Different antigenic distance metrics generate similar predictions of influenza vaccine response breadth despite moderate correlation

Authors

- W. Zane Billings[1,2,*] (ORCID: 0000-0002-0184-6134);
- Yang Ge[1] (0000-0001-5100-0703);
- Amanda L. Skarlupka[3] (0000-0002-3654-9076);
- Savannah L. Miller[1,2] (0000-0003-2231-3510);
- Hayley Hemme[1,2] (0009-0002-6609-1390);
- Murphy John[1,2] (0009-0009-9591-1117);
- Natalie E. Dean[5] (0000-0003-3884-0921);
- Sarah Cobey[6] (0000-0001-5298-8979);
- Benjamin J. Cowling[7] (0000-0002-6297-7154);
- Ye Shen[1];
- Ted M. Ross[4,8] (0000-0003-1947-7469);
- Andreas Handel[1,2,*] (0000-0002-4622-1146).

Affiliations

- 1: Department of Epidemiology and Biostatistics, University of Georgia, Athens, GA, USA.
- 2: Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA, USA.
- 3: Vivli, Cambridge, MA, USA.
- 4: Center for Vaccines and Immunology, Department of Veterinary Medicine, University of Georgia, Athens, GA, USA.
- 5: Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, USA.
- 6: Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA.
- 7: Division of Epidemiology and Biostatistics, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong.
- 8: Florida Research & Innovation Center, Cleveland Clinic, Port St. Lucie, FL, USA.

∗: Corresponding author: W. Zane Billings; 131 B.S. Miller Hall, 101 Buck Road, Athens, GA 30602, USA; email wesley.billings@uga.edu. Alternate corresponding author: Andreas Handel; 124 B.S. Miller Hall, 101 Buck Road, Athens, GA 30602, USA; email ahandel@uga.edu.

Author contributions (CRediT)

- ZB: conceptualization, data curation, formal analysis, methodology, software, visualization, writing (original draft preparation), writing (review and editing)
- YG: data curation, methodology, writing (review and editing)
- ALS: data curation, methodology, writing (review and editing)
- SLM: methodology, validation, writing (review and editing)
- HH: methodology, validation, writing (review and editing)
- MJ: methodology, validation, writing (review and editing)
- NED: writing (review and editing)
- SC: writing (review and editing)
- BJC: writing (review and editing)
- YS: writing (review and editing)
- TMR: data curation, investigation, resources, writing (review and editing)
- AH: conceptualization, methodology, supervision, writing (original draft preparation), writing (review and editing), funding acquisition

Funding sources

- NED: partial funding from NIH contract(s)/grant(s) R01-AI139761.
- TMR: partial funding from the Georgia Research Alliance as an Eminent Scholar.
- AH: partial funding from NIH contract(s)/grant(s) U01AI150747, R01AI170116, and 75N93019C00052.
- YS: partial funding from NIH contract(s)/grant(s) R35GM146612, R01AI170116, and 75N93019C00052.
- SC: partial funding from NIH contract(s)/grant(s) R01AI170116.
- All other authors: none declared

The funders had no role in the decision to conduct or publish the study or in the formulation of the study.

Potential conflicts of interest: BJC has consulted for AstraZeneca, Fosun Pharma, GlaxoSmithKline, Haleon, Moderna, Novavax, Pfizer, Roche, and Sanofi Pasteur. None of these companies were involved in the formulation of the study or the decision to publish or conduct the study. All other authors declare no potential conflicts of interest.

Abstract

Introduction: Influenza continuously evolves to escape population immunity, which makes formulating a vaccine challenging. Antigenic differences between vaccine strains and circulating strains can affect vaccine effectiveness (VE). Quantifying the antigenic difference between vaccine strains and circulating strains can aid interpretation of VE, and several antigenic distance metrics have been discussed in the literature. Here, we compare how the predicted breadth of vaccine-induced antibody response varies when different metrics are used to calculate antigenic distance.

Methods: We analyzed data from a seasonal influenza vaccine cohort which collected serum samples from 2013/14 – 2017/18 at three study sites. The data include pre- and post-vaccination HAI titers to the vaccine strains and a panel of heterologous strains. We used that data to calculate four different antigenic distance measures between assay strains and vaccine strains: difference in year of isolation (temporal), p-Epitope (sequence), Grantham's distance (biophysical), and antigenic cartography distance (serological). We analyzed agreement between the four metrics using Spearman's correlation and intraclass correlation. We then fit Bayesian generalized additive mixed-effects models to predict the effect of antigenic distance on post-vaccination titer after controlling for confounders and analyzed the pairwise difference in predictions between metrics.

Results: The four antigenic distance metrics had low or moderate correlation for influenza subtypes A(H1N1), B/Victoria, and B/Yamagata. A(H3N2) distances were highly correlated. We found that after accounting for pre-vaccination titer, study site, and repeated measurements across individuals, the predicted post-vaccination titers conditional on antigenic distance and subtype were nearly identical across antigenic distance metrics, with A(H1N1) showing the only notable deviation between metrics.

Discussion: Despite moderate correlation among metrics, we found that different antigenic distance metrics generated similar predictions about breadth of vaccine response. Costly titer assays for antigenic cartography may not be needed when simpler sequence-based metrics suffice for quantifying vaccine breadth.

Introduction

Influenza viruses constantly evolve over time. As host immunity induces selective pressure, new influenza strains accumulate mutations, a phenomenon called antigenic drift (1–6). As mutations accumulate, antigenic drift leads to vaccine escape (7–9). Seasonal influenza vaccines are formulated based on the strains that are expected to circulate, but imperfect matches occur between selected vaccine strains and circulating strains in some years, and vaccine effectiveness (VE) varies annually (10). A major determinant of VE is the similarity between vaccine strains and circulating influenza strains (11–20). While previous studies have analyzed how mismatch between a circulating strain and the vaccine reduces VE, a full understanding of how viral changes affect vaccine response requires quantitative antigenic distance calculations (21–25). If our goal is the development of a broadly-protective (or even "universal") influenza vaccine, which induces a robust immune response to both historical and future influenza strains, defining a broad response is a crucial first step. Defining a broad response relies on accurate measurements of antigenic distance.

The most common method for quantifying antigenic distance between influenza strains is antigenic cartography, which relies on extensive serological data generated against a wide panel of strains (26). Briefly, statistical dimension reduction techniques are used to reduce large panels of serological data to fewer dimensions, and pairwise distances are calculated between strains in the reduced space. Serum samples from many individuals with wide assay panels are necessary to create stable cartographic maps. Cartographic distance has proven useful in understanding influenza evolution, but validating the ability of cartography to estimate population-level protection is difficult because of the required data (27–29). Sequence-based methods can accurately predict cartographic distance based on genetic sequences of influenza strains, but still rely on accurate serological data for calibration (30–38). Furthermore, multiple cartography methods yield different maps on the same data (26,27,39–41). Maps based on HAI titers also incorporate bias from HAI assays, which are often not replicable between labs (42,43) and do not always accurately reflect differences in common antigenic phenotypes, also called antigenic clusters (21–25,35,39,44–46). While cartographies can be generated from alternative assays (47–49), HAI is still the most common immunological assay used for influenza and the majority of highly-cited cartographies in use are based on HAI (26–28,50,51).

We can also compute antigenic distance without serological data. Simpler antigenic distance metrics calculated from genetic or amino acid sequences correlate with vaccine effectiveness at a population level (52–54), even though they only weakly correlate with antigenic distances derived from serological data (27,32,55). Influenza strains that evolve to escape prior immune response typically have mutations at the same important genetic sites (56–58), and advanced predictive models consistently identify properties of the amino acid sequence of the major antigens as important predictors of vaccine escape (59–62). Analyses of vaccine response or immunogenicity based on temporal (63–68) or sequence-based distances can provide information about the breadth of the vaccine

response (30,44,52–55,69–71). Taken together, these results imply that genetic analyses should provide important information about antigenic evolution without the need for serology. A direct comparison of antigenic distance methods is necessary to determine whether serological and sequence-based antigenic distance calculations can provide the same information in a practical setting. Specifically, we compare temporal distance (difference in the years of strain isolation), p-Epitope sequence distance (52), Grantham's sequence distance (72), and cartographic distance.

To compare the implications of multiple antigenic distance metrics on practical outcomes, we perform a secondary data analysis of an influenza vaccine cohort with a panel of HAI measurements to historical strains for each individual. We aim to assess whether low-cost measurements of antigenic distance between the vaccine strain and circulating strain may be similarly informative of the post-vaccination immune response. We find that, despite the modest correlation in antigenic distance metrics, these different metrics lead to similar conclusions about vaccine response to antigenically distant strains. Our results suggest that implementing costly antigenic analyses may not be necessary, as simple sequence-based measures lead to similar predictions about vaccine response as antigenic distance varies.

Methods

Data source

Immunological data

The data for our study are from a human vaccination cohort study which has been described in detail previously (73–75). Briefly, the study recruited participants at three study sites. The first two sites were Pittsburgh, PA, USA (PA site), and Port St. Lucie, FL, USA (FL site), beginning in the 2013/14 influenza season (approximately September through March (76)) and continuing through the 2016/17 influenza season. In January 2017, the study moved to Athens, GA, USA (UGA site). Participants visited the study site at least two times. At the first visit, patients completed a demographic questionnaire, gave a pre-vaccination serum sample, and received a Fluzone (Sanofi-Pasteur) seasonal influenza vaccine. At a follow-up visit approximately 21 days after the first visit, individuals returned and donated a post-vaccination serum sample. Individuals under 65 years of age received a standard dose Fluzone vaccination, and individuals aged 65 and older were given the choice between standard dose (SD) and high dose (HD) Fluzone vaccines. The study was a prospective, open cohort design where individuals could enroll in multiple years in the study, but were not required to re-enroll in every consecutive year.

Researchers tested the pre- and post-vaccination serum samples with a panel of hemagglutination inhibition (HAI) assays to the homologous strains (the strains included in the seasonal vaccine formulation), and a panel of historical, heterologous influenza virus strains. HAI assays are a common measurement for the strength of the antibody response,

and correlate with the amount of antibodies in a serum sample that bind to the receptor-binding domain of the influenza hemagglutinin protein (77,78). Strains included in the historical panel represented major lineages of circulating influenza viruses. See the Supplement for details on the Fluzone vaccine formulation and for a list of strains used in each season.

Each HAI assay in our dataset can be defined by its (1) subtype, (2) vaccine strain, and (3) assay strain. The broadest grouping is "subtype", which we use to describe both influenza A subtypes (H1N1 and H3N2) and influenza B lineages (Pre-divergence or Pre-split, Victoria-like and Yamagata-like). The vaccine strains associated with an HAI assay are the strains used in the Fluzone vaccine formulation in the season when the serum sample was collected. Each assay has three or four associated vaccine strains, depending on whether the individual who gave the serum sample received a trivalent or quadrivalent vaccine (see Supplement for details on the vaccine formulations). Finally, the assay strain for a given HAI assay is the strain of the actual virus added to the serum sample during the assay. We only compared vaccine strains and assay strains of the same subtype in our analysis.

For our secondary data analysis, we extracted previously deidentified records from the 2013/14 through 2017/18 influenza seasons. The study is ongoing and more assays are available, but the size of the historical panel was reduced after the 2017/18 season, and there would not be enough heterologous strains to estimate stable cartographic maps, so we limit our analysis to these seasons of data. Since examining the effect of vaccine dose was not our main focus here, and we previously observed dose-dependent differences in the heterologous response (79), we only included individuals who received SD vaccines in our study. We included all participants from the specified seasons who received SD vaccine and had records for both pre-vaccination and post-vaccination serum samples in our analysis. Our primary outcome of interest was the post-vaccination HAI titer (the reciprocal of the highest serum dilution which shows agglutination), which we log transformed:

$$\text{transformed titer} = \log_2\left(\frac{\text{HAI titer}}{5}\right).$$

Our final dataset for analysis contained one pair of transformed titers (pre- and post-vaccination) per person-year per assay strain in the historical panel, along with corresponding covariate measurements.

We divided the titer by 5 before taking the log because the HAI assay had a lower limit of detection (LoD) of 10, and an upper LoD of $20{,}480$. Values below the LoD were coded as titers of 5 in the dataset, corresponding to a transformed titer of 0. All observed values in our dataset were below the upper LoD. We used the same outcome definitions defined in our previous work on this dataset (79).

Sequence data

We computed the pairwise antigenic distance for all strains used in the dataset (see the Supplement for a complete list). We used four different methods to compute the antigenic distance: temporal distance, dominant p-Epitope distance (52), Grantham's distance (72), and cartographic distance (26). We calculated the temporal difference as the difference in the year of isolation between two strains (we had no assay strains with years of isolation subsequent to the vaccine strain, so all distances are positive). The dominant p-Epitope distance is the maximum of the Hamming distances (80) calculated for each of the five major epitope sites on the hemagglutinin head. Grantham's distance is similar to the Hamming distance on the entire HA sequence, but weights each substitution between strains by a score which is larger for amino acids with very different biochemical or biophysical properties. Finally, we conducted antigenic cartography using Racmacs (81) and reduced all of our cartographic maps to two dimensions. For complete details on antigenic distance calculation, see the Supplement.

To calculate the sequence-based, distances, we obtained sequences for the HA amino acid sequences for each of the strains used in the UGAFluVac data from either the U.S. National Center for Biotechnology Information (NCBI)'s GenBank database (82,83), the UniProt dataset (84), or GISAID's EpiFlu database (85,86). Accession numbers and sources for the sequence for each strain are shown in the supplement.

Statistical analyses

We first summarized demographic information about the cohort in a descriptive analysis, stratifying by measurements, individuals, and person-years to demonstrate the multilevel structure of our data.

We calculated reliability statistics between the difference antigenic distance metrics, using antigenic distances for all pairs of vaccine strains and assay strains that were present in the study design (instead of examining the reliability between all strains pairwise). As an omnibus test of measurement reliability, we calculated the intraclass correlation (ICC) using a Bayesian two-way mixed effects model for consistency and a single score, i.e., ICC(3, 1) in the Shrout-Fleiss taxonomy (87–89). The Supplement shows the exact model we fit and formula for calculating the ICC. To analyze which metrics drove disagreement or agreement, we also calculated the Spearman rank correlation coefficient (90) between each pair of antigenic distance metrics. We show credible intervals for the Spearman correlations in the Supplement, calculated with the Bayesian bootstrap (91).

We built generalized additive mixed-effects models (GAMMs) and linear mixed-effects models (LMMs) with the transformed post-vaccination titers as the otucome, (92,93) and adjusted for interval censoring (94) (see Supplement for details). To answer our primary question, we modeled antigenic distance in two ways. For the LMM, we included a linear effect of antigenic distance which was allowed to vary by subtype. For the GAMM, we modeled antigenic distance using a flexible semi-parametric spline which allows the relationship to be nonlinear, but constrained. We adjusted for effects of birth year, age,

sex, race/ethnicity, effects of the vaccine and assay strain, differences between study sites, and repeated measurements from the same individual.

We fit the models in a Bayesian framework using weakly informative priors chosen by a prior predictive simulation (92,95). We obtained posterior samples of the model parameters using the No U-Turn Sampler (NUTS) algorithm implemented by Stan (96,97), via the `brms` (98–100) and `cmdstanr` (101) packages for `R` (102). After obtaining the posterior samples, we calculated marginal posterior predictions for interpolated values of the normalized antigenic distance (103). We summarized the posterior prediction samples with a mean point estimate and 95% highest density continuous interval (HDCI). We compared the GAMM and LMM for each antigenic distance metric using the leave-one-out expected log pointwise predictive density (LOO-ELPD) which is conceptually similar to model selection using cross-validation in a frequentist scenario (105). See the Supplement for extensive details on our models.

To examine the differences in predictions across each of the antigenic distance metrics, we compared the slope and intercept for LMMs and the fold change in predicted post-vaccination HAI titer for the LMM and GAMM since the GAMM has no equivalent simple parametrization (fold change comparisons are shown in the Supplement). We extracted the fixed effects coefficients from the models, along with the random effects and residual variance components. We computed the variance contribution of the fixed effects (106) and calculated the proportion of variance explained by each of the variance components, defining the total variance as the sum of the residual variance parameter, the fixed effects variance contribution, and all random effects variance components.

Implementation

We conducted our analysis with R version 4.4.1 (2024-06-14 ucrt) (102) in RStudio version 2024.09.0+375 (107). Our analysis pipeline was implemented in `targets` (108). We used the packages `here` (109), `renv` (110), and the `tidyverse` (111) suite for data curation and project management and the packages `marginaleffects` (103), `tidybayes` (112), `ggdist` (113,114), `bayesboot` (115), and `loo` (105) for formal analysis. We used the packages `ggplot2` (117) and `GGally` (118) for generating figures; and the packages `gtsummary` (119) and `flextable` (120) for generating tables. We generated the manuscript using Quarto version 1.6.40 (121) along with the `R` packages `knitr` (122–124) and `softbib` (125). We implemented our Bayesian models with the `brms` package (98–100) using the `cmdstanr` backend and cmdstan version 2.34.1 (101) as the interface to the Stan programming language for Bayesian modeling. The Supplement contains more exhaustive details on our methodology, including instructions for reproducing our results. Our dataset and code are archived on GitHub (https://github.com/ahgroup/billings-comp-agdist-public) and Zenodo (DOI: 10.5281/zenodo.15522148).

Study ethics

Study participants in the UGAFluVac study were enrolled into the study with written informed consent at their respective study site. The study procedures, informed consent, and data collection documents were previously reviewed by the University of Georgia Institutional Review Board (IRB), and by WCG IRB. We only used deidentified data from UGAFluVac, and our study was determined to be not human research and exempt from review by the University of Georgia IRB.

Results

Data description

Our dataset included 54,101 pairs of pre-vaccination and post-vaccination HAI titer measurements drawn from 677 individuals who contributed 1,163 person-years to the study across three different study sites. The contributions of paired measurements, person-years, and unique participants from each study site are shown in Table 1. In a given year, each individual contributed three (trivalent vaccine in 2013/14 and 14/15) or four (quadrivalent vaccine from 2015/16 onward) homologous HAI assay pairs, along with a number of heterologous assay pairs, which varied by season due to the change in historical panels each year, and by individual due to random lab and assay issues. Each person-year represented in the data contributed a median of 48 HAI assay pairs (range: 8 to 52 pairs). Additional demographic information about our cohort is provided in the Supplement (summaries of race/ethnicity, sex assigned at birth, contributed person-years, age at enrollment, and pre-vaccination titer).

Table 1: Counts of HAI assay pairs, person-years, and new participants who enrolled for the first time that season contributed by each study site for the duration of the study. The PA and FL study sites operated from September 2013 to December 2016 and the GA study site began operating in January 2017 (during the 2016/17 influenza season).

	Season					
	2013/14	2014/15	2015/16	2016/17	2017/18	Total
Paired HAI assays, n						
FL	2459	6597	6656	6188	0	21900
PA	2163	3716	4131	3136	0	13146
UGA	0	0	0	6815	12240	19055
Overall	4622	10313	10787	16139	12240	54101
Person years, n						
FL	60	150	128	119	0	457
PA	73	88	81	64	0	306
UGA	0	0	0	145	255	400
Overall	133	238	209	328	255	1163

New participants, n						
FL	60	113	37	31	0	241
PA	73	46	2	12	0	133
UGA	0	0	0	145	158	303
Overall	133	159	39	188	158	677

Antigenic distance metrics have low or moderate correlation for all subtypes except A(H3N2)

First, we examined the overall agreement between the different distance metrics. We analyzed agreement using the intraclass correlation (ICC), shown in Table 2. ICC was low for all subtypes except A(H3N2), and the credible interval included zero for all subtypes except A(H3N2), so despite the moderate point estimate for B/Yamagata with a high upper limit, there was low consistency in antigenic distance measurements across methods. For A(H3N2), we observed moderate agreement across methods. Our ICC results indicate for each subtype except A(H3N2), at least one of the antigenic distance metrics systematically disagrees from the other.

Table 2: Intraclass correlation (ICC) across all antigenic distance measurements, calculated separately for each subtype or lineage (strain type). The posterior distribution for each ICC was calculated as the ratio of variance components for vaccine strain and assay strain divided by the sum of all variance components, estimated with a Bayesian model. Numbers shown are the mean and 95% highest density credible interval (HDCI) of the posterior distribution of ICCs.

Strain Type	ICC
H1N1	0.09 (0.00, 0.24)
H3N2	0.35 (0.20, 0.53)
B-Yam	0.21 (0.00, 0.42)
B-Vic	0.03 (0.00, 0.12)

To better understand the lack of overall agreement, we computed the Spearman rank correlation between each pair of metrics (again, separately for each subtype). Figure 1 shows the pairwise scatterplots and correlation coefficients. The pairwise correlations between distance measurements varied widely across subtypes and combinations, indicating that low agreement was not driven by a specific metric or subtype. All distance metrics tended to correlate well for H3N2. Distance metrics correlated highly for both influenza B lineages with the exception of the cartographic distance, which had a moderately high correlation with the other three distances for B/Yamagata and a low correlation with the other three distances for B/Victoria. The only high correlation for A(H1N1) was between Grantham and p-Epitope distance, with small correlations between

the other distance metrics. Grantham and *p*-Epitope distances correlated well for all strains (although it was notably lower for A(H1N1)), which we expected given the similarity between the measures. The Supplement contains a table with credible intervals for each correlation.

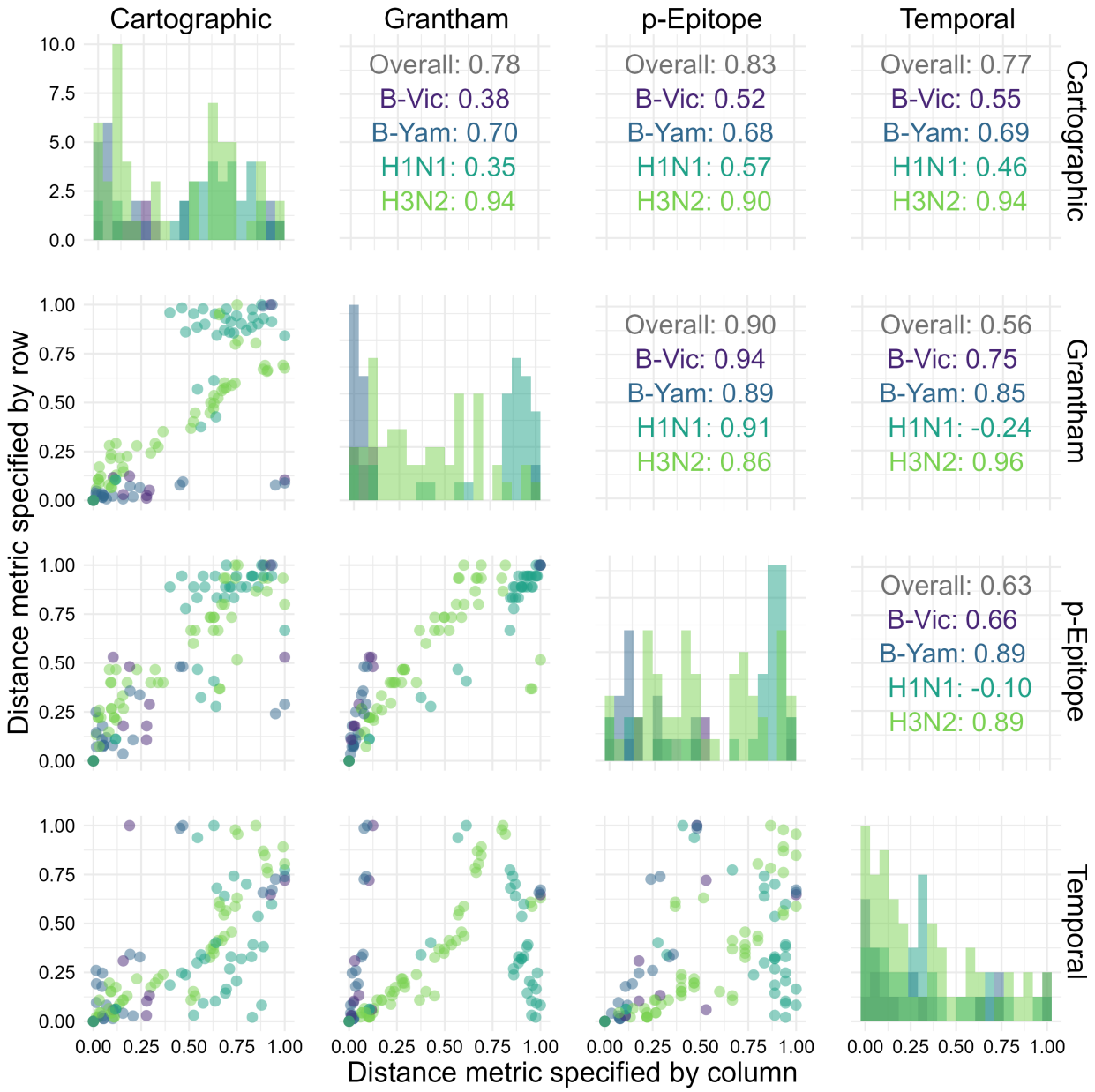

Figure 1: Distribution and correlation plots for each of the antigenic distance metrics. For each HAI assay in the dataset, we calculated the antigenic distance between the vaccine and assay strains with four different methods. We examined the distribution (shown along the diagonal) and the correlation between the different metrics for the same pairwise comparisons (we show pairwise scatterplots in the plots below the diagonal, and overall Spearman's correlation values in the plots above the diagonal). We include

each unique combination as only one point in this plot. We calculated correlation coefficients separately for each subtype – colors in the plot indicate subtype.

Predicted vaccine response breadth is similar across antigenic distance metrics, despite the low between-metric correlation

Examining the agreement and pairwise correlations between the different distance metrics is useful for understanding which metrics disagree most, but these disagreements do not necessarily translate into different predictions about vaccine response. We built LMMs and GAMMs to model the effect of antigenic distance after controlling for multiple host and assay features.

To quantify whether the effect of antigenic distance deviated strongly from a linear effect, we calculated the LOO-ELPD for the GAMM and LMM models fit with each antigenic distance metric, shown in Table 3. LOO-ELPD is comparable to (W)AIC or BIC, and differences in ELPD strongly supported the linear model for every antigenic distance metric. The ratio of the difference in ELPD was always much greater than its standard error, so the difference between models can be trusted for model selection. Including spline terms to account for nonlinearity did not improve the model fit.

Table 3: Differences in expected log pointwise predictive density (ELPD) from the best-fitting model, estimated by the leave-one-out (LOO) method for all models and all antigenic distance metrics. We fit the models separately for each antigenic distance metric, so comparisons are shown separately. The ΔELPD is the difference in ELPD between the LMM and the GAMM, so a positive number indicates the LMM performed better than the GAMM, and a larger number means the LMM outperforms the GAMM more. We show the ΔELPD ± its standard error, along with the ratio of the estimate to its standard error.

Metric	ΔELPD (LMM - GAMM)	ΔELPD / SE
Cartographic	108.14 (±19.96)	5.4
Grantham	203.64 (±24.23)	8.4
Temporal	47.16 (±11.17)	4.2
p-Epitope	290.25 (±35.48)	8.2

Figure 2 shows how the average post-vaccination titer predicted by the model changes along with antigenic distance for each subtype. For both influenza B lineages, the data were sparsely measured across the span of any of the antigenic distance metrics, making the GAMM predictions difficult to distinguish from the LMMs. Both influenza A subtypes showed a larger difference in predictions made by the GAMMs vs. the LMMs where the GAMMs predicted non-monotone relationships between post-vaccination titer and

antigenic distance. The LMM and GAMM were most similar for cartographic distance for both A(H1N1) and A(H3N2), perhaps suggesting that cartographic distance partially accounts for nonlinear effects of antigenic distance. There were some interesting trends in the shape of the spline curves, but the nonlinear effects for the p-epitope and Grantham distance did not appear to match the data well. Combined with the lack of ELPD support (Table 3), the spline models are likely picking up random fluctuations which may be partially driven by gaps in antigenic distance space rather than by true non-monotone signals (see the Supplement for an analysis of the gaps in antigenic distance space).

Since the linear model had better ELPD support for all metrics (Table 3), we focused on attempting to understand the effects in the linear model. Other than the normalized antigenic distance effect, the other effects were similar across the four models (which is what we expect). Table 4 shows the estimated fixed effects coefficients from our models. The effects of sex and race/ethnicity were negligible, and the effects of age and birth year appear to be highly negative because they are not identifiable in our dataset, but together they provide a non-negligible contribution for each individual. Log pre-vaccination titer had a strong positive effect on post-vaccination titer as expected. We did not interpret those effects further since we have not casually adjusted for effects other than antigenic distance. The effect of antigenic distance was negative for all four models, as we would expect, but the magnitude of the effect varied. While the point estimates were similar, the effect size for p-epitope was the smallest and the effect size for cartographic distance was the largest. The effect size for the cartographic distance also had the most density away from zero. Only the temporal distance model had an HDCI for the distance effect that included zero.

We also attempted to understand the variance contributions in the model by decomposing the variance (Table 5). The fixed effects explained the most variance of the three model components in all four models. The contribution of the residual variance was nearly identical in all four models, suggesting that the random effects are more important in some models than others, without explaining any additional variance. The variance explained by the assay strain, vaccine strain, study site, and subject variance components was similar across the four models, with the most noticeably different contribution being the effect of the subtype. The subtype apparently explained more variance in the temporal and Grantham distance models than in the cartographic and p-Epitope distance models, suggesting that those metrics might be more affected by differences in subtypes. Overall, the fixed effects were typically slightly more important than the random effects, but the variance explained by the random effects was still large for each model.

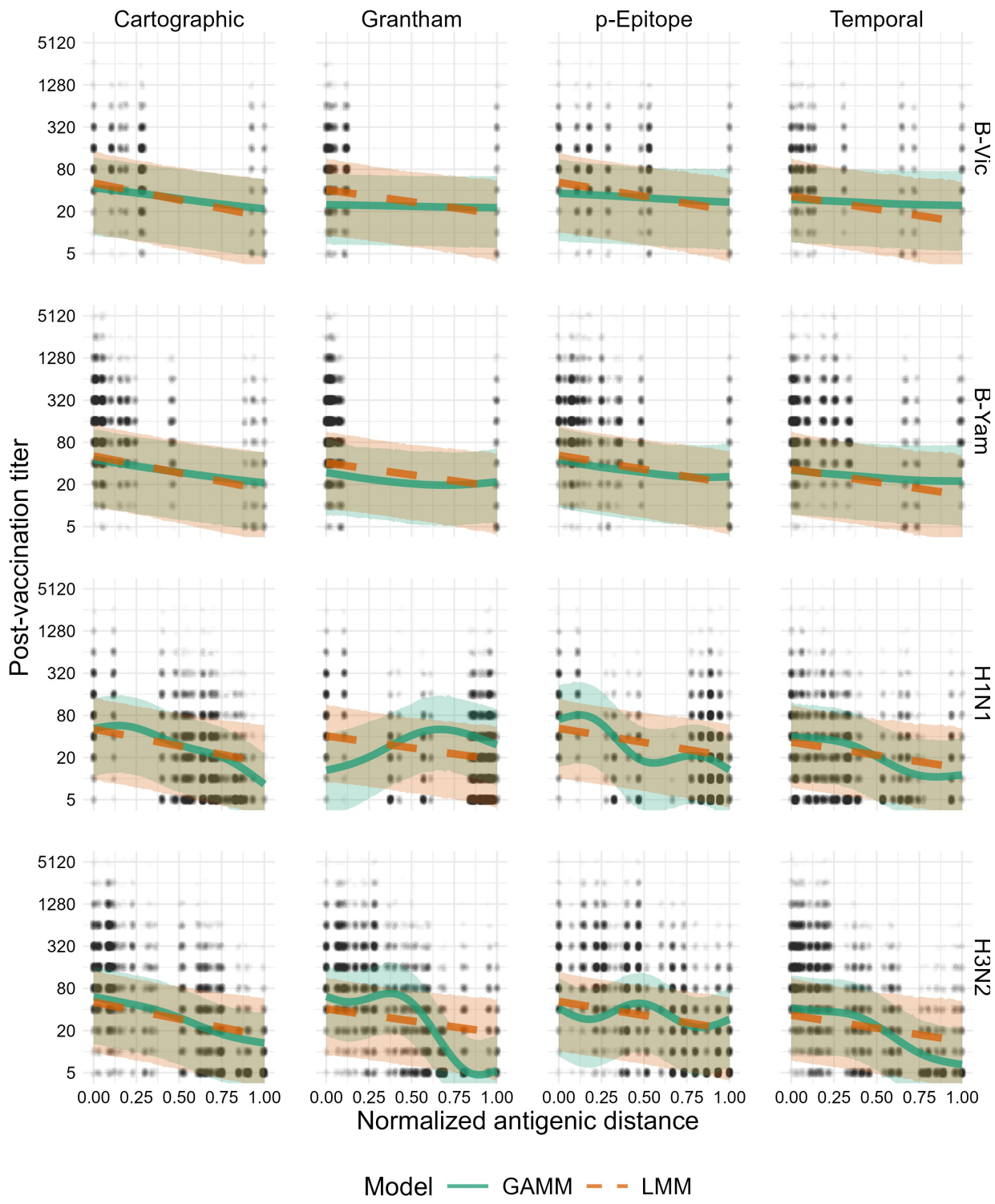

Figure 2: Model predictions for both the GAMM and LMM. Solid green lines and green ribbons show the mean and 95% highest density continuous interval (HDCI) for GAMM predictions. Dashed orange lines and orange ribbons show the mean and 95% HDCI for LMM predictions. Circular points show the data values. Each subplot shows the model

predictions for a particular subtype (changes by row) and distance metric (changes by column). Outcomes shown on the plot are predicted post-vaccination titers for an average individual to an average strain (see Supplement for computational details).

Table 4: Coefficients for all of the fixed effects included in our primary models. The model coefficients for scaled birth year, scaled age, sex (effect of being male relative to being female), race/ethnicity (effect of being non-Hispanic white or Caucasian vs. any other self-reported identity), log pre-vaccination titer, and normalized antigenic distance. We fit a separate model for each of the metrics, but the variables are standardized the same way across all four models so the coefficients are on the same scale across all models.

Metric	Birth Year	Age	Sex[1]	Race/Ethnicity[2]	Log pre-vaccination HAI titer	Normalized antigenic distance
Cartographic	-3.14 (-4.08,-2.20)	-3.46 (-4.37,-2.55)	0.01 (-0.05, 0.06)	0.03 (-0.02, 0.08)	0.78 (0.77, 0.79)	-1.61 (-2.42,-0.58)
Grantham	-3.19 (-4.12,-2.26)	-3.50 (-4.41,-2.60)	0.01 (-0.05, 0.07)	0.03 (-0.02, 0.09)	0.78 (0.78, 0.79)	-1.14 (-2.03,-0.02)
p-Epitope	-3.13 (-4.08,-2.19)	-3.45 (-4.36,-2.53)	0.01 (-0.05, 0.06)	0.03 (-0.02, 0.08)	0.78 (0.78, 0.79)	-1.33 (-2.12,-0.50)
Temporal	-3.17 (-4.11,-2.24)	-3.48 (-4.41,-2.58)	0.01 (-0.05, 0.06)	0.03 (-0.02, 0.08)	0.78 (0.78, 0.79)	-1.24 (-2.38, 0.16)

[1]Reference: Male (vs. female)

[2]Reference: Non-Hispanic white (vs. other)

Table 5: Variance contributions to the total variance estimated in the model. To estimate the fixed effects variance contribution as the variance of the estimated linear predictor, while the residual variance and random effects variance contributions (all variance contributions other than the fixed effects and residual variance) are estimated as model parameters. All contributions are rounded to the nearest percent and may not sum (rowwise) to 100 due to rounding error.

				Specific random effects				
Metric	Residual variance	Fixed effects	Total random effects	Subtype	Assay strain	Vaccine strain	Study site	Subject
Cartographic	12% (9, 15)	50% (38, 62)	36% (22, 52)	13% (4, 26)	1% (1, 2)	3% (1, 5)	11% (0, 31)	6% (4, 7)
p-Epitope	12% (9, 15)	50% (38, 60)	36% (24, 51)	11% (3, 24)	5% (3, 6)	3% (1, 5)	10% (0, 28)	6% (4, 7)
Grantham	11% (9, 14)	48% (37, 58)	40% (27, 53)	17% (5, 33)	4% (2, 5)	3% (1, 7)	8% (0, 25)	6% (4, 7)
Temporal	11% (8, 13)	44% (33, 54)	44% (32, 57)	23% (9, 40)	4% (2, 5)	3% (1, 7)	7% (0, 23)	5% (4, 6)

Predictions made by different antigenic distance metrics are similar after accounting for host factors

Finally, we directly compared estimates from the models across normalized antigenic distance metrics for each subtype (Figure 3). Since the LMM is easier to interpret and was supported by our ELPD analysis, we examined the slope and intercept for each subtype across the four antigenic distance metrics. The intercepts (representing the predicted post-vaccination titer to the homologous strain of the specified subtype for an individual with no pre-vaccination antibodies) were similar across all metrics regardless of the subtype. The slopes varied more, indicating that the antigenic distance had a stronger effect on predicted titer for some metrics and subtypes. For both B lineages, estimates of the slope were nearly identical across antigenic distance metrics. For A(H1N1), the cartographic distance model had a lower slope than the other three antigenic distance metrics, but the credible interval still overlapped with the credible interval for the temporal distance. For A(H3N2), the slope for the p-Epitope distance was much smaller than the other slopes (reflecting our results in Figure 2), despite the high correlation between the antigenic distances for A(H3N2) (Figure 1). We can only perform a visual inspection of these overlaps, because there is no existing approach to combine posterior distributions across the four models.

Furthermore, these estimates do not take variance from the random effects in our model into account. To analyze predictions for both the LMM and GAMM, with the random effects variances included in uncertainty calculations, we directly compared predictions from the models and saw much higher overlap (shown in the Supplement), as we would expect when we include all of the variance in the data.

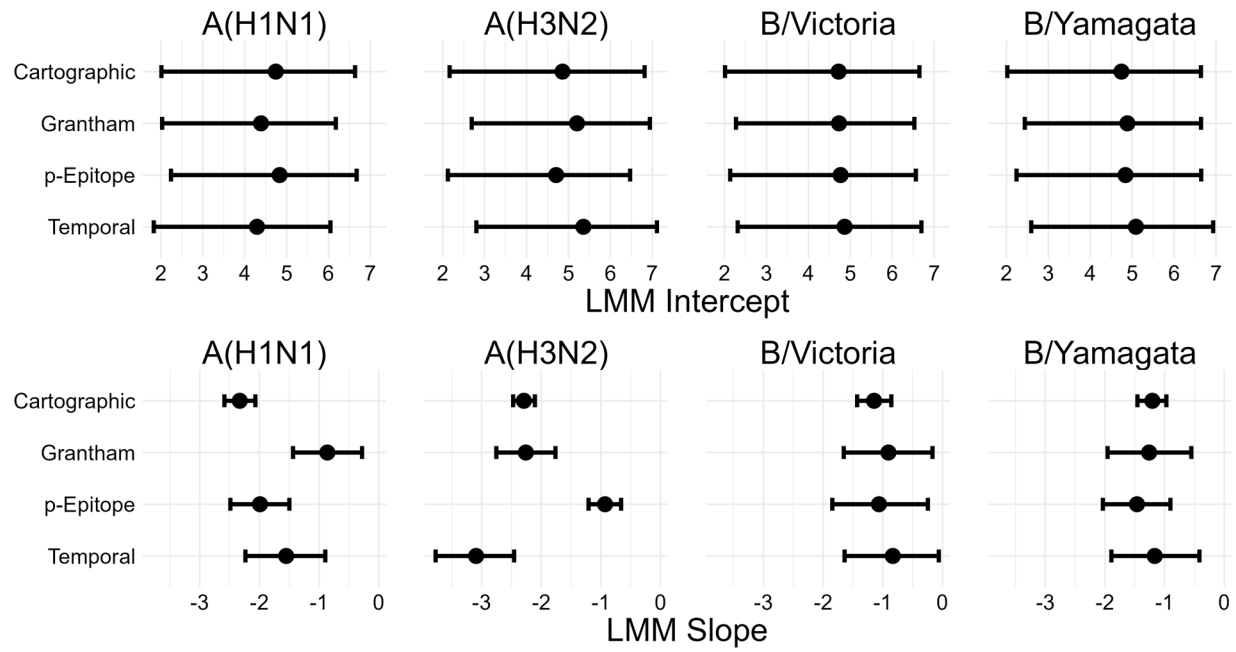

Figure 3: Intercept and slope estimates stratified by subtype for each LMM (one for each distance metric). Points and intervals show the mean and 95% HDCI of posterior samples of the indicated parameter. The top row of plots shows the mean and CI for estimates of the intercept, and the bottom row of plots shows the mean and CI for estimates of the slope. Columns of plots indicate which subtype the slope and intercept are for.

We compare the relative LOO-ELPD for each model in Table 6. The models are fit to the same set of predictors and data points and the antigenic distances are normalized, so the ELPDs are on the same scale and we can directly compare them. We found that all of the models had very similar performances – while the ELPDs were different between the four models, each contrast was smaller than the SE for either ELPD. For example, while the cartographic model had an ELPD around 150 points lower than the p-Epitope model, the SE for both estimates was around 470, so we cannot assume that these contrasts are meaningful differences. All of the models appeared to fit the data equally well.

Table 6: Expected log pointwise predictive density (ELPD) calculated for each of the linear mixed-effects models (LMMs) using the leave-one-out (LOO) method. For each metric, we show the estimated ELPD ± its standard error. The differences between the model ELPDs were negligible.

Metric	LMM LOO-IC
Cartographic	151131.5 ± 471.6
p-Epitope	151188.6 ± 471.7
Grantham	151250.6 ± 472.1
Temporal	151188.4 ± 472.2

Discussion

We computed multiple antigenic distance metrics on the same set of influenza strains. Using immunogenicity data from a human cohort, we were able to compare cartographic data to sequence-based, biophysical, and temporal antigenic distance measures which have been used before for analyzing vaccine breadth. We then fit linear mixed-effects models (LMMs) and generalized additive mixed models (GAMMs) to the immunological data separately for each cohort, controlling for subtype, pre-vaccination titer, and multiple sources of random variation. By comparing the predictions and parameters from the estimated models across the four antigenic distance metrics, we were able to assess the similarity of the metrics in a more practical context.

Despite moderate correlations between the four antigenic distance measures for all subtypes except A(H3N2), we found that all four antigenic distance measures produced similar predictions about the heterologous vaccine response, regardless of subtype. Unexpectedly, the subtype generating the most different predictions was A(H3N2), which had the highest correlation between metrics. After we account for important confounders and other sources of variation, the differences between metrics seemed to disappear, with the exception of the unusually small slope for p-Epitope distance for influenza A(H3N2). Along with our pointwise prediction comparisons (shown in the supplement), these results suggest a systematic disagreement on the vaccine outcome scale between p-Epitope distance and other metrics for A(H3N2), which contrasts with the high pairwise correlations between p-Epitope and other metrics for this subtype. Perhaps important antigenic changes for H3N2 have occurred outside of the immunodominant epitopes, or features like glycosylation which might be more easily captured by Grantham or cartographic distance are important, but we were unable to identify specific strains driving this effect. Alternatively, the difference could be due to some form of noise or sampling error in our study — we have no data from equivalent human cohort studies with wide heterologous panels to compare our results to, so we do not know if this result is consistent.

Our overall results could imply that the differences between antigenic distance metrics can appear large but are small compared to between-subject and between-study variability in real life, or that accounting for interindividual differences or pre-vaccination titer helps to explain the differences between metrics. We also found that a linear model was sufficient for explaining the relationship between post-vaccination titer and antigenic distance, rather than a nonlinear model. For example, we might expect a tapering effect or a sharp drop-off, which could both be produced by the GAMMs. Notably, we even found that temporal distance tends to produce similar predictions to cartographic distance in this setting, despite the evidence for epochal antigenic evolution and emergence or circulation of multiple clusters in a single year (3,9,59,126).

While we used data from a multicenter study with tens of thousands of measurements and over one thousand contributed person-years, our study still has some weaknesses. First, as a secondary data analysis, none of the data were designed with our questions in mind. While we have attempted to control for as much confounding as possible, we lack data on the exposure histories, including infections and prior vaccinations outside of the study, of individuals in our cohort which could confound our results (9,127). Our results also only apply to the split-inactivated Fluzone standard dose vaccine. Higher doses can either help or hinder heterologous responses (128–130), and in a previous study we found that the heterologous antibody response varied by Fluzone vaccine dose (79), so our results might change for other vaccine doses or formulations. A balanced design with randomized vaccine design would be preferable for understanding the impact of vaccine design on agreement between antigenic distance metrics.

We also used cartographies based on our pre-immune human data, which were generated on the same data we analyzed. With access to multiple cartographies on the same data set

or imputation techniques (131,132) we could treat different cartographies as different antigenic distance metrics and compare cartographic distances in the same way. Our metrics also did not all cover antigenic distance space evenly as the strains in the historical panel were selected to cover a wide variety of years. However, there were several "gaps" between discrete antigenic distance values for A(H1N1) and the two B lineages, which could impact our estimates (see Supplement for details), and a broader panel with more evenly spaced strains would make our effect estimates more precise. Finally, we have no real proxy for the response to "future" strains. We could get a better predictive understanding of how the vaccine generates immune responses to future strains by testing serum samples from, say, 2016, to novel vaccine strains which have emerged since the samples were collected. Such measurements would allow us to validate the use of the historical panel as a proxy for future vaccine response. Longitudinal studies designed with long-term collection and multiplex assays in mind would be beneficial for answering similar questions about antigenic distance and vaccine breadth.

Overall, we found that simple antigenic distance metrics like Grantham's distance generated very similar predictions about vaccine breadth to distances based on antigenic cartography in our study. While some distance metrics potentially deviated, the effect was subtype specific (p-Epitope for A(H3N2) strains). While cartography is important for understanding the antigenic diversity and evolution of influenza, researchers analyzing vaccine breadth should not be afraid to use easier, potentially less biased metrics of antigenic distance.

Acknowledgements

We thank William Michael Landau (Eli Lilly and Company, Indianapolis, IN, USA) and Eric R. Scott (University of Arizona, Tucson, AZ, USA) for their generous help with computational issues and pipeline development. Additionally, we thank Michael A. Carlock (Cleveland Clinic Florida Research & Innovation Center, Port St. Lucie, FL, USA) for assistance with obtaining data. This study was supported in part by resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia's Office of the Vice President for Research and Office of the Vice President for Information Technology.

We gratefully acknowledge all data contributors, i.e., the Authors and their Originating laboratories responsible for obtaining the specimens, and their Submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which this research is based. Similarly, we gratefully acknowledge all data contributors for generating the genetic sequence and metadata and sharing via NCBI GenBank and UniProt, on which this research is based.

References

1. Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, et al. Influenza. Nature Reviews Disease Primers. 2018 Jun;4(1):1–21.

2. van de Sandt CE, Kreijtz JHCM, Rimmelzwaan GF. Evasion of Influenza A Viruses from Innate and Adaptive Immune Responses. Viruses. 2012 Sep;4(9):1438–76.

3. Petrova VN, Russell CA. The evolution of seasonal influenza viruses. Nature Reviews Microbiology. 2018 Jan;16(1):47–60.

4. Morens DM, Taubenberger JK, Fauci AS. The 2009 H1N1 Pandemic Influenza Virus: What Next? mBio. 2010 Sep;1(4):e00211–10.

5. Paules CI, Fauci AS. Influenza Vaccines: Good, but We Can Do Better. The Journal of Infectious Diseases. 2019 Apr;219(Suppl 1):S1–4.

6. Kim H, Webster RG, Webby RJ. Influenza Virus: Dealing with a Drifting and Shifting Pathogen. Viral Immunology. 2018 Mar;31(2):174–83.

7. Koelle K, Cobey S, Grenfell B, Pascual M. Epochal Evolution Shapes the Phylodynamics of Interpandemic Influenza A (H3N2) in Humans. Science. 2006 Dec;314(5807):1898–903.

8. Zinder D, Bedford T, Gupta S, Pascual M. The Roles of Competition and Mutation in Shaping Antigenic and Genetic Diversity in Influenza. PLOS Pathogens. 2013 Jan;9(1):e1003104.

9. Oidtman RJ, Arevalo P, Bi Q, McGough L, Russo CJ, Cruz DV, et al. Influenza immune escape under heterogeneous host immune histories. Trends in microbiology. 2021 Dec;29(12):1072–82.

10. Centers for Disease Control and Prevention (CDC). National, Regional, and State Level Outpatient Illness and Viral Surveillance. https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html;

11. Smith DJ, Forrest S, Ackley DH, Perelson AlanS. Variable efficacy of repeated annual influenza vaccination. Proceedings of the National Academy of Sciences of the United States of America. 1999 Nov;96(24):14001–6.

12. Skowronski DM, Chambers C, De Serres G, Sabaiduc S, Winter AL, Dickinson JA, et al. Serial Vaccination and the Antigenic Distance Hypothesis: Effects on Influenza Vaccine Effectiveness During A(H3N2) Epidemics in Canada, 2010–2011 to 2014–2015. The Journal of Infectious Diseases. 2017 Apr;215(7):1059–99.

13. Lim WW, Cowling BJ, Nakafero G, Feng S, Nguyen-Van-Tam JS, Bolt H. The impact of repeated vaccination on relative influenza vaccine effectiveness among vaccinated adults in the United Kingdom. Epidemiology and Infection. 2022 Nov;150:e198.

14. Jones-Gray E, Robinson EJ, Kucharski AJ, Fox A, Sullivan SG. Does repeated influenza vaccination attenuate effectiveness? A systematic review and meta-analysis. The Lancet Respiratory Medicine. 2023 Jan;11(1):27–44.

15. Xie H, Wan XF, Ye Z, Plant EP, Zhao Y, Xu Y, et al. H3N2 Mismatch of 2014–15 Northern Hemisphere Influenza Vaccines and Head-to-head Comparison between Human and Ferret Antisera derived Antigenic Maps. Scientific Reports. 2015 Oct;5:15279.

16. Sandbulte MR, Westgeest KB, Gao J, Xu X, Klimov AI, Russell CA, et al. Discordant antigenic drift of neuraminidase and hemagglutinin in H1N1 and H3N2 influenza viruses. Proceedings of the National Academy of Sciences of the United States of America. 2011 Dec;108(51):20748–53.

17. Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature. 2009 Jun;459(7250):1122–5.

18. Morimoto N, Takeishi K. Change in the efficacy of influenza vaccination after repeated inoculation under antigenic mismatch: A systematic review and meta-analysis. Vaccine. 2018 Feb;36(7):949–57.

19. Okoli GN, Racovitan F, Abdulwahid T, Righolt CH, Mahmud SM. Variable seasonal influenza vaccine effectiveness across geographical regions, age groups and levels of vaccine antigenic similarity with circulating virus strains: A systematic review and *Meta*-analysis of the evidence from test-negative design studies after the 2009/10 influenza pandemic. Vaccine. 2021 Feb;39(8):1225–40.

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

21. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, et al. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science (New York, NY). 2009 Oct;326(5953):734–6.

22. Hensley SE. Challenges of selecting seasonal influenza vaccine strains for humans with diverse pre-exposure histories. Current Opinion in Virology. 2014 Oct;8:85–9.

23. Wang Y, Tang CY, Wan XF. Antigenic characterization of influenza and SARS-CoV-2 viruses. Analytical and Bioanalytical Chemistry. 2021-12-14, 2021-12;414(9):2841–81.

24. Parker L, Wharton SA, Martin SR, Cross K, Lin Y, Liu Y, et al. Effects of egg-adaptation on receptor-binding and antigenic properties of recent influenza A (H3N2) vaccine viruses. The Journal of General Virology. 2016 Jun;97(6):1333–44.

25. Gambaryan AS, Robertson JS, Matrosovich MN. Effects of Egg-Adaptation on the Receptor-Binding Properties of Human Influenza A and B Viruses. Virology. 1999 Jun;258(2):232–9.

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

27. Bedford T, Suchard MA, Lemey P, Dudas G, Gregory V, Hay AJ, et al. Integrating influenza antigenic dynamics with molecular evolution. eLife. 2014;3:e01914.

28. Fonville JM, Wilks SH, James SL, Fox A, Ventresca M, Aban M, et al. Antibody landscapes after influenza virus infection or vaccination. Science (New York, NY). 2014 Nov;346(6212):996–1000.

29. Hay JA, Laurie K, White M, Riley S. Characterising antibody kinetics from multiple influenza infection and vaccination events in ferrets. PLoS Computational Biology. 2019 Aug;15(8):e1007294.

30. Sun H, Yang J, Zhang T, Long LP, Jia K, Yang G, et al. Using sequence data to infer the antigenicity of influenza virus. Perlman S, Biron C, editors. mBio. 2013-08-30, 2013-08;4(4).

31. Neher RA, Russell CA, Shraiman BI. Predicting evolution from the shape of genealogical trees. eLife. 2014 Nov;3:e03568.

32. Neher RA, Bedford T, Daniels RS, Russell CA, Shraiman BI. Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. Proceedings of the National Academy of Sciences. 2016-03-07, 2016-03;113(12).

33. Huang L, Li X, Guo P, Yao Y, Liao B, Zhang W, et al. Matrix completion with side information and its applications in predicting the antigenicity of influenza viruses. Bioinformatics. 2017 Oct;33(20):3195–201.

34. Han L, Li L, Wen F, Zhong L, Zhang T, Wan XF. Graph-guided multi-task sparse learning model: A method for identifying antigenic variants of influenza A(H3N2) virus. Bioinformatics. 2019 Jan;35(1):77–87.

35. Li L, Chang D, Han L, Zhang X, Zaia J, Wan XF. Multi-task learning sparse group lasso: A method for quantifying antigenicity of influenza A(H1N1) virus using mutations and variations in glycosylation of Hemagglutinin. BMC Bioinformatics. 2020-05-11, 2020-05;21(1):182.

36. Harvey WT, Davies V, Daniels RS, Whittaker L, Gregory V, Hay AJ, et al. A Bayesian approach to incorporate structural data into the mapping of genotype to antigenic phenotype of influenza A(H3N2) viruses. PLOS Computational Biology. 2023 Mar;19(3):e1010885.

37. Jia Q, Xia Y, Dong F, Li W. MetaFluAD: Meta-learning for predicting antigenic distances among influenza viruses. Briefings in Bioinformatics. 2024 Aug;25(5):bbae395.

38. Wang L, Yao S, Pei Y, Huang A, lefrancq N, Sun K, et al. Using pathogen genomics to predict antigenic changes of influenza H3N2 virus. Research Square; 2024.

39. Cai Z, Zhang T, Wan XF. A computational framework for influenza antigenic cartography. Fraser C, editor. PLoS Computational Biology. 2010-10-07, 2010-10;6(10):e1000949.

40. Barnett JL, Yang J, Cai Z, Zhang T, Wan XF. AntigenMap 3D: An online antigenic cartography resource. Bioinformatics. 2012 May;28(9):1292–3.

41. Arhami O, Rohani P. Topolow: A mapping algorithm for antigenic cross-reactivity and binding affinity assays. Bioinformatics. 2025 Jun;btaf372.

42. Waldock J, Zheng L, Remarque EJ, Civet A, Hu B, Jalloh SL, et al. Assay harmonization and use of biological standards to improve the reproducibility of the hemagglutination inhibition assay: A FLUCOP collaborative study. Pasetti MF, editor. mSphere. 2021-08-25, 2021-08;6(4):10.1128/msphere.00567–21.

43. Zacour M, Ward BJ, Brewer A, Tang P, Boivin G, Li Y, et al. Standardization of hemagglutination inhibition assay for influenza serology allows for high reproducibility between laboratories. Hodinka RL, editor. Clinical and Vaccine Immunology. 2016 Mar;23(3):236–42.

44. Adabor ES, Ndifon W. Bayesian inference of antigenic and non-antigenic variables from haemagglutination inhibition assays for influenza surveillance. Royal Society Open Science. 2018 Jul;5(7):180113.

45. Ndifon W. New methods for analyzing serological data with applications to influenza surveillance. Influenza and Other Respiratory Viruses. 2011-01-05, 2011-01;5(3):206–12.

46. Forghani M, Khachay M. Convolutional Neural Network Based Approach to In Silico Non-Anticipating Prediction of Antigenic Distance for Influenza Virus. Viruses. 2020 Sep;12(9):1019.

47. Einav T, Creanga A, Andrews SF, McDermott AB, Kanekiyo M. Harnessing low dimensionality to visualize the antibody–virus landscape for influenza. Nature Computational Science. 2023 Feb;3(2):164–73.

48. Azulay A, Cohen-Lavi L, Friedman LM, McGargill MA, Hertz T. Mapping antibody footprints using binding profiles. Cell Reports Methods. 2023 Aug;3(8):100566.

49. Catani JPP, Smet A, Ysenbaert T, Vuylsteke M, Bottu G, Mathys J, et al. The antigenic landscape of human influenza N2 neuraminidases from 2009 until 2017. Neher RA, Diamond B, editors. eLife. 2024 May;12:RP90782.

50. Cai Z, Zhang T, Wan XF. Antigenic Distance Measurements for Seasonal Influenza Vaccine Selection. Vaccine. 2011 Nov;30(2):448.

51. Vijaykrishna D, Holmes EC, Joseph U, Fourment M, Su YC, Halpin R, et al. The contrasting phylodynamics of human influenza B viruses. Neher RA, editor. eLife. 2015 Jan;4:e05055.

52. Gupta V, Earl DJ, Deem MW. Quantifying influenza vaccine efficacy and antigenic distance. Vaccine. 2006 May;24(18):3881–8.

53. Pan K, Subieta KC, Deem MW. A novel sequence-based antigenic distance measure for H1N1, with application to vaccine effectiveness and the selection of vaccine strains. Protein Engineering Design and Selection. 2011 Mar;24(3):291–9.

54. Pan Y, Deem MW. Prediction of influenza B vaccine effectiveness from sequence data. Vaccine. 2016 Aug;34(38):4610–7.

55. Anderson CS, McCall PR, Stern HA, Yang H, Topham DJ. Antigenic cartography of H1N1 influenza viruses using sequence-based antigenic distance calculation. BMC bioinformatics. 2018 Feb;19(1):51.

56. Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GCM, Vervaet G, et al. Substitutions Near the Receptor Binding Site Determine Major Antigenic Change During Influenza Virus Evolution. Science. 2013 Nov;342(6161):976–9.

57. Harvey WT, Benton DJ, Gregory V, Hall JPJ, Daniels RS, Bedford T, et al. Identification of Low- and High-Impact Hemagglutinin Amino Acid Substitutions That Drive Antigenic Drift of Influenza A(H1N1) Viruses. PLoS pathogens. 2016 Apr;12(4):e1005526.

58. Mögling R, Richard MJ, Vliet S van der, Beek R van, Schrauwen EJA, Spronken MI, et al. Neuraminidase-mediated haemagglutination of recent human influenza A(H3N2) viruses is determined by arginine 150 flanking the neuraminidase catalytic site. Journal of General Virology. 2017;98(6):1274–81.

59. Castro LA, Bedford T, Ancel Meyers L. Early prediction of antigenic transitions for influenza A/H3N2. Kouyos RD, editor. PLOS Computational Biology. 2020-02-18, 2020-02;16(2):e1007683.

60. Borkenhagen LK, Allen MW, Runstadler JA. Influenza virus genotype to phenotype predictions through machine learning: A systematic review: Computational Prediction of Influenza Phenotype. Emerging Microbes & Infections. 2021 Jan;10(1):1896–907.

61. Forna A, Weedop KB, Damodaran L, Hassell N, Kondor R, Bahl J, et al. Sequence-based detection of emerging antigenically novel influenza A viruses. Proceedings of the Royal Society B: Biological Sciences. 2024 Aug;291(2028):20240790.

62. Lefrancq N, Duret L, Bouchez V, Brisse S, Parkhill J, Salje H. Learning the fitness dynamics of pathogens from phylogenies. Nature. 2025 Jan;637(8046):683–90.

63. Auladell M, Phuong HVM, Mai LTQ, Tseng YY, Carolan L, Wilks S, et al. Influenza virus infection history shapes antibody responses to influenza vaccination. Nature Medicine. 2022 Feb;28(2):363–72.

64. Boyoglu-Barnum S, Ellis D, Gillespie RA, Hutchinson GB, Park YJ, Moin SM, et al. Quadrivalent influenza nanoparticle vaccines induce broad protection. Nature. 2021-03-24, 2021-03;592(7855):623–8.

65. Hinojosa M, Shepard SS, Chung JR, King JP, McLean HQ, Flannery B, et al. Impact of immune priming, vaccination, and infection on influenza A(H3N2) antibody landscapes in children. The Journal of Infectious Diseases. 2020-10-22, 2020-10;

66. Li ZN, Liu F, Gross FL, Kim L, Ferdinands J, Carney P, et al. Antibody landscape analysis following influenza vaccination and natural infection in humans with a high-throughput multiplex influenza antibody detection assay. Subbarao K, editor. mBio. 2021-02-23, 2021-02;12(1).

67. Yang B, Lessler J, Zhu H, Jiang CQ, Read JM, Hay JA, et al. Life course exposures continually shape antibody profiles and risk of seroconversion to influenza. PLOS Pathogens. 2020 Jul;16(7):e1008635.

68. Carlock MA, Ross TM. A computationally optimized broadly reactive hemagglutinin vaccine elicits neutralizing antibodies against influenza B viruses from both lineages. Scientific Reports. 2023 Sep;13:15911.

69. Nachbagauer R, Choi A, Hirsh A, Margine I, Iida S, Barrera A, et al. Defining the antibody cross-reactome directed against the influenza virus surface glycoproteins. Nature Immunology. 2017-02-13, 2017-02;18(4):464–73.

70. Anderson CS, Sangster MY, Yang H, Mariani TJ, Chaudhury S, Topham DJ. Implementing sequence-based antigenic distance calculation into immunological shape space model. BMC Bioinformatics. 2020-06-19, 2020-06;21(1):256.

71. Lee MS, Chen JSE. Predicting antigenic variants of influenza A/H3N2 viruses. Emerging Infectious Diseases. 2004 Aug;10(8):1385–90.

72. Grantham R. Amino Acid Difference Formula to Help Explain Protein Evolution. Science. 1974 Sep;185(4154):862–4.

73. Nuñez IA, Carlock MA, Allen JD, Owino SO, Moehling KK, Nowalk P, et al. Impact of age and pre-existing influenza immune responses in humans receiving split inactivated influenza vaccine on the induction of the breadth of antibodies to influenza A strains. PLOS ONE. 2017 Nov;12(11):e0185666.

74. Carlock MA, Allen JD, Hanley HB, Ross TM. Longitudinal assessment of human antibody binding to hemagglutinin elicited by split-inactivated influenza vaccination over six consecutive seasons. PLOS ONE. 2024 Jun;19(6):e0301157.

75. Abreu RB, Kirchenbaum GA, Sautto GA, Clutter EF, Ross TM. Impaired memory B-cell recall responses in the elderly following recurrent influenza vaccination. Huber VC, editor. PLOS ONE. 2021 Aug;16(8):e0254421.

76. CDC Seasonal Flu Vaccine Effectiveness Studies Flu Vaccines Work CDC. https://web.archive.org/web/20250311230637/https://www.cdc.gov/flu-vaccines-work/php/effectiveness-studies/index.html; 2025.

77. Potter CW, Oxford JS. Determinants of immunity to influenza infection in man. British Medical Bulletin. 1979 Jan;35(1):69–75.

78. Noah DL, Hill H, Hines D, White EL, Wolff MC. Qualification of the Hemagglutination Inhibition Assay in Support of Pandemic Influenza Vaccine Licensure. Clinical and Vaccine Immunology : CVI. 2009 Apr;16(4):558–66.

79. Billings WZ, Ge Y, Knight JH, Hemme H, Hammerton SM, Skarlupka AL, et al. High-Dose Inactivated Influenza Vaccine Inconsistently Improves Heterologous Antibody Responses in an Older Human Cohort. The Journal of Infectious Diseases. 2025 Jan;jiaf003.

80. Hamming RW. Error detecting and error correcting codes. The Bell System Technical Journal. 1950 Apr;29(2):147–60.

81. Wilks S. Racmacs: Antigenic cartography macros. 2024.

82. National Library of Medicine (US), National Center for Biotechnology Information. GenBank. Bethesda, MD; 1982.

83. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Research. 2016 Jan;44(D1):D67–72.

84. The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2025. Nucleic Acids Research. 2025 Jan;53(D1):D609–17.

85. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data – from vision to reality. Eurosurveillance. 2017 Mar;22(13):30494.

86. Elbe S, Buckland-Merrett G. Data, disease and diplomacy: GISAID's innovative contribution to global health. Global Challenges. 2017;1(1):33–46.

87. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin. 1979 Mar;86(2):420–8.

88. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychological Methods. 1996;1(1):30–46.

89. Liljequist D, Elfving B, Skavberg Roaldsen K. Intraclass correlation - A discussion and demonstration of basic features. PloS One. 2019;14(7):e0219854.

90. Spearman C. The proof and measurement of association between two things. The American Journal of Psychology [Internet]. 1904 [cited 2025 May 26];15(1):72–101. Available from: https://www.jstor.org/stable/1412159

91. Rubin DB. The Bayesian Bootstrap. The Annals of Statistics. 1981 Jan;9(1):130–4.

92. McElreath R. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. 2nd ed. Boca Raton: Taylor and Francis, CRC Press; 2020. (CRC Texts in Statistical Science).

93. Gelman A, Hill J. Data analysis using regression and multilevel/hierarchical models. Cambridge ; New York: Cambridge University Press; 2007. (Analytical methods for social research).

94. Breen R. Regression Models: Censored, Sample Selected, or Truncated Data. SAGE Publications, Inc.; 1996.

95. Gelman A, Vehtari A, Simpson D, Margossian CC, Carpenter B, Yao Y, et al. Bayesian Workflow [Internet]. arXiv; 2020 [cited 2025 Apr 15]. Available from: https://arxiv.org/abs/2011.01808

96. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. Stan: A Probabilistic Programming Language. Journal of Statistical Software. 2017 Jan;76(1):1–32.

97. Stan Development Team. Stan Modeling Language Users Guide and Reference Manual. 2024.

98. Bürkner PC. brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software. 2017;80(1):1–28.

99. Bürkner PC. Advanced Bayesian multilevel modeling with the R package brms. The R Journal. 2018;10(1):395–411.

100. Bürkner PC. Bayesian item response modeling in R with brms and Stan. Journal of Statistical Software. 2021;100(5):1–54.

101. Gabry J, Češnovar R, Johnson A, Bronder S. Cmdstanr: R interface to 'CmdStan' [Internet]. 2024. Available from: https://github.com/stan-dev/cmdstanr

102. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2024. Available from: https://www.R-project.org/

103. Arel-Bundock V, Greifer N, Heiss A. How to interpret statistical models using marginaleffects for R and Python. Journal of Statistical Software. 2024;111(9):1–32.

104. Vehtari A, Gelman A, Gabry J. Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing. 2017;27:1413–32.

105. Yao Y, Vehtari A, Simpson D, Gelman A. Using stacking to average bayesian predictive distributions. Bayesian Analysis. 2017;

106. Nakagawa S, Johnson PCD, Schielzeth H. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of The Royal Society Interface. 2017 Sep;14(134):20170213.

107. RStudio Team. RStudio: Integrated development environment for R. Boston, MA: RStudio, PBC.; 2024.

108. Landau WM. The targets r package: A dynamic make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. Journal of Open Source Software [Internet]. 2021;6(57):2959. Available from: https://doi.org/10.21105/joss.02959

109. Müller K. Here: A simpler way to find your files [Internet]. 2020. Available from: https://here.r-lib.org/

110. Ushey K, Wickham H. Renv: Project environments [Internet]. 2025. Available from: https://rstudio.github.io/renv/

111. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the tidyverse. Journal of Open Source Software. 2019;4(43):1686.

112. Kay M. tidybayes: Tidy data and geoms for Bayesian models [Internet]. 2024. Available from: http://mjskay.github.io/tidybayes/

113. Kay M. ggdist: Visualizations of distributions and uncertainty in the grammar of graphics. IEEE Transactions on Visualization and Computer Graphics. 2024;30(1):414–24.

114. Kay M. ggdist: Visualizations of distributions and uncertainty [Internet]. 2024. Available from: https://mjskay.github.io/ggdist/

115. Bååth R. Bayesboot: An implementation of rubin's (1981) bayesian bootstrap [Internet]. 2025. Available from: https://github.com/rasmusab/bayesboot

116. Vehtari A, Gabry J, Magnusson M, Yao Y, Bürkner PC, Paananen T, et al. Loo: Efficient leave-one-out cross-validation and WAIC for bayesian models [Internet]. 2024. Available from: https://mc-stan.org/loo/

117. Wickham H. ggplot2: Elegant graphics for data analysis [Internet]. Springer-Verlag New York; 2016. Available from: https://ggplot2.tidyverse.org

118. Schloerke B, Cook D, Larmarange J, Briatte F, Marbach M, Thoen E, et al. GGally: Extension to 'ggplot2' [Internet]. 2024. Available from: https://ggobi.github.io/ggally/

119. Sjoberg DD, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible summary tables with the gtsummary package. The R Journal [Internet]. 2021;13:570–80. Available from: https://doi.org/10.32614/RJ-2021-053

120. Gohel D, Skintzos P. Flextable: Functions for tabular reporting [Internet]. 2024. Available from: https://ardata-fr.github.io/flextable-book/

121. Allaire J, Dervieux C. Quarto: R interface to 'quarto' markdown publishing system [Internet]. 2024. Available from: https://github.com/quarto-dev/quarto-r

122. Xie Y. Knitr: A general-purpose package for dynamic report generation in r [Internet]. 2024. Available from: https://yihui.org/knitr/

123. Xie Y. Dynamic documents with R and knitr [Internet]. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC; 2015. Available from: https://yihui.org/knitr/

124. Xie Y. Knitr: A comprehensive tool for reproducible research in R. In: Stodden V, Leisch F, Peng RD, editors. Implementing reproducible computational research. Chapman; Hall/CRC; 2014.

125. Arel-Bundock V, McCrain J. Software citations in political science. PS: Political Science & Politics. 2023;April:1–4.

126. Bedford T, Riley S, Barr IG, Broor S, Chadha M, Cox NJ, et al. Global circulation patterns of seasonal influenza viruses vary with antigenic drift. Nature. 2015-06-08, 2015-06;523(7559):217–20.

127. Cobey S. Vaccination against rapidly evolving pathogens and the entanglements of memory. Nature Immunology. 2024 Nov;25(11):2015–23.

128. Hilleman MR. ANTIBODY RESPONSE IN VOLUNTEERS TO ASIAN INFLUENZA VACCINE. Journal of the American Medical Association. 1958 Mar;166(10):1134.

129. Couch RB, Winokur P, Brady R, Belshe R, Chen WH, Cate TR, et al. Safety and immunogenicity of a high dosage trivalent influenza vaccine among elderly subjects. Vaccine. 2007 Nov;25(44):7656–63.

130. Angeletti D, Yewdell JW. Understanding and Manipulating Viral Immunity: Antibody Immunodominance Enters Center Stage. Trends in Immunology. 2018 Jul;39(7):549–61.

131. Einav T, Cleary B. Extrapolating missing antibody-virus measurements across serological studies. Cell Systems. 2022 Jul;13(7):561–573.e5.

132. Stacey H, Carlock MA, Allen JD, Hanley HB, Crotty S, Ross TM, et al. Leveraging pre-vaccination antibody titres across multiple influenza H3N2 variants to forecast the post-vaccination response. EBioMedicine. 2025 May;116:105744.