Antigenic drift and subtype interference shape A(H3N2) epidemic dynamics in the United States

Amanda C Perofsky[1,2*], John Huddleston[3], Chelsea Hansen[1,2], John R Barnes[4], Thomas Rowe[4], Xiyan Xu[4], Rebecca Kondor[4], David E Wentworth[4], Nicola Lewis[5], Lynne Whittaker[5], Burcu Ermetal[5], Ruth Harvey[5], Monica Galiano[5], Rodney Stuart Daniels[5], John W McCauley[5], Seiichiro Fujisaki[6], Kazuya Nakamura[6], Noriko Kishida[6], Shinji Watanabe[6], Hideki Hasegawa[6], Sheena G Sullivan[7], Ian G Barr[7], Kanta Subbarao[7], Florian Krammer[8,9], Trevor Bedford[2,3,10,11], Cécile Viboud[1]

Affiliations

1. Fogarty International Center, National Institutes of Health, United States
2. Brotman Baty Institute for Precision Medicine, University of Washington, United States
3. Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, United States
4. Virology Surveillance and Diagnosis Branch, Influenza Division, National Center for Immunization and Respiratory Diseases (NCIRD), Centers for Disease Control and Prevention (CDC), United States
5. WHO Collaborating Centre for Reference and Research on Influenza, Crick Worldwide Influenza Centre, The Francis Crick Institute, United Kingdom
6. Influenza Virus Research Center, National Institute of Infectious Diseases, Japan
7. WHO Collaborating Centre for Reference and Research on Influenza, The Peter Doherty Institute for Infection and Immunity, Department of Microbiology and Immunology, The University of Melbourne, The Peter Doherty Institute for Infection and Immunity, Australia
8. Center for Vaccine Research and Pandemic Preparedness (C-VaRPP), Icahn School of Medicine at Mount Sinai, United States
9. Department of Pathology, Molecular and Cell-Based Medicine, Icahn School of Medicine at Mount Sinai, United States
10. Department of Genome Sciences, University of Washington, United States
11. Howard Hughes Medical Institute, Seattle, United States

*Correspondence to amanda.perofsky@nih.gov

Abstract

Influenza viruses continually evolve new antigenic variants, through mutations in epitopes of their major surface proteins, hemagglutinin (HA) and neuraminidase (NA). Antigenic drift potentiates the reinfection of previously infected individuals, but the contribution of this process to variability in annual epidemics is not well understood. Here we link influenza A(H3N2) virus evolution to regional epidemic dynamics in the United States during 1997—2019. We integrate phenotypic measures of HA antigenic drift and sequence-based measures of HA and NA fitness to infer antigenic and genetic distances between viruses circulating in successive seasons. We estimate the magnitude, severity, timing, transmission rate, age-specific patterns, and subtype dominance of each regional outbreak and find that genetic distance based on broad sets of epitope sites is the strongest evolutionary predictor of A(H3N2) virus epidemiology. Increased HA and NA epitope distance between seasons correlates with larger, more intense epidemics, higher transmission, greater A(H3N2) subtype dominance, and a greater proportion of cases in adults relative to children, consistent with increased population susceptibility. Based on random forest models, A(H1N1) incidence impacts A(H3N2) epidemics to a greater extent than viral evolution, suggesting that subtype interference is a major driver of influenza A virus infection dynamics, presumably via heterosubtypic cross-immunity.

Introduction

Influenza viruses continually accumulate genetic changes in epitopes of two major surface proteins, hemagglutinin (HA) and neuraminidase (NA), in a process known as "antigenic drift." Though individual hosts develop long-lasting immunity to specific influenza virus strains after infection, antigenic drift helps the virus to escape immune recognition, leaving previously exposed hosts susceptible to reinfection and necessitating the regular updates to the antigens included in the influenza vaccine [1]. While antigenic drift aids immune escape, prospective cohort studies and modeling of surveillance data also indicate that reinfection by antigenically homologous viruses occurs on average every 1-4 years, due to the waning of protection over time and antigenic drift [2,3].

Among the influenza virus types that routinely co-circulate in humans (A and B), type A viruses, particularly subtype A(H3N2), experience the fastest rates of antigenic evolution and cause the most substantial morbidity and mortality [4-7]. Seasonal influenza A viruses (IAV) cause annual winter epidemics in temperate zones of the Northern and Southern Hemispheres and circulate year-round in tropical regions [8]. Influenza A epidemic burden fluctuates substantially from year to year [9], and there is much scientific interest in disentangling the relative roles of viral evolution, prior immunity, human behavior, and climatic factors in driving this seasonal variability. Climatic factors, such as humidity and temperature, have been implicated in the seasonality and timing of winter outbreaks in temperate regions [10-14], while contact and mobility patterns contribute to the seeding of new outbreaks and geographic spread [10,15-19]. A principal requirement for the recurrence of epidemics is a sufficient and continuous source of susceptible individuals, which is determined by the degree of cross-immunity between the surface antigens of currently circulating viruses and functional antibodies elicited by prior infection or vaccination in a population.

Because mutations to the HA1 region of the HA protein are considered to drive the majority of antigenic drift [20,21], influenza virus genetic and antigenic surveillance have focused primarily on HA, and official influenza vaccine formulations prescribe the amount of HA [22]. Yet, evidence for the effect of HA drift on influenza epidemic dynamics remains conflicting. Theoretical and empirical studies have shown that HA drift between currently circulating viruses and the previous season's viruses is expected to cause earlier, larger, more severe, or more synchronized epidemics; however, the majority of these studies were limited to the pre 2009 influenza pandemic period [6,17,23-28]. Information on HA evolution has been shown to improve forecasts of seasonal influenza dynamics in Israel and the United States [29,30], but recent research has also found that HA evolution is not predictive of epidemic size in Australia [31] or epidemic timing in the United States [16]. A caveat is that many of these studies used binary indicators to study seasonal antigenic change, defined as seasons in which circulating viruses were antigenically distinct from the vaccine reference strain [16,17,24,31,32]. This may obscure epidemiologically relevant patterns, as positive selection in HA and NA is both episodic and continuous [6,32-37]. Past research has also typically focused on serological and sequence-based measures of viral evolution in isolation, and the relative importance of these two approaches in predicting epidemic dynamics has not been systematically assessed. Further, to the best of our knowledge, the epidemiologic impact of NA evolution has not been explored.

There has been recent recognition of NA's role in virus inhibiting antibodies and its potential as a vaccine target [38-40]. Though antibodies against NA do not prevent influenza infection, NA immunity attenuates the severity of infection by limiting viral replication [41-46], and NA-specific antibody titers are an independent correlate of protection in both field studies and human challenge trials [47-49]. Lastly, the phenomenon of interference between influenza A subtypes, modulated by immunity to conserved T-cell epitopes [50-52], has long been debated [53,54]. Interference effects are most pronounced during pandemic seasons, leading to troughs or even replacement of the resident subtype in some pandemics [55], but the contribution of heterosubtypic interference to annual dynamics is unclear [2,56-59].

Here, we link A(H3N2) virus evolutionary dynamics to epidemiologic surveillance data in the United States over the course of 22 influenza seasons prior to the coronavirus disease (COVID-19) pandemic, considering the full diversity of viruses circulating in this period. We analyze a variety of antigenic and genetic markers of HA and NA evolution against multiple indicators characterizing the epidemiology and disease burden of annual outbreaks. We find a signature of both HA and NA antigenic drift in surveillance data, with a more pronounced relationship in epitope change rather than the serology-based indicator, along with a major effect of subtype interference. Our study has implications for surveillance of evolutionary indicators that are most relevant for population impact and for prediction of influenza burden on inter-annual timeframes.

Results

Our study focuses on the impact of A(H3N2) virus evolution on seasonal epidemics from seasons 1997-1998 to 2018-2019 in the US; whenever possible, we make use of regionally disaggregated indicators and analyses. We start by identifying multiple indicators of influenza evolution each season based on changes in HA and NA. Next, we compile influenza virus subtype-specific incidence time series for US Department of Health and Human Service (HHS) regions and estimate multiple indicators characterizing influenza A(H3N2) epidemic dynamics each season, including epidemic burden, severity, intensity, type/subtype dominance, timing, and the age distribution of cases. We then assess univariate relationships between indicators of evolution and epidemic characteristics. Lastly, we measure the relative importance of viral evolution, heterosubtypic interference, and prior immunity in predicting regional A(H3N2) epidemic dynamics, using multivariable regression models and random forest models.

Indicators of influenza A(H3N2) evolution

We characterized seasonal patterns of genetic and antigenic evolution among A(H3N2) viruses circulating from 1997 to 2019, using HA and NA sequence data shared via the Global Initiative on Sharing Avian Influenza Data (GISAID) EpiFlu database [60] and ferret hemagglutination inhibition (HI) assay data shared by the WHO Global Influenza Surveillance and Response System (GISRS) Collaborating Centers in London, Melbourne, Atlanta, and Tokyo. Prior to constructing phylogenetic trees, we subsampled sequences to representative sets of 50 viruses per month, with preferential sampling for North American sequences. Although our study is US-focused, we used a global dataset because US-collected sequences and HI titers were sometimes sparse during the earlier seasons of the study. Time-resolved phylogenies of HA and NA genes are shown in Figure 1.

Our choice of evolutionary indicators builds on earlier studies that found hemagglutination inhibition (HI) phenotype or HA sequence data beneficial in forecasting seasonal influenza virus evolution [35,61-63] or annual epidemic dynamics [27,29,30] (Table 1). Historically, HI serological assays were considered the gold standard for measuring immune cross-reactivity between viruses, yet measurements are available for only a subset of viruses. To overcome this limitation, we used a computational approach that maps HI titer measurements onto the HA phylogenetic tree to infer antigenic phenotypes [35,63]. Importantly, this model infers the antigenicity of virus isolates that lack HI titer measurements, which comprise the majority of HA sequences in GISAID. Our sequence-based measures of drift counted substitutions at epitope sites in the globular head domains of HA and NA, identified through monoclonal antibody escape or protein crystal structure: 129 sites in HA epitope regions A to E [21,64-67], 7 sites adjacent to the HA receptor binding site (RBS) [68], and 223 or 53 sites in NA epitope regions A to C [34,69].

We included other indicators of viral fitness for HA and NA, including the number of substitutions at non-epitope sites (mutational load) [35,61] and the average rate of phylogenetic branching in a season (local branching index, LBI) [35,62]. We also calculated the Shannon entropy of LBI values, which considers the richness and relative abundances of viral clades with different growth rates. Lastly, we counted the number of substitutions at epitope sites in the HA stalk domain (stalk footprint distance) [70]. Although the majority of the antibody-mediated response to HA is directed to the immunodominant HA head,

antibodies towards the highly conserved immunosubdominant stalk domain of HA are widely prevalent in older individuals, although at low levels [71-73]. We considered stalk footprint distance to be our "control" metric for drift, given the HA stalk evolves at a significantly slower rate than the HA head [70].

To measure antigenic distances between consecutive seasons, we calculated mean genetic distances at epitope sites or mean $\log_2$ titer distances from HI titer measurements (Figure 1), between viruses circulating in the current season *t* and the prior season *t-1* year (one season lag) or two prior seasons ago *t-2* years (two season lag). These time windows generated seasonal antigenic distances consistent with empirical and theoretical studies characterizing transitions between H3 or N2 antigenic clusters [6,32,35,55,62,74], with H3 epitope distance and HI $\log_2$ titer distance, at two-season lags, and N2 epitope distance, at one-season lags, capturing expected "jumps" in antigenic drift during key seasons that have been previously associated with major antigenic transitions [32], such as the seasons dominated by A/Sydney/5/1997-like strains (SY97) (1997-1998, 1998-1999, 1999-2000) and the 2003-2004 season dominated by A/Fujian/411/2002-like strains (FU02) (Figures S1-S2). Prior studies explicitly linking antigenic drift to epidemic size or severity also support a one-year [6] or two-year time window of drift [26,27]. Given that protective immunity wanes after 1-4 years, we would also expect these timeframes to return the greatest signal in epidemiological surveillance data.

We measured pairwise correlations between seasonal indicators of HA and NA evolution to assess their degree of concordance. As expected, we found moderate-to-strong associations between HA epitope distance and HI $\log_2$ titer distance and HA RBS distance and HI $\log_2$ titer distance (Figure S1-S3). Consistent with prior serological studies [39,75,76], epitope distances in HA and NA were not correlated (one-season lag: Spearman's $\rho = 0.25$, $P = 0.26$; two-season lag: $\rho = 0.15$, $P = 0.5$; Figures S2-S4). Seasonal diversity of HA and NA LBI values was negatively correlated with NA epitope distance (Figure S3), suggesting that selective sweeps follow the emergence of drifted variants.

Associations between A(H3N2) evolution and epidemic dynamics

We explored relationships between viral evolution and variation in A(H3N2) epidemic dynamics from seasons 1997-1998 to 2018-2019, excluding the 2009 A(H1N1) pandemic, using syndromic and virologic surveillance data collected by the US CDC and WHO.

We estimated weekly incidences of influenza A(H3N2), A(H1N1), and B in 10 HHS regions by multiplying the influenza-like illness (ILI) rate – the proportion of outpatient encounters for ILI, weighted by regional population size – by the regional proportion of respiratory samples testing positive for each influenza type/subtype (percent positive) [57,77]. We combined pre-2009 seasonal A(H1N1) viruses and A(H1N1)pdm09 viruses as A(H1N1) and the Victoria and Yamagata lineages of influenza B viruses as influenza B. Weekly incidences of influenza A(H3N2), A(H1N1), and type B, averaged across the 10 HHS regions, are shown in Figure 2. Weekly regional incidences, which show variability in the timing and intensity of annual epidemics, are shown in Figure 2 and Figure S5. Based on these incidence time series, we measured indicators of epidemic burden, intensity, severity, subtype dominance, timing, and age-specific patterns during each non-pandemic season and assessed their univariate relationships with each indicator of HA and NA evolution, which we describe in turn below. Seasonal characteristics of A(H3N2) epidemic dynamics were based on epidemic size, defined as the cumulative weekly incidence; peak incidence, defined as the maximum weekly incidence; excess mortality attributable to A(H3N2), an indicator of epidemic severity; transmissibility, defined as the maximum time-varying effective reproductive number, effective Rt; and epidemic intensity, defined as the inverse Shannon entropy of the weekly incidence distribution (i.e., the sharpness of the epidemic curve). See methods and Table 2 for details on all epidemic metrics and Figure S6 for pairwise correlations between metrics.

Two sequence-based measures based on broad sets of epitope sites exhibited stronger relationships with seasonal epidemic burden and transmissibility than the serology-based measure, HI $\log_2$ titer distance. Both H3 epitope distance ($t-2$) and N2 epitope distance ($t-1$) correlated with increased epidemic size

(linear models, LMs: H3, adjusted $R^2 = 0.37$, $P = 0.03$; N2: $R^2 = 0.26$, $P = 0.08$) and peak incidence (LMs, H3: $R^2 = 0.4$, $P = 0.02$; N2: $R^2 = 0.33$, $P = 0.04$) and higher effective Rt (generalized linear models, GLMs: H3, $R^2 = 0.38$, $P = 0.05$; N2, $R^2 = 0.32$, $P = 0.03$) (regression results: Figure 3, Spearman's correlations: Figure S7). HI $\log_2$ titer distance ($t-2$) exhibited positive but non-significant associations with different measures of epidemic impact (Figure 3, Figure S7). Seasonal diversity in the growth rates of circulating lineages in the current t or prior season ($t-1$) had strong negative correlations with effective Rt (GLMs, H3 ($t-1$): $R^2 = 0.49$, $P = 0.009$; N2, t: $R^2 = 0.46$, $P = 0.006$) and epidemic intensity (Beta GLMs, H3 ($t-1$): $R^2 = 0.45$, $P = 0.003$; N2, t: $R^2 = 0.51$, $P = 0.001$) (Figures S7-S8). Seasonal mean LBI exhibited similar but slightly weaker correlations with effective Rt and epidemic intensity. Pneumonia and influenza excess mortality attributable to A(H3N2) also increased with H3 epitope distance, though this relationship was not statistically significant (Figure S9). The remaining indicators of viral evolution, including H3 and N2 non-epitope distance (mutational load), H3 RBS distance, and H3 stalk footprint distance had weak, non-significant correlations with the different measures of epidemic impact (Figure S7).

We explored whether evolutionary changes in A(H3N2) may predispose this subtype to dominate influenza virus circulation in a given season. A(H3N2) subtype dominance – the proportion of influenza positive samples typed as A(H3N2) – increased with H3 epitope distance ($t-2$) and N2 epitope distance ($t-1$) (Beta GLMs, H3: $R^2 = 0.32$, $P = 0.05$; N2: $R^2 = 0.34$, $P = 0.03$; Figure 4, Figure S7). Figure 4 illustrates this relationship at the regional level across two seasons in which A(H3N2) was nationally dominant, but where antigenic change differed. In 2003-2004, we observed widespread dominance of A(H3N2) viruses after the emergence of the novel antigenic cluster, FU02 (A/Fujian/411/2002-like strains). In contrast, there was substantial regional heterogeneity in subtype circulation during 2007-2008, a season in which A(H3N2) viruses were antigenically similar to those from the previous season. Patterns in type/subtype circulation across all influenza seasons in our study period are shown in Figure S10. As observed for the 2003-2004 season, widespread A(H3N2) dominance tends to coincide with major antigenic transitions (e.g., A/Sydney/5/1997 (SY97) seasons, 1997-1998 to 1999-2000; A/California/7/2004 (CA04) season, 2004-2005), though this was not universally the case (e.g., A/Perth/16/2009 (PE09) season, 2010-2011).

Next, we tested for associations between A(H3N2) evolution and epidemic timing, including onset week, defined as the winter changepoint in incidence [16], and peak week, defined as the first week of maximum incidence; spatiotemporal synchrony, measured as the variation (standard deviation, s.d.) in regional onset and peak timing; and epidemic speed, including seasonal duration and the number of weeks from onset to peak (Table 2, Figure S11). Seasonal duration increased with H3 or N2 LBI diversity in the current t or prior season ($t-1$) (Gamma GLMs, H3, t: $R^2 = 0.6$; $P = 0.001$; N2, t: $R^2 = 0.6$; $P = 0.002$; Figures S11-S12), while the number of days from epidemic onset to peak shortened with increasing N2 epitope distance ($t-1$) (Gamma GLM, $R^2 = 0.31$, $P = 0.04$; Figure S11, Figure S13). Onset and peak timing tended to be earlier in seasons with increased H3 and N2 antigenic novelty, but correlations were not statistically significant (Figure S14). A(H3N2) evolution did not correlate with the degree of spatiotemporal synchrony across HHS regions.

Lastly, we considered the effects of antigenic change on the age distribution of outpatient ILI cases, with the expectation that the proportion of cases in children would decrease in seasons with greater antigenic novelty, due to drifted variants' increased ability to infect more immunologically experienced adults [7,78]. Consistent with this hypothesis, N2 epitope distance from prior seasons was negatively correlated with the fraction of cases in children aged < 5 years (LMs, one-season lag: $R^2 = 0.29$, $P = 0.1$; two-season lag: $R^2 = 0.59$, $P = 0.003$) and individuals aged 5-24 years (one-season lag: $R^2 = 0.38$, $P = 0.04$; two-season lag: $R^2 = 0.17$, $P = 0.18$) and negatively correlated with the fraction of cases in adults aged 25-64 years (one-season lag: $R^2 = 0.36$, $P = 0.05$; two-season lag: $R^2 = 0.49$, $P = 0.01$) and ≥65 years (one-season lag: $R^2 = 0.39$, $P = 0.01$; two-season lag: $R^2 = 0.33$, $P = 0.05$) (Figures S15-S16). As observed in Gostic et al. [78], H3 epitope distance ($t-2$) had negative but non-significant associations with the fraction of cases in children and positive but non-significant associations with the fraction of cases in adult age groups (Figures S15-S16).

Effects of heterosubtypic viral interference on A(H3N2) epidemic burden and timing

We investigated the effects of influenza type/subtype interference – proxied by influenza A(H1N1) and B epidemic size – on A(H3N2) incidence during annual outbreaks. Across the entire study period, we observed moderate-to-strong, non-linear relationships between A(H1N1) epidemic size and A(H3N2) epidemic size (GLM, $R^2 = 0.65$, $P = 0.01$), peak incidence ($R^2 = 0.66$, $P = 0.02$), and excess mortality (all age groups and $\geq$ 65 years, $R^2 = 0.57$, $P = 0.01$) (Figure 5, Figure S17), wherein A(H3N2) epidemic burden and excess mortality decreased as A(H1N1) incidence increased. A(H1N1) epidemic size was also significantly correlated with A(H3N2) effective Rt, exhibiting a negative, approximately linear relationship (GLM, $R^2 = 0.45$, $P = 0.01$) (Figure 5). A(H3N2) epidemic intensity was negatively associated with A(H1N1) epidemic size, but this relationship was not statistically significant (Beta GLM, $R^2 = 0.21$, $P = 0.15$). Influenza B epidemic size was not significantly correlated with any A(H3N2) epidemic metrics (Figure 5, Figure S17).

The internal gene segments NS, M, NP, PA, and PB2 of A(H3N2) viruses and pre-2009 seasonal A(H1N1) viruses share a common ancestor [79] whereas A(H1N1)pdm09 viruses have a combination of gene segments derived from swine and avian reservoirs that were not reported prior to the 2009 pandemic [80,81]. Because pre-2009 seasonal A(H1N1) viruses and A(H3N2) are more closely related, seasonal A(H1N1) viruses may limit the circulation of A(H3N2) viruses to a greater extent than A(H1N1)pdm09 viruses. As a sensitivity analysis, we measured correlations between A(H1N1) incidence and A(H3N2) epidemic metrics separately for pre- and post-2009 pandemic time periods. Relationships between different A(H3N2) epidemic metrics and A(H1N1) epidemic size were broadly similar for both periods, with slightly stronger correlations observed during the pre-2009 period (Figure S18).

We compared A(H3N2) epidemic timing across A(H3N2) and A(H1N1) dominant seasons, which we defined as when $\geq$70% of influenza A positive samples are typed as A(H3N2) or A(H1N1)/A(H1N1)pdm09, respectively. We applied a strict threshold for subtype dominance because seasons with < 70% samples of one IAV subtype tended to have greater geographic heterogeneity in circulation, resulting in regions with dominant subtypes that were not nationally dominant. A(H3N2) epidemic onsets and peaks occurred, on average, three weeks earlier in A(H3N2) dominant seasons (Wilcoxon test, $P < 0.0001$). In A(H1N1) dominant seasons, regional A(H3N2) epidemics exhibited greater heterogeneity in epidemic timing (onset s.d.: H3 dominant seasons, 12.4 weeks versus H1 dominant seasons, 16.3 weeks; peak s.d., H3 dominant seasons, 13.3 weeks versus H1 dominant seasons, 22.6 weeks; Wilcoxon tests, $P < 0.0001$) and were significantly shorter in duration compared to A(H3N2) dominant seasons (median duration: H3 dominant seasons, 29 weeks versus H1 dominant seasons, 21 weeks; Wilcoxon test, $P < 0.0001$).

We applied a wavelet approach [82] to weekly time series of type/subtype-specific incidences to measure more fine-scale differences in the relative timing of type/subtype circulation (Figure S19). A(H3N2) incidence preceded A(H1N1) incidence during most seasons prior to 2009 and during the two seasons in which A(H1N1)pdm09 was dominant, potentially because A(H3N2) viruses are more globally prevalent and migrate between regions more frequently than A(H1N1) viruses [7]. There was not a clear relationship between the direction of seasonal phase lags and A(H1N1) epidemic size (LM, $R^2 = 0.23$, $P = 0.1$; Figure S19). A(H3N2) incidence led influenza B incidence in all influenza seasons (positive phase lag), irrespective of influenza B epidemic size (LM, $R^2 = 0.05$, $P = 0.5$; Figure S19).

The relative impacts of viral evolution, heterosubtypic interference, and prior immunity on A(H3N2) epidemic dynamics

We implemented conditional inference random forest models to assess the relative importance of viral evolution, type/subtype co-circulation, prior population immunity, and vaccine-related parameters in predicting regional A(H3N2) epidemic metrics (Figure 6). We limited viral evolutionary indicators to H3 epitope distance ($t-2$), N2 epitope distance ($t-1$), HI $\log_2$ titer distance ($t-2$), and H3 and N2 LBI

diversity in the current and prior season, due to weaker or non-significant correlations between the other evolutionary metrics and epidemic burden (Figure S7). To account for potential type or subtype interference, we included A(H1N1) epidemic size (A(H1N1) or A(H1N1)pdm09) and B epidemic size in the current and prior season and the dominant IAV subtype in the prior season. We included A(H3N2) epidemic size in the prior season as a proxy of natural prior immunity to A(H3N2). To account for vaccine-induced immunity, we considered four categories of predictors and included estimates for the current and prior seasons: seasonal vaccination coverage among adults (18-49 years coverage $\times \geq$ 65 years coverage), adjusted A(H3N2) vaccine effectiveness (VE), a combined metric of vaccination coverage and A(H3N2) VE (18-49 years coverage $\times \geq$ 65 years coverage $\times$ VE), and H3 and N2 epitope distance between currently circulating strains and the US vaccine reference strain. We could not include a predictor for vaccination coverage in children or consider clade-specific VE estimates, because data were not available for most seasons in our study. We did not predict excess mortality attributable to A(H3N2), due to data limitation (one national estimate per season) and omitted models predicting epidemic timing, due to weak or non-significant correlations between timing-related measures and most indicators of viral evolution (Figure S11). Lastly, we could not separate our analysis into pre- and post-2009 pandemic periods due to small sample sizes.

Based on variable importance scores, A(H1N1) epidemic size in the current season was the most informative predictor of A(H3N2) epidemic size and peak incidence, followed by H3 epitope distance, and the dominant IAV subtype in the previous season or N2 epitope distance (Figure 6). For A(H3N2) subtype dominance, the highest ranked predictors were H3 epitope distance, N2 epitope distance, and the dominant IAV subtype in the previous season (Figure 6). We note that we did not include A(H1N1) epidemic size as a predictor in this model, due to its confounding with the target variable. For models of A(H3N2) effective Rt and epidemic intensity, we observed less discernable differences in variable importance scores across the set of candidate predictors (Figure 6). For the model of effective Rt, N2 LBI diversity in the current season, A(H1N1) epidemic size in the current season, and N2 epitope distance between circulating strains and the vaccine strain were the highest ranked variables, while the most important predictors of epidemic intensity were H3 and N2 LBI diversity in the current season and adult vaccination coverage in the current and prior season. Variable importance rankings from LASSO (least absolute shrinkage and selection operator) regression models were qualitatively similar to those from random forest models, with A(H1N1) epidemic size in the current season, H3 and N2 epitope distance, and the dominant IAV subtype in the prior season consistently retained across the best-tuned models of epidemic size, peak incidence, and subtype dominance (Figure S20). Vaccine-related parameters and H3 antigenic drift (either H3 epitope distance or HI $\log_2$ titer distance) were retained in the best-tuned LASSO models of effective Rt and epidemic intensity (Figure S20).

We measured correlations between observed values and model-predicted values at the HHS region level. Among our various epidemic metrics, random forest models produced the most accurate predictions of A(H3N2) subtype dominance (ρ = 0.94, regional range = 0.8 – 0.98), peak incidence (Spearman's ρ = 0.91, regional range = 0.73 – 0.95), and epidemic size (ρ = 0.9, regional range = 0.73 – 0.94), while predictions of effective Rt and epidemic intensity were less accurate (ρ = 0.8, regional range = 0.65 – 0.91; ρ = 0.78, regional range = 0.63 – 0.91, respectively) (Figure 7). Random forest models tended to underpredict most epidemic targets in seasons with substantial H3 antigenic transitions, in particular the SY97 cluster seasons (1998-1999, 1999-2000) and the FU02 cluster season (2003-2004) (Figure 7).

For epidemic size and peak incidence, seasonal predictive error – root-mean-square error (RMSE) across all regional predictions in a season – increased with H3 epitope distance (size, Spearman's ρ = 0.51, P = 0.02; peak, ρ = 0.61, P = 0.007) and N2 epitope distance (size, ρ = 0.43, P = 0.06; peak, ρ = 0.46, P = 0.04). For models of epidemic intensity, seasonal RMSE increased with N2 epitope distance (ρ = 0.62, P = 0.006) but not H3 epitope distance (ρ = 0.07, P = 0.8) (Figures S21-S22). The RMSE of effective Rt and subtype dominance predictions were not significantly correlated with H3 or N2 epitope distance (Figures S21-22).

To further refine our set of informative predictors, we performed multivariable regression with the top 10 ranked predictors from each random forest model and used Bayesian Information Criterion (BIC) to select the best fit model for each epidemic metric, allowing each metric's regression model to include up to three independent variables. This additional step of variable selection demonstrated that models with few predictors fit the observed data relatively well (epidemic size, adjusted $R^2 = 0.69$; peak incidence, adj. $R^2 = 0.63$; effective Rt, adj. $R^2 = 0.65$; epidemic intensity, adj. $R^2 = 0.75$), except for subtype dominance (adj. $R^2 = 0.48$) (Table 3). The set of variables retained after model selection were similar to those with high importance rankings in random forest models and LASSO regression models, with the exception that HI $\log_2$ titer distance, rather than H3 epitope distance, was included in the minimal models of effective Rt and epidemic intensity.

Discussion

Antigenic drift between currently circulating influenza viruses and the previous season's viruses is expected to confer increased viral fitness, leading to earlier, larger, or more severe epidemics. However, prior evidence for the impact of antigenic drift on seasonal influenza outbreaks is mixed. Here, we systematically compare experimental and sequence-based measures of A(H3N2) evolution in predicting regional epidemic dynamics in the United States across 22 seasons, from 1997 to 2019. We also consider the effects of other co-circulating influenza viruses, prior immunity, and vaccine-related parameters, such as coverage and effectiveness, on A(H3N2) incidence. Our findings indicate that evolution in both major surface proteins – hemagglutinin (HA) and neuraminidase (NA) – contributes to variability in epidemic magnitude across seasons, though viral fitness appears to be secondary to subtype interference in shaping annual outbreaks.

The first question of this study sought to determine which metrics of viral fitness have the strongest relationships with A(H3N2) epidemic burden and timing. Among our set of candidate evolutionary predictors, genetic distances based on broad sets of epitope sites (HA = 129 sites; NA = 223 epitope sites) had the strongest, most consistent associations with A(H3N2) epidemic size, transmission rate, severity, subtype dominance, and age-specific patterns. Increased epitope distance in both H3 and N2 correlated with larger epidemics and increased transmissibility, with univariate analyses finding H3 distance more strongly correlated with epidemic size, peak incidence, transmissibility, and excess mortality, and N2 distance more strongly correlated with epidemic intensity (i.e., the "sharpness" of the epidemic curve) and subtype dominance patterns. However, we note that minor differences in correlative strength between H3 and N2 epitope distance are not necessarily biologically relevant and could be attributed to noise in epidemiological or virological data or the limited number of influenza seasons in our study. The fraction of ILI cases in children relative to adults was negatively correlated with N2 epitope distance, consistent with the expectation that cases are more restricted to immunologically naïve children in seasons with low antigenic novelty [7,78]. Regarding epidemic timing, the number of days from epidemic onset to peak (a proxy for epidemic speed) decreased with N2 epitope distance, but other measures of epidemic timing, such as peak week, onset week, and spatiotemporal synchrony across HHS regions, were not significantly correlated with H3 or N2 antigenic change.

The local branching index (LBI) is traditionally used to predict the success of individual clades, with a high LBI value indicating high viral fitness [35,62]. In our epidemiological analysis, low diversity of H3 or N2 LBI values, in the current or prior season, correlated with greater epidemic intensity, higher transmission rates, and shorter seasonal duration. This outcome suggests that low LBI diversity is indicative of a rapid selective sweep by one successful clade and that high LBI diversity is indicative of multiple co-circulating clades with variable seeding times over the course of an epidemic. A caveat is that LBI estimation is more sensitive to sequence sub-sampling schemes than strain-level measures. If an epidemic is very short and intense (e.g., 1-2 months), a phylogenetic tree with our sub-sampling scheme (50 sequences per month) may not incorporate enough sequences to capture the true diversity of LBI values in that season.

Positive associations between H3 antigenic drift and population-level epidemic burden are consistent with previous observations from theoretical models [25,26,83]. For example, phylodynamic models of punctuated antigenic evolution have reproduced key features of A(H3N2) phylogenetic patterns and case dynamics, such as the sequential replacement of antigenic clusters, the limited standing diversity in HA after a cluster transition, and higher incidence and attack rates in cluster transition years [25,26,83]. Our results also corroborate empirical analyses of surveillance data [6,27,28,66] and forecasting models of annual epidemics [29,30] that found direct, quantitative links between HA antigenic novelty and the number of influenza cases or deaths in a season. Moving beyond the paradigm of antigenic clusters, Wolf et al., 2010 and Bedford et al., 2014 demonstrated that smaller, year-to-year changes in H3 antigenic drift also correlate with seasonal severity and incidence [6,27]. A more recent study did not detect an association between antigenic drift and city-level epidemic size in Australia [31], though the authors used a binary indicator to signify seasons with major HA antigenic transitions and did not consider smaller, more gradual changes in antigenicity. While Lam and colleagues did not observe a consistent effect of antigenic change on epidemic magnitude, they found a negative relationship between the cumulative prior incidence of an antigenic variant and its probability of successful epidemic initiation in a city [31].

We did not observe a clear relationship between H3 receptor binding site (RBS) distance and epidemic burden, even though single substitutions at these seven amino acid positions are implicated in major antigenic transitions [68,84]. The outperformance of the RBS distance metric by a broader set of epitope sites could be attributed to the tempo of antigenic cluster changes. A(H3N2) viruses are characterized by both continuous and punctuated antigenic evolution, with transitions between antigenic clusters occurring every 2 to 8 years [6,26,32,33,36,37,67,68,85]. Counting substitutions at only a few sites may fail to capture more modest, gradual changes in antigenicity that are on a time scale congruent with annual outbreaks. Further, a broader set of epitope sites may better capture the epistatic interactions that underpin antigenic change in HA [86]. Although the 7 RBS sites were responsible for the majority of antigenic phenotype in Koel et al.'s experimental study [68], their findings do not necessarily contradict studies that found broader sets of sites associated with antigenic change. Mutations at other epitope sites may collectively add to the decreased recognition of antibodies or affect viral fitness through alternate mechanisms (e.g., compensatory or permissive mutations) [26,32,36,62,68,86-88].

A key result from our study is the direct link between NA antigenic drift and A(H3N2) incidence patterns. Although HA and NA both contribute to antigenicity [20,89] and undergo similar rates of positive selection [34], we expected antigenic change in HA to exhibit stronger associations with seasonal incidence, given its immunodominance relative to NA [90]. H3 and N2 epitope distance were both moderately correlated with epidemic size, peak incidence, and subtype dominance patterns, but, except for subtype dominance, H3 epitope distance had higher variable importance rankings in random forest models and N2 epitope distance was not retained after post-hoc model selection of top ranked random forest features. However, N2 epitope distance but not H3 epitope distance was associated with faster epidemic speed and a greater fraction of ILI cases in adults relative to children. Antigenic changes in H3 and N2 were independent across the 22 seasons of our study, consistent with previous research [34,74,76]. Thus, the similar predictive performance of HA and NA epitope distance for some epidemic metrics does not necessarily stem from the coevolution of HA and NA.

HI $\log_2$ titer distance was positively correlated with different measures of epidemic impact yet underperformed in comparison to H3 and N2 epitope distances. This outcome was surprising given that we expected our method for generating titer distances to produce more realistic estimates of immune cross-protection between viruses than epitope-based measures. Our computational approach for inferring HI phenotype dynamically incorporates newer titer measurements and assigns antigenic weight to phylogenetic branches rather than fixed sequence positions [35,63], while our method for calculating epitope distance assumes that the contributions of specific sites to antigenic drift are constant through time, even though beneficial mutations previously observed at these sites are contingent on historical patterns of viral fitness and host immunity [26,35,62]. HI titer measurements have been more useful than

epitope substitutions in predicting future A(H3N2) viral populations [35] and vaccine effectiveness [91], with the caveat that these targets are more proximate to viral evolution than epidemic dynamics.

HI titer measurements may be more immunologically relevant than epitope-based measures, yet several factors could explain why substitutions at epitope sites outperformed HI titer distances in epidemiological predictions. First, epitope distances may capture properties that affect viral fitness (and in turn outbreak intensity) but are unrelated to immune escape, such as intrinsic transmissibility, ability to replicate, or epistatic interactions. A second set of factors concern methodological issues associated with HI assays. The reference anti-sera for HI assays are routinely produced in ferrets recovering from their first influenza virus infection. Most humans are infected by different influenza virus strains over the course of their lifetimes, and one's immune history influences the specificity of antibodies generated against drifted influenza virus strains [92-95]. Thus, human influenza virus antibodies, especially those of adults, have more heterogeneous specificities than anti-sera from immunologically naïve ferrets [92].

A related methodological issue is that HI assays disproportionately measure anti-HA antibodies that bind near the receptor binding site and, similar to the RBS distance metric, may capture only a partial view of the antigenic change occurring in the HA protein [31,78,96,97]. A recent study of longitudinal serological data found that HI titers are a good correlate of protective immunity for children, while time since infection is a better predictor of protection for adults [97]. This outcome is consistent with the concept of antigenic seniority, in which an individual's first exposure to influenza virus during childhood leaves an immunological "imprint", and exposure to new strains "back boosts" one's antibody response to strains of the same subtype encountered earlier in life [78,98,99]. Ranjeva et al.'s study and others suggest that human influenza virus antibodies shift focus from the HA head to other more conserved epitopes as individuals age [78,96]. Given that HI assays primarily target epitopes adjacent to the RBS, HI assays using ferret or human serological data are not necessarily suitable for detecting the broader immune responses of adults. A third explanation for the underperformance of HI titers concerns measurement error. Recent A(H3N2) viruses have reduced binding efficiency in HI assays, which can skew estimates of immune cross-reactivity between viruses [100]. These combined factors could obfuscate the relationship between the antigenic phenotypes inferred from HI assays and population-level estimates of A(H3N2) incidence.

Novel antigenic variants are expected to have higher infectivity in immune populations, leading to earlier epidemics and more rapid geographic spread [19], but few studies have quantitatively tied antigenic drift to epidemic timing or geographic synchrony. Previous studies of pneumonia and influenza-associated mortality observed greater severity or geographic synchrony in seasons with major antigenic transitions [21,24]. A more recent Australian study of lab-confirmed cases also noted greater spatiotemporal synchrony during seasons in which novel H3 antigenic variants emerged, although their assessment was based on virus typing alone (i.e., influenza A or B) [17]. A subsequent Australian study with finer-resolution data on subtype incidence and variant circulation determined that more synchronous epidemics were not associated with drifted A(H3N2) strains [31], and a US-based analysis of ILI data also failed to detect a relationship between HA antigenic cluster transitions and geographic synchrony [16]. In our study, the earliest epidemics tended to occur in seasons with transitions between H3 antigenic clusters (e.g., the emergence of the FU02 cluster in 2003-2004) or vaccine mismatches (e.g., N2 mismatch in 1999-2000, H3 mismatch in 2014-2015) [32,74,101], but there was not a statistically significant correlation between antigenic change and earlier epidemic onsets or peaks. Regarding epidemic speed, the length of time from epidemic onset to peak decreased with N2 epitope distance but not H3 epitope distance. The relationship between antigenic drift and epidemic timing may be ambiguous because external seeding events or climatic factors, such as temperature and absolute humidity, are more important in driving influenza seasonality and the onsets of winter epidemics [7,10-14,16]. Alternatively, the resolution of our epidemiological surveillance data (HHS regions) may not be granular enough to detect a signature of antigenic drift in epidemic timing, though studies of city-level influenza dynamics were also unable to identify a clear relationship [16,31].

After exploring individual correlations between evolutionary indicators and annual epidemics, we considered the effects of influenza A(H1N1) incidence and B incidence on A(H3N2) virus circulation within a season. We detected strong negative associations between A(H1N1) incidence and A(H3N2) epidemic size, peak incidence, transmissibility, and excess mortality, consistent with previous animal, epidemiological, phylodynamic, and theoretical studies that found evidence for cross-immunity between IAV subtypes [53-55,57,59,102]. For example, individuals recently infected with seasonal influenza A viruses are less likely to become infected during subsequent pandemic waves [52,53,102-104], and the early circulation of one influenza virus type or subtype is associated with a reduced total incidence of the other type/subtypes within a season [31,57]. Due to the shared evolutionary history of their internal genes [79], pre-2009 seasonal A(H1N1) viruses may impact A(H3N2) virus circulation to a greater extent than A(H1N1)pdm09 viruses, which have a unique combination of genes that were not identified in animals or humans prior to 2009 [81,105]. We observed similar relationships between A(H3N2) epidemic metrics and A(H1N1) incidence during pre- and post-2009 pandemic seasons, with slightly stronger correlations observed during the pre-2009 period. However, given the small sample size (12 pre-2009 seasons and 9 post-2009 seasons), we cannot fully answer this question.

In our study, univariate correlations between A(H1N1) and A(H3N2) incidence were more pronounced than those observed between A(H3N2) incidence and evolutionary indicators, and A(H1N1) epidemic size was the highest ranked feature by random forest models predicting epidemic size and peak incidence. Consequently, interference between the two influenza A subtypes may be more impactful than viral evolution in determining the size of annual A(H3N2) outbreaks. Concerning epidemic timing, we did not detect a relationship between A(H3N2) antigenic change and the relative timing of A(H3N2) and A(H1N1) cases; specifically, A(H3N2) incidence did not consistently lead A(H1N1) incidence in seasons with greater H3 or N2 antigenic change. Overall, we did not find any indication that influenza B incidence affects A(H3N2) epidemic burden or timing, which is not unexpected, given that few T and B cell epitopes are shared between the two virus types [106].

Lastly, we used random forest models and multivariable linear regression models to assess the relative importance of viral evolution, prior population immunity, co-circulation of other influenza viruses, and vaccine-related parameters in predicting regional A(H3N2) epidemic dynamics. We chose conditional inference random forest models as our primary method of variable selection because several covariates were collinear, relationships between some predictors and target variables were nonlinear, and our goal was inferential rather than predictive. We performed leave-one-season-out cross-validation to tune each model, but, due to the limited number of seasons in our dataset, we were not able to test predictive performance on an independent test set. With the caveat that models were likely overfit to historical data, random forest models produced accurate predictions of regional epidemic size, peak incidence, and subtype dominance patterns, while predictions of epidemic intensity and transmission rates were less exact. The latter two measures could be more closely tied to climatic factors, the timing of influenza case importations from abroad, or mobility patterns [7,13,14,16] or they may be inherently more difficult to predict because their values are more constrained. Random forest models tended to underpredict epidemic burden in seasons with major antigenic transitions, particularly the SY97 seasons (1998-1999, 1999-2000) and the FU02 season (2003-2004), potentially because antigenic jumps of these magnitudes were infrequent during our 22-season study period. An additional step of post-hoc model selection demonstrated that models with only three covariates could also produce accurate fits to observed epidemiological data.

Our study is subject to several limitations, specifically regarding geographic resolution and data availability. First, our analysis is limited to one country with a temperate climate and its findings concerning interactions between A(H3N2), A(H1N1), and type B viruses may not be applicable to tropical or subtropical countries, which experience sporadic epidemics of all three viruses throughout the year [107]. Second, our measure of population-level influenza incidence is derived from regional CDC outpatient data because those data are publicly available starting with the 1997-1998 season. State level outpatient data are not available until after the 2009 A(H1N1) pandemic, and finer resolution data from

electronic health records are accessible in theory but not in the public domain. Access to ILI cases aggregated at the state or city level, collected over the course of decades, would increase statistical power and enable us to add more location-specific variables to our analysis, such as climatic and environmental factors. A third limitation is that we measured influenza incidence by multiplying the rate of influenza-like illness by the percentage of tests positive for influenza, which does not completely eliminate the possibility of capturing the activity of other co-circulating respiratory pathogens [11]. Surveillance data based on more specific diagnosis codes would ensure the exclusion of patients with non-influenza respiratory conditions. Fourth, our data on the age distribution of influenza cases were derived from ILI encounters across four broad age groups and did not include test positivity status, virus type/subtype, or denominator information. Despite the coarseness of these data, we found statistically significant correlations in the expected directions between N2 antigenic change and the fraction of cases in children relative to adults. Lastly, a serological assay exists for NA, but NA titer measurements are not widely available because the assay is labor-intensive and inter-lab variability is high. Thus, we could not test the performance of NA antigenic phenotype in predicting epidemic dynamics.

Beginning in early 2020, non-pharmaceutical interventions (NPIs), including lockdowns, school closures, physical distancing, and masking, were implemented in the United States and globally to slow the spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus responsible for the COVID-19 pandemic. These mitigation measures disrupted the transmission of seasonal influenza viruses and other directly-transmitted respiratory viruses throughout 2020 and 2021 [108-113], and population immunity to influenza is expected to have decreased substantially during this period of low circulation [114,115]. COVID-19 NPIs relaxed during 2021 and 2022, and co-circulation of A(H3N2) and A(H1N1)pdm09 viruses in the United States resumed during the 2022-2023 influenza season. Our study concludes with the 2018-2019 season, and thus it is unclear whether our modeling approach would be useful in projecting seasonal burden during the post-pandemic period, without an additional component to account for COVID-19-related perturbations to influenza transmission. Further studies will need to determine whether ecological interactions between influenza viruses have changed or if the effects of viral evolution and subtype interference on seasonal outbreaks are different in the post-pandemic period.

In conclusion, relationships between A(H3N2) antigenic drift, epidemic impact, and age dynamics are moderate, with genetic distances based on broad sets of H3 and N2 epitope sites having greater predictive power than serology-based antigenic distances for the timeframe analyzed. Influenza epidemiological patterns are consistent with increased population susceptibility in seasons with high antigenic novelty, and our study is the first to link NA antigenic drift to epidemic burden, timing, and the age distribution of cases. It is well established that anti-HA and anti-NA antibodies are independent correlates of immunity [45,47-49,116-118], and the influenza research community has advocated for NA-based vaccines [39,119]. The connection between NA drift and seasonal incidence further highlights the importance of monitoring evolution in both HA and NA to inform vaccine strain selection and epidemic forecasting efforts. Although antigenic change in both HA and NA was correlated with epidemic dynamics, ecological interactions between influenza A subtypes appear to be more influential than viral evolution in determining the intensity of annual A(H3N2) epidemics. The aim of our study was to retrospectively assess the potential drivers of annual A(H3N2) epidemics, yet we cautiously suggest that one could project the size or intensity of future epidemics based on sequence data and A(H1N1)pdm09 incidence alone [27,57].

Methods

Unless otherwise noted, data processing and statistical analyses were performed using R version 4.3.0.

Influenza epidemic timing and burden

Influenza-like illness and virological surveillance data

We obtained weekly epidemiological and virological data for influenza seasons 1997-1998 to 2018-2019, at the U.S. HHS region level [120]. We defined influenza seasons as calendar week 40 in a given year to calendar week 20 in the following year, with the exception of the 2008-2009 season, which ended in 2009 week 16 due to the emergence of the A(H1N1)pdm09 virus [57].

We extracted syndromic surveillance data for the ten HHS regions from the U.S. Outpatient Influenza-like Illness Surveillance Network (ILINet) [120]. ILINet consists of approximately 3,200 sentinel outpatient healthcare providers throughout the United States that report the total number of consultations for any reason and the number of consultations for influenza-like illness (ILI) every week. ILI is defined as fever (temperature of 100°F [37.8°C] or greater) and a cough and/or a sore throat. The indicator is based on the weekly proportion of outpatient consultations for influenza-like illness and is available weighted or unweighted by regional population size. The number of ILI encounters by age group are also provided (0-4, 5-24, 25-64, and ≥65), but these data are not weighted by total encounters or population size.

Data on weekly influenza virus type and subtype circulation were obtained from the US CDC's World Health Organization (WHO) Collaborating Center for Surveillance, Epidemiology and Control of Influenza [121]. We estimated the weekly number of respiratory samples testing positive for influenza A(H1N1), A(H1N1)pdm09, A(H3N2), or B at the HHS region level (see Supplementary Methods for details on data processing). We defined influenza A subtype dominance in each season based on the proportion of influenza A virus (IAV) positive samples typed as A(H3N2). We defined seasons as A(H3N2) or A(H1N1) or A(H1N1)pdm09 dominant when ≥70% of IAV positive samples were typed as one IAV subtype and co-dominant when one IAV subtype comprised 50-69% of IAV positive samples.

For each HHS region, we estimated weekly incidences of influenza A(H3N2), A(H1N1), and B by multiplying the percentage of influenza-like illness among outpatient visits, weighted by regional population, with the percentage of respiratory samples testing positive for a particular type/subtype [57,77]. We combined pre-2009 seasonal A(H1N1) and A(H1N1)pdm09 viruses as A(H1N1) and the Victoria and Yamagata lineages of influenza B as influenza B. ILI x percent positive (ILI^+) is considered a robust estimate of influenza activity and has been used in multiple prior modeling studies [6,18,57,77,122]. We used linear interpolation to estimate missing values for time spans of up to four consecutive weeks.

The emergence of A(H1N1)pdm09 in 2009 altered influenza testing and reporting patterns. We adjusted weekly incidences for differences in reporting rates between the pre-2009 pandemic period – defined as 1997 week 40 to 2009 week 17 – and the post-pandemic period – defined as the weeks after 2010 week 33. For each region, we scaled pre-pandemic incidences by the ratio of mean weekly ILI^+ (for all influenza type/subtypes combined) in the post-pandemic period to that of the pre-pandemic period. Incidences for HHS Region 10 were not adjusted for pre- and post-pandemic reporting because surveillance data for this region were not available before 2009. To account for differences in reporting rates across HHS regions, we next scaled each region's type/subtype incidences by its mean weekly ILI^+ for the entire study period. Scaled incidences were used in all downstream analyses of epidemic burden and timing.

Epidemic burden and timing

Epidemic burden: We considered three complementary indicators of epidemic burden, separately for each influenza type/subtype, HHS region, and season. We defined *peak incidence* as the maximum weekly scaled incidence and *epidemic size* as the cumulative weekly scaled incidence. We also estimated *epidemic intensity* based on a method previously developed to study variation in the shape (i.e., sharpness) of influenza epidemics across US cities [123]. Epidemic intensity was based on the inverse Shannon entropy of the weekly incidence distribution. Epidemic intensity increases when incidence is more concentrated in particular weeks and decreases when incidence is more evenly spread across weeks.

Specifically, we defined the incidence distribution p_{ij} as the fraction of influenza incidence in season j that occurred during week i in a given region, and epidemic intensity v_j as the inverse of the Shannon entropy of the incidence distribution:

$$v_j = \left(-\sum_i p_{ij} \ln p_{ij}\right)^{-1}$$

Epidemic intensity values were normalized to fall between 0 and 1.

Transmission intensity: For each region, we used the Epidemia R package to model annual A(H3N2) epidemics and to estimate time-varying (instantaneous) reproduction numbers, effective Rt [124,125](see Supplementary Methods for model details). Epidemia implements a semi-mechanistic Bayesian approach using the probabilistic programming language Stan [126].

To generate seasonal indicators of transmission intensity, we extracted posterior draws of daily Rt estimates for each region and season, calculated the median value for each day, and averaged daily median values by epidemic week. For each region and season, we averaged Rt estimates from the weeks spanning epidemic onset to epidemic peak (*initial Rt*) and averaged the two highest Rt estimates (*maximum Rt*). Initial Rt and maximum Rt produced qualitatively similar results in downstream analyses; we opted to report results for maximum Rt.

Excess pneumonia and influenza deaths attributable to A(H3N2): To measure the epidemic severity each season, we obtained estimates of seasonal excess mortality attributable to influenza A(H3N2) from Hansen et al., 2022 [127]. Excess mortality is a measure of the mortality burden of a given pathogen in excess of a seasonally adjusted baseline, obtained by regressing weekly deaths from broad disease categories against indicators of influenza virus circulation. Hansen et al. used pneumonia and influenza (P&I) excess deaths, which is considered the most specific indicator of influenza burden [128]. Deaths with a mention of P&I (ICD 10: J00-J18) were aggregated by week and age group (<1, 1-4, 5-49, 50-64, and ≥65) for 1998-2018. Age-specific generalized linear models were fit to observed weekly P&I death rates, while accounting for influenza and respiratory syncytial virus (RSV) activity and seasonal and temporal trends. Hansen et al. estimated the weekly national number of excess A(H3N2)-associated deaths by subtracting the baseline death rate expected in the absence of A(H3N2) circulation (A(H3N2) model terms set to zero) from the observed P&I death rate. We summed the number of excess A(H3N2) deaths per 100,000 people from October to May to obtain seasonal age-specific estimates.

Epidemic onset and peak timing: We estimated the regional onsets of A(H1N1), A(H1N1)pdm09, A(H3N2), and B epidemics each season by fitting piecewise linear models to subtype-specific incidence curves from week 30 to the first week of maximum incidence. We did not estimate epidemic onsets for regions with insufficient signal, which we defined as fewer than three weeks of consecutive incidence and/or greater than 30% of weeks with missing data in a particular season. The timing of the changepoint in incidence represents epidemic establishment (i.e., sustained transmission) rather the timing of influenza introduction or arrival [16]. We were able to estimate A(H3N2) onset timing for most seasons, except for three A(H1N1) dominant seasons: 2000-2001 (0 regions), 2002-2003 (3 regions), and 2009-2010 (0 regions). We defined epidemic peak timing as the first week of maximum incidence. To measure spatiotemporal synchrony, we calculated seasonal variation (standard deviation, s.d.) in regional onset and peak timing [19,27]. To measure the speed of viral spread, we calculated the number of days between onset and peak and seasonal duration (the number of weeks with non-zero incidence) for each region. As a sensitivity analysis, we used wavelets to estimate timing differences between A(H3N2), A(H1N1), A(H1N1)pdm09, and B epidemics (see Supplementary Methods).

Age patterns: We calculated the seasonal proportion of ILI encounters in each age group (0-4 years, 5-24 years, 25-64 years, and ≥65 years). Data for more narrow age groups are available after 2009, but we chose these four categories to increase the number of seasons in our analysis.

Influenza vaccination coverage and A(H3N2) vaccine effectiveness

Influenza vaccination coverage and effectiveness vary between years and would be expected to affect the population impact of seasonal outbreaks, and in turn our epidemiologic indicators. We obtained seasonal estimates of national vaccination coverage for adults 18-49 years and adults ≥65 years from studies utilizing vaccination questionnaire data collected by the National Health Interview Survey [129-135]. We did not consider the effects of vaccination coverage in children, due to our inability to find published estimates for most influenza seasons in our study.

We obtained seasonal estimates of adjusted A(H3N2) vaccine effectiveness (VE) from 32 observational studies [136-167]. Most of these studies had case-control test-negative designs (N = 30) and took place in North America (N = 25) or Europe (N = 6). When possible, we limited VE estimates to those for healthy adults or general populations. When multiple VE studies were available for a given season, we calculated mean VE as the weighted average of *m* different VE point estimates:

$$\frac{\sum_{i=1}^{m} \delta_{VE_i}{}^{-1/2} VE_i}{\sum_{i=1}^{m} \delta_{VE_i}{}^{-1/2}}$$

Wherein δ_{VE} denotes the width of the 95% confidence interval (CI) for VE_i [91].

The 95% CI for the weighted mean VE was calculated as:

$$\frac{1}{m}\sqrt{\sum_{i=1}^{m} \left(\delta_{VE_i}\right)^2}$$

Correlations among epidemic metrics

We used Spearman's correlation coefficients to measure pairwise relationships between A(H3N2) epidemiological indictors. We adjusted P-values for multiple testing using the Benjamini and Hochberg method [168].

Indicators of influenza A(H3N2) evolution

We considered multiple indicators of influenza evolution based on genetic and phenotypic (serologic) data, separately for HA and NA.

HA and NA sequence data

We downloaded all H3 sequences and associated metadata from the Global Initiative on Sharing Avian Influenza Data (GISAID) EpiFlu database [60]. We focused our analysis on complete H3 sequences that were sampled between January 1, 1997, and October 1, 2019. We prioritized viruses with corresponding HI titer measurements provided by the WHO Global Influenza Surveillance and Response System (GISRS) Collaborating Centers and excluded all egg-passaged viruses and sequences with ambiguous year, month, and day annotations. To account for variation in sequence availability across global regions, we subsampled the selected sequences five times to representative sets of 50 viruses per month, with preferential sampling for North America. Each month 25 viruses (when available) were selected from North America, with even sampling across nine other global regions (Africa, Europe, China, South Asia, Japan and Korea, Oceania, South America, Southeast Asia, and West Asia) for the remaining 25 viruses. To ensure proper topology early in the phylogeny, we included reference strains that had been collected no earlier than 5 years prior to January 1, 1997. The resultant sets of H3 sequences included 10,088 to 10,090 sequences spanning December 25, 1995 – October 1, 2019.

720 As with the H3 analysis, we downloaded all N2 sequences and associated metadata from GISAID and
721 selected complete N2 sequences that were sampled between January 1, 1997, and October 1, 2019. We
722 excluded all sequences with ambiguous year, month, and day annotations, forced the inclusion of
723 reference strains collected no earlier than 5 years prior to January 1, 1997, and compiled five replicate
724 subsampled datasets with preferential sampling for North America (9,007 to 9,009 sequences; June 8,
725 1995 – October 1, 2019).

726 ***HA serologic data***

727 Hemagglutination inhibition (HI) measurements from ferret sera were provided by WHO GISRS
728 Collaborating Centers in London, Melbourne, Atlanta, and Tokyo. We converted these raw two-fold
729 dilution measurements to log_2 titer drops normalized by the corresponding log_2 autologous measurements
730 [35,63].

731 Although a phenotypic assay exists for NA, NA inhibiting antibody titers are not routinely measured for
732 influenza surveillance. Therefore, we could not include a phenotypic marker of NA evolution in our study.

733 ***Phylogenetic inference***

734 For each set of H3 and N2 sequences, we aligned sequences with the augur align command [169] and
735 MAFFT v7.407 [170]. We inferred initial phylogenies with IQ-TREE v1.6.10 [171]. To reconstruct time-
736 resolved phylogenies, we applied TreeTime v0.5.6 [172] with the augur refine command [173].

737 ***Viral fitness metrics***

738 Following Huddleston et al., 2020 [35], we defined the following fitness metrics for each influenza season:

739 Antigenic drift**:** We estimated antigenic drift for each H3 strain using either genetic or serologic data. We
740 implemented three sequence-based metrics based on substitutions at putative epitope sites: 129 sites in
741 HA1 [21,64,66,67,174], 7 sites adjacent to the receptor-binding site (RBS) [68], and 34 sites in the HA
742 stalk [70], hereon *HA epitope distance*, *HA RBS distance*, and *HA stalk footprint distance*. To estimate
743 antigenic drift with hemagglutination inhibition (HI) titer data, hereon *HI log_2 titer distance*, we applied the
744 phylogenetic tree model from Neher et al., 2016 [63] to the H3 phylogeny and available HI data for its
745 sequences. The tree model estimates the antigenic drift per branch in units of log_2 titer change.

746 To estimate N2 antigenic drift, we implemented two sequence-based metrics that count substitutions at
747 putative epitope sites in the NA head: 223 sites [34] or 53 sites [69], hereon *NA epitope distance.*

748 Mutational load: To estimate mutational load for each H3 and N2 strain, an inverse proxy of viral fitness
749 [61], we implemented metrics that count substitutions at putative non-epitope sites in HA (N = 200) and
750 NA (N = 246), hereon *HA non-epitope distance* and *NA non-epitope distance*. Mutational load metrics
751 produce higher values for strains that are less fit compared to previously circulating strains.

752 Clade growth: We estimated the seasonal growth of H3 clades and N2 clades with the local branching
753 index (LBI) [62]. To calculate LBI for each H3 and N2 strain, we applied the LBI heuristic algorithm as
754 originally described by Neher et al., 2014 [62] to H3 and N2 phylogenetic trees, respectively. We set the
755 neighborhood parameter, τ, to 0.4 and only considered viruses sampled between the current season t
756 and the previous season t - 1 as contributing to recent clade growth in the current season t. To estimate
757 the diversity of clade growth rates in each season, we binned LBI values by units of 2 into 10 categories
758 ((0-2],(2-4], (4-6], (6-8], (8-10], (10-12],(12-14], (14-16],(16-18], (18-20]) and estimated the Shannon
759 entropy of LBI categories. Here, the Shannon entropy [175] considers both the richness and relative
760 abundance of viral clades with different growth rates and is calculated as follows:

$$H' = -\sum_i p_i \ln p_i$$

wherein p_i is the proportion of LBI values belonging to the ith bin.

Antigenic and genetic distance relative to prior seasons

We estimated genetic and antigenic distances between influenza viruses circulating in consecutive seasons by calculating the mean distance between viruses circulating in the current season t and viruses circulating during the prior season ($t-1$ year; one season lag) or two prior seasons ago ($t-2$ years; two season lag). Seasonal genetic and antigenic distances are greater when currently circulating strains are more antigenically distinct from previously circulating strains. We used Spearman's correlation coefficients to measure pairwise relationships between scaled H3 and N2 evolutionary indictors. We adjusted P-values for multiple testing using the Benjamini and Hochberg method [168].

Univariate relationships between viral fitness, (sub)type interference and A(H3N2) epidemic impact

We measured univariate associations between national indicators of A(H3N2) viral fitness and regional A(H3N2) epidemic parameters – peak incidence, epidemic size, effective Rt, epidemic intensity, subtype dominance, excess P&I deaths, onset timing, peak timing, spatiotemporal synchrony, the number of weeks from onset to peak, and seasonal duration. We first measured Spearman correlation coefficients between pairs of scaled fitness indicators and epidemic metrics using 1000 bootstrap replicates of the original dataset (1000 samples with replacement).

Next, we fit regression models with different distribution families (Gaussian or Gamma) and link functions (identity, log, or inverse) to observed data and used Bayesian information criterion (BIC) to select the best fit model, with lower BIC values indicating a better fit to the data. For subtype dominance, epidemic intensity, and age-specific proportions of ILI cases, we fit Beta regression models with logit links. For each epidemic metric, we fit the best-performing regression model to the resampled dataset. To measure the effects of sub(type) interference on A(H3N2) epidemics, the same approach was applied to measure the univariate relationships between A(H1N1) or B epidemic size and A(H3N2) peak incidence, epidemic size, effective Rt, epidemic intensity, and excess mortality. As a sensitivity analysis, we tested univariate relationships between A(H3N2) epidemic metrics and A(H1N1) epidemic size during pre-2009 seasons (seasonal A(H1N1) viruses) and post-2009 seasons (A(H1N1)pdm09 viruses) separately.

All predictors were centered and scaled prior to measuring Spearman's correlations or fitting regression models.

Selecting relevant predictors of A(H3N2) epidemic impact

Next, we explored multivariable approaches that would shed light on the potential mechanisms driving annual epidemic impact. Considering that we had many predictors and relatively few observations (22 seasons x 9-10 HHS regions), several covariates were collinear, and our goal was explicative rather than predictive, we settled on methods that tend to select few covariates.

We first used conditional inference random forest models to select relevant predictors of A(H3N2) epidemic size, peak incidence, effective Rt, epidemic intensity, and subtype dominance (party and caret R packages) [176-179]. Candidate predictors included viral fitness indicators: genetic and antigenic distance from previously circulating strains and the Shannon entropy of H3 and N2 LBI values in the current and prior season; proxies for prior natural immunity: A(H3N2) epidemic size in the prior season, influenza A(H1N1) epidemic size and B epidemic size in the current and prior seasons, and the dominant sub(type) in the prior season [12]; and vaccine-related parameters: national adult vaccination coverage in

the current and previous season, A(H3N2) vaccine effectiveness in the current and previous season, and H3 and N2 epitope distances between circulating A(H3N2) viruses in the United States and the A(H3N2) vaccine strain in the same season. We did not conduct variable selection analysis for excess A(H3N2) mortality due to data limitations (one national estimate per season). Metrics related to epidemic timing were also excluded from this analysis because we found weak or non-statistically significant associations with most of the candidate evolutionary predictors in univariate analyses.

We created each forest by generating 3,000 regression trees from 10 repeats of a leave-one-season-out (jackknife) cross-validated sample of the data. Due to the small size of our dataset, evaluating the predictive accuracy of random forest models on a quasi-independent test set produced unstable estimates. Consequently, we included all data in the training set and report root mean squared error (RMSE) and R^2 values from the best tuned model. We used permutation importance (N = 50 permutations) to estimate the relative importance of each predictor in determining model outcomes. Permutation importance is the decrease in prediction accuracy when a single feature (predictor) is randomly permuted, with larger values indicating more important variables. Because our features were collinear, we used conditional permutation importance to compute feature importance scores, rather than the standard marginal procedure [177,178,180,181].

As an alternative method for variable selection, we performed LASSO regression on the same cross-validated dataset and report RMSE and R^2 values from the best tuned model (glmnet and caret R packages)[179,182]. Unlike random forest models, this approach assumes linear relationships between predictors and the target variable. LASSO models (L1 penalty) are more restrictive than ridge models (L2 penalty) and elastic net models (combination of L1 and L2 penalties) and will arbitrarily select one variable from a set of collinear variables.

To further reduce the set of predictors for each epidemic metric, we performed model selection with linear regression models that considered all combinations of the top 10 ranked predictors from conditional inference random forest models. Candidate models were limited to three independent variables, and models were compared using BIC. We did not include HHS region or season as fixed or random effects in these models because these variables either did not improve model fit (region) or caused convergence issues (season).

All predictors were centered and scaled prior to fitting random forest or regression models.

Data availability

Sequence data are available from GISAID using accession ids provided in Supplementary file 1. Source code for phylogenetic analyses, inferred HI titers from serological measurements, and evolutionary fitness measurements are available in the GitHub repository https://github.com/blab/perofsky-ili-antigenicity. The five replicate trees for HA and NA can be found at https://nextstrain.org/groups/blab/ under the keyword "perofsky-ili-antigenicity". Epidemiological data, datasets combining seasonal evolutionary fitness measurements and epidemic metrics, and source code for calculating epidemic metrics and performing statistical analyses are available in the GitHub repository https://github.com/aperofsky/H3N2_Antigenic_Epi. Raw serological measurements are restricted from public distribution by previous data sharing agreements.

Acknowledgements

We thank the Influenza Division at the US Centers for Disease Control and Prevention, the Victorian Infectious Diseases Reference Laboratory at the Australian Peter Doherty Institute for Infection and Immunity, the Influenza Virus Research Center at the Japan National Institute of Infectious Diseases, the Crick Worldwide Influenza Centre at the UK Francis Crick Institute for sharing HI titer data. We gratefully acknowledge the authors, originating and submitting laboratories of the sequences from the GISAID

EpiFlu Database on which this research is based (listed in Appendix 1). We thank members of the Fogarty International Center's Division of International Epidemiology and Population Studies (DIEPS) and the Bedford Lab for useful discussions.

Funding information

ACP, CH, and CV were supported by the in-house research division of the Fogarty International Center, US National Institutes of Health. ACP was supported by the NSF Infectious Disease Evolution Across Scales (IDEAS) Research Collaboration Network. JH was supported by NIH NIAID awards F31 AI140714 and R01 AI165821. The work done at the Crick Worldwide Influenza Centre was supported by the Francis Crick Institute receiving core funding from Cancer Research UK (FC001030), the Medical Research Council (FC001030) and the Wellcome Trust (FC001030). SF, KN, NK, SW and HH were supported by the Ministry of Health, Labour and Welfare, Japan (10110400 and 10111800). SW was supported by the Japan Agency for Medical Research and Development (JP22fk0108118 and JP23fk0108662). The WHO Collaborating Centre for Reference and Research on Influenza is supported by the Australia Government Department of Health and Aged Care. The Melbourne WHO Collaborating Centre for Reference and Research on Influenza is supported by the Australian Government Department of Health. Influenza virus work in the Krammer laboratory was partially supported by the NIAID Centers of Excellence for Influenza Research and Surveillance (CEIRS) contract HHSN272201400008C, NIAID Centers of Excellence for Influenza Research and Response (CEIRR) contract 75N93021C00014 (FK), and NIAID CIVIC contract (75N93019C00051). TB was supported by NIH awards NIGMS R35 GM119774 and NIAID R01 AI127893. TB is an Investigator of the Howard Hughes Medical Institute. Funding sources were not involved in study design, data collection and interpretation, or the decision to submit the work for publication.

Disclaimer

The conclusions of this study do not necessarily represent the views of the National Institutes of Health, the Centers for Disease Control and Prevention, or the US government.

Author contributions

Amanda C Perofsky: Conceptualization, Data curation, Software, Formal analysis, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing - original draft, Project administration, Writing - review and editing; John Huddleston: Data curation, Software, Formal Analysis, Validation, Investigation, Visualization, Methodology, Writing - review and editing; Chelsea Hansen: Data curation, Software, Formal Analysis, Investigation, Writing – review and editing; John R Barnes, Thomas Rowe, Xiyan Xu, Rebecca Kondor, David E Wentworth, Nicola Lewis, Lynne Whittaker, Burcu Ermetal, Ruth Harvey, Monica Galiano, Rodney Stuart Daniels, John W McCauley, Seiichiro Fujisaki, Kazuya Nakamura, Noriko Kishida, Shinji Watanabe, Hideki Hasegawa, Sheena G Sullivan, Ian Barr, Kanta Subbarao: Resources, Investigation, Methodology, Writing - review and editing; Florian Krammer: Data curation, Resources, Investigation, Funding acquisition, Writing - review and editing; Trevor Bedford: Conceptualization, Resources, Software, Supervision, Methodology, Project administration, Funding acquisition; Cécile Viboud: Conceptualization, Resources, Supervision, Methodology, Project administration, Funding acquisition, Writing - review and editing

Competing interests

The WHO Collaborating Centre for Reference and Research on Influenza in Melbourne has a collaborative research and development agreement (CRADA) with CSL Seqirus for isolation of candidate vaccine viruses in cells and an agreement with IFPMA for isolation of candidate vaccine viruses in eggs. SGS reports honoraria from CSL Seqirus, Moderna, Pfizer, and Evo Health. The Icahn School of Medicine at Mount Sinai has filed patent applications relating to influenza virus vaccines, SARS-CoV-2 serological assays, and SARS-CoV-2 vaccines which list FK as co-inventor. Mount Sinai has spun out companies, Kantaro and Castlevax, to market the SARS-CoV-2 related technologies. FK has consulted for Merck and Pfizer (before 2020), and is currently consulting for Pfizer, Seqirus, 3rd Rock Ventures, GSK and Avimex. The Krammer laboratory is also collaborating with Pfizer on animal models of SARS-

CoV-2 and with Dynavax on universal influenza virus vaccines. All other authors declare no competing interests.

Supplementary Methods

Influenza virological surveillance data

Data on weekly influenza type and subtype circulation were obtained from the US CDC's World Health Organization (WHO) Collaborating Center for Surveillance, Epidemiology and Control of Influenza [121]. Approximately 100 public health laboratories and 300 clinical laboratories located throughout the United States report influenza test results to the US CDC, through either the US WHO Collaborating Laboratories Systems or the National Respiratory and Enteric Virus Surveillance System (NREVSS). Clinical laboratories test respiratory specimens for diagnostic purposes whereas public health laboratories primarily test specimens to characterize influenza virus type, subtype, and lineage circulation. Public health laboratories often receive samples that have already tested positive for influenza at a clinical laboratory.

We estimated the weekly number of respiratory samples testing positive for influenza A(H1N1), A(H3N2), or B at the HHS region level. Beginning in the 2015/2016 season, reports from public health and clinical laboratories are presented separately in the CDC's weekly influenza updates. From 2015 week 40 onwards, we used clinical laboratory data to estimate the proportion of respiratory samples testing positive for any influenza type/subtype and the proportion of samples testing positive for influenza A or B. We used public health laboratory data to estimate the proportion of influenza A isolates typed as A(H3N2) or A(H1N1)pdm09 in each week. Untyped influenza A-positive isolates were assigned to either A(H3N2) or A(H1N1) according to their proportions among typed isolates. We combined seasonal and pandemic A(H1N1) as seasonal A(H1N1) influenza and the Victoria and Yamagata lineages of influenza B as influenza B. We defined influenza A subtype dominance in each season based on the proportion of influenza A positive samples typed as A(H1N1) or A(H3N2).

A(H3N2) epidemiological model

Prior to R_t estimation, we computed daily case counts by disaggregating weekly A(H3N2) incidence rates to daily rates (tempdisagg package) [183] and rounding the resultant values to integers. Observed cases were modelled as a function of latent infections in the population, assuming a negative binomial distribution. We assumed an infection ascertainment rate of 0.45 [184], a lognormal-distributed infection-to-symptom-onset time period with mean 1.4 days and standard deviation 1.5 days [185], and a lognormal-distributed onset-to-case-observation time period with mean 2 days and standard deviation 1.5 days [186]. Thus, the time distribution for infection-to-case-observation was

$$\pi \sim \mathrm{lognormal}(1.4,1.5) + \mathrm{lognormal}(2,1.5)$$

Instead of using the renewal equation to propagate infections, we treated infections as latent parameters in the model, because the additional variance around infections leads to a posterior distribution that is easier to sample [125]. For the generation time, we assumed a discretized Weibull distribution with mean 3.6 days and standard deviation 1.6 days [187]. To control for temporal autocorrelation, we modelled R_t as a daily random walk. We assigned the intercept a normal prior with mean $\log 2$ and variance 0.2, which gives the initial reproduction number R_0 a prior mean of approximately 2.

Epidemic trajectories for each region and season were fit independently using Stan's Hamiltonian Monte Carlo sampler [188]. For each model, we ran 4 chains, each for 10,000 iterations (including a burn-in period of 2,000 iterations that was discarded), producing a total posterior sample size of 32,000. We verified convergence by confirming that all parameters had sufficiently low R hat values (all R hat < 1.1) and sufficiently large effective sample sizes (>15% of the total sample size).

957 **Wavelet analysis**

958 We applied a wavelet approach to quantify the relative timing of influenza A(H3N2), A(H1N1), and B
959 epidemics in each HHS region. Incidence time series were square root transformed and normalized and
960 then padded with zeros to reduce edge effects. Wavelet coherence was used to determine the degree of
961 synchrony between A(H3N2) versus A(H1N1) incidence and A(H3N2) versus B incidence within each
962 region at multi-year time scales. Statistical significance was assessed using 10,000 Monte Carlo
963 simulations. Coherence measures time- and frequency-specific associations between two wavelet
964 transforms, with high coherence indicating that two non-stationary signals (time series) are associated at
965 a particular time and frequency [82].

966 Following methodology developed for influenza and other viruses [19,82,189-191], we used continuous
967 wavelet transformations (Morlet) to calculate the phase of seasonal A(H3N2), A(H1N1), and B epidemics.
968 We reconstructed weekly time series of phase angles using wavelet reconstruction [19,192] and extracted
969 the major one-year seasonal component (period 0.8 to 1.2 years) of the Morlet decomposition of
970 A(H3N2), A(H1N1), and B time series. To estimate the relative timing of A(H3N2) and A(H1N1) incidence
971 or A(H3N2) and B incidence in each region, phase angle differences were calculated as phase in
972 A(H3N2) minus phase in A(H1N1) (or B), with a positive value indicating that A(H1N1) (or B) lags
973 A(H3N2).

974 **References**

975

976 1. Gerdil, C. (2003) The annual production cycle for influenza vaccine. *Vaccine* 21, 1776-1779.
977 10.1016/s0264-410x(03)00071-9
978 2. He, D. *et al.* (2015) Global Spatio-temporal Patterns of Influenza in the Post-pandemic Era. *Sci*
979 *Rep* 5, 11013. 10.1038/srep11013
980 3. Wraith, S. *et al.* (2022) Homotypic protection against influenza in a pediatric cohort in Managua,
981 Nicaragua. *Nat Commun* 13, 1190. 10.1038/s41467-022-28858-9
982 4. Hay, A.J. *et al.* (2001) The evolution of human influenza viruses. *Philos Trans R Soc Lond B Biol*
983 *Sci* 356, 1861-1870. 10.1098/rstb.2001.0999
984 5. Ferguson, N.M. *et al.* (2005) Strategies for containing an emerging influenza pandemic in
985 Southeast Asia. *Nature* 437, 209-214. 10.1038/nature04017
986 6. Bedford, T. *et al.* (2014) Integrating influenza antigenic dynamics with molecular evolution. *Elife* 3,
987 e01914. 10.7554/eLife.01914
988 7. Bedford, T. *et al.* (2015) Global circulation patterns of seasonal influenza viruses vary with
989 antigenic drift. *Nature* 523, 217-220. 10.1038/nature14460
990 8. Simonsen, L. (1999) The global impact of influenza on morbidity and mortality. *Vaccine* 17, S3-
991 S10. 10.1016/S0264-410X(99)00099-7
992 9. Viboud, C. *et al.* (2004) Influenza epidemics in the United States, France, and Australia, 1972-
993 1997. *Emerg Infect Dis* 10, 32-39. 10.3201/eid1001.020705
994 10. Chattopadhyay, I. *et al.* (2018) Conjunction of factors triggering waves of seasonal influenza. *Elife*
995 7. 10.7554/eLife.30756
996 11. Kramer, S.C. and Shaman, J. (2019) Development and validation of influenza forecasting for 64
997 temperate and tropical countries. *PLoS Comput Biol* 15, e1006742. 10.1371/journal.pcbi.1006742
998 12. Lee, E.C. *et al.* (2018) Deploying digital health data to optimize influenza surveillance at national
999 and local scales. *PLoS Comput Biol* 14, e1006020. 10.1371/journal.pcbi.1006020
1000 13. Shaman, J. *et al.* (2010) Absolute humidity and the seasonal onset of influenza in the continental
1001 United States. *PLoS Biol* 8, e1000316. 10.1371/journal.pbio.1000316
1002 14. Shaman, J. and Kohn, M. (2009) Absolute humidity modulates influenza survival, transmission,
1003 and seasonality. *Proc Natl Acad Sci U S A* 106, 3243-3248. 10.1073/pnas.0806852106
1004 15. Bedford, T. *et al.* (2010) Global migration dynamics underlie evolution and persistence of human
1005 influenza A (H3N2). *PLoS Pathog* 6, e1000918. 10.1371/journal.ppat.1000918
1006 16. Charu, V. *et al.* (2017) Human mobility and the spatial transmission of influenza in the United
1007 States. *PLoS Comput Biol* 13, e1005382. 10.1371/journal.pcbi.1005382

17. Geoghegan, J.L. *et al.* (2018) Continental synchronicity of human influenza virus epidemics despite climatic variation. *PLoS Pathog* 14, e1006780. 10.1371/journal.ppat.1006780
18. Pei, S. *et al.* (2018) Forecasting the spatial transmission of influenza in the United States. *Proc Natl Acad Sci U S A* 115, 2752-2757. 10.1073/pnas.1708856115
19. Viboud, C. *et al.* (2006) Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* 312, 447-451. 10.1126/science.1125237
20. Nelson, M.I. and Holmes, E.C. (2007) The evolution of epidemic influenza. *Nat Rev Genet* 8, 196-205. 10.1038/nrg2053
21. Wiley, D.C. *et al.* (1981) Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* 289, 373-378. 10.1038/289373a0
22. Fiore, A.E. *et al.* (2009) Prevention and control of seasonal influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP), 2009. *MMWR Recomm Rep* 58, 1-52
23. Boni, M.F. *et al.* (2004) Influenza drift and epidemic size: the race between generating and escaping immunity. *Theor Popul Biol* 65, 179-191. 10.1016/j.tpb.2003.10.002
24. Greene, S.K. *et al.* (2006) Patterns of influenza-associated mortality among US elderly by geographic region and virus subtype, 1968-1998. *Am J Epidemiol* 163, 316-326. 10.1093/aje/kwj040
25. Koelle, K. *et al.* (2009) Understanding the dynamics of rapidly evolving pathogens through modeling the tempo of antigenic change: influenza as a case study. *Epidemics* 1, 129-137. 10.1016/j.epidem.2009.05.003
26. Koelle, K. *et al.* (2006) Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science* 314, 1898-1903. 10.1126/science.1132745
27. Wolf, Y.I. *et al.* (2010) Projection of seasonal influenza severity from sequence and serological data. *PLoS Curr* 2, RRN1200. 10.1371/currents.RRN1200
28. Wu, A. *et al.* (2010) Correlation of influenza virus excess mortality with antigenic variation: application to rapid estimation of influenza mortality burden. *PLoS Comput Biol* 6. 10.1371/journal.pcbi.1000882
29. Axelsen, J.B. *et al.* (2014) Multiannual forecasting of seasonal influenza dynamics reveals climatic and evolutionary drivers. *Proc Natl Acad Sci U S A* 111, 9538-9542. 10.1073/pnas.1321656111
30. Du, X. *et al.* (2017) Evolution-informed forecasting of seasonal influenza A (H3N2). *Sci Transl Med* 9. 10.1126/scitranslmed.aan5325
31. Lam, E.K.S. *et al.* (2020) The impact of climate and antigenic evolution on seasonal influenza virus epidemics in Australia. *Nat Commun* 11, 2741. 10.1038/s41467-020-16545-6
32. Smith, D.J. *et al.* (2004) Mapping the antigenic and genetic evolution of influenza virus. *Science* 305, 371-376. 10.1126/science.1097211
33. Bedford, T. *et al.* (2011) Strength and tempo of selection revealed in viral gene genealogies. *BMC Evol Biol* 11, 220. 10.1186/1471-2148-11-220
34. Bhatt, S. *et al.* (2011) The genomic rate of molecular adaptation of the human influenza A virus. *Mol Biol Evol* 28, 2443-2451. 10.1093/molbev/msr044
35. Huddleston, J. *et al.* (2020) Integrating genotypes and phenotypes improves long-term forecasts of seasonal influenza A/H3N2 evolution. *Elife* 9. 10.7554/eLife.60067
36. Shih, A.C. *et al.* (2007) Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. *Proc Natl Acad Sci U S A* 104, 6283-6288. 10.1073/pnas.0701396104
37. Suzuki, Y. (2008) Positive selection operates continuously on hemagglutinin during evolution of H3N2 human influenza A virus. *Gene* 427, 111-116. 10.1016/j.gene.2008.09.012
38. Chen, Y.Q. *et al.* (2018) Influenza Infection in Humans Induces Broadly Cross-Reactive and Protective Neuraminidase-Reactive Antibodies. *Cell* 173, 417-429 e410. 10.1016/j.cell.2018.03.030

1059 39. Eichelberger, M.C. *et al.* (2018) Neuraminidase as an influenza vaccine antigen: a low hanging
1060 fruit, ready for picking to improve vaccine effectiveness. *Curr Opin Immunol* 53, 38-44.
1061 10.1016/j.coi.2018.03.025
1062 40. Wohlbold, T.J. *et al.* (2015) Vaccination with adjuvanted recombinant neuraminidase induces
1063 broad heterologous, but not heterosubtypic, cross-protection against influenza virus infection in
1064 mice. *mBio* 6, e02556. 10.1128/mBio.02556-14
1065 41. Brett, I.C. and Johansson, B.E. (2005) Immunization against influenza A virus: comparison of
1066 conventional inactivated, live-attenuated and recombinant baculovirus produced purified
1067 hemagglutinin and neuraminidase vaccines in a murine model system. *Virology* 339, 273-280.
1068 10.1016/j.virol.2005.06.006
1069 42. Couch, R.B. *et al.* (1974) Induction of partial immunity to influenza by a neuraminidase-specific
1070 vaccine. *J Infect Dis* 129, 411-420. 10.1093/infdis/129.4.411
1071 43. Johansson, B.E. *et al.* (1993) Infection-permissive immunization with influenza virus
1072 neuraminidase prevents weight loss in infected mice. *Vaccine* 11, 1037-1039. 10.1016/0264-
1073 410x(93)90130-p
1074 44. Kilbourne, E.D. (1976) Comparative efficacy of neuraminidase-specific and conventional
1075 influenza virus vaccines in induction of antibody to neuraminidase in humans. *J Infect Dis* 134,
1076 384-394. 10.1093/infdis/134.4.384
1077 45. Murphy, B.R. *et al.* (1972) Association of serum anti-neuraminidase antibody with resistance to
1078 influenza in man. *N Engl J Med* 286, 1329-1332. 10.1056/NEJM197206222862502
1079 46. Schulman, J.L. *et al.* (1968) Protective effects of specific immunity to viral neuraminidase on
1080 influenza virus infection of mice. *J Virol* 2, 778-786. 10.1128/JVI.2.8.778-786.1968
1081 47. Couch, R.B. *et al.* (2013) Antibody correlates and predictors of immunity to naturally occurring
1082 influenza in humans and the importance of antibody to the neuraminidase. *J Infect Dis* 207, 974-
1083 981. 10.1093/infdis/jis935
1084 48. Memoli, M.J. *et al.* (2016) Evaluation of Antihemagglutinin and Antineuraminidase Antibodies as
1085 Correlates of Protection in an Influenza A/H1N1 Virus Healthy Human Challenge Model. *mBio* 7,
1086 e00417-00416. 10.1128/mBio.00417-16
1087 49. Monto, A.S. *et al.* (2015) Antibody to Influenza Virus Neuraminidase: An Independent Correlate of
1088 Protection. *J Infect Dis* 212, 1191-1199. 10.1093/infdis/jiv195
1089 50. Grebe, K.M. *et al.* (2008) Heterosubtypic immunity to influenza A virus: where do we stand?
1090 *Microbes Infect* 10, 1024-1029. 10.1016/j.micinf.2008.07.002
1091 51. Ulmer, J.B. *et al.* (1998) Protective CD4+ and CD8+ T cells against influenza virus induced by
1092 vaccination with nucleoprotein DNA. *J Virol* 72, 5648-5653. 10.1128/JVI.72.7.5648-5653.1998
1093 52. Sridhar, S. *et al.* (2013) Cellular immune correlates of protection against symptomatic pandemic
1094 influenza. *Nat Med* 19, 1305-1312. 10.1038/nm.3350
1095 53. Epstein, S.L. (2006) Prior H1N1 influenza infection and susceptibility of Cleveland Family Study
1096 participants during the H2N2 pandemic of 1957: an experiment of nature. *J Infect Dis* 193, 49-53.
1097 10.1086/498980
1098 54. Sonoguchi, T. *et al.* (1985) Cross-subtype protection in humans during sequential, overlapping,
1099 and/or concurrent epidemics caused by H3N2 and H1N1 influenza viruses. *J Infect Dis* 151, 81-
1100 88. 10.1093/infdis/151.1.81
1101 55. Ferguson, N.M. *et al.* (2003) Ecological and immunological determinants of influenza evolution.
1102 *Nature* 422, 428-433. 10.1038/nature01509
1103 56. Cowling, B.J. *et al.* (2014) Incidence of influenza virus infections in children in Hong Kong in a 3-
1104 year randomized placebo-controlled vaccine study, 2009-2012. *Clin Infect Dis* 59, 517-524.
1105 10.1093/cid/ciu356
1106 57. Goldstein, E. *et al.* (2011) Predicting the epidemic sizes of influenza A/H1N1, A/H3N2, and B: a
1107 statistical method. *PLoS Med* 8, e1001051. 10.1371/journal.pmed.1001051
1108 58. Steinhoff, M.C. *et al.* (1993) Effect of heterosubtypic immunity on infection with attenuated
1109 influenza A virus vaccines in young children. *J Clin Microbiol* 31, 836-838. 10.1128/jcm.31.4.836-
1110 838.1993

1111 59. Gatti, L. *et al.* (2022) Cross-reactive immunity potentially drives global oscillation and opposed
1112 alternation patterns of seasonal influenza A viruses. *Sci Rep* 12, 8883. 10.1038/s41598-022-
1113 08233-w
1114 60. Shu, Y. and McCauley, J. (2017) GISAID: Global initiative on sharing all influenza data - from
1115 vision to reality. *Euro Surveill* 22. 10.2807/1560-7917.ES.2017.22.13.30494
1116 61. Luksza, M. and Lassig, M. (2014) A predictive fitness model for influenza. *Nature* 507, 57-61.
1117 10.1038/nature13087
1118 62. Neher, R.A. *et al.* (2014) Predicting evolution from the shape of genealogical trees. *Elife* 3.
1119 10.7554/eLife.03568
1120 63. Neher, R.A. *et al.* (2016) Prediction, dynamics, and visualization of antigenic phenotypes of
1121 seasonal influenza viruses. *Proc Natl Acad Sci U S A* 113, E1701-1709.
1122 10.1073/pnas.1525578113
1123 64. Bush, R.M. *et al.* (1999) Predicting the evolution of human influenza A. *Science* 286, 1921-1925.
1124 10.1126/science.286.5446.1921
1125 65. Webster, R.G. and Laver, W.G. (1980) Determination of the number of nonoverlapping antigenic
1126 areas on Hong Kong (H3N2) influenza virus hemagglutinin with monoclonal antibodies and the
1127 selection of variants with potential epidemiological significance. *Virology* 104, 139-148.
1128 10.1016/0042-6822(80)90372-4
1129 66. Wilson, I.A. and Cox, N.J. (1990) Structural basis of immune recognition of influenza virus
1130 hemagglutinin. *Annu Rev Immunol* 8, 737-771. 10.1146/annurev.iy.08.040190.003513
1131 67. Wolf, Y.I. *et al.* (2006) Long intervals of stasis punctuated by bursts of positive selection in the
1132 seasonal evolution of influenza A virus. *Biol Direct* 1, 34. 10.1186/1745-6150-1-34
1133 68. Koel, B.F. *et al.* (2013) Substitutions near the receptor binding site determine major antigenic
1134 change during influenza virus evolution. *Science* 342, 976-979. 10.1126/science.1244730
1135 69. Krammer, F. (2023) Unpublished.
1136 70. Kirkpatrick, E. *et al.* (2018) The influenza virus hemagglutinin head evolves faster than the stalk
1137 domain. *Sci Rep* 8, 10432. 10.1038/s41598-018-28706-1
1138 71. Krammer, F. (2019) The human antibody response to influenza A virus infection and vaccination.
1139 *Nat Rev Immunol* 19, 383-397. 10.1038/s41577-019-0143-6
1140 72. Margine, I. *et al.* (2013) H3N2 influenza virus infection induces broadly reactive hemagglutinin
1141 stalk antibodies in humans and mice. *J Virol* 87, 4728-4737. 10.1128/JVI.03509-12
1142 73. Nachbagauer, R. *et al.* (2016) Age Dependence and Isotype Specificity of Influenza Virus
1143 Hemagglutinin Stalk-Reactive Antibodies in Humans. *mBio* 7, e01996-01915.
1144 10.1128/mBio.01996-15
1145 74. Sandbulte, M.R. *et al.* (2011) Discordant antigenic drift of neuraminidase and hemagglutinin in
1146 H1N1 and H3N2 influenza viruses. *Proc Natl Acad Sci U S A* 108, 20748-20753.
1147 10.1073/pnas.1113801108
1148 75. Kilbourne, E.D. *et al.* (1990) Independent and disparate evolution in nature of influenza A virus
1149 hemagglutinin and neuraminidase glycoproteins. *Proc Natl Acad Sci U S A* 87, 786-790.
1150 10.1073/pnas.87.2.786
1151 76. Schulman, J.L. and Kilbourne, E.D. (1969) Independent variation in nature of hemagglutinin and
1152 neuraminidase antigens of influenza virus: distinctiveness of hemagglutinin antigen of Hong
1153 Kong-68 virus. *Proc Natl Acad Sci U S A* 63, 326-333. 10.1073/pnas.63.2.326
1154 77. Goldstein, E. *et al.* (2012) Improving the estimation of influenza-related mortality over a seasonal
1155 baseline. *Epidemiology* 23, 829-838. 10.1097/EDE.0b013e31826c2dda
1156 78. Gostic, K.M. *et al.* (2019) Childhood immune imprinting to influenza A shapes birth year-specific
1157 risk during seasonal H1N1 and H3N2 epidemics. *PLoS Pathog* 15, e1008109.
1158 10.1371/journal.ppat.1008109
1159 79. Webster, R.G. *et al.* (1992) Evolution and ecology of influenza A viruses. *Microbiological Reviews*
1160 56, 152-179. doi:10.1128/mr.56.1.152-179.1992
1161 80. Garten, R.J. *et al.* (2009) Antigenic and Genetic Characteristics of Swine-Origin 2009 A(H1N1)
1162 Influenza Viruses Circulating in Humans. *Science* 325, 197-201. doi:10.1126/science.1176225

1163 81. Smith, G.J. *et al.* (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1
1164 influenza A epidemic. *Nature* 459, 1122-1125. 10.1038/nature08182
1165 82. Johansson, M.A. *et al.* (2009) Multiyear climate variability and dengue--El Nino southern
1166 oscillation, weather, and dengue incidence in Puerto Rico, Mexico, and Thailand: a longitudinal
1167 data analysis. *PLoS Med* 6, e1000168. 10.1371/journal.pmed.1000168
1168 83. Bedford, T. *et al.* (2012) Canalization of the evolutionary trajectory of the human influenza virus.
1169 *BMC Biol* 10, 38. 10.1186/1741-7007-10-38
1170 84. Petrova, V.N. and Russell, C.A. (2018) The evolution of seasonal influenza viruses. *Nat Rev*
1171 *Microbiol* 16, 47-60. 10.1038/nrmicro.2017.118
1172 85. Koelle, K. and Rasmussen, D.A. (2015) The effects of a deleterious mutation load on patterns of
1173 influenza A/H3N2's antigenic evolution in humans. *Elife* 4, e07361. 10.7554/eLife.07361
1174 86. Kryazhimskiy, S. *et al.* (2011) Prevalence of epistasis in the evolution of influenza A surface
1175 proteins. *PLoS Genet* 7, e1001301. 10.1371/journal.pgen.1001301
1176 87. Gong, L.I. *et al.* (2013) Stability-mediated epistasis constrains the evolution of an influenza
1177 protein. *Elife* 2, e00631. 10.7554/eLife.00631
1178 88. Myers, J.L. *et al.* (2013) Compensatory hemagglutinin mutations alter antigenic properties of
1179 influenza viruses. *J Virol* 87, 11168-11172. 10.1128/JVI.01414-13
1180 89. Webster, R.G. *et al.* (1982) Molecular mechanisms of variation in influenza viruses. *Nature* 296,
1181 115-121. 10.1038/296115a0
1182 90. Altman, M.O. *et al.* (2015) Lamprey VLRB response to influenza virus supports universal rules of
1183 immunogenicity and antigenicity. *Elife* 4. 10.7554/eLife.07467
1184 91. Ndifon, W. *et al.* (2009) On the use of hemagglutination-inhibition for influenza surveillance:
1185 surveillance data are predictive of influenza vaccine effectiveness. *Vaccine* 27, 2447-2452.
1186 10.1016/j.vaccine.2009.02.047
1187 92. Hensley, S.E. (2014) Challenges of selecting seasonal influenza vaccine strains for humans with
1188 diverse pre-exposure histories. *Curr Opin Virol* 8, 85-89. 10.1016/j.coviro.2014.07.007
1189 93. Lee, J.M. *et al.* (2019) Mapping person-to-person variation in viral mutations that escape
1190 polyclonal serum targeting influenza hemagglutinin. *Elife* 8. 10.7554/eLife.49324
1191 94. Li, Y. *et al.* (2013) Immune history shapes specificity of pandemic H1N1 influenza antibody
1192 responses. *J Exp Med* 210, 1493-1500. 10.1084/jem.20130212
1193 95. Miller, M.S. *et al.* (2013) Neutralizing antibodies against previously encountered influenza virus
1194 strains increase over time: a longitudinal analysis. *Sci Transl Med* 5, 198ra107.
1195 10.1126/scitranslmed.3006637
1196 96. Henry, C. *et al.* (2019) Influenza Virus Vaccination Elicits Poorly Adapted B Cell Responses in
1197 Elderly Individuals. *Cell Host Microbe* 25, 357-366 e356. 10.1016/j.chom.2019.01.002
1198 97. Ranjeva, S. *et al.* (2019) Age-specific differences in the dynamics of protective immunity to
1199 influenza. *Nat Commun* 10, 1660. 10.1038/s41467-019-09652-6
1200 98. Cobey, S. and Hensley, S.E. (2017) Immune history and influenza virus susceptibility. *Curr Opin*
1201 *Virol* 22, 105-111. 10.1016/j.coviro.2016.12.004
1202 99. Zhang, A. *et al.* (2019) Original Antigenic Sin: How First Exposure Shapes Lifelong Anti-Influenza
1203 Virus Immune Responses. *J Immunol* 202, 335-340. 10.4049/jimmunol.1801149
1204 100. Zost, S.J. *et al.* (2017) Contemporary H3N2 influenza viruses have a glycosylation site that alters
1205 binding of antibodies elicited by egg-adapted vaccine strains. *Proc Natl Acad Sci U S A* 114,
1206 12578-12583. 10.1073/pnas.1712377114
1207 101. Xie, H. *et al.* (2015) H3N2 Mismatch of 2014-15 Northern Hemisphere Influenza Vaccines and
1208 Head-to-head Comparison between Human and Ferret Antisera derived Antigenic Maps. *Sci Rep*
1209 5, 15279. 10.1038/srep15279
1210 102. Cowling, B.J. *et al.* (2010) Protective efficacy of seasonal influenza vaccination against seasonal
1211 and pandemic influenza virus infection during 2009 in Hong Kong. *Clin Infect Dis* 51, 1370-1379.
1212 10.1086/657311
1213 103. Fox, S.J. *et al.* (2017) Seasonality in risk of pandemic influenza emergence. *PLoS Comput Biol*
1214 13, e1005749. 10.1371/journal.pcbi.1005749

104. Laurie, K.L. *et al.* (2015) Interval Between Infections and Viral Hierarchy Are Determinants of Viral Interference Following Influenza Virus Infection in a Ferret Model. *J Infect Dis* 212, 1701-1710. 10.1093/infdis/jiv260
105. Garten, R.J. *et al.* (2009) Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* 325, 197-201. 10.1126/science.1176225
106. Terajima, M. *et al.* (2013) Cross-reactive human B cell and T cell epitopes between influenza A and B viruses. *Virol J* 10, 244. 10.1186/1743-422x-10-244
107. Yang, W. *et al.* (2020) Dynamic interactions of influenza viruses in Hong Kong during 1998-2018. *PLoS Comput Biol* 16, e1007989. 10.1371/journal.pcbi.1007989
108. Cowling, B.J. *et al.* (2020) Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *Lancet Public Health* 5, e279-e288. 10.1016/S2468-2667(20)30090-6
109. Huang, Q.S. *et al.* (2021) Impact of the COVID-19 nonpharmaceutical interventions on influenza and other respiratory viral infections in New Zealand. *Nat Commun* 12, 1001. 10.1038/s41467-021-21157-9
110. Olsen, S.J. *et al.* (2020) Decreased influenza activity during the COVID-19 pandemic-United States, Australia, Chile, and South Africa, 2020. *Am J Transplant* 20, 3681-3685. 10.1111/ajt.16381
111. Olsen, S.J. *et al.* (2021) Changes in Influenza and Other Respiratory Virus Activity During the COVID-19 Pandemic - United States, 2020-2021. *MMWR Morb Mortal Wkly Rep* 70, 1013-1019. 10.15585/mmwr.mm7029a1
112. Qi, Y. *et al.* (2021) Quantifying the Impact of COVID-19 Nonpharmaceutical Interventions on Influenza Transmission in the United States. *J Infect Dis* 224, 1500-1508. 10.1093/infdis/jiab485
113. Tempia, S. *et al.* (2021) Decline of influenza and respiratory syncytial virus detection in facility-based surveillance during the COVID-19 pandemic, South Africa, January to October 2020. *Euro Surveill* 26. 10.2807/1560-7917.ES.2021.26.29.2001600
114. Ali, S.T. *et al.* (2022) Prediction of upcoming global infection burden of influenza seasons after relaxation of public health and social measures during the COVID-19 pandemic: a modelling study. *Lancet Glob Health* 10, e1612-e1622. 10.1016/S2214-109X(22)00358-8
115. Baker, R.E. *et al.* (2020) The impact of COVID-19 nonpharmaceutical interventions on the future dynamics of endemic infections. *Proc Natl Acad Sci U S A* 117, 30547-30553. 10.1073/pnas.2013182117
116. Gaglani, M. *et al.* (2016) Influenza Vaccine Effectiveness Against 2009 Pandemic Influenza A(H1N1) Virus Differed by Vaccine Type During 2013-2014 in the United States. *J Infect Dis* 213, 1546-1556. 10.1093/infdis/jiv577
117. Gill, P.W. and Murphy, A.M. (1977) Naturally acquired immunity to influenza type A: a further prospective study. *Med J Aust* 2, 761-765. 10.5694/j.1326-5377.1977.tb99276.x
118. Hope-Simpson, R.E. (1971) Hong Kong influenza variant. *Br Med J* 3, 531. 10.1136/bmj.3.5773.531-b
119. Krammer, F. *et al.* (2018) NAction! How Can Neuraminidase-Based Immunity Contribute to Better Influenza Virus Vaccines? *mBio* 9. 10.1128/mBio.02332-17
120. Centers for Disease Control and Prevention (2023). FluView Interactive. Centers for Disease Control and Prevention
121. World Health Organization (2023). FluNet.
122. Pei, S. *et al.* (2021) Optimizing respiratory virus surveillance networks using uncertainty propagation. *Nat Commun* 12, 222. 10.1038/s41467-020-20399-3
123. Dalziel, B.D. *et al.* (2018) Urbanization and humidity shape the intensity of influenza epidemics in U.S. cities. *Science* 362, 75-79. 10.1126/science.aat6030
124. Cori, A. *et al.* (2013) A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am J Epidemiol* 178, 1505-1512. 10.1093/aje/kwt133
125. Scott, J.A. *et al.* (2021) Epidemia: An R package for semi-mechanistic bayesian modelling of infectious diseases using point processes. *arXiv preprint arXiv:2110.12461*,

126. Carpenter, B. *et al.* (2017) Stan: A Probabilistic Programming Language. *Journal of Statistical Software* 76, 1 - 32. 10.18637/jss.v076.i01
127. Hansen, C.L. *et al.* (2022) Mortality Associated With Influenza and Respiratory Syncytial Virus in the US, 1999-2018. *JAMA Netw Open* 5, e220527. 10.1001/jamanetworkopen.2022.0527
128. Simonsen, L. and Viboud, C. (2012) The art of modeling the mortality impact of winter-seasonal pathogens. *J Infect Dis* 206, 625-627. 10.1093/infdis/jis419
129. Centers for Disease Control and Prevention (2019). Flu Vaccination Coverage, United States, 2018–19 Influenza Season | FluVaxView | Seasonal Influenza (Flu) | CDC.
130. Jang, S.H. and Kang, J. (2021) Factors Associated with Influenza Vaccination Uptake among U.S. Adults: Focus on Nativity and Race/Ethnicity. *Int J Environ Res Public Health* 18. 10.3390/ijerph18105349
131. Lu, P.J. *et al.* (2019) Seasonal Influenza Vaccination Coverage Trends Among Adult Populations, U.S., 2010-2016. *Am J Prev Med* 57, 458-469. 10.1016/j.amepre.2019.04.007
132. Lu, P.J. *et al.* (2013) Seasonal influenza vaccination coverage among adult populations in the United States, 2005-2011. *Am J Epidemiol* 178, 1478-1487. 10.1093/aje/kwt158
133. National Center for Health Statistics (2008). TABLE: Self-reported influenza vaccination coverage trends 1989-2008 among adults by age group, risk group, race/ethnicity, health-care worker status, and pregnancy status. National Health Interview Survey (NHIS)
134. Ward, B. *et al.* (2014). Early Release of Selected Estimates Based on Data From the 2014 National Health Interview Survey. In Statistics, N.C.f.H., ed.
135. Ward, B. *et al.* (2016). Early Release of Selected Estimates Based on Data From the 2015 National Health Interview Survey (05/2016). In Statistics, N.C.f.H., ed.
136. Belongia, E.A. *et al.* (2011) Influenza vaccine effectiveness in Wisconsin during the 2007-08 season: comparison of interim and final results. *Vaccine* 29, 6558-6563. 10.1016/j.vaccine.2011.07.002
137. Bridges, C.B. *et al.* (2000) Effectiveness and cost-benefit of influenza vaccination of healthy working adults: A randomized controlled trial. *JAMA* 284, 1655-1663. 10.1001/jama.284.13.1655
138. Castilla, J. *et al.* (2016) Effectiveness of subunit influenza vaccination in the 2014-2015 season and residual effect of split vaccination in previous seasons. *Vaccine* 34, 1350-1357. 10.1016/j.vaccine.2016.01.054
139. Flannery, B. *et al.* (2019) Influenza Vaccine Effectiveness in the United States During the 2016-2017 Season. *Clin Infect Dis* 68, 1798-1806. 10.1093/cid/ciy775
140. Flannery, B. *et al.* (2020) Spread of Antigenically Drifted Influenza A(H3N2) Viruses and Vaccine Effectiveness in the United States During the 2018-2019 Season. *J Infect Dis* 221, 8-15. 10.1093/infdis/jiz543
141. Flannery, B. *et al.* (2016) Enhanced Genetic Characterization of Influenza A(H3N2) Viruses and Vaccine Effectiveness by Genetic Group, 2014-2015. *J Infect Dis* 214, 1010-1019. 10.1093/infdis/jiw181
142. Jackson, M.L. *et al.* (2017) Influenza Vaccine Effectiveness in the United States during the 2015-2016 Season. *N Engl J Med* 377, 534-543. 10.1056/NEJMoa1700153
143. Janjua, N.Z. *et al.* (2012) Estimates of influenza vaccine effectiveness for 2007-2008 from Canada's sentinel surveillance system: cross-protection against major and minor variants. *J Infect Dis* 205, 1858-1868. 10.1093/infdis/jis283
144. Kawai, N. *et al.* (2003) A prospective, Internet-based study of the effectiveness and safety of influenza vaccination in the 2001-2002 influenza season. *Vaccine* 21, 4507-4513. 10.1016/s0264-410x(03)00508-5
145. Kissling, E. *et al.* (2013) Low and decreasing vaccine effectiveness against influenza A(H3) in 2011/12 among vaccination target groups in Europe: results from the I-MOVE multicentre case-control study. *Euro Surveill* 18. 10.2807/ese.18.05.20390-en
146. Lester, R.T. *et al.* (2003) Use of, effectiveness of, and attitudes regarding influenza vaccine among house staff. *Infect Control Hosp Epidemiol* 24, 839-844. 10.1086/502146
147. McLean, H.Q. *et al.* (2014) Impact of repeated vaccination on vaccine effectiveness against influenza A(H3N2) and B during 8 seasons. *Clin Infect Dis* 59, 1375-1385. 10.1093/cid/ciu680

148. Ohmit, S.E. *et al.* (2014) Influenza vaccine effectiveness in the 2011-2012 season: protection against each circulating virus and the effect of prior vaccination on estimates. *Clin Infect Dis* 58, 319-327. 10.1093/cid/cit736
149. Pebody, R. *et al.* (2017) End-of-season influenza vaccine effectiveness in adults and children, United Kingdom, 2016/17. *Euro Surveill* 22. 10.2807/1560-7917.ES.2017.22.44.17-00306
150. Centers for Disease Control and Prevention (2004) Assessment of the effectiveness of the 2003-04 influenza vaccine among children and adults--Colorado, 2003. *MMWR Morb Mortal Wkly Rep* 53, 707-710
151. Rolfes, M.A. *et al.* (2019) Effects of Influenza Vaccination in the United States During the 2017-2018 Influenza Season. *Clin Infect Dis* 69, 1845-1853. 10.1093/cid/ciz075
152. Simpson, C.R. *et al.* (2015) Trivalent inactivated seasonal influenza vaccine effectiveness for the prevention of laboratory-confirmed influenza in a Scottish population 2000 to 2009. *Euro Surveill* 20. 10.2807/1560-7917.es2015.20.8.21043
153. Skowronski, D. *et al.* (2005) Effectiveness of vaccine against medical consultation due to laboratory-confirmed influenza: results from a sentinel physician pilot project in British Columbia, 2004-2005. *Can Commun Dis Rep* 31, 181-191
154. Skowronski, D.M. *et al.* (2007) Estimating vaccine effectiveness against laboratory-confirmed influenza using a sentinel physician network: results from the 2005-2006 season of dual A and B vaccine mismatch in Canada. *Vaccine* 25, 2842-2851. 10.1016/j.vaccine.2006.10.002
155. Skowronski, D.M. *et al.* (2009) Component-specific effectiveness of trivalent influenza vaccine as monitored through a sentinel surveillance network in Canada, 2006-2007. *J Infect Dis* 199, 168-179. 10.1086/595862
156. Skowronski, D.M. *et al.* (2010) Association between the 2008-09 seasonal influenza vaccine and pandemic H1N1 illness during Spring-Summer 2009: four observational studies from Canada. *PLoS Med* 7, e1000258. 10.1371/journal.pmed.1000258
157. Skowronski, D.M. *et al.* (2012) A sentinel platform to evaluate influenza vaccine effectiveness and new variant circulation, Canada 2010-2011 season. *Clin Infect Dis* 55, 332-342. 10.1093/cid/cis431
158. Skowronski, D.M. *et al.* (2014) Low 2012-13 influenza vaccine effectiveness associated with mutation in the egg-adapted H3N2 vaccine strain not antigenic drift in circulating viruses. *PLoS One* 9, e92153. 10.1371/journal.pone.0092153
159. Skowronski, D.M. *et al.* (2014) Influenza A/subtype and B/lineage effectiveness estimates for the 2011-2012 trivalent vaccine: cross-season and cross-lineage protection with unchanged vaccine. *J Infect Dis* 210, 126-137. 10.1093/infdis/jiu048
160. Skowronski, D.M. *et al.* (2017) Interim estimates of 2016/17 vaccine effectiveness against influenza A(H3N2), Canada, January 2017. *Euro Surveill* 22. 10.2807/1560-7917.ES.2017.22.6.30460
161. Skowronski, D.M. *et al.* (2022) Influenza Vaccine Effectiveness by A(H3N2) Phylogenetic Subcluster and Prior Vaccination History: 2016-2017 and 2017-2018 Epidemics in Canada. *J Infect Dis* 225, 1387-1398. 10.1093/infdis/jiaa138
162. Skowronski, D.M. *et al.* (2016) A Perfect Storm: Impact of Genomic Variation and Serial Vaccination on Low Influenza Vaccine Effectiveness During the 2014-2015 Season. *Clin Infect Dis* 63, 21-32. 10.1093/cid/ciw176
163. Skowronski, D.M. *et al.* (2017) Serial Vaccination and the Antigenic Distance Hypothesis: Effects on Influenza Vaccine Effectiveness During A(H3N2) Epidemics in Canada, 2010-2011 to 2014-2015. *J Infect Dis* 215, 1059-1099. 10.1093/infdis/jix074
164. Treanor, J.J. *et al.* (2012) Effectiveness of seasonal influenza vaccines in the United States during a season with circulation of all three vaccine strains. *Clin Infect Dis* 55, 951-959. 10.1093/cid/cis574
165. Valenciano, M. *et al.* (2018) Exploring the effect of previous inactivated influenza vaccination on seasonal influenza vaccine effectiveness against medically attended influenza: Results of the European I-MOVE multicentre test-negative case-control study, 2011/2012-2016/2017. *Influenza Other Respir Viruses* 12, 567-581. 10.1111/irv.12562

166. van Doorn, E. *et al.* (2017) Influenza vaccine effectiveness estimates in the Dutch population from 2003 to 2014: The test-negative design case-control study with different control groups. *Vaccine* 35, 2831-2839. 10.1016/j.vaccine.2017.04.012
167. Zimmerman, R.K. *et al.* (2016) 2014-2015 Influenza Vaccine Effectiveness in the United States by Vaccine Type. *Clin Infect Dis* 63, 1564-1573. 10.1093/cid/ciw635
168. Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B* 57, 289-300. DOI 10.1111/j.2517-6161.1995.tb02031.x
169. Hadfield, J. *et al.* (2018) Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34, 4121-4123. 10.1093/bioinformatics/bty407
170. Katoh, K. *et al.* (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30, 3059-3066. 10.1093/nar/gkf436
171. Nguyen, L.T. *et al.* (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32, 268-274. 10.1093/molbev/msu300
172. Sagulenko, P. *et al.* (2018) TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol* 4, vex042. 10.1093/ve/vex042
173. Huddleston, J. *et al.* (2021) Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens. *J Open Source Softw* 6. 10.21105/joss.02906
174. Munoz, E.T. and Deem, M.W. (2005) Epitope analysis for influenza vaccine design. *Vaccine* 23, 1144-1148. 10.1016/j.vaccine.2004.08.028
175. Shannon, C.E. (1948) A mathematical theory of communication. *The Bell system technical journal* 27, 379-423
176. Hothorn, T. *et al.* (2006) Survival ensembles. *Biostatistics* 7, 355-373. 10.1093/biostatistics/kxj011
177. Strobl, C. *et al.* (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 8, 25. 10.1186/1471-2105-8-25
178. Strobl, C. *et al.* (2008) Conditional variable importance for random forests. *BMC Bioinformatics* 9, 307. 10.1186/1471-2105-9-307
179. Kuhn, M. (2008) Building Predictive Models in R Using the caret Package. *Journal of Statistical Software* 28, 1 - 26. 10.18637/jss.v028.i05
180. Altmann, A. *et al.* (2010) Permutation importance: a corrected feature importance measure. *Bioinformatics* 26, 1340-1347. 10.1093/bioinformatics/btq134
181. Debeer, D. and Strobl, C. (2020) Conditional permutation importance revisited. *BMC Bioinformatics* 21, 307. 10.1186/s12859-020-03622-2
182. Friedman, J. *et al.* (2010) Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* 33, 1-22. 10.18637/jss.v033.i01
183. Sax, C. and Steiner, P. (2013) Temporal Disaggregation of Time Series. *R J.* 5, 80
184. Biggerstaff, M. *et al.* (2014) Influenza-like illness, the time to seek healthcare, and influenza antiviral receipt during the 2010-2011 influenza season-United States. *J Infect Dis* 210, 535-544. 10.1093/infdis/jiu224
185. Lessler, J. *et al.* (2009) Incubation periods of acute respiratory viral infections: a systematic review. *Lancet Infect Dis* 9, 291-300. 10.1016/s1473-3099(09)70069-6
186. Russell, K.E. *et al.* (2018) Comparison of outpatient medically attended and community-level influenza-like illness-New York City, 2013-2015. *Influenza Other Respir Viruses* 12, 336-343. 10.1111/irv.12540
187. Cowling, B.J. *et al.* (2009) Estimation of the serial interval of influenza. *Epidemiology* 20, 344-347. 10.1097/EDE.0b013e31819d1092
188. Hoffman, M.D. and Gelman, A. (2014) The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* 15, 1593-1623
189. Grenfell, B.T. *et al.* (2001) Travelling waves and spatial hierarchies in measles epidemics. *Nature* 414, 716-723. 10.1038/414716a
190. Liebhold, A. *et al.* (2004) Spatial Synchrony in Population Dynamics. *Annual Review of Ecology, Evolution, and Systematics* 35, 467-490. 10.1146/annurev.ecolsys.34.011802.132516

191. Weinberger, D.M. *et al.* (2012) Influenza epidemics in Iceland over 9 decades: changes in timing and synchrony with the United States and Europe. *Am J Epidemiol* 176, 649-655. 10.1093/aje/kws140

192. Torrence, C. and Compo, G.P. (1998) A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society* 79, 61-78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

Figures

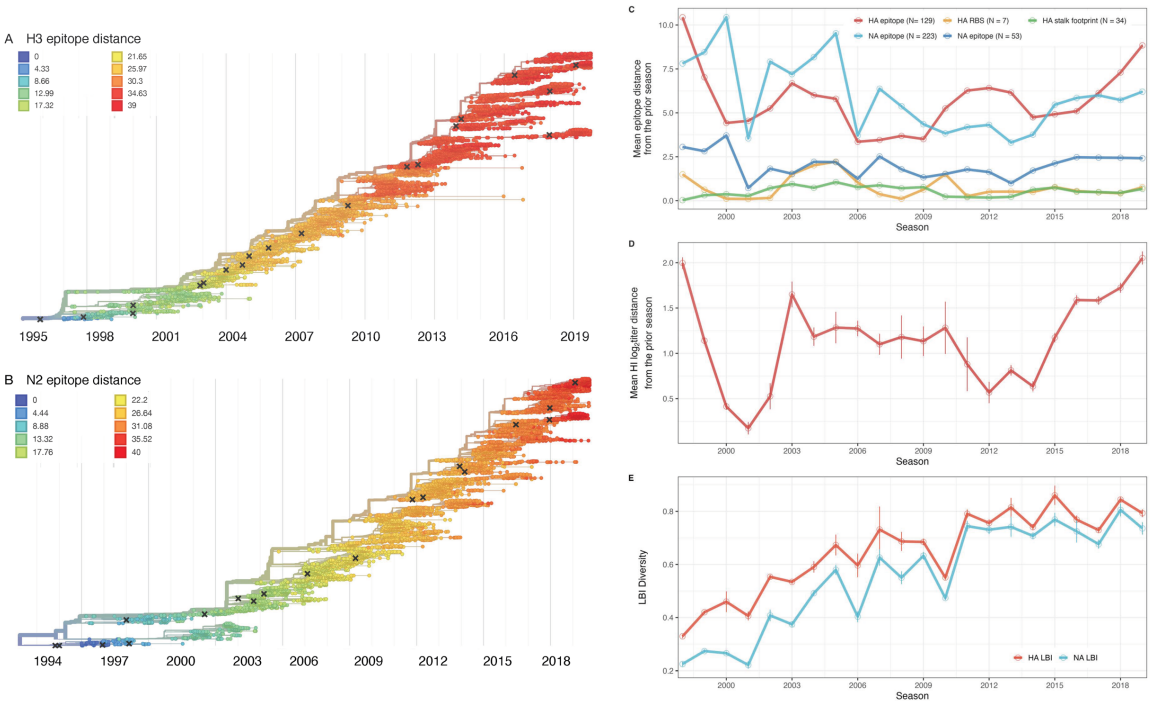


Figure 1. Antigenic and genetic evolution of seasonal influenza A(H3N2) viruses, 1997 - 2019. A-B. Temporal phylogenies of hemagglutinin (H3) and neuraminidase (N2) gene segments. Tip color denotes the Hamming distance from the root of the tree, based on the number of substitutions at epitope sites in H3 (N = 129 sites) and N2 (N = 223 sites). "X" marks indicate the phylogenetic positions of US recommended vaccine strains. **C-D.** Seasonal genetic and antigenic distances are the mean distance between A(H3N2) viruses circulating in the current season t versus the prior season (t – 1), measured by **C.** four sequence-based metrics (HA receptor binding site (RBS), HA stalk footprint, HA epitope, and NA epitope) and **D.** hemagglutination inhibition (HI) titer measurements. **E.** The Shannon entropy of H3 and N2 local branching index (LBI) values in each season. Vertical bars in **C, D,** and **E** and are 95% confidence intervals of seasonal estimates from five bootstrapped phylogenies.

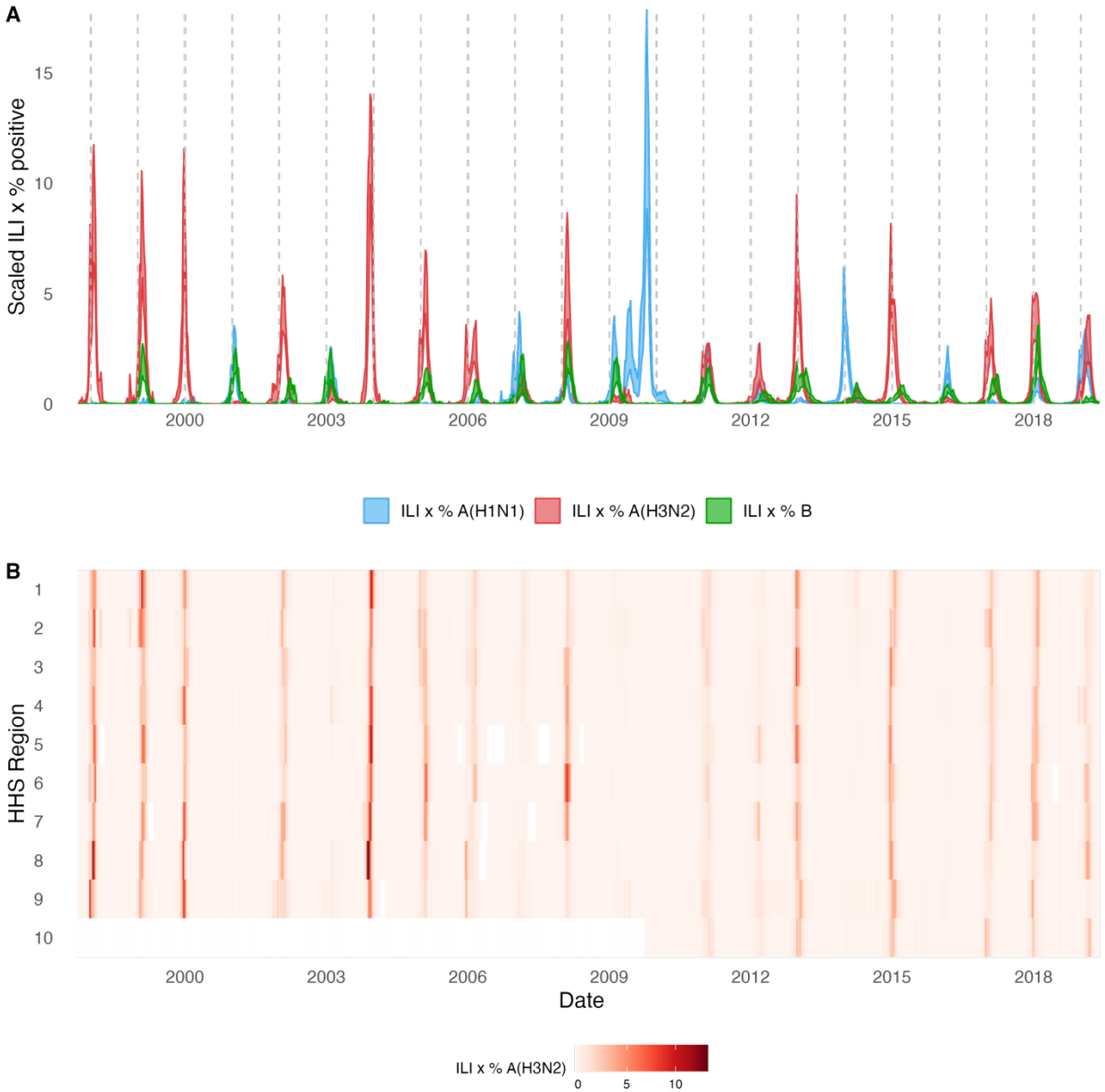


Figure 2. Annual influenza A(H3N2) epidemics in the United States, 1997 - 2019. A. Weekly incidence of influenza A(H3N2) (red), A(H1N1) (blue), and B (green) averaged across ten HHS regions (Region 1: Boston; Region 2: New York City; Region 3: Washington, DC; Region 4: Atlanta; Region 5: Chicago; Region 6: Dallas, Region 7: Kansas City; Region 8: Denver; Region 9: San Francisco; Region 10: Seattle). Time series are 95% confidence intervals of regional incidence estimates. Incidences are the proportion of influenza-like illness (ILI) visits among all outpatient visits, multiplied by the proportion of respiratory samples testing positive for each influenza type/subtype. Vertical dashed lines indicate January 1 of each year. **B.** Intensity of weekly influenza A(H3N2) incidence in ten HHS regions. White tiles indicate weeks when influenza-like-illness data or virological data were not reported. Weekly time series for A(H1N1) and B are in Figure S5.

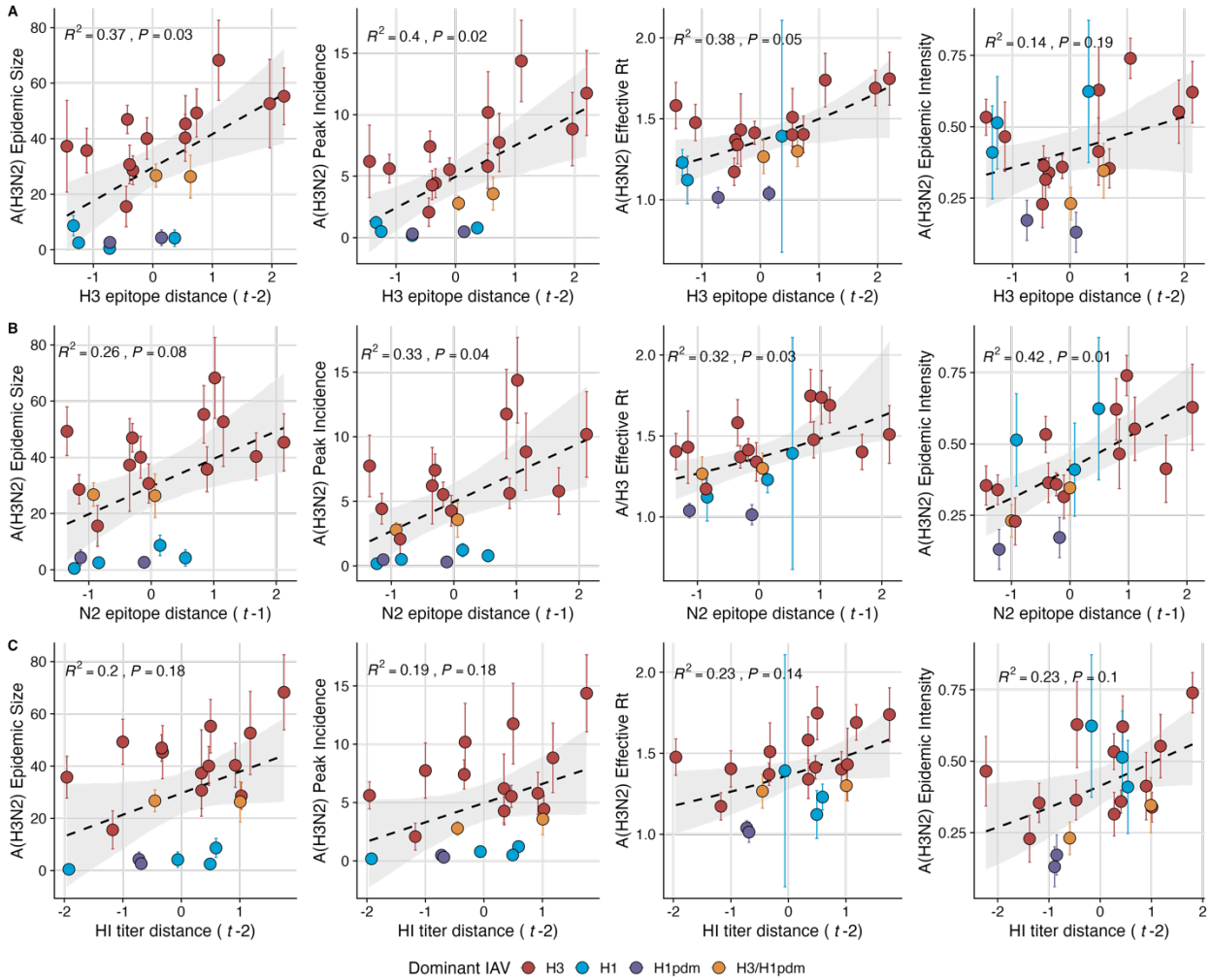


Figure 3. A(H3N2) antigenic drift correlates with larger, more intense annual epidemics. A(H3N2) epidemic size, peak incidence, epidemic intensity, and transmissibility (effective reproduction number, R_t) increase with antigenic drift, measured by **A.** hemagglutinin (H3) epitope distance, and **B.** neuraminidase (N2) epitope distance, and **C.** hemagglutination inhibition (HI) $\log_2$ titer distance. Seasonal antigenic drift is the mean titer distance or epitope distance between viruses circulating in the current season t versus the prior season (t – 1) or two prior seasons (t – 2). Distances are scaled to aid in direct comparison of evolutionary indicators. Point color indicates the dominant influenza A virus (IAV) subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant), and vertical bands are 95% confidence intervals of regional estimates. Seasonal mean A(H3N2) epidemic metric values were fit as a function of antigenic or genetic distance using LMs (epidemic size, peak incidence), Gaussian GLMs (effective Rt: inverse link), or Beta GLMs (epidemic intensity) with 1000 bootstrap resamples.

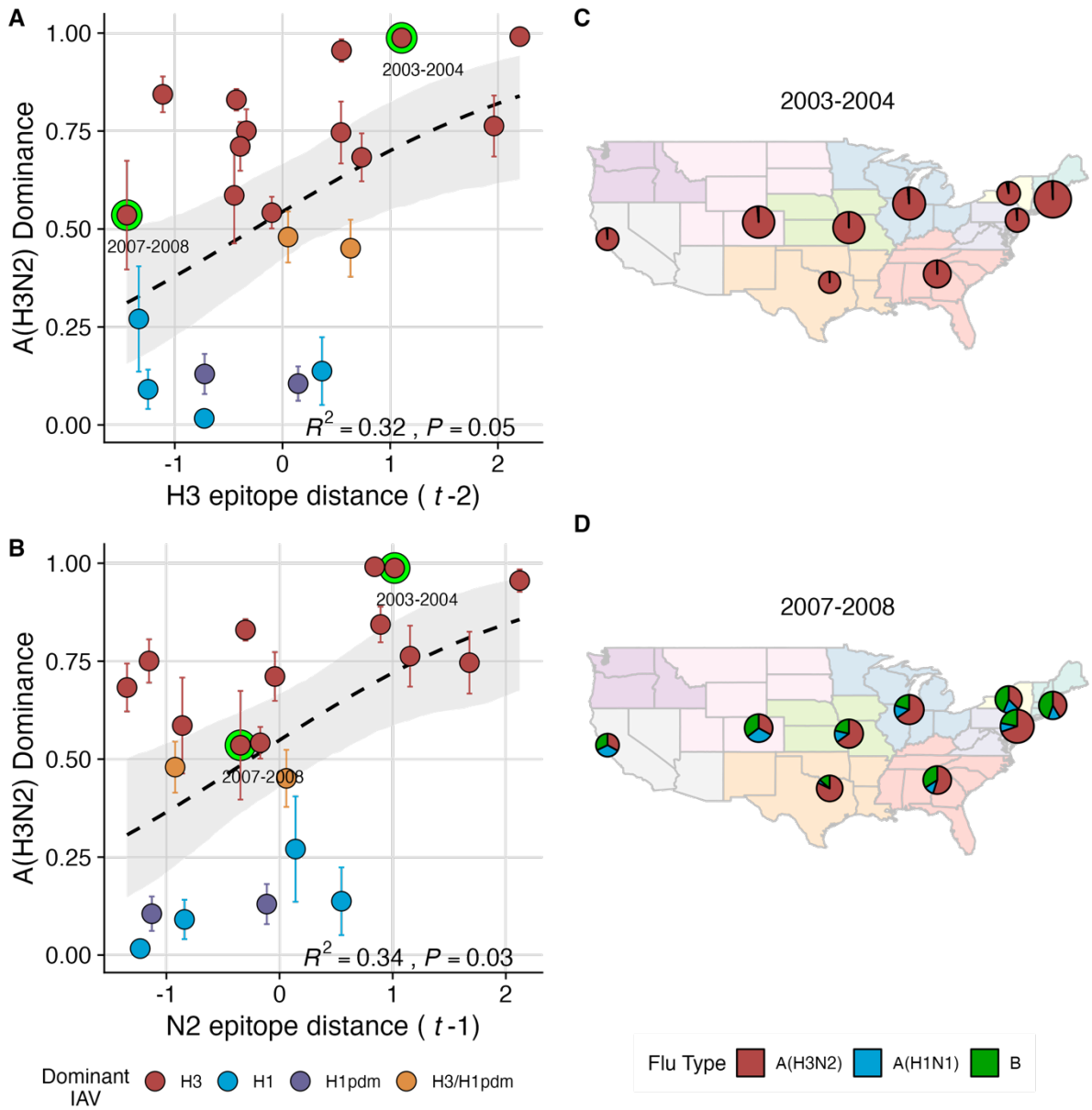


Figure 4. The proportion of influenza positive samples typed as A(H3N2) increases with antigenic drift. A-B. Seasonal A(H3N2) subtype dominance increases with H3 and N2 epitope distance. Seasonal epitope distance is the mean epitope distance between viruses circulating in the current season *t* versus the prior season ($t-1$) or two prior seasons ($t-2$). Distances were scaled to aid in direct comparison of evolutionary indicators. Point color indicates the dominant influenza A virus (IAV) subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant), and vertical bands are 95% confidence intervals of regional estimates. Seasonal mean A(H3N2) dominance was fit as a function of H3 or N2 epitope distance using Beta GLMs with 1000 bootstrap resamples. **C-D.** Regional patterns of influenza type and subtype incidence during two seasons when A(H3N2) was nationally dominant. **C.** Widespread A(H3N2) dominance during 2003-2004 after the emergence of a novel antigenic cluster, FU02 (A/Fujian/411/2002-like strains). **D.** Spatial heterogeneity in subtype circulation during 2007-2008, a season with low A(H3N2) antigenic novelty relative to the prior season. Pie charts represent the proportion of influenza positive samples typed as A(H3N2) (red), A(H1N1) (blue), or B (green) in each HHS region. Data for Region 10 (purple) were not available for seasons prior to 2009. The sizes of regional pie charts are proportional to the total number of influenza positive samples.

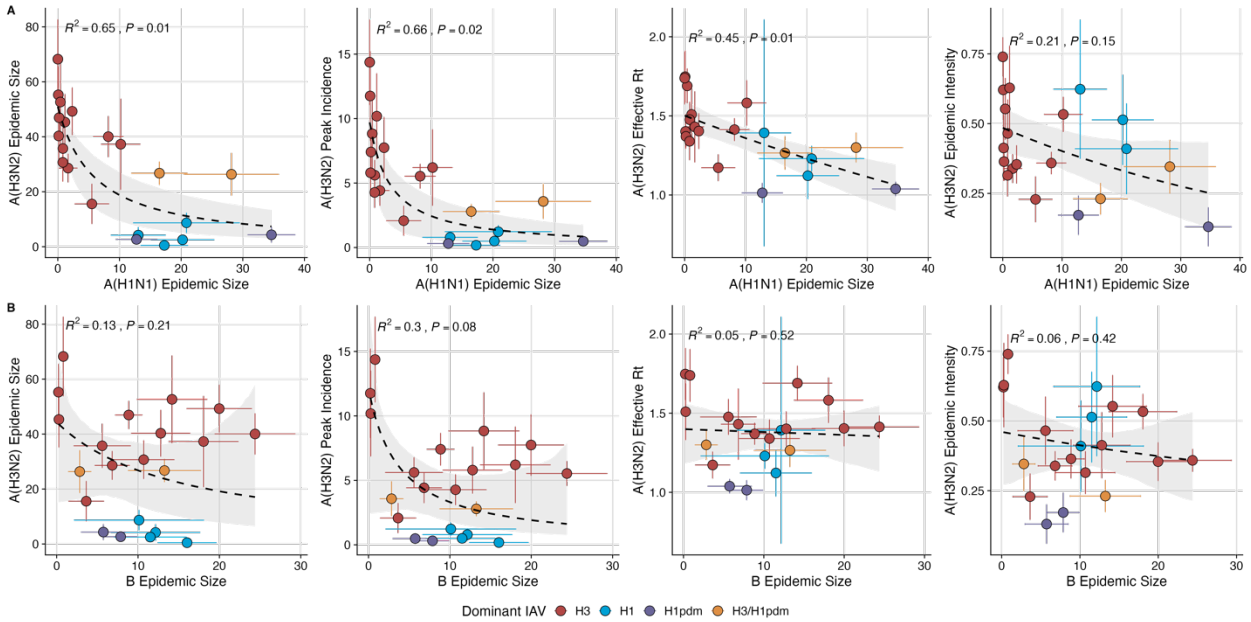


Figure 5. The effects of influenza A(H1N1) and B epidemic size on A(H3N2) epidemic burden. A. Influenza A(H1N1) epidemic size negatively correlates with A(H3N2) epidemic size, peak incidence, transmissibility (effective reproduction number, R_t), and epidemic intensity. **B.** Influenza B epidemic size does not significantly correlate with A(H3N2) epidemic metrics. Point color indicates the dominant influenza A virus (IAV) subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant), and vertical and horizontal bands are 95% confidence intervals of regional estimates. Seasonal mean A(H3N2) epidemic metrics were fit as a function of mean A(H1N1) or B epidemic size using Gaussian GLMs (inverse link: epidemic size, peak incidence; log link: effective Rt) or Beta GLMs (epidemic intensity) with 1000 bootstrap resamples.

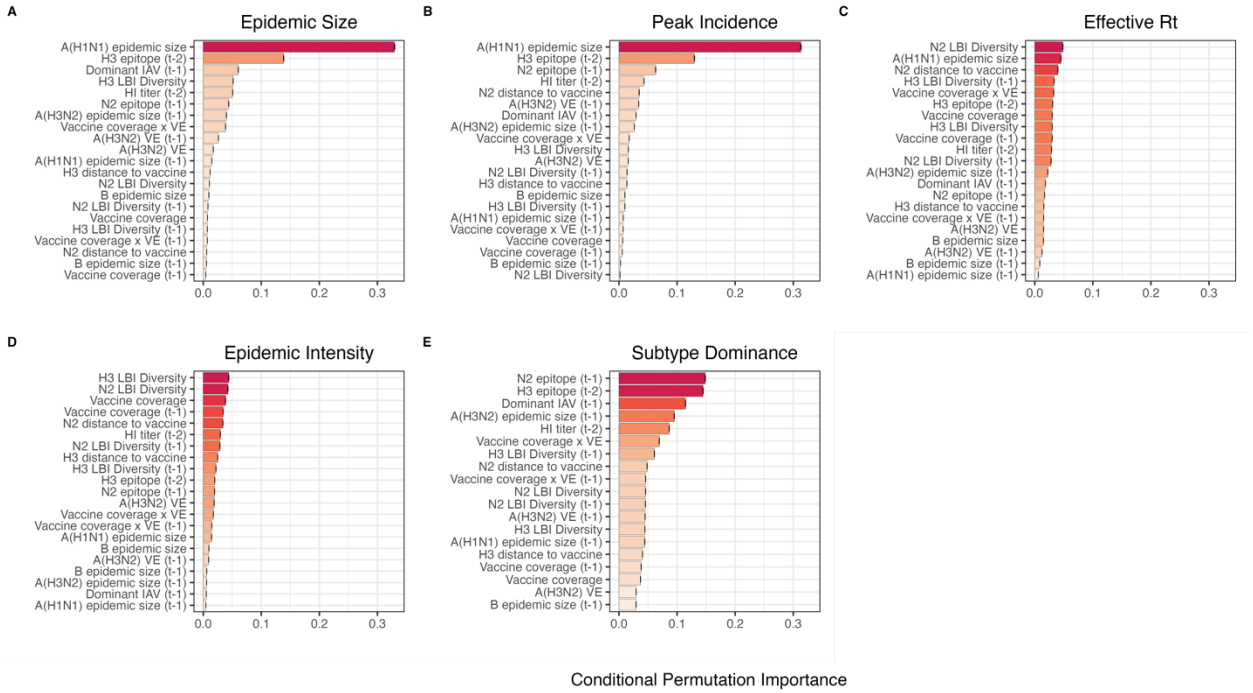


Figure 6. Variable importance rankings from conditional inference random forest models predicting A(H3N2) epidemic dynamics. Ranking of variables in predicting regional A(H3N2) **A.** epidemic size, **B.** peak incidence, **C.** effective reproduction number, Rt, **D.** epidemic intensity, and **E.** subtype dominance. Each forest was created by generating 3,000 regression trees from a repeated leave-one-season-out cross-validated sample of the data. Variables are ranked by their conditional permutation importance, with differences in prediction accuracy scaled by the total (null model) error. Black error bars are 95% confidence intervals of conditional permutation scores. Abbreviations: HI titer = hemagglutination inhibition $\log_2$ titer distance, t - 1 = one-season lag, t - 2 = two-season lag, LBI = local branching index, peak = peak incidence, distance to vaccine = epitope distance between currently circulating strains and the recommended vaccine strain, VE = vaccine effectiveness.

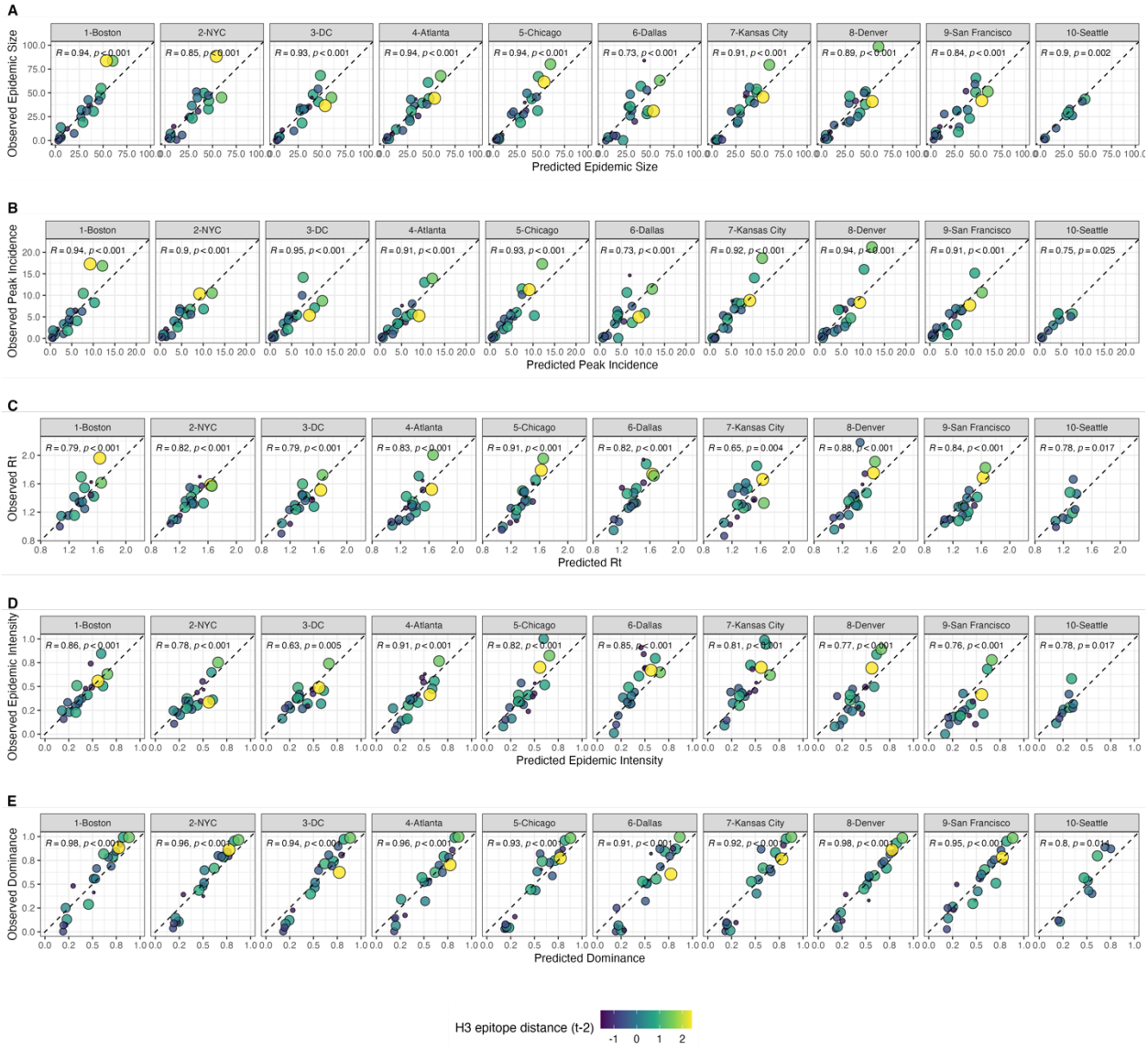


Figure 7. Observed versus predicted values of seasonal region-specific A(H3N2) A. epidemic size, B. peak incidence, C. effective reproduction number, Rt, D. epidemic intensity, and E. subtype dominance from conditional random forest models. Results are facetted by HHS region and epidemic metric. Point color and size corresponds to the degree of hemagglutinin (H3) epitope distance in viruses circulating in season *t* versus viruses circulating two seasons ago ($t-2$). Large, yellow points indicate seasons with high antigenic novelty, and small blue points indicate seasons with low antigenic novelty. Regional Spearman's correlation coefficients and associated P-values are in the top left section of each facet.

Tables

Table 1. Evolutionary indicators of seasonal viral fitness. Evolutionary indicators are labeled by the influenza gene for which data are available (hemagglutinin, HA or neuraminidase, NA), the type of data they are based on, and the component of influenza fitness they represent. Table format is adapted from Huddleston et al., 2020 [35].

Evolutionary indicator	Influenza gene	Data type	Fitness category	Citations
Mean HI titer $\log_2$ distance from the prior season	HA	Hemagglutinin inhibition assays using ferret sera	Antigenic drift	Huddleston et al., 2020; Neher et al., 2016
Mean epitope distance from the prior season	HA and NA	Sequences	Antigenic drift	Bhatt et al., 2011; Bush et al., 1999; Krammer, unpublished; Webster and Laver, 1980; Wiley et al., 1981; Wilson and Cox, 1990; Wolf et al., 2010
Mean receptor binding site distance from the prior season	HA	Sequences	Antigenic drift	Koel et al., 2013
Mutational load (mean non-epitope distance from the prior season)	HA and NA	Sequences	Functional constraint	Łuksza and Lässig, 2014
Mean stalk “footprint” distance from the prior season	HA	Sequences	Negative control	Kirkpatrick et al., 2018
Mean local branching index	HA and NA	Sequences	Clade growth	Huddleston et al., 2020; Łuksza and Lässig, 2014
Shannon entropy of local branching index	HA and NA	Sequences	Diversity of clade growth rates	Huddleston et al., 2020; Neher et al., 2014

1527 **Table 2. Seasonal metrics of A(H3N2) epidemic dynamics.** Epidemic metrics are defined and labeled
1528 by which outcome category they represent.
1529

Epidemic Outcome	Definition	Outcome category	Citations
Epidemic size	Cumulative weekly incidence	Burden	
Peak incidence	Maximum weekly incidence	Burden	
Maximum time-varying effective reproduction number, Rt	The number of secondary cases arising from a symptomatic index case, assuming conditions remain the same	Transmissibility	Cori et al., 2013; Scott et al., 2021
Epidemic intensity	Inverse Shannon entropy of the weekly incidence distribution (i.e., the spread of incidence across the season)	Sharpness of the epidemic curve	Dalziel et al., 2018
Subtype dominance	The proportion of influenza positive samples typed as A(H3N2)	Viral activity	
Excess pneumonia and influenza mortality attributable to A(H3N2) virus	Mortality burden in excess of a seasonally adjusted baseline	Severity	Hansen et al., 2022; Simonsen and Viboud, 2012
Onset week	Winter changepoint in incidence	Timing	Charu et al., 2017
Peak week	First week of maximum incidence	Timing	
Spatiotemporal synchrony	Variation (s.d.) in regional onset or peak timing	Speed	Viboud et al., 2006
Onset to peak	Number of days between onset week and peak week	Speed	
Seasonal duration	Number of weeks with non-zero incidence	Speed	

1530

1531 **Table 3. Predictors of seasonal A(H3N2) epidemic burden, transmissibility, intensity, and subtype**
1532 **dominance.** Variables retained in the best fit model for each epidemic outcome were determined by BIC.
1533

Outcome	Best Minimal Model[1]	R^2	Adj. R^2	RMSE
Epidemic Size	H3 epitope distance (t-2) + H1 epidemic size + H3 epidemic size (t-1)	0.74	0.69	9.88
Peak Incidence	H3 epitope distance (t-2) + H1 epidemic size + Dominant IAV Subtype (t-1)	0.69	0.63	2.09
Effective Rt	HI titer distance (t-2) + H1 epidemic size + H3 LBI Diversity (t-1)	0.71	0.65	0.1
Epidemic Intensity	HI titer distance (t-2) + N2 distance to vaccine strain + vaccination coverage (t-1)	0.79	0.75	0.07
Subtype Dominance	H3 epitope distance (t-2) + N2 epitope distance (t-1) + Dominant IAV Subtype (t-1)	0.56	0.48	0.2

1534

1535 [1]Candidate models were limited to 3 independent variables and considered all combinations of the top 10
1536 ranked predictors from conditional inference random forest models (Figure 6).

Supplementary Figures

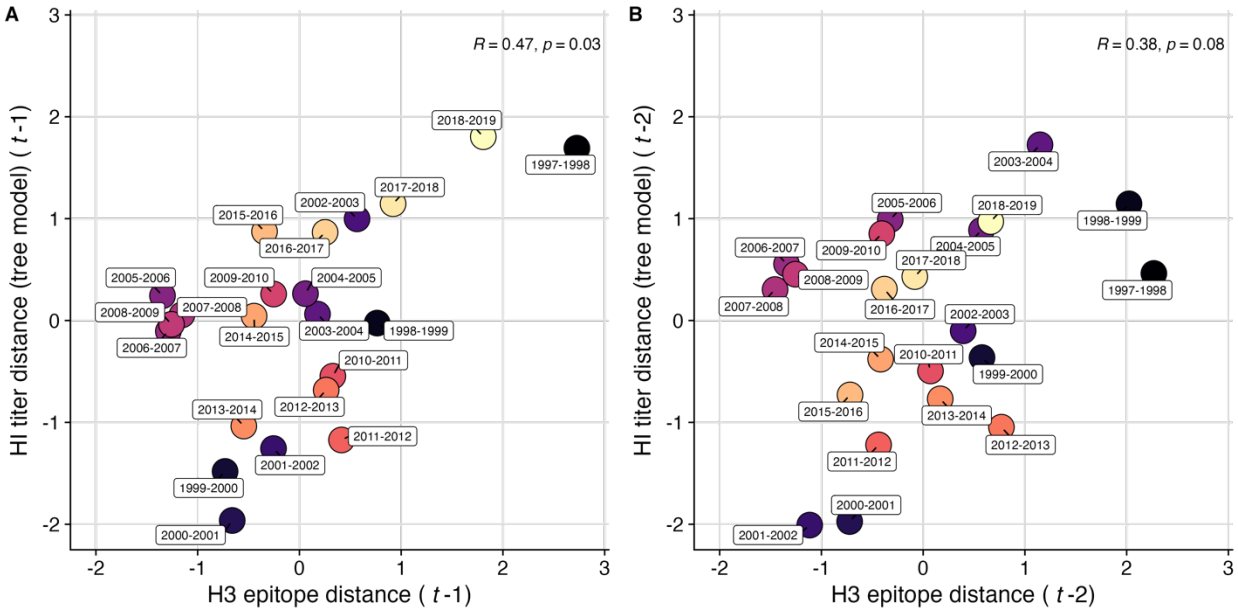


Figure S1. Comparison of seasonal antigenic drift measured by substitutions at hemagglutinin (H3) epitope sites and HI titer measurements, from 1997-1998 to 2018-2019. We used Spearman correlation tests to measure associations between H3 epitope distance and HI titer distance at **A.** one-season lags and **B.** two-season lags. Seasonal antigenic distance is the mean distance between strains circulating in season *t* and strains circulating in the prior season *t* – 1 year (one season lag) or two seasons ago *t* – 2 years (two season lag). Seasonal distances are scaled because epitope distance and HI titer distance use different units of measurement. Point labels indicate the current influenza season, and point color denotes the relative timing of influenza seasons, with earlier seasons shaded dark purple (e.g., 1997-1998) and later seasons shaded light yellow (e.g., 2018-2019). H3 epitope distance and HI titer (tree model) distance at two-season lags capture expected "jumps" in antigenic drift during key seasons previously associated with major antigenic transitions [32], such as the SY97 cluster seasons (1997-1998, 1998-1999, 1999-2000) and the FU02 cluster season (2003-2004).

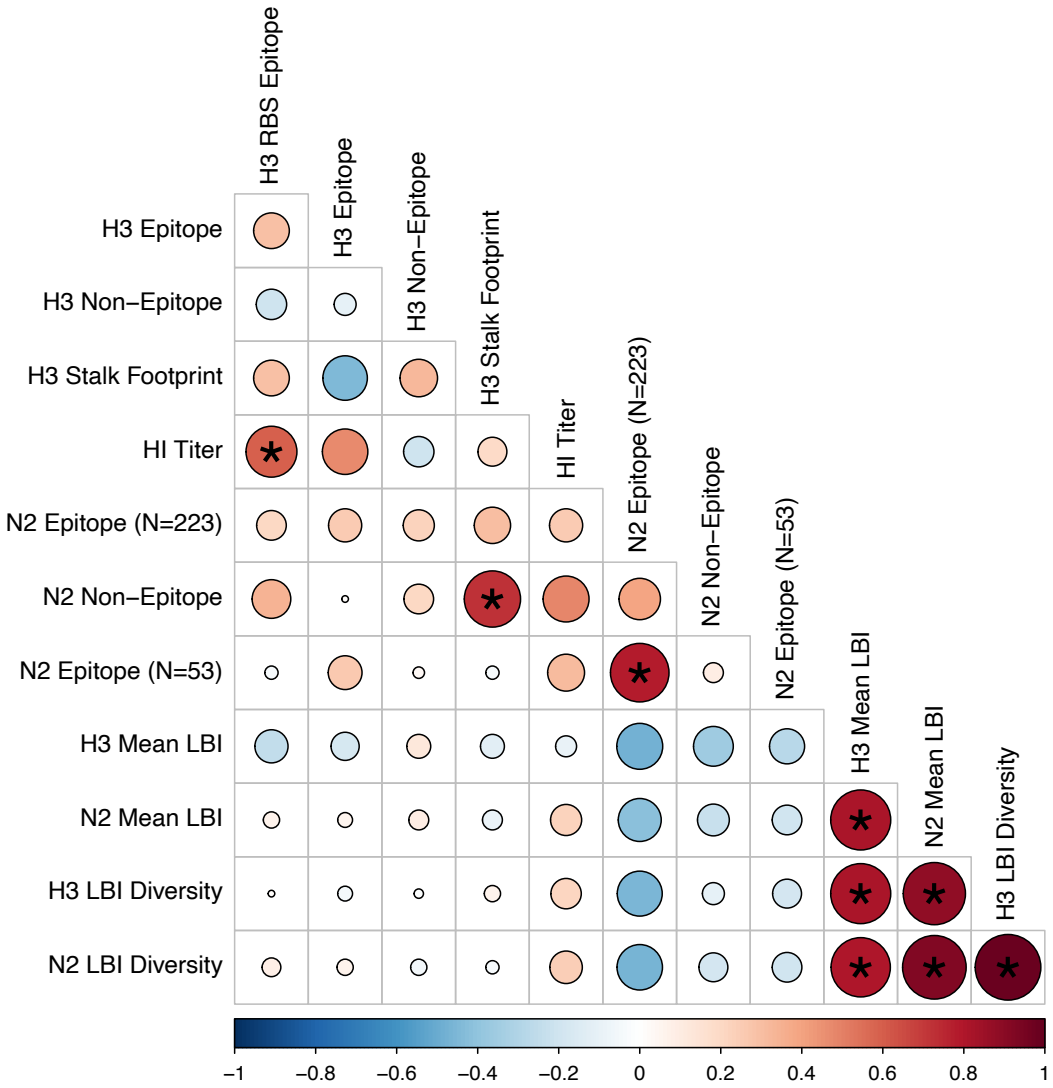


Figure S2. Pairwise correlations between H3 and N2 evolutionary indicators (one season lags). We measured Spearman's correlations between seasonal measures of H3 and N2 evolution, including H3 RBS distance, H3 epitope distance, H3 non-epitope distance, H3 stalk footprint distance, HI titer distance, N2 epitope distance based on 223 or 53 epitope sites, N2 non-epitope distance, mean clade growth of H3 and N2 (local branching index, LBI), and the Shannon entropy of H3 and N2 LBI values. Seasonal distances were estimated as the mean distance between strains circulating in the current season *t* and those circulating in the prior season ($t-1$). The Benjamini and Hochberg method was used to adjust P-values for multiple testing. The color of each circle indicates the strength and direction of the association, from dark red (strong positive correlation) to dark blue (strong negative correlation). Stars within circles indicate statistical significance (adjusted $P < 0.05$).

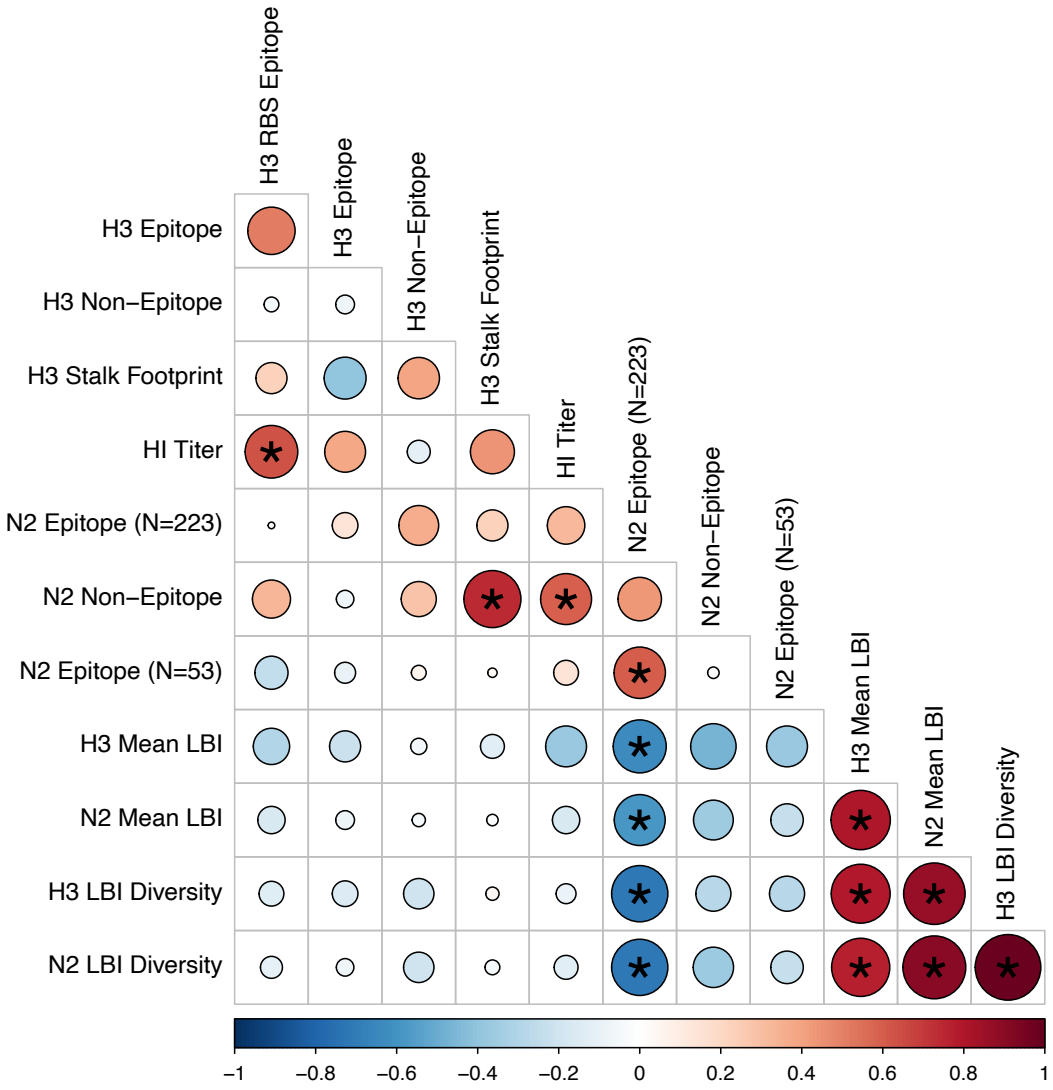


Figure S3. Pairwise correlations between H3 and N2 evolutionary indicators (two season lags). We measured Spearman's correlations between seasonal measures of H3 and N2 evolution, including H3 RBS distance, H3 epitope distance, H3 non-epitope distance, H3 stalk footprint distance, HI titer distance (tree model), N2 epitope distance based on 223 or 53 epitope sites, N2 non-epitope distance, mean clade growth of H3 and N2 (local branching index, LBI), and the Shannon entropy of H3 and N2 LBI values. Seasonal distances were estimated as the mean distance between strains circulating in the current season t and those circulating in the prior season ($t-1$). The Benjamini and Hochberg method was used to adjust P-values for multiple testing. The color of each circle indicates the strength and direction of the association, from dark red (strong positive correlation) to dark blue (strong negative correlation). Stars within circles indicate statistical significance (adjusted $P < 0.05$).

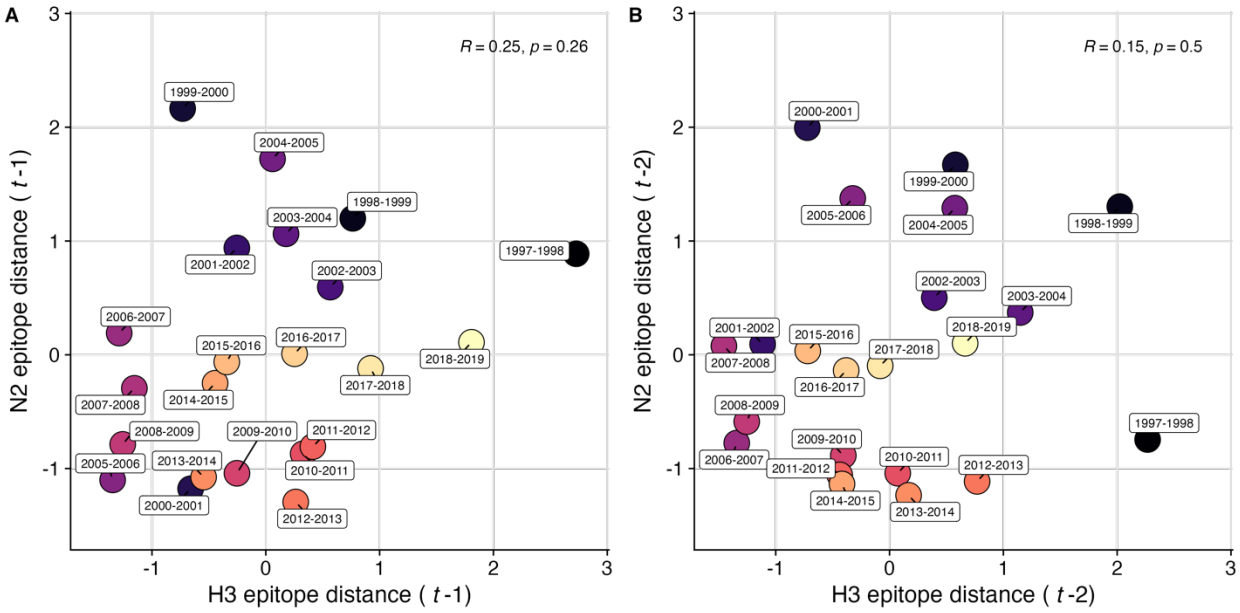


Figure S4. Comparison of seasonal antigenic drift measured by substitutions at hemagglutinin (H3) and neuraminidase (N2) epitope sites, from 1997-1998 to 2018-2019. We used Spearman correlation tests to measure associations between H3 epitope distance and N2 epitope distance at **A.** one-season lags and **B.** two-season lags. Seasonal epitope distance is the mean distance between strains circulating in season *t* and strains circulating in the prior season $t-1$ (one season lag) or two seasons ago $t-2$ (two season lag). Point labels indicate the current influenza season, and point color denotes the relative timing of influenza seasons, with earlier seasons shaded dark purple (e.g., 1997-1998) and later seasons shaded light yellow (e.g., 2018-2019). N2 epitope distance at one-season lags captures expected "jumps" in antigenic drift during key seasons previously associated with major antigenic transitions [32], such as the SY97 cluster seasons (1997-1998, 1998-1999, 1999-2000) the FU02 cluster season (2003-2004), and the CA04 cluster season (2004-2005).

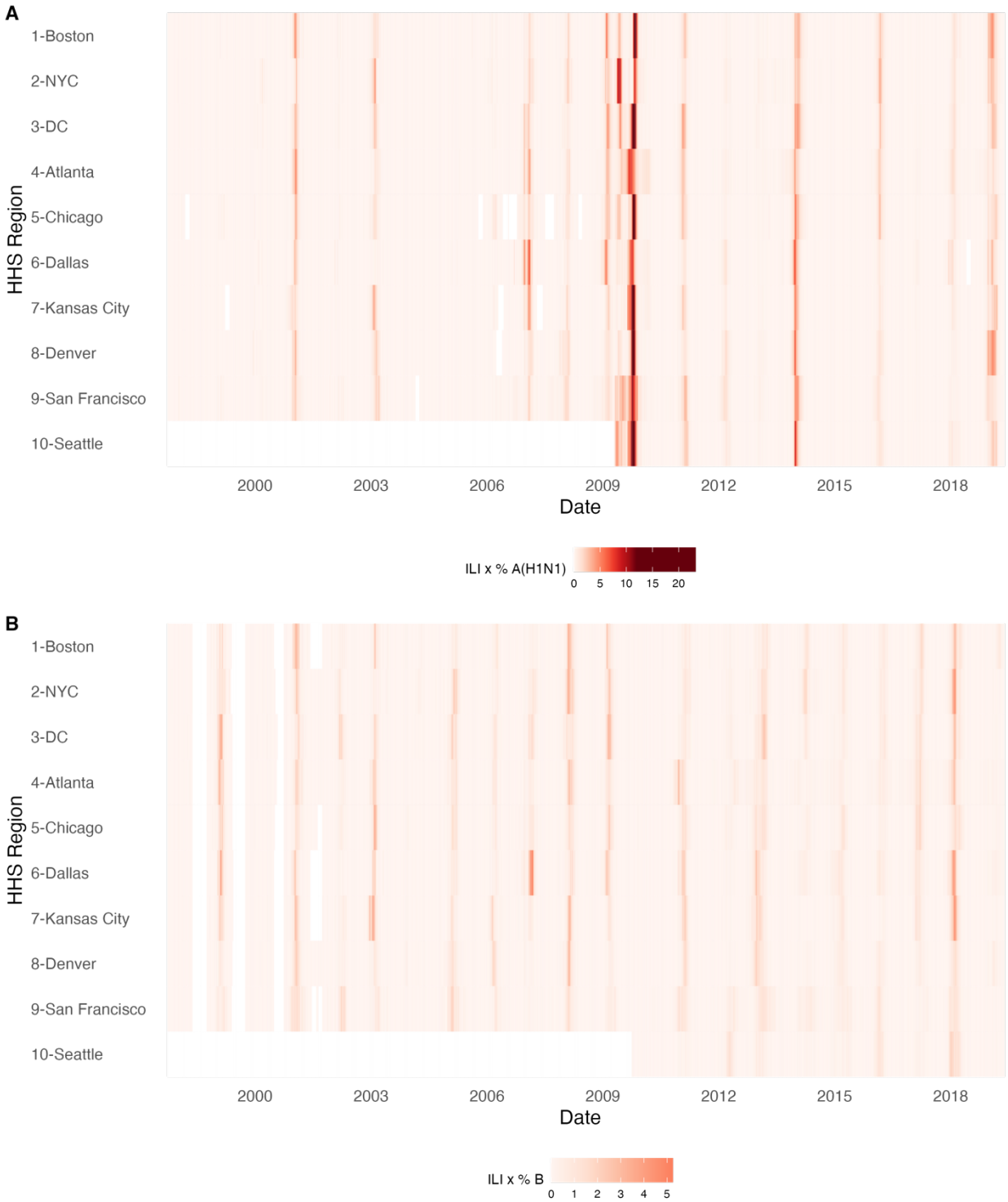


Figure S5. Intensity of weekly incidence of A. influenza A(H1N1) and B. influenza B in ten HHS regions, 1997 - 2019. Seasonal and pandemic A(H1N1) were combined as A(H1N1), and the Victoria and Yamagata lineages of influenza B were combined as influenza B. White tiles indicate weeks when either influenza-like-illness cases or virological data were not reported. Data for Region 10 were not available in seasons prior to 2009.

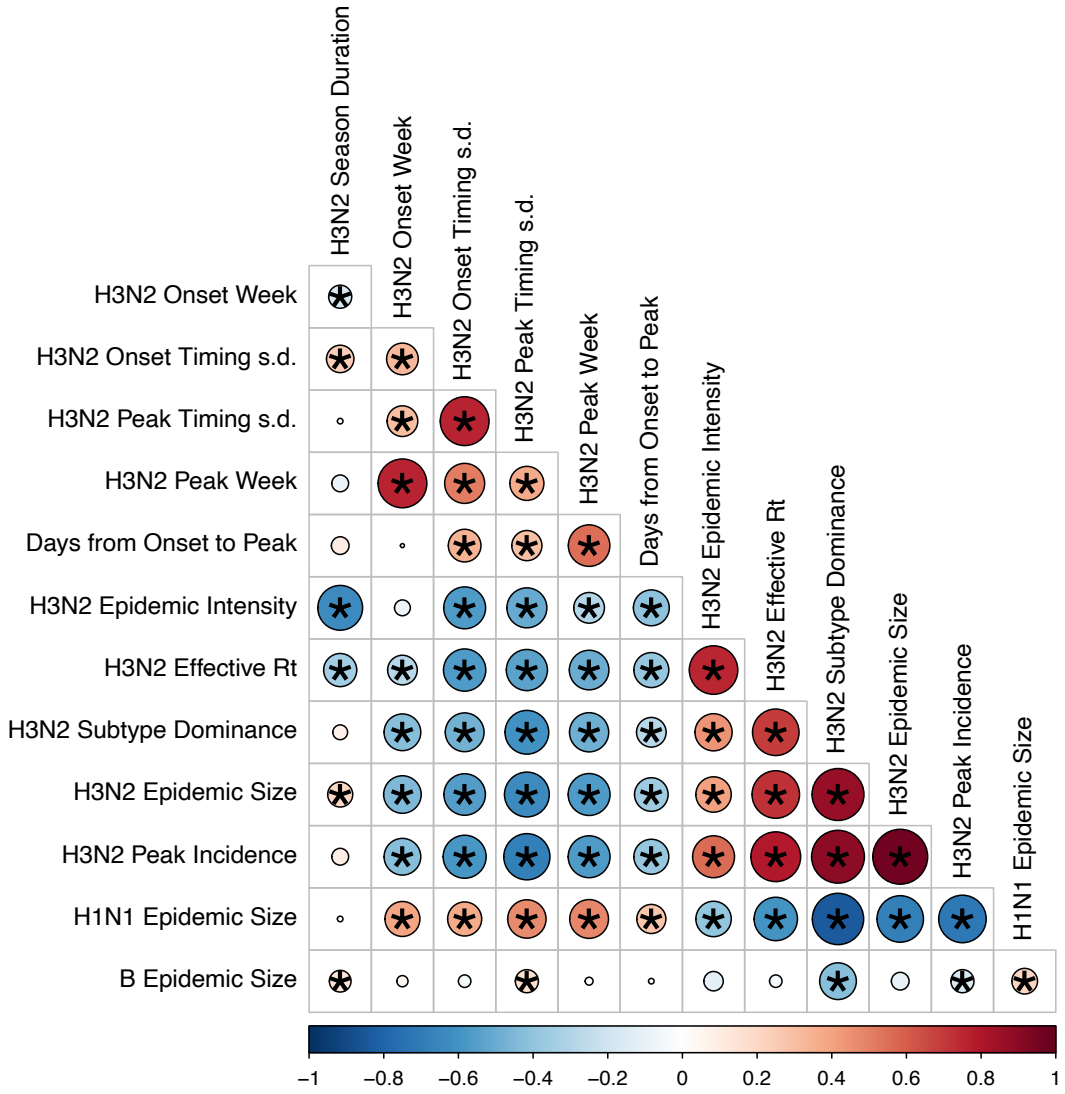


Figure S6. Pairwise correlations between seasonal A(H3N2), A(H1N1), and B epidemic metrics. We measured Spearman's correlations among indicators of A(H3N2) epidemic timing, including onset week, peak week, regional variation (s.d.) in onset and peak timing, and the number of days from onset to peak, indicators of A(H3N2) epidemic magnitude, including epidemic intensity (i.e., the "sharpness" of the epidemic curve), transmissibility (maximum effective reproduction number, Rt), subtype dominance patterns, epidemic size, and peak incidence. We also considered relationships between the circulation of other types/subtypes and A(H3N2) epidemic burden and timing. The Benjamini and Hochberg method was used to adjust P-values for multiple testing. The color of each circle indicates the strength and direction of the association, from dark red (strong positive correlation) to dark blue (strong negative correlation). Stars within circles indicate statistical significance (adjusted $P < 0.05$).

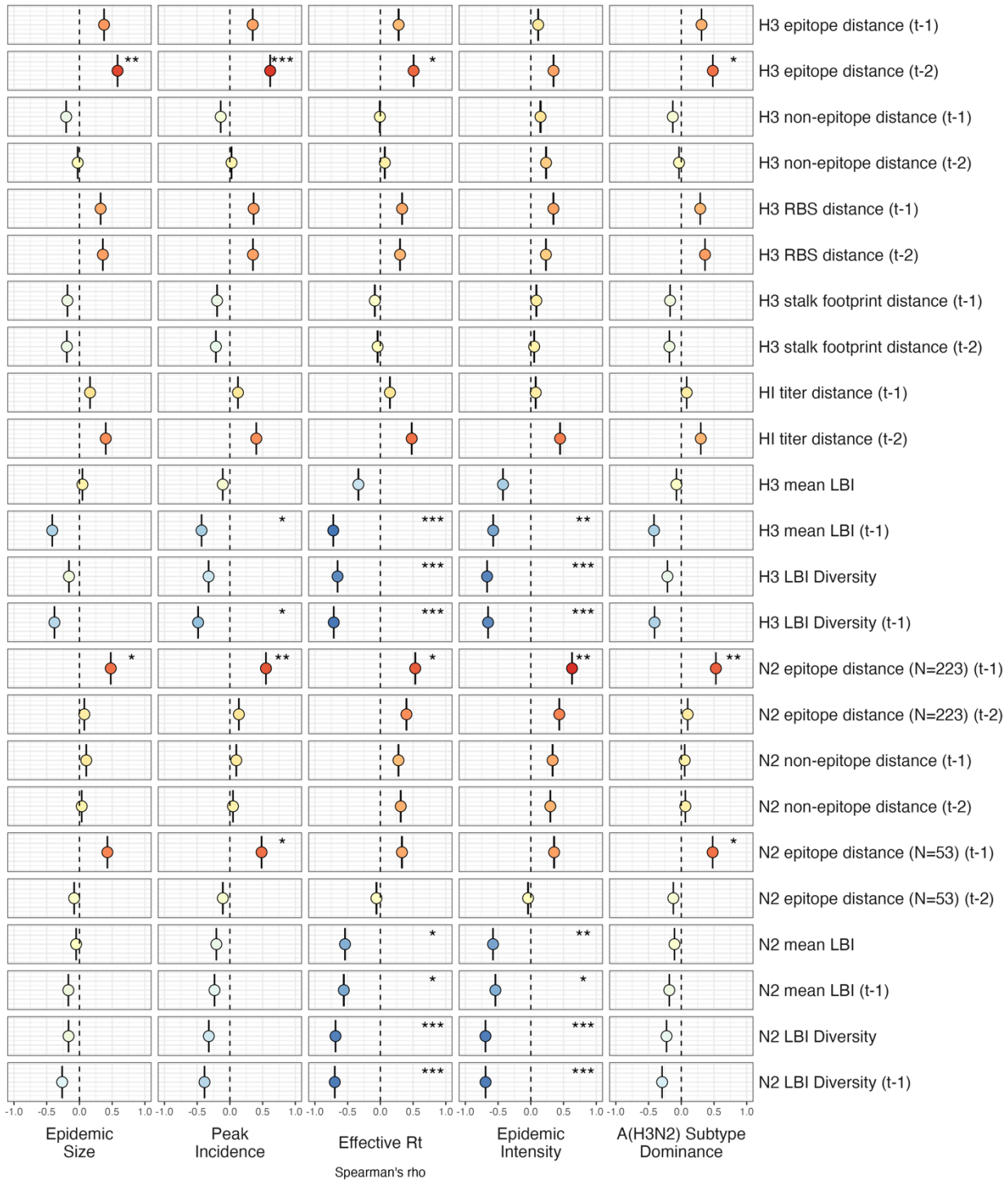


Figure S7. Univariate correlations between A(H3N2) viral fitness and epidemic impact. Mean Spearman correlation coefficients, 95% confidence intervals of correlation coefficients, and corresponding p-values of bootstrapped (N = 1000) viral fitness indicators (rows) and epidemic metrics (columns). Point color indicates the strength and direction of the association, from dark red (strong positive correlation) to dark blue (strong negative correlation), and stars indicate statistical significance (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). Abbreviations: HI = hemagglutination inhibition, RBS: receptor binding site, t - 1 = one-season lag, t - 2 = two-season lag, LBI = local branching index.

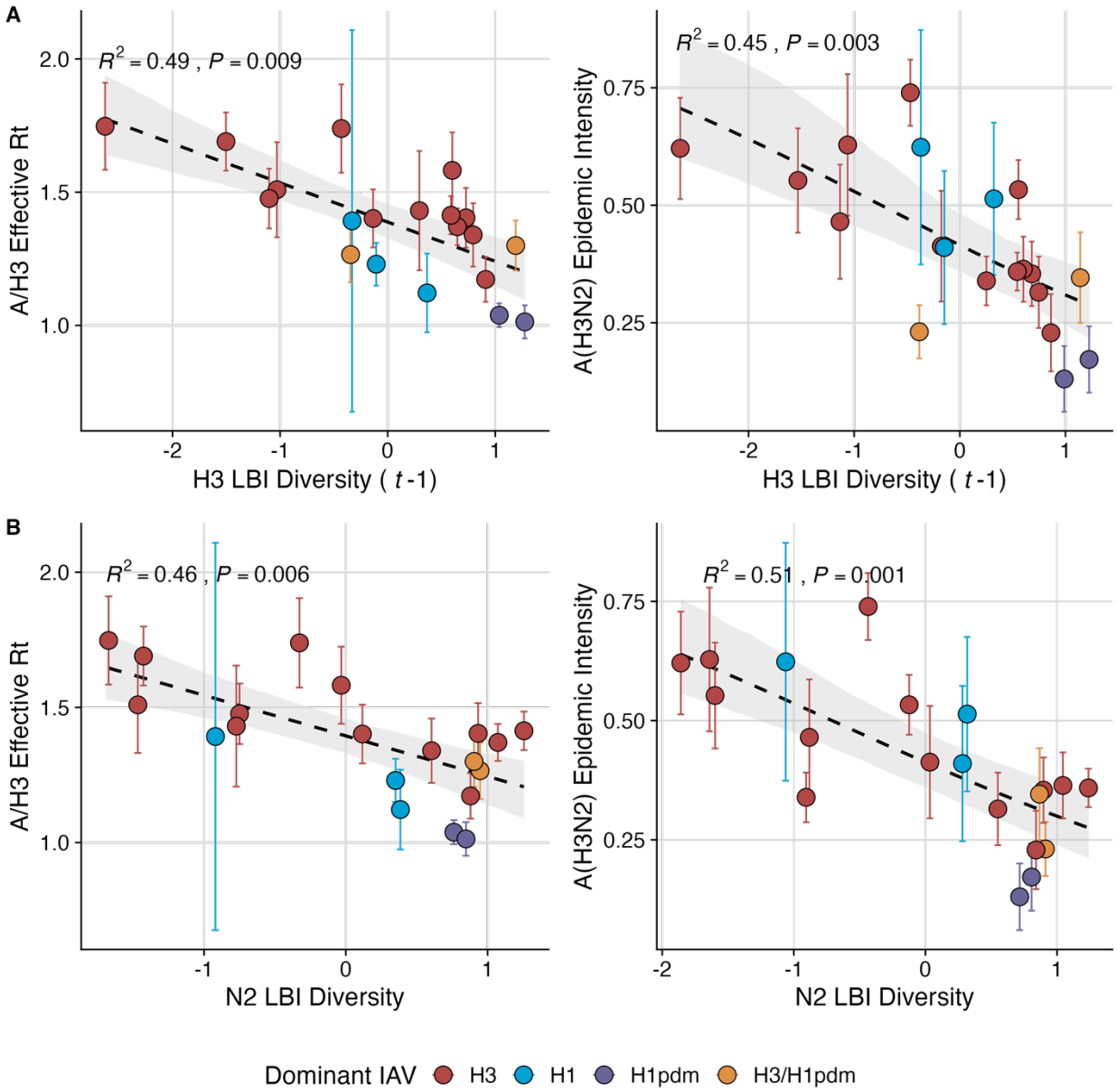


Figure S8. Low diversity in the growth rates of circulating A(H3N2) clades is associated with more intense epidemics and higher transmissibility. A(H3N2) effective Rt and epidemic intensity negatively correlate with the diversity of LBI values among circulating A(H3N2) lineages in the current or prior season, measured by the Shannon entropy of **A.** H3 local branching index (LBI) values in the prior season (t – 1), and **B.** the Shannon entropy of N2 LBI values in the current season t. LBI values are scaled to aid in direct comparisons of H3 and N2 LBI diversity. Point color indicates the dominant influenza A subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant), and vertical bands are 95% confidence intervals of regional estimates. Mean A(H3N2) epidemic metric values were fit as a function of seasonal LBI diversity using Gaussian GLMs (effective Rt: inverse link) or Beta GLMs (epidemic intensity: logit link) with 1000 bootstrap resamples.

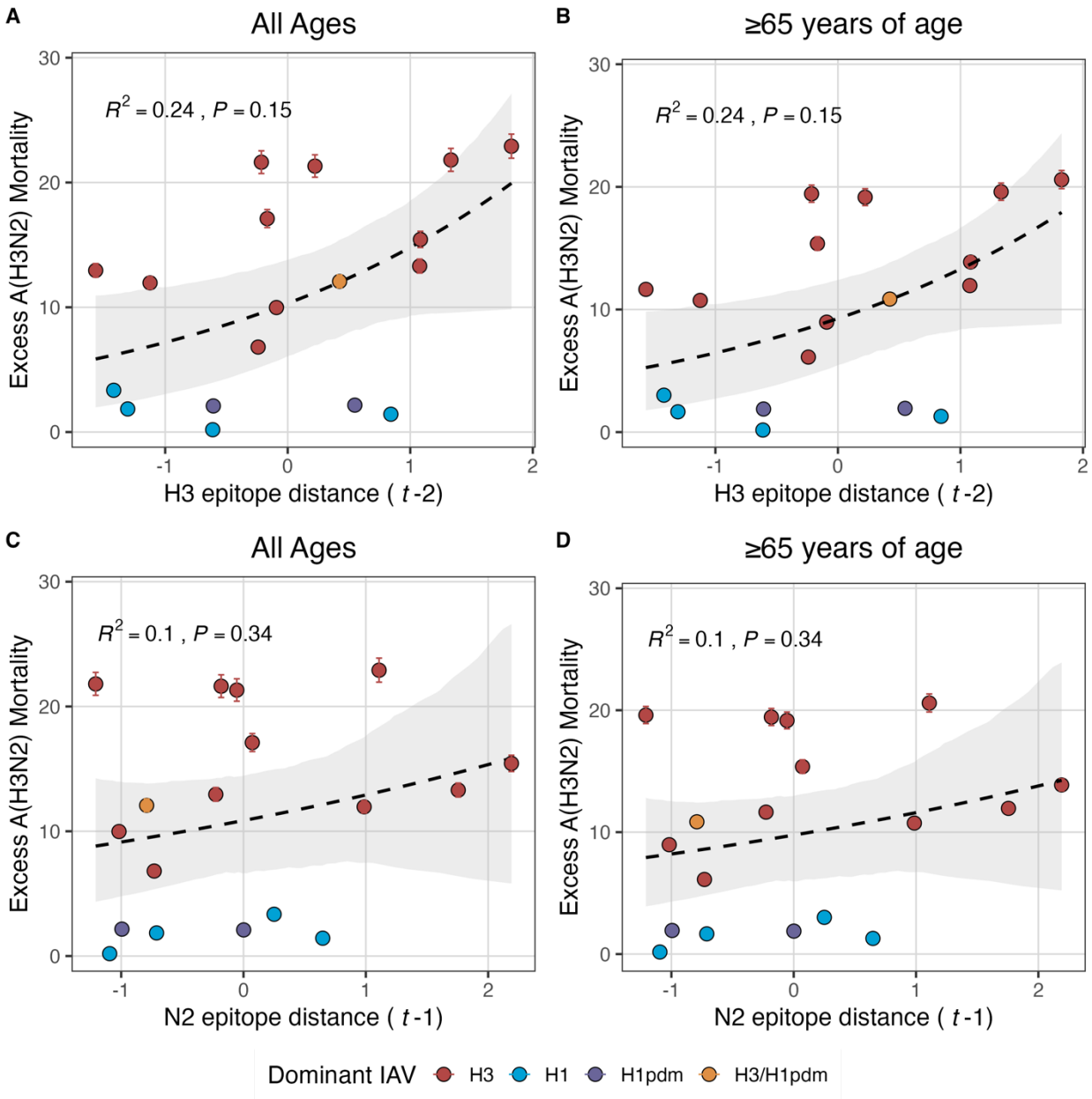


Figure S9. Excess influenza A(H3N2) mortality increases with H3 and N2 antigenic drift, but correlations are not statistically significant. The number of excess influenza deaths attributable to A(H3N2) (per 100,000 people) were estimated from a seasonal regression model fit to weekly pneumonia and influenza-coded deaths [127]. Seasonal epitope distance is the mean distance between strains circulating in season t and those circulating in the prior season ($t-1$) or two seasons ago ($t-2$). Distances are scaled to aid in direct comparison of evolutionary indicators. Point color indicates the dominant influenza A subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant), and vertical bars are 95% confidence intervals of excess mortality estimates. National excess mortality estimates were fit as a function of seasonal H3 or N2 epitope distance using Gaussian GLMs (log link) with 1000 bootstrap resamples.

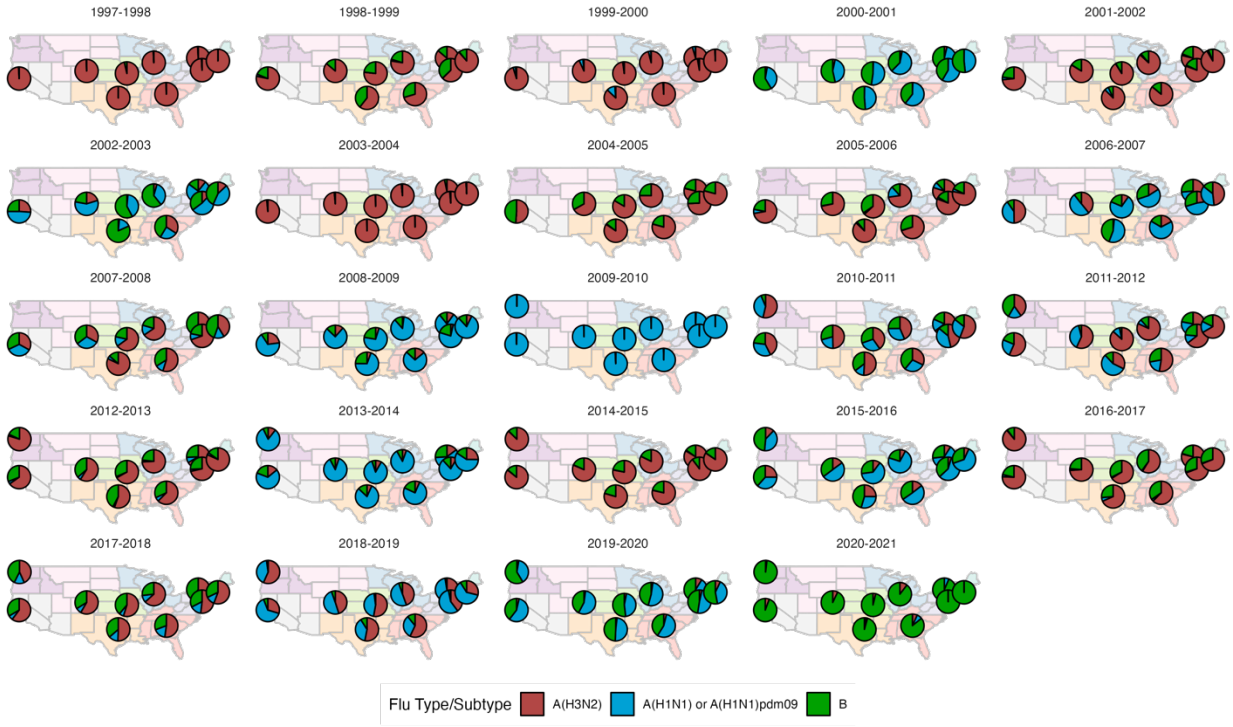


Figure S10. Regional patterns of influenza type and subtype incidence from seasons 1997-1998 to 2018-2019. Pie charts represent the proportion of influenza positive samples that were typed as A(H3N2), A(H1N1) or A(H1N1)pdm09, and B in each HHS region. Data for Region 10 (purple) were not available in seasons prior to the 2009 A(H1N1) pandemic.

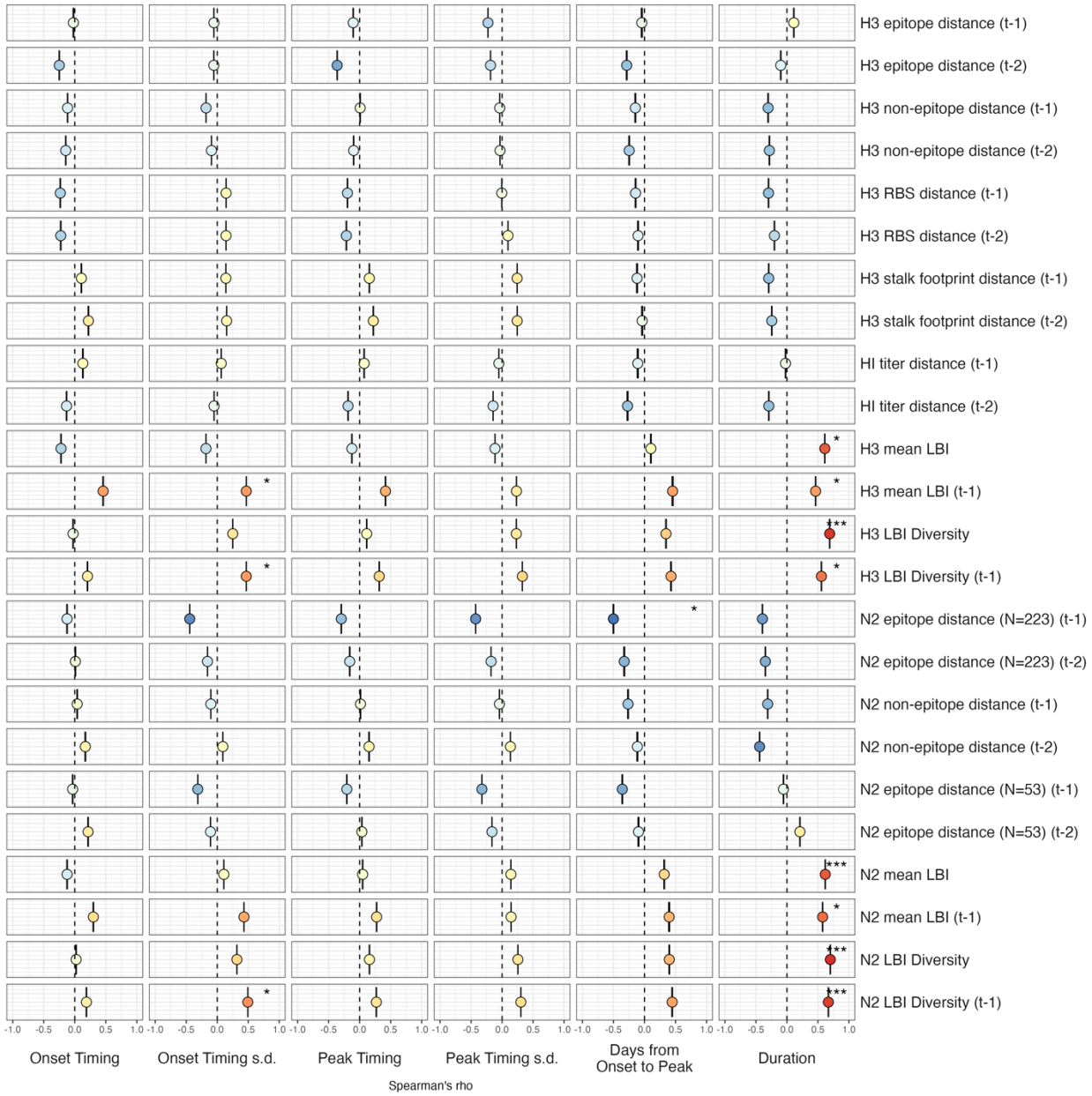


Figure S11. Univariate correlations between A(H3N2) viral fitness and epidemic timing. Mean Spearman correlation coefficients, 95% confidence intervals of correlation coefficients, and corresponding p-values of bootstrapped (N = 1000) viral fitness indicators (columns) and epidemic timing metrics (rows). Epidemic timing metrics are the week of epidemic onset, regional variation (s.d.) in onset timing, the week of epidemic peak, regional variation (s.d.) in peak timing, the number of days between epidemic onset and peak, and seasonal duration. Color indicates the strength and direction of the association, from dark red (strong positive correlation) to dark blue (strong negative correlation), and stars indicate statistical significance (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). Abbreviations: HI = hemagglutination inhibition, RBS: receptor binding site, t - 1 = one-season lag, t - 2 = two-season lag, LBI = local branching index.

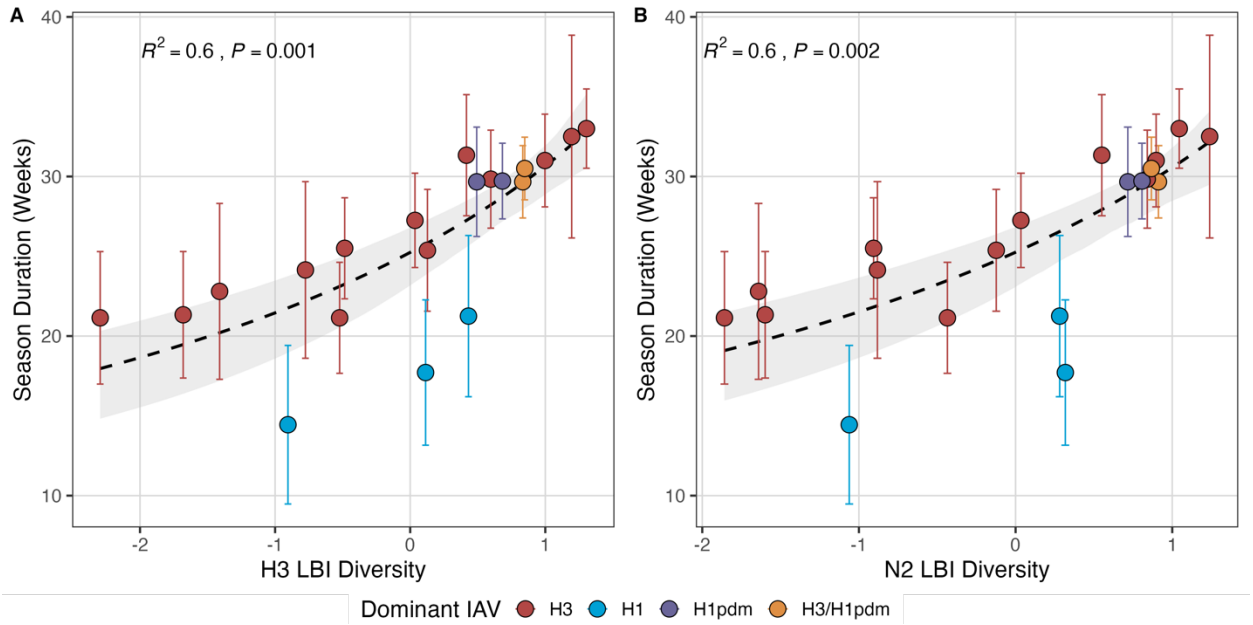


Figure S12. Seasonal duration increases with diversity in clade growth rates of circulating H3 and N2 lineages, measured as the Shannon entropy of local branching index (LBI) values. A. H3 LBI diversity and **B.** N2 LBI diversity during the current season positively correlate with seasonal duration. LBI values are scaled to aid in direct comparisons of H3 and N2 LBI diversity. Point color indicates the dominant influenza A subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant). Mean values of regional season duration were fit as a function of H3 LBI diversity or N2 LBI diversity using Gaussian GLMs (inverse link) with 1000 bootstrap resamples.

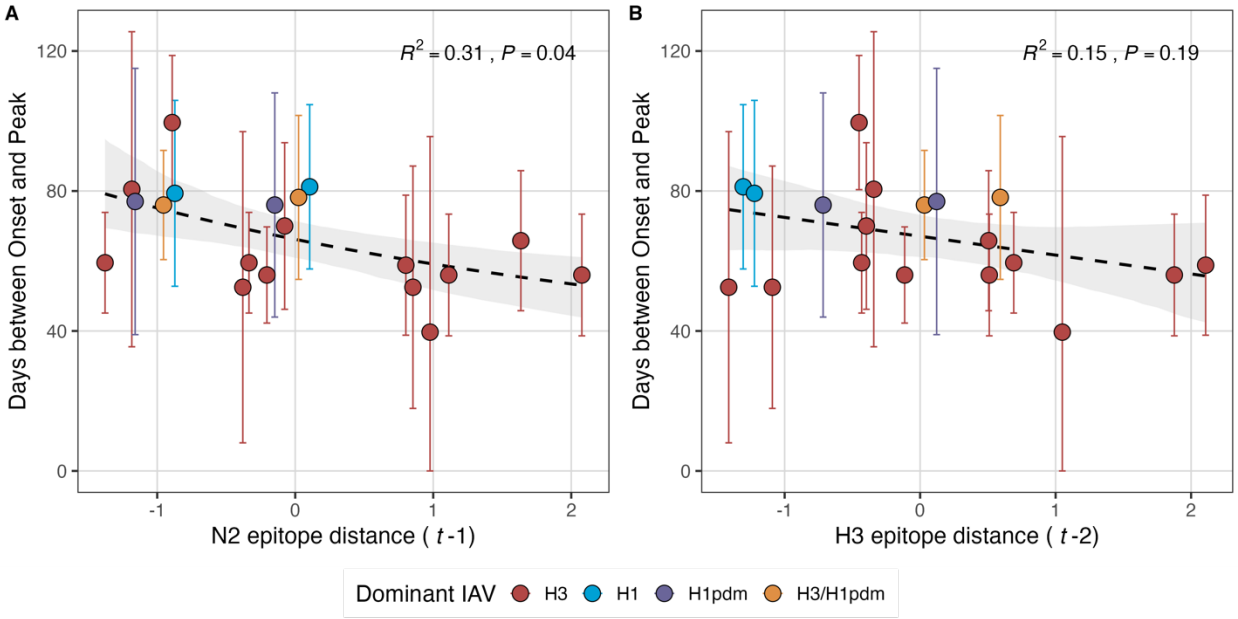


Figure S13. Epidemic speed increases with N2 antigenic drift. N2 epitope distance correlates with fewer days from epidemic onset to peak (**A**), while the relationship between H3 epitope distance and epidemic speed is less apparent (**B**). Seasonal epitope distance is the mean distance between strains circulating in season t and those circulating in the prior season ($t-1$) or two seasons ago ($t-2$). Distances are scaled to aid in direct comparison of evolutionary indicators. Point color indicates the dominant influenza A subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant). Mean values of regional days from onset to peak were fit as a function of H3 or N2 epitope distance using Gamma GLMs (inverse link) with 1000 bootstrap resamples.

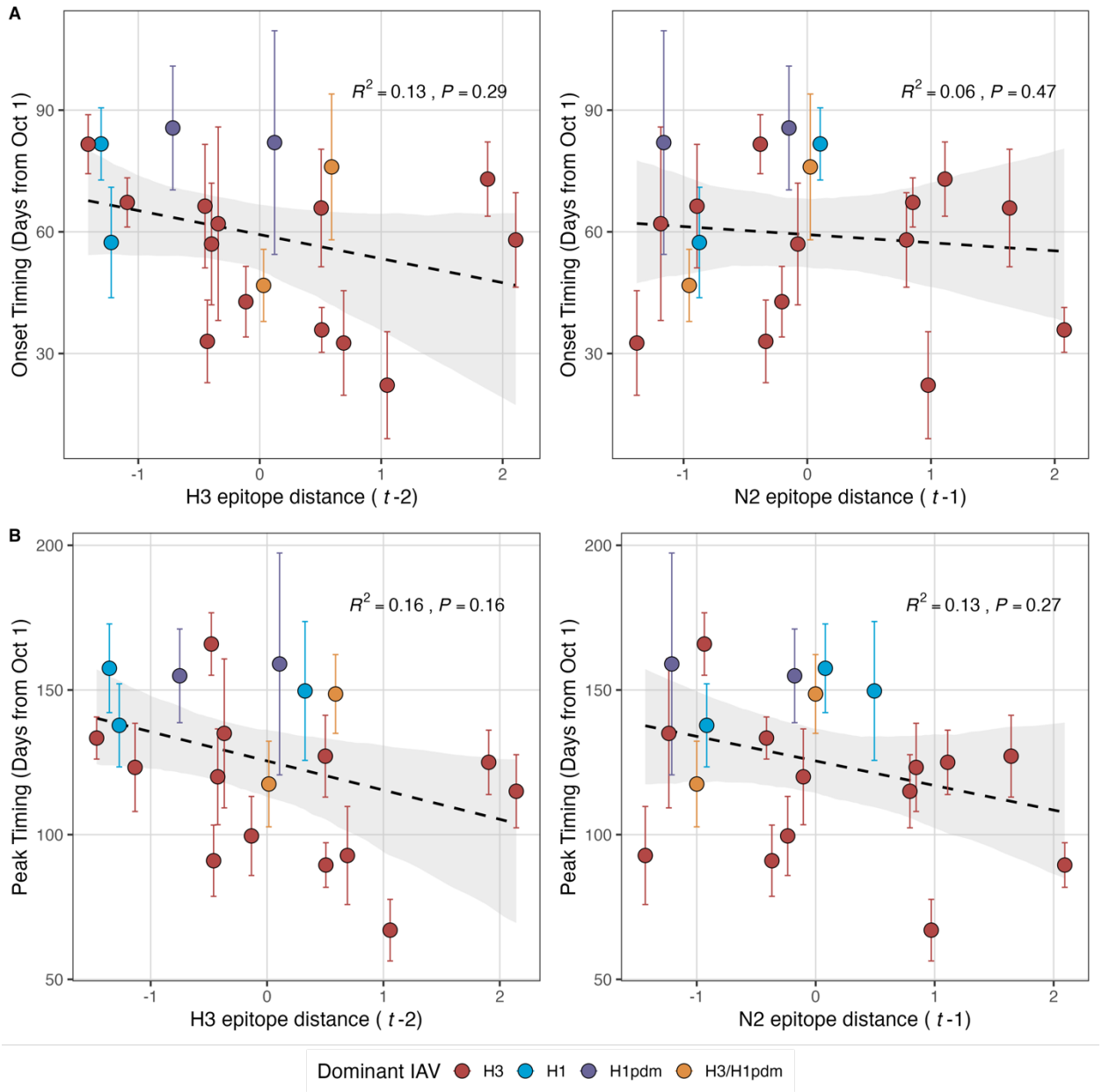


Figure S14. The timing of epidemic onsets and peaks are weakly correlated with H3 and N2 antigenic change. A. Epidemic onsets are earlier in seasons with increased H3 epitope distance ($t-2$), but the correlation is not statistically significant. **B.** Epidemic peaks are earlier in seasons with increased H3 epitope distance ($t-2$) or increased N2 epitope distance ($t-1$), but correlations are not statistically significant. Seasonal epitope distance is the mean distance between strains circulating in season t and those circulating in the prior season ($t-1$) or two seasons ago ($t-2$). Distances are scaled to aid in direct comparison of evolutionary indicators. Point color indicates the dominant influenza A subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant). Mean values of regional epidemic onsets and peaks were fit as a function of H3 or N2 epitope distance using LMs with 1000 bootstrap resamples.

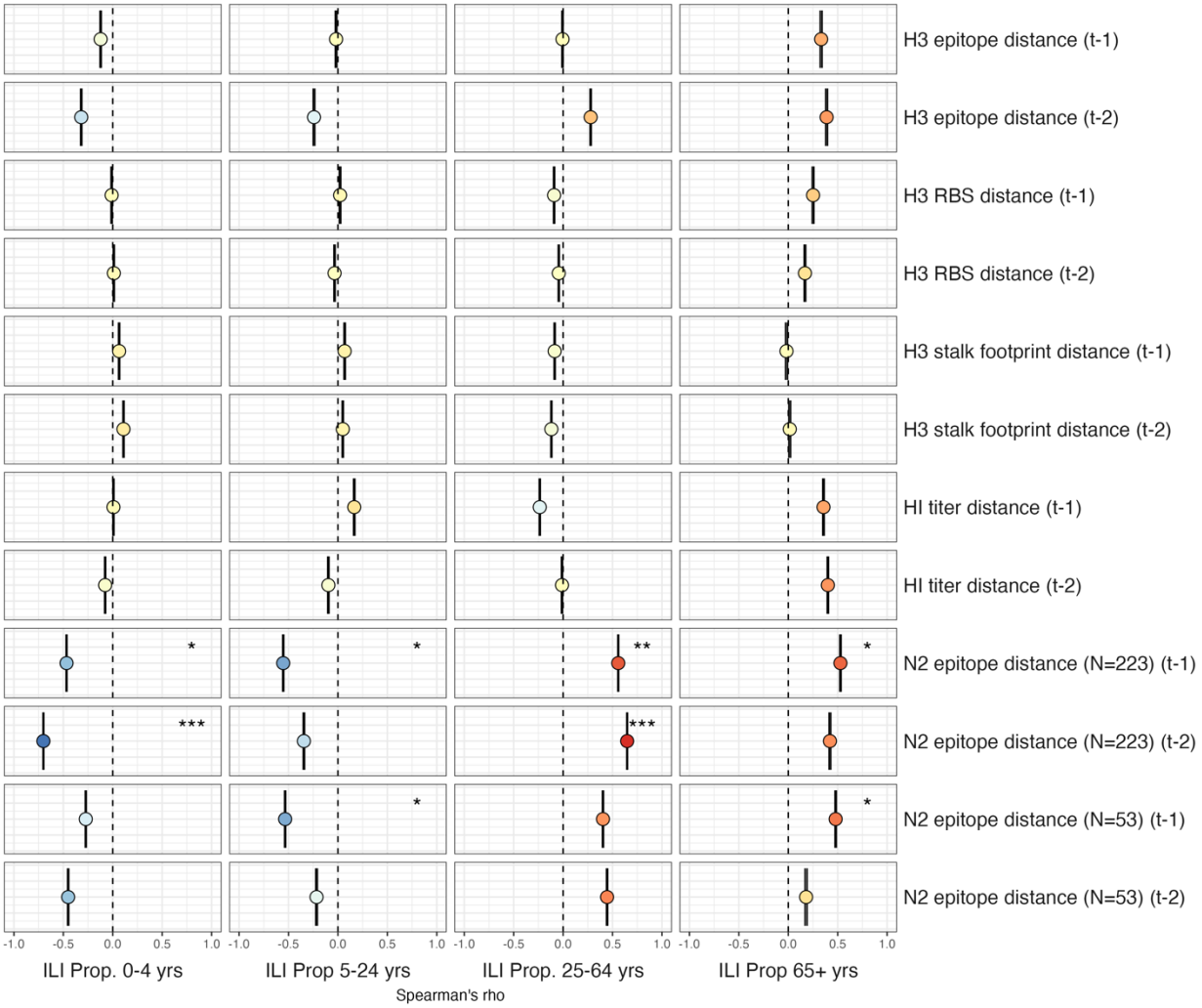


Figure S15. Univariate correlations between A(H3N2) antigenic change and the age distribution of outpatient influenza-like illness (ILI) cases. Mean Spearman correlation coefficients, 95% confidence intervals of correlation coefficients, and corresponding p-values of bootstrapped (N = 1000) evolutionary indicators (rows) and the proportion of ILI cases in individuals aged < 5 years, 5-24 years, 25-64 years, and ≥ 65 years (columns). Color indicates the strength and direction of the association, from dark red (strong positive correlation) to dark blue (strong negative correlation), and stars indicate statistical significance (* P < 0.05, ** P < 0.01, *** P < 0.001). Abbreviations: HI = hemagglutination inhibition, RBS: receptor binding site, t - 1 = one-season lag, t - 2 = two-season lag.

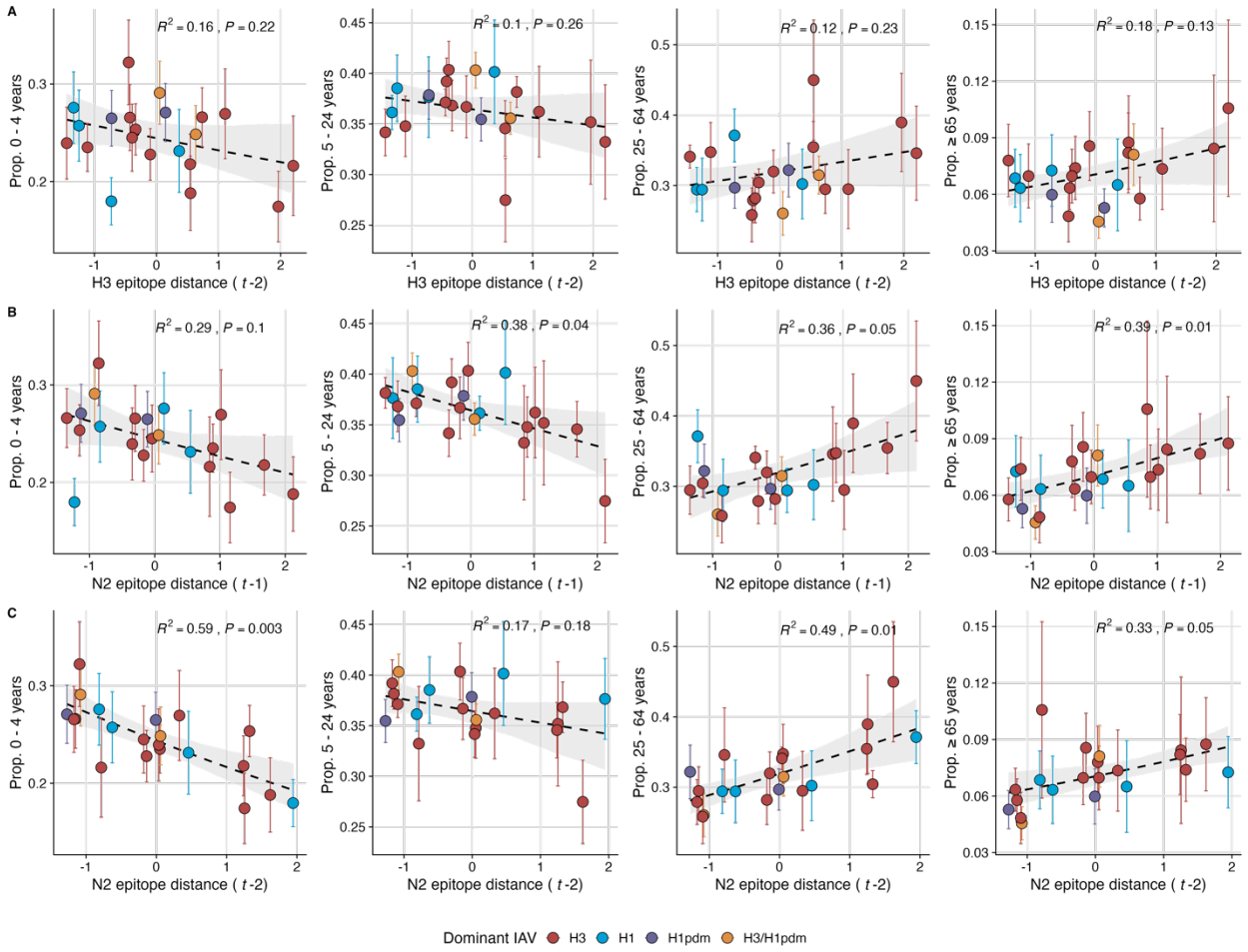


Figure S16. N2 epitope distance correlates with the age distribution of outpatient influenza-like illness (ILI) cases. Seasonal epitope distance is the mean distance between strains circulating in season t and those circulating in the prior season ($t-1$) or two seasons ago ($t-2$). Distances are scaled to aid in direct comparison of evolutionary indicators. Point color indicates the dominant influenza A subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant), and vertical bars are 95% confidence intervals of regional age distribution estimates. The fraction of cases in each age group were fit as a function of seasonal H3 or N2 epitope distance using Beta GLMs (logit link) with 1000 bootstrap resamples.

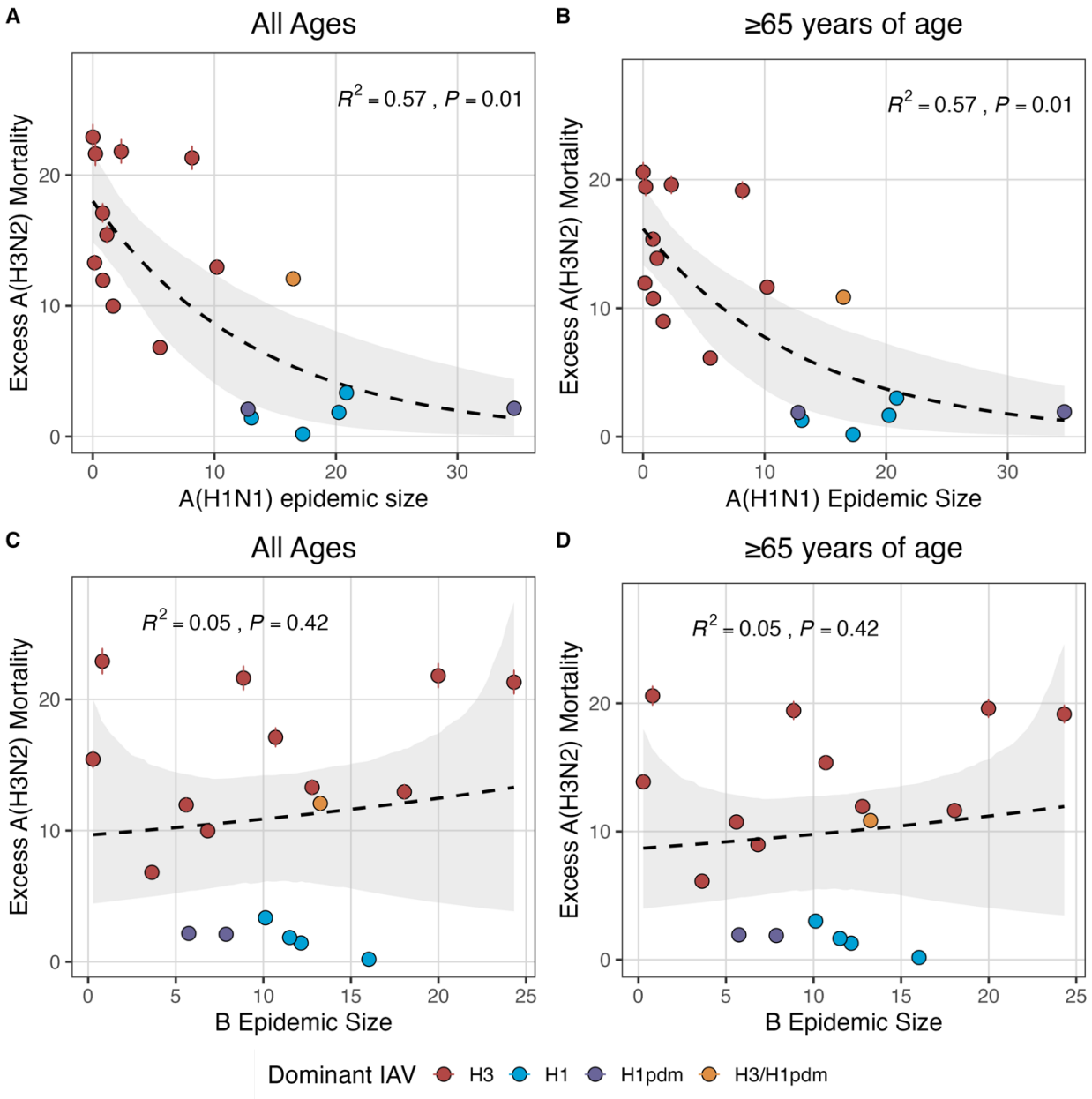


Figure S17. National excess influenza A(H3N2) mortality decreases with A(H1N1) epidemic size but not B epidemic size. Excess influenza deaths attributable to A(H3N2) (per 100,000 people) were estimated from a seasonal regression model fit to weekly pneumonia and influenza-coded deaths . Point color indicates the dominant influenza A subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant), and vertical bands are 95% confidence intervals of model estimates. National excess mortality estimates were fit as a function of seasonal A(H1N1) or B epidemic size using Gaussian GLMs (log link) with 1000 bootstrap resamples.

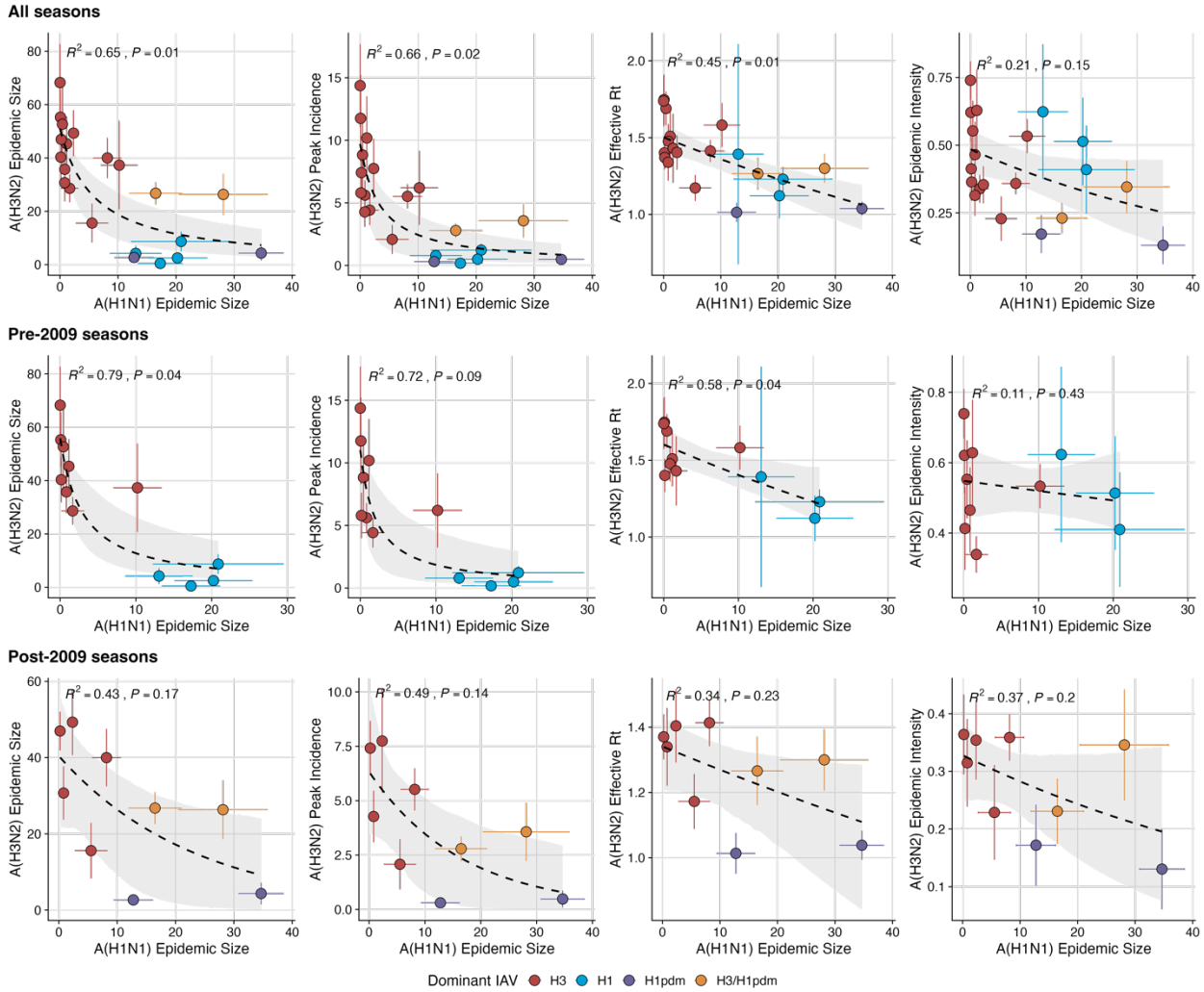


Figure S18. The effect of influenza A(H1N1) epidemic size on A(H3N2) epidemic burden during the entire study period (1997-2019) (top), pre-2009 seasons (middle), and post-2009 seasons (bottom). Influenza A(H1N1) epidemic size inversely correlates with A(H3N2) epidemic size, peak incidence, transmissibility (maximum effective reproduction number, Rt), and epidemic intensity. Point color indicates the dominant influenza A virus (IAV) subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant), and vertical and horizontal bands are 95% confidence intervals of regional estimates. Seasonal mean A(H3N2) epidemic metrics were fit as a function of mean A(H1N1) epidemic size using Gaussian GLMs (epidemic size, peak incidence: inverse link; effective Rt: log link) or Beta GLMs (epidemic intensity: logit link) with 1000 bootstrap resamples.

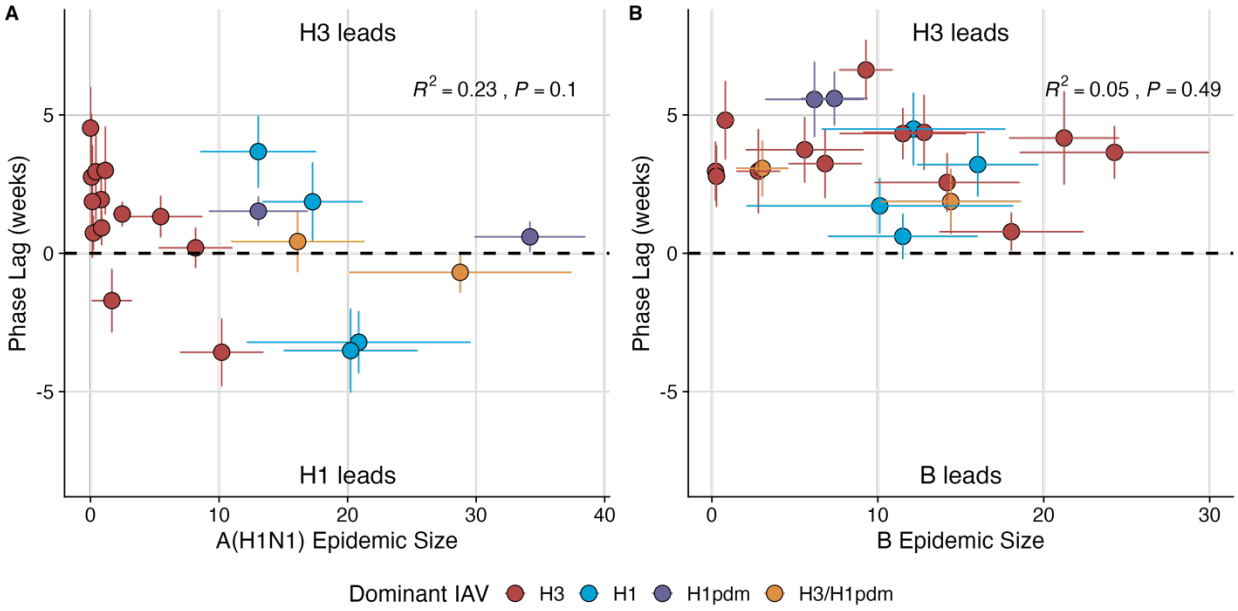


Figure S19. Wavelet analysis of influenza A and B epidemic timing. A. A(H3N2) incidence precedes A(H1N1) incidence in most seasons. Although A(H1N1) incidence sometimes leads or is in phase with A(H3N2) incidence (negative or zero phase lag), the direction of seasonal phase lags is not clearly associated with A(H1N1) epidemic size. **B.** A(H3N2) incidence leads B incidence (positive phase lag) during each season, irrespective of B epidemic size. Point color indicates the dominant influenza A subtype based on CDC influenza season summary reports (red: A(H3N2), blue: A(H1N1), purple: A(H1N1)pdm09, orange: A(H3N2)/A(H1N1)pdm09 co-dominant), and vertical bars are 95% confidence intervals of regional estimates. To estimate the relative timing of influenza subtype incidences, phase angle differences were calculated as phase in A(H3N2) minus phase in A(H1N1) (or B), with a positive value indicating that A(H1N1) (or B) incidence lags A(H3N2) incidence. To calculate seasonal phase lags, we averaged pairwise phase angle differences from epidemic week 40 to epidemic week 20. Seasonal phase lags were fit as a function of seasonal A(H1N1) or B epidemic size using LMs with 1000 bootstrap resamples.

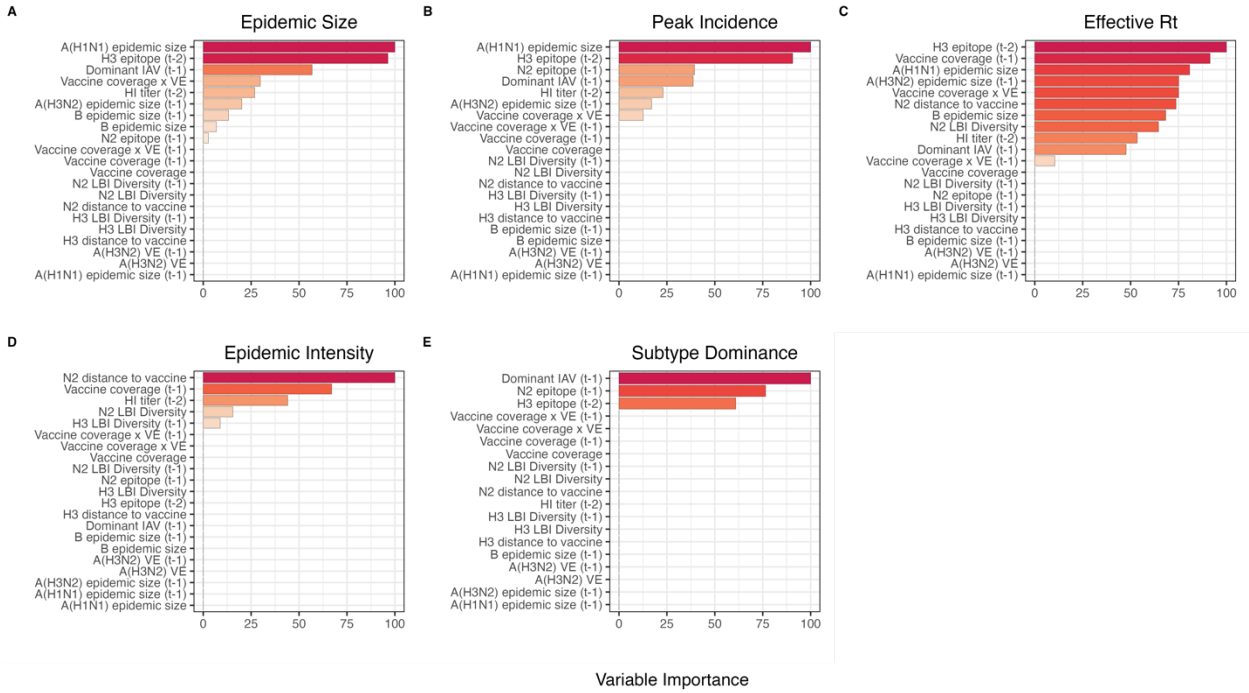
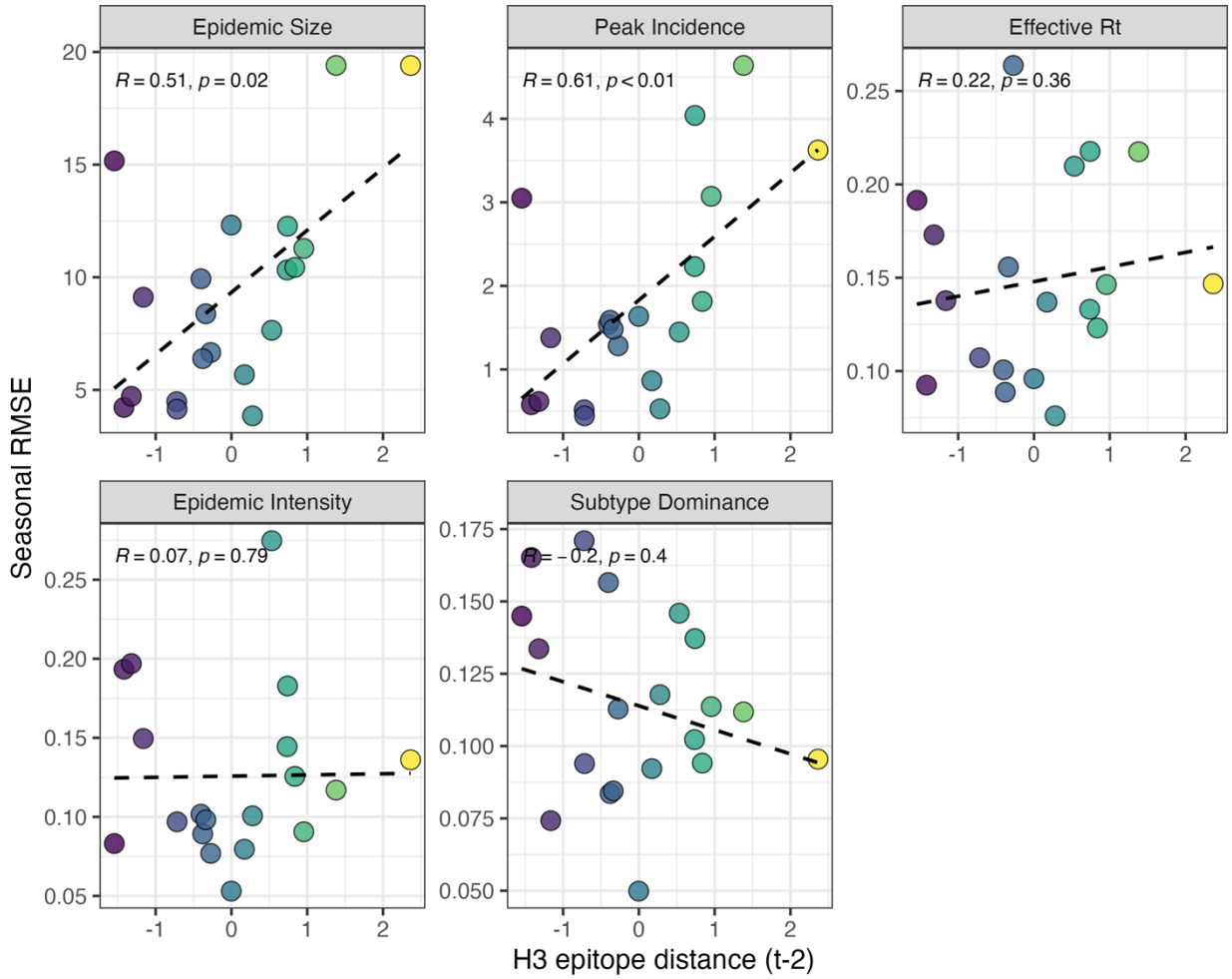
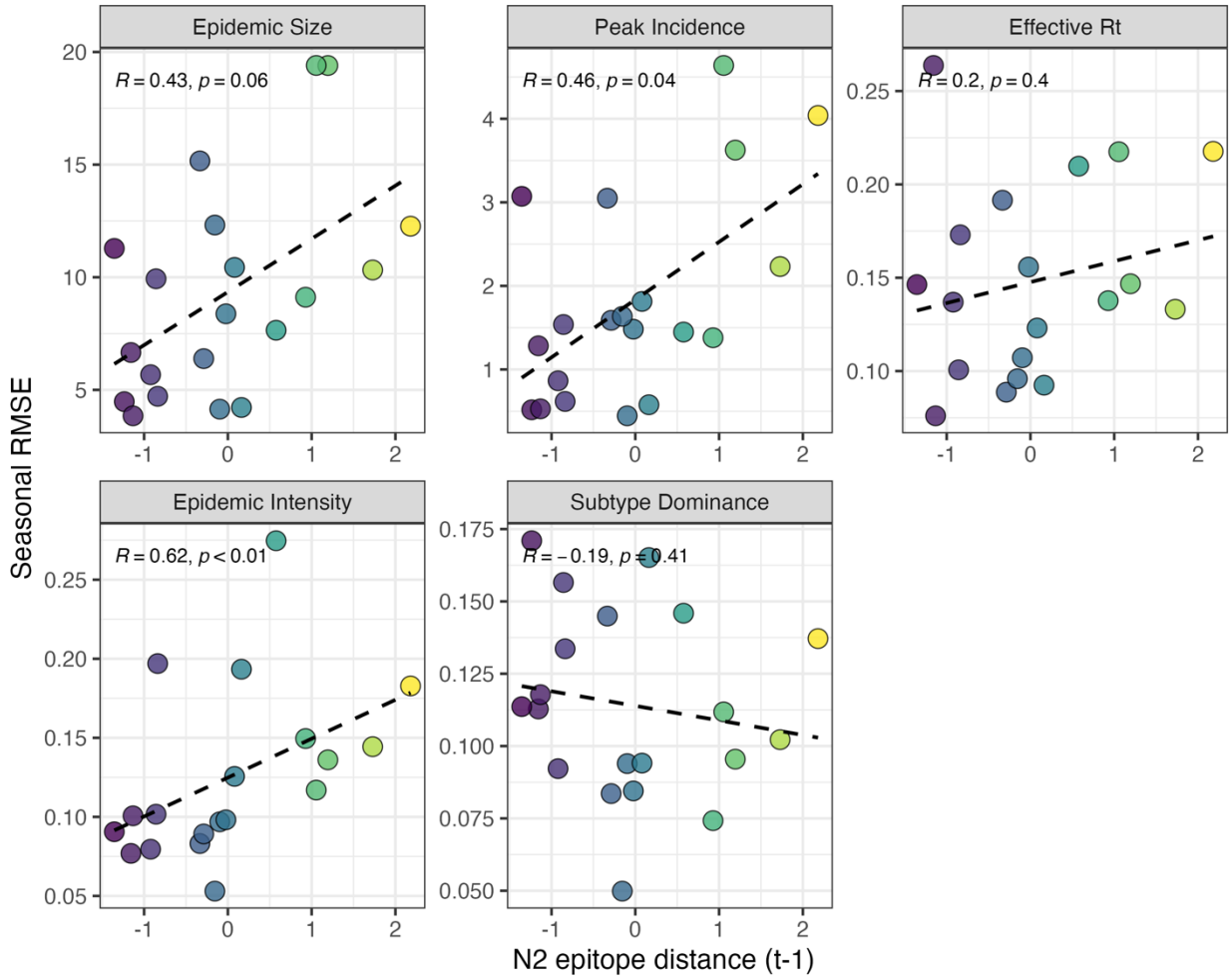


Figure S20. Variable importance rankings from LASSO models predicting A(H3N2) epidemic dynamics. Ranking of variables in predicting seasonal A(H3N2) **A.** epidemic size, **B.** peak incidence, **C.** transmissibility (effective reproduction number, Rt), **D.** epidemic intensity (inverse Shannon entropy), and **E.** subtype dominance. Models were tuned using a repeated leave-one-season-out cross-validated sample of the data. Variables are ranked by their coefficient estimates, with differences in prediction accuracy scaled by the total (null model) error. Abbreviations: HI titer = hemagglutination inhibition $\log_2$ titer distance, t - 1 = one-season lag, t - 2 = two-season lag, LBI = local branching index, peak = peak incidence, distance to vaccine = epitope distance between currently circulating strains and the recommended vaccine strain, VE = vaccine effectiveness.



Figures S21. Relationships between the predictive accuracy of random forest models and H3 epitope distance. Root mean squared errors between observed and model-predicted values were averaged across regions for each season, and results are facetted according to epidemic metric. Point color corresponds to the degree of H3 epitope distance in viruses circulating in season *t* relative to those circulating two seasons ago (*t – 2*), with bright yellow points indicating seasons with greater antigenic novelty. Spearman correlation coefficients and associated P-values are provided in the top left section of each facet.



Figures S22. Relationships between the predictive accuracy of random forest models and N2 epitope distance Root mean squared errors between observed and model-predicted values were averaged across regions for each season, and results are facetted according to epidemic metric. Point color corresponds to the degree of N2 epitope distance in viruses circulating in season t relative to those circulating in the prior season ($t-1$), with bright yellow points indicating seasons with greater antigenic novelty. Spearman correlation coefficients and associated P-values are provided in the top left section of each facet.