1	Childhood immune imprinting to influenza A shapes
2	birth year-specific risk during seasonal H1N1 and H3N2
3	epidemics
4	-
5	Katelyn M Gostic ¹ , Rebecca Bridge ² , Shane Brady ² , Cécile Viboud ³ , Michael Worobey ⁴ , and
6	James O Lloyd-Smith ^{1,3*}
7	
8	
9	¹ Dept. of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles,
10	CA, USA ² Arizona Demontment of Health Services, Dheaniy, A.Z. USA
11	³ Eggarty International Center National Institutes of Health Bethesda MD USA
12	⁴ Dent of Ecology and Evolutionary Biology University of Arizona Tucson AZ USA
14	[†] Current address: Dept. of Ecology and Evolution. University of Chicago. Chicago. IL, USA
15	* illovdsmith@ucla.edu
16	
17	
18	
19	
20	
21	Abstract
22	Across decades of co-circulation in humans, influenza A subtypes H1N1 and H3N2 have caused
23	seasonal epidemics characterized by different age distributions of cases and mortality. H3N2 causes the
24	majority of relatively severe, clinically attended cases in high-risk elderly cohorts, and the majority of
25 26	overall deaths, whereas HINI causes fewer deaths overall, and cases shifted towards young and middle-
20 27	aged adults. These contrasting age promes may result from differences in childhood exposure to HTNT and H3N2 or from differences in evolutionary rate between subtypes. Here we analyze a large
27	enidemiological surveillance dataset to test whether childhood immune imprinting shapes seasonal
29	influenza epidemiology, and if so, whether it acts primarily via immune memory of a particular influenza
30	subtype or via broader immune memory that protects across subtypes. We also test the impact of
31	evolutionary differences between influenza subtypes on age distributions of cases. Likelihood-based
32	model comparison shows that narrow, within-subtype imprinting shapes seasonal influenza risk alongside
33	age-specific risk factors. The data do not support a strong effect of evolutionary rate, or of broadly
34	protective imprinting that acts across subtypes. Our findings emphasize that childhood exposures can
35	imprint a lifelong immunological bias toward particular influenza subtypes, and that these cohort-specific
36	biases shape epidemic age distributions. As a result, newer and less "senior" antibody responses acquired
5/ 20	later in file do not provide the same strength of protection as responses imprinted in childhood. Finally,
38 39	cohorts that lack H1N1-specific imprinting eventually reach old age.

41

Author Summary

42 Influenza viruses of subtype H1N1 and H3N2 both cause seasonal epidemics in humans, but with 43 different age-specific impacts. H3N2 causes a greater proportion of cases in older adults than H1N1, and 44 more deaths overall. People tend to gain the strongest immune memory of influenza viruses encountered 45 in childhood, and so differences in H1N1 and H3N2's age-specific impacts may reflect that individuals 46 born in different eras of influenza circulation have been imprinted with different immunological risk 47 profiles. Another idea is that H3N2 may be more able to infect immunologically experienced adults 48 because it evolves slightly faster than H1N1 and can more quickly escape immune memory. We analyzed 49 a large epidemiological data set and found the clearest signal that birth year-specific differences in 50 childhood immune imprinting, not differences in evolutionary rate, explain differences in H1N1 and 51 H3N2's age-specific impacts. These results can help epidemiologists understand how epidemic risk from 52 specific influenza subtypes is distributed across the population and predict how population risk may shift 53 as differently imprinted birth years grow older. Further, these results provide immunological clues to 54 which facets of immune memory become biased in childhood, and then later play a strong role in 55 protection during seasonal influenza epidemics.

56

Introduction

57 Childhood influenza exposures leave an immunological imprint, which has reverberating, lifelong 58 impacts on immune memory. Foundational work on original antigenic sin (1) and antigenic seniority (2) 59 showed that individuals maintain the highest antibody titers against influenza strains encountered in 60 childhood. But how these serological patterns map to functional immune protection, and shape birth year-61 specific risk during outbreaks, remains an active area of inquiry. One open question is the breadth of 62 cross-protection provided by immune memory imprinted in childhood.

63 We define immune imprinting as a lifelong bias in immune memory of, and protection against, 64 the strains encountered in childhood. Such biases most likely become entrenched as subsequent exposures 65 back-boost existing memory responses, rather than stimulating true *de novo* responses (3). By providing 66 particularly robust protection against certain antigenic subtypes, or clades, imprinting can provide 67 immunological benefits, but perhaps at the cost of equally strong protection against variants encountered 68 later in life. For example, every modern influenza pandemic has spared certain birth cohorts, presumably 69 due to cross-protective memory primed in childhood (4-10). Recently, we showed that imprinting also 70 protects against novel, emerging avian influenza viruses of the same phylogenetic group as the first 71 childhood exposure (9,11). Imprinting may additionally shape birth year-specific risk from seasonal 72 influenza (12-14), but the importance of broadly protective immunity remains unclear in this context. 73 Until recently, narrow cross-protective immunity specific to variants of a single hemagglutinin 74 (HA) subtype has been considered the primary mode of defense against seasonal influenza. Lymphocyte 75 memory of variable epitopes on the HA head (i.e. sites at which hemagglutinin antigens of different 76 subtypes show limited homology) drives this narrow, within-subtype protection, which is the main 77 mechanism of protection from the inactivated influenza vaccine. But a growing body of evidence shows 78 protection may also be driven by memory of other influenza antigens (e.g. neuraminidase, NA) (15–17), 79 or by immune response to conserved epitopes, many of which are found on the HA stalk (11,18-21). 80 (Antibodies that target conserved HA epitopes can provide broad protection across multiple HA subtypes

in the same phylogenetic group (18,20,22), where HA group 1 contains hemagglutinin subtypes H1 and
H2, while group 2 contains H3 (11,19,23). H1, H2 and H3 are the only HA subtypes that have circulated
seasonally in humans since 1918.)

Together, these insights suggest that within a single host, imprinting probably induces multiple levels of bias in immune memory, to both conserved (broadly protective) and variable (narrowly protective) sites on various influenza antigens. The functional role of any single layer of imprinted immune memory depends on both immunodominance hierarchies and epidemic context. Here, we examine which layers of imprinted memory impact risk from seasonal influenza.

89 Within-subtype immunity to HA is known to shape seasonal influenza's epidemiology and 90 evolution (24). But because this type of narrow immunity decays rapidly in the face of antigenic drift, it 91 would not be expected to shape cohort-specific imprinting protection across an entire human lifetime 92 (25,26). Conversely, broad, HA group-level immune memory arises when lymphocytes target conserved 93 HA epitopes. Responses to these conserved epitopes should be stable over time, and can play a strong 94 role in defense against unfamiliar influenza strains (e.g. novel, avian or pandemic subtypes (11,18,20,22,27,28)). Broad, HA group-level responses are not traditionally thought to play a strong role 95 96 in defense against familiar, seasonal influenza subtypes, but have recently been identified as an 97 independent correlate of protection against seasonal influenza (21), and might play a particularly strong 98 role against drifted seasonal strains whose variable HA epitopes have become unrecognizable. Thus, 99 childhood immune imprinting may determine which birth cohorts are primed for effective defense against 100 seasonal strains with conserved HA epitopes characteristic of group 1 or group 2, or with variable HA 101 epitopes characteristic of a particular subtype (H1, H2, etc.). A similar line of reasoning may apply to 102 immunity against NA, although much less attention has been paid to this antigen. 103 Since 1977, two distinct subtypes of influenza A, H1N1 and H3N2, have circulated seasonally in

humans, with striking but poorly understood differences in their age-specific impact (9,12–14,29). These differences could be associated with childhood imprinting: older cohorts were almost certainly exposed to H1N1 in childhood (since it was the only subtype circulating in humans from 1918-1957), and now seem

107 to be preferentially protected against modern seasonal H1N1 variants (9,12-14). Likewise, younger adults 108 have the highest probabilities of childhood imprinting to H3N2 (Fig. 1), which is consistent with 109 relatively low numbers of clinically attended H3N2 cases in these cohorts. Alternatively, differences in 110 the evolutionary dynamics of H1N1 and H3N2 could explain the observed age profiles. Subtype H3N2 111 exhibits slightly faster drift in its antigenic phenotype than H1N1, and as a result, H3N2 may be better 112 able to escape pre-existing immunity in immunologically experienced adults, whereas H1N1 may be 113 relatively restricted to causing cases in immunologically inexperienced children (30). 114 We analyzed a large surveillance data set of relatively severe, clinically attended influenza cases 115 to test whether cohort effects from childhood imprinting primarily act against variable epitopes, only 116 providing narrow cross-protection against closely related HA or NA variants of the same subtype, or 117 against more conserved epitopes, providing broad cross-protection across HA subtypes in the same 118 phylogenetic group (Fig. 1A-B). We fitted a suite of models to data using maximum likelihood and 119 compared models using AIC. In a separate analysis, we considered the hypothesis that differences in 120 evolutionary rate of H1N1 and H3N2, rather than imprinting effects, shape differences in age distribution. 121 Our results have implications for long-term projections of seasonal influenza risk in elderly cohorts (13), 122 who suffer the heaviest burdens of influenza-related morbidity and mortality, and whose imprinting status 123 will shift through time as cohorts born during different inter-pandemic eras grow older. 124 The Data 125 126 The Arizona Department of Health Services (ADHS) provided a dataset containing 9,510 127 seasonal H1N1 and H3N2 cases from their statewide surveillance system. Cases of all ages were 128 confirmed to subtype by PCR and/or culture, primarily from virologic testing at the Arizona State Public 129 Health Laboratory (ASPHL). The ADHS surveillance system aims to characterize circulating strains from 130 patients across the state with medically attended influenza. Although surveillance does not target specific 131 at-risk groups, relatively severe cases (especially those tested at hospital labs) are overrepresented in our

data, as these cases are most likely to be medically attended and confirmed to subtype, thus meeting our
study inclusion criteria. This is because confirmation to subtype requires a second-line test (PCR or
culture); rapid tests are more common, but do not indicate subtype.

135 Although the exact collection setting of individual specimens was not always recorded, ADHS 136 staff internally reviewed reporting source and provider organization of cases in our data to estimate that 137 roughly 76% were reported and/or submitted by hospital labs or may have originated at hospital-138 associated outpatient clinics. A previous ADHS analysis of subtyped 2016 influenza A data matched to 139 hospital discharge data found that nearly half of cases reported from hospital labs (if extrapolated to other 140 seasons, roughly 38% of the overall data) were severe enough to warrant hospital admission. The rest 141 (also roughly 38% of the overall data, if extrapolated) were discharged without admission. To obtain a 142 broader representation of clinically attended cases from across the state, ADHS collaborates with county 143 health departments, commercial laboratories, and outpatient clinics to receive specimens. We estimate 144 that roughly 8% of the data originated in outpatient settings. The remaining 17% of cases were either 145 tested at commercial labs, or were tested at ASPHL, but with unknown origin. Ultimately, these data 146 allow us to analyze drivers of relatively severe, clinically attended cases, but our results cannot be 147 assumed to generalize to mild or asymptomatic cases. 148 Cases were observed across 22 years of influenza surveillance, from the 1993-1994 influenza

150 pandemic (*Table 1*). Sampling changed slightly starting in 2004, when commercial labs were first

151 mandated to report positive tests to the state (31), but the vast majority of cases analyzed (9150/9451)

season through the 2014-2015 season, although sample sizes increased dramatically after the 2009

152 were observed from the 2004-2005 season onwards, after this change had been implemented.

Following CDC standards, ADHS defines the influenza season as epidemiological week 40 (around early October) through week 39 of the following year (32). 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 cases and molecular

- 157 drivers of immune memory differed during the 2009 pandemic from the normal drivers of seasonal
- 158 influenza's immuno-epidemiology of interest to this study (14,18,22). From the dataset of 9,510 seasonal
- 159 cases (defined as any case observed outside the 2008-2009 or 2009-2010 season), we excluded 58 cases
- 160 with birth years before 1918 (whose imprinting status could not be inferred unambiguously), and one case
- 161 whose year of birth was recorded in error. Ultimately, we analyzed 9,541 cases.
- 162
- 163
- 164 165
- 166
- 167
- 168
- 169
- 170
- 171

172	Table	1.

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

173 Confirmed cases in surveillance data from Arizona Department of Health Services. Data

- 174 representing the first and second waves of the 2009 H1N1 pandemic (2008-2009 and 2009-
- 175 2010 seasons) were excluded.
- 176
- 177
- 178

179

The Model

180 **Reconstructed imprinting patterns**

We reconstructed birth year-specific probabilities of childhood imprinting to H1N1, H2N2 or 181 182 H3N2 using methods described previously (11). These probabilities are based on patterns of first 183 childhood exposure to influenza A and reflect historical circulation (Fig. 1A). Most individuals born 184 between pandemics in 1918 and 1957 experienced a first influenza A virus (IAV) infection by H1N1, and 185 middle-aged cohorts born between pandemics in 1957 and 1968 almost all were first infected by H2N2 186 (note that because the first influenza exposure may occur after the first year of life, individuals born in the 187 years leading up to a pandemic have some probability of first infection by the new pandemic subtype, 188 Fig. 1A). Ever since its emergence in 1968, H3N2 has dominated seasonal circulation in humans, and 189 caused the majority of first infections in younger cohorts. However, H1N1 has also caused some seasonal 190 circulation since 1977, and has imprinted a fraction of all cohorts born since the mid-1970s (*Fig. 1A*). 191 Reconstructions assumed children age 0-12 in the year of case observation might not yet have 192 been exposed to any influenza virus. Interactions between imprinting and vaccination of naïve infants are 193 plausible, but poorly understood (11,33). We did not consider childhood vaccination effects here; only a 194 small percentage of individuals in the ADHS data were born at a time when healthy infants were routinely 195 vaccinated against influenza.

196

197

198 Expected age distributions under alternate imprinting models

199 If HA subtype-level imprinting protection shapes seasonal influenza risk, primary exposure to 200 HA subtype H1 or H3 in childhood should provide lifelong protection against modern variants of the 201 same HA subtype. If imprinting protection acts primarily against specific NA subtypes, lifelong 202 protection will be specific to N1 or to N2 (*Fig. 1B*). Alternatively, if broad HA group-level imprinting 203 shapes seasonal influenza risk, then cohorts imprinted to HA subtype H1 or H2 (both group 1) should be

protected against modern, seasonal H1N1 (also group 1), while only cohorts imprinted to H3 (group 2)
would be protected against modern, seasonal H3N2 (also group 2) (*Fig. 1B*). Collinearities between the
predictions of different imprinting models (*Fig. 1D-1*) were inevitable, given the limited diversity of
influenza antigenic subtypes circulating in humans over the past century (reflected in *Fig. 1A*). Note that
middle-aged cohorts, which were first infected by H2N2, are crucial, because they provide the only
leverage to differentiate between imprinting at the HA subtype, NA subtype or HA group-level level (*Fig. 1B*).

211 Our approach distinguishes between age-specific risk factors related to health and social 212 behavior, and birth year-specific effects related to imprinting. Specifically, age-specific risk could be 213 influenced by medical factors like age-specific vaccine coverage, age-specific risk of severe disease, age-214 related changes in endocrinology and immunosenescence, or by behavioral factors like age-assorted 215 social mixing, and age-specific healthcare seeking behavior. These factors should have similar impacts on 216 any influenza subtype. In contrast, imprinting effects are subtype-specific. Thus, we fit a step function to 217 characterize the shape of age-specific risk of any confirmed influenza case. Simultaneously, we modeled 218 residual, subtype-specific differences in risk as a function of birth year, to focus on the possible role of 219 childhood imprinting. Each tested model used a linear combination of age-specific risk (Fig. 1C) and 220 birth year-specific risk (Fig. 1D-F) to generate an expected distribution of H1N1 or H3N2 cases (Fig. 221 **1G-I**). Note that for a given birth cohort, age-specific risk changed across progressive years of case 222 observation (as the cohort got older), whereas birth year-specific risk was constant over time.

To test quantitatively whether observed subtype-specific differences in incidence were most consistent with imprinting at the HA subtype, NA subtype or HA group level, or with no contribution of imprinting, we fitted a suite of models to each data set using a multinomial likelihood and then performed model selection using 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 (34,35). Technical details are provided in the *Methods*.

229

230 Tested models

231	We fit a set of four models to the ADHS data set. The simplest model contained only age-specific
232	risk (abbreviated A), and more complex models added effects from imprinting at the HA subtype level
233	(S), at the HA group level (G), or at the NA subtype level (N): abbreviated AS, AG, and AN,
234	respectively. The age-specific risk curve took the form of a step function, in which relative risk was fixed
235	to 1 in age bin 0-4, and one free parameter was fit to represent relative risk in each of the following 12
236	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
237	models that contained imprinting effects, the fraction of individuals in each single year of birth with
238	protective childhood imprinting was assumed proportional to reductions in risk. Two additional free
239	parameters quantified the relative risk of a confirmed H1N1 or H3N2 case, given imprinting protection
240	against that seasonal subtype.
241	
242	Effect of influenza evolutionary rate on age profiles
243	We used publicly available data from Nextstrain (36,37), and from one previously published
244	study (38), to calculate annual antigenic advance, which we defined as the antigenic distance between
245	strains of a given lineage (pre-2009 H1N1, post-2009 H1N1 or H3N2) that circulated in consecutive
246	seasons (Methods). The "antigenic distance" between two influenza strains is used as a proxy for
247	similarity in antigenic phenotype, and potential for immune cross-protection. A variety of methods have
248	been developed to estimate antigenic distance using serological data, genetic data, or both (37-39).
249	To assess the impact of antigenic evolutionary rate on the epidemic age distribution, we tested
250	whether the proportion of cases in children increased in seasons associated with large antigenic changes.
251	If the rate of antigenic drift is a strong driver of age-specific influenza risk, then the fraction of influenza
252	cases observed in children should be negatively related to annual antigenic advance (30). In other words,
253	strains that have not changed much antigenically since the previous season should be unable to escape
254	pre-existing immunity in immunologically experienced adults, and more restricted to causing cases in

255	immunologically inexperienced children; strains that have changed substantially will be less restricted to
256	children.
257	
258	Results
259	Subtype-specific differences in age distribution
260	Seasonal H3N2 epidemics consistently caused more clinically attended cases in older cohorts,
261	while H1N1 caused a greater proportion of cases in young and middle-aged adults (Figs. 2, S1-S2). These
262	patterns were apparent whether we compared H3N2 epidemic age distributions with those caused by the
263	pre-2009 seasonal H1N1 lineage, or with the post-2009 lineage. Observed patterns are consistent with the
264	predicted effects of cohort-specific imprinting (Fig. 1), and with previously reported differences in age
265	distribution of seasonal H1N1 and H3N2 incidence (12–14,29). See <i>Fig. 2</i> for seasons where H1N1 and
266	H3N2 co-circulated in substantial numbers, and Figs. S1-S2 for the entire dataset and alternate smoothing
267	parameters.
268	
269	Imprinting model selection
270	The data showed a strong preference for NA subtype-level imprinting over HA subtype-level
271	imprinting ($\Delta AIC=34.54$), and effectively no statistical support for broad, HA group-level imprinting
272	(Δ AIC=249.06), or for an absence of imprinting effects (Δ AIC=385.42) (<i>Fig. 3, Table 2</i>). Visual
273	assessment of model fits (Fig. 3C-D) confirmed that models containing imprinting effects at the narrow,
274	NA or HA subtype levels provided the best fits to data. The lack of support for the no-imprinting model
275	suggests that imprinting from the first exposure shapes lifelong seasonal influenza risk, just as it does
276	avian-origin influenza (10, 12). However, imprinting appears to act more narrowly against seasonal
277	influenza than against avian influenza, providing cross protection only to a specific NA or HA subtype,
278	instead of broader, HA group-level protection. This result is consistent with the idea that

279 immunodominance of variable HA epitopes limits the breadth of immune cross protection deployed

against familiar, seasonal influenza subtypes (20,22).

As expected (see *Fig. 1G-I*), predictions from the two best models were highly collinear, except

in their risk predictions among middle-aged, H2N2-imprinted cohorts (birth years 1957-1968), and some

283 other minor differences arising from normalization across birth-years.

284

285

286

287 **Table 2**.

Model	AN	AS	AG	А
ΔΑΙC	0.00	34.54	249.06	385.42
H1N1 impr. protection	0.36 (0.30-0.44)	0.29 (0.24-0.35)	0.65 (0.56-0.76)	
H3N2 impr. protection	0.66 (0.58-0.76)	0.90 (0.78-1.04)	0.70 (0.62-0.82)	
Ages 0-4		Reference group:	Value fixed to 1	
Ages 5-10	0.67 (0.62-0.73)	0.65 (0.60-0.70)	0.65 (0.60-0.71)	0.61 (0.56-0.66)
Ages 11-17	0.33 (0.30-0.37)	0.30 (0.28-0.34)	0.32 (0.30-0.36)	0.29 (0.27-0.33)
Ages 18-24	0.37 (0.34-0.42)	0.34 (0.32-0.38)	0.37 (0.34-0.42)	0.34 (0.31-0.38)
Ages 25-31	0.35 (0.32-0.40)	0.33 (0.30-0.38)	0.34 (0.32-0.38)	0.32 (0.29-0.36)
Ages 32-38	0.3 (0.28-0.35)	0.28 (0.26-0.32)	0.3 (0.27-0.34)	0.27 (0.25-0.31)
Ages 39-45	0.25 (0.22-0.30)	0.22 (0.20-0.26)	0.25 (0.22-0.29)	0.23 (0.21-0.26)
Ages 46-52	0.27 (0.24-0.30)	0.22 (0.20-0.26)	0.26 (0.23-0.29)	0.24 (0.22-0.28)
Ages 53-59	0.25 (0.23-0.30)	0.22 (0.20-0.26)	0.23 (0.21-0.27)	0.23 (0.20-0.26)
Ages 60-66	0.27 (0.24-0.30)	0.29 (0.26-0.33)	0.24 (0.22-0.28)	0.23 (0.21-0.27)
Ages 67-73	0.37 (0.33-0.43)	0.42 (0.37-0.48)	0.34 (0.30-0.38)	0.33 (0.30-0.38)
Ages 74-80	0.57 (0.50-0.64)	0.64 (0.57-0.74)	0.52 (0.46-0.59)	0.5 (0.46-0.57)
Ages 81+	0.99 (0.88-1.11)	1.12 (1.00-1.26)	0.9 (0.81-1.01)	0.87 (0.80-0.96)

288

289 Maximum likelihood parameter estimates and 95% profile confidence intervals from each

model fit. All estimated parameters represent the relative risk of a confirmed case, given the factors listed in the left-hand column. Age-specific risk parameters could take any value, but we only considered imprinting protection values between 0 and 1, as it would be illogical for protective imprinting to cause an increase in relative risk. Model name abbreviations specific which factors were included. A = age-specific risk after normalizing to demographic age distribution, N = NA subtype-level imprinting, S = HA subtype-level imprinting, G = HA group-

296 level imprinting.

297

299 Fitted risk patterns

300	Fitted age-specific risk curves took similar forms in all tested models. After controlling for
301	demographic age distribution, estimated age-specific risk was highest in children and the elderly,
302	consistent with the buildup of immune memory across childhood, and waning immune function in the
303	aged (Fig. 3A shows the fitted curve from the best model). Estimates of imprinting parameters were less
304	than one, indicating some reduction in relative risk (Table 2). Within the best model, estimated reductions
305	in relative risk from childhood imprinting were stronger for H1N1 (0.34, 95% CI 0.29-0.42) than for
306	H3N2 (0.71, 95% CI 0.62-0.82). In the second-best model, AS (subtype-specific imprinting), estimated
307	reductions in H3N2 risk were particularly weak, and the confidence interval overlapped the null value of
308	1. <i>Table 2</i> shows parameter estimates and 95% profile confidence intervals from all models fitted.
309	
310	
311	Effect of evolutionary rate
312	To test for effects of evolutionary rate on epidemic age distribution, we searched for decreases in
313	the proportion of cases among children in seasons associated with antigenic novelty, when highly drifted
314	strains might be more able to infect immunologically experienced adults. (We defined children as ages 0-
315	10, and verified internally that our analysis of evolutionary rate was insensitive to our exact choice of age
316	range for children). Consistent with this expectation, the data showed a slight negative but not significant
317	association between annual antigenic advance and the fraction of H3N2 cases observed in children (Fig.
318	4A). However, note that no clear relationship emerged between antigenic novelty and the fraction of cases
319	observed in any age group including older children (>10) and adults (<i>Fig. 4A</i>). These are the cohorts in
320	which epidemiological data show the clearest differences between H1N1 and H3N2's age-specific
321	impacts (<i>Fig. 2</i>); if rate of antigenic evolution is a dominant driver of age-specific differences in
322	incidence, we would have expected to see clearer evidence of evolutionary rate effects within adults
323	cohorts, not just between adults and the youngest children. The data contained too few influenza seasons

324	with sufficient numbers of confirmed H1N1 cases to support meaningful Spearman correlation
325	coefficients for either pre-2009 or post-2009 seasonal H1N1 lineages.
326	Furthermore, if evolutionary rate is the dominant driver of subtype-specific differences in
327	epidemic age distribution, then when subtypes H1N1 and H3N2 show similar degrees of annual antigenic
328	advance, their age distributions of cases should appear more similar. However, the data showed that
329	differences in H1N1 and H3N2's age-specific impacts did not converge when lineages showed similar
330	annual advance. When comparing the fraction of cases observed in specific age classes, H1N1 data
331	consistently clustered separately from H3N2, with H1N1 consistently causing fewer cases at the extremes
332	of age (children 0-10 and elderly adults 71-85), but more cases in middle-aged adults, regardless of
333	antigenic novelty (Fig. 4A). Smoothed density plots showed no clear relationship between annual
334	antigenic advance and age distribution (Fig. 4B). Overall, the data showed a weak, but not significant
335	signal that relatively severe, clinically attended cases may be more restricted to young children when
336	antigenic novelty is low, but the data did not show strong evidence that the magnitude of annual antigenic
337	drift is a systematic driver of epidemic age distribution across the entire population.
220	

- 338
- 339

2	Λ	n
Э	4	·U

Discussion

341	We analyzed a large epidemiological surveillance dataset and found that seasonal influenza
342	subtypes H1N1 and H3N2 cause different age distributions of relatively severe, clinically attended cases,
343	confirming previously reported patterns (12-14). We analyzed several possible drivers of these
344	differences side-by-side, and found the greatest support for imprinting protection against seasonal
345	influenza viruses of the same NA or HA subtype as the first influenza strain encountered in childhood
346	(12,13). The data did not support strong effects from broader HA group-level imprinting, as recently
347	detected for novel zoonotic or pandemic viruses (9,11), or from differences in rates of antigenic evolution
348	(30). Our results suggest individuals retain a lifelong bias in immune memory, and that this imprint is not
349	erased even after decades of exposure to or vaccination against dissimilar influenza subtypes.
350	External evidence corroborates the idea that birth year, rather than age, drives subtype-specific
351	differences in seasonal influenza risk. When H3N2 first emerged in 1968, it caused little or no excess
352	mortality in the elderly, who had putatively been exposed, as children or young adults, to an H3 virus that
353	had circulated in the late 1800s (7,9). Meanwhile, H1N1-imprinted cohorts (those ~10-50 years old at the
354	time) experienced considerable excess mortality in the 1968 pandemic (7). Now, fifty years later, the
355	same H1N1-imprinted cohorts continue to experience excess H3N2 morbidity and mortality as older
356	adults (12–14,29) (<i>Fig. 2</i>).

357 In model comparison, the data supported childhood imprinting to NA. Although NA is not as 358 intensively studied as HA, these results emphasize the increasingly recognized importance of both 359 antigens as drivers of protection against seasonal influenza (15-17). Realistically, some combination of 360 effects from both HA and NA subtype-level imprinting probably shapes seasonal influenza risk; both 361 models of imprinting produced similar fits to data, and far outperformed other models in terms of AIC 362 (Fig. 3). Unfortunately, due to the limited diversity of seasonal influenza strains that have circulated in 363 humans over the past century, collinearities between even the relatively simple models tested here 364 prevented us from testing more complicated models of combined effects from imprinting to multiple

365 antigens. Deeper insights into the respective roles of HA and NA will most likely need to come from 366 focused immunological cohort studies, in which individual histories of influenza infection are recorded 367 and can be studied alongside changes in serology, PBMCs, and/or the B cell repertoire (33). 368 Alternatively, the development of immunological biomarkers for diagnosis of imprinting status in 369 individual patients could substantially increase the power of epidemiological inference. 370 We did not detect a clear relationship between annual antigenic advance and epidemic age 371 distribution, although small sample sizes may have limited our statistical power. We did detect a weak 372 trend, consistent with the idea that influenza cases are more restricted to immunologically inexperienced 373 children in seasons of low antigenic advance, as previously proposed (30). But the data did not reveal a 374 clear relationship between antigenic advance and the fraction of cases occurring in adult age groups, 375 where epidemiological data reveal distinct subtype-specific differences in impact. Perhaps antigenic 376 advance shapes how cases are distributed between children and adults, but has small or inconsistent 377 impacts within the adult population. We speculate that clearer relationships between antigenic advance 378 and epidemic age distribution might emerge if methods to estimate antigenic distance were able to 379 incorporate effects such as immune history (40), glycosylation (40,41), and immunity to antigens other 380 than HA (16,17,42).

381 The exact immunological drivers of imprinting protection against seasonal influenza remain 382 unclear, but these results provide some new clues. Traditionally, within-subtype cross-protection is 383 thought to decay quickly with antigenic drift. Strains that circulated more than 14 years apart rarely show 384 measurable cross-protective titers by the hemagglutination inhibition (HI) assay (38). The short timescale 385 of immune memory to variable HA head epitopes stands in contrast to patterns observed in our study and 386 others (12–14), where within-subtype immune memory imprinted in childhood appears to persist for an 387 entire human lifetime, remaining evident even in the oldest cohorts. We speculate that within-subtype 388 imprinting protection may involve epitopes that are more conserved, and stable over time, than those 389 typically measured in HI assays (which inform most existing estimates of antigenic distance, but may 390 disproportionately measure antibodies to variable, immunodominant epitopes on the HA head (21,22)).

Across a lifetime of exposures to diverse H1N1 and H3N2 variants, repeated back-boosting of antibodies to intermediately conserved sites on HA or NA (i.e. sites conserved within but not across HA and NA subtypes), could explain the longevity of subtype-level imprinting protection. This is consistent with recent evidence that the immune repertoire shifts to focus on more conserved influenza epitopes as we age (25,26).

396 Another possibility is that, memory B cell clones developed during the first childhood influenza 397 exposure may later adapt via somatic hypermutation to "follow" antigenic targets as they drift over time 398 (25,28). However, this would be inconsistent with new evidence suggesting memory B cells are relatively 399 fixed in phenotype, and have little potential for ongoing affinity maturation (43,44). Thus, the first 400 influenza exposure in life may fill a child's memory B cell repertoire with clones that serve in the future 401 as prototypes that can be rapidly and effectively tailored to recognize drifted influenza strains of the same 402 subtype. Finally, the role of CD4+ T cells in imprinting is unclear, but T cell memory and T cell help to B 403 cells within germinal centers both play at least some role in the development of the immune repertoire 404 (45).

405 Signals of imprinting protection are anomalously strong in the current cohort of elderly adults, as 406 reflected by higher estimates of imprinting protection to H1N1 than H3N2. The oldest subjects in our 407 data, born slightly after 1918, and would not have encountered an influenza virus of any subtype other 408 than H1N1 until roughly age 30. Repeated early-life exposures to diverse H1N1 variants may have 409 reinforced and expanded the breadth of H1N1-specific immune memory (5,46). But this strong H1N1 410 protection seems to come at a cost; even after decades of seasonal H3N2 exposure, and vaccination, older 411 cohorts have evidently failed to develop equally strong protection against H3N2. HA group 1 antigens 412 (e.g. H1) appear to induce narrower immune responses, and less cross-group protection than structurally 413 distinct HA group 2 antigens (e.g. H3) (23). Perhaps elderly cohorts imprinted to group 1 antigens have 414 been trapped in narrower responses that offer exceptional protection against strains similar to that of first 415 exposure but relatively poor adaptability to other subtypes.

416 We speculate that imprinting protection, which currently limits the number of severe, clinically-417 attended H1N1 cases in the elderly, also limits the mortality impact of H1N1 viruses. Although pre- and 418 post- 2009 H1N1 lineages have caused slightly different profiles of age-specific mortality (13), neither 419 H1N1 lineage causes nearly as many deaths as H3N2 in high-risk elderly cohorts (13,29,47). On the one 420 hand, if strong subtype-specific biases from imprinting remain in future cohorts of elderly adults, our 421 results would corroborate the idea that mortality from H1N1 may increase as protection in the elderly 422 shifts from H1N1 toward other subtypes (9,13). On the other hand, given that cohorts born after 1968 423 have had much more varied early life exposures to both H1N1 and H3N2, these cohorts may show a 424 greater ability to act as immunological generalists as they become elderly, capable of effective defense 425 against multiple subtypes. 426 Our study has several limitations. Relatively severe, clinically attended cases are much more 427 likely to be detected, confirmed to subtype, and included in our data than mild cases. Thus, while our 428 results show a clear relationship between subtype-level imprinting and risk of relatively severe, clinically 429 attended influenza, the relationship between imprinting and mild or asymptomatic cases could not be 430 determined from available data. 431 Given the limited number of variables recorded in the data, we could not model explicitly the 432 impact of individual risk factors such as the presence of comorbidities, patient sex, or vaccination status. 433 All these factors are known to shape immunity and influenza risk (48), and all may cause individual 434 imprinting outcomes to vary from the average, population trends measured by our study. Understanding 435 how these patient-level covariates modulate imprinting and other aspects of immunity is the next frontier 436 in this line of research. For now, working within the constraints of the available data, we designed the 437 age-specific risk component of the model to capture empirically the combined effects of several risk 438 factors that could not be modeled individually. Additionally, we analyzed the relative count of H1N1 to 439 H3N2 cases within each single year of birth, not absolute incidence, to control for minor age-specific 440 biases in sampling, which are almost inevitably present in any large data set.

441 Another limitation was the low number of confirmed cases available in the pre-2009 era. Large, 442 detailed data sets collected continuously over decades provide the greatest power to separate the effects of 443 age from birth year. We emphatically echo earlier calls (49) for more systematic sharing of single year-of-444 age influenza surveillance data, standardization of sampling effort, and reporting of age-specific 445 denominators, which could substantially boost the scientific community's ability to link influenza's 446 genetic and antigenic properties with epidemiological outcomes. Additionally, collection and reporting of 447 covariates such as sex, vaccination status and the presence of comorbidities in surveillance data would 448 help us understand how patient-level variables modulate imprinting, and immunity in general (50,51). 449 Altogether, this analysis confirms that the epidemiological burden of H1N1 and H3N2 is shaped 450 by cohort-specific differences in childhood imprinting (9,12,13,52), and that this imprinting acts at the 451 HA or NA subtype level against seasonal influenza. The lack of support for broader, HA group-level 452 imprinting effects emphasizes the consequences of immunodominance of influenza's most variable 453 epitopes, and the difficulty of deploying broadly protective memory B cell responses against familiar, 454 seasonal strains. Overall, these findings advance our understanding of how antigenic seniority shapes 455 cohort-specific risk during epidemics. The fact that elderly cohorts show relatively weak immune 456 protection against H3N2, even after living through decades of seasonal exposure to or vaccination against 457 H3N2, suggests that antibody responses acquired in adulthood do not provide the same strength or 458 durability of immune protection as responses primed in childhood. Immunological experiments that 459 consider multiple viral exposures, and cohort studies in which individual histories of influenza infection 460 are tracked from birth, promise to illuminate how B cell and T cell memory develop across a series of 461 early life exposures. In particular, these studies may provide clearer insights than epidemiological data 462 into which influenza antigens, epitopes and immune effectors play the greatest role in immune imprinting, 463 and how quickly subtype-specific biases become entrenched across the first or the first few exposures. 464

Materials and Methods 466 Estimation of age from birth year in ADHS data 467 468 The data contained three variables, influenza season, birth year and confirmed subtype. For most 469 cases, birth year was extracted directly from the reported date of birth in patient medical records, but age 470 was not known. We estimated patient age at the time case observation using the formula [year of 471 observation]-[birth year]. To ensure that the minimum estimated age was 0, the second year in the 472 influenza season of case observation was considered the calendar year of observation (e.g. 2013 for the 473 2012-2013 season). 474 475 **Splines** 476 In *Figure 2*, smoothing splines were fit to aid visual interpretation of noisy data. We fit splines 477 using the command *smooth.spline(x = AGE, v = FRACTIONS, spar = 0.8)* in R version 3.5.0. Variables 478 AGE and FRACTIONS were vectors whose entries represented single years of age, and the fraction of 479 cases observed in the corresponding age group. The smoothing parameter 0.8 was chosen to provide a 480 visually smooth fit. Alternative smoothing parameter choices (0.6 & 1.0) are shown in Figs. S1-S2. 481 Although the choice of smoothing parameter changed the shape of each fitted spline, qualitative 482 differences between splines fitted to H1N1 or H3N2 were insensitive. 483 484 **Model formulation** 485 For each unique season in which cases were observed, define p as a vector whose entries 486 represent the expected probability that a randomly drawn H1N1 or a randomly drawn H3N2 case was 487 observed in an individual born in year b. Each model defined, p as a linear combination of age-specific 488 risk, birth year-specific risk (i.e. imprinting effects). All tested models were nested within the equation:

490
$$p = DA * \mathbf{1}_{H1N1}(I_{H1N1}) * \mathbf{1}_{H3N2}(I_{H3N2})$$
 1

491

492	To include risk factors that only modulated risk from one subtype, we included indicator
493	functions I_{H1N1} and I_{H3N2} , which took value 1 if p described the expected age distribution of H1N1 or
494	H3N2 cases, respectively, and 0 otherwise.
495	
496	
497	Demographic age distribution (D)
498	The population of Arizona aged slightly across the study period, so we controlled for shifting
499	demography in all tested models. Demographic age distribution was obtained from intercensal estimates
500	of total population (both sexes) for the state of Arizona, based on the 2000 and 2010 census (53). The US
501	Census Bureau reports population estimates for ages 0-84, but only provides an aggregate estimate for
502	ages 85+. We impute the number of individuals in each single year of age over 85 using a linear model fit
503	to data on age 75-84, with a minimum threshold of 1000 individuals per single year of age. State-specific
504	population estimates were not available prior to the 2000 census, so we substituted estimates from the
505	year 2000 for cases observed in the 1993-94, and the 1994-95 seasons. Vector D represented the fraction
506	of the total population at the time of case observation that fell in a given birth year.
507	
508	Age-specific risk (A)
509	Age-specific risk was defined as a step function, in which relative risk was fixed to value 1 in an
510	arbitrarily chosen age bin, and then z-1 free parameters, denoted r_2 to r_z , were fit to describe relative risk
511	in all other age bins. Below, I_i are indicator functions specifying whether each vector entry is a member
512	of age bin <i>i</i> .
513	$A = 1_1 + 1_2 r_2 + \cdots 1_z r_z $
514	
515	To obtain the predicted fraction of cases observed in each single year of birth, we normalized so

516 that the product of vectors representing demographic age distribution, and age-specific risk, (DA in

517	equation 1) summed to 1. Thus, vector DA can be interpreted as the expected distribution of cases of any
518	influenza case (either subtype), in the absence of birth year-specific biases from imprinting.

- 519
- 520 Imprinting (I)

521 An indicator function defined whether a given prediction vector described risk of confirmed 522 H1N1 or H3N2. Let f_{IHxNy} be vectors describing the fraction of cases of each birth year that were protected 523 against strain HxNy by their childhood imprinting. We defined r_{IHxNy} as free parameters describing the risk 524 of confirmed HxNy, given imprinting protection. Finally, the factor describing the effect of imprinting (I) 525 was defined as:

3

4

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

526

527

528

- 529

530 Likelihood

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

538
$$\mathcal{L} = \prod_{s} \frac{n_{s}!}{x_{0s}! \dots x_{ms}!} p_{0s}^{x_{0s}} \dots p_{ms}^{x_{ms}}$$

- 539
- 540
- 541

542 Model fitting and model comparison

We fit models containing all possible combinations of the above factors to influenza data from each season in the data. We simultaneously estimated all free parameter values using the optim() function in R, with method L-BFGS-B. Imprinting parameters could take values in [0,1], representing the possibility of a reduction in risk. Age-specific risk parameters could take any value greater than 0. We calculated likelihood profiles and 95% profile confidence intervals for each free parameter. Confidence intervals were defined using the method of likelihood ratios (34).

549

550 Antigenic advance

551 We obtained antigenic distance estimates from Nextstrain (nextstrain.org) (36,54), and from 552 source data from Figure 3 in Bedford et al. (38). Nextstrain calculates antigenic distance using genetic 553 data from GISAID (55), and using methods described by Neher et al. (37). We analyzed "CTiter" 554 estimates from *Nextstrain*, which correspond to Neher et al.'s "tree model" method, and are most directly 555 comparable to pre-2009 H1N1 estimates from (38). We repeated analyses using estimates from the 556 similar "substitution model" method and verified that our choice of antigenic distance metric did not 557 meaningfully impact our results. The negative Spearman correlation between antigenic advance and 558 proportion of cases in children was lower, but still non-significant when using the substitution model 559 (p=0.06); all other differences were unremarkable. Datasets from *Nextstrain* and Bedford et al. both 560 contained redundant antigenic distance estimates for the H3N2 lineage, but only Bedford et al. analyzed 561 the pre-2009 H1N1 lineage, and only *Nextstrain* data analyzed the post-2009 H1N1 lineage. The antigenic 562 distance estimates reported by Bedford et al. were roughly proportional to those reported on Nextstrain, 563 but greater in absolute magnitude (37). To enable visualization of all three lineages on the same plot axes, we rescaled pre-2009 H1N1 estimates from Bedford et al. using the formula $d_{Nextstrain} = 0.47 d_{Bedford}$. The 564 565 scaling factor was chosen so that directly-comparable H3N2 distance estimates obtained using each 566 method spanned the same range (Fig. S3). The Nextstrain data files used in this analysis are archived 567 within our analysis code.

568	Acknowledgements
569	We are grateful to Ken Komatsu and Kristen Herrick for their assistance with data access, and to
570	Trevor Bedford for assistance accessing and interpreting antigenic distance data from Nextstrain. We
571	thank Lone Simonsen for helpful discussions. KG was supported by the National Institutes of Health
572	(F31AI134017, T32-GM008185). JLS was supported by NSF grants OCE-1335657 and DEB-1557022,
573	SERDP RC-2635, and DARPA PREEMPT D18AC00031. MW was supported by the David and Lucile
574	Packard Foundation. The funders had no role in study design, data collection and analysis, decision to
575	publish, or preparation of the manuscript.
576	
577	Code and data availability
578	Code to perform all reported analyses and construct all plots, and all relevant data (Arizona surveillance
579	data and relevant antigenic advance data) is archived at https://zenodo.org/badge/latestdoi/160883450.
580	
581	Disclaimer
582	This work does not necessarily represent the views of the US government or the NIH.
583	
584	Competing interests
585	The authors declare no competing interests.
586	
587	Ethics Statement
588	This study analyzed only existing epidemiological data, which was completely anonymized.
589	
590	

591	Author contributions
592	MW, KG and JLS conceived of the questions and modeling analysis. CV and MW provided
593	crucial assistance with data access and study design. SB and RB supervised data curation and advised the
594	modeling arm of our team about compatibility between the data and analysis strategy. KG wrote the code
595	and performed analyses, with supervision from JLS, and drafted the manuscript. All authors provided
596	input on analysis and interpretation of the results, and helped revise and edit the manuscript text.
597	
598	

599		References
600	1.	Francis T. On the Doctrine of Original Antigenic Sin. Proc Am Philos Soc. 1960;104(6):572-8.
601 602 603	2.	Lessler J, Riley S, Read JM, Wang S, Zhu H, Smith GJD, et al. Evidence for Antigenic Seniority in Influenza A (H3N2) Antibody Responses in Southern China. PLOS Pathog. 2012 Jul 19;8(7):e1002802.
604 605	3.	Henry C, Palm A-KE, Krammer F, Wilson PC. From Original Antigenic Sin to the Universal Influenza Virus Vaccine. Trends Immunol. 2018;39(1):70–9.
606 607	4.	Xu R, Ekiert DC, Krause JC, Hai R, Crowe JE, Wilson IA. Structural Basis of Preexisting Immunity to the 2009 H1N1 Pandemic Influenza Virus. Science. 2010 Apr 16;328(5976):357–60.
608 609 610	5.	Hancock K, Veguilla V, Lu X, Zhong W, Butler EN, Sun H, et al. Cross-Reactive Antibody Responses to the 2009 Pandemic H1N1 Influenza Virus. N Engl J Med Boston. 2009 Nov 12;361(20):1945–52.
611 612 613	6.	Simonsen L, Spreeuwenberg P, Lustig R, Taylor RJ, Fleming DM, Kroneman M, et al. Global Mortality Estimates for the 2009 Influenza Pandemic from the GLaMOR Project: A Modeling Study. PLOS Med. 2013 Nov 26;10(11):e1001558.
614 615	7.	Simonsen L, Reichert TA, Miller MA. The virtues of antigenic sin: consequences of pandemic recycling on influenza-associated mortality. Int Congr Ser. 2004 Jun 1;1263:791–4.
616 617	8.	Ma J, Dushoff J, Earn DJD. Age-specific mortality risk from pandemic influenza. J Theor Biol. 2011 Nov 7;288:29–34.
618 619	9.	Worobey M, Han G-Z, Rambaut A. Genesis and pathogenesis of the 1918 pandemic H1N1 influenza A virus. Proc Natl Acad Sci. 2014 Jun 3;111(22):8107–12.
620 621 622	10.	Gagnon A, Miller MS, Hallman SA, Bourbeau R, Herring DA, Earn DJD, et al. Age-Specific Mortality During the 1918 Influenza Pandemic: Unravelling the Mystery of High Young Adult Mortality. PLoS ONE. 2013 Aug 5;8(8).
623 624	11.	Gostic KM, Ambrose M, Worobey M, Lloyd-Smith JO. Potent protection against H5N1 and H7N9 influenza via childhood hemagglutinin imprinting. Science. 2016 Nov 11;354(6313):722–6.
625 626	12.	Khiabanian H, Farrell GM, George KS, Rabadan R. Differences in Patient Age Distribution between Influenza A Subtypes. PLOS ONE. 2009 Aug 31;4(8):e6832.
627 628 629	13.	Budd AP, Beacham L, Smith CB, Garten RJ, Reed C, Kniss K, et al. Birth Cohort Effects in Influenza Surveillance Data: Evidence that First Influenza Infection Affects Later Influenza-Associated Illness. J Infect Dis.
630 631	14.	Lemaitre M, Carrat F. Comparative age distribution of influenza morbidity and mortality during seasonal influenza epidemics and the 2009 H1N1 pandemic. BMC Infect Dis. 2010 Jun 9;10(1):162.
632 633	15.	Huang QS, Bandaranayake D, Wood T, Newbern EC, Seeds R, Ralston J, et al. Risk Factors and Attack Rates of Seasonal Influenza Infection: Results of the Southern Hemisphere Influenza and

- Vaccine Effectiveness Research and Surveillance (SHIVERS) Seroepidemiologic Cohort Study. J
 Infect Dis. 2019 Jan 9;219(3):347–57.
- 636 16. Cowling BJ, Sullivan SG. The Value of Neuraminidase Inhibition Antibody Titers in Influenza
 637 Seroepidemiology. J Infect Dis. 2019 Jan 9;219(3):341–3.
- Memoli MJ, Shaw PA, Han A, Czajkowski L, Reed S, Athota R, et al. Evaluation of
 Antihemagglutinin and Antineuraminidase Antibodies as Correlates of Protection in an Influenza
 A/H1N1 Virus Healthy Human Challenge Model. mBio. 2016 Apr 19;7(2).
- 18. Wrammert J, Koutsonanos D, Li G-M, Edupuganti S, Sui J, Morrissey M, et al. Broadly crossreactive antibodies dominate the human B cell response against 2009 pandemic H1N1 influenza
 virus infection. J Exp Med. 2011 Jan 17;208(1):181–93.
- Pica N, Hai R, Krammer F, Wang TT, Maamary J, Eggink D, et al. Hemagglutinin stalk antibodies
 elicited by the 2009 pandemic influenza virus as a mechanism for the extinction of seasonal H1N1
 viruses. Proc Natl Acad Sci U S A. 2012;109(7):2573–8.
- 647 20. Krammer F. Novel universal influenza virus vaccine approaches. Curr Opin Virol. 2016 Apr;17:95–
 648 103.
- Ng S, Nachbagauer R, Balmaseda A, Stadlbauer D, Ojeda S, Patel M, et al. Novel correlates of
 protection against pandemic H1N1 influenza A virus infection. Nat Med. 2019 Jun;25(6):962–7.
- Andrews SF, Huang Y, Kaur K, Popova LI, Ho IY, Pauli NT, et al. Immune history profoundly affects broadly protective B cell responses to influenza. Sci Transl Med. 2015 Dec 2;7(316):316ra192-316ra192.
- 23. Zost SJ, Wu NC, Hensley SE, Wilson IA. Immunodominance and Antigenic Variation of Influenza
 Virus Hemagglutinin: Implications for Design of Universal Vaccine Immunogens. J Infect Dis.
 2019 Apr 8;219(Supplement_1):S38–45.
- 657 24. Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, et al. Unifying the
 658 Epidemiological and Evolutionary Dynamics of Pathogens. Science. 2004 Jan 16;303(5656):327–
 659 32.
- 660 25. Henry C, Zheng N-Y, Huang M, Cabanov A, Rojas KT, Kaur K, et al. Influenza Virus Vaccination
 661 Elicits Poorly Adapted B Cell Responses in Elderly Individuals. Cell Host Microbe. 2019
 662 Mar;25(3):357-366.e6.
- Ranjeva S, Subramanian R, Fang VJ, Leung GM, Ip DKM, Perera RAPM, et al. Age-specific
 differences in the dynamics of protective immunity to influenza. Nat Commun. 2019 Apr
 10;10(1):1–11.
- Miller MS, Gardner TJ, Krammer F, Aguado LC, Tortorella D, Basler CF, et al. Neutralizing
 Antibodies Against Previously Encountered Influenza Virus Strains Increase over Time: A
 Longitudinal Analysis. Sci Transl Med. 2013 Aug 14;5(198):198ra107-198ra107.
- Tesini BL, Kanagaiah P, Wang J, Hahn M, Halliley JL, Chaves FA, et al. Broad HemagglutininSpecific Memory B Cell Expansion by Seasonal Influenza Virus Infection Reflects Early-Life
 Imprinting and Adaptation to the Infecting Virus. J Virol. 2019 Apr 15;93(8):e00169-19.

- Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, et al. Mortality
 associated with influenza and respiratory syncytial virus in the United States. JAMA. 2003 Jan
 8;289(2):179–86.
- 67530.Bedford T, Riley S, Barr IG, Broor S, Chadha M, Cox NJ, et al. Global circulation patterns of676seasonal influenza viruses vary with antigenic drift. Nature. 2015 Jul;523(7559):217–20.
- Arizona Department of Health Services. 2015–2016 Influenza Summary [Internet]. [cited 2019 May
 Available from: https://www.azdhs.gov/documents/preparedness/epidemiology-diseasecontrol/flu/surveillance/2015-2016-influenza-summary.pdf
- 880 32. National Notifiable Diseases Surveillance System, Division of Health Informatics and Surveillance,
 881 National Center for Surveillance, Epidemiology and Laboratory Services. MMWR Week Fact Sheet
 882 [Internet]. [cited 2019 May 23]. Available from:
 883 https://wwwn.cdc.gov/nndss/document/MMWR Week overview.pdf
- 684 33. Erbelding EJ, Post DJ, Stemmy EJ, Roberts PC, Augustine AD, Ferguson S, et al. A Universal
 685 Influenza Vaccine: The Strategic Plan for the National Institute of Allergy and Infectious Diseases.
 686 J Infect Dis. 2018 Jul 2;218(3):347–54.
- 687 34. Bolker BM. Ecological Models and Data in R. Princeton University Press; 2008. 409 p.
- Burnham KP, Anderson DR. Model Selection and Multimodel Inference: A Practical Information Theoretic Approach. 2nd ed. New York: Springer-Verlag; 2002.
- 690 36. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. Nextstrain: real-time
 691 tracking of pathogen evolution. Bioinformatics. 2018 Dec 1;34(23):4121–3.
- Neher RA, Bedford T, Daniels RS, Russell CA, Shraiman BI. Prediction, dynamics, and
 visualization of antigenic phenotypes of seasonal influenza viruses. Proc Natl Acad Sci. 2016 Mar
 22;113(12):E1701–9.
- Bedford T, Suchard MA, Lemey P, Dudas G, Gregory V, Hay AJ, et al. Integrating influenza antigenic dynamics with molecular evolution. Losick R, editor. eLife. 2014 Feb 4;3:e01914.
- Smith DJ, Lapedes AS, Jong JC de, Bestebroer TM, Rimmelzwaan GF, Osterhaus ADME, et al.
 Mapping the Antigenic and Genetic Evolution of Influenza Virus. Science. 2004 Jul
 16;305(5682):371–6.
- Cobey S, Hensley SE. Immune history and influenza virus susceptibility. Curr Opin Virol. 2017 Feb
 1;22:105–11.
- 41. Linderman SL, Chambers BS, Zost SJ, Parkhouse K, Li Y, Herrmann C, et al. Potential antigenic
 explanation for atypical H1N1 infections among middle-aged adults during the 2013–2014
 influenza season. Proc Natl Acad Sci. 2014 Nov 4;111(44):15798–803.
- Chen Y-Q, Wohlbold TJ, Zheng N-Y, Huang M, Huang Y, Neu KE, et al. Influenza Infection in
 Humans Induces Broadly Cross-Reactive and Protective Neuraminidase-Reactive Antibodies. Cell.
 2018 Apr 5;173(2):417-429.e10.

708 43. Andrews SF, Chambers MJ, Schramm CA, Plyler J, Raab JE, Kanekivo M, et al. Activation 709 Dynamics and Immunoglobulin Evolution of Pre-existing and Newly Generated Human Memory B 710 cell Responses to Influenza Hemagglutinin. Immunity. 2019 Aug 20;51(2):398-410.e5. 711 Matsuda K, Huang J, Zhou T, Sheng Z, Kang BH, Ishida E, et al. Prolonged evolution of the 44. 712 memory B cell response induced by a replicating adenovirus-influenza H5 vaccine. Sci Immunol. 713 2019 Apr 19;4(34):eaau2710. 714 45. Ramiscal RR, Vinuesa CG. T-cell subsets in the germinal center. Immunol Rev. 2013;252(1):146– 715 55. 716 Rozo M, Gronvall GK. The Reemergent 1977 H1N1 Strain and the Gain-of-Function Debate. mBio. 46. 717 2015 Sep 1;6(4):e01013-15. 718 Dushoff J, Plotkin JB, Viboud C, Earn DJD, Simonsen L. Mortality due to Influenza in the United 47. 719 States-An Annualized Regression Approach Using Multiple-Cause Mortality Data. Am J 720 Epidemiol. 2006 Jan 15;163(2):181-7. 721 48. Lewnard JA, Cobey S. Immune History and Influenza Vaccine Effectiveness. Vaccines. 2018 722 Jun;6(2):28. 723 49. Gagnon A, Acosta E, Miller MS. Reporting and evaluating influenza virus surveillance data: An 724 argument for incidence by single year of age. Vaccine. 2018 Oct 8;36(42):6249-52. 725 50. Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol. 2016 726 Oct;16(10):626–38. 727 51. vom Steeg LG, Klein SL. Sex and sex steroids impact influenza pathogenesis across the life course. 728 Semin Immunopathol. 2019 Mar 1;41(2):189-94. 729 Glezen WP, Keitel WA, Taber LH, Piedra PA, Clover RD, Couch RB. Age Distribution of Patients 52. 730 with Medically-Attended Illnesses Caused by Sequential Variants of Influenza A/H1N1: 731 Comparison to Age-Specific Infection Rates, 1978–1989. Am J Epidemiol. 1991 Feb 1;133(3):296– 732 304. 733 53. US Census Bureau. Index of program surveys, population estimates [Internet]. [cited 2019 Aug 21]. 734 Available from: https://www2.census.gov/programs-surveys/popest/ 735 54. Sagulenko P, Puller V, Neher RA. TreeTime: Maximum-likelihood phylodynamic analysis. Virus 736 Evol. 2018 Jan 8;4(1). 737 Bogner P, Capua I, Lipman DJ, Cox NJ. A global initiative on sharing avian flu data. Nature. 2006 55. 738 Aug:442(7106):981. 739

740

Figure captions

741 Figure 1. Model and expectations under different imprinting hypotheses. (A) Reconstructed, birth year-specific probabilities of imprinting (representative example specific to cases observed in 2015). 742 743 Throughout the manuscript, group 1 HA subtypes are represented in blue and group 2 subtypes in red. (B) 744 Expected imprinting protection against H1N1 or H3N2 under the three tested models. (C) Cartoon of 745 expected age distribution of any influenza case, before controlling for subtype-specific imprinting. The 746 shape of this curve is purely hypothetical, but each of our tested models combined demographic age 747 distribution with a fitted, age-specific risk step function to generate similar, data-driven curves. (D-F) 748 Fraction of each birth year unprotected by their childhood imprinting (from A) determines the shape of 749 birth year-specific risk. (G-I) A linear combination of age-specific risk (C), and birth year-specific risk 750 (D-F) give the expected age distribution of H1N1 or H3N2 cases under each model. 751 752 Figure 2. Observed age distributions, Arizona. Points show fraction of confirmed H1N1 or H3N2 cases 753 observed in each single year of age. Lines show a smoothing spline fit to observed distributions. (A) All 754 confirmed cases in the data (aggregate across all seasons). (B-G) Age distributions from individual 755 seasons in which both H1N1 and H3N2 circulated (seasons with \geq 50 confirmed cases of each subtype are 756 shown here. See Fig. S1 for all seasons). 757 758 Figure 3. Model fits and model selection. (A) Fitted effects of age, after normalization to demographic 759 age distribution and (B) imprinting effects from model AN, which provided the best fit to data. (C-D) 760 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 761 762 subtype-level imprinting, G = HA group-level imprinting. All models included demographic age 763 distribution. 764 765 Figure 4. Effect of antigenic advance on age distribution. (A) Relationship between annual antigenic 766 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 767 768 observed. Blue label shows Spearman correlation between the fraction of H3N2 cases observed in each 769 age group and annual antigenic advance. Blue dashes show linear trend fitted using lm() in R. (B) Season-770 specific age distributions of casses, colored by antigenic advance since the previous season. 771 772 773 **Table legends** 774 775 Table 1. Confirmed cases in surveillance data from Arizona Department of Health Services. Data 776 representing the first and second waves of the 2009 H1N1 pandemic (2008-2009 and 2009-2010 seasons) 777 were excluded. 778 779 Table 2. Maximum likelihood parameter estimates and 95% profile confidence intervals from each 780 model fit to ADHS data. All estimated parameters represent the relative risk of a confirmed case, given 781 the factors listed in the left-hand column. Model name abbreviations specific which factors were included. 782 A = age-specific risk, N = NA subtype-level imprinting, S = HA subtype-level imprinting, G = HA group-783 level imprinting.

Figures

784

В Α H3N2 = H2N2 = H1N1 = naive prob imprinting to subtype 0.0 0.4 0.8 Child imprinted to H1N1 H3N2 H2N2 protected protected protected HA group level H1N1 H1N1 H3N2 protected protected no HA subtype level H1N1 H3N2 protection protected protected protected NA subtype level H3N2 H1N1 H3N2 40 50 60 age in 2015 0 10 20 30 70 80 90 H1N1 risk = H3N2 risk H1N1 risk = H3N2 risk G D HA subtype-level fraction of cases 0.000 0.010 0.020 P(unprotected) 0.0 0.4 0.8 2010 1980 1950 1920 0 20 40 60 80 100 birth year age С Ε Η HA group-level fraction of cases 000 0.010 0.020 P(unprotected) 0.0 0.4 0.8 0.000 0 60 80 100 2010 1980 1950 1920 20 60 80 100 20 40 0 40 age birth year age F I NA subtype-level 0.020 P(unprotected) 0.0 0.4 0.8 0.010 0.000

Figure 1. Model and expectations under different imprinting hypotheses. (A) Reconstructed, birth 786 787 year-specific probabilities of imprinting (representative example specific to cases observed in 2015). 788 Throughout the manuscript, group 1 HA subtypes are represented in blue and group 2 subtypes in red. (B) 789 Expected imprinting protection against H1N1 or H3N2 under the three tested models. (C) Cartoon of 790 expected age distribution of any influenza case, before controlling for subtype-specific imprinting. The 791 shape of this curve is purely hypothetical, but each of our tested models combined demographic age 792 distribution with a fitted, age-specific risk step function to generate similar, data-driven curves. (D-F) 793 Fraction of each birth year unprotected by their childhood imprinting (from A) determines the shape of 794 birth year-specific risk. (G-I) A linear combination of age-specific risk (C), and birth year-specific risk 795 (D-F) give the expected age distribution of H1N1 or H3N2 cases under each model.



medRxiv preprint doi: https://doi.org/10.1101/19001834; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.



796ageage797Figure 2. Observed age distributions, Arizona. Points show fraction of confirmed H1N1 or H3N2 cases798observed in each single year of age. Lines show a smoothing spline fit to observed distributions. (A) All799confirmed cases in the data (aggregate across all seasons). (B-G) Age distributions from individual800seasons in which both H1N1 and H3N2 circulated (seasons with \geq 50 confirmed cases of each subtype are801shown here. See Fig. S1 for all seasons).



Figure 3. Model fits and model selection. (A) Fitted effects of age, after normalization to demographic age distribution and (B) imprinting effects 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. All models included demographic age distribution.



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 Spearman 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 cases, colored by antigenic advance since the previous season.