate in the population was assumed to be constant at 20%.

As a positive control, we also ran simulations that mimicked the hypothesised imprinting effect, by increasing the antigenic-targeted boosting of participants born on or after 1968 by 50% (on the log-scale) relative to participants born prior to 1968. We refer to this model as the “positive control model”.

From the simulation output, we extract the titers corresponding to two visits (2004 and 2014) against 21 H3N2 strains circulating in the same years as those present in the data.

Output from the models were analysed using the same approach described above (“Statistical approach”). Here, however, we have 100 stochastic realisations of the model, yielding 100 sets of coefficients for the non-probabilistic imprinting variable. For each of the 100 model realisations including the probabilistic imprinting variable, we drew 10 samples for the subtypes of first exposure, producing a total of 1000 sets of coefficients. These 1000 coefficients therefore not only incorporate uncertainty in the sampling of the imprinting variable with the 10 samples per realisation (as in the analogous analyses of the data) but also uncertainties in the simulation through the 100 realisations of the model.

Sensitivity analyses

We tested the robustness of our results with further analyses. The GAMs we fit assume normally-distributed residuals; we also fit negative binomial GAMs. Several of the parameters we fit are correlated with the imprinting variable, and raises the question of parameter identifiability, and whether or how much of a possible underlying imprinting signal may be absorbed in parameters other than the imprinting variable. To this end, we re-fit the same models to data generated from the GAMs. We also assess whether our results change when accounting for maternal antibodies, and when fitting models to subsets of the data. Finally, we also check whether our results based on the simulation output are specific to the demographic profile of the participants. All of these analyses, and the corresponding results, are described in Section “Sensitivity analyses” in SI.

Additional information

Data availability

Data used in the manuscript are available at https://github.com/UF-IDD/influenza_imprinting .

Funding

J.A.H. is supported by a Wellcome Trust Early Career Award (grant 225001/Z/22/Z); J.R. is supported by a U.K. Medical Research Council grant (MR/S004793/1) and a Physical Sciences Research Council grant (EP/N014499/1); S.R. is supported by a Wellcome Trust Award (200861/Z/16/Z); C.Q.J., S.R., J.L., and D.A.T.C. are supported by a U.S. National Institutes of Health grant (R01TW0008246); J.L. and D.A.T.C. are supported by a U.S. National Institute on Aging grant (R56AG048075); D.A.T.C. is supported by a U.S. National Institute of Allergy and Infectious Diseases grant (R01AI114703).

Author ORCID iDs

Bernardo García-Carreras: <https://orcid.org/0000-0002-0714-300X>

Bingyi Yang: <https://orcid.org/0000-0002-0811-8332>

James A Hay: <https://orcid.org/0000-0002-1998-1844>

Siyu Chen: <https://orcid.org/0000-0001-9418-2527>

Steven Riley: <https://orcid.org/0000-0001-7904-4804>

Sarah Cobey: <https://orcid.org/0000-0001-5298-8979>

Author notes

Competing interests: B.G.-C., M.D.T.H., and D.A.T.C. received financial research support through their institution from Merck for unrelated work; D.A.T.C. receives financial research support through a grant to his institution from Pfizer for unrelated work; J.L. receives research support from CDC and NIH-NIGMS for unrelated work; S.R. receives grants from the Wellcome Trust; The other authors declare no competing interests.

Additional files

[Supplementary material](#) 

References

Arevalo C. P., Le Sage V., Bolton M. J., Eilola T., Jones J. E., Kormuth K. A., Nturibi E., Balmaseda A., Gordon A., Lakdawala S. S., *et al.* (2020a) Original antigenic sin priming of influenza virus hemagglutinin stalk antibodies. *Proceedings of the National Academy of Sciences* **117**:17221-17227

Arevalo P., McLean H. Q., Belongia E. A., Cobey S. (2020b) Earliest infections predict the age distribution of seasonal influenza A cases. *eLife* **9**:e50060 <https://doi.org/10.7554/eLife.50060>

Brouwer A. F., Balmaseda A., Gresh L., Patel M., Ojeda S., Schiller A. J., Lopez R., Webby R. J., Nelson M. I., Kuan G., *et al.* (2022) Birth cohort relative to an influenza A virus's antigenic cluster introduction drives patterns of children's antibody titers. *PLOS Pathogens* **18**:e1010317

Edler P., Schwab L. S. U., Aban M., Wille M., Spirason N., Deng Y.-M., Carlock M. A., Ross T. M., Juno J. A., Rockman S., *et al.* (2024) Immune imprinting in early life shapes cross-reactivity to influenza B virus haemagglutinin. *Nature Microbiology* **9**:2073-2083

Fonville J. M., Wilks S. H., James S. L., Fox A., Ventresca M., Aban M., Xue L., Jones T. C., Le N. M. H., Pham Q. T., *et al.* (2014) Antibody landscapes after influenza virus infection or vaccination. *Science* **346**:996-1000

Fox A., Mai L. Q., Thanh L. T., Wolbers M., Hang N. L. K., Thai P. Q., Yen N. T. T., Hoa L. N. M., Bryant J. E., Duong T. N., *et al.* (2015) Hemagglutination inhibiting antibodies and protection against seasonal and pandemic influenza infection. *Journal of Infection* **70**:187-196

- Gagnon A.**, Acosta E., Hallman S., Bourbeau R., Dillon L. Y., Ouellette N., Earn D. J. D., Herring D. A., Inwood K., Madrenas J., *et al.* (2018) Pandemic paradox: Early life H2N2 pandemic influenza infection enhanced susceptibility to death during the 2009 H1N1 pandemic. *mBio* **9**
<https://doi.org/10.1128/mbio.02091-17>
- Gostic K. M.**, Ambrose M., Worobey M., Lloyd-Smith J. O. (2016) Potent protection against H5N1 and H7N9 influenza via childhood hemagglutinin imprinting. *Science* **354**:722-726
- Gostic K. M.**, Bridge R., Brady S., Viboud C., Worobey M., Lloyd-Smith J. O. (2019) Childhood immune imprinting to influenza A shapes birth year-specific risk during seasonal H1N1 and H3N2 epidemics. *PLOS Pathogens* **15**:e1008109
- Hoang L. N. M.**, Sullivan S. G., Mai L. Q., Khvorov A., Phuong H. V. M., Hang N. L. K., Thai P. Q., Thanh L. T., Carolan L., Anh D. D., *et al.* (2022) Influenza A(H1N1)pdm09 But Not A(H3N2) Virus Infection Induces Durable Seroprotection: Results From the Ha Nam Cohort. *The Journal of Infectious Diseases* **226**:59-69
- Hobson D.**, Curry R. L., Beare A. S., Ward-Gardner A. (1972) The role of serum haemagglutination-inhibiting antibody in protection against challenge infection with influenza A2 and B viruses. *Epidemiology & Infection* **70**:767-777
- Jiang C. Q.**, Lessler J., Kim L., Kwok K. O., Read J. M., Wang S., Tan L., Hast M., Zhu H., Guan Y., *et al.* (2016) Cohort profile: A study of influenza immunity in the urban and rural Guangzhou region of China: the Fluscape Study. *International Journal of Epidemiology* **46**:e16
- Knight M.**, Changrob S., Li L., Wilson P. C. (2020) Imprinting, immunodominance, and other impediments to generating broad influenza immunity. *Immunological Reviews* **296**:191-204
- Kucharski A. J.**, Lessler J., Cummings D. A. T., Riley S. (2018) Timescales of influenza A/H3N2 antibody dynamics. *PLOS Biology* **16**:e2004974
- Lessler J.**, Riley S., Read J. M., Wang S., Zhu H., Smith G. J. D., Guan Y., Jiang C. Q., Cummings D. A. T. (2012) Evidence for antigenic seniority in influenza A (H3N2) antibody responses in Southern China. *PLOS Pathogens* **8**:e1002802
- Miller M. S.**, Gardner T. J., Krammer F., Aguado L. C., Tortorella D., Basler C. F., Palese P. (2013) Neutralizing antibodies against previously encountered influenza virus strains increase over time: A longitudinal analysis. *Science Translational Medicine* **5**:198ra107-198ra107
- Moritzky S. A.**, Richards K. A., Glover M. A., Krammer F., Chaves F. A., Topham D. J., Branche A., Nayak J. L., Sant A. J. (2023) The Negative Effect of Preexisting Immunity on Influenza Vaccine Responses Transcends the Impact of Vaccine Formulation Type and Vaccination History. *The Journal of Infectious Diseases* **227**:381-390
- Nelson S. A.**, Sant A. J. (2019) Imprinting and Editing of the Human CD4 T Cell Response to Influenza Virus. *Frontiers in Immunology* **10**
- Oidtman R. J.**, Arevalo P., Bi Q., McGough L., Russo C. J., Cruz D. V., Vieira M. C., Gostic K. M. (2021) Influenza immune escape under heterogeneous host immune histories. *Trends in Microbiology* **29**:1072-1082
- R Development Core Team** (2023) R: A Language and Environment for Statistical Computing.
- Ranjeva S.**, Subramanian R., Fang V. J., Leung G. M., Ip D. K. M., Perera R. A. P. M., Peiris J. S. M., Cowling B. J., Cobey S. (2019) Age-specific differences in the dynamics of protective immunity to influenza. *Nature Communications* **10**:1660
- Reber A. J.**, Kim J. H., Biber R., Talbot H. K., Coleman L. A., Chirkova T., Gross F. L., Steward-Clark E., Cao W., Jefferson S., *et al.* (2015) Preexisting Immunity, More Than Aging, Influences Influenza Vaccine Responses. *Open Forum Infectious Diseases* **2**:ofv052
- Wood S. N.** (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* **73**:3-36

Yang B., García-Carreras B., Lessler J., Read J. M., Zhu H., Metcalf C. J. E., Hay J. A., Kwok K. O., Shen R., Jiang C. Q., *et al.* (2022) Long term intrinsic cycling in human life course antibody responses to influenza A(H3N2): an observational and modeling study. *eLife* **11**:e81457
<https://doi.org/10.7554/eLife.81457>

Yang B., Lessler J., Zhu H., Jiang C. Q., Read J. M., Hay J. A., Kwok K. O., Shen R., Guan Y., Riley S., *et al.* (2020) Life course exposures continually shape antibody profiles and risk of seroconversion to influenza. *PLOS Pathogens* **16**:e1008635

Peer reviews

Reviewer #1 (Public review):

Summary:

This manuscript uses serological data to quantify the effects of imprinting on subsequent influenza antibody responses. While this is an admirable goal, the HI dataset sounds impressive, and the authors developed a number of models, the manuscript came off as very dense and technical. One of the biggest pitfalls is that it is not easy to understand the lessons learned. The two Results section headers make clear statements - there was an imprinting signal in the HI titers, but much of this signal could also be seen in an imprinting-free simulation - and then the Discussion states a number of limitations. This is fine, but it leaves the reader wondering exactly how large an error would be introduced by ignoring imprinting effects altogether; alternatively, if imprinting is purposefully added, what would the expected effect size be? The comments below will provide some concrete steps to help clarify these points.

Major comments:

(1) Lines 107-133: The first Results section is a dense slog of information, and the reader is never given a good overview of what the imprinting coefficients exactly are. As the paper currently stands, if you do not start by reading the Methods, you will take away very little. I suggest adding a schematic for any of your models, showing what HI titers would be expected with/without imprinting effects. or age effects, or both, to tie in your modeling coefficients with quantities that all readers are familiar with.

(1.1) Clarify what the imprinting coefficient (y-axis in Figure 1A) looks like in this schematic.

(1.2) Another aspect that I missed: In addition to stating which models were best by BIC, what is the absolute effect size in the HI titers? During my initial reading, I had hoped that Figure 3 would answer this question, but it turned out to be just an overview of the dataset. I strongly suggest having such a figure to show the imprinting effect inferred by different models. What would the expected effect be if you kept someone's birth year constant but tuned their age? What if you kept their age at collection constant but tuned their birth year?

(1.3) It would also help to explain in your schematic what the x-axis labels (H1, H2, H1/H3) would look like in these scenarios, and what imprinting relative to H3 means.

(2) As mentioned above, it was hard to understand the takeaway messages, such as:

(2.1) A similar question would be: If you model antibody titers without imprinting, how far off would you be from the actual measurements (2x off, 4x off...)? If you add the imprinting effect, how much closer do you get?

(2.2) Are there specific age groups that require imprinting to be taken into consideration, since otherwise HI analyses will be markedly off?

(2.3) Are there age groups where imprinting can be safely ignored?

(3) HI titers against multiple H1 and H3 variants were measured, but it is unclear how these are used, and why titers against a single variant each season would not have worked equally well.

<https://doi.org/10.7554/eLife.109291.1.sa1>

Reviewer #2 (Public review):

Summary:

In this study, the authors were testing the hypothesis that hemagglutination inhibition antibody titers, measured later in life, might be higher against influenza viruses that belong to the same hemagglutinin classification group as the influenza virus that a person was likely first exposed to early in life. This is one conceptualization of a phenomenon termed immune imprinting, which may explain previously observed differences in susceptibility to severe influenza infection between cohorts that were likely first exposed to different hemagglutinin groups. The results of the analysis provide some support for this analysis. However, support for the hypothesis is not consistently observed across sensitivity analyses, and a simulation study finds that antibody patterns consistent with immune imprinting may arise due to other factors in the absence of true imprinting effects. Therefore, overall support for the hypothesis is weak. Nonetheless, this study is important in that it provides guidance and has developed an analytic methodology for additional studies in this area of research. These findings and methods may also be useful for other infectious diseases for which patterns consistent with immune imprinting have been observed.

Strengths:

The strengths of this study include the relatively large cohort data source with broad age representation, rigorous statistical methods, and the use of sensitivity and simulation analyses to assess the robustness of the results.

Weaknesses:

The model outcome includes antibody titers measured against many different viruses, and the imprinting parameter was defined at the subtype level. This may obscure specific imprinting effects related to finer structural similarities between first and subsequent virus exposures. This analysis focuses only on one component of the immune response to influenza; immune imprinting may also involve other immune mechanisms. The analysis was carried out in a Chinese cohort, and vaccination status of the cohort is not discussed; the results may not be generalizable to other populations, particularly if vaccination patterns differ.

<https://doi.org/10.7554/eLife.109291.1.sa0>