Vaccine 30 (2012) 448–453

Contents lists available at SciVerse ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

# Antigenic distance measurements for seasonal influenza vaccine selection

Zhipeng Cai [a], Tong Zhang [b], Xiu-Feng Wan [a,*]

[a] Department of Basic Sciences, College of Veterinary Medicine, Mississippi State University, Mississippi State, MS, USA
[b] Department of Statistics, Rutgers University, Piscataway, NJ 08854, USA

ARTICLE INFO

*Article history:*
Received 15 February 2011
Accepted 23 October 2011
Available online 7 November 2011

*Keywords:*
Influenza
Vaccination
Antigenic distance
Antigenic cartography
Vaccine strain selection
Antigenic drift

ABSTRACT

Influenza vaccination is one of the major options to counteract the effects of influenza diseases. Selection of an effective vaccine strain is the key to the success of an effective vaccination program since vaccine protection can only be achieved when the selected influenza vaccine strain matches the antigenic variants causing future outbreaks. Identification of an antigenic variant is the first step to determine whether vaccine strain needs to be updated. Antigenic distance derived from immunological assays, such as hemagglutination inhibition, is commonly used to measure the antigenic closeness between circulating strains and the current influenza vaccine strain. Thus, consensus on an explicit and robust antigenic distance measurement is critical in influenza surveillance. Based on the current seasonal influenza surveillance procedure, we propose and compare three antigenic distance measurements, including Average antigenic distance (A-distance), Mutual antigenic distance (M-distance), and Largest antigenic distance (L-distance). With the assistance of influenza antigenic cartography, our simulation results demonstrated that M-distance is a robust influenza antigenic distance measurement. Experimental results on both simulation and seasonal influenza surveillance data demonstrate that M-distance can be effectively utilized in influenza vaccine strain selection.



## 1. Introduction

Influenza viruses cause both pandemics and seasonal influenza and continue to present a threat to public health [1–4]. Vaccination is one of the major practical options to prevent and reduce the burdens from influenza outbreaks. To identify an optimal influenza vaccine strain, the World Health Organization (WHO) Global Influenza Surveillance Network (GISN), which consists of more than 120 national influenza centers from all over the world, performs large scale influenza viral sampling and characterization. Four basic criteria are generally used in influenza vaccine strain selection: (1) the candidate strain is an influenza antigenic variant based on the results from hemagglutination inhibition (HI) and/or microneutralization (MN); (2) the candidate strain-like viruses are becoming more predominant; (3) the candidate strain-like viruses are geographically distributed; and (4) the candidate strain can be propagated efficiently in chicken embryonic eggs.

The immunological assays, such as HI and MN assays [7,8,5] are generally used to identify the antigenic variants among those influenza viruses isolated from the influenza surveillance. In the experimental results from HI assay, each value (titre) measures the affinity between an influenza virus and an antiserum. The antigenic relations between two viruses are usually determined indirectly from their reactions against the same panel of antisera, which are generated against the testing influenza viruses or other related influenza viruses. The more similar their HI profiles are, the closer their antigenic relationships are. In order to decide whether an influenza virus is an influenza variant, it will be critical to quantify the difference among antigenic profiles, with the so called antigenic distance [10,9] in the immunological assays. However, there has been no formal definitions of antigenic distances based on HI values. In practice, a two-way analysis is commonly used: the antigenic distance between two antigens is a certain averaged ratio for the pair-wise titers, e.g. HI values, between two antigen–antiserum pairs, in which each antiserum corresponds to each of the testing antigens, respectively. For instance, in Table 1, the antigenic distance is 12-fold between A/New York/55/2001 (H3N2) and A/Fujian/411/2002(H3N2). If their antigenic distance is less than 4-fold, previous vaccine could have negative interference on the subsequent vaccine [10,9,11,12]. Thus, a minimum of a 4-fold antigenic distance is used in vaccine strain selection. However, there is no agreement on what should be the correct definition of antigenic distance given an HI table, especially when the table contains many antisera. The purpose of this study is to investigate what should be a suitable definition of antigenic distance. This issue deserves extra attention because different notations of antigenic distance may lead to different conclusions in the influenza vaccine selection process.

* Correspondence author. Tel.: +1 662 325 3559.
*E-mail addresses:* wan@cvm.msstate.edu, wanhenry@yahoo.com (X.-F. Wan).

0264-410X/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.vaccine.2011.10.051

**Table 1**
Two-way analysis in influenza antigenic distance measurement.

| | A/NEW YORK/55/01 | A/FUJIAN/411/02 |
|---|---|---|
| A/NEW YORK/55/01 | 1280 | 80 |
| A/FUJIAN/411/02 | 160 | 1280 |

The concept of antigenic distance is also important in antigenic cartography, which projects all antigens into a two or three dimensional graph [15,13,14]. In general, the Euclidean distance in the antigenic cartography should reflect the antigenic distance in the HI tables, so that the cartography will facilitate the interpretation and measurements of antigenic differences among influenza viruses. In the metric MDS approach [15], the notation of antigenic distance is not explicitly based on HI table values, but rather on the theory of shape space [6,16]. After constructing the cartography, the normalized HI values between antigens and antisera are treated as their distances in the cartography. That is, the antisera are utilized as the reference substances to measure the antigenic distances between antigens [15]. In the more recently developed Matrix Completion-Multidimensional Scaling (MC-MDS) method [13], the antigenic distance between two antigens corresponds to a certain explicitly defined distance between the HI values of these two antigens against all the antisera. The specific distance used in [13] corresponds to the Average antigenic distance (A-distance) described in this study. Since different definitions of HI distances can lead to different cartographies, it is important to understand the effect of different definitions. In fact, our previous study showed that there were some discrepancies between distance measurement in metric MDS and that in MC-MDS [13]. This issue will be further explained in the current paper.

In this study, we propose and compare three antigenic distance measurements that are explicitly defined based on HI values: Average antigenic distance (A-distance), Mutual antigenic distance (M-distance), and Largest antigenic distance (L-distance). We also argue that the cartographic distance from the metric MDS approach [15] is more similar to the L-distance. Our simulation results show that M-distance is the most robust among different definitions. The merits and drawbacks of those definitions are discussed. Experimental results on some real influenza surveillance data demonstrate the suitable of M-distance for cartography and influenza vaccine strain selection.

## 2. Materials and methods

### 2.1. Influenza antigenic distance definition

Three antigenic distance measurements, A-distance, M-distance, and L-distance, are proposed. Assume we have an immunological dataset, e.g. an HI matrix, with $n$ antigens (viruses) and $m$ antisera. Let $H_{ij}$ be the HI value between antigen $v_i$ and antiserum $a_j$. We also assume that both antigens and antisera can be partitioned in to $k$ clusters by their temporal information, e.g. according to influenza season information. Denote by $P_x$ the $x$-th cluster, where $x$ is the temporal cluster index. The A-distance between two antigens $v_i$ and $v_j$ is defined as $(1/n)\sum_{t=1}^{n}|H_{it}-H_{jt}|$. This quantity measures the average difference of antigen–antiserum interaction effect of the two antigens. The L-distance is defined as $\max_{t=1,2,\ldots,n}|H_{it}-H_{jt}|$, which measures the maximum difference of antigen–antiserum interaction effect of the two antigens. The definition of M-distance between antigens $v_i$ and $v_j$ requires the cluster information, where we assume that $v_i \in P_x$ and $v_j \in P_y$. The M-distance between $v_i$ and $v_j$ is defined as

$$\frac{\sum_{a_t \in P_x \vee a_t \in P_y}|H_{it}-H_{jt}|}{|\{a_t \in P_x \vee a_t \in P_y\}|},$$

where $|\{a_t \in P_x \vee a_t \in P_y\}|$ is the number of antisera in the union of clusters $P_x$ and $P_y$. The distance between two antigens will be defined as A-distance in the rare situation that $|\{a_t \in P_x \vee a_t \in P_y\}|$ is zero (that is, no antisera are produced in the time periods corresponding to antigens $v_i$ or $v_j$). The motivation behind this definition is that for vaccine selection purpose we should focus our attention on the part of antigen–antiserum interaction assay during the most relevant time periods, which are the periods that contain antigens $v_i$ or $v_j$. M-distance thus only considers the interaction of antigens $v_i$ and $v_j$ to antisera $a_t$ in the same periods as these two antigens (the union of antisera in clusters $P_x$ and $P_y$).

In the simulation data, the cluster membership is pre-defined. In the seasonal influenza data, the clusters of antigens and antisera are defined based on seasonal information of the influenza viruses. Specifically, antigens and antisera collected from the same influenza season will be assigned into the same cluster.

### 2.2. Influenza antigenic distance definition in metric MDS

In the metric MDS [15], the antigenic distance is calculated indirectly. This method attempts to minimize an error function $\sum e(D_{ij}^E, X_{ij}^E)$, where $D_{ij}^E$ is the normalized observed HI values between antigen $i$ and antiserum $j$. Metric MDS requires a pre-defined dimension $L$. Each antigen (and each antiserum) is represented by an $L$-dimension vector. Let $v_{ij}$ (or $a_{ij}$) denote the coordinate values of the $j$th dimension in $v_i$ (or $a_i$), then $X_{ij}$ is the Euclidean distance between antigen $i$ and antiserum $j$ defined as: $X_{ij} = \sqrt{\sum_{l=1}^{L}(v_{il}-a_{jl})^2}$. Basically, metric MDS tries to minimize the difference between the Euclidean distances of all embedded antigen and antiserum pairs in the cartography and the corresponding HI table values. The distance between each pair of antigens is calculated as the Euclidean distance between them.

### 2.3. Antigenic distance between influenza vaccine strain and the emerging antigenic variants

During influenza epidemic, the vaccine strain needs to be updated for the next influenza season if antigenic properties of the circulating viruses have changed significantly so that the vaccine strain used in current influenza season could not generate adequate immunological protection. Let antigen $c$ be the vaccine strain for current influenza season and $CI$ is the set of the circulating antigens in an HI dataset. In the antigenic cartography, we calculate $d_{c\hat{j}}$, the distance between the current vaccine strain in use $c$ and the center $\hat{j}$ of the circulating viruses:

$$\hat{j} = \arg\max_j \frac{\sum_{t\in CI} p_{jt}}{|CI|}, \quad s.t. \quad p_{jt} = 2^{-d_{jt}}$$

The quantity $p_{jt}$ measures the similarity between antigen $j$ and antigen $t$ (larger means more similar). Therefore $\hat{j}$ is the antigen that is on average most similar to antigens in $CI$. Since the distance in antigenic cartography measures $\log_2$-scale change of the original unnormalized HI assay values, the quantity $2^{-d_{jt}}$ can also be considered as a measure that is inversely proportional to the physical HI-value difference (in folds) between antigens $j$ and $t$. A large $d_{c\hat{j}}$ indicates that the center of the antigens strays away from the vaccine strain used in current influenza season. This means on average the circulating viruses have different antigenic properties from the currently used vaccine strain, and thus new vaccine strains shall be recommended for the next influenza season. The decision can be based on testing a condition $d_{c,\hat{j}} > T$, where $T$ is a pre-defined threshold for vaccine strain selection. Usually, $T$ is defined as 4-fold change in the original HI assay. Equivalently it is 2 units in antigenic cartography [10,9] because the distance in antigenic cartography

measures $\log_2$-scale change of the original unnormalized HI assay values.

### 2.4. Antigenic cartography construction

Given any distance that is explicitly defined based on HI values (including the three distances proposed in the paper), the corresponding antigenic cartography can be constructed using the MC-MDS method of [13]. Briefly, a matrix completion algorithm is applied first to estimate the low reactors and possible missing entries in the HI matrix. After that, the pairwise distances among antigens are calculated by using the given distance definition, which could be A-distance, M-distance, or L-distance. Finally, a Multidimensional scaling (MDS) method is utilized to project all the antigens into a two or three dimensional graph. The HI values are normalized as follows: each observed value or reactor is transformed using the formula $\lceil \max(T_{ij}) \rceil - \log_2(\max(H_j)/H_{ij})$, where $\max(T_{ij})$ represents the largest $\log_2(\max(H_j)/H_{ij})$ value among the table and $\max(H_j)$ represents the maximum HI value for antibody $j$.

### 2.5. Generation of simulated HI data

During an influenza epidemic, if antigenic drift occurs, we expect the influenza antigenic variants to become more predominant. To mimic the emerging of new antigenic variants, we simulate three scenarios: (1) the antigenicity of the influenza viruses from the previous season is similar to that in the current season, and thus no new antigenic variant emerges (Fig. 1A); (2) the antigenicity of the influenza viruses from the previous season is slightly different from that in the current season, and thus new antigenic variants are starting to emerge (Fig. 1B); (3) the antigenicity of the influenza viruses from the previous season is significantly different from that in the current season, and thus new antigenic variants have already formed (Fig. 1C). These scenarios are representative of real influenza evolutionary trends from the historic data. For example, the H3N2 influenza antigenic changes from 1991–1992 (Fig. 2A), 1993–1994 (Fig. 2B), and 1992–1993 (Fig. 2C) resemble these three scenarios.

In order to investigate the performance of different antigenic distance measurements, we generated simulated HI matrices, each has $n$ antigens and $m$ antisera. Both antigens and antisera can be partitioned into $t$ clusters by their temporal information. We assume that the HI values between antigens in cluster $p_i$ and antisera in cluster $p_j$ follow a Gaussian distribution $N(8 - |i - j|, \sigma^2)$. Here $\sigma$ is chosen from a range [0.1, 1.5] with an increment 0.1 in our simulation. Besides these observed values, there are two other types of values in each HI matrix: missing values and low reactors [13]. A low reactor indicates a weak immunological reaction between an antigen and an antiserum. A value less than 20 is usually regarded as a low reactor in an HI assay. We randomly select 5% of the HI values off the diagonal of the simulated HI matrix as low reactors or missing values. The HI data contain noises due to reasons such as variations in the experimental conditions from different laboratory and even the same laboratory. These variations could come from but not limited to inconsistencies in virus titration, red blood cells, and antisera. To simulate this type of noise, we assigned 10% HI values in off diagonal entries in the simulation data to random values. The results from the data with and without noises will be compared.

In order to check the robustness of various antigenic distance measurements, simulation datasets were generated with three clusters of antigens and three clusters of corresponding antisera. The three clusters can be regarded as three consecutive influenza seasons, e.g. the 2008–2009 influenza season in Northern Hemisphere, the 2009 influenza season in Southern Hemisphere, and the 2009–2010 influenza season in Northern Hemisphere. We can assume that the vaccine strain currently in use belongs to cluster 1, but the vaccine candidate strain will be located from cluster 3. If cluster 3 is significantly different from cluster 1, then there is a significant antigenic drift, which implies that an update of the influenza vaccine may be necessary. The degree of changes from cluster 1 to cluster 3 will be varied in the simulation study.

**Table 2**
HI simulation datasets.

| | P1 | P2 | P3 |
|---|---|---|---|
| I | 50 | 0 | 50 |
| II | 30 | 40 | 30 |
| III | 10 | 80 | 10 |

There are three clusters in each simulation data. P1, P2 and P3 represent the percentage of antigens in cluster 1, cluster 2 and cluster 3, respectively.

In the influenza sample collection process, there are often large variations in sample sizes obtained from different locations or during different time periods. Such variations could be the result of sample biases or the availability of samples during disease epidemics. A robust antigenic distance measurement should be less sensitive to such variations. Therefore in our simulation, we generate three sets of data, representing variations of the same underlying data distribution, according to Table 2. For instance, Dataset I contains 50%, 0%, and 50% of clusters 1, 2, and 3. In our simulation of Dataset I, II, and III, the number of antigens in cluster 1 and cluster 3 is set to be 10. The different proportions are achieved by changing the number of antigens in cluster 2. These simulation programs can be downloaded from http://sysbio.cvm.msstate.edu/AntigenMap/AntigenicDistance.

### 2.6. Bootstrap and statistical analysis

Since most of the experimental variations are heavily affected by the selection of antisera, we performed bootstrap based on random sampling of antisera in the H3N2 datasets. That is, we sample (with replacement) the antisera from the seed HI matrix to form new HI matrices of the same size. An antibody might be selected more than once in each matrix. We generated 100 HI matrices this way, and the average pairwise antigenic distances are calculated. To test the hypotheses on four antigenic measurements within HI datasets, a two-sample multivariate test (Student's $t$ test) was used.

## 3. Results

### 3.1. Comparison of antigenic distance measurements on simulated HI datasets

This simulation study shows that the distance from M-distance measurement is the most robust among the four distance measurements compared in Fig. 3. Metric MDS has the largest variation among all four measurements. The A-distance has relatively large variations when $\sigma$ on Gaussian distribution is small, whereas the L-distance has relatively large variations when $\sigma$ is large. The results from M-distance are the closest to the benchmark distance, and they also have the smallest variations. Instead, the A-distance, L-distance, and the distance measurement from metric MDS are dependent on the data distribution to different degrees. The distance from metric MDS was least stable since it has the largest variations. In general, L-distance > Metric MDS > M-distance > A-distance.

The simulation on the dataset with 10% noisy data also demonstrated that M-distance is the most robust distance measurement among the four compared measurements (Fig. 3B). That is, the M-distance is least affected when we increase the noise level. Our

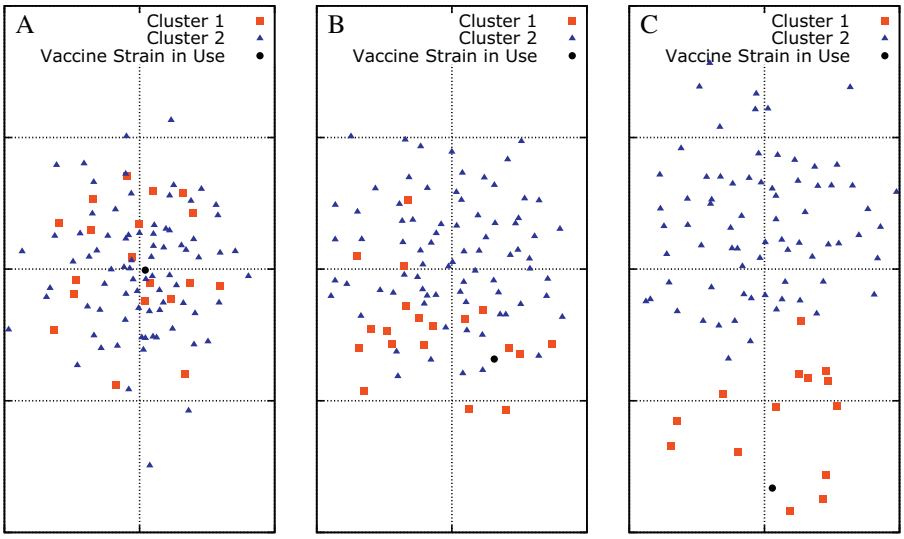


**Fig. 1.** Antigenic cartography for three possible scenarios in influenza surveillance: (A) the antigenicity of the influenza viruses from previous season is similar to that in current season; (B) the antigenicity of the influenza viruses from previous season is slightly different from that in current season; (C) the antigenicity of the influenza viruses from previous season is different from that in current season, and antigenic variants are emerging.

study suggested also that A-distance is not as stable as M-distance; this is because A-distance is heavily affected by the number of antisera used in the simulation datasets. L-distance has the largest antigenic distance because L-distance always selects the maximum distance between two antigens.

### 3.2. Application of antigenic distance measurements in H3N2 influenza A viruses

To demonstrate the impact of different antigenic distance measurements on the real influenza data, we applied the three definitions (A-distance, M-distance, and L-distance) and the antigenic distance measurement from metric MDS to evaluate the antigenic distances for H3N2 influenza A viruses. The HI dataset contains 273 antigens isolated from 1968 to 2003 [15]. To minimize the potential biases from data size, we focused on the HI data from 1991 to 1993, which has a relatively large number of antigens and antisera compared to those from other years.

The bootstrap results from this real dataset are consistent with those from the simulation: in general, L-distance > Metric MDS > M-distance > A-distance (Table 3). A two-sample multivariate test indicated that the differences are statistically significant because $p$ values between each pair of measurements are smaller than 0.001 (Table S1). Similar to our simulation results, the M-distance has the smallest variation among the four antigenic distance measurements compared. This means that M-distance is the most robust because the selection of antisera has the least impact on M-distance measurement.

### 3.3. Application in influenza vaccine strain selection for 2009–2010 season

The WHO GISN committee meets twice a year to discuss whether the seasonal influenza vaccine strains need to be updated for the next influenza season: in February for the Northern Hemisphere influenza season and in September for the Southern Hemisphere influenza season. The influenza vaccine selection includes H1N1, H3N2, and influenza B viruses before 2009–2010 season. Starting from the 2010–2011 season, 2009 H1N1 virus is also added. In order to further confirm the results from our simulation data, we applied M-distance to available real data. Due to the limitation of HI data availability, we can only study the 2009–2010 influenza season. Fig. 4A and B showed the antigenic cartography for H3N2 (Fig. 4A) and H1N1 (Fig. 4B) influenza viruses in the season of 2009–2010, where antigenic drift values for H3N2 and H1N1 were 0.8000 and 1.0000, respectively. In the season of 2009–2010,

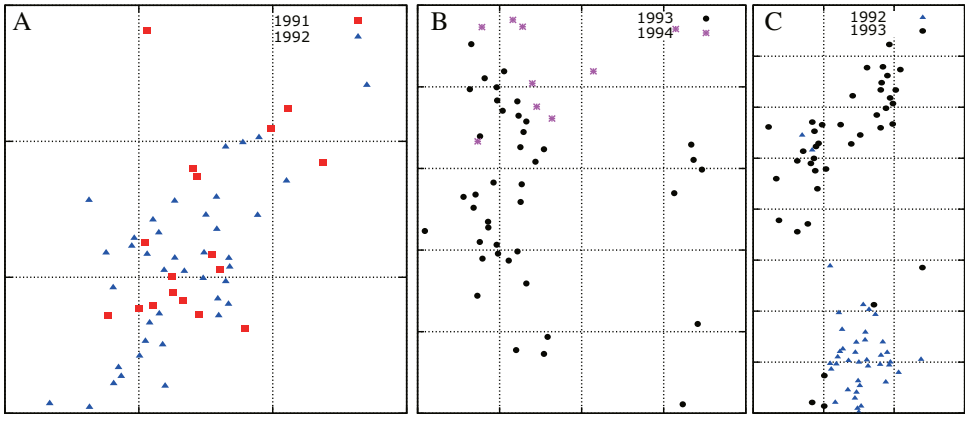


**Fig. 2.** Antigenic cartography for H3N2 influenza A viruses using the HI dataset from [15]: (A) 1991–1992 season; (B) 1993–1994 season; (C) 1992–1993 season. These three seasons resemble the three scenarios described in Fig. 1.

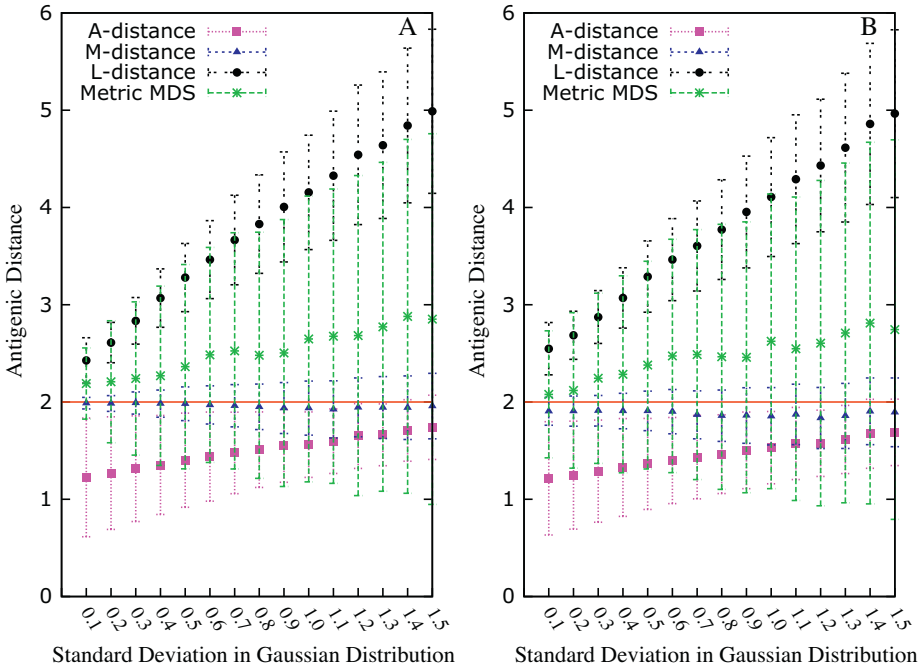


**Fig. 3.** Comparison of antigenic distance measurements on simulated HI datasets. The antigenic distances were computed based on 100 runs of different distributions shown in data I, II, or III: (A) no random values were assigned in the simulation data; (B) the 10% HI values in off diagonal entries in the simulation data were set to be random values. The red line represents the benchmark antigenic distance from the simulated HI data. *y*-axis represents the antigenic distance and its variations, and *x*-axis represents the $\sigma$ (standard deviation) in the simulation dataset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

**Table 3**
Application of antigenic distance measurements in H3N2 influenza A viruses.

| | A-distance | M-distance | L-distance | Metric MDS |
|---|---|---|---|---|
| 1991 vs 1992 | 1.4661(0.3854) | 1.7713(0.4799) | 4.1508(0.9549) | 3.0171(1.1625) |
| 1992 vs 1993 | 3.8376(0.9190) | 3.9313(0.8598) | 9.3012(1.7496) | 6.1302(0.9229) |
| 1991 vs 1993 | 4.0328(0.9712) | 4.6765(0.8476) | 9.6652(1.7893) | 6.1107(1.0830) |

The value in the bracket denotes the standard deviation in the results from 100 runs.

WHO GISN recommended not to update the vaccine strains for H3N2 and H1N1 [17,20]. The antigenic drift value for B was 3.7281, indicating the vaccine strain of B shall be updated (Fig. 4C), and in fact, WHO GISN recommended to update the vaccine influenza B [17,20]. Therefore, our antigenic distance analysis is consistent with the WHO recommendations. We would like to mention that such a large distance in influenza B was a switch of circulating strain from Yamagata lineage to Victoria lineage [17]

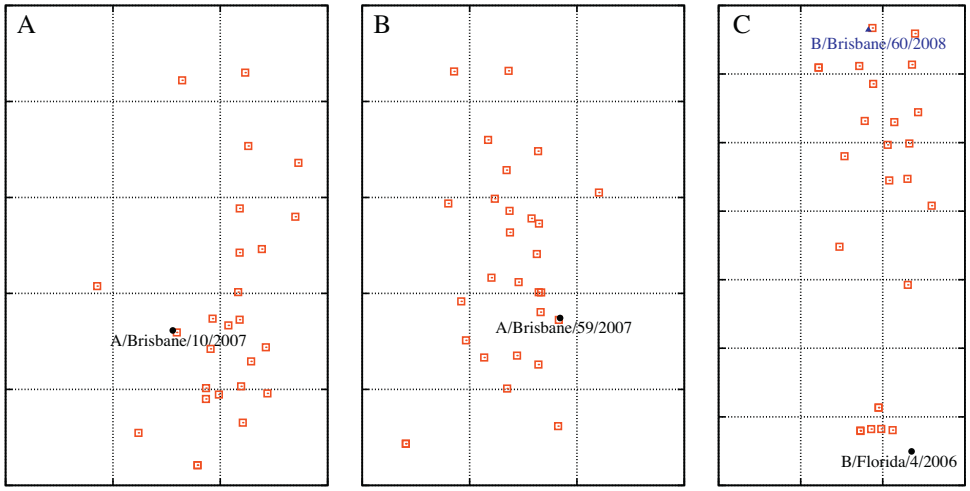


**Fig. 4.** Antigenic cartography for seasonal influenza viruses in the 2009–2010 season, Northern Hemisphere. Each square represents an influenza virus isolated from the season of 2009–2010. (A) The A/Brisbane/10/2007(H3N2) is the H3N2 vaccine strain used for the 2008–2009 season; (B) the A/Brisbane/57/2007(H1N1) is the H1N1 vaccine strain used for the 2008–2009 season; (C) the B/Florida/4/2006 is the influenza B vaccine strain used for the 2008–2009 season while the B/Brisbane/60/2008 is the influenza B vaccine strain used for the 2009–2010 season.

## 4. Discussion

In this study, we proposed and compared three antigenic distance measurements. Our simulation study suggested that the mutual antigenic distance, M-distance, is the most robust measurement. Since our simulation data are generated based on historical influenza data, we believe M-distance is the most suitable distance definition for vaccine selection among the several definitions examined in this study. This is confirmed with additional experiments on some real influenza data. The antigenic distance measurements are proposed specifically for human seasonal influenza vaccine strain selection. Unlike seasonal influenza viruses in human, influenza viruses in animals, such as avian influenza viruses, do not have similar antigenic drift patterns as those in human seasonal influenza viruses. For instance, the H3N2 human influenza A virus has a major trunk in their phylogenetic tree [18], which suggests that this virus has gradual changes in genetic properties but punctual changes in antigenic properties [15]. On the other hand, different lineages of the same subtype of avian influenza viruses (e.g. H5N1) could co-exist [19] without clear clustering structures. Therefore M-distance, which requires clustering based on influenza season information, might not be the best choice for studying these influenza viruses. In the general case where clearly defined clustering structures are not available, the A-distance might be the best default option.

The selection of influenza vaccine is the first but also the most critical step toward an effective influenza vaccination program. The proposed methods can be applied in HI, MN, and ELISA experiments. Due to the variations of HI assays, e.g. the red blood cells, the interpretation should be made cautiously. Integration of results from multiple assays, especially HI and MN, should be considered [17]. In addition, as a practical point, more than one vaccine candidate should be selected since additional factors not included in this study need to be considered. For example, the propagation ability in chicken embryonic eggs is important because it indicates the usefulness of the selected virus for vaccine manufacturing.

## Acknowledgements

The project described was supported by award number RC1AI086830 from the National Institute Of Allergy And Infectious Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute Of Allergy And Infectious Diseases or the National Institutes of Health. This study is also supported by the start-up package from Mississippi State University and NSF-EPS-0903787 subaward 012156-007 to XFW; NSF DMS-1007527, NSF IIS-1016061, and AFOSR-10097389 to TZ. ZC is supported partially by NSERC postdoc fellowship.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.vaccine.2011.10.051.

## References

[1] Simonsen L, Fukuda K, Schonberger LB, Cox NJ. The impact of influenza epidemics on hospitalizations. J Infect Dis 2010;181(3):831–7.
[2] Thompson WW, Shay DK, Weintraub E, Brammer L, Cox NJ, Anderson LJ, et al. Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA 2003;289(2):179–86.
[3] Lave JR, Lin CJ, Fine MJ, Hughes-Cromick P. The cost of treating patients with community-acquired pneumonia. Sem Respir Crit Care Med 1999;20:189–97.
[4] CDC. Estimates of deaths associated with seasonal Influenza – United States, 1976–2007. JAMA 2010;304(16):1778–80.
[5] Frank AL, Puck J, Hughes BJ, Cate TR. Microneutralization test for influenza A and B and parainfluenza 1 and 2 viruses that uses continuous cell lines and fresh serum enhancement. J Clin Microbiol 1980;12(3):426–32.
[6] Perelson A, Oster G. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. J Theor Biol 1979;81:645–67.
[7] Potter CW, Oxford JS. Determinants of immunity to influenza infection in man. Br Med Bull 1979;35:69–75.
[8] Arnold SM, Maassab HF. Ether treatment of type B influenza virus antigen for the Hemagglutination inhibition Test. J Clin Microbiol 1981;13:54–7.
[9] Smith DJ, Lapedes AS, Forrest S, de Jong J, Osterhausa DME, Fouchier AM, et al. Modeling the effects of updating the influenza vaccine on the efficacy of repeated vaccination. In: Osterhaus AD, Cox NJ, Hampson A, editors. Options for the control of influenza IV. Crete, Greece: Excerpta Medica, International Congress Series 1219. 2001. p. 655–60.
[10] Smith DJ, Forrest S, Ackley DH, Perelson AS. Variable efficacy of repeated annual influenza vaccination. Proc Natl Acad Sci USA 1999;94(24):14001–6.
[11] Fazekas de St Groth S, Webster RG. Disquisitions on original antigenic sin. II. Proof in lower creatures. J Exp Med 1966;124(3):347–61.
[12] Webster RG, Kasel JA, Couch RB, Laver WG. Influenza virus subunit vaccines. II. Immunogenicity and original antigenic sin in humans. J Exp Med 1976;134(1):48–58.
[13] Cai Z, Zhang T, Wan X-F. A Computational framework for influenza antigenic cartography. PLoS Comput Biol 2010;6(10):e1000949.
[14] Cai, Z, Zhang, T and Wan X-F. Concepts and applications for influenza antigenic cartography. Influenza and Other Respiratory Viruses 2011; 5 (Suppl. 1): 204-207.
[15] Smith D, Lapedes A, Jong J, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, et al. Mapping the antigenic and genetic evolution of influenza virus. Science 2004;35:371–6.
[16] Lapedes A, Farber R. The geometry of shape space: application to influenza. J Theor Biol 2001;212(1):57–69.
[17] Barr I, McCauley J, Cox NJ, Daniels R, Engelhardt OG, Fukuda K, et al. Epidemiological, antigenic and genetic characteristics of seasonal influenza A(H1N1), A(H3N2) and B influenza viruses: Basis for the WHO recommendation on the composition of influenza vaccines for use in the 2009–2010 Northern Hemisphere season. Vaccine 2010;28(5):1156–67.
[18] Bush R, Bender C, Subbarao K, Cox NJ, Fitch W. Emergence of a novel swine-origin influenza a (H1N1) virus in humans. Science 1999;286: 1921–5.
[19] WHO and OIE. Fao HN and Evolution Working Group. Toward a unified nomenclature system for highly pathogenic avian influenza virus (H5N1). Emerg Infect Dis 2008;14(7):e1.
[20] WHO. Recommended composition of influenza virus vaccines for use in the 2009–2010 influenza season. WHO report; 2009.