

BREADTH QUANTIFICATION OF THE HUMAN ANTIBODY RESPONSE TO INFLUENZA VACCINATION

by

AMANDA LEE SKARLUPKA

(Under the Direction of Andreas Handel)

ABSTRACT

Influenza viruses are endemic in human and animal populations and have caused numerous pandemics. The diversity of the circulating human viruses includes four subtypes (H1N1, H3N2, Yamagata-lineage, Victoria-lineage). Within the subtypes, there are multiple antigenic clusters. Influenza vaccine initiatives push for broadly protective vaccine development. However, currently, there is no consensus on the definition of the breadth of response. Different methods are used, including both qualitative and quantitative methods. Here we introduced influenza distance measures (time-, sequence-, and cartographic-based) to quantify the change in distance and the change in HAI titer in a cohort of individuals who received the standard dose influenza vaccine. Further, the area under the curve was quantified to measure breadth. A proof-of-concept compared standard and high dose vaccination responses in an elderly cohort. Overall, a consistent breadth measure allows for quantitative antibody, cellular, and antiviral activity descriptions.

INDEX WORDS: Influenza Antibody Response; Antigenic Distance; Breadth Quantification

BREADTH QUANTIFICATION OF THE HUMAN ANTIBODY RESPONSE TO INFLUENZA
VACCINATION

by

AMANDA LEE SKARLUPKA

B.S., University of Wisconsin - Madison, 2014

A Thesis Submitted to the Graduate Faculty of the
University of Georgia in Partial Fulfillment of the Requirements for the Degree.

MASTER OF SCIENCE

ATHENS, GEORGIA

2022

©2022

Amanda Lee Skarlupka
All Rights Reserved

BREADTH QUANTIFICATION OF THE HUMAN ANTIBODY RESPONSE TO INFLUENZA
VACCINATION

by

AMANDA LEE SKARLUPKA

Major Professor: Andreas Handel

Committee: Ye Shen
Ted M. Ross

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2022

CONTENTS

List of Figures	v
List of Tables	vii
1 Introduction	i
1.1 Background	i
1.2 Influenza vaccines	2
1.3 Breadth of a vaccine response	3
1.4 Distance measures	4
1.5 Weighting	9
1.6 Purpose of this study	9
2 Data description and viral pair-wise distance calculations	ii
2.1 Analysis tools and cleaning	ii
2.2 Cohort information	13
2.3 Influenza viruses	16
2.4 Pair-wise distance measures	16
2.5 Normalization of measures	21
2.6 Comparison of measures	22
2.7 Linear regression, area calculations, and weighting	22
3 Antigenic cartography using pre-immune human sera	26
3.1 Dataset and number of dimensions selection	27
3.2 Comparison to previous studies	30
3.3 Conclusion	37
4 Application of distance measures to vaccine responses	38
4.1 Pre- and post-vaccination titers	38
4.2 Titer Increase	40
4.3 Quantification of breadth and application of weighting	44
4.4 Case-study	46

4.5 Limitations	52
5 Discussion	53
Bibliography	55
Appendices	62
A Type B investigation	62
B Supplemental Figures	64

LIST OF FIGURES

1.1	Example antigenic landscapes are created using the year of isolation	5
1.2	Schematic of antigenic cartography from HAI results	9
2.1	Schematic of FluZone Cohorts	12
2.2	Comparison of the reordering of influenza H ₁ N ₁ strains by the normalized distances . .	24
2.3	Comparison of the reordering of influenza H ₃ N ₂ strains by the normalized distances .	25
3.1	Validation of two-dimensional antigenic cartography with human sera	29
3.2	Procrustes maps of different datasets compared to the original H ₃ N ₂ cartography.	31
3.3	Correlations of the relative euclidean distances from A/H ₃ N ₂ /Panama/2007/1999	33
3.4	Relative distances to the vaccine strains for each season for pre- and post-vaccination standard dose sera datasets of the H ₃ N ₂ viruses.	34
3.5	Relative distances to the vaccine strains for each season for pre- and post-vaccination standard dose sera datasets of the H ₁ N ₁ viruses.	35
3.6	Antigenic cartographies were created using pre-immune human sera for the H ₁ N ₁ and H ₃ N ₂ subtypes	36
4.1	Pre- and post-vaccination titers to the H ₁ N ₁ and H ₃ N ₂ virus panels by season of all individuals who received SD vaccination for the 2017 season.	39
4.2	The titer increase to the H ₁ N ₁ and H ₃ N ₂ virus panels for the 2017 season of all individuals who received SD vaccination	41
4.3	The titer increase to the H ₁ N ₁ strains for all individuals who received SD vaccination . .	42
4.4	The titer increase to the H ₃ N ₂ strains for all individuals who received SD vaccination .	43
4.5	Examples of the weighting schemes applied to data	46
4.6	The titer increase to the H ₁ N ₁ and H ₃ N ₂ virus panels for the 2017 season of all individuals greater than or equal to 65 years of age	47
4.7	The titer increase to the H ₁ N ₁ strains for HD and SD recipients	49
4.8	The titer increase to the H ₃ N ₂ strains for HD and SD recipients	50
A.1	Cross-reactivity between Type B influenza lineages	63
B.1	Pre- and post- vaccination titers to the H ₁ N ₁ and H ₃ N ₂ virus panels by season of all individuals who received SD vaccination for the 2014-2016 and 2018-2019 seasons.	65

B.2	The titer increase to the H ₁ N ₁ and H ₃ N ₂ virus panels for the 2014-2016 and 2018-2019 seasons of all individuals who received SD vaccination	66
B.3	The titer increase (TI) to the H ₁ N ₁ and H ₃ N ₂ virus panels for the 2014-2016 and 2018-2019 season of all individuals greater than or equal to 65 years of age	67

LIST OF TABLES

1.1	Antigenic distance calculation using experimental HAI values and the one-sided equation.	8
2.1	Cohort demographics for individuals who received the SD dose by season.	14
2.2	Cohort demographics for individuals of equal to or greater than 65 years of age, stratified by dose and season.	15
2.3	Influenza virus strain components within the Fluzone inactivated virus vaccines.	15
2.4	Strain names and source of HA amino acid sequence used for analysis	17
2.5	Amino acid residues used for the sequence-based distance measure, p-epitope.	19
2.6	Maximum normalized distances when grouping by season or strain	23
3.1	Dimensional analysis of H ₁ N ₁ and H ₃ N ₂ human cartography maps pre- and post-vaccination using all sera available (all) or only sera from SD individuals (SD) and comparison of the mean root mean square error (RMSE) and variance (Var).	28
3.2	Procrustes analysis RMSD values under different alignment conditions	32
3.3	Correlation coefficients of pair-wise cartographic distances when varying the relative viruses	33
3.4	Number of viruses and sera samples included in each antigenic cartography	36
4.1	Area under the curve (AUC) of the 2017 season with different weighting schemes applied	44
4.2	Area under the curve (AUC) of different distance measures by influenza vaccine strain with different weighting schemes applied.	45
4.3	Area under the curve for three weighting schemes for 65+ individuals for the 2017 season	48
4.4	Area under the curve for three weighting schemes for 65+ individuals for different strains	51

CHAPTER I

INTRODUCTION

1.1 Background

Influenza viruses have been responsible for seasonal epidemics of influenza and four global pandemics since 1918. The virus infects nearly 1 billion individuals annually, with 290,000 to 650,000 cases resulting in death. Endemically, two species of influenza virus seasonally infect humans, Type A and Type B. Only Type A has proven pandemic potential due to its copious zoonotic reservoirs and all previously known pandemics being of Type A origin. Most viral neutralizing antibodies target the surface proteins, hemagglutinin (HA), and neuraminidase (NA).

Over time, the antibodies elicited to one influenza virus strain become inadequate to current circulating strains due to *antigenic drift*. Host antibodies bind to antigenic sites located on the surface proteins, and these antigenic sites mutate through selective pressure from the host and a high viral RNA polymerase mutation rate. Hence, antibodies that once bound to and neutralized the influenza virus before the antigenic drift are no longer highly protective to the newly circulating antigenically drifted strain.

Further, reassortment between two or more Type A influenza viruses results in progeny with distinct gene constellations due to the eight RNA genome segments. These reassorted viruses may have unique surface protein combinations. In 1968, the H₃N₂ pandemic was the result of reassortment of the H₂N₂ human seasonal virus and an H₃NX avian virus (Scholtissek et al., 1978; Valkenburg & Poon, 2022). This

drastic change in surface protein composition resulted in most antibodies directed to H₂ protein being ineffective. This phenomenon is termed an *antigenic shift*. Antigenic shifts can also occur within protein subtypes. For example, the 2009 H1N1 pandemic resulted from an antigenic shift from the human seasonal lineage H1 HA to a swine lineage H1 HA. Antigenic shift and drift allow evolutionary drifted viruses and new reassortments to infect hosts even though those hosts had previous influenza virus exposure.

The antibodies that inhibit the HA activity are the primary correlate of protection in humans for seasonal Type A influenza. Infection elicits these HA-specific antibodies, and fortunately, vaccination does as well. Vaccination remains the primary preventative method of the disease. Influenza vaccines are most commonly split-inactivated viruses that contain four components: two Type A strains – one each of an H1N1 and H3N2 – and two Type B – one each of the Victoria-lineage and the Yamagata-lineage. The World Health Organization (WHO) and other governmental agencies re-evaluate and update the vaccine strains annually depending on factors, including surveillance data and antigenic distance between the vaccine strains and current circulating strains. Antigenic characterization assesses how similar two viruses are to antibody-mediated inhibition. When the antigenic distances of the viruses are close, two antigenic distance units (AU) (4-fold hemagglutinin inhibition (HAI) titer difference), the data suggests no vaccine change. However, three or greater (8-fold HAI titer difference) suggests that antibodies elicited to the vaccine strain do not adequately inhibit the HA receptor-binding of the circulating strain, and data support a vaccine strain change.

1.2 Influenza vaccines

The split-inactivated virus vaccines elicit strain-specific neutralizing antibodies to the HA and NA (Dong et al., 2018). In addition to the time necessary to survey the circulating strains, create anti-sera in ferrets, calculate antigenic distance, and propose a vaccine strain change for next season, the new vaccine strain then needs to be optimized for growth for vaccine production. Overall, it takes time and resources to change a vaccine strain.

Therefore, influenza vaccine initiatives have focused on developing broadly protective candidates that elicit protection across subtypes and ideally across all Type A influenza (Wei et al., 2020). Multiple candidates are in development and evaluation for protective responses (Vogel & Manicassamy, 2020). Different assays for detecting and quantifying the vaccine response are proposed and utilized based upon the mechanism of action and the research group's interest (Krammer et al., 2019). However, HA inhibiting antibodies (HAI) remain the main biomarker associated with protection. Thus, most commonly, the magnitude of the vaccine response is measured by the titer increase of the serum HAI antibodies. An HAI titer of 1:40 correlated with 50% protection from clinical disease for Type A human seasonal influenza (Hobson et al., 1972). Yet, the HAI results for the recent H₃N₂ isolates are unreliable due to difficulties in conducting the HAI assay with these newer strains from a combination of increased NA mediated agglutination and decreased HA agglutination (Auladell et al., 2022; Mögling et al., 2017).

Scientists assess vaccine candidates for efficacy in pre-clinical animal models and human clinical trials. From these studies, different assays measure the magnitude and functionality of the antibody response (Krammer et al., 2019). The viral breadth of these responses is qualitatively analyzed. The broadly reactive responses are compared to those of a strain-specific vaccine to conclude that the new candidate has a broader profile to a select panel of viruses.

1.3 Breadth of a vaccine response

The viral breadth panel is commonly quantified by counting the positive responses. If the vaccine candidate has inhibited more viruses than the comparator, it is seen as more broadly reactive (Allen & Ross, 2021; Andrews et al., 2022; Dugan et al., 2020; Hinojosa et al., 2020; Jang & Ross, 2021; G.-M. Li et al., 2012; Z.-N. Li et al., 2021; Nachbagauer et al., 2014). This method remains the easiest to conduct. The selection of a panel by the researchers opens the door for bias in testing. Virus selection bias includes over-representation of easily accessible viruses or viruses from the same antigenic group, regional differences in virus strains, and subconscious or conscious selection for viruses that provide a desired positive response. Further, when comparing multiple vaccine candidates from different research groups, the breadth pan-

els may not overlap. Therefore, before determining if a vaccine candidate is more broadly reactive than another, the method for calculating the breadth of a vaccine response needs to be determined.

The correlate of protection assay used for determining the breadth of an antibody response is most commonly the hemagglutinin inhibition assay (HAI). The HA protein of influenza agglutinates red blood cells by binding sialic acid with the HA's receptor-binding site (RBS). This mechanism quantifies the amount of anti-RBS HA antibodies in a sample. Blood samples are separated into serum and clotted red blood cells. The serum contains the individual's circulating peripheral antibodies. Sera are diluted in a two-fold dilution starting at 1:10 (1:10, 1:20, 1:40), up to 1:10,240 or 1:20,480. The diluted serum incubates with virus and red blood cells. After, results are interpreted by eye, and the last dilution that inhibits the agglutination of the virus and red blood cells is that individual's HAI titer to the tested virus.

The HAI is difficult to interpret quantitatively. The combination of the subjectivity of interpreting the results, the measurement of an indirect relationship between the virus and antibody, and the variability of the viral binding affinity between viruses and red blood cells (Adabor & Ndifon, 2018; Cai et al., 2010; Ndifon, 2011). These different variables make the HAI results unreliable when comparing minor differences in HAI titers with a geometric coefficient of variation ranging between 0-117% for results between other labs (Waldock et al., 2021; Zacour et al., 2016). Although the same laboratory conducted all HAIs for this study, the assays were performed over six years with different personnel and different reagents. Further, due to the assay only quantifying antibodies that bind to the RBS, the assay may not detect antigenic changes in other locations of the HA protein (Ndifon, 2011).

1.4 Distance measures

Sera breadth measurements mainly entail conducting a correlation of protection assay, such as the HAI assay, against a panel of viruses. The proportion of strains that had a positive result of different serum samples are compared together; the one with a greater proportion is classified as having more breadth. Antibody landscapes have addressed limitations in categorizing viruses. Antibody landscapes assign a distance between viruses in the panel, thus creating a continuous value for further quantification.

I.4.1 Time-based

Researchers have incorporated the year it was isolated from the human population (year of isolation) into antigenic landscapes. The simplest technique places each virus an equal distance from each other and orders them in chronological order (Figure I.4.1A) (Boyoglu-Barnum et al., 2021; Hinojosa et al., 2020; Z.-N. Li et al., 2021). The other assigns the distance based upon year of isolation (Figure I.4.1) (Auladell et al., 2022; B. Yang et al., 2020). Depending upon the viruses in the panel, the results of these two methods can vary greatly. Further, assigning an equal distance between each year increment assumes that genetic or antigenic evolution occurs at a constant rate over time. However, different antigenic clusters arise sporadically every 2-5 years for H₃N₂ and 3-8 years for H₁N₁ and Type B (Bedford et al., 2015; Castro et al., 2020; Petrova & Russell, 2017). The analysis is complicated further since multiple antigenic clusters are isolated within the same year (Castro et al., 2020). But with this method, they would be assigned a distance of zero.

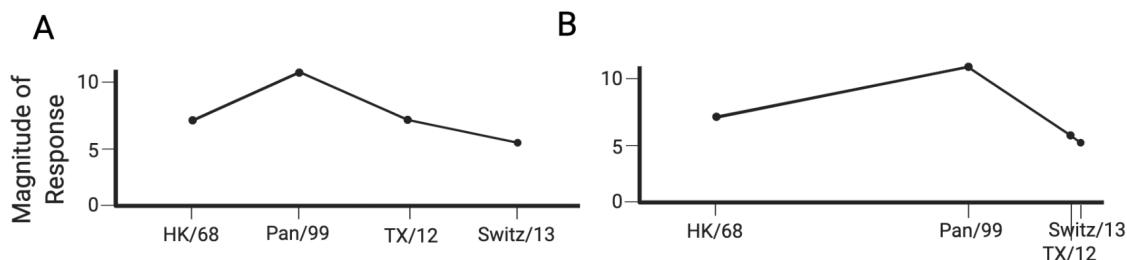


Figure I.1: Examples of how antigenic landscapes are created using the year of isolation. A) Representative antigenic landscape where the influenza viruses are treated as a continuous variable. Each virus is equidistant to the neighboring virus. They are ordered chronologically. B) Representative antigenic landscape where each distance between the virus is the year it was isolated. Created with Biorender.com

I.4.2 Genetic and amino acid sequence-based

Sequence-based methods for determining distance between viruses have been used consistently in bioinformatics and phylogenetic analyses. The amino acid sequence is usually preferred over nucleic acid sequence for comparing distances for antibody responses due to multiple codons encoding the same amino acid.

Thus although a difference may be observed in the nucleic acids, the amino acid sequence remains unchanged, and there will be no difference in antibody recognition.

The amino acid differences have been determined using the full-length HA protein (Nachbagauer et al., 2017), the head region (Smith et al., 2004), and specific antigenic sites (Anderson et al., 2018; Gupta et al., 2006). The portion of the HA used for amino acid distances changes based on where the antibodies bind. Nachbagauer et al., 2017 was investigating anti-HA antibody titers as measured through an enzyme-linked immunosorbent assay (ELISA). As such, the research question included stem binding antibodies. However, other outcomes, such as HAI titers, quantify only antibodies that bind to specific antigenic sites located on the head of the HA. Therefore, other techniques have removed the highly conserved or internal amino acid residues from the distance comparison (Anderson et al., 2018; Anderson et al., 2020; Gupta et al., 2006; Lee & Chen, 2004; Sun et al., 2013). After defining the residues of interest, the Hamming distance is most commonly used for determining the distance (Anderson et al., 2018; Anderson et al., 2020; Gupta et al., 2006; Lee & Chen, 2004). Briefly, the Hamming distance is the sum of the number of differences when comparing two strains. A difference is assigned a value of 1, and a matching residue is 0. This method does not take into account physicochemical similarities between amino acids. Sun et al., 2013 combined the genetic distance matrix with experimental antigenic distance matrix data to develop a method that weights each residue depending on its influence on antigenicity, and Neher et al., 2016 combined the two data types to assist with predicting antigenic properties of newly isolated viruses (please see Y. Wang et al., 2021 for an extensive recent review on antigenic characterization). Models have begun to include more features of the protein, such as physicochemical properties or glycosylation residues, which, when present, block antibody binding (Forghani & Khachay, 2020; L. Li et al., 2020).

One of the benefits of sequence-based distances is that they can be determined with relatively little resources. Surveillance sequences are readily available in influenza databases. However, the genetic cluster of a virus does not always match its antigenic cluster (Durviaux et al., 2014; Smith et al., 2004). Mutation of even only one amino acid has been implicated in changing the antigenic profiles between the two very similar proteins. Sequence-based methods are being used to develop prediction models to determine if a

new viral strain will be antigenically distinct. However, the most accurate models require HAI antigenic distance data for training (Forghani & Khachay, 2020; L. Li et al., 2020; Neher et al., 2016; J. Yang et al., 2014).

I.4.3 Antigenic distance-based

Experimental antigenic distance

HAI results are the gold standard for measuring the antigenic distance between two strains. The experimental pair-wise antigenic distance is determined by creating anti-sera specific to viruses A and B. The HAI inhibition to each virus with each anti-sera is measured and creates a matrix (Table I.4.3). The Archetti-Horsfall measure was introduced in 1950 and requires all four HAI titers HAI_{AA} , HAI_{BB} , HAI_{AB} , and HAI_{BA} if $HAI_{Virus, Antisera}$ (Equation I.1). The inclusion of both homologous titers makes the value less dependent on non-antigenic factors associated with the HAI assay, such as viral avidity for red blood cells (Archetti & Horsfall, 1950; Lee & Chen, 2004; Ndifon et al., 2009).

$$distance = \sqrt{\frac{H_{AA}H_{BB}}{H_{BA}H_{AB}}} \quad (I.1)$$

The Archetti-Horsfall equation requires both homologous and heterologous titers for each virus. The number of resources needed for anti-sera generation, virus propagation, and HAI conduction is too significant to justify the creation of a full matrix. Therefore, the one-side antigenic distance is the most used (Equation I.2). The one-sided measure allows sera generated to a limited set of viruses to be tested against a larger HAI panel. The experimental antigenic distance was initially used for vaccine strain selection to determine if the current circulating viruses had drifted too far from the vaccine strain.

$$distance = \sqrt{\frac{H_{AA}}{H_{BA}}} \quad (I.2)$$

In addition to vaccine selection, Hay et al., 2019 used experimental antigenic distances for antigenic landscapes in ferrets with H₃N₂ and H₁N₁ strains. The antigenic distance hypothesis also relies on the

Table 1.1: Antigenic distance calculation using experimental HAI values and the one-sided equation.

HAI Virus	HAI Titer		Calculation		Antigenic Distance	
	Serum A	Serum B	Serum A	Serum B	Serum A	Serum B
Virus A	2560	320	$\log_2(\frac{2560}{2560})$	$\log_2(\frac{5120}{320})$	0	4
Virus B	40	5120	$\log_2(\frac{2560}{40})$	$\log_2(\frac{5120}{5120})$	6	0

empirical measure (Skowronski et al., 2017). Briefly, the antigenic distance hypothesis that if vaccinated with a strain that is too closely related (≤ 2 antigenic distance units), then there will be negative interference between the previous vaccine and the new vaccine, and vaccination with this new strain is not suggested (Belongia et al., 2017).

Antigenic cartography distance

Antigenic cartography utilizes the experimental antigenic distance (Equation 1.2). The seminal study by Smith et al., 2004 transformed the HAI titer matrix into an antigenic map, which placed the viruses and anti-sera onto a two-dimensional map (Figure 1.4.3). The dimension reduction used metric multidimensional scaling (MDS), allowing the experimental antigenic distance information to be retained. The difference between the cartographic Euclidean distances between viruses and sera and experimental antigenic distances are minimized during metric MDS.

The original cartographies were created with ferret sera specific to only one virus (Smith et al., 2004). Fonville et al., 2015 then compared the Smith et al., 2004 to cartographies created with human sera collected from children before multiple exposures to influenza had occurred. The positioning of the viruses correlated between the different serum origins. In addition, antigenic cartography has been used for H3N2, swine H1, and neuraminidase of influenza (Fonville et al., 2014; Hinojosa et al., 2020; Kucharski et al., 2018; J. Wang et al., 2021). Antigenic cartography has been applied to influenza, norovirus (Kendra et al., 2021), dengue virus (Katzelnick et al., 2021), and *Plasmodium falciparum* membrane protein (Tuju et al., 2019).

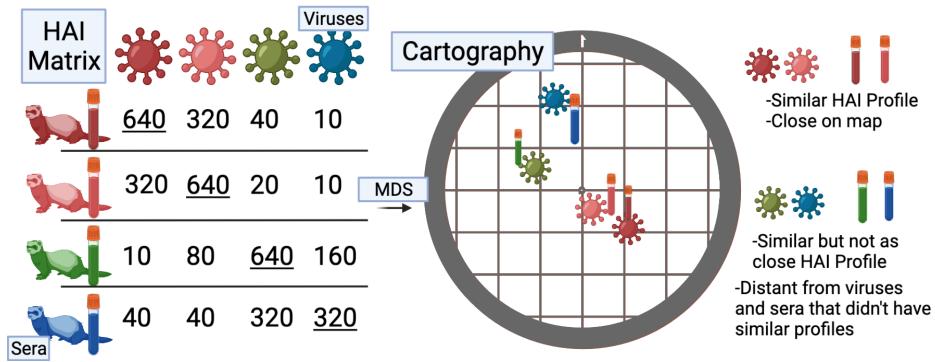


Figure 1.2: Schematic of antigenic cartography from HAI results. An HAI matrix is created using anti-sera (rows) and viruses (columns). The matching sera and virus are underlined and color-matched. After creating the HAI matrix, multidimensional scaling reduces the data to two dimensions. The HAI fold change and the units of the antigenic cartography grid are linearly related so that one antigenic unit is one 2-fold change in HAI titer. Created with Biorender.com

1.5 Weighting

Weighting of the overall antibody response can be applied for specific research questions. For instance, multiple H₂N₂ HA vaccine candidates developed in Reneer et al., 2020; Reneer et al., 2021, produced a broad response against the panel of viruses. However, by introducing a weighting scheme to the results, no longer circulating viruses can be weighted lower than the current circulating viruses that individuals are more at danger of being infected with. This provides an opportunity for fine-tuning the selection based upon the outcome goal of the vaccine. It has also been noted that if the vaccine strain is too close to the previous vaccine strain, then there is negative interference. Therefore, the application of negatively weighting vaccines that too similar is applicable (Skowronski et al., 2017).

1.6 Purpose of this study

Although the goal remains to be developing a broadly reactive influenza vaccine, there is no consensus on the definition of breadth. With an objective, reproducible method for quantifying breadth, researchers

and officials will be able to select an appropriate candidate for future use. Further, the breadth measure allows for more quantitative descriptions of the breadth of antibody, cellular, and antiviral activity.

Therefore, this study determined which distance measures between viruses were appropriate, looking at H₁N₁ and H₃N₂ subtypes. A human cohort's antibody responses to the influenza FluZone vaccine were analyzed to determine the proper virus distance measure. Antigenic landscapes were created using time-, sequence-, and cartographic-based distance measures and change in HAI titer. The breadth response of the vaccine was quantified. After which, we propose different weighting schemes that could be applied to focus on the desired breadth outcome.

CHAPTER 2

DATA DESCRIPTION AND VIRAL PAIR-WISE DISTANCE CALCULATIONS

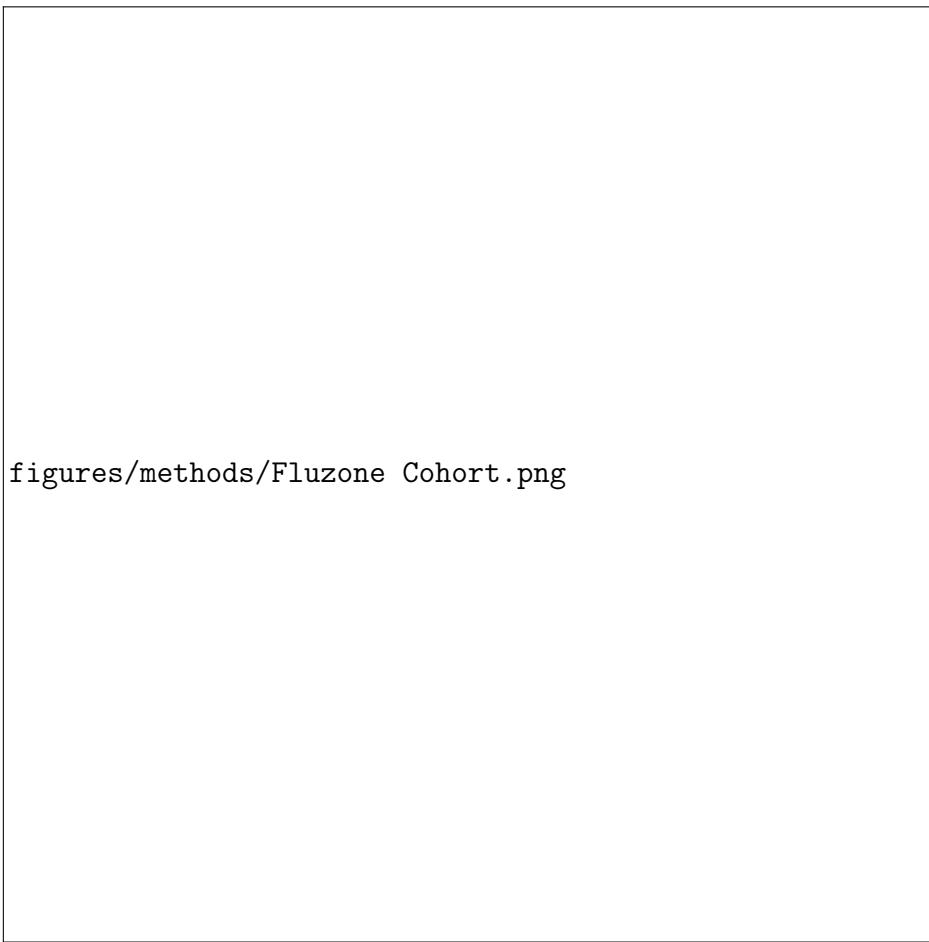
2.1 Analysis tools and cleaning

Data cleaning and analysis were conducted with R version 4.1.2 (2021-11-01) and Rstudio version Prairie Trillium 2022.2.0.443 (R Core Team, 2021; RStudio Team, 2022). Individuals who did not have both pre- and post-vaccination titers were removed during cleaning. Further, one individual from the 2014 season was removed from the SD/HD comparisons because they received the HD vaccine with an age of 61 (participant id: pa2014_id_1218). This was due to the analysis cutoff being individuals of age 65 years or greater. Each individual's sera were tested in an HAI panel of H1N1 and H3N2 viruses pre- and post-vaccination (Do and D21-D28) (Figure 2.1). The reciprocal HAI titers were transformed before analysis (Equation 2.1).

$$\text{Transformed Titer} = \log_2\left(\frac{\text{titer}}{5}\right) \quad (2.1)$$

After transformation, the titer increase was determined by subtracting the pre-vaccination titer from the post-vaccination titer for each strain in the virus panel (Equation 2.2).

$$\text{Titer Increase} = \text{Transformed Titer}_{post} - \text{Transformed Titer}_{pre} \quad (2.2)$$



figures/methods/Fluzone Cohort.png

Figure 2.1: Individuals were enrolled starting in September at the start of the Northern Hemispheres influenza season. Individuals greater than or equal to 65 years of age were given the option of either the standard dose (SD) or high dose (HD) Fluzone inactivated influenza vaccine. All individuals less than 65 years received the SD formulation. Sera was collected on the day of vaccination and then between 21 and 28 days after vaccination. Pre- and post-vaccination sera was then tested in H1N1 and H3N2 influenza virus HAI panels.

2.2 Cohort information

The CIVIC cohort data comprised volunteers enrolled from September 2014 to 2019 for the Northern Hemisphere influenza season. From 2014 to 2016, the individuals were enrolled from Pittsburgh, PA, and Stuart, FL. Individuals between the ages of 18 and 85 years of age. From 2017 to 2019, the study site was moved to Athens, GA, and individuals between the ages of 11 and 85 were enrolled. If possible, individuals were given the ability to re-enroll for the following season leading to involvement in multiple seasons ('number of year enrolled' in Tables 2.1 and 2.2).

All individuals less than 65 years of age received the recommended influenza vaccine composition in the standard dose (SD; 15 μ g of each viral HA protein) formulation. Individuals greater than or equal to 65 years of age were given a choice between SD and high dose (HD; 60 μ g of each viral HA protein) formulations. The intramuscular vaccines administered were FluzoneTM (Sanofi Pasteur, Swiftwater, PA, USA) inactivated influenza trivalent (2014 SD and HD; 2015-2019 HD) or quadrivalent (2015-2019 SD) formulations. Exclusion criteria for individuals included: receipt of influenza vaccine for the specific season, previous history of Guillain-Barré, dementia, Alzheimer's disease, immunocompromised status, allergies to egg or egg products, concurrent enrollment in another influenza vaccine study, or less than two years of life expectancy. Blood was collected from individuals at day 0 pre-vaccination and 21-28 days post-vaccination. Sera was extracted from blood samples and stored at -150°C until use in the hemagglutinin inhibition (HAI) assay.

The data was divided into two datasets. The individuals who received only the standard dose vaccination (Table 2.1) and the individuals who are equal to or greater than 65 years of age separated by the dose received (Table 2.2). The recommended influenza vaccine compositions changed over the seasons dependent on the circulating strains at the time and the dose (Table 2.3). In 2014, both doses were trivalent formulations.

Table 2.1: Cohort demographics for individuals who received the SD dose by season.

Category	Season	2014	2015	2016	2017	2018	2019
	Total Individuals	238	209	328	255	242	392
	Median Age (Years)	57	56	44	25	16	36.5
	Age Range	20-83	21-84	18-82	12-83	12-75	11-85
Sex	Female	172 (72.3%)	156 (74.6%)	221 (67.4%)	142 (55.7%)	137 (56.6%)	232 (59.2%)
	Male	66 (27.7%)	53 (25.4%)	107 (32.6%)	113 (44.3%)	105 (43.4%)	160 (40.8%)
Race	Black	44 (18.5%)	38 (18.2%)	46 (13.7%)	19 (7.5%)	11 (4.5%)	33 (8.4%)
	Hispanic	27 (11.3%)	17 (8.1%)	24 (7.3%)	16 (6.3%)	11 (4.5%)	13 (3.3%)
	White	158 (66.4%)	147 (70.3%)	235 (71.6%)	205 (80.4%)	203 (83.9%)	324 (82.7%)
	Not Available	9 (3.8%)	7 (3.3%)	10 (3%)			
	Other			14 (4.3%)	15 (5.9%)	17 (7%)	22 (5.6%)
# Years Enrolled	1	238 (100%)	42 (19.8%)	194 (57.9%)	157 (61.6%)	100 (41.3%)	245 (62.5%)
	2		170 (80.2%)	21 (6.3%)	98 (38.4%)	73 (30.2%)	56 (14.3%)
	3			120 (35.8%)		69 (28.5%)	52 (13.3%)
	4					39 (9.9%)	

Percentages are by column

Table 2.2: Cohort demographics for individuals of equal to or greater than 65 years of age, stratified by dose and season.

Season		2014		2015		2016		2017		2018		2019	
Category	Dose	SD	HD	SD	HD	SD	HD	SD	HD	SD	HD	SD	HD
	Total Individuals	54	39	40	58	34	77	22	16	11	8	21	69
	Median Age (Years)	69	71	67	70	68	70	67	69	68	70	69	70
	Age Range	65-83	65-80	65-84	66-81	65-82	65-85	65-83	66-83	65-75	68-80	65-85	65-81
Sex	Female	37 (68.5%)	30 (76.9%)	26 (65%)	44 (75.9%)	22 (64.7%)	54 (70.1%)	10 (45.5%)	6 (37.5%)	4 (36.4%)	4 (50%)	10 (47.6%)	42 (60.9%)
	Male	17 (31.5%)	9 (23.1%)	14 (35%)	14 (24.1%)	12 (35.3%)	23 (29.9%)	12 (54.5%)	10 (62.5%)	7 (63.6%)	4 (50%)	11 (52.4%)	27 (39.1%)
Race	Black	10 (18.5%)	18 (46.2%)	11 (27.5%)	16 (27.6%)	5 (14.7%)	20 (26%)	3 (13.6%)		1 (9.1%)		1 (4.8%)	1 (1.4%)
	Hispanic	1 (1.9%)	1 (2.6%)	1 (2.5%)	1 (1.7%)		1 (1.3%)	1 (4.5%)				1 (4.8%)	2 (2.9%)
	White	42 (77.8%)	20 (51.3%)	26 (65%)	41 (70.7%)	28 (82.4%)	55 (71.4%)	18 (81.8%)	16 (100%)	9 (81.8%)	8 (100%)	18 (85.7%)	65 (94.2%)
	Not Available	1 (1.9%)		2 (5%)		1 (2.9%)	1 (1.3%)						
	Other									1 (9.1%)		1 (4.8%)	1 (1.4%)
# Years Enrolled	1	54 (100%)	39 (100%)	2 (5%)	3 (5.2%)	17 (50%)	9 (11.7%)	10 (45.5%)	12 (75%)		1 (12.5%)	15 (71.4%)	55 (79.7%)
	2			38 (95%)	55 (94.8%)	1 (2.9%)	5 (6.5%)	12 (54.5%)	4 (25%)		3 (37.5%)		1 (1.4%)
	3					18 (50%)	64 (78%)			II (100%)	4 (50%)	1 (4.8%)	3 (4.3%)
	4										5 (23.8%)	10 (14.5%)	

Percentages are by column

Table 2.3: Influenza virus strain components within the FluZone inactivated virus vaccines.

Season	H1N1	H3N2	B-Victoria	B-Yamagata
2014	California-2009	Texas-2012		Massachusetts-2012*
2015	California-2009	Switzerland-2013	Brisbane-2008	Phuket-2013*
2016	California-2009	Hong Kong-2014	Brisbane-2008*	Phuket-2013
2017	Michigan-2015	Hong Kong-2014	Brisbane-2008*	Phuket-2013
2018	Michigan-2015	Singapore-2016	Colorado-2017*	Phuket-2013
2019	Brisbane-2018	Kansas-2017	Colorado-2017*	Phuket-2013

*Virus strains included in the high dose vaccine formulations

2.3 Influenza viruses

The Do and D₂₁-D₂₈ sera from each individual were tested against a virus panel to determine pre- and post-vaccination titers. HAI titers were measured against the vaccine strain for a subtype as well as against heterologous strains. A broad selection of H₁N₁, H₃N₂, and Type B influenza viruses were included in the HAI panel over the different seasons (Table 2.4). Due to the cross-reactivity of the HAI titers of the Type B lineages, i.e., vaccination with one lineage elicits an antibody response to the vaccination lineage and the heterologous lineage, only the Type A influenza strains were analyzed thoroughly. The cross-reactivity of the Type B viruses and the preliminary investigation into breadth quantification are included in the Appendix A.

2.4 Pair-wise distance measures

The differences between the vaccine and heterologous strains were determined within subtypes. Thus the distances of an H₁N₁ strain to an H₃N₂ strain were not calculated. Each distance is a pair-wise distance from a specified vaccine strain to another strain in the HAI virus panel. The vaccine strains used for the relative distance measures were listed in Table 2.3. The HAI panel strains were all of the strains in Table 2.4. Three different measures were compared, time-, genetic sequence-, and cartographic-based distances. With relative distances, the distance of the vaccine strain to itself has a distance of 0, regardless of measure used, and serves as the reference point.

2.4.1 Time-based measures

The time-based measure required the years of isolation of a select vaccine strain and a virus from the HAI panel strain of the shared subtype. The pair-wise distance was determined by taking the absolute value of the difference between the two strains' years of isolation (Equation 2.3).

Table 2.4: Strain names and source of HA amino acid sequence used for analysis

Strain Name	Short Name	Abbreviation	Full Length?	HA Sequence Source
A/H1Ni/South Carolina/1/1918	H1Ni-South Carolina-1918	SC/18	Yes	UniProt: Q9WFX3
A/H1Ni/Weiss/43	H1Ni-Weiss-1943	Wei/43	Yes	UniProt: Q2oN27
A/H1Ni/Fort Monmouth/1/1947	H1Ni-Fort Monmouth-1947	FM/47	Yes	UniProt: Q84110
A/H1Ni/Denver/1957	H1Ni-Denver-1957	Den/57	Yes	UniProt: Q2IBI1
A/H1Ni/New Jersey/8/1976	H1Ni-New Jersey-1976	NJ/76	Yes	UniProt: LoL141
A/H1Ni/Ussr/90/1977	H1Ni-Ussr-1977	USSR/77	Yes	UniProt: Po3453
A/H1Ni/Brazil/II/1978	H1Ni-Brazil-1978	Bra/78	Yes	UniProt: A4GBX7
A/H1Ni/California/10/1978	H1Ni-California-1978	CA/78	Yes	UniProt: A4U6W3
A/H1Ni/Chile/1/1983	H1Ni-Chile-1983	Chi/83	Yes	UniProt: A4GCH5
A/H1Ni/Singapore/6/1986	H1Ni-Singapore-1986	Sing/86	Yes	UniProt: A4GCNo
A/H1Ni/Texas/36/1991	H1Ni-Texas-1991	TX/91	Yes	UniProt: B4UPL3
A/H1Ni/Beijing/262/1995	H1Ni-Beijing-1995	Bei/95	Yes	UniProt: B4UPF7
A/H1Ni/New Caledonia/20/1999	H1Ni-New Caledonia-1999	NC/99	Yes	UniProt: Q6WG00
A/H1Ni/Solomon Islands/3/2006	H1Ni-Solomon Islands-2006	SI/06	Yes	UniProt: A7Y8Ir
A/H1Ni/Brisbane/59/2007	H1Ni-Brisbane-2007	Bris/07	Yes	UniProt: D5FiQ8
A/H1Ni/California/07/2009	H1Ni-California-2009	CA/09	Yes	UniProt: C3W5X2
A/H1Ni/Michigan 45/2015	H1Ni-Michigan-2015	MI/15	Yes	UniProt: AoAi44YDV8
A/H1Ni/Brisbane/02/2018	H1Ni-Brisbane-2018	Bris/18	Yes	Gisaid: EPII415369
A/H3N2/Hong Kong/8/1968	H3N2-Hong Kong-1968	HK/68	Yes	UniProt: A6YBG1
A/H3N2/Port Chalmers/1/1973	H3N2-Port Chalmers-1973	PC/73	Yes	UniProt: Q1PUD9
A/H3N2/Texas/1/1977	H3N2-Texas-1977	TX/77	Yes	UniProt: I6RX51
A/H3N2/Mississippi/1/1985	H3N2-Mississippi-1985	MI/85	No (17-345AA)	Gisaid: EPII29066; Uniprot: Q67178
A/H3N2/Sichuan/2/1987	H3N2-Sichuan-1987	Sich/87	Yes	UniProt: H9XCU1
A/H3N2/Shandong/9/1993	H3N2-Shandong-1993	Shan/93	Yes	UniProt: H9XM74
A/H3N2/Nanchang/933/1995	H3N2-Nanchang-1995	Nan/95	Yes	UniProt: H9XED1
A/H3N2/Sydney/5/1997	H3N2-Sydney-1997	Syd/97	Yes	UniProt: C3PR59
A/H3N2/Panama/2007/1999	H3N2-Panama-1999	Pan/99	Yes	UniProt: QiK9M3
A/H3N2/Fujian/411/2002	H3N2-Fujian-2002	Fuj/02	Yes	UniProt: H9XEX9
A/H3N2/New York/55/2004	H3N2-New York-2004	NY/04	Yes	UniProt: B4UPJ0
A/H3N2/Wisconsin/67/2005	H3N2-Wisconsin-2005	WI/05	Yes	UniProt: WoRXT2
A/H3N2/Uruguay/716/2007	H3N2-Uruguay-2007	Uru/07	Yes	UniProt: B2ZV32
A/H3N2/Perth/16/2009	H3N2-Perth-2009	Per/09	Yes	UniProt: C6KNH7
A/H3N2/Victoria/361/2011	H3N2-Victoria-2011	Vic/11	Yes	UniProt: AoAo97PF39
A/H3N2/Texas/50/2012	H3N2-Texas-2012	TX/12	Yes	UniProt: R4L4F3
A/H3N2/Switzerland/9715293/2013	H3N2-Switzerland-2013	Switz/13	Yes	Gisaid: EPII30687
A/H3N2/Hong Kong/4801/2014	H3N2-Hong Kong-2014	HK/14	Yes	Gisaid: EPII834581
A/H3N2/Alaska/232/2015	H3N2-Alaska-2015	AK/15	Yes	UniProt: AoAi20I466
A/H3N2/Singapore/infimh-16-0019/2016	H3N2-Singapore-2016	Sing/16	Yes	Gisaid: EPII780183
A/H3N2/Stockholm/28/2016	H3N2-Stockholm-2016	Stock/16	Yes	Gisaid: EPII781636
A/H3N2/Kansas/14/2017	H3N2-Kansas-2017	KS/17	Yes	UniProt: AoA2L2FM43
A/H3N2/Switzerland/8060/2017	H3N2-Switzerland-2017	Switz/17	Yes	Gisaid: EPII266285
A/H3N2/South Australia/34/2019	H3N2-South Australia-2019	SA/19	Yes	Gisaid: EPII387331
B/Presplit-Lineage/Lee/1940	B-Lee-1940	P/Lee/40	Yes	UniProt: Po3460
B/Presplit-Lineage/Maryland/59	B-Maryland-1959	P/MD/59	No (11-584AA)	UniProt: Po3461
B/Presplit-Lineage/Singapore/1964	B-Singapore-1964	P/Sing/64	No (5-379AA)	UniProt: Pi2443
B/Victoria-Lineage/Victoria/02/1987	B-Victoria-1987	V/Vic/87	Yes	UniProt: A4D5N9
B/Victoria-Lineage/Hong Kong/330/2001	B-Hong-Kong-2001	V/HK/01	Yes	UniProt: Q596G1
B/Victoria-Lineage/Malaysia/27127/2004	B-Malaysia-2004	V/Mal/04	Yes	UniProt: I2DDB7
B/Victoria-Lineage/Victoria/326/2006	B-Victoria-2006	V/Vic/06	Yes	UniProt: U3RTN3
B/Victoria-Lineage/Brisbane/60/2008	B-Brisbane-2008	V/Bris/08	Yes	UniProt: IoB7A3
B/Victoria-Lineage/Colorado/06/2017	B-Colorado-2017	V/CO/17	Yes	UniProt: AoAiX9RX59
B/Victoria-Lineage/Washington/02/2019	B-Washington-2019	V/WA/19	Yes	UniProt: AoA4Y5N4V8
B/Yamagata-Lineage/Yamagata/16/88	B-Yamagata-1988	Y/Yam/88	No (16-360AA)	UniProt: Pi8880
B/Yamagata-Lineage/Harbin/7/1994	B-Harbin-1994	Y/Har/94	Yes	UniProt: C4LQ32
B/Yamagata-Lineage/Sichuan/379/1999	B-Sichuan-1999	Y/Sich/99	No (16-361AA)	UniProt: A5A5P7
B/Yamagata-Lineage/Florida/4/2006	B-Florida-2006	Y/FL/06	Yes	UniProt: R9RV94
B/Yamagata-Lineage/Wisconsin/01/2010	B-Wisconsin-2010	Y/WI/10	Yes	UniProt: G8JJL9
B/Yamagata-Lineage/Texas/06/2011	B-Texas-2011	Y/TX/11	Yes	UniProt: M9QX33
B/Yamagata-Lineage/Massachusetts/02/2012	B-Massachusetts-2012	Y/MA/12	Yes	UniProt: R4KWJ1
B/Yamagata-Lineage/Phuket/3073/2013	B-Phuket-2013	Y/Phu/13	Yes	Gisaid: EPII649072

$$Time_{ab} = | \text{Year}_a - \text{Year}_b | \quad (2.3)$$

where:

Year = Year of isolation as classified in the strain name

a = Vaccine Strain

b = HAI panel strain

2.4.2 Sequence-based measures

The hemagglutinin amino acid sequence for each vaccine strain and HAI virus were retrieved from either GISAID or UniProtKB databases. The HA amino acid sequences were either extracted from the UniProt protein database using the Rcpip package `getSeqFromUniProt()` function and the UniProt accession numbers or from manually downloading the sequence from the Gisaid database if the sequence was not available on UniProt (Cao et al., 2015; Elbe & Buckland-Merrett, 2017; UniProt Consortium et al., 2020). The previously described dominant p-epitope was used as the sequence-based distance measure (Gupta et al., 2006). Antigenic site-specific p-epitopes are selected by calculating the pair-wise hamming distance of an antigenic site and dividing by the total amino acid residues that comprise that site (Equation 2.4). After calculating the site-specific p-epitopes for all antigenic sites, the dominant p-epitope measure was defined as the maximum of those (Equation 2.5). The antigenic sites used for calculating the sequence-based P-epitope were previously defined (Deem & Pan, 2009; Munoz & Deem, 2005). The specific residues used are included in Table 2.5. The numbering begins at the beginning of the N-terminal sequence of the mature protein after the signal peptide had been removed (Burke & Smith, 2014). The extraction of the defined antigenic sites was done in R using the 'seqinr' and 'Biostrings' packages (Charif & Lobry, 2007; Pagès et al., 2021).

$$p_{epitopeX} = \frac{\text{Hamming distance between vaccine strain and HAI panel virus in epitope X}}{\text{total number of amino acids in epitope X}} \quad (2.4)$$

where: X = Antigenic sites A, B, C, D or E

$$P_{epitope} = \max(p_{epitopeA}, p_{epitopeB}, p_{epitopeC}, p_{epitopeD}, p_{epitopeE}) \quad (2.5)$$

Table 2.5: Amino acid residues used for the sequence-based distance measure, p-epitope.

Subtype	Antigenic Site	Epitope Residues
H1N1	A	118, 120, 121, 122, 126, 127, 128, 129, 132, 133, 134, 135, 137, 139, 140, 141, 142, 143, 146, 147, 149, 165, 252, 253
	B	124, 125, 152, 153, 154, 155, 156, 157, 160, 162, 183, 184, 185, 186, 187, 189, 190, 191, 193, 194, 195, 196
	C	34, 35, 36, 37, 38, 40, 41, 43, 44, 45, 269, 270, 271, 272, 273, 274, 276, 277, 278, 283, 288, 292, 295, 297, 298, 302, 303, 305, 306, 307, 308, 309, 310
	D	89, 94, 95, 96, 113, 117, 163, 164, 166, 167, 168, 169, 170, 171, 172, 173, 174, 176, 179, 198, 200, 202, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 222, 223, 224, 225, 226, 227, 235, 237, 241, 243, 244, 245
	E	47, 48, 50, 51, 53, 54, 56, 57, 58, 66, 68, 69, 70, 71, 72, 73, 74, 75, 78, 79, 80, 82, 83, 84, 85, 86, 102, 257, 258, 259, 260, 261, 263, 267
H3N2	A	122, 124, 126, 130, 131, 132, 133, 135, 137, 138, 140, 142, 143, 144, 145, 146, 150, 152, 168
	B	128, 129, 155, 156, 157, 158, 159, 160, 163, 165, 186, 187, 188, 189, 190, 192, 193, 194, 196, 197, 198
	C	44, 45, 46, 47, 48, 50, 51, 53, 54, 273, 275, 276, 278, 279, 280, 294, 297, 299, 300, 304, 305, 307, 308, 309, 310, 311, 312
	D	96, 102, 103, 117, 121, 167, 170, 171, 172, 173, 174, 175, 176, 177, 179, 182, 201, 203, 207, 208, 209, 212, 213, 214, 215, 216, 217, 218, 219, 226, 227, 228, 229, 230, 238, 240, 242, 244, 246, 247, 248
	E	57, 59, 62, 63, 67, 75, 78, 80, 81, 82, 83, 86, 87, 88, 91, 92, 94, 109, 260, 261, 262, 265

H1N1 numbering follows A/California/04/2009 numbering scheme.

All numbering starts at the beginning of the N-terminal sequence of the mature protein after removal of the signal peptide.

2.4.3 Antigenic cartography-based measures

Each of the Type A influenza subtypes (H_1N_1 , H_3N_2) were analyzed individually. Antigenic cartographies were created with the human HAI datasets using the post-vaccinated HAI titers of all standard dose recipients. The R package Racmcs was used to create the HAI-based antigenic cartographies (Wilks, 2021). Sera samples and HAI viruses with less than $n + 1$ titers for the dimension (n) tested were underconstrained for mapping and removed before cartography creation. Two-dimensional cartography was found to be appropriate for use with both subtypes. The HAI titers from standard dose vaccination were used to limit the number of variables that could affect antigen positioning. Mapping with the resulting data set was done with 100 optimizations. The map with the lowest resulting stress was used for calculating the pair-wise distance between vaccine virus and HAI panel viruses. More details regarding the antigenic cartography with human sera are available in Chapter 3.

Antigenic cartography was utilized to provide antigenic distances between different viruses. Distances relative to the different vaccine strains were determined by calculating the Euclidean distances using the map coordinates of the vaccine strain and the other strains within the HAI panel (Equation 2.6) (Cai et al., 2012).

$$\text{Cartography Distance}_{ab} = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2} \quad (2.6)$$

where:

(x_a, y_a) = Vaccine strain coordinates

(x_b, y_b) = HAI panel strain coordinates

2.5 Normalization of measures

The pair-wise distances had different ranges depending on the measure used (Table 2.5). Therefore, to allow for comparison between them, they were normalized by the maximum value to rescale the values from 0 to 1. Two different normalization schemes were presented due to the variation in the HAI panel used over the different seasons.

Season-based normalization used when comparing season-level associations

$$\text{Normalized Distance}_{ijklm} = \frac{d_{ijklm}}{\max(D_{ijkl})} \quad (2.7)$$

Strain-based normalization used when comparing strain-level associations

$$\text{Normalized Distance}_{ijklm} = \frac{d_{ijklm}}{\max(D_{ijk})} \quad (2.8)$$

where:

d = Untransformed Pair-wise Distance

i = Distance Method (Year, P-epitope, Cartography)

j = Vaccine Subtype (H₁N₁, H₃N₂)

k = Vaccine Strain

l = Influenza Season (2014-2019)

m = HAI Panel Virus

The maximum distances used to normalize change depend upon the grouping and comparisons used (Table 2.5). For the H₁N₁ viruses, the HAI panel did not vary much for the CA/09 and MI/15 vaccine viruses since the panel did not change significantly over those seasons (Figure 2.6). For the H₃N₂ viruses,

the normalization denominator was similar up until 2018 and 2019, when the HAI panel was truncated. The difference in virus panels over time are more clearly depicted in Figures 2.6 and 2.7.

2.6 Comparison of measures

One of the key features of using different distance measures to quantify pair-wise virus distances is that the relative ordering of the virus panel is adjusted depending on which measure is used. For example, Bris/07 has a very short year distance from CA/09 (2 years) but a larger sequence- and cartographic-based distance (Figure 2.6). Whereas for the H₃N₂, rankings between viruses do not change as often as the H₁N₁ strains due to the relationship between the year, sequence, and cartography measures (Figure 2.7).

2.7 Linear regression, area calculations, and weighting

Simple linear regression used the distance measure as the independent variable and the outcome as the dependent variable. The outcomes for linear regression included: pre- and post-vaccination titer and titer increase. The linear regression equations were provided with the 95% confidence interval bands. The area under the linear regression was quantified between the minimum and maximum distance values. Area below $y = 0$ was considered negative area in the summation.

Three weighting schemes included unweighted, linear decrease, and antigenic unit cutoff. The unweighted scheme was the area under the regression with no transformation. This weighting scheme applies to situations where antibody responses to all viruses in the panel are equally important. The linear decrease weighting applied a greater weight to the closely related strains and linearly decreased that weight to 0 for the furthest strains. This scheme applies to situations where an antibody response to the vaccine virus and closely related ones is more important than the furthest strain in the panel. The last weighting incorporates the antigenic unit threshold used for vaccine selection. An antigenic unit of 2 was used as the cutoff to represent a 4-fold HAI titer difference. All antibody responses within this measure were equally

Table 2.6: Maximum normalized distances when grouping by season or strain

Method	Subtype	Vaccine	Season	Season	Strain
				$\max(D_{ijkl})$	$\max(D_{ijk})$
Year	H1N1	CA/09	2014		
			2015	91.00	91.00
			2016		
		MI/15	2017	97.00	97.00
			2018	32.00	
	Bris/18		2019	35.00	35.00
		TX/12	2014	44.00	44.00
		Switz/13	2015	45.00	45.00
	H3N2	HK/14	2016		
			2017	46.00	46.00
		Sing/16	2018	17.00	17.00
		KS/17	2019	18.00	18.00
P-epitope	H1N1	CA/09	2014		
			2015	0.77	0.77
			2016		
		MI/15	2017		
			2018	0.82	0.82
	Bris/18		2019		
		TX/12	2014	0.71	0.71
		Switz/13	2015	0.67	0.67
	H3N2	HK/14	2016		
			2017	0.71	0.71
		Sing/16	2018		
		KS/17	2019	0.43	0.43
Cartography	H1N1	CA/09	2014		
			2015	4.20	4.24
			2016	4.24	
		MI/15	2017	4.73	
			2018	4.51	4.73
	Bris/18		2019	4.15	4.15
		TX/12	2014	5.46	5.46
		Switz/13	2015	5.43	5.43
	H3N2	HK/14	2016		
			2017	6.24	6.24
		Sing/16	2018	3.45	3.45
		KS/17	2019	3.57	3.57

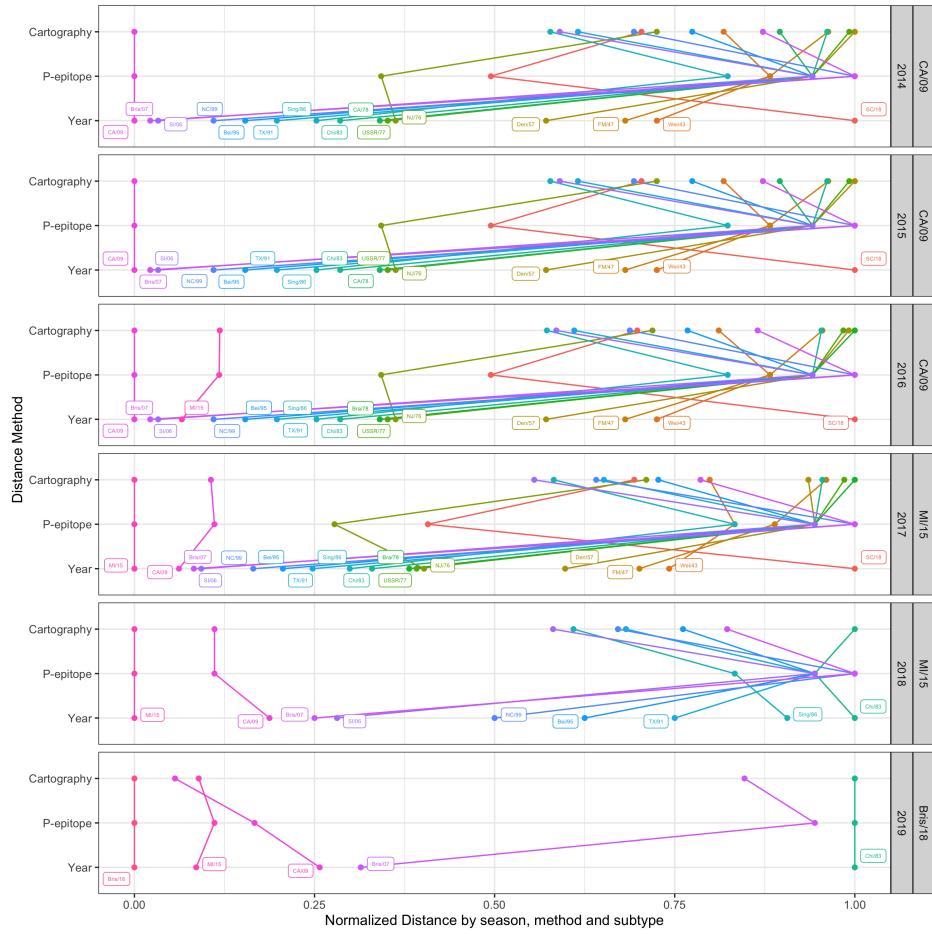


Figure 2.2: Comparison of the reordering of influenza H1N1 strains by the normalized distances. Depending upon the measure, the position on the x-axis varies by season, method, and subtype. The virus panel for each season had minor variations leading to different normalized values per season. These x-axis positions correspond to the axis positions for the figures investigating the by season analysis. The vaccine strain and season are listed for each plot.

weighted, and all responses greater than it were weighted at 0. The cutoff value can be adjusted based on the research question.

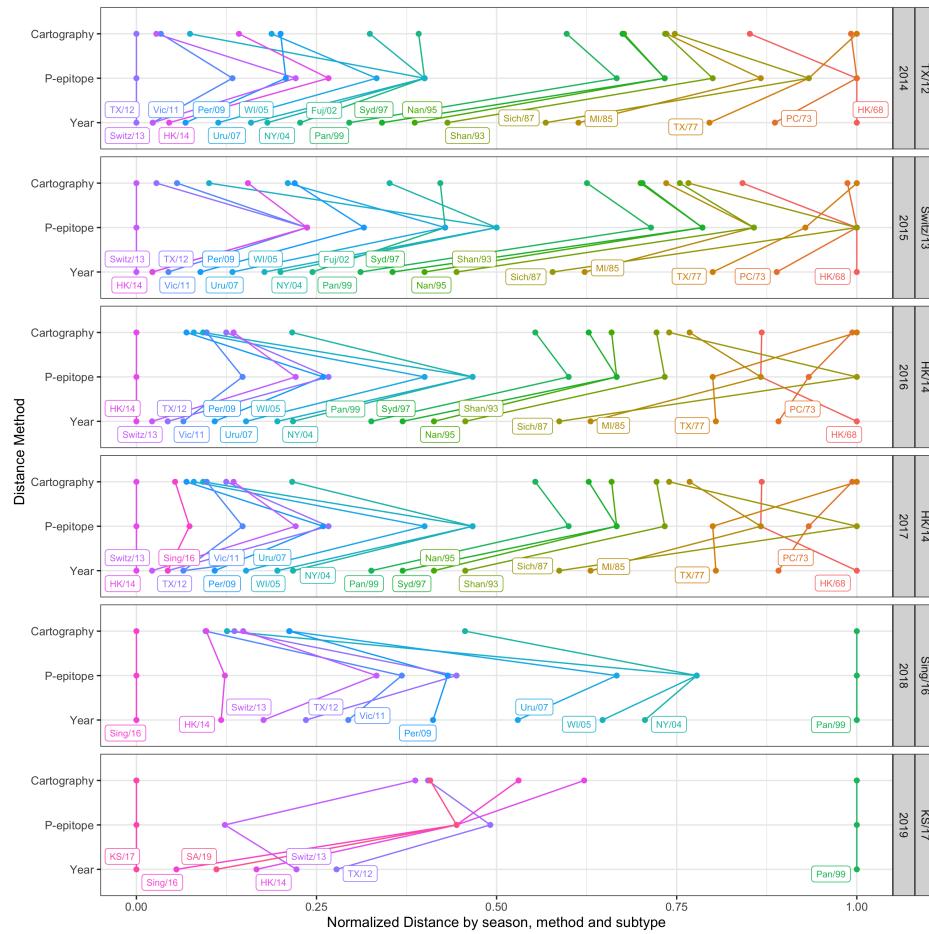


Figure 2.3: Comparison of the reordering of influenza H₃N₂ strains by the normalized distances. Depending upon the measure, the position on the x-axis varies by season, method, and subtype. The virus panel for each season had minor variations leading to different normalized values per season. These x-axis positions correspond to the axis positions for the figures investigating the by season analysis. The vaccine strain and season are listed for each plot.

CHAPTER 3

ANTIGENIC CARTOGRAPHY USING PRE-IMMUNE HUMAN SERA

Vaccine strain selection utilizes influenza-naive ferrets for anti-sera generation. These animals have no pre-existing immunity to influenza. This model provides a platform free of the imprinting bias imparted from initial and multiple exposures to influenza. Humans, however, have been imprinted during childhood with the first strain that infected them. Imprinting bias has been observed between subtypes; H₁N₁ imprinted individuals do not have the same immune response to an H₃N₂ or H₇N_X infection as an H₃N₂ imprinted individual would (Gostic et al., 2016). Therefore, using an influenza-naive ferret model for immunological responses may not be appropriate to model an influenza pre-immune background which is most commonly seen in adolescents and adults (For a review of pre-immune ferret models, please refer to Skarupka and Ross, 2020).

Most humans older than four years have been immunologically exposed to influenza virus either through an initial vaccination or an infection. Individuals born before 1956 were imprinted with H₁N₁, individuals from 1956 to 1968 with H₂N₂, and individuals from 1968 to 1977 with H₃N₂. After 1977, H₁N₁ re-emerged in the human population. Due to the co-circulation of the Type A subtypes, the imprinting history of individuals after 1977 is difficult to discern. The H₁N₁ pandemic of 2009 further complicates the imprinting histories. Individuals imprinted with H₁N₁ after 1977 have been imprinted with either

human-seasonal or pandemic H1N1 strain. Birth year and the surveillance levels of circulating subtypes are used to determine the probability of one's imprinting status (Gostic et al., 2019). Other methods include analyzing an individual's present-day influenza antibody profile and using that information to infer the previous exposure history (Hay et al., 2020).

3.1 Dataset and number of dimensions selection

Due to the lack of preliminary data on the reliability of antigenic cartography from pre-immune human sera, antigenic cartographies were created using different subsets of the cohort sera and comparing the results to determine if there were viable differences between the maps. The other comparisons included maps created with pre-vaccination HAI titers and post-vaccination HAI titers, maps created with sera from only SD vs. all influenza vaccine recipients, and season-based maps produced with only sera from a specific season compared to all seasons from 2014-2019. The following section outlines the selection process of the dataset used for pair-wise cartography distances between viruses. Further, the resulting antigenic cartographies were validated with published data.

First, the appropriate number of dimensions for analysis was determined through a dimension test (using the Racmacs package and dimensionTestMap() function). Each map was cross-validated to determine how well the predicted titers matched the observed titers. When comparing maps created with all and only SD sera-virus pairs, there was minimal difference between the root mean square error (RMSE) of the maps (Table 3.1). The RMSE was calculated separately for detectable and nondetectable titers (titers below the limit of detection <1:10). When comparing the different dimensions, the detectable RMSE of the H1N1 viruses plateaued at two dimensions (Figure 3.1). For the H3N2 viruses the RMSE plateaued at three dimensions, with the decrease from two to three dimensions being minimal compared to the decline from one to two dimensions (Table 3.1).

Table 3.1: Dimensional analysis of H1N1 and H3N2 human cartography maps pre- and post-vaccination using all sera available (all) or only sera from SD individuals (SD) and comparison of the mean root mean square error (RMSE) and variance (Var).

Subtype	Timing	Dimensions	Vaccine	Detectable		Nondetectable	
				All	Mean RMSE	Var RMSE	Mean RMSE
H1N1	Pre	1	All	1.634	0.002	1.582	0.001
			SD	1.689	0.002	1.620	0.001
		2	All	1.501	0.002	1.372	0.002
			SD	1.524	0.002	1.365	0.003
		3	All	1.502	0.002	1.245	0.003
			SD	1.524	0.002	1.240	0.002
		4	All	1.501	0.002	1.142	0.002
			SD	1.520	0.003	1.138	0.002
		5	All	1.509	0.002	1.110	0.002
			SD	1.531	0.003	1.113	0.002
H3N2	Post	1	All	1.675	0.002	1.846	0.001
			SD	1.726	0.002	1.833	0.002
		2	All	1.552	0.002	1.427	0.003
			SD	1.554	0.002	1.405	0.004
		3	All	1.543	0.002	1.314	0.002
			SD	1.562	0.002	1.277	0.003
		4	All	1.532	0.002	1.239	0.002
			SD	1.556	0.002	1.207	0.003
		5	All	1.539	0.002	1.226	0.002
			SD	1.564	0.003	1.192	0.003
	Pre	1	All	1.644	0.002	2.325	0.008
			SD	1.712	0.005	2.418	0.016
		2	All	1.376	0.001	1.443	0.003
			SD	1.396	0.001	1.430	0.003
		3	All	1.344	0.001	1.338	0.003
			SD	1.363	0.001	1.345	0.003
		4	All	1.341	0.001	1.311	0.003
			SD	1.363	0.001	1.325	0.003
		5	All	1.341	0.001	1.303	0.003
			SD	1.363	0.001	1.318	0.002
	Post	1	All	1.563	0.010	2.029	0.110
			SD	1.578	0.011	2.032	0.119
		2	All	1.355	0.001	1.555	0.003
			SD	1.356	0.001	1.539	0.005
		3	All	1.323	0.001	1.494	0.003
			SD	1.326	0.001	1.485	0.003
		4	All	1.328	0.001	1.489	0.002
			SD	1.330	0.001	1.478	0.003
		5	All	1.327	0.001	1.488	0.002
			SD	1.330	0.001	1.477	0.004

Each dimension was tested with 100 replicates,
with each of those maps undergoing 1000 optimizations.
The test proportion was 10% of the full data set.

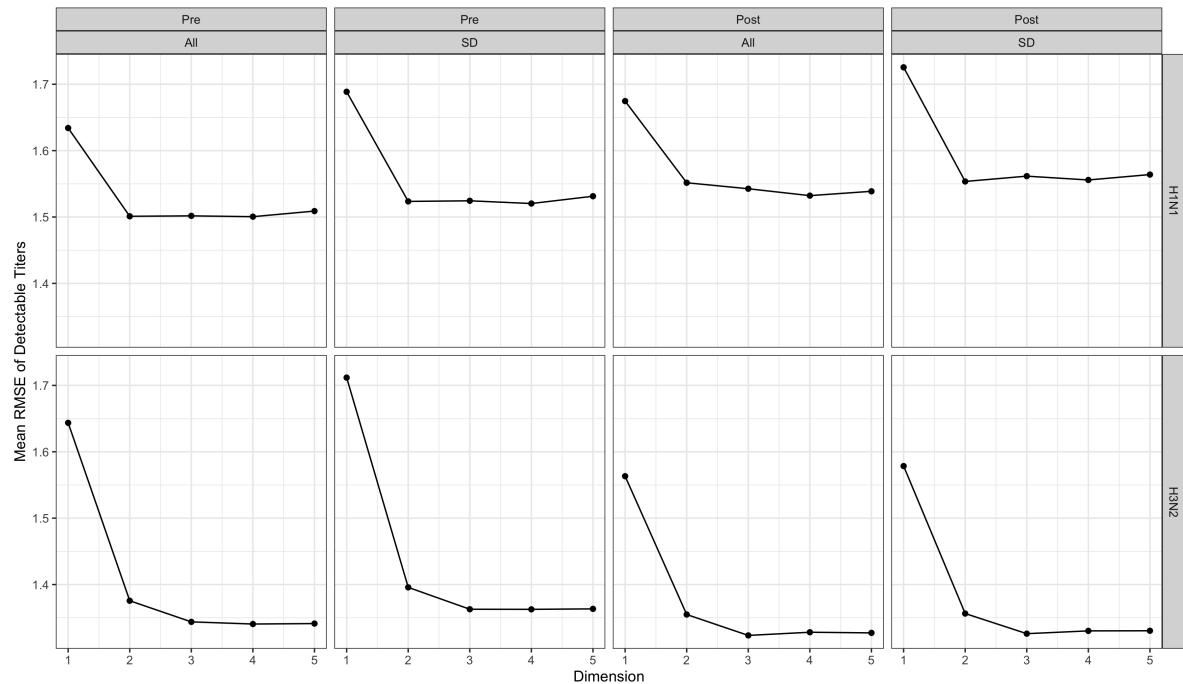


Figure 3.1: Validation of two-dimensional antigenic cartography with human sera. Each of the datasets was tested through five dimensions. Two dimensions were determined to be appropriate for visualizing the maps as indicated by the leveling off of the root mean square prediction error (RMSE). Each dimension was tested with 100 replicates, with each of those maps undergoing 1000 optimizations. The test proportion was 10% of the full data set

3.2 Comparison to previous studies

Antigenic cartography was introduced with the H₃N₂ viruses using ferret sera (Smith et al., 2004). Fonville et al., 2015 then expanded the initial map, which previously ended in 2002, to 2014. The viruses in this study (this thesis) do not overlap entirely with the viruses used for Smith et al., 2004 or Fonville et al., 2015 maps, and the virus distances from these maps cannot be used. Generation of the ferret HAI titers necessary to provide the virus distances was outside the scope of this study in resources and time.

In addition to the H₃N₂ maps being incomplete, only a limited set of antigenic maps for the H₁N₁ viruses have been completed. The antigenic maps for the H₁N₁ viruses have been done using serological HAI data for H₁ human and swine viruses (Lorusso et al., 2010; Smith et al., 2004; Tapia et al., 2020) and using sequence-based antigenic distance estimates (Anderson et al., 2018).

Therefore, the HAI titer matrix from this human cohort was used to create both H₁N₁ and H₃N₂ maps. Before using distances from these cartographies, the H₃N₂ pre-immune human maps were confirmed to match H₃N₂ cartographies created with ferret sera (Fonville et al., 2015; Smith et al., 2004) and child human sera. Child human sera were used for the Fonville et al., 2015 study to minimize the pre-existing antibodies to influenza in the sera. Adolescents and adults have a history of exposure; therefore, it was confirmed that these pre-existing antibodies did not overtly affect the position of the viral antigens on the cartographies.

The Smith et al., 2004 HAI data was sourced from the Racmacs R package (Wilks, 2021). The Fonville et al., 2015 HAI data was sourced from the manuscript's supplementary data, specifically supplementary dataset 5 for the ferret H₃N₂ maps and supplementary dataset 4 for the high responding human H₃N₂ maps. The map from this study used as the comparator was the H₃N₂ map created with the post-vaccination standard dose only sera. The Fonville et al., 2015 and current studies maps were compared to the Smith et al., 2004 maps (Figure 3.2). All maps were aligned to each other using Procrustes analysis, which allows for translation, scaling, or both (Table 3.2). Scaling when aligning maps was recommended for combining data created using different assays and multiplying all coordinates by a constant scalar.

Translation allows for the movement of the origin of one map. Both of these techniques are used to minimize the distance between matching points. Scaling and translation were both investigated because it was observed that the human cartographies, either pre-immune or child, had a smaller radius than the ferret maps. When comparing the Smith et al., 2004 map to the current study maps, the root means square distance (RMSD) of the antigens was smaller when scaling and translation were both applied to the pre-immune human cartography.

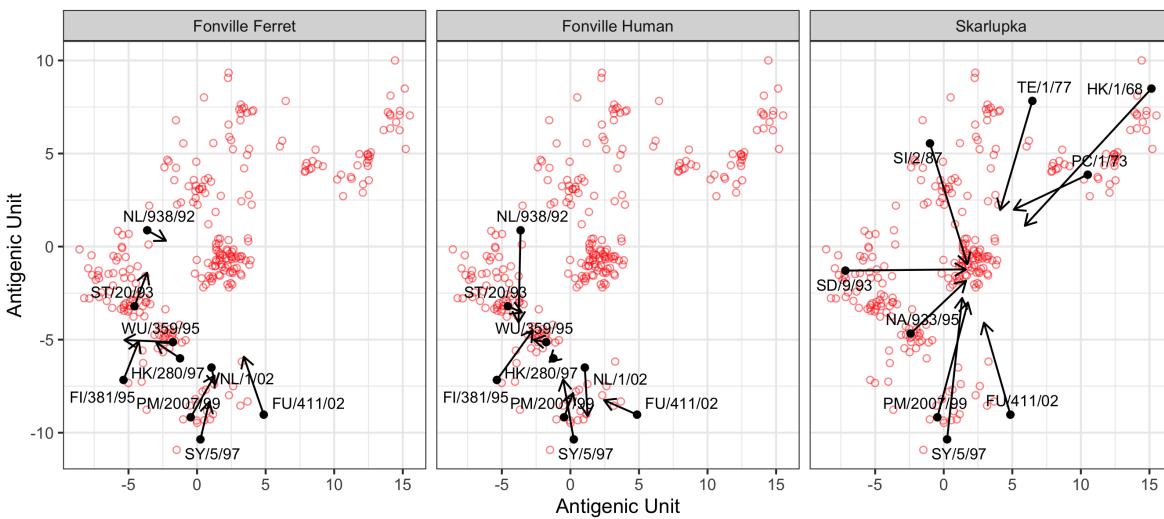


Figure 3.2: Procrustes maps of different datasets compared to the original H₃N₂ cartography. The Smith et al., 2004 cartography had nine viruses that overlapped with Fonville et al., 2015 and current study. The viruses that did not have a match are shown with red open circles. Arrows indicate the positioning of the virus in the other cartography. Maps were aligned, allowing for only translation.

After confirming that the maps produced with human pre-immune sera correspond to the previously published maps, the distances derived from these maps were investigated. Due to the interest in the cartographic pair-wise distance from a specified vaccine strain, the relative cartographic distance introduced previously was used (J. Wang et al., 2021). After selecting the vaccine virus to use for the reference point of the relative antigenic distance, the Euclidean distance from the vaccine virus to the HAI panel virus was calculated by extracting the position coordinates of both viruses and calculating the Euclidean distance between the two points (Equation 2.6). For validation, all pair-wise distances were found regardless of whether a virus was a vaccine virus. Thus, for each map and virus, the relative antigenic distances were

Table 3.2: Procrustes analysis RMSD values under different alignment conditions

Map 1	Map 2	# Ag	RMSD of matched viruses		
			Translation = T Scaling = F	Translation = F Scaling = T	Translation = T Scaling = T
Smith	Fonville Ferret	9	3.314	4.123	3.604
	Fonville Human	9	2.395	3.34	3.07
	Skarlupka	9	8.027	3.341	2.704
Fonville Ferret	Fonville Human	14	3.211	3.237	2.321
	Skarlupka	5	2.924	2.043	2.255
Fonville Human	Skarlupka	5	2.397	2.252	1.976

Map 1: Map that is procrustes against

Map 2: Comparison map

RMSD: root mean square distance

Ag: Number of shared viruses (antigens) between maps

calculated. The correlation of the Smith et al., 2004 distances to the other maps for each reference virus was determined (Figure 3.3, Table 3.3). The relative distances with the pre-vaccination and all sera dataset maps were compared and did not differ significantly from the post-vaccination SD sera dataset results.

Of the different virus pairs, there was only one that was not significantly correlated with Smith et al., 2004 distances, SI/2/87 (Sing/87) with a Pearson correlation coefficient of 0.555. The Fonville et al., 2015 dataset did not contain this virus to provide a comparison.

The datasets were then separated even further into individual seasons to determine if there was a season-based effect. However, after stratifying by season, the high correlation between the pre- and post-vaccination relative distances remained for the H₃N₂ viruses (Figure 3.4) and the H₁N₁ viruses (Figure 3.5). The total number of viral antigens and sera samples used for each map is included in Table 3.4.

Antigenic distances between the H₁N₁ and H₃N₂ strains were determined using antigenic cartography created using human post-vaccination HAI titers to a panel of H₁N₁ viruses (Figure 3.2A) and the H₃N₂ viruses (Figure 3.2B).

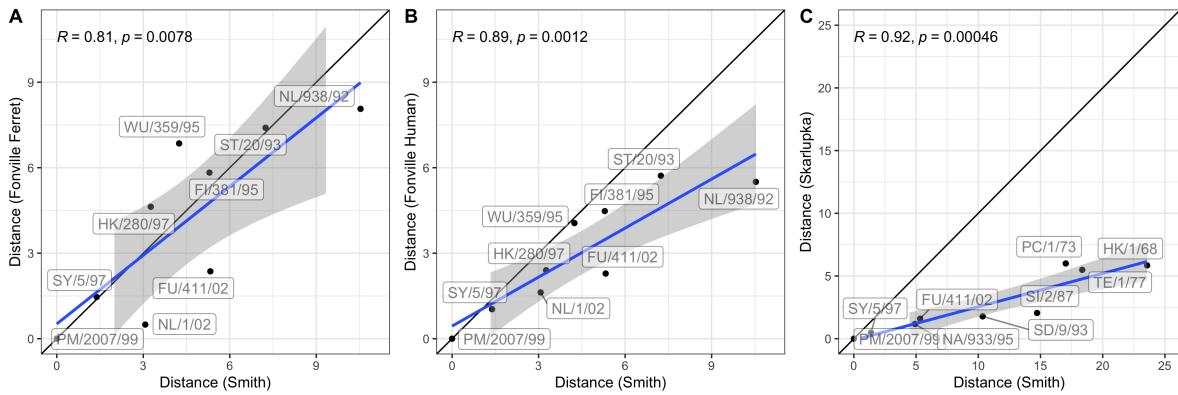


Figure 3.3: Correlations of the relative Euclidean distances from A/H₃N₂/Panama/2007/1999. The relative distances to PM/99 from the Smith et al., 2004 dataset were compared with the matching antigens in the ferret, and human datasets from Fonville et al., 2015 and in this current study (SD dataset). All maps were aligned to the Smith et al., 2004 maps, allowing for only translation. The matched viruses between the Smith et al., 2004 dataset and the other datasets are labeled on the respective plots. The linear regression between the paired points is shown with 95% confidence intervals with the blue line. The black line indicates the line of identity ($y = x$). Pearson's correlation coefficient (R) and the two-sided p -value are included on each plot.

Table 3.3: Correlation coefficients of pair-wise cartographic distances when varying the relative viruses

Relative Virus	Fonville Ferret		Fonville Human		Skarupka	
	r	p-value	r	p-value	r	p-value
HK/1/68					0.939	<0.001
PC/1/73					0.964	<0.001
TE/1/77					0.930	<0.001
SI/2/87					0.555	0.121
NL/938/92	0.916	<0.001	0.930	<0.001		
SD/9/93					0.781	0.013
ST/20/93	0.921	<0.001	0.903	<0.001		
FI/381/95	0.944	<0.001	0.754	0.019		
NA/933/95					0.955	<0.001
WU/359/95	0.935	<0.001	0.804	0.009		
HK/280/97	0.907	<0.001	0.827	0.006		
SY/5/97	0.901	<0.001	0.925	<0.001	0.947	<0.001
PM/2007/99	0.815	0.007	0.890	0.001	0.973	<0.001
FU/411/02	0.915	<0.001	0.949	<0.001	0.978	<0.001
NL/1/02	0.853	0.003	0.843	0.004		

Pearson's correlation coefficient (r) and two-sided p-value

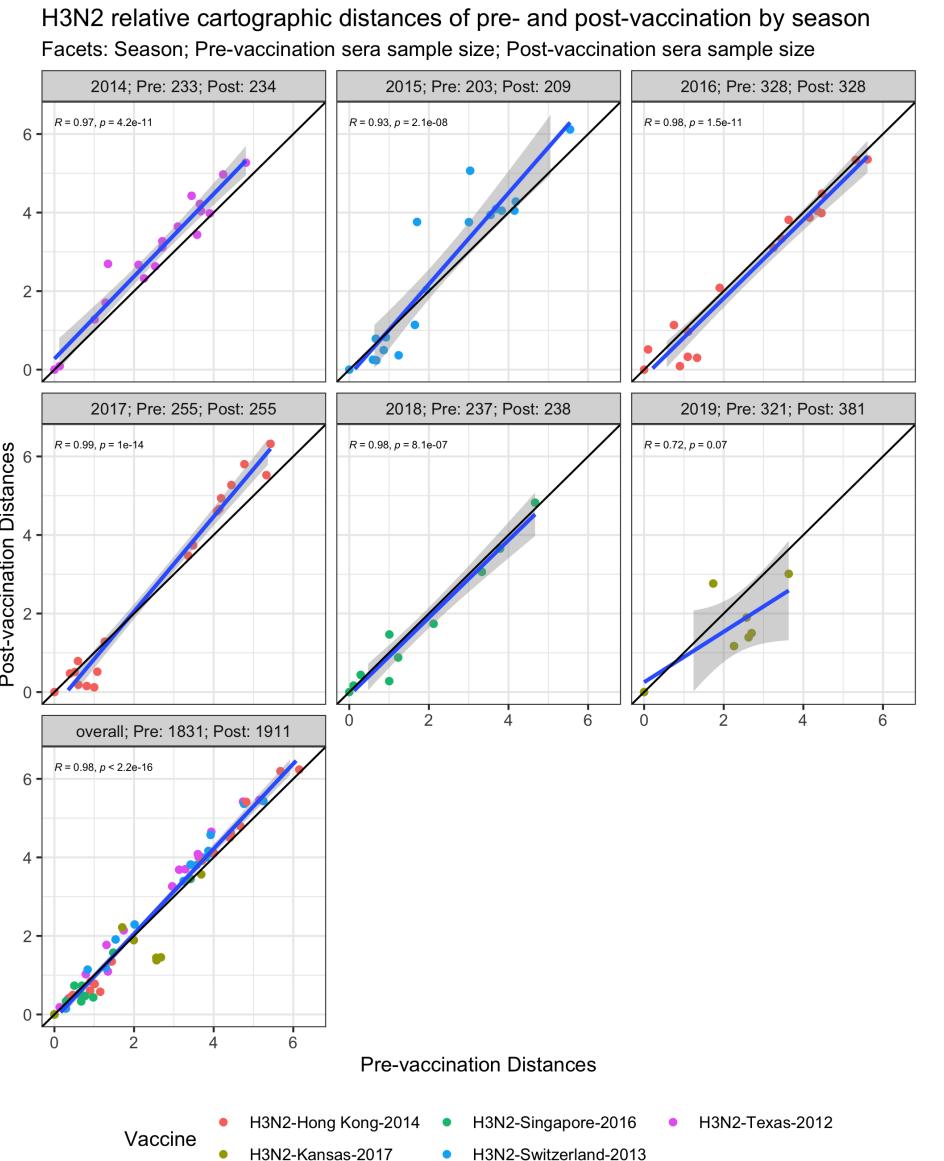


Figure 3.4: Relative distances to the vaccine strains for each season for pre- and post-vaccination standard dose sera datasets of the H₃N₂ viruses. The vaccine strains for the given season were used as the relative virus to calculate the distances. Plots were faceted by season with the sample size of individuals used to create each map. The number of individuals varied due to some being underconstrained for map placement. The Pearson's correlation coefficient (R) and the two-sided p -value are shown. The blue line indicates the linear regression with 95% confidence intervals, and the black line is the line of identity ($y = x$).

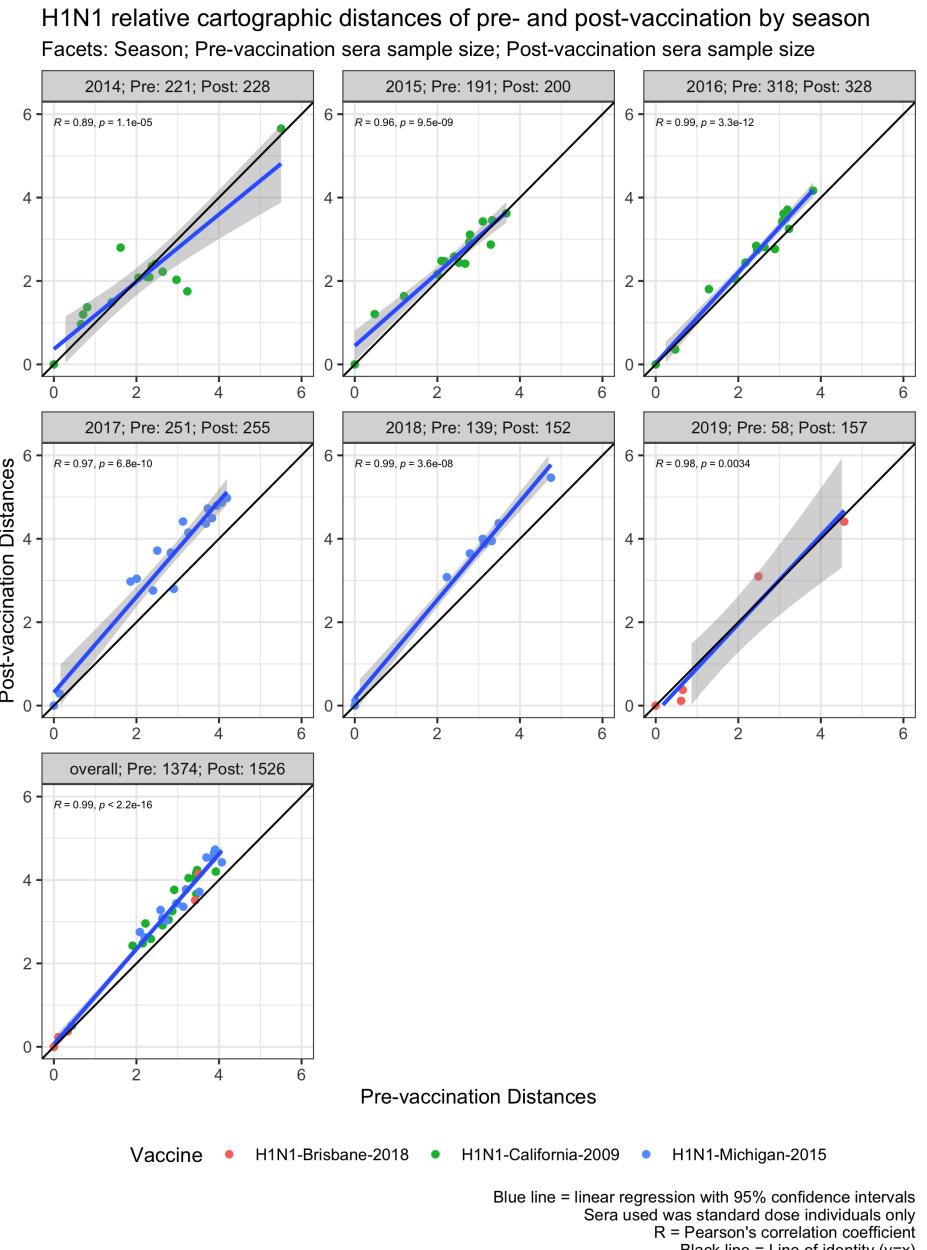


Figure 3.5: Relative distances to the vaccine strains for each season for pre- and post-vaccination standard dose sera datasets of the H₃N₂ viruses. The vaccine strains for the given season were used as the relative virus to calculate the distances. Plots were faceted by season with the sample size of individuals used to create each map. The number of individuals varied due to some being underconstrained for map placement. The Pearson's correlation coefficient (R) and the two-sided p-value are shown. The blue line indicates the linear regression with 95% confidence intervals, and the black line is the line of identity ($y = x$).

Table 3.4: Number of viruses and sera samples included in each antigenic cartography

Season	Subtype	Virus		Sera	
		Pre	Post	Pre	Post
2014	H1N1	15	15	221	228
	H3N2	18	18	233	234
2015	H1N1	15	15	191	200
	H3N2	18	18	203	209
2016	H1N1	16	16	318	328
	H3N2	17	17	328	328
2017	H1N1	16	16	251	255
	H3N2	18	18	255	255
2018	H1N1	9	9	139	152
	H3N2	10	10	237	238
2019	H1N1	5	5	58	157
	H3N2	7	7	321	381
Overall		H1N1	18	1,188	1,316
		H3N2	21	1,585	1,645

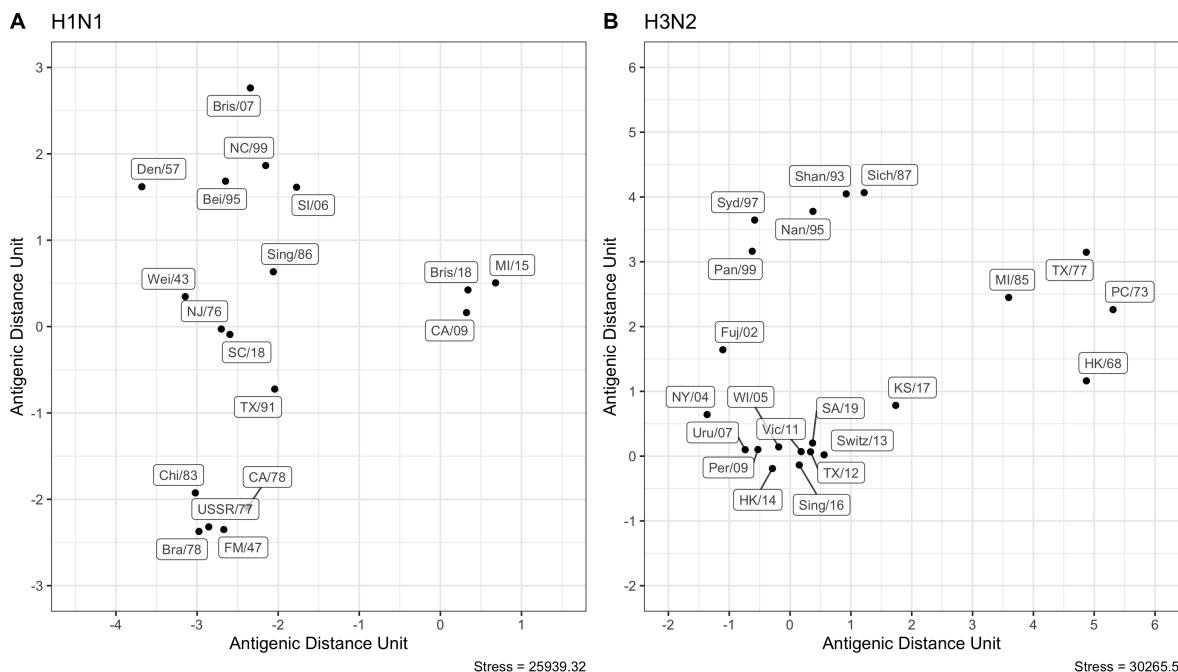


Figure 3.6: Antigenic cartographies were created using pre-immune human sera for the H1N1 and H3N2 subtypes. (A) The H1N1 panel included viruses from 1918 to 2018 with vaccine viruses of CA/09, MI/15, and Bris/18. (B) The H3N2 viruses ranged from 1968 to 2019 with vaccine viruses of TX/12, Switz/13, HK/14, Sing/16, and KS/17. All viruses were of human origin. Each map was created using the post-vaccination HAI titers from SD cohort participants. Each map underwent 100 optimizations. 1 antigenic distance unit = \log_2 increase in HAI = 2-fold increase in HAI.

3.3 Conclusion

The distances of the H₃N₂ viruses using pre-immune human maps were validated using pre-existing data created with a naive ferret and human sera. The relative antigenic distances correlated highly between the different datasets. Therefore, it was concluded that the use of the pre-immune human sera maps was appropriate. For the H₁N₁ viruses, the lack of previously published antigenic cartographies using naive ferret sera limited the ability for validation. Therefore, the H₃N₂ map conclusion was applied to the H₁N₁ cartographies.

The post-vaccination standard dose antigenic cartography maps were used to calculate each virus's relative distance from a vaccine strain for both the H₁N₁ and H₃N₂ subtypes.

CHAPTER 4

APPLICATION OF DISTANCE MEASURES TO VACCINE RESPONSES

Three methods of calculating pair-wise distances between viruses were compared to determine which was more strongly associated with vaccination outcomes. The three methods were 1.) difference in the year of isolation, 2.) p-epitope sequence-based measure, and 3.) antigenic cartographic Euclidean distance. The investigated vaccine outcomes were 1.) pre- and post-vaccination titers and 2.) titer increase.

4.1 Pre- and post-vaccination titers

The relationship between the pre- and post-vaccination titers and the distances were first investigated by stratifying by season and vaccine strain (Figure 4.1 and Supplementary Figure B). All individuals who received the SD vaccination were included for the respective seasons. For H1N1 viruses, antigenic and p-epitope distances had greater slopes and y-intercepts than year distance for each season until 2019, when A/Brisbane/2018 was first introduced as a vaccine antigen. For the H3N2 viruses, the difference between the antigenic and p-epitope regressions was not as great as the year method. There was stronger association with these measures for some seasons, though, especially the p-epitope (2014-2017 seasons). The slopes of the pre- and post-vaccination titers linear regressions were not parallel. Individuals did not have an

equal response across all strains in the panel, which was most prominent in the p-epitope and cartography measures.

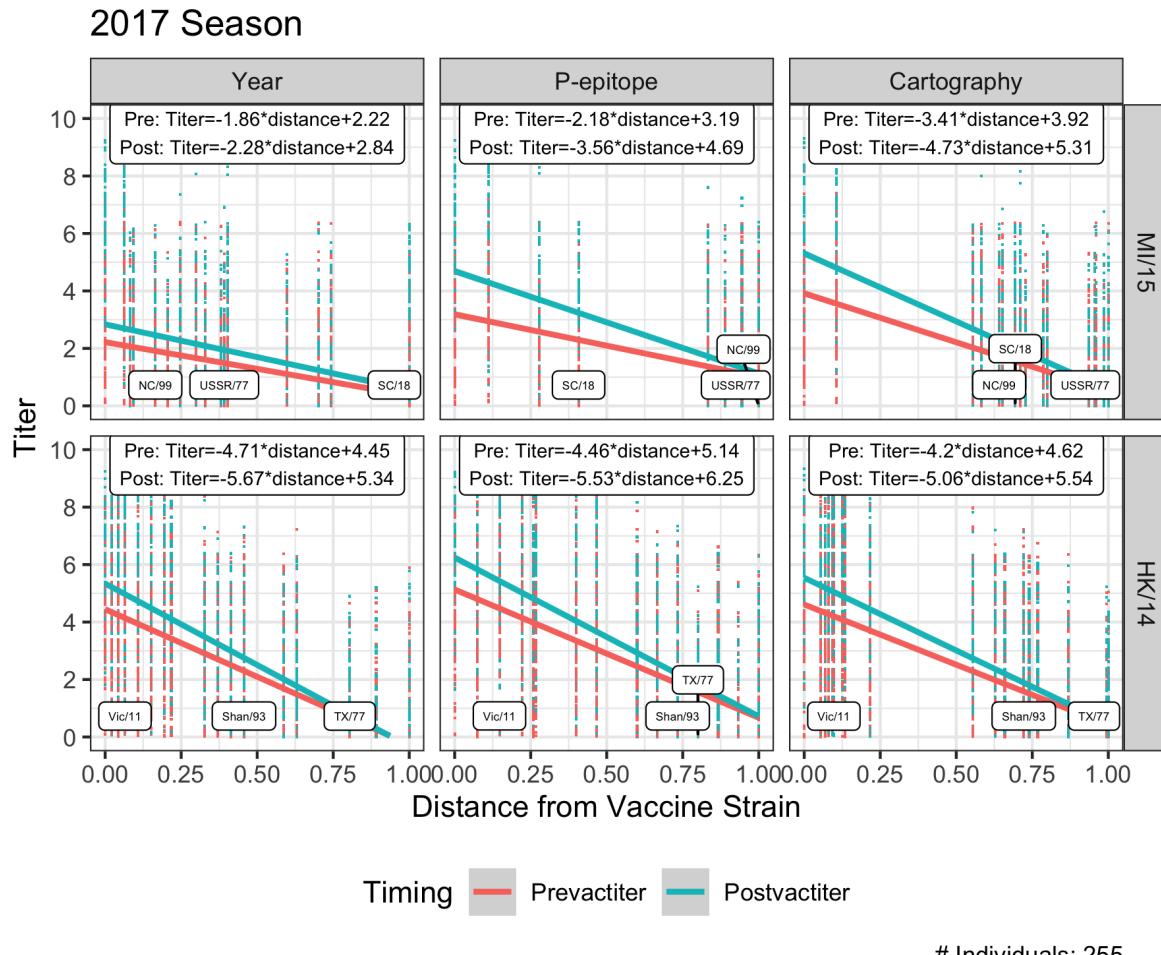


Figure 4.1: Pre- and post- vaccination titers to the H₁N₁ and H₃N₂ virus panels by season of all individuals who received SD vaccination for the 2017 season (sample size per season included). The vaccination strain of each subtype per season is shown in the row label. The linear regression of the pre- and post-vaccination titers with 95% confidence intervals are shown. The three distance measures are shown in the column label. The resulting linear regression equations of the pre- and post-vaccination linear regressions are included. Select viruses are included to depict the change in positioning based upon distance measure. The distances were normalized by season (Equation 2.7).

4.2 Titer Increase

Pre- and post-vaccination titers provide information regarding where the individual's titer begins and ends. However, since change in titer after vaccination is the primary outcome of interest, the titer increase was also investigated.

4.2.1 Season level analysis

When comparing titer increase to relative distance, the relationship previously observed with pre- and post-vaccination remained. Initially, each season was analyzed separately. For the 2017 season, the p-epitope and cartographic sequence measures had a stronger association with the titer increase induced by the H1N1 vaccine strain (Figure 4.2.1). Whereas for the H3N2, the three measures were again similar. The other season results are shown in Supplementary Figure B.

4.2.2 Strain level analysis

The influenza vaccine strains were included in multiple seasons. Therefore, the breadth of a vaccine strain over the tested HAI panel was investigated. HAI panels varied across vaccine strains; as such, comparisons are to be made only across various distance measures, not across different strains. For instance, the CA/09 vaccine strain was tested against a broader panel than the Bris/18.

On a strain basis, the CA/09 and MI/15 strains had a stronger relationship between titer increase and distance from vaccine strain (Figure 4.2.2). For the Bris/18 strain, the linear regressions were similar across the different measures, and the Bris/18 strain panel was not as extensive as the other two strains. However, all three measures indicate an association between the titer increase and distance.

Of the H3N2 strains, distance measures remain similar across the different vaccine strains (Figure 4.2.2). The strain-based results are nearly identical to the season-based results due to the H3N2 vaccine component changing every year except for 2016 and 2017, where HK/14 was the vaccine strain.

2017 Season

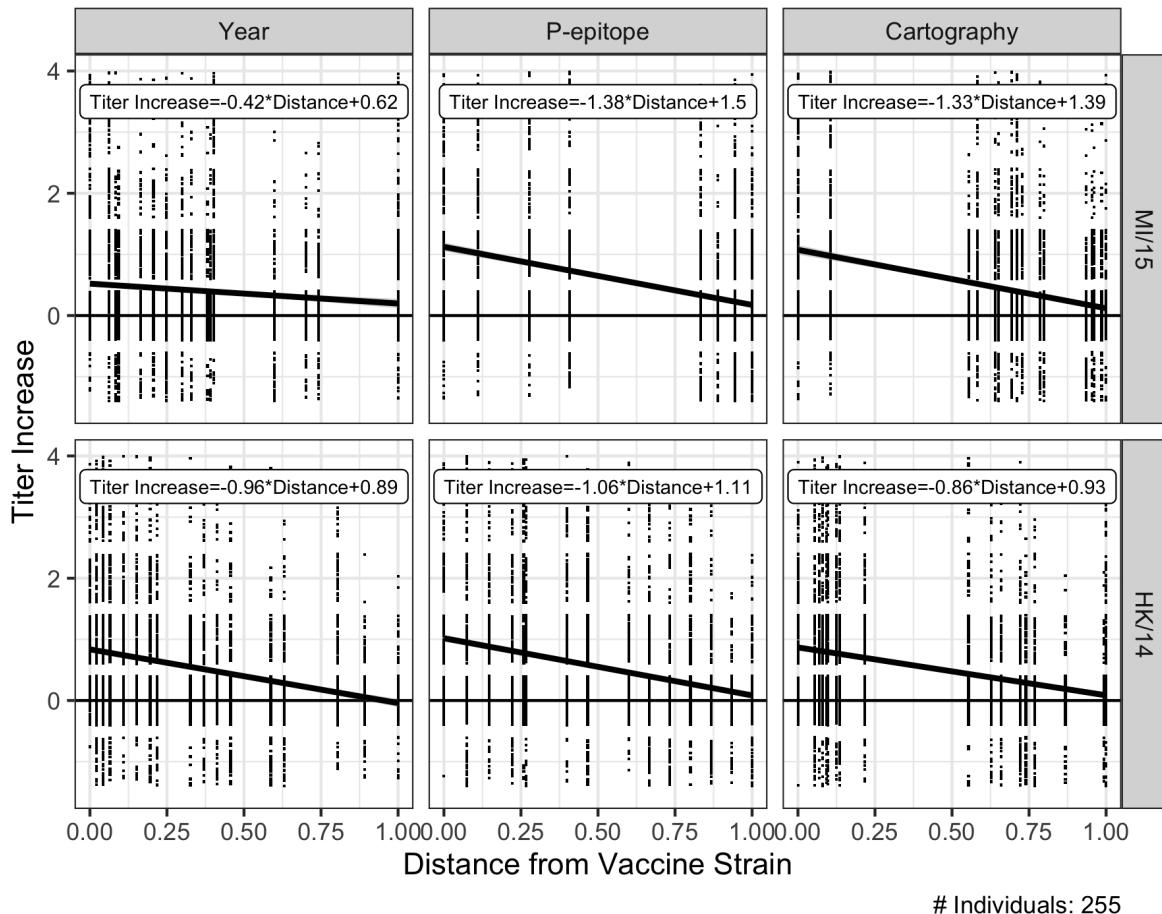


Figure 4.2: The titer increase to the H₁N₁ and H₃N₂ virus panels for the 2017 season of all individuals who received SD vaccination. The columns are separated by vaccine strain. The linear regression of the titer increase with 95% confidence intervals is shown for each distance method. The distances were normalized by season (Equation 2.7). The raw data points had jitter applied with +/- 0.4 in the y-axis. Raw data points that fell outside of the y-axis bounds are not shown.

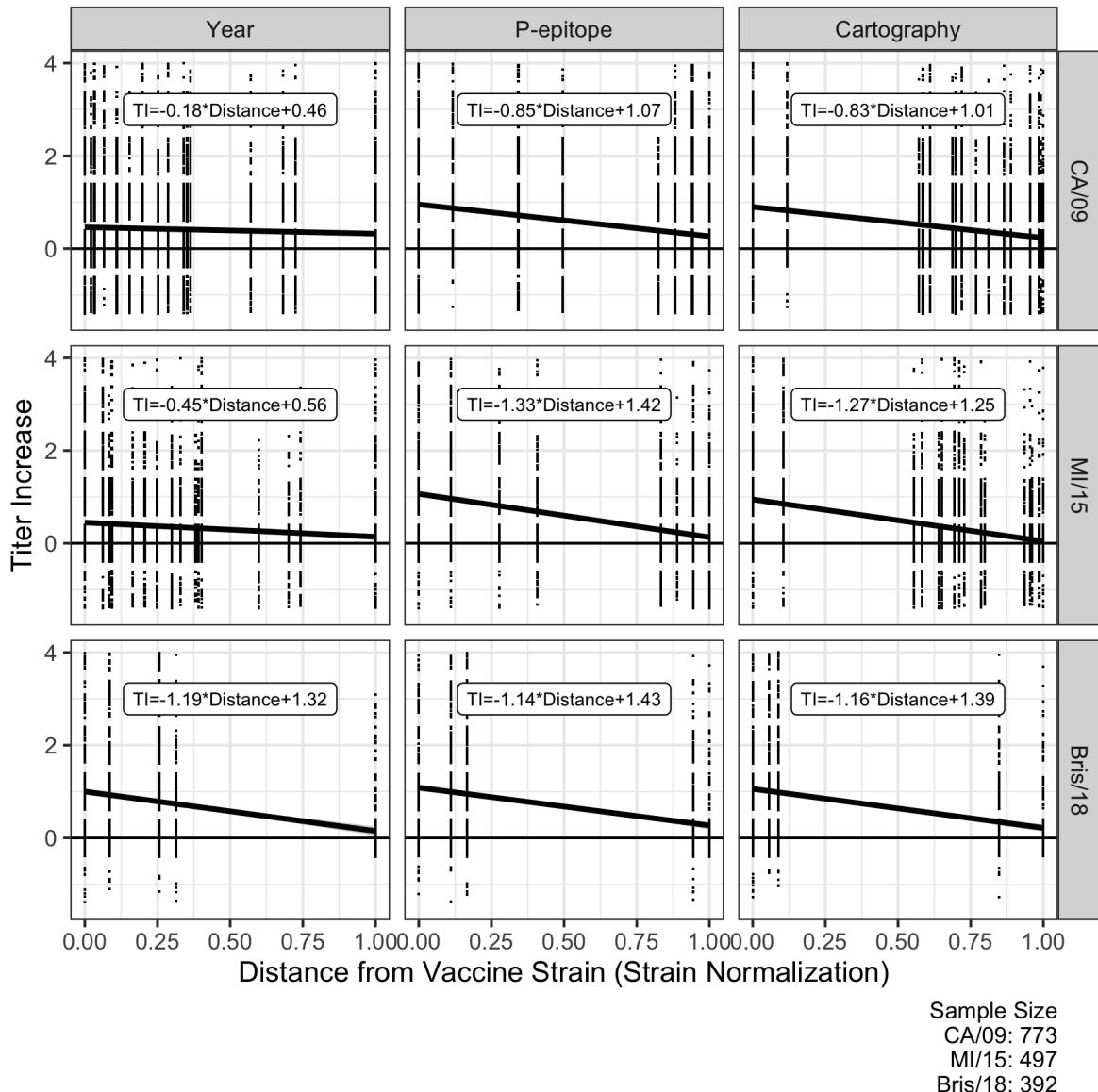


Figure 4.3: The titer increase to the H1N1 vaccine viruses, including all individuals who received SD vaccination. The columns are separated by vaccine strain and include all seasons that strain was used as the vaccine virus. The linear regression of the titer increase with 95% confidence intervals is shown for each distance method. The distances were normalized by strain (Equation 2.8). The raw data points had jitter applied with +/- 0.4 in the y-axis. Raw data points that fell outside of the y-axis bounds are not shown.

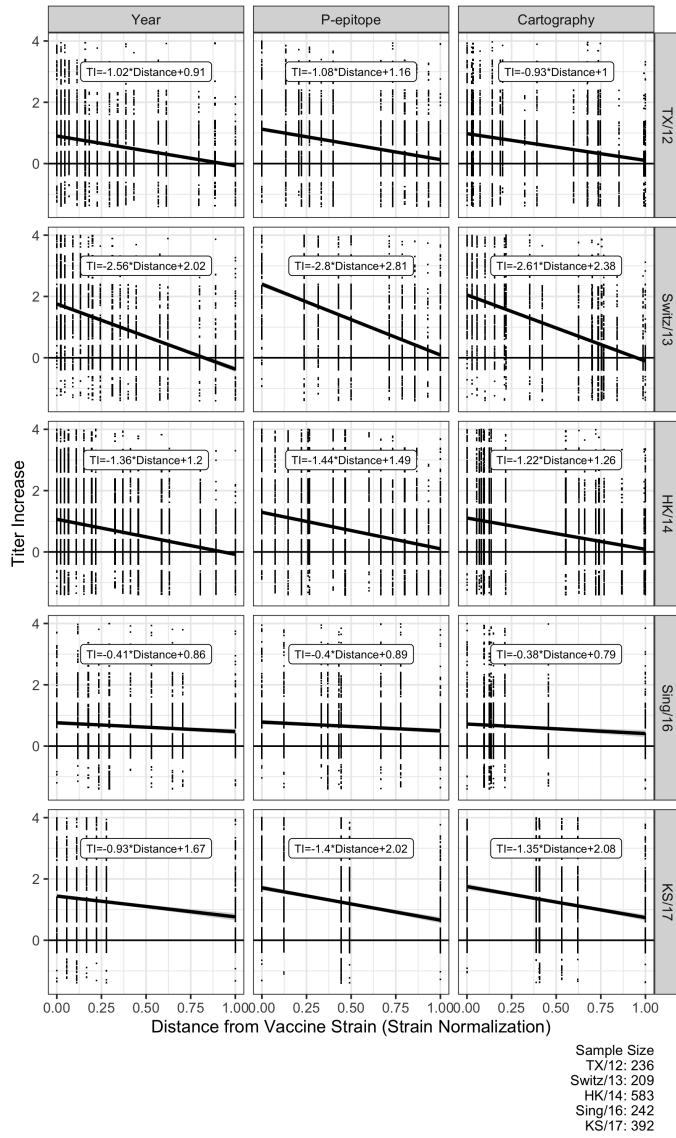


Figure 4.4: The titer increase to the H₃N₂ vaccine viruses, including all individuals who received SD vaccination. The columns are separated by vaccine strain and include all seasons that strain was used as the vaccine virus. The linear regression of the titer increase with 95% confidence intervals is shown for each distance method. The distances were normalized by strain (Equation 2.8). The raw data points had jitter applied with +/- 0.4 in the y-axis. Raw data points that fell outside of the y-axis bounds are not shown.

4.3 Quantification of breadth and application of weighting

Although the linear regression provides information regarding the association of the distance measures and vaccine outcomes, the area under the curve was used to quantify the breadth of the vaccine response. The area included all areas bounded between the minimum and maximum distance. For the 2017 SD cohort, the AUC was greater for the p-epitope and cartography measures, as expected for H1N1 (Table 4.3). For the H3N2, the p-epitope and cartography measures are still greater than a year, although by a smaller magnitude.

Table 4.1: Area under the curve (AUC) of the 2017 season with different weighting schemes applied

		AUC		
Vaccine	Method	Unweighted	Linear	2 AU
MI/15	Year	0.409	0.240	0.223
	P-epitope	0.812	0.521	0.509
	Cartography	0.726	0.474	0.466
HK/14	Year	0.409	0.285	0.235
	P-epitope	0.578	0.378	0.299
	Cartography	0.495	0.319	0.251

AU: Antigenic Distance

Distance normalization by strain (Equation 2.7)

Different weighting approaches were applied to the AUC (Figure 4.3). These weightings schemes are examples of different techniques that could be applied to data. The first is unweighted (original AUC), representing a scenario where all viruses in the panel are equally important in terms of eliciting an antibody response (Figure 4.3A). The second scenario places a linearly decreasing weight on the virus panel. Distances that are closer to the vaccine have a higher weight, whereas viruses further away are weighted less (Figure 4.3B). The third scenario is a step function based on the antigenic distance of the viruses. This method is to, for example, determine the vaccine breadth within a specified range (Figure 4.3). Since two antigenic units define antigenically similar viruses, it was used as the cutoff for the examples. The weighting schemes were applied to the SD 2017 cohort (Table 4.3) and for each strain (Table 4.3).

Table 4.2: Area under the curve (AUC) of different distance measures by influenza vaccine strain with different weighting schemes applied.

Subtype	Vaccine	Method	AUC		
			Unweighted	Linear	ω AU
H1N1	CA/09	Year	0.374	0.202	0.198
		P-epitope	0.645	0.394	0.409
		Cartography	0.592	0.365	0.381
	MI/15	Year	0.332	0.204	0.194
		P-epitope	0.756	0.489	0.479
		Cartography	0.612	0.412	0.411
	Bris/18	Year	0.729	0.464	0.498
		P-epitope	0.866	0.528	0.558
		Cartography	0.812	0.503	0.534
H3N2	TX/12	Year	0.402	0.286	0.267
		P-epitope	0.624	0.402	0.355
		Cartography	0.531	0.343	0.304
	Switz/13	Year	0.743	0.585	0.572
		P-epitope	1.404	0.936	0.844
		Cartography	1.072	0.754	0.699
	HK/14	Year	0.524	0.375	0.314
		P-epitope	0.768	0.504	0.401
		Cartography	0.646	0.425	0.338
	Sing/16	Year	0.657	0.362	0.431
		P-epitope	0.690	0.378	0.449
		Cartography	0.597	0.330	0.393
	KS/17	Year	1.209	0.682	0.792
		P-epitope	1.319	0.776	0.912
		Cartography	1.401	0.813	0.952

Distance normalization by strain (Equation 2.8)

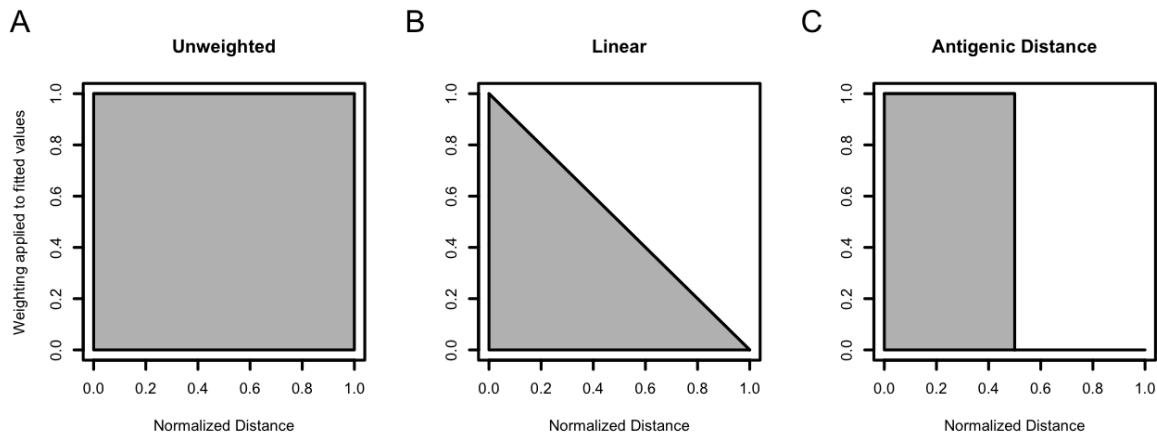


Figure 4.5: Examples of the weighting schemes applied to data. Weighting of the results can emphasize the desired antibody responses while reducing the emphasis on undesired ones. Three schemes presented here included the A) unweighted scheme, where each distance value is equally weighted after fitting. B) The linear scheme weights the antibody responses at the minimum distance one and the antibody responses at the maximum distance zero. The furthest virus antibody response is thus reduced to have no impact on the breadth measurement. In a negative linear relationship, the contributions of the viruses with increasing distance have less weight. C) The antigenic unit step weighting is based on the two antigenic units used for vaccine selection. For viruses within two antigenic units of the vaccine strain, they are weighted at one. The strains outside of the cutoff have a final weight of zero.

4.4 Case-study

Overall it was found that the p-epitope and cartographic distances were more appropriate for determining quantification of the breadth. Although it is known that the HD vaccine elicits a stronger antibody response to the vaccine strain, the comparative breadth is not well characterized. Therefore, these distance measures and breadth quantification were applied to compare HD and SD vaccines to determine differences in breadth.

Individuals with age greater than or equal to 65 were included in the dose analysis (Table 2.2). The titer increase for the different seasons (Figure 4.4 and Supplementary Figure B) was quantified for the area under the curve and with the different weightings applied (Table 4.4). Within the 2017 season, the HD vaccine had greater breadth than the SD vaccine. Even with different weightings, the HD vaccine had

a greater AUC than SD for both H₁N₁ and H₃N₂ components. Overall, the p-epitope and cartography AUC were greater than the year AUC, similar to the previous SD only measurements.

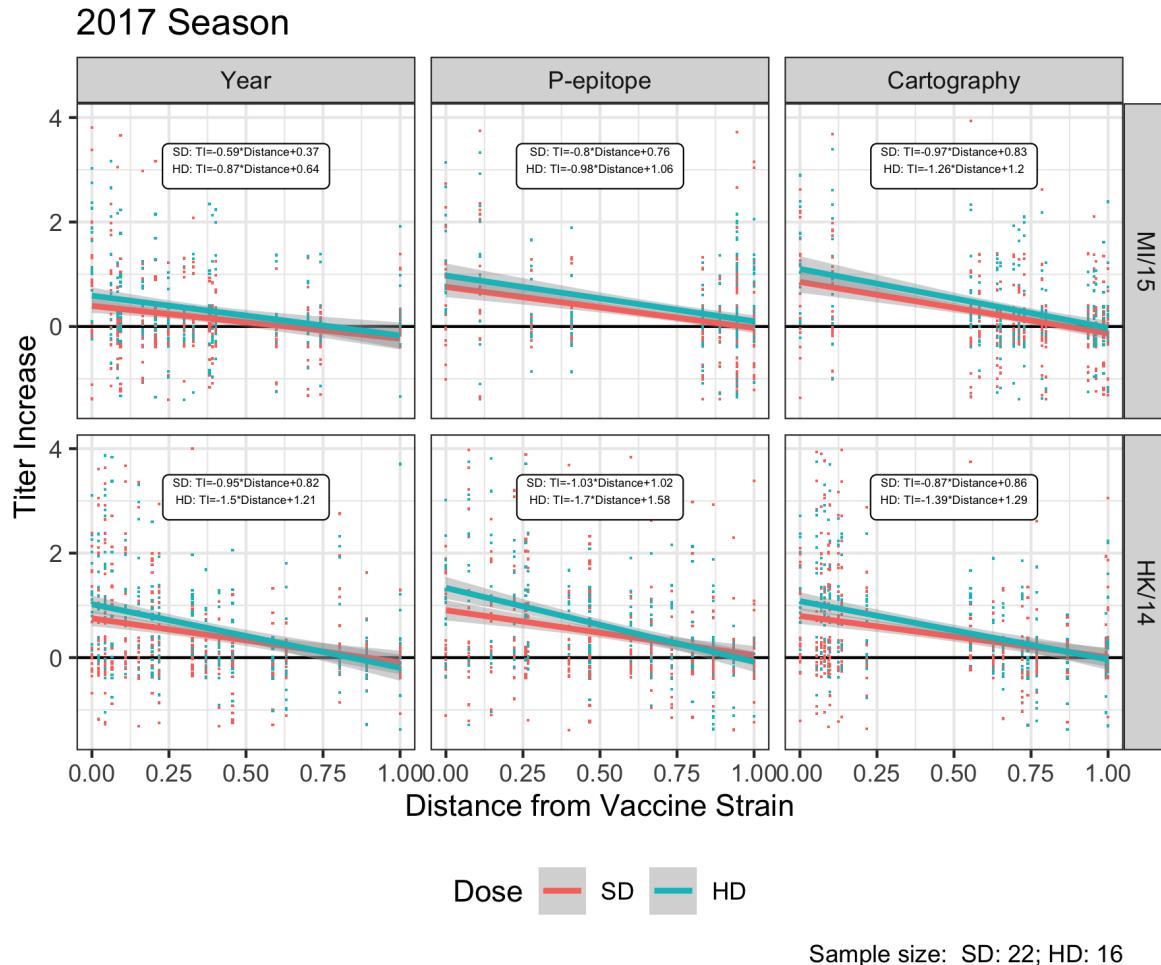


Figure 4.6: The titer increase (TI) to the H₁N₁ and H₃N₂ virus panels for the 2017 season of all individuals greater than or equal to 65 years of age. The columns are separated by type of distance measure. The linear regression of the titer increase with 95% confidence intervals is shown for the dose received. The distances were normalized by season (Equation 2.7). The black line indicates no titer increase. The raw data points had jitter applied with +/- 0.4 in the y-axis. Raw data points that fell outside of the y-axis bounds are not shown.

On a strain level, the CA/09 and MI/15 strains elicited a greater breadth in HD vaccine (Table 4.4). Whereas, for the Bris/18 strain, the SD formulation elicited a greater breadth with the unweighted changes (Year: -0.189; P-epitope: -0.184; Cartography: -0.186). These values were consistent across measures. For the H₃N₂ vaccine strains, the TX/12 HD vaccine strains elicited an increased breadth. Whereas the other

Table 4.3: Area under the curve for three weighting schemes for 65+ individuals for the 2017 season

Vaccine	Method	Unweighted			Linear			2 Antigenic Unit		
		HD	SD	Change	HD	SD	Change	HD	SD	Change
MI/15	Year	0.203	0.077	0.126	0.174	0.088	0.086	0.191	0.104	0.087
	P-epitope	0.570	0.360	0.210	0.366	0.247	0.119	0.358	0.249	0.109
	Cartography	0.570	0.349	0.221	0.390	0.255	0.135	0.392	0.264	0.128
HK/14	Year	0.465	0.341	0.124	0.357	0.250	0.107	0.310	0.212	0.098
	P-epitope	0.729	0.507	0.222	0.506	0.340	0.166	0.416	0.273	0.143
	Cartography	0.595	0.424	0.171	0.413	0.285	0.128	0.340	0.230	0.110

HD: High dose

SD: Standard dose

Distance normalization by season (Equation 2.7)

strains, Switz/13, HK/14, and Sing/16, had minimal change in breadth, with KS/17 strain SD having a greater response.

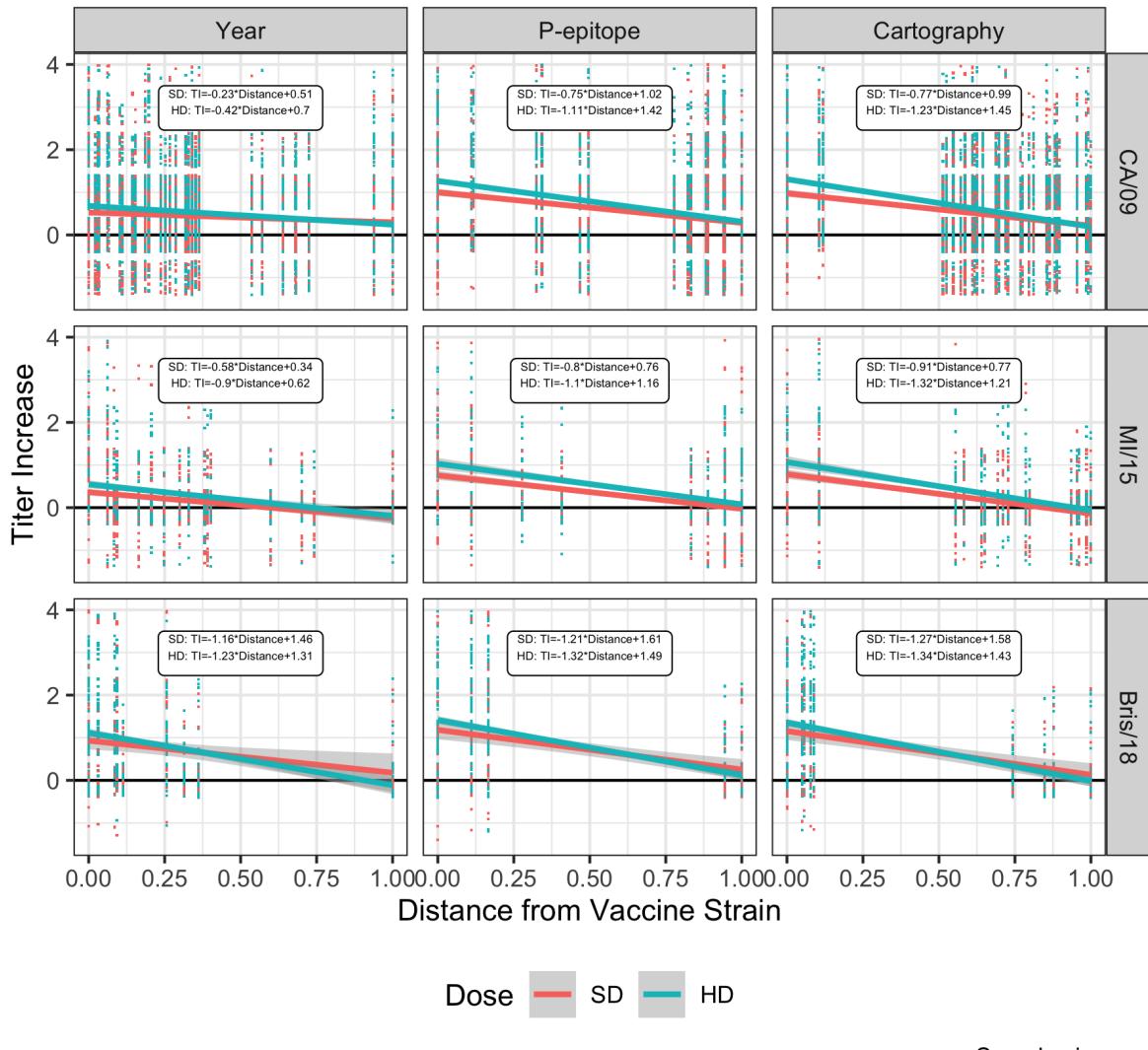


Figure 4.7: The titer increase to the H1N1 strains for HD and SD recipients. The columns are separated by vaccine strain and include all seasons that strain was used as the vaccine virus. The linear regression of the titer increase with 95% confidence intervals is shown for each distance method. The distances were normalized by strain (Equation 2.8). The raw data points had jitter applied with +/- 0.4 in the y-axis. Raw data points that fell outside of the y-axis bounds are not shown.

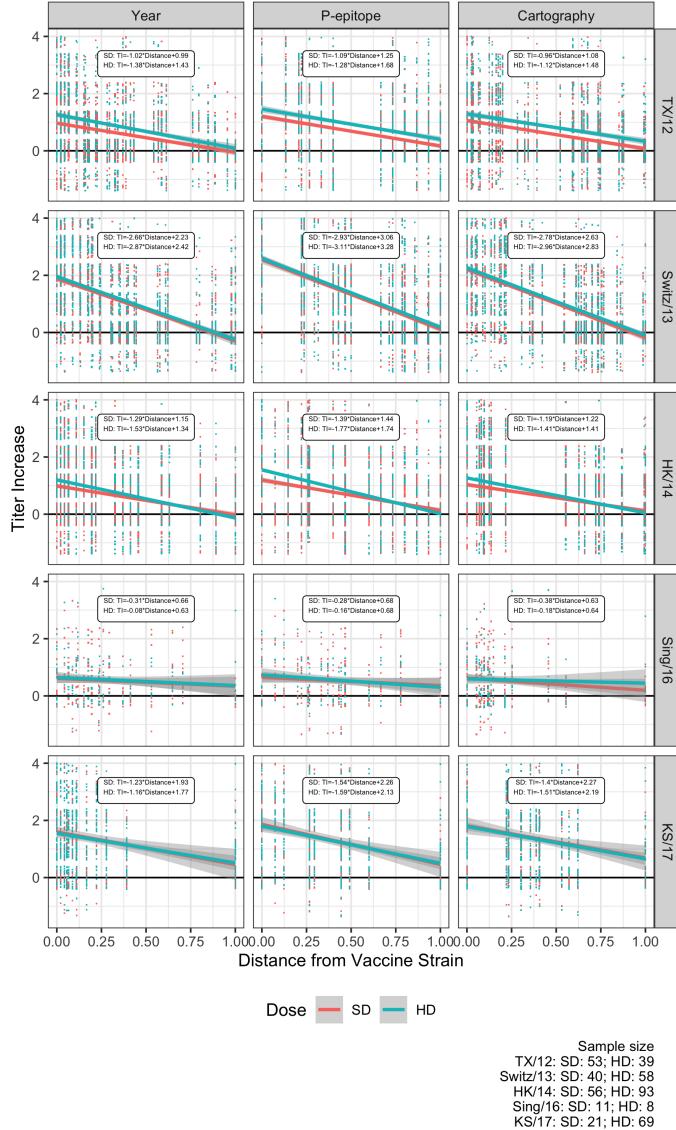


Figure 4.8: The titer increase to the H₃N₂ strains for HD and SD recipients. The columns are separated by vaccine strain and include all seasons that strain was used as the vaccine virus. The linear regression of the titer increase with 95% confidence intervals is shown for each distance method. The distances were normalized by strain (Equation 2.8). The raw data points had jitter applied with +/- 0.4 in the y-axis. Raw data points that fell outside of the y-axis bounds are not shown.

Table 4.4: Area under the curve for three weighting schemes for 65+ individuals for different strains

Subtype	Vaccine	Method	Unweighted			Linear			2 Antigenic Units		
			HD	SD	Change	HD	SD	Change	HD	SD	Change
CA/09	Year		0.486	0.394	0.092	0.278	0.216	0.062	0.281	0.213	0.068
	P-epitope		0.868	0.647	0.221	0.526	0.386	0.140	0.546	0.397	0.149
	Cartography		0.834	0.609	0.225	0.519	0.368	0.151	0.545	0.382	0.163
HnN1	Year		0.169	0.052	0.117	0.159	0.074	0.085	0.180	0.092	0.088
	P-epitope		0.609	0.358	0.251	0.396	0.246	0.150	0.390	0.248	0.142
	Cartography		0.550	0.309	0.241	0.385	0.231	0.154	0.391	0.241	0.150
Bris/18	Year		0.691	0.880	-0.189	0.448	0.537	-0.089	0.485	0.567	-0.082
	P-epitope		0.825	1.009	-0.184	0.523	0.605	-0.082	0.561	0.635	-0.074
	Cartography		0.762	0.948	-0.186	0.493	0.579	-0.086	0.533	0.613	-0.080
TX/12	Year		0.744	0.478	0.266	0.487	0.324	0.163	0.434	0.295	0.139
	P-epitope		1.034	0.701	0.333	0.624	0.441	0.183	0.531	0.385	0.146
	Cartography		0.923	0.606	0.317	0.555	0.383	0.172	0.471	0.335	0.136
Switz/13	Year		0.982	0.904	0.078	0.730	0.673	0.057	0.696	0.642	0.054
	P-epitope		1.719	1.592	0.127	1.119	1.040	0.079	0.996	0.928	0.068
	Cartography		1.350	1.244	0.106	0.921	0.854	0.067	0.842	0.782	0.060
H ₃ N ₂	Year		0.573	0.508	0.065	0.414	0.362	0.052	0.348	0.302	0.046
	P-epitope		0.859	0.741	0.117	0.477	0.487	0.090	0.465	0.387	0.078
	Cartography		0.709	0.622	0.087	0.472	0.410	0.062	0.379	0.327	0.052
Sing/16	Year		0.593	0.509	0.084	0.303	0.280	0.023	0.354	0.333	0.021
	P-epitope		0.599	0.534	0.065	0.312	0.291	0.021	0.367	0.345	0.022
	Cartography		0.556	0.441	0.115	0.292	0.252	0.040	0.344	0.302	0.042
KS/17	Year		1.189	1.313	-0.124	0.691	0.759	-0.068	0.810	0.887	-0.077
	P-epitope		1.340	1.484	-0.144	0.803	0.870	-0.067	0.947	1.022	-0.075
	Cartography		1.434	1.576	-0.142	0.843	0.904	-0.061	0.989	1.056	-0.067

Distance normalization by strain (Equation 2.8)

4.5 Limitations

One of the limitations of using titer increase and post-vaccination titer as a measurement is a physiological maximum. A maximum threshold of HAI titer limits the antibody response observed (Attias et al., 2021; Krammer et al., 2021). Therefore, if an individual has pre-existing titers to a virus through multiple influenza exposures, a smaller titer increase can be observed than a naive individual. This smaller increase may have no bearing on the efficacy of the vaccine. To overcome this limitation, future studies can investigate the outcome of seroconversion, which is an HAI antibody titer greater than or equal to

I:40.

CHAPTER 5

DISCUSSION

Overall, sequence- or cartographic-based distances provided the strongest association with vaccine outcome than the year-based measure. Further, even though the three measures varied minimally between some strains (H_3N_2 strains) for some strains, the year method never produced a greater association than the other two. Therefore, the sequence and cartographic measures performed equal to or better than the year-based measure. Consequently, it was determined that sequence and cartographic methods are preferred over using the year of isolation.

The year of isolation correlates highly with the sequence and cartographic measures for the H_3N_2 viruses. The H_3N_2 viruses have undergone constant antigenic drift over time with no antigenic shifts, leading to this correlation. For now, using the time-based measure yields similar results as the sequence and cartographic measures. However, only human isolated viruses that underwent this antigenic drift were included in the panel. Developing a vaccine that protects from potential zoonotic pandemics requires a virus panel that includes viruses isolated from animal reservoirs. Antigenically distinct H_3 influenza is common in the swine population, and variants have been isolated from human infection (Bangaru et al., 2016). These viruses would overlap in time of isolation with the human isolated viruses; the year, sequence, and cartographic values would no longer correlate. Hence, although the different measures were not significantly different in this study, these methods would be preferred in a study looking at a more comprehensive panel.

The differences observed between the sequence- and cartographic-based methods were minimal for both the H₁N₁ and H₃N₂ for season and strain comparisons. In fact, for the H₃N₂ subtype, the p-epitope measure had consistently greater AUC than the cartographic measure. It is reassuring that the sequence-based method provided results consistent with the cartographic ones due to the number of resources necessary for calculating the cartographies. The generation of sera and testing the panel of viruses provides direct antigenic data but is not feasible to conduct for a large-scale project. Therefore, using the sequence data, even though there is no perfect correlation with antigenicity, may provide an appropriate surrogate. Further, the HAI assay results were used as the outcome in this study. HAI cartography may not be suitable if using a different assay such as total antibody binding data (HAI only measures RBS binding antibodies), and new cartographies would need to be created to match the outcome.

Vaccine candidates and immune responses of interest are becoming more diverse. Many research groups are investigating multiple subtype breadth and breadth with other proteins such as the NA (Chen et al., 2018; Skarlupka et al., 2021). These methods can be applied to these components as well. Overall, the techniques described here can assist in selecting vaccine candidates, delivery platforms, and adjuvants when the goal is to expand the breadth of the immune response. Adopting an objective, reproducible breadth value would be beneficial for the research community.

BIBLIOGRAPHY

- Adabor, E. S., & Ndifon, W. (2018). Bayesian inference of antigenic and non-antigenic variables from haemagglutination inhibition assays for influenza surveillance. *Royal Society Open Science*, 5(7), 180113. <https://doi.org/10.1098/rsos.180113>
- Allen, J. D., & Ross, T. M. (2021). Evaluation of next-generation h3 influenza vaccines in ferrets pre-immune to historical h3n2 viruses. *Frontiers in Immunology*, 12. <https://doi.org/10.3389/fimmu.2021.707339>
- Anderson, C. S., McCall, P. R., Stern, H. A., Yang, H., & Topham, D. J. (2018). Antigenic cartography of h1n1 influenza viruses using sequence-based antigenic distance calculation. *BMC Bioinformatics*, 19(1). <https://doi.org/10.1186/s12859-018-2042-4>
- Anderson, C. S., Sangster, M. Y., Yang, H., Mariani, T. J., Chaudhury, S., & Topham, D. J. (2020). Implementing sequence-based antigenic distance calculation into immunological shape space model. *BMC Bioinformatics*, 21(1). <https://doi.org/10.1186/s12859-020-03594-3>
- Andrews, S. F., Raab, J. E., Gorman, J., Gillespie, R. A., Cheung, C. S. F., Rawi, R., Cominsky, L. Y., Boyington, J. C., Creanga, A., Shen, C.-H., Harris, D. R., Olia, A. S., Nazzari, A. F., Zhou, T., Houser, K. V., Chen, G. L., Mascola, J. R., Graham, B. S., Kanekiyo, M., ... McDermott, A. B. (2022). A single residue in influenza virus h2 hemagglutinin enhances the breadth of the b cell response elicited by h2 vaccination. *Nature Medicine*, 28(2), 373–382. <https://doi.org/10.1038/s41591-021-01636-8>
- Archetti, I., & Horsfall, F. L. (1950). Persistent antigenic variation of influenza a viruses after incomplete neutralization in ovo with heterologous immune serum. *Journal of Experimental Medicine*, 92(5), 441–462. <https://doi.org/10.1084/jem.92.5.441>
- Attias, P., Sakhi, H., Rieu, P., Soorkia, A., Assayag, D., Bouhroum, S., Nizard, P., & El Karoui, K. (2021). Antibody response to the bnt162b2 vaccine in maintenance hemodialysis patients. *Kidney International*, 99(6), 1490–1492. <https://doi.org/10.1016/j.kint.2021.04.009>
- Auladell, M., Phuong, H. V. M., Mai, L. T. Q., Tseng, Y.-Y., Carolan, L., Wilks, S., Thai, P. Q., Price, D., Duong, N. T., Hang, N. L. K., Thanh, L. T., Thuong, N. T. H., Huong, T. T. K., Diep, N. T. N., Bich, V. T. N., Khvorov, A., Hensen, L., Duong, T. N., Kedzierska, K., ... Fox, A. (2022). Influenza virus infection history shapes antibody responses to influenza vaccination. *Nature Medicine*, 28(2), 363–372. <https://doi.org/10.1038/s41591-022-01690-w>
- Bangaru, S., Nieusma, T., Kose, N., Thornburg, N. J., Finn, J. A., Kaplan, B. S., King, H. G., Singh, V., Lampley, R. M., Sapparapu, G., Cisneros, A., Edwards, K. M., Slaughter, J. C., Edupuganti, S., Lai, L., Richt, J. A., Webby, R. J., Ward, A. B., & Crowe, J. E. (2016). Recognition of influenza h3n2 variant virus by human neutralizing antibodies. *JCI Insight*, 1(10). <https://doi.org/10.1172/jci.insight.86673>

- Bedford, T., Riley, S., Barr, I. G., Broor, S., Chadha, M., Cox, N. J., Daniels, R. S., Gunasekaran, C. P., Hurt, A. C., Kelso, A., Klimov, A., Lewis, N. S., Li, X., McCauley, J. W., Odagiri, T., Potdar, V., Rambaut, A., Shu, Y., Skepner, E., ... Russell, C. A. (2015). Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature*, 523(7559), 217–220. <https://doi.org/10.1038/nature14460>
- Belongia, E. A., Skowronski, D. M., McLean, H. Q., Chambers, C., Sundaram, M. E., & De Serres, G. (2017). Repeated annual influenza vaccination and vaccine effectiveness: Review of evidence. *Expert Review of Vaccines*, 16(7), 723–736. <https://doi.org/10.1080/14760584.2017.1334554>
- Boyoglu-Barnum, S., Ellis, D., Gillespie, R. A., Hutchinson, G. B., Park, Y.-J., Moin, S. M., Acton, O. J., Ravichandran, R., Murphy, M., Pettie, D., Matheson, N., Carter, L., Creanga, A., Watson, M. J., Kephart, S., Ataca, S., Vaile, J. R., Ueda, G., Crank, M. C., ... Kanekiyo, M. (2021). Quadrivalent influenza nanoparticle vaccines induce broad protection. *Nature*, 592(7855), 623–628. <https://doi.org/10.1038/s41586-021-03365-x>
- Burke, D. F., & Smith, D. J. (2014). A recommended numbering scheme for influenza a ha subtypes (P. Digard, Ed.). *PLoS ONE*, 9(11), e112302. <https://doi.org/10.1371/journal.pone.0112302>
- Cai, Z., Zhang, T., & Wan, X.-F. (2010). A computational framework for influenza antigenic cartography (C. Fraser, Ed.). *PLoS Computational Biology*, 6(10), e1000949. <https://doi.org/10.1371/journal.pcbi.1000949>
- Cai, Z., Zhang, T., & Wan, X.-F. (2012). Antigenic distance measurements for seasonal influenza vaccine selection. *Vaccine*, 30(2), 448–453. <https://doi.org/10.1016/j.vaccine.2011.10.051>
- Cao, D.-S., Xiao, N., Xu, Q.-S., & Chen, A. F. (2015). Rcipi: R/bioconductor package to generate various descriptors of proteins, compounds and their interactions. 31. <https://doi.org/10.1093/bioinformatics/btu624>
- Castro, L. A., Bedford, T., & Ancel Meyers, L. (2020). Early prediction of antigenic transitions for influenza a/h3n2 (R. D. Kouyos, Ed.). *PLOS Computational Biology*, 16(2), e1007683. <https://doi.org/10.1371/journal.pcbi.1007683>
- Charif, D., & Lobry, J. (2007). Seqin{r} 1.0-2: A contributed package to the {r} project for statistical computing devoted to biological sequences retrieval and analysis. *Structural approaches to sequence evolution: Molecules, networks, populations*, 207–232.
- Chen, Y.-Q., Wohlbold, T. J., Zheng, N.-Y., Huang, M., Huang, Y., Neu, K. E., Lee, J., Wan, H., Rojas, K. T., Kirkpatrick, E., Henry, C., Palm, A.-K. E., Stamper, C. T., Lan, L. Y.-L., Topham, D. J., Treanor, J., Wrammert, J., Ahmed, R., Eichelberger, M. C., ... Wilson, P. C. (2018). Influenza infection in humans induces broadly cross-reactive and protective neuraminidase-reactive antibodies. *Cell*, 173(2), 417–429.e10. <https://doi.org/10.1016/j.cell.2018.03.030>
- Deem, M. W., & Pan, K. (2009). The epitope regions of h1-subtype influenza a, with application to vaccine efficacy. *Protein Engineering Design and Selection*, 22(9), 543–546. <https://doi.org/10.1093/protein/gzp027>
- Dong, W., Bhide, Y., Sicca, F., Meijerhof, T., Guilfoyle, K., Engelhardt, O. G., Boon, L., de Haan, C. A. M., Carnell, G., Temperton, N., de Vries-Idema, J., Kelvin, D., & Huckriede, A. (2018). Cross-protective immune responses induced by sequential influenza virus infection and by sequential vaccination with inactivated influenza vaccines. *Frontiers in Immunology*, 9. <https://doi.org/10.3389/fimmu.2018.02312>

- Dugan, H. L., Guthmiller, J. J., Arevalo, P., Huang, M., Chen, Y.-Q., Neu, K. E., Henry, C., Zheng, N.-Y., Lan, L. Y.-L., Tepora, M. E., Stovicek, O., Bitar, D., Palm, A.-K. E., Stamper, C. T., Changrob, S., Utset, H. A., Coughlan, L., Krammer, F., Cobey, S., & Wilson, P. C. (2020). Preexisting immunity shapes distinct antibody landscapes after influenza virus infection and vaccination in humans. *Science Translational Medicine*, 12(573). <https://doi.org/10.1126/scitranslmed.abd3601>
- Durviaux, S., Treanor, J., Beran, J., Duval, X., Esen, M., Feldman, G., Frey, S. E., Launay, O., Leroux-Roels, G., McElhaney, J. E., Nowakowski, A., Ruiz-Palacios, G. M., van Essen, G. A., Oostvogels, L., Devaster, J.-M., & Walravens, K. (2014). Genetic and antigenic typing of seasonal influenza virus breakthrough cases from a 2008-2009 vaccine efficacy trial (R. L. Hodinka, Ed.). *Clinical and Vaccine Immunology*, 21(3), 271–279. <https://doi.org/10.1128/cvi.00544-13>
- Elbe, S., & Buckland-Merrett, G. (2017). Data, disease and diplomacy: Gisaid's innovative contribution to global health. *Global Challenges*, 1(1), 33–46. <https://doi.org/10.1002/gch2.1018>
- Fonville, J. M., Wilks, S. H., James, S. L., Fox, A., Ventresca, M., Aban, M., Xue, L., Jones, T. C., H., L. N. M., T., P. Q., D., T. N., Wong, Y., Mosterin, A., Katzelnick, L. C., Labonte, D., T., L. T., van der Net, G., Skepner, E., Russell, C. A., ... Smith, D. J. (2014). Antibody landscapes after influenza virus infection or vaccination. *Science*, 346(6212), 996–1000. <https://doi.org/10.1126/science.1256427>
- Fonville, J. M., Fraaij, P. L. A., de Mutsert, G., Wilks, S. H., van Beek, R., Fouchier, R. A. M., & Rimmelzwaan, G. F. (2015). Antigenic maps of influenza a(h3n2) produced with human antisera obtained after primary infection. *Journal of Infectious Diseases*, 213(1), 31–38. <https://doi.org/10.1093/infdis/jiv367>
- Forghani, M., & Khachay, M. (2020). Convolutional neural network based approach to in silico non-anticipating prediction of antigenic distance for influenza virus. *Viruses*, 12(9), 1019. <https://doi.org/10.3390/v12091019>
- Gostic, K. M., Ambrose, M., Worobey, M., & Lloyd-Smith, J. O. (2016). Potent protection against h5n1 and h7n9 influenza via childhood hemagglutinin imprinting. *Science*, 354(6313), 722–726. <https://doi.org/10.1126/science.aag1322>
- Gostic, K. M., Bridge, R., Brady, S., Viboud, C., Worobey, M., & Lloyd-Smith, J. O. (2019). Childhood immune imprinting to influenza a shapes birth year-specific risk during seasonal h1n1 and h3n2 epidemics (S. L. Klein, Ed.). *PLOS Pathogens*, 15(12), e1008109. <https://doi.org/10.1371/journal.ppat.1008109>
- Gupta, V., Earl, D. J., & Deem, M. W. (2006). Quantifying influenza vaccine efficacy and antigenic distance. *Vaccine*, 24(18), 3881–3888. <https://doi.org/10.1016/j.vaccine.2006.01.010>
- Hay, J. A., Laurie, K., White, M., & Riley, S. (2019). Characterising antibody kinetics from multiple influenza infection and vaccination events in ferrets (R. Antia, Ed.). *PLOS Computational Biology*, 15(8), e1007294. <https://doi.org/10.1371/journal.pcbi.1007294>
- Hay, J. A., Minter, A., Ainslie, K. E. C., Lessler, J., Yang, B., Cummings, D. A. T., Kucharski, A. J., & Riley, S. (2020). An open source tool to infer epidemiological and immunological dynamics from serological data: Serosolver (R. R. Regoes, Ed.). *PLOS Computational Biology*, 16(5), e1007840. <https://doi.org/10.1371/journal.pcbi.1007840>
- Hinojosa, M., Shepard, S. S., Chung, J. R., King, J. P., McLean, H. Q., Flannery, B., Belongia, E. A., & Levine, M. Z. (2020). Impact of immune priming, vaccination, and infection on influenza

- a(h3n2) antibody landscapes in children. *The Journal of Infectious Diseases*. <https://doi.org/10.1093/infdis/jiaa665>
- Hobson, D., Curry, R. L., Beare, A. S., & Ward-Gardner, A. (1972). The role of serum haemagglutination-inhibiting antibody in protection against challenge infection with influenza a2 and b viruses. *Epidemiology and Infection*, 70(4), 767–777. <https://doi.org/10.1017/s0022172400022610>
- Jang, H., & Ross, T. M. (2021). Hemagglutination inhibition (hai) antibody landscapes after vaccination with h7nx virus like particles (V. C. Huber, Ed.). *PLOS ONE*, 16(3), e0246613. <https://doi.org/10.1371/journal.pone.0246613>
- Katzelnick, L. C., Coello Escoto, A., Huang, A. T., Garcia-Carreras, B., Chowdhury, N., Maljkovic Berry, I., Chavez, C., Buchy, P., Duong, V., Dussart, P., Gromowski, G., Macareo, L., Thaisomboonsuk, B., Fernandez, S., Smith, D. J., Jarman, R., Whitehead, S. S., Salje, H., & Cummings, D. A. T. (2021). Antigenic evolution of dengue viruses over 20 years. *Science*, 374(6570), 999–1004. <https://doi.org/10.1126/science.abk0058>
- Kendra, J. A., Tohma, K., Ford-Siltz, L. A., Lepore, C. J., & Parra, G. I. (2021). Antigenic cartography reveals complexities of genetic determinants that lead to antigenic differences among pandemic gii.4 noroviruses. *Proceedings of the National Academy of Sciences*, 118(11). <https://doi.org/10.1073/pnas.2015874118>
- Krammer, F., Srivastava, K., Alshammary, H., Amoako, A. A., Awawda, M. H., Beach, K. F., Bermúdez-González, M. C., Bielak, D. A., Carreño, J. M., Chernet, R. L., Eaker, L. Q., Ferreri, E. D., Floda, D. L., Gleason, C. R., Hamburger, J. Z., Jiang, K., Kleiner, G., Jurczyszak, D., Matthews, J. C., ... Simon, V. (2021). Antibody responses in seropositive persons after a single dose of sars-cov-2 mrna vaccine. *New England Journal of Medicine*, 384(14), 1372–1374. <https://doi.org/10.1056/nejmci2101667>
- Krammer, F., Weir, J. P., Engelhardt, O., Katz, J. M., & Cox, R. J. (2019). Meeting report and review: Immunological assays and correlates of protection for next-generation influenza vaccines. *Influenza and Other Respiratory Viruses*, 14(2), 237–243. <https://doi.org/10.1111/irv.12706>
- Kucharski, A. J., Lessler, J., Cummings, D. A. T., & Riley, S. (2018). Timescales of influenza a/h3n2 antibody dynamics (S. Rowland-Jones, Ed.). *PLOS Biology*, 16(8), e2004974. <https://doi.org/10.1371/journal.pbio.2004974>
- Lau, Y. C., Perera, R. A. P. M., Fang, V. J., Luk, L. H., Chu, D. K. W., Wu, P., Barr, I. G., Peiris, J. S. M., & Cowling, B. J. (2020). Variation by lineage in serum antibody responses to influenza b virus infections (R. Borrow, Ed.). *PLOS ONE*, 15(11), e0241693. <https://doi.org/10.1371/journal.pone.0241693>
- Lee, M.-S., & Chen, J. S.-E. (2004). Predicting antigenic variants of influenza a/h3n2 viruses. *Emerging Infectious Diseases*, 10(8), 1385–1390. <https://doi.org/10.3201/eid1008.040107>
- Li, G.-M., Chiu, C., Wrammert, J., McCausland, M., Andrews, S. F., Zheng, N.-Y., Lee, J.-H., Huang, M., Qu, X., Edupuganti, S., Mulligan, M., Das, S. R., Yewdell, J. W., Mehta, A. K., Wilson, P. C., & Ahmed, R. (2012). Pandemic h1n1 influenza vaccine induces a recall response in humans that favors broadly cross-reactive memory b cells. *Proceedings of the National Academy of Sciences*, 109(23), 9047–9052. <https://doi.org/10.1073/pnas.118979109>
- Li, L., Chang, D., Han, L., Zhang, X., Zaia, J., & Wan, X.-F. (2020). 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*, 21(1). <https://doi.org/10.1186/s12859-020-3527-5>
- Li, Z.-N., Liu, F., Gross, F. L., Kim, L., Ferdinands, J., Carney, P., Chang, J., Stevens, J., Tumpey, T., & Levine, M. Z. (2021). Antibody landscape analysis following influenza vaccination and natural infection in humans with a high-throughput multiplex influenza antibody detection assay (K. Subbarao, Ed.). *mBio*, 12(1). <https://doi.org/10.1128/mbio.02808-20>
- Lorusso, A., Vincent, A. L., Harland, M. L., Alt, D., Bayles, D. O., Swenson, S. L., Gramer, M. R., Russell, C. A., Smith, D. J., Lager, K. M., & Lewis, N. S. (2010). Genetic and antigenic characterization of h1 influenza viruses from united states swine from 2008. *Journal of General Virology*, 92(4), 919–930. <https://doi.org/10.1099/vir.0.027557-0>
- Mögling, R., Richard, M. J., Vliet, S. v. d., Beek, R. v., Schrauwen, E. J. A., Spronken, M. I., Rimmelzwaan, G. F., & Fouchier, R. A. M. (2017). Neuraminidase-mediated haemagglutination of recent human influenza a(h3n2) viruses is determined by arginine 150 flanking the neuraminidase catalytic site. *Journal of General Virology*, 98(6), 1274–1281. <https://doi.org/10.1099/jgv.0.000809>
- Munoz, E. T., & Deem, M. W. (2005). Epitope analysis for influenza vaccine design. *Vaccine*, 23(9), 1144–1148. <https://doi.org/10.1016/j.vaccine.2004.08.028>
- Nachbagauer, R., Choi, A., Hirsh, A., Margine, I., Iida, S., Barrera, A., Ferres, M., Albrecht, R. A., García-Sastre, A., Bouvier, N. M., Ito, K., Medina, R. A., Palese, P., & Krammer, F. (2017). Defining the antibody cross-reactome directed against the influenza virus surface glycoproteins. *Nature Immunology*, 18(4), 464–473. <https://doi.org/10.1038/ni.3684>
- Nachbagauer, R., Wohlbold, T. J., Hirsh, A., Hai, R., Sjursen, H., Palese, P., Cox, R. J., & Krammer, F. (2014). Induction of broadly reactive anti-hemagglutinin stalk antibodies by an h5ni vaccine in humans (D. S. Lyles, Ed.). *Journal of Virology*, 88(22), 13260–13268. <https://doi.org/10.1128/jvi.02133-14>
- Ndifon, W. (2011). New methods for analyzing serological data with applications to influenza surveillance. *Influenza and Other Respiratory Viruses*, 5(3), 206–212. <https://doi.org/10.1111/j.1750-2659.2010.00192.x>
- Ndifon, W., Dushoff, J., & Levin, S. A. (2009). On the use of hemagglutination-inhibition for influenza surveillance: Surveillance data are predictive of influenza vaccine effectiveness. *Vaccine*, 27(18), 2447–2452. <https://doi.org/10.1016/j.vaccine.2009.02.047>
- Neher, R. A., Bedford, T., Daniels, R. S., Russell, C. A., & Shraiman, B. I. (2016). Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. *Proceedings of the National Academy of Sciences*, 113(12). <https://doi.org/10.1073/pnas.1525578113>
- Pagès, H., Aboyoun, P., Gentleman, R., & DebRoy, S. (2021). Biostrings: Efficient manipulation of biological strings. <https://bioconductor.org/packages/Biostrings>
- Petrova, V. N., & Russell, C. A. (2017). The evolution of seasonal influenza viruses. *Nature Reviews Microbiology*, 16(1), 47–60. <https://doi.org/10.1038/nrmicro.2017.118>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Reneer, Z. B., Jamieson, P. J., Skarupka, A. L., Huang, Y., & Ross, T. M. (2020). Computationally optimized broadly reactive h2 ha influenza vaccines elicited broadly cross-reactive antibodies and protected mice from viral challenges (K. Subbarao, Ed.). *Journal of Virology*, 95(2). <https://doi.org/10.1128/jvi.01526-20>

- Reneer, Z. B., Skarupka, A. L., Jamieson, P. J., & Ross, T. M. (2021). Broadly reactive h2 hemagglutinin vaccines elicit cross-reactive antibodies in ferrets preimmune to seasonal influenza a viruses (M. J. Imperiale, Ed.). *mSphere*, 6(2). <https://doi.org/10.1128/msphere.00052-21>
- RStudio Team. (2022). *Rstudio: Integrated development environment for r*. RStudio, PBC. Boston, MA. <http://www.rstudio.com/>
- Scholtissek, C., Rohde, W., Von Hoyningen, V., & Rott, R. (1978). On the origin of the human influenza virus subtypes h2n2 and h3n2. *Virology*, 87(1), 13–20. [https://doi.org/10.1016/0042-6822\(78\)90153-8](https://doi.org/10.1016/0042-6822(78)90153-8)
- Skarupka, A. L., Bebin-Blackwell, A.-G., Sumner, S. F., & Ross, T. M. (2021). Universal influenza virus neuraminidase vaccine elicits protective immune responses against human seasonal and pre-pandemic strains (K. Subbarao, Ed.). *Journal of Virology*, 95(17). <https://doi.org/10.1128/jvi.00759-21>
- Skarupka, A. L., & Ross, T. M. (2020). Immune imprinting in the influenza ferret model. *Vaccines*, 8(2), 173. <https://doi.org/10.3390/vaccines8020173>
- Skowronski, D. M., Chambers, C., De Serres, G., Sabaiduc, S., Winter, A.-L., Dickinson, J. A., Gubbay, J. B., Fonseca, K., Drews, S. J., Charest, H., Martineau, C., Krajden, M., Petric, M., Bastien, N., Li, Y., & Smith, D. J. (2017). 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*, 215(7), 1059–1099. <https://doi.org/10.1093/infdis/jix074>
- Smith, D. J., Lapedes, A. S., de Jong, J. C., Bestebroer, T. M., Rimmelzwaan, G. F., Osterhaus, A. D. M. E., & Fouchier, R. A. M. (2004). Mapping the antigenic and genetic evolution of influenza virus. *Science*, 305(5682), 371–376. <https://doi.org/10.1126/science.1097211>
- Sun, H., Yang, J., Zhang, T., Long, L.-P., Jia, K., Yang, G., Webby, R. J., & Wan, X.-F. (2013). Using sequence data to infer the antigenicity of influenza virus (S. Perlman & C. Biron, Eds.). *mBio*, 4(4). <https://doi.org/10.1128/mbio.00230-13>
- Tapia, R., Torremorell, M., Culhane, M., Medina, R. A., & Neira, V. (2020). Antigenic characterization of novel h1 influenza a viruses in swine. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-61315-5>
- Tuju, J., Mackinnon, M. J., Abdi, A. I., Karanja, H., Musyoki, J. N., Warimwe, G. M., Gitau, E. N., Marsh, K., Bull, P. C., & Urban, B. C. (2019). Antigenic cartography of immune responses to plasmodium falciparum erythrocyte membrane protein 1 (pfemp1) (L. Hviid, Ed.). *PLOS Pathogens*, 15(7), e1007870. <https://doi.org/10.1371/journal.ppat.1007870>
- UniProt Consortium, Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., Da Silva, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., ... Teodoro, D. (2020). Uniprot: The universal protein knowledgebase in 2021. *Nucleic Acids Research*, 49(D1), D480–D489. <https://doi.org/10.1093/nar/gkaa1100>
- Valkenburg, S. A., & Poon, L. L. M. (2022). Exploring the landscape of immune responses to influenza infection and vaccination. *Nature Medicine*, 28(2), 239–240. <https://doi.org/10.1038/s41591-021-01656-4>
- Vogel, O. A., & Manicassamy, B. (2020). Broadly protective strategies against influenza viruses: Universal vaccines and therapeutics. *Frontiers in Microbiology*, 11. <https://doi.org/10.3389/fmicb.2020.00135>

- Waldock, J., Zheng, L., Remarque, E. J., Civet, A., Hu, B., Jalloh, S. L., Cox, R. J., Ho, S., Hoschler, K., Ollinger, T., Trombetta, C. M., Engelhardt, O. G., & Caillet, C. (2021). Assay harmonization and use of biological standards to improve the reproducibility of the hemagglutination inhibition assay: A flucop collaborative study (M. F. Pasetti, Ed.). *mSphere*, 6(4). <https://doi.org/10.1128/msphere.oo567-21>
- Wang, J., Li, D., Perry, S. P., Hilchey, S. P., Wiltse, A., Treanor, J. J., Sangster, M. Y., & Zand, M. S. (2021). Broadly reactive IgG responses to heterologous H5 prime-boost influenza vaccination are shaped by antigenic relatedness to priming strains (R. Rappuoli, Ed.). *mBio*, 12(4). <https://doi.org/10.1128/mBio.oo449-21>
- Wang, Y., Tang, C. Y., & Wan, X.-F. (2021). Antigenic characterization of influenza and SARS-CoV-2 viruses. *Analytical and Bioanalytical Chemistry*, 414(9), 2841–2881. <https://doi.org/10.1007/s00216-021-03806-6>
- Wei, C.-J., Crank, M. C., Shiver, J., Graham, B. S., Mascola, J. R., & Nabel, G. J. (2020). Next-generation influenza vaccines: Opportunities and challenges. *Nature Reviews Drug Discovery*, 19(4), 239–252. <https://doi.org/10.1038/s41573-019-0056-x>
- Wilks, S. (2021). Racmacs: R antigenic cartography macros. <https://github.com/acorg/Racmacs>
- Yang, B., Lessler, J., Zhu, H., Jiang, C. Q., Read, J. M., Hay, J. A., Kwok, K. O., Shen, R., Guan, Y., Riley, S., & Cummings, D. A. T. (2020). Life course exposures continually shape antibody profiles and risk of seroconversion to influenza (C. A. Russell, Ed.). *PLOS Pathogens*, 16(7), e1008635. <https://doi.org/10.1371/journal.ppat.1008635>
- Yang, J., Zhang, T., & Wan, X.-F. (2014). Sequence-based antigenic change prediction by a sparse learning method incorporating co-evolutionary information (G. P. S. Raghava, Ed.). *PLoS ONE*, 9(9), e106660. <https://doi.org/10.1371/journal.pone.0106660>
- Zacour, M., Ward, B. J., Brewer, A., Tang, P., Boivin, G., Li, Y., Warhuus, M., McNeil, S. A., LeBlanc, J. J., & Hatchette, T. F. (2016). Standardization of hemagglutination inhibition assay for influenza serology allows for high reproducibility between laboratories (R. L. Hodinka, Ed.). *Clinical and Vaccine Immunology*, 23(3), 236–242. <https://doi.org/10.1128/cvi.oo613-15>

APPENDIX A

TYPE B INVESTIGATION

The Type B influenza have evolved differently overtime than the Type A influenza. The Type B Yamagata- and Victoria-lineages originated from a most recent common ancestor in the 1980's. Strains isolated previous to this phylogenetic split are within the pre-split lineage. The two circulating lineages named for the strains that marked the divergence (Yamagata-1988 and Victoria-1987). The two lineages have continued to evolve independently of each other.

In this cohort of individuals cross-reactivity of lineages was observed (Figure A). Cross-reactivity similar to this was observed previously for a longitudinal cohort in Hong Kong after infection with a B/Victoria lineage (Lau et al., 2020).

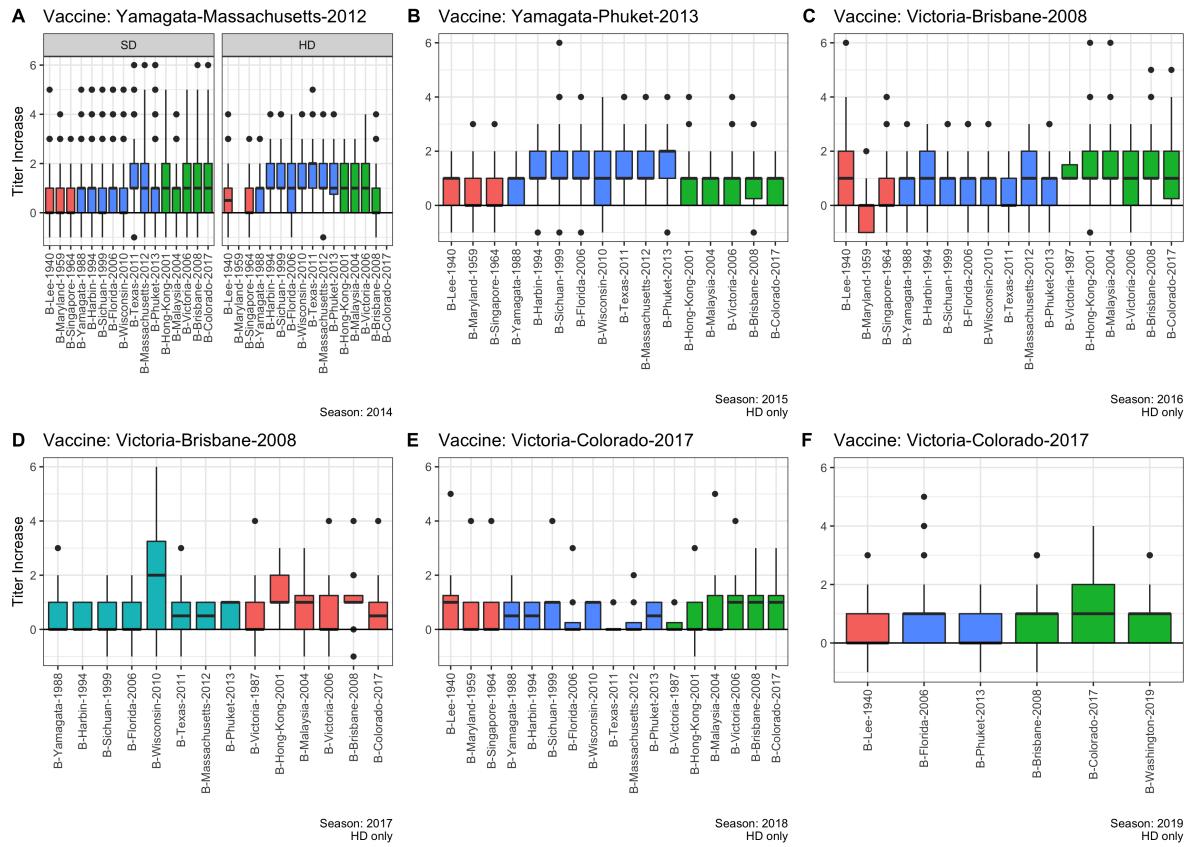


Figure A.i: Cross-reactivity between Type B influenza lineages. Individuals who were vaccinated with Fluzone vaccines that contained only one Type B component were analyzed for changes in Type B HAI titers. All three antigenic clusters were included in the panel (pre-, Yamagata-, and Victoria-lineages). Vaccination with one lineage elicited an increase across all three groups. The SD formulation for the 2014 season was trivalent, and quadrivalent for all other seasons.

APPENDIX B

SUPPLEMENTAL FIGURES

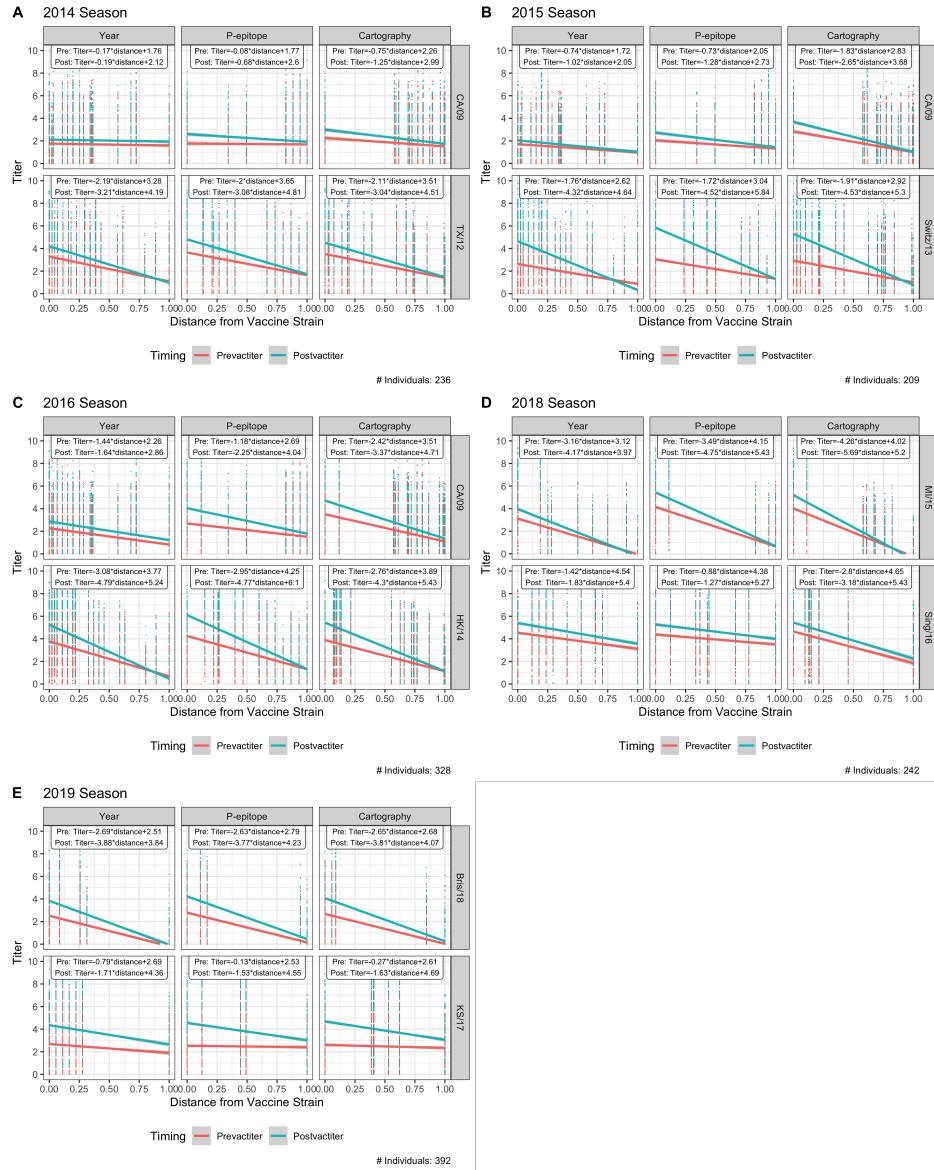


Figure B.1: Pre- and post- vaccination titers to the H1N1 and H3N2 virus panels by season of all individuals who received SD vaccination for the 2014-2016 and 2018-2019 seasons (sample size per season included). The vaccination strain of each subtype per season is shown in the row label. The linear regression of the pre- and post-vaccination titers with 95% confidence intervals are shown. The three distance measures are shown in the column label. The resulting linear regression equations of the pre- and post-vaccination linear regressions are included. Distances were season-based normalized (Equation 2.7).

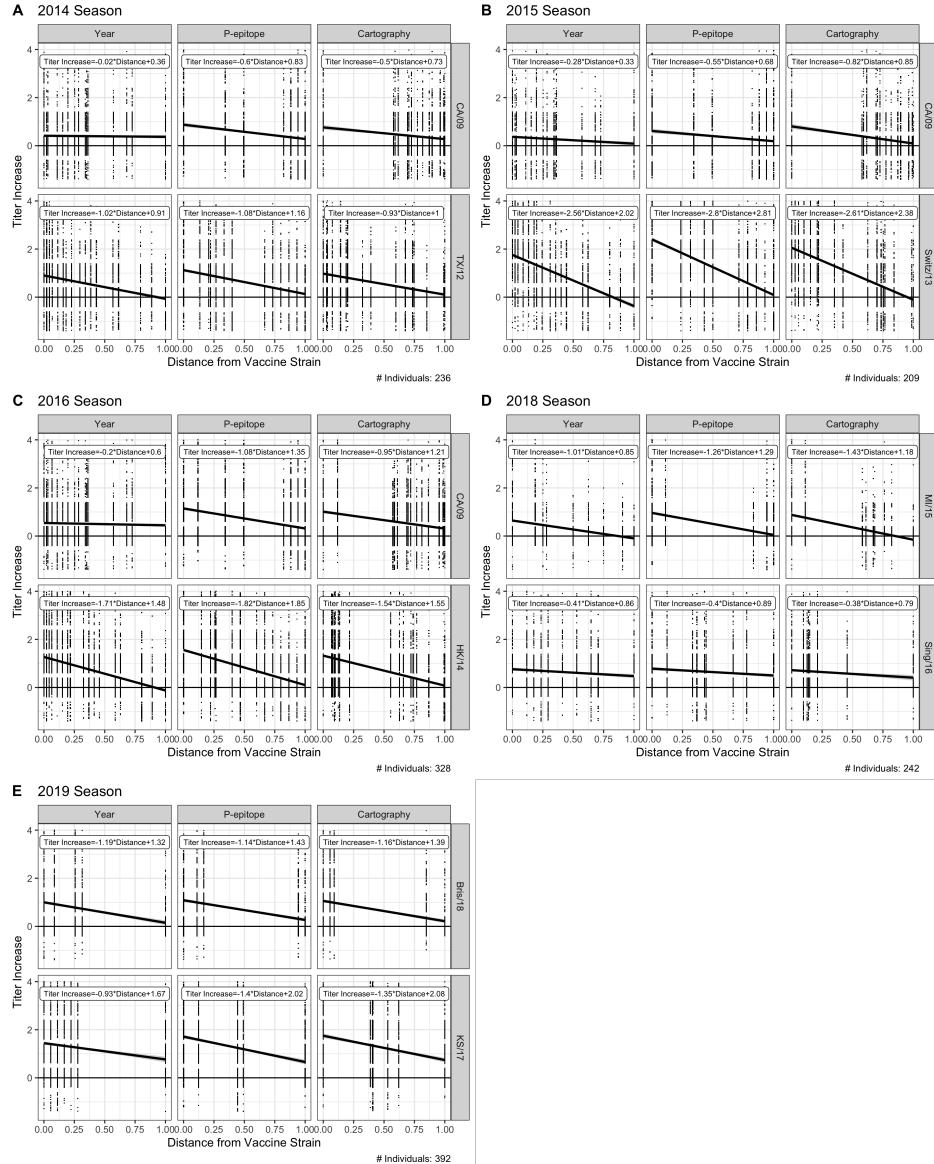


Figure B.2: The titer increase to the H1N1 and H3N2 virus panels for the 2014-2016 and 2018-2019 seasons of all individuals who received SD vaccination. The columns are separated by vaccine strain. The linear regression of the titer increase with 95% confidence intervals are shown for each distance method. Distances were season-based normalized (Equation 2.7). The raw datapoints had jitter applied with +/- 0.4 in the y-axis. Raw data points that fell outside of the y-axis bounds are not shown.

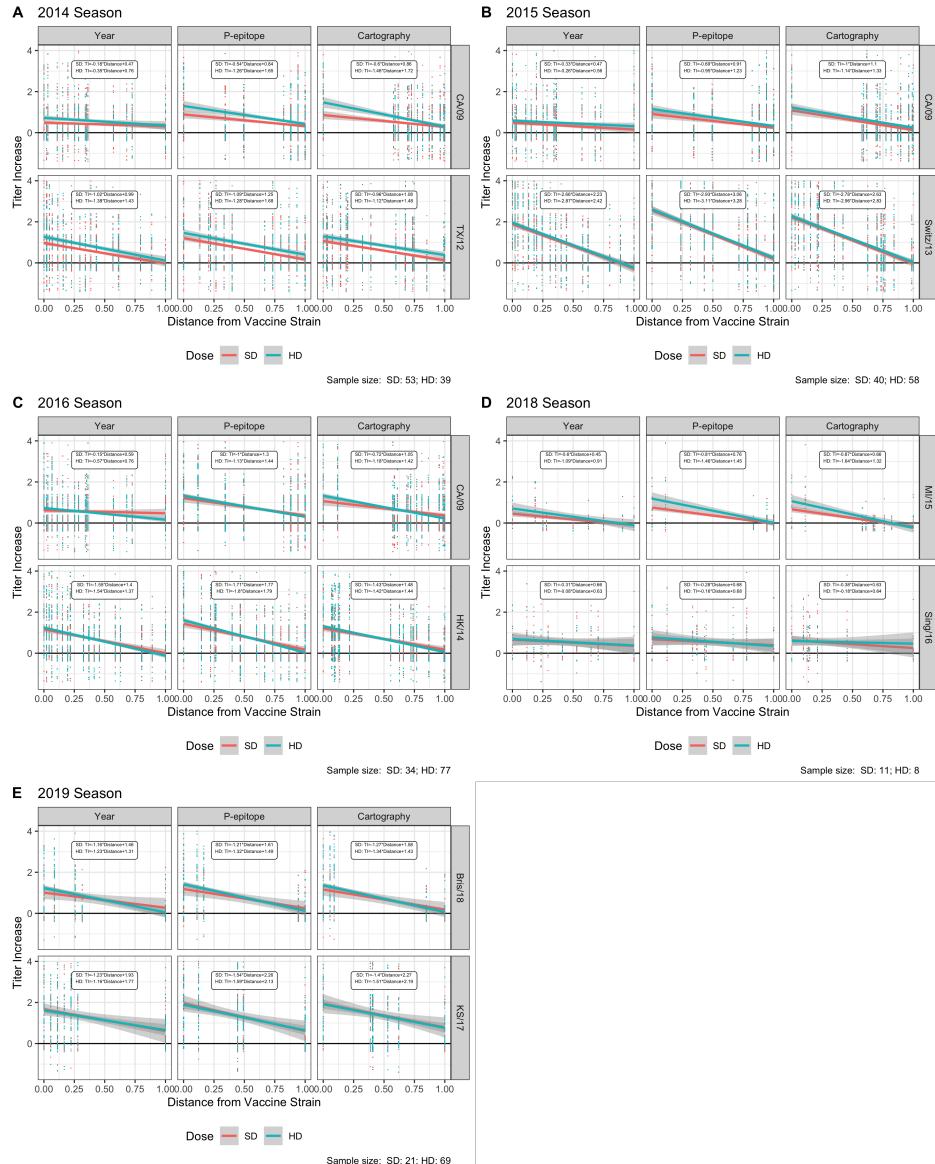


Figure B.3: The titer increase (TI) to the H1N1 and H3N2 virus panels for the 2014-2016 and 2018-2019 season of all individuals greater than or equal to 65 years of age. The columns are separated by distance method. The linear regression of the titer increase with 95% confidence intervals are shown for the dose received. Distances were normalized by season (Equation 2.7). The black line indicates no titer increase. The raw datapoints had jitter applied with ± 0.4 in the y-axis. Raw data points that fell outside of the y-axis bounds are not shown.