Childhood immune imprinting to influenza A shapes birth year-specific risk during seasonal H1N1 and H3N2 epidemics

Katelyn M Gostic[1], Rebecca Bridge[2], Shane Brady[2], Cécile Viboud[3], Michael Worobey[4], and James O Lloyd-Smith[1,3*]

[1]Dept. of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA, USA
[2]Arizona Department of Health Services, Phoenix AZ, USA
[3]Fogarty International Center, National Institutes of Health, Bethesda, MD, USA
[4]Dept. of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA
* jlloydsmith@ucla.edu

Abstract

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

Introduction

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

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

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

60 Within-subtype cross-protection is known to shape seasonal influenza's epidemiology and
61 evolution (22). But because this type of narrow immunity decays rapidly in the face of antigenic drift, it
62 would not be expected to shape cohort-specific protection across an entire human lifetime (23,24).
63 Conversely, broad, HA group-level immune memory arises when lymphocytes target conserved HA
64 epitopes, and can play a strong role in defense against unfamiliar influenza strains (e.g. novel, avian or
65 pandemic subtypes (10,17,19,20,25,26)). Broad, HA group-level responses are not traditionally thought to
66 play a strong role in defense against familiar, seasonal influenza subtypes, but could plausibly be
67 deployed against drifted seasonal strains whose variable HA epitopes have become unrecognizable.
68 Because the conserved antigenic targets involved in broad, HA group-level protection are relatively stable
69 over time, they could in theory, facilitate lifelong imprinting-related biases in immune memory (10,26).
70 Thus, childhood immune imprinting may determine which birth cohorts are primed for effective defense
71 against seasonal strains with conserved HA epitopes characteristic of group 1 or group 2, or with variable
72 HA epitopes characteristic of a particular subtype (H1, H2, etc.). A similar line of reasoning may apply to
73 immunity against NA, although much less attention has been paid to this antigen.
74 Since 1977, two distinct subtypes of influenza A, H1N1 and H3N2, have circulated seasonally in
75 humans, with striking but poorly understood differences in their age-specific impact (8,11–13,27). These
76 differences could be associated with childhood imprinting: older cohorts were almost certainly exposed to
77 H1N1 in childhood (since it was the only subtype circulating in humans from 1918-1957), and now seem
78 to be preferentially protected against modern seasonal H1N1 variants (8,11–13). Likewise, younger adults
79 have the highest probabilities of childhood imprinting to H3N2, which is consistent with relatively low
80 incidence of seasonal H3N2 in these cohorts (***Fig. 1-2***). Alternatively, differences in the evolutionary
81 dynamics of H1N1 and H3N2 could explain the observed age profiles. Subtype H3N2 exhibits slightly
82 faster drift in its antigenic phenotype than H1N1, and as a result, H3N2 may be more able to escape pre-
83 existing immunity and infect older, immunologically experienced adults, whereas H1N1 may be relatively
84 restricted to infecting immunologically naïve children (28).

We analyzed a large data set on seasonal influenza incidence to test whether cohort effects from childhood imprinting primarily act against variable epitopes, only providing cross-protection against closely related HA or NA variants of the same subtype, or against more conserved epitopes, providing broad cross-protection across HA subtypes in the same phylogenetic group (***Fig. 1A-B***). We fitted a suite of models to data using maximum likelihood and compared models using AIC. In a separate analysis, we considered the hypothesis that differences in evolutionary rate of H1N1 and H3N2, rather than imprinting effects, shape differences in age distribution. Our results have implications for long-term projections of seasonal influenza risk in elderly cohorts (12), who suffer the heaviest burdens of influenza-related morbidity and mortality, and whose imprinting status will shift through time as cohorts born during different inter-pandemic eras grow older.

The Data

The Arizona Department of Health Services (ADHS) provided a dataset containing 9,510 seasonal H1N1 and H3N2 cases from their statewide surveillance system. Cases of all ages were confirmed to subtype by PCR and/or culture, primarily from virologic testing at the Arizona State Public Health Laboratory. A smaller number of positive influenza tests were obtained through reporting by other clinical labs, which has been mandatory in Arizona since 2004 (29). Cases were observed across 22 years of influenza surveillance, from the 1993-1994 influenza season through the 2014-2015 season, although sample sizes increased dramatically after the 2009 pandemic (***Table 1***). The data included positive test results from patients in hospitals, long-term care facilities, correctional facilities, and outpatient clinics, and thus captured a range of case severities.

Following CDC standards, ADHS defines the influenza season as epidemiological week 40 (around early October) through week 39 of the following year (30). The 2008-2009 and 2009-2010 influenza seasons spanned the first and second wave, respectively, of the 2009 H1N1 pandemic. We did not analyze cases observed during this time period, because age distributions of infection and molecular

110 drivers of immune memory differed during the 2009 pandemic from the normal, drivers of seasonal
111 influenza's immuno-epidemiology of interest to this study (13,17,20). From the dataset of 9,510 seasonal
112 cases (defined as any case observed outside the 2008-2009 or 2009-2010 season), we excluded 58 cases
113 with birth years before 1918 (whose imprinting status could not be inferred unambiguously), and one case
114 whose year of birth was recorded in error. Ultimately, we analyzed 9,541 cases.
115

116 # The Model

117 **Reconstructed imprinting patterns**

118 We reconstructed birth year-specific probabilities of childhood imprinting to H1N1, H2N2 or
119 H3N2 using methods described previously (10). These probabilities are based on patterns of first
120 childhood exposure to influenza A and reflect historical circulation (***Fig. 1A***). Most individuals born
121 between pandemics in 1918 and 1957 experienced a first influenza A virus (IAV) infection by H1N1, and
122 middle-aged cohorts born between pandemics in 1957 and 1968 almost all were first infected by H2N2
123 (note that because the first influenza exposure may occur after the first year of life, individuals born in the
124 years leading up to a pandemic have some probability of first infection by the new pandemic subtype,
125 ***Fig. 1A***). Ever since its emergence in 1968, H3N2 has dominated seasonal circulation in humans, and
126 caused the majority of first infections in younger cohorts. However, H1N1 has also caused some seasonal
127 circulation since 1977, and has imprinted a fraction of all cohorts born since the mid-1970s (***Fig. 1A***).
128 Reconstructions assumed children age 0-12 in the year of case observation might not yet have
129 been exposed to any influenza virus. Interactions between imprinting and vaccination of naïve infants are
130 plausible, but poorly understood (10,31). We did not consider childhood vaccination effects here; only a
131 small percentage of individuals in the ADHS data were born at a time when healthy infants were routinely
132 vaccinated against influenza.
133
134

135 **Expected age distributions under alternate imprinting models**

136 If HA subtype-level imprinting protection shapes seasonal influenza risk, primary exposure to
137 HA subtype H1 or H3 in childhood should provide lifelong protection against modern variants of the
138 same HA subtype. If imprinting protection acts primarily against specific NA subtypes, lifelong
139 protection will be specific to N1 or to N2 (***Fig. 1B***). Alternatively, if broad HA group-level imprinting
140 shapes seasonal influenza risk, then cohorts imprinted to HA subtype H1 or H2 (both group 1) should be
141 protected against modern, seasonal H1N1 (also group 1), while only cohorts imprinted to H3 (group 2)
142 would be protected against modern, seasonal H3N2 (also group 2) (***Fig. 1B***). Collinearities between the
143 predictions of different imprinting models (***Fig. 1D-I***) were inevitable, given the limited diversity of
144 influenza antigenic subtypes circulating in humans over the past century (reflected in ***Fig. 1A***). Note that
145 middle-aged cohorts, which were first infected by H2N2, are crucial, because they provide the only
146 leverage to differentiate between imprinting at the HA subtype, NA subtype or HA group-level level (***Fig.***
147 ***1B***).

148 Our approach distinguishes between age-specific risk factors of influenza infection related to
149 health and social behavior, and birth year-specific effects related to imprinting. Specifically, age-specific
150 risk could be influenced by medical factors like age-specific vaccine coverage, age-specific risk of severe
151 disease, and immunosenescence, or by behavioral factors like age-assorted social mixing, and age-
152 specific healthcare seeking behavior. These factors should have similar impacts on any influenza subtype.
153 In contrast, imprinting effects are subtype-specific. Thus, we fit a step function to characterize the shape
154 of age-specific risk of any confirmed influenza infection. Simultaneously, we modeled residual, subtype-
155 specific differences in risk as a function of birth year, to focus on the possible role of childhood
156 imprinting in H1N1 or H3N2 infections. Each tested model used a linear combination of age-specific risk
157 (***Fig. 1C***) and birth year-specific risk (***Fig. 1D-F***) to generate an expected distribution of H1N1 or H3N2
158 incidence (***Fig. 1G-I***). Note that for a given birth cohort, age-specific risk changed across progressive
159 years of case observation (as the cohort got older), whereas birth year-specific risk was constant over
160 time.

To test quantitatively whether observed subtype-specific differences in incidence were most consistent with imprinting at the HA subtype, NA subtype or HA group level, or with no contribution of imprinting, we fitted a suite of models to each data set using a multinomial likelihood and then performed model selection using AIC. AIC is used to compare the relative strength of statistical support for a set of candidate models, each fitted to the same data, and favors parsimonious models that fit the data well (32,33). Technical details are provided in the ***Methods***.

Tested models

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

Effect of influenza evolutionary rate on age profiles

We used publicly available data from *Nextstrain* (34,35), and from one previously published study (36), to calculate annual antigenic advance, which we defined as the antigenic distance between strains of a given lineage (pre-2009 H1N1, post-2009 H1N1 or H3N2) that circulated in consecutive seasons (***Methods***). The "antigenic distance" between two influenza strains is used as a proxy for similarity in antigenic phenotype, and potential for immune cross-protection. A variety of methods have been developed to estimate antigenic distance using serological data, genetic data, or both (35–37).

To assess the impact of antigenic evolutionary rate on the epidemic age distribution, we tested whether the proportion of cases in children increased in seasons associated with large antigenic changes.

If the rate of antigenic drift is a strong driver of age-specific influenza risk, then the fraction of influenza cases observed in children should be negatively related to annual antigenic advance (28). In other words, strains that have not changed much antigenically since the previous season should be unable to escape pre-existing immunity in immunologically experienced adults, and more restricted to causing cases in immunologically naïve children; strains that have changed substantially will be less restricted to children.

Results

Subtype-specific differences in age distribution

Seasonal H3N2 epidemics consistently caused more cases in older cohorts, while H1N1 caused a greater proportion of cases in young and middle-aged adults (***Figs. 2, S1-S2***). These patterns were apparent whether we compared H3N2 epidemic age distributions with those caused by the pre-2009 seasonal H1N1 lineage, or with the post-2009 lineage. Observed patterns are consistent with the predicted effects of cohort-specific imprinting (***Fig. 1***), and with previously reported differences in age distribution of seasonal H1N1 and H3N2 incidence (11–13,27). See ***Fig. 2*** for seasons where H1N1 and H3N2 co-circulated in substantial numbers, and ***Figs. S1-S2*** for the entire dataset and alternate smoothing parameters.

Imprinting model selection

The model containing NA subtype-level imprinting received the most statistical support, and the model containing HA subtype-level imprinting was the second most preferred in terms of AIC ***(Fig. 3, Table 2***). The ADHS data showed a strong preference for NA subtype-level imprinting over HA subtype-level imprinting (ΔAIC=23.42), and effectively no statistical support for broad, HA group-level imprinting (ΔAIC=245.18), or for an absence of imprinting effects (ΔAIC=380.47). Visual assessment of model fits (***Fig. 3C-D***) confirmed that models containing imprinting effects at the narrow, NA or HA subtype levels provided the best fits to data. The lack of support for the no-imprinting model supports the

212 hypothesis that imprinting from the first exposure shapes lifelong seasonal influenza risk, just as it does
213 avian-origin influenza (10, 12). However, imprinting appears to act more narrowly against seasonal
214 influenza than against avian influenza, providing cross protection only to a specific NA or HA subtype,
215 instead of broader, HA group-level protection. This result is consistent with the idea that
216 immunodominance of variable HA epitopes limits the breadth of immune cross protection deployed
217 against familiar, seasonal influenza subtypes (19,20).
218 As expected (see ***Fig. 1G-I***), predictions from the two best models were highly collinear, except
219 in their risk predictions among middle-aged, H2N2-imprinted cohorts (birth years 1957-1968), and some
220 other minor differences arising from normalization across birth-years.
221
222 **Fitted risk patterns**
223 Fitted age-specific risk curves took similar forms in all tested models, with risk decreasing
224 rapidly from birth through adolescence, and then decreasing much more slowly until the end of life (***Fig.***
225 ***2A*** shows the fitted curve from the best model). Estimates of imprinting parameters were less than one,
226 indicating some reduction in relative risk of infection (***Table 2***). Within the best model, estimated
227 reductions in relative risk from childhood imprinting were stronger for H1N1 (0.34, 95% CI 0.29-0.42)
228 than for H3N2 (0.71, 95% CI 0.62-0.82). ***Table 2*** shows parameter estimates and 95% profile confidence
229 intervals from all models fitted.
230
231 **Effect of evolutionary rate**
232 To test for effects of evolutionary rate on epidemic age distribution, we searched for decreases in
233 the proportion of cases among children in seasons associated with antigenic novelty, when highly drifted
234 strains might be more able to infect immunologically experienced adults. Consistent with this expectation,
235 the data showed a slight negative but not significant association between annual antigenic advance and
236 the fraction of H3N2 cases observed in children (***Fig. 4A***). However, note that no clear relationship
237 emerged between antigenic novelty and the fraction of cases observed in children and adults older than 10

238 (***Fig. 4A***). These are the cohorts in which epidemiological data show the clearest differences between
239 H1N1 and H3N2's age-specific impacts (***Fig. 2***); if rate of antigenic evolution is a dominant driver of
240 age-specific differences in incidence, we would have expected to see clearer evidence of evolutionary rate
241 effects within adults cohorts, not just in the youngest children. The data contained too few influenza
242 seasons with sufficient numbers of confirmed H1N1 cases to support meaningful Pearson correlation
243 coefficients for either pre-2009 or post-2009 seasonal H1N1 lineages.

244 Furthermore, if evolutionary rate is the dominant driver of subtype-specific differences in
245 epidemic age distribution, then when subtypes H1N1 and H3N2 show similar degrees of annual antigenic
246 advance, their age distributions of infection should appear more similar. However, the data showed that
247 differences in H1N1 and H3N2's age-specific impacts did not converge when lineages showed similar
248 annual advance. When comparing the fraction of cases observed in specific age classes, H1N1 data
249 consistently clustered separately from H3N2, with H1N1 consistently causing fewer cases at the extremes
250 of age (children 0-10 and elderly adults 71-85), but more cases in middle-aged adults, regardless of
251 antigenic novelty (***Fig. 4A***). Smoothed density plots showed no clear relationship between annual
252 antigenic advance and age distribution (***Fig. 4B***). Overall, the data showed a weak, but not significant
253 signal that incidence may be more restricted to young children when antigenic novelty is low, but the data
254 did not show strong evidence that the magnitude of annual antigenic drift is a systematic driver of
255 epidemic age distribution across the entire population.

256

257
Discussion

258 We analyzed a large epidemiological surveillance dataset and found that seasonal influenza
259 subtypes H1N1 and H3N2 cause different age distributions of infection, confirming previously reported
260 patterns (11–13). We analyzed several possible drivers of these differences side-by-side, and found
261 greatest support for the hypothesis that immunological imprinting leads to lasting protection against the
262 NA or HA subtype of the first influenza strain encountered in childhood (11,12). The data did not support

263 strong effects from broader HA group-level imprinting, as recently detected for novel zoonotic or
264 pandemic viruses (8,10), or from differences in rates of antigenic evolution (28). Our results suggest that
265 the first childhood infection leaves a lifelong imprint of immune memory to seasonal influenza, and that
266 this imprint is not erased even after decades of exposure to or vaccination against dissimilar influenza
267 subtypes.

268 As additional evidence that birth year, rather than age, drives subtype-specific differences in
269 seasonal influenza risk, we note that H3N2's impacts have not always been focused in elderly cohorts.
270 When H3N2 first emerged in 1968, it caused little or no excess mortality in the elderly, putatively
271 because those who were elderly in 1968 had been exposed, as children or young adults, to an H3 virus
272 that had circulated in the late 1800s (6,8). Meanwhile, H1N1-imprinted cohorts (those ~10-50 years old at
273 the time of the H3N2 pandemic), experienced considerable excess mortality in the H3N2 pandemic (6),
274 and continue to experience excess H3N2 morbidity and mortality today as elderly adults ((11–13,27), ***Fig.***
275 ***2***). In short, comparing data from H3N2's emergence in 1968 to its seasonal circulation today shows
276 impacts that have remained consistent with respect to birth year, but that have shifted with respect to age.

277 In model comparison, the data showed the strongest support for effects from childhood imprinting
278 to NA. Although NA is not as intensively studied as HA, these results emphasize the increasingly
279 recognized importance of both antigens as drivers of protection against seasonal influenza (14–16).
280 Realistically, some combination of effects from both HA and NA subtype-level imprinting probably
281 shape seasonal influenza risk. The models containing NA and HA subtype-level imprinting produced very
282 similar fits to data and emerged as the top two models in terms of AIC. Unfortunately, collinearities
283 between predictions of the simple, single-antigen models considered here arose inevitably from
284 influenza's limited diversity of circulation in humans over the past century. These collinearities prevented
285 us from testing more complicated models of combined effects from imprinting to HA and NA, or to other
286 antigens such as internal proteins. Because analysis of population-level data can support only a limited
287 scope of inference, deeper insights into the respective roles of HA, NA and other influenza antigens as
288 drivers of cohort effects will most likely need to come from focused immunological cohort studies in

289 which individual histories of influenza infection are known, such as those recently funded by the National
290 Institutes of Health (38). Alternatively, the development of immunological biomarkers for diagnosis of
291 imprinting status in individual patients could substantially increase the power of epidemiological
292 inference, which (as in this study) currently relies instead on probabilistic reconstructions of imprinting
293 histories according to birth year.

294 Small sample sizes may have limited our power to detect a statistically significant relationship
295 between annual antigenic advance and epidemic age distribution. The data did show a weak trend
296 supporting the idea that in seasons where antigenic advance is low, the seasonal influenza cases may be
297 more restricted to the youngest, immunologically naïve children (28). But the data did not reveal a clear
298 relationship between antigenic advance and the fraction of cases occurring in adult age groups, the same
299 age groups where epidemiological data reveals distinct subtype-specific differences in incidence
300 proportion. This lack of clear signal is consistent with growing recognition that existing methods to map
301 antigenic distance, which rely heavily on hemagglutination inhibition (HI) assays performed in laboratory
302 ferrets, do not always capture realistic patterns of cross-reactivity in humans (reviewed in 39,40). Further,
303 glycosylation of HA can cause antigenic escape in large subsets of the human population, yet such
304 posttranslational modifications may be perceived as neutral in existing antigenic maps (40,41). Moreover,
305 existing metrics of evolutionary and antigenic advance are based on properties of HA (34–36), but our
306 epidemiologic data support an equal if not stronger role of NA. We speculate that a clearer relationship
307 between epidemic age distribution and antigenic drift might emerge if antigenic distance measures were
308 able to incorporate cohort-specific variation in immune history, and impacts from multiple antigens.

309 While our results provide some valuable new clues about the underlying immune drivers of
310 imprinting protection against seasonal influenza, we can only speculate as to the exact mechanism.
311 Traditionally, narrow, within-subtype influenza immunity is thought to decay quickly in the face of
312 antigenic drift. Signals of rapid drift are largely based on HI data, which measures antibody responses to
313 just a handful of immunodominant, variable epitopes found near the receptor-binding domain on the HA
314 head. These epitopes accumulate substitutions rapidly, and so strains that circulated more than 14 years

315 apart rarely show measurable cross-protective HI titers (36). The short timescale of immune memory to
316 variable HA head epitopes stands in contrast to patterns observed in our study and others (11–13), where
317 within-subtype immune memory imprinted in childhood appears to persist for an entire human lifetime,
318 remaining evident even in the oldest cohorts in the data. Thus, we speculate that within-subtype
319 imprinting protection arises via different immune mechanisms than the well-studied antibody responses
320 measured by the HI assay.

321 One possibility is that within-subtype imprinting protection is driven by antibody responses to
322 intermediately conserved epitopes, which might remain stable over time, but lack structural homology
323 across different HA and NA subtypes. We rule out a strong role from antibody responses against the best-
324 studied conserved epitopes (e.g. those on the stalk), which tend to provide broader, cross-subtype
325 protection (10,17,19) than supported by model comparison. But recent studies show that B cell memory
326 shifts to focus on conserved influenza epitopes as we grow older, presumably because a lifetime of
327 exposures to drifted, seasonal H1N1 or H3N2 variants repeatedly back-boosts memory of unchanged
328 epitopes (23,24). Repeat boosting of intermediately conserved HA or NA antigens could explain the
329 longevity of subtype-level imprinting protection.

330 Another potential explanation supported by recent immunological data (26), is that the memory B
331 cell clones developed during the first childhood influenza exposure later adapt via somatic hypermutation
332 to “follow” antigenic targets as they drift over time. Thus, the first influenza exposure in life may fill a
333 child's memory B cell repertoire with clones that will serve in the future, not as final products but as
334 prototypes that can be rapidly and effectively tailored to recognize drifted influenza strains of the same
335 subtype. The adaptability of the B cell repertoire would not be detectable in traditional HI panels, which
336 are collected using sera from ferrets exposed to a single influenza variant, and do not reflect the
337 development of the human B cell repertoire across repeated, seasonal influenza exposures. A final
338 possibility is that cellular immunity (e.g. CD4+ T cell memory), which would not be captured in
339 serological assays, plays an underappreciated role in imprinting protection.

340 Signals of imprinting protection are anomalously strong in the current cohort of elderly adults, as
341 reflected by higher estimates of imprinting protection to H1N1 than H3N2. For nearly four decades from
342 1918-1957, H1N1 persisted as the only strain circulating in humans. The oldest subjects in our data were
343 born slightly after its emergence in 1918, and would not have encountered an influenza virus of any
344 subtype but H1N1 until after age 30. Decades of early-life exposures to H1N1 variants may have
345 reinforced and expanded the breadth of H1N1-specific immune memory in these oldest cohorts. But this
346 strong protection against H1N1 seems to come at a cost; even after decades of seasonal H3N2 exposure,
347 and vaccination, older cohorts have evidently failed to develop equally strong protection against H3N2.
348 Antigenic similarity between H1N1 strains that circulated earlier in the 20th century (which caused
349 imprinting in older cohorts), and modern H1N1 lineages that emerged in 1977 and in 2009, may also have
350 amplified the strength and longevity of H1N1 protection (4,42). One additional consideration is that HA
351 group 1 antigens appear to induce narrower immune responses than structurally distinct HA group 2
352 antigens, which may be better able to induce cross-group responses (21). Perhaps elderly cohorts
353 imprinted to group 1 antigens have been trapped in narrower responses that offer exceptional protection
354 against strains similar to that of first exposure but relatively poor adaptability to other subtypes.
355 Given that cohorts born after 1968 have had much more varied early life exposures to both H1N1
356 and H3N2, it is unclear whether equally strong, subtype-specific biases in protection will persist when
357 post-1968 birth cohorts eventually become elderly. Determining the precise immune mechanism(s)
358 responsible for subtype-level imprinting is necessary to project long-term shifts in influenza-related
359 incidence, and possibly in mortality. The vast majority of influenza-related deaths occur in adults over
360 age 65, and H3N2 has caused many times the number of fatalities in high-risk elderly cohorts as seasonal
361 H1N1, even in the post-2009 pandemic period (12,27,43). These patterns may arise because H3N2 is
362 intrinsically more virulent than H1N1, but we speculate that imprinting protection, which currently limits
363 the incidence of clinically-attended H1N1 infection in the elderly, may also explain these differences. In
364 the future, cohorts imprinted to H2N2 (born c. 1950-1968) will become elderly, and would expect
365 protection against H3N2 via NA subtype-level imprinting, while HA H2-level imprinting would not be of

366 much use against any currently circulating seasonal subtype. If future elderly cohorts continue to show
367 strong subtype-specific biases from imprinting, our results would corroborate the idea that mortality from
368 H1N1 may increase in the future (8,12) as protection in the elderly shifts toward other subtypes. On the
369 other hand, future generations of elderly adults, especially those born after H1N1 and H3N2 began to co-
370 circulate in 1977, may show a greater ability to act as immunological generalists, with effective defenses
371 against multiple influenza subtypes.

372 One limitation of this study was that we could not model the impact of seasonal influenza
373 vaccination explicitly, as the vaccination status of subjects in the ADHS data was unknown. We note that
374 the influenza vaccine contains both an H1N1 and H3N2 strain, and so on average, influenza vaccination
375 should protect individuals similarly against both subtypes. However, we also acknowledge that influenza
376 vaccine effectiveness varies by season, age group, and subtype, in ways that are poorly understood and
377 difficult to measure (44). These asynchronous and multi-dimensional shifts in vaccine effectiveness may
378 contribute to variability in H1N1 and H3N2's age distributions across influenza seasons.

379 Another limitation of this study was the low number of confirmed cases available in the pre-2009
380 era. To separate age-specific risk effects from birth year-specific cohort effects, the greatest power will
381 come from large data sets collected continuously over decades, so that individual birth cohorts can be
382 followed as they become considerably older. We emphatically echo earlier calls (45) for more systematic
383 sharing of single year-of-age in influenza surveillance data, standardization of sampling effort, and
384 reporting of age-specific denominators, which could substantially boost the scientific community's ability
385 to link influenza's genetic and antigenic properties with epidemiological outcomes.

386 Altogether, this analysis confirms that the epidemiological burden of H1N1 and H3N2 is shaped
387 by cohort-specific differences in childhood imprinting (8,11,12,46). The finding that such imprinting acts
388 at the HA or NA subtype level informs prediction of the future epidemiological impact of specific
389 seasonal subtypes in high-risk elderly cohorts. The lack of support for broader, HA group-level imprinting
390 effects emphasizes the consequences of immunodominance of influenza's most variable epitopes, and the
391 difficulty of deploying broadly protective memory B cell responses against familiar, seasonal strains.

392 Overall, these findings further our understanding of how antigenic seniority shapes cohort-specific risk
393 during epidemics. The fact that elderly cohorts show relatively weak immune protection against H3N2,
394 even after living through decades of seasonal exposure to or vaccination against H3N2, suggests that
395 antibody responses acquired in adulthood do not provide the same strength of immune protection as
396 responses primed in childhood. Immunological experiments that consider multiple viral exposures, and
397 cohort studies in which individual histories of influenza infection are tracked from birth, promise to
398 illuminate how B cell and T cell memory develop across a series of early life exposures. In particular,
399 these studies may provide clearer insights than epidemiological data into which influenza antigens,
400 epitopes and immune effectors play the greatest role in immune imprinting, and how quickly subtype-
401 specific biases become entrenched across the first or the first few exposures.
402
403
404

Methods

405 **Estimation of age from birth year in ADHS data**

406 The data contained three variables, influenza season, birth year and confirmed subtype. For most
407 cases, birth year was extracted directly from the reported date of birth in patient medical records, but age
408 was not known. We estimated patient age at the time case observation using the formula [year of
409 observation]-[birth year]. To ensure that the minimum estimated age was 0, the second year in the
410 influenza season of case observation was considered the calendar year of observation (e.g. 2013 for the
411 2012-2013 season).
412

413 **Splines**

414 In ***Figure 2***, smoothing splines were fit to aid visual interpretation of noisy data. We fit splines
415 using the command *smooth.spline(x = AGE, y = FRACTIONS, spar = 0.8)* in R version 3.5.0. Variables
416 *AGE* and *FRACTIONS* were vectors whose entries represented single years of age, and the fraction of

417 cases observed in the corresponding age group. The smoothing parameter 0.8 was chosen to provide a
418 visually smooth fit. Alternative smoothing parameter choices (0.6 & 1.0) are shown in ***Figs. S1-S2***.
419 Although the choice of smoothing parameter changed the shape of each fitted spline, qualitative
420 differences between splines fitted to H1N1 or H3N2 were robust.
421
422 **Model formulation**
423 For each unique season in which cases were observed, define p as a vector whose entries
424 represent the expected probability that a randomly drawn H1N1 or a randomly drawn H3N2 case was
425 observed in an individual of age a. Each model defined, p as a linear combination of age-specific risk,
426 birth year-specific risk (i.e. imprinting effects). All tested models were nested within the equation:
427
428 $$p = A * \mathbf{1}_{H1N1}(I_{H1N1}) * \mathbf{1}_{H3N2}(I_{H3N2}) \quad \mathbf{1}$$
429
430 To include risk factors that only modulated risk from one subtype, we included indicator
431 functions $\boldsymbol{1}_{H1N1}$ and $\boldsymbol{1}_{H3N2}$, which took value 1 if p described the expected age distribution of H1N1 or
432 H3N2 cases, respectively, and 0 otherwise.
433
434 ***Age-specific risk (A)***
435 Age-specific risk was defined as a step function, in which relative risk was fixed to value 1 in an
436 arbitrarily chosen age bin, and then z-1 free parameters, denoted r_2 to r_z, were fit to describe relative risk
437 in all other age bins. Below, $\boldsymbol{1}_i$ are indicator functions specifying whether each vector entry is a member
438 of age bin i. To obtain the predicted fraction of cases observed in each single year of age, we normalized
439 the risk distribution so that predicted risk across all age groups summed to 1.
440
441 $$A = norm(\mathbf{1}_1 + \mathbf{1}_2 r_2 + \cdots \mathbf{1}_z r_z) \quad \mathbf{2}$$
442

443 ***Imprinting (I)***

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

449

450 $$I_{HxNy} = \mathbf{1}_{HxNy} * [f_{IHxNy} r_{IHxNy} + (1 - f_{IHxNy})] \quad \mathbf{3}$$

451

452 **Likelihood**

453 We used equations 1-3 to generate predicted case age distributions (*p*) for each influenza season
454 (s) in which cases were observed in the data. Then, the likelihood was obtained as a product of
455 multinomial densities across all seasons. If n_s represents the total number of cases observed in a given
456 season, $x_{0cs}, \ldots x_{mcs}$ each represent the number of cases observed in each single year of age/single year of
457 birth, and if $p_{0cs} \ldots p_{mcs}$ each represent entries in the model's predicted age/birth year-distribution of cases,
458 then the likelihood is given by:

459

460 $$\mathcal{L} = \prod_s \frac{n_s!}{x_{0s}! \ldots x_{ms!}} p_{0s}^{x_{0s}} \ldots p_{ms}^{x_{ms}} \quad \mathbf{4}$$

461

462 **Model fitting and model comparison**

463 We fit models containing all possible combinations of the above factors to influenza data from
464 each season in the data. We simultaneously estimated all free parameter values using the optim() function
465 in R. We calculated likelihood profiles and 95% profile confidence intervals for each free parameter.
466 Confidence intervals were defined using the method of likelihood ratios (32).

467

468 **Antigenic advance**

469 We obtained antigenic distance estimates from *Nextstrain (nextstrain.org)*(34,47), and from
470 source data from Figure 3 in Bedford et al. (36). *Nextstrain* calculates antigenic distance using genetic
471 data from GISAID (48), and using methods described by Neher et al. (35). We analyzed "CTiter"
472 estimates from *Nextstrain*, which correspond to Neher et al.'s "tree model" method. We repeated analyses
473 using estimates from the similar "substitution model" method and verified that our choice of antigenic
474 distance metric did not meaningfully impact our results (results not shown). Datasets from *Nextstrain* and
475 Bedford et al. both contained redundant antigenic distance estimates for the H3N2 lineage, but only
476 Bedford et al. analyzed the pre-2009 H1N1 lineage, and only *Nextstrain* data analyzed the post-2009
477 H1N1 lineage. The antigenic distance estimates reported by Bedford et al. were roughly proportional to
478 those reported on *Nextstrain*, but greater in absolute magnitude (35). To enable visualization of all three
479 lineages on the same plot axes, we rescaled pre-2009 H1N1 estimates from Bedford et al. using the
480 formula $d_{Nextstrain} = 0.47 d_{Bedford}$. The scaling factor was chosen so that directly-comparable H3N2 distance
481 estimates obtained using each method spanned the same range (***Fig. S3***). The *Nextstrain* data files used in
482 this analysis are archived within our analysis code.

483 # References

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

485 2. Lessler J, Riley S, Read JM, Wang S, Zhu H, Smith GJD, et al. Evidence for Antigenic Seniority in
486 Influenza A (H3N2) Antibody Responses in Southern China. PLOS Pathog. 2012 Jul
487 19;8(7):e1002802.

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

490 4. Hancock K, Veguilla V, Lu X, Zhong W, Butler EN, Sun H, et al. Cross-Reactive Antibody
491 Responses to the 2009 Pandemic H1N1 Influenza Virus. N Engl J Med Boston. 2009 Nov
492 12;361(20):1945–52.

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

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

498 7. Ma J, Dushoff J, Earn DJD. Age-specific mortality risk from pandemic influenza. J Theor Biol.
499 2011 Nov 7;288:29–34.

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

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

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

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

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

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

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

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

522 16. Memoli MJ, Shaw PA, Han A, Czajkowski L, Reed S, Athota R, et al. Evaluation of
523 Antihemagglutinin and Antineuraminidase Antibodies as Correlates of Protection in an Influenza
524 A/H1N1 Virus Healthy Human Challenge Model. mBio [Internet]. 2016 Apr 19 [cited 2019 May
525 31];7(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4959521/

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

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

532 19. Krammer F. Novel universal influenza virus vaccine approaches. Curr Opin Virol. 2016 Apr;17:95–
533 103.

534 20. Andrews SF, Huang Y, Kaur K, Popova LI, Ho IY, Pauli NT, et al. Immune history profoundly
535 affects broadly protective B cell responses to influenza. Sci Transl Med. 2015 Dec
536 2;7(316):316ra192-316ra192.

537 21. Zost SJ, Wu NC, Hensley SE, Wilson IA. Immunodominance and Antigenic Variation of Influenza
538 Virus Hemagglutinin: Implications for Design of Universal Vaccine Immunogens. J Infect Dis.
539 2019 Apr 8;219(Supplement_1):S38–45.

540 22. Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, et al. Unifying the
541 Epidemiological and Evolutionary Dynamics of Pathogens. Science. 2004 Jan 16;303(5656):327–
542 32.

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

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

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

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

555 27. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, et al. Mortality
556 associated with influenza and respiratory syncytial virus in the United States. JAMA. 2003 Jan
557 8;289(2):179–86.

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

560 29. Arizona Department of Health Services. 2015–2016 Influenza Summary [Internet]. [cited 2019 May
561 23]. Available from: https://www.azdhs.gov/documents/preparedness/epidemiology-disease-
562 control/flu/surveillance/2015-2016-influenza-summary.pdf

563 30. National Notifiable Diseases Surveillance System, Division of Health Informatics and Surveillance,
564 National Center for Surveillance, Epidemiology and Laboratory Services. MMWR Week Fact Sheet
565 [Internet]. [cited 2019 May 23]. Available from:
566 https://wwwn.cdc.gov/nndss/document/MMWR_Week_overview.pdf

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

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

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

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

576 35. Neher RA, Bedford T, Daniels RS, Russell CA, Shraiman BI. Prediction, dynamics, and
577 visualization of antigenic phenotypes of seasonal influenza viruses. Proc Natl Acad Sci. 2016 Mar
578 22;113(12):E1701–9.

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

581 37. Smith DJ, Lapedes AS, Jong JC de, Bestebroer TM, Rimmelzwaan GF, Osterhaus ADME, et al.
582 Mapping the Antigenic and Genetic Evolution of Influenza Virus. Science. 2004 Jul
583 16;305(5682):371–6.

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

587 39. Morris DH, Gostic KM, Pompei S, Bedford T, Łuksza M, Neher RA, et al. Predictive Modeling of
588 Influenza Shows the Promise of Applied Evolutionary Biology. Trends Microbiol. 2018 Feb
589 1;26(2):102–18.

590 40. Cobey S, Hensley SE. Immune history and influenza virus susceptibility. Curr Opin Virol. 2017 Feb
591 1;22:105–11.

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

595 42. Rozo M, Gronvall GK. The Reemergent 1977 H1N1 Strain and the Gain-of-Function Debate. mBio.
596 2015 Sep 1;6(4):e01013-15.

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

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

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

604 46. Glezen WP, Keitel WA, Taber LH, Piedra PA, Clover RD, Couch RB. Age Distribution of Patients
605 with Medically-Attended Illnesses Caused by Sequential Variants of Influenza A/H1N1:
606 Comparison to Age-Specific Infection Rates, 1978–1989. Am J Epidemiol. 1991 Feb 1;133(3):296–
607 304.

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

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

613

Code and data availability

Code to perform all reported analyses and construct all plots is included as a supplementary file. All data is available in the supplementary materials.

Acknowledgements

We are grateful to Ken Komatsu and Kristen Herrick for their assistance with data access, and to Trevor Bedford for assistance accessing and interpreting antigenic distance data from *Nextstrain*. We thank Lone Simonsen for helpful discussions. KG was supported by the National Institutes of Health (F31AI134017, T32-GM008185). JLS was supported by NSF grants OCE-1335657 and DEB-1557022, SERDP RC-2635, and DARPA PREEMPT D18AC00031. MW was supported by the David and Lucile Packard Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclaimer

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

Competing interests

The authors declare no competing interests.

Ethics Statement

This study analyzed only existing epidemiological data, which did not include identifying patient details.

Author contributions

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

Figures

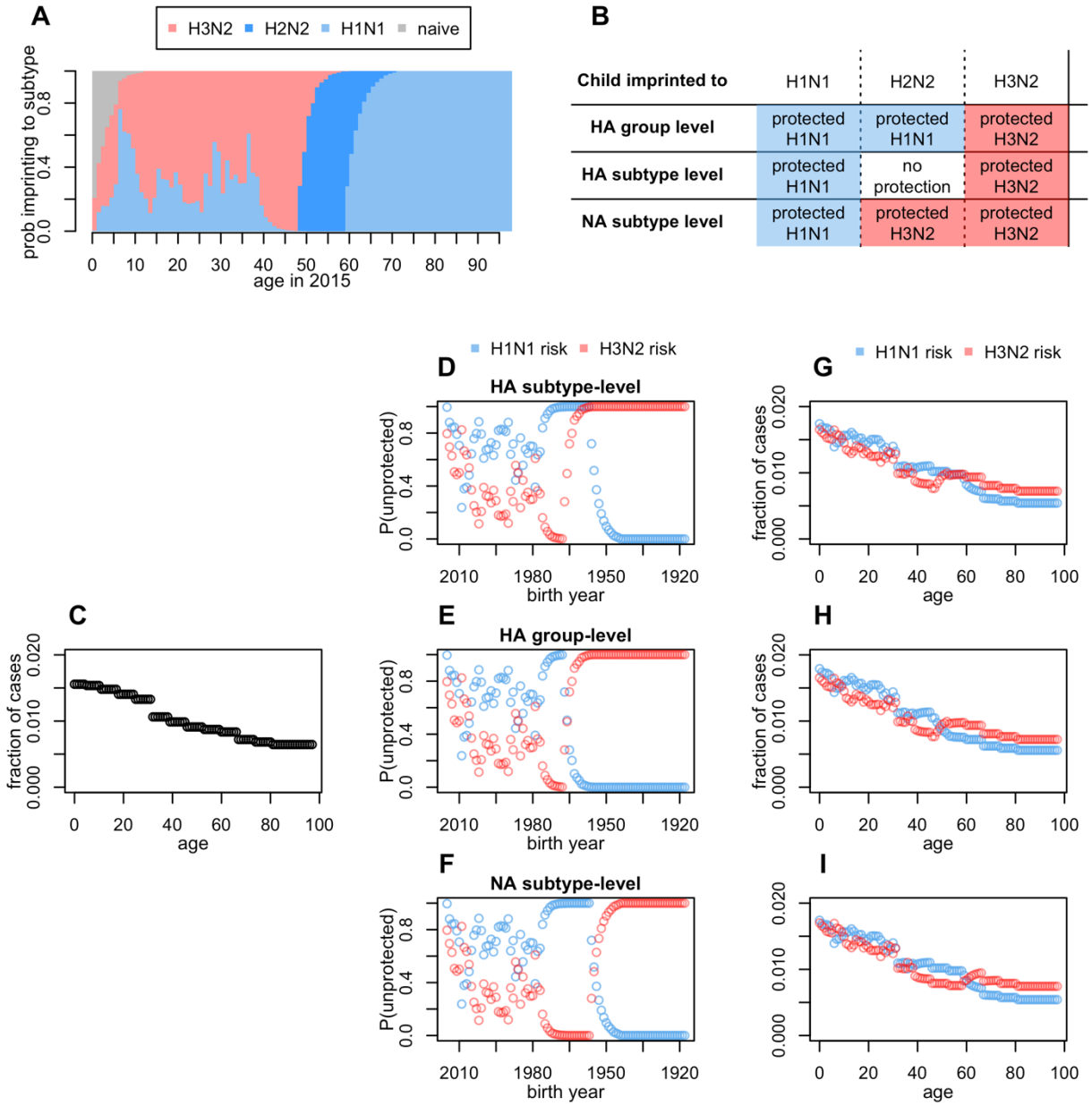


Figure 1. Model and expectations under different imprinting hypotheses. (A) Reconstructed, birth year-specific probabilities of imprinting (representative example specific to USA for cases observed in 2015). Throughout the manuscript, group 1 HA subtypes are represented in blue and group 2 subtypes in red. **(B)** Expected imprinting protection against H1N1 or H3N2 under the three tested models. (**C**) Cartoon of age-specific risk curve. The shape of this curve is purely hypothetical, but each tested model fit a similar step function to data. (**D-F)** Fraction of each birth year unprotected by their childhood imprinting (from A) determines the shape of birth year-specific risk. (**G-I)** A linear combination of age-specific risk (C), and birth year-specific risk (D-F) give the expected age distribution of H1N1 or H3N2 cases under each model.

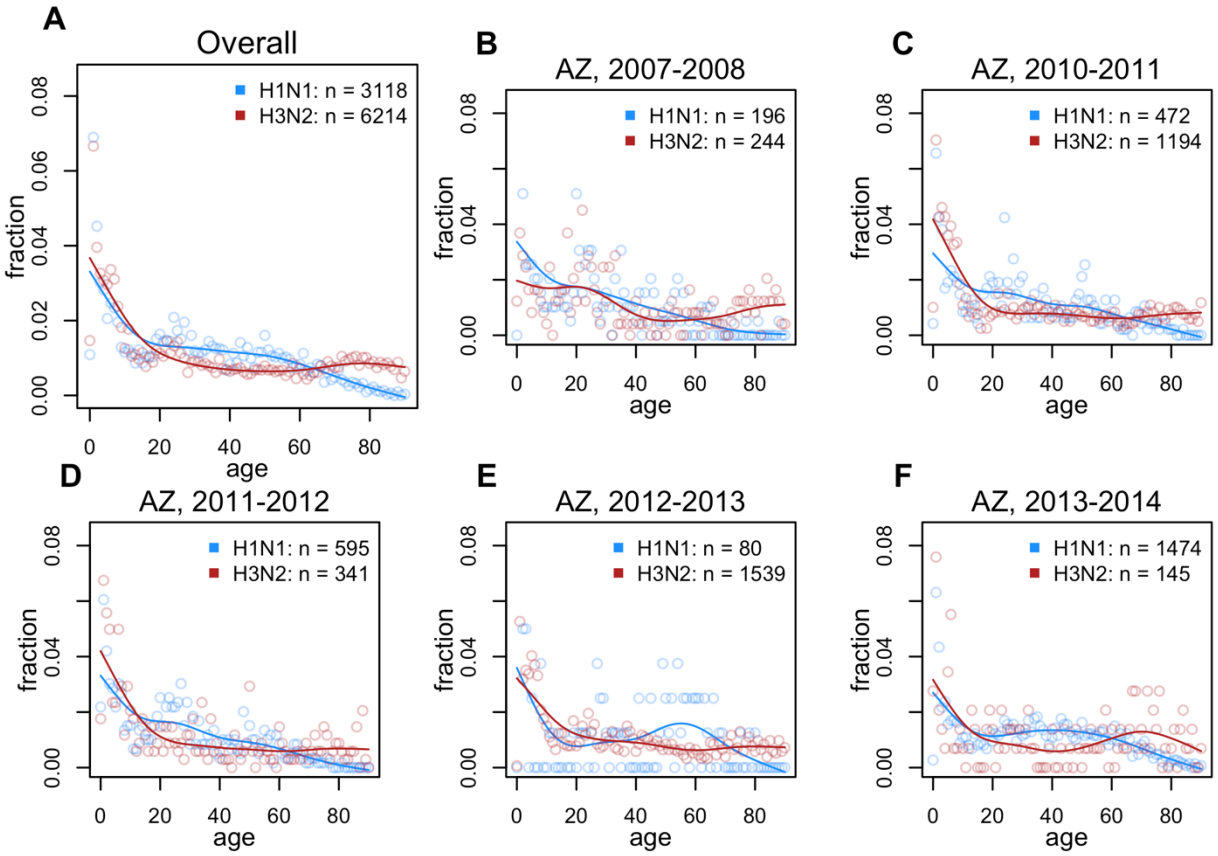


Figure 2. Observed age distributions, Arizona. Points show fraction of confirmed H1N1 or H3N2 cases observed in each single year of age. Lines show a smoothing spline fit to observed distributions. **(A)** All confirmed cases in the data (aggregate across all seasons). (**B-G**) Age distributions from individual seasons in which both H1N1 and H3N2 circulated (seasons with ≥ 50 confirmed cases of each subtype are shown here. See Fig. S1 for all seasons).

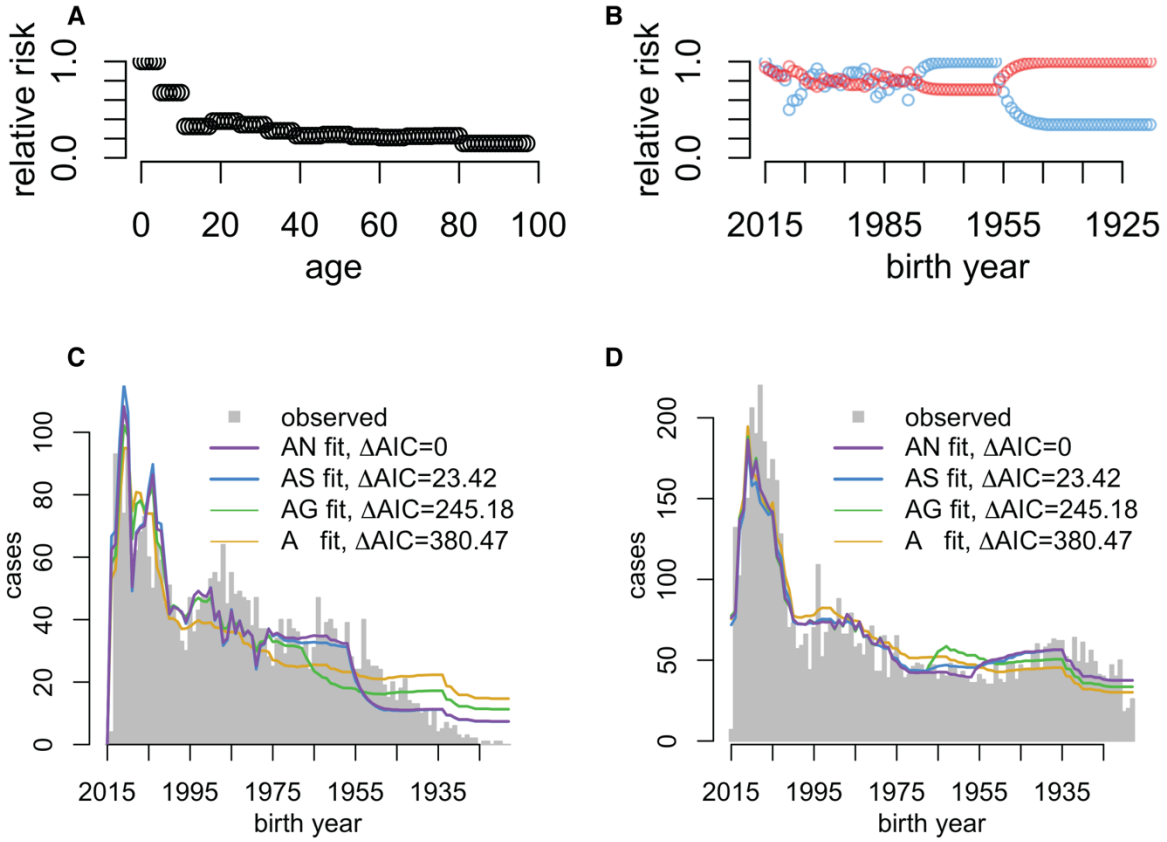


Figure 3. **Model fits and model selection**. **(A)** Fitted effects of age and **(B)** imprinting from model AN, which provided the best fit to data**. (C-D)** Model fits to observed age distributions of H1N1 (**C**) and H3N2 (**D**) cases**.** Model name abbreviations indicate which factors were included: A = age-specific risk, N = NA subtype-level imprinting, S = HA subtype-level imprinting, G = HA group-level imprinting.

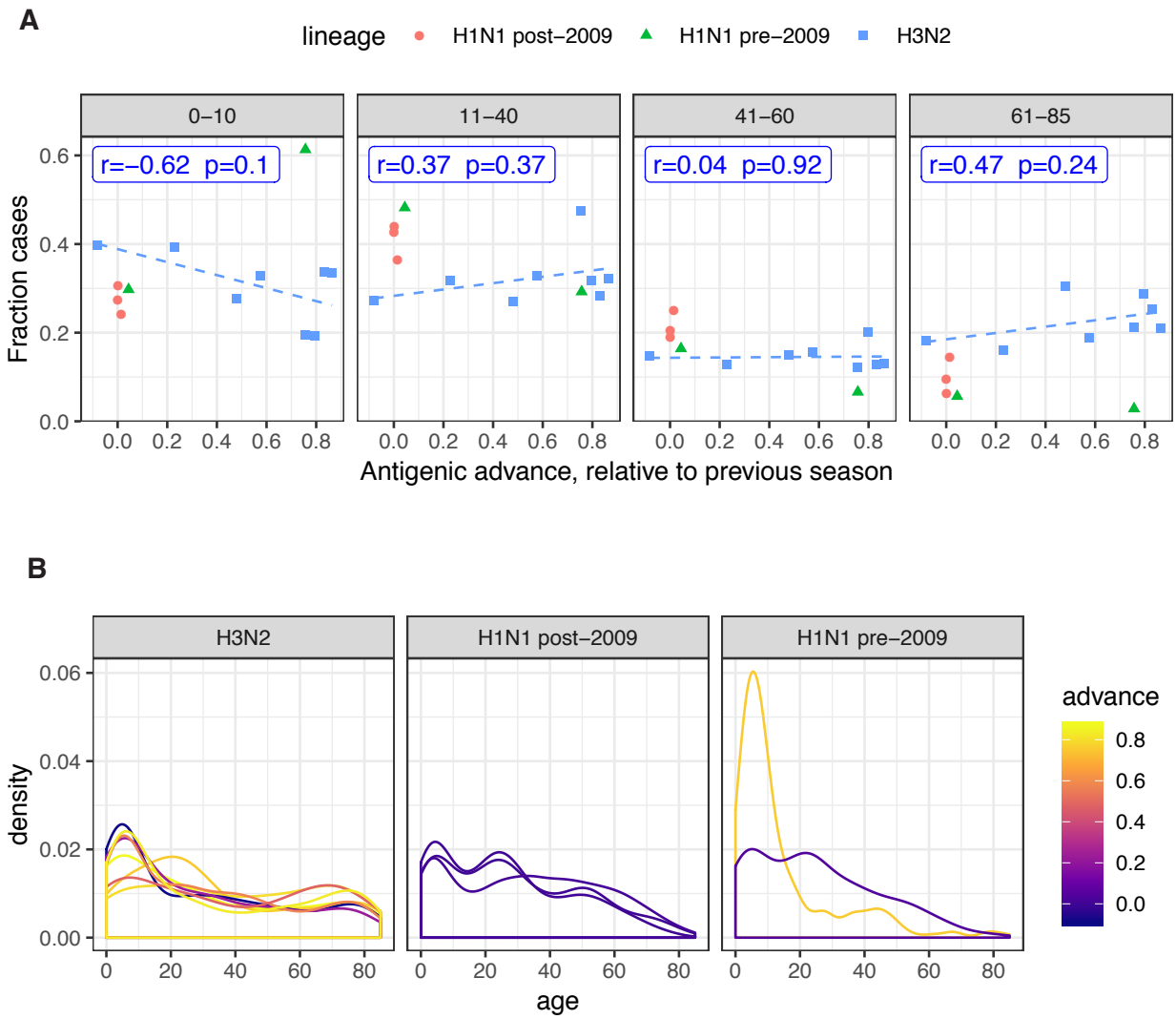


Figure 4. Effect of antigenic advance on age distribution. (A) Relationship between annual antigenic advance and the fraction of cases observed in children (0-10), or in adult age groups. Each data point represents a single influenza season in which at least 100 confirmed cases of a given subtype were observed. Blue label shows Pearson correlation between the fraction of H3N2 cases observed in each age group and annual antigenic advance. Blue dashes show linear trend fitted using lm() in R. **(B)** Season-specific age distributions of infection, colored by antigenic advance since the previous season.

669 # Tables

670 **Table 1. Confirmed cases in surveillance data from Arizona Department of Health Services.** Data
671 representing the first and second waves of the 2009 H1N1 pandemic (2008-2009 and 2009-2010 seasons)
672 were excluded.

Season	Confirmed H1N1	Confirmed H3N2
1993-94	0	101
1994-95	12	38
2002-03	71	8
2003-04	0	71
2004-05	0	131
2005-06	1	321
2006-07	212	28
2007-08	196	244
2010-11	472	1204
2011-12	595	348
2012-13	80	1578
2013-14	1475	151
2014-15	5	2109
Total	**3119**	**6332**

673

674 **Table 2. Maximum likelihood parameter estimates and 95% profile confidence intervals from each**
675 **model fit to ADHS data.** All estimated parameters represent the relative risk of confirmed infection,
676 given the factors listed in the left-hand column. Model name abbreviations specific which factors were
677 included. A = age-specific risk, N = NA subtype-level imprinting, S = HA subtype-level imprinting, G =
678 HA group-level imprinting.
679

Model	AN	AS	AG	A
ΔAIC	0.00	23.42	245.18	380.47
H1N1 impr. protection	0.34 (0.29-0.42)	0.29 (0.24-0.35)	0.67 (0.58-0.78)	
H3N2 impr. protection	0.71 (0.62-0.82)	0.9 (0.78- >1)	0.69 (0.6-0.8)	
Ages 0-4	Reference group: Value fixed to 1			
Ages 5-10	0.68 (0.63-0.74)	0.66 (0.61-0.72)	0.66 (0.62-0.72)	0.62 (0.57-0.68)
Ages 11-17	0.33 (0.3-0.36)	0.31 (0.28-0.34)	0.33 (0.3-0.37)	0.3 (0.28-0.34)
Ages 18-24	0.38 (0.35-0.42)	0.36 (0.32-0.4)	0.39 (0.35-0.43)	0.35 (0.32-0.39)
Ages 25-31	0.34 (0.32-0.38)	0.33 (0.3-0.37)	0.34 (0.31-0.38)	0.31 (0.28-0.35)
Ages 32-38	0.28 (0.26-0.32)	0.26 (0.24-0.3)	0.28 (0.26-0.32)	0.26 (0.24-0.29)
Ages 39-45	0.23 (0.2-0.27)	0.21 (0.18-0.24)	0.24 (0.22-0.28)	0.21 (0.2-0.24)
Ages 46-52	0.24 (0.22-0.28)	0.21 (0.19-0.24)	0.24 (0.22-0.28)	0.23 (0.2-0.26)
Ages 53-59	0.22 (0.2-0.26)	0.2 (0.18-0.23)	0.2 (0.18-0.24)	0.2 (0.18-0.23)
Ages 60-66	0.21 (0.19-0.24)	0.22 (0.2-0.26)	0.19 (0.16-0.22)	0.18 (0.16-0.21)
Ages 67-73	0.22 (0.2-0.26)	0.25 (0.22-0.29)	0.2 (0.18-0.23)	0.19 (0.18-0.22)
Ages 74-80	0.23 (0.2-0.26)	0.25 (0.22-0.3)	0.2 (0.18-0.24)	0.2 (0.18-0.23)
Ages 81+	0.15 (0.14-0.18)	0.17 (0.15-0.2)	0.13 (0.12-0.16)	0.13 (0.12-0.15)

680

Supplementary Figures

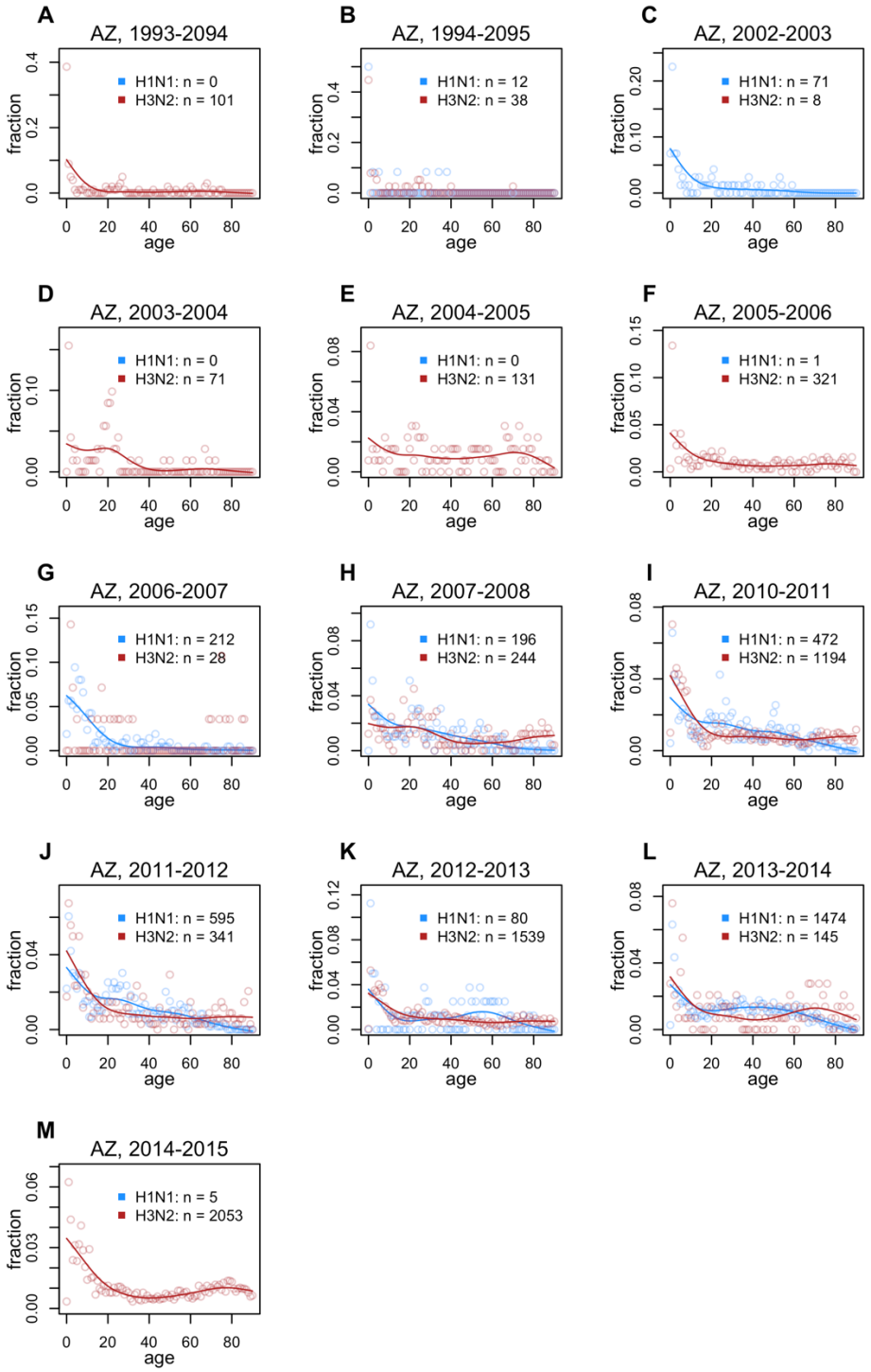


Figure S1. ADHS age distributions, all seasons. Supplement to ***Fig. 2*** showing observed age distributions from all influenza seasons. Observed case fractions (points) were only plotted if 10 or more cases of a given subtype were confirmed, to avoid extreme stretching of the y axis. Smoothing splines 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.

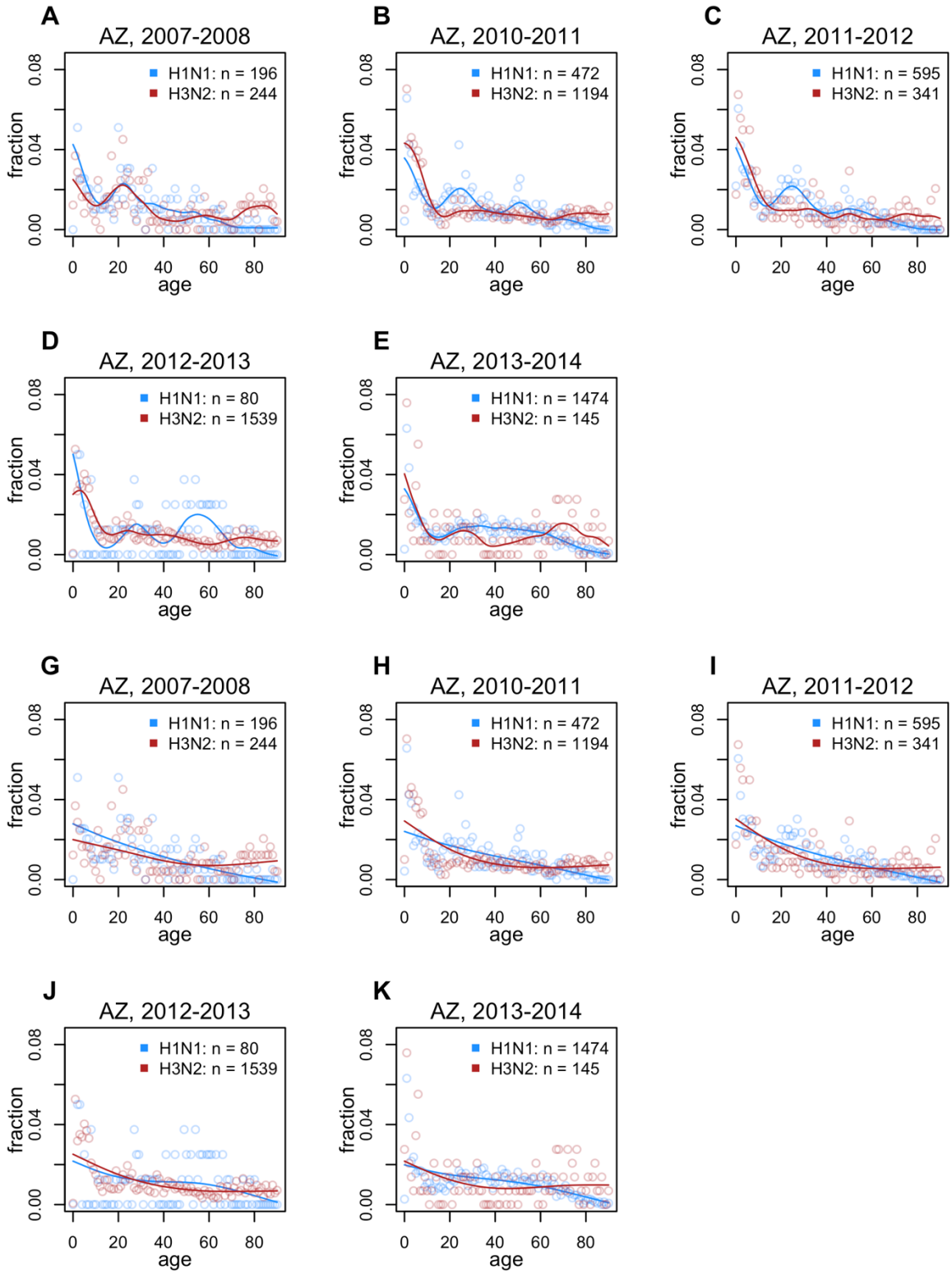
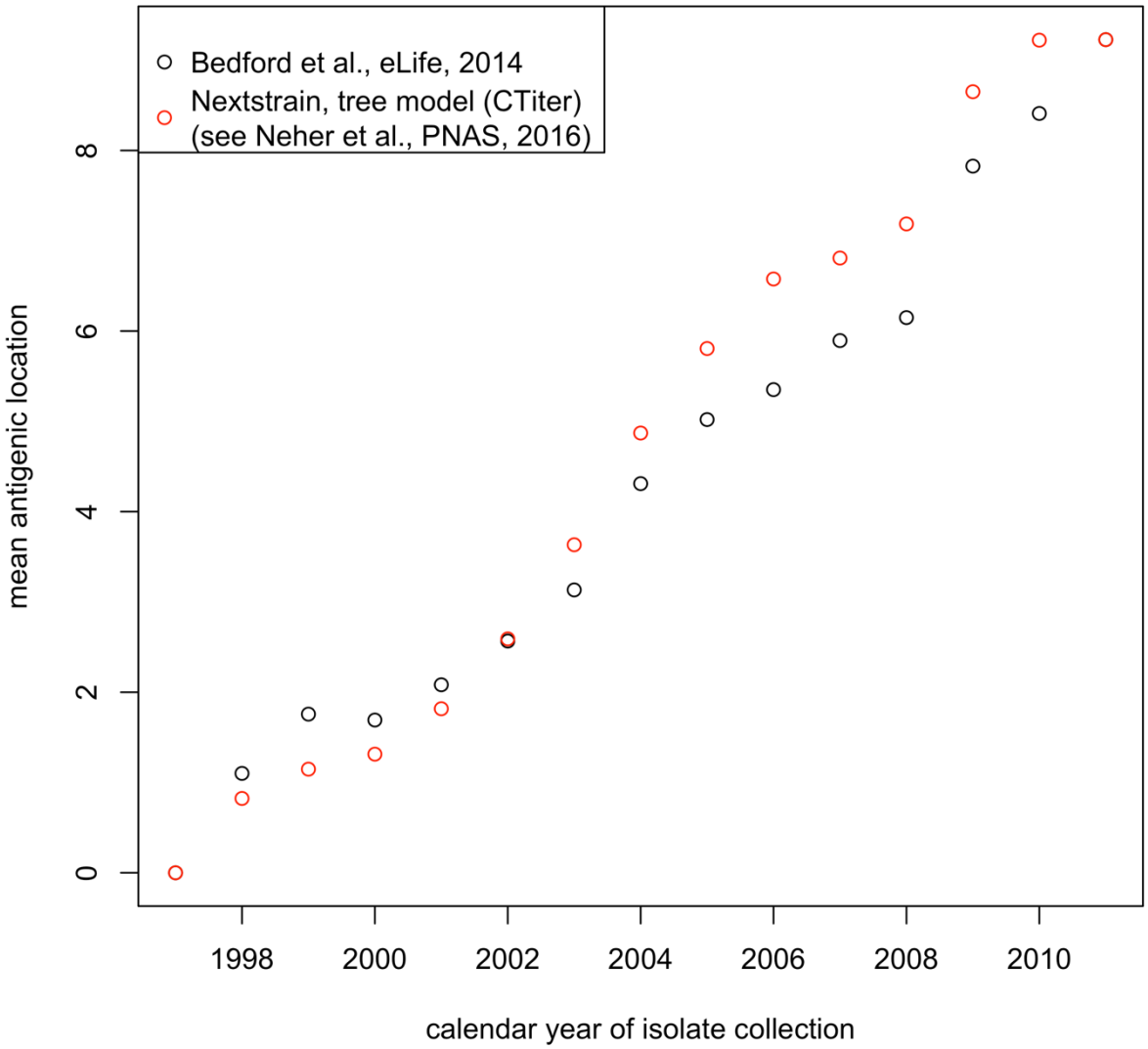


Figure S2. Alternate smoothing parameters, AZDHS data. Supplement to ***Fig. 2***, with smoothing parameters chosen to fit splines that are less (**A-F**), or more (**G-L**) smooth than the splines shown in the main text. Differences between H1N1 and H3N2's age-specific impacts remain evident, especially in the oldest cohorts, regardless of smootness.



694
695 **Figure S3. Comparison of rescaled antigenic distance estimates from the Bedford et al., and**
696 ***Nextstrain* datasets.** Points represent average antigenic position of all isolates from a given calendar year.