Antigenic Distance in Consequence to Avian Influenza Vaccine Strain Selection

Student name	: Syabilla Rivenia (1064215)
Program	: MSc Animal Sciences
Chair group	: Quantitative Veterinary Epidemiology
Course code	: QVE80436
ECTS	: 36 ECTS
Supervisor	: Mart de Jong and Boris V. Schmid

ABSTRACT

Avian influenza poses a high-risk threat to poultry and human health. Monitoring and surveillance of the disease to acquire knowledge of antigenic variability is important to evaluate the current vaccine strain by examining the antigenic distance. Antigenic distance is calculated using R-acmac package in R-studio that provide cartography to help evaluate the distance visually. However, this mapping causes distortion as they reduce the dimensionality into two or three dimensions. Calculating an accurate antigenic distance is essential, especially for selecting the strain used for the vaccine update. Inaccurate antigenic distance calculation may result in unnecessary vaccine updates or the selection of the wrong strain. To provide a better view of the antigenic distance calculation method, this research aims to evaluate the distortion between the R-acmac method and real distance in consequence to the vaccine strain selection. Three datasets were used in this research to evaluate both methods. This research has shown that distortion is evident in R-acmac distance, potentially impacting the selection of vaccine strain. This shows the significance of accurately calculating antigenic distance to enhance precision in vaccine strain selection. Practical implication of these outcome can be utilized for vaccine development and avian influenza control strategies to formulate a broad protection vaccine that is aligns effectively with circulating strain.

Keywords: Avian influenza, antigenic distance, distortion, vaccine strain selection

INTRODUCTION

Avian flu is a highly contagious disease caused by *Alphainfluenzavirus influenzae* that belongs to the Orthomyxoviridae family (ICTV, 2022). Avian flu is recognised by severe depression, ruffled feathers, facial oedema extending to the comb, wattles and neck, and congested conjunctiva in chickens (Gaide et al., 2022; Theary et al., 2012). This disease creates severe economic losses in the poultry industry worldwide due to the massive mortality and morbidity caused by Highly Pathogenic Avian Influenza (HPAI), mainly caused by H5N1 (Swayne and Suarez, 2000). For instance, 30% of farms discontinued operating in three regions of Indonesia, West Java, East Java, and Lampung, during the HPAI outbreak in 2004 and 2005 (Basuno et al. 2010). It is also reported that the HPAI outbreak caused two to eleven times gross margin losses in independent broiler farms in West Java, Indonesia (Pramuwidyatama et al., 2023).

Besides the threat in the poultry industry, HPAI H5N1 has emerged as a global concern for human health, including in Indonesia. These concerns are traced back to the first human infection with Avian Influenza (AI) in Hong Kong in 1997 (Wan, 2012). Since then, the disease has spread worldwide, with 907 human H5N1 cases reported between 1997 and 2015, as reported by Lai et al. (2016). A recent update from the World Health Organization (WHO) documented 461 fatalities out of 882 cases of human infection with avian influenza from 2003 to 2023 (WHO, 2023). Taking a closer look at Indonesia, the initial HPAI in human cases was identified on Java Island in 2003 and reached the peak of disease detection in 2005-2006, with 100% case fatality in 2012 (Sumiarto and Arifin, 2008). Despite subsequent declines in the reported HPAI cases, no human cases have been reported to WHO from Indonesia from 2020 to 2023. Nevertheless, the persistent threat of HPAI to human infection remains evident as the occurrence of four deaths in humans caused by HPAI in Cambodia in 2023 (WHO, 2023). Given its proximity to the affected country, this situation might pose a new pandemic threat to Indonesia.

Due to the high risk to animal and human health, control measures are essential. Aside from culling infected chickens, vaccination is used to control Avian Influenza in various countries (FAO, 2012). However, the use of vaccination for avian influenza is challenging. Avian influenza is a negative-sense, segmented, and single-stranded RNA virus(Fonville et al., 2016). The virus has two important surface glycoproteins, Hemagglutination (HA) and Neuraminidase (NA), which undergo continuous mutation and cause numerous genetic and antigenic variations (Cai et al., 2012; Sandie & Aris-Brosou, 2014). Swayne et al. (2014) have reviewed that antigenic matching between vaccine and field virus is critical in achieving optimal vaccine efficacy. Antigenic variation is used as the strategy of infectious pathogens to evade recognition of host antibodies by modifying their surface proteins, leading to inadequate vaccine protection against challenge viruses (Oli et al., 2020). A mismatch of antigenic variation between vaccine and field virus will not adequately protect against challenge viruses (Eggert & Swayne, 2010; Grund et al., 2011; Pfeiffer et al., 2010). Thus, monitoring Avian Influenza's antigenic variant is important to ensure vaccine efficacy.

Antigenic analysis takes place as a crucial step in selecting vaccine strains and designing the type of vaccine for Avian influenza. The steps for selecting vaccine strains and developing the vaccine are similar to the seasonal influenza vaccine in humans, as shown in Figure 1 (WHO, 2007). Details of requirements specific to the Avian Influenza vaccine are written in the Manual of Diagnostic Tests and Vaccines for Terrestrial Animals, twelfth edition 2023 by WOAH(2023). After isolation, viruses must

undergo screening tests such as Hemagglutination (HA), Haemagglutination inhibition (HI) test or enzyme-linked immunosorbent assays (ELISA). Further, molecular tests such as the RT-PCR using a specific primer are performed to confirm HA and NA sequences. Vaccine strains must be well-characterised and fulfil quality criteria (sterility, purity, freedom from extraneous agents). The final vaccine product must also comply with the safety and potency test. Over time, due to the improper use of vaccines or the immune escape ability of the virus due to the vaccine pressure, the current vaccine might inadequately protect the new circulating strain (Escorcia et al., 2008; Grund et al., 2011; Sitaras et al., 2014). To address this issue, several vaccine design strategies are used, such as creating a Differentiating Infected from Vaccinated Animals strategy (DIVA) (Capua et al., 2003), using several strains of AI in a vaccine (Bivalent) (Kim et al., 2023), subunit or DNA vaccine (Harder et al., 2023; Nielsen et al., 2023). Thus, surveillance programs focusing on antigenic variety have become necessary to evaluate vaccine efficacy.

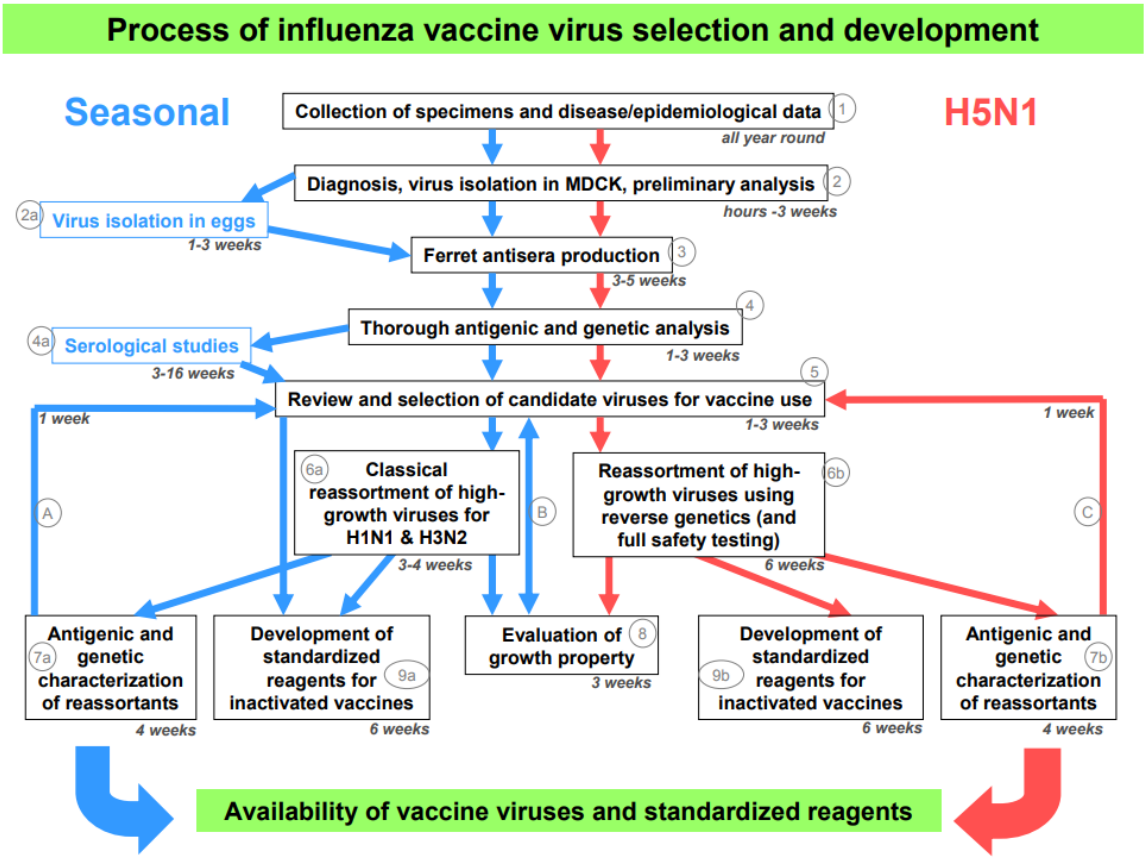

Figure 1 Diagram Process of Influenza H5N1 vaccine virus selection based on the World Health Organization (). Specimens and epidemiological data are first collected (1). The specimens are then diagnosed and grown in Embryonated Eggs (2, 2a). Ferret antisera are produced to undergo antigenic and genetic analysis (3, 4, 4a). Next, the high-growth isolates are reviewed for virus reassortment, evaluate the growth property, and developed the standardized reagents for the inactivated vaccine.

The vaccine's efficacy in protecting against transmission depends upon factors such as the antigen content, administration procedure and antigenic distance between the vaccine and challenge strain (Sitaras et al., 2016). The antigenic distance is the quantitative measure to determine the relations between two viruses from the reactions against the same antibody, for example, derived from the Hemagglutinin Inhibition (HI) assay, which determines the affinity of the antigen-antibody of AI virus

(Cai, Zhang, and Wan 2012; Lapedes and Farber, 2001). The antigenic distance between the viruses with similar HI profiles is closer and, in this case, smaller (Anderson et al. 2020). The distance is measured by antigenic units (AU), in which 1 AU equals log2 dilution in the HI assay (1AU = 2-fold dilution) (Zheng et al., 2022). Immunological assays are very crude as the results are usually averaged, and there are variations among individuals, often said to be low-resolution data (Barnett et al., 2012). To increase the resolution, antigenic distance is visualised with antigenic cartography, commonly used for influenza viruses (human, swine, equine, and avian) (Zheng et al., 2022). Antigenic cartography is projecting the antigens-antibody on a map (2 or 3 dimensions) by applying the multidimensional scaling method (MDS) (Cai et al., 2010). Antigen-antibody affinity is quantified to construct numerical spatial coordinates (Lapedes & Farber, 2001).

The antigenic cartography claimed by Smith et al., (2004) to offer a greater resolution than the raw HI table by visualization of antigen-antibody in a 2- or 3- Dimension that provides easier analysis of the HI values. The algorithm was based on (Lapedes & Farber, 2001) ordinal multidimensional scaling (MDS) that was modified to modification metric MDS. An antigenic evolution could also be observed by seeing clustering on the map). Smith et al. 2004 further modified the ordinal MDS used by Lapedes to" modification metric MDS" to position the antigen and antibody in the map. This new approach offers computational advantages, allowing us to work with large datasets and reduce running time. Antigenic cartography is used by the World Health Organization (WHO) influenza reference laboratories and Indonesia Virus Monitoring (IVM) Online as the standard tool in influenza surveillance. The latter is a web-based animal health laboratory network system in Indonesia dedicated to managing antigenic and genetic data related to the country's circulating Highly Pathogenic Avian Influenza (HPAI) viruses.

The vaccine is updated when there is a difference of at least two antigenic units with the recently circulated strain virus (Smith et al., 2004). The current method by Smith et al. (2004) is used as the basis of the website https://acmacs-web.antigenic-cartography.org, which is freely accessible. However, the website is sometimes out of order due to maintenance. Thus, this creates a hassle and delays when a new vaccine should be released. Following the website, a version of the R-acmac package has been published for use in R-studio. This R-studio version tackles technical problems such as low internet connection or website maintenance. Thus, allowing to researcher to analysis antigenic distance and cartography faster. However, distance distortion may occur since this method tries to project every point and reduce the dimension to 2- or 3-dimension. Cai et al. (2010) also reported that this method may contain relatively large errors due to the dependency on initial value. Consequently, it impacted the accuracy of this method for two-dimensional projection.

Calculating an accurate antigenic distance is essential, especially for selecting the strain used for the vaccine update. Inaccurate antigenic distance calculation may result in unnecessary vaccine updates or the selection of the wrong strain (Sitaras et al. 2017). To provide a better view of the antigenic distance calculation method, this research aims to evaluate the distortion between the R-acmac method and real distance in consequence to the vaccine strain selection.

MATERIAL AND METHODS

Dataset

The dataset used in this research is obtained from previous research. Raw Hemagglutination Inhibition (HI), Virus Neutralization (VN) or both titers are extracted from the supplementary materials and used to calculate antigenic distance.

The first dataset is obtained from Peeters et al. (2017). This study investigated the relation of genetic and antigenic differences among different HPAI H5N1. The datasets that were used are HI and VN titer. In brief, the titer was generated in chickens after vaccination with recombinant PR8-H5 influenza (PR8) inserted with HA genes from different low pathogenic H5N1 viruses. The information on the virus's name, abbreviation and clade for the first dataset used in this study were listed in Table 1. In addition, another set of antisera was obtained by vaccinating chickens with soluble trimeric HA (SHA3) protein. Detailed methods for generating antisera and conducting serological tests are outlined in the paper. Titers were averaged before calculating the antigenic distance. This dataset was selected for the primary analysis because this paper provides a full cross-immunity table. Twelve virus strains from various clades were used to analyse the antigenic distance. This condition represents a situation where there are multiple clades of virus circulating. The complete cross-immunity table presented an ideal scenario for analyzing antigenic distance calculation. Moreover, this paper provides the R-script for calculating the real distance without having any distortion.

Table 1. List of viruses from Peeters et al. 2017

Clade	Abbreviation	Full name
1	VN04	A/Vietnam/1194/04
1.1	Cam07	A/Cambodia/R0405050/2007
2.2	Hunan02	A/duck/Hunan/795/2002
2.2	MSD05	consensus based on clade 2.2
2.2	TT05	A/turkey/Turkey/1/2005
2.2.1	Eg10	A/Egypt/N03072/2010
2.2.1.1	Eg08	A/chicken/Egypt/0879-NLQP/2008
2.3.2.1	Hubei10	A/Hubei/1/2010
2.3.2.1	VN11	A/duck/Vietnam/QT801/2011
2.3.4	Anhui05	A/Anhui/1/2005
2.3.4.2	VN08	A/Vietnam/HN31432M/2008
7.1	VN0408	A/chicken/Vietnam/NCVD-04/2008

The second dataset was obtained from Pawestri et al. (2020). This paper characterized the genetic and antigenic properties using HI assays for twenty-five isolates of H5N1 Influenza from human patients in Indonesia between 2008 and 2015. The information on the virus's name, abbreviation, and clade for the second dataset used in this study were listed in Table 2. The isolates belong to the lineage of clade 2.1.3.x including subclade 2.1.3.2, 2.1.3.2a, 2.1.3.2b and 2.1.3.3. However, the twenty-five isolates were not tested against themselves. Instead, a panel of ferret antisera was raised against clade 2.1.3.1 (A/Chicken/South Sulawesi/157/2011), different clade 2.1.3.2 viruses (A/Indonesia/5/2005, A/Chicken/Central Java/51/2009, A/Chicken/East Java/121/2010, A/Chicken/West Java/30/2007 and A/Chicken/West Java/119/2010) and against clade 2.1.3.3 (A/Chicken/North Sumatra/72/2010). Methodologies for generating antisera and conducting serological tests are explained in detail in the

paper. This situation provides examples of a more realistic scenario in which disease monitoring was conducted. Although more viruses were obtained, there are only four genetic differences, all within one monophyletic lineage 2.1.3.x. The viruses were not tested with the full panel of antisera, resulting in incomplete cross-immunity HI data. This condition highlighted the different conditions with Peeters et al., (2017) data set .

Table 2. List of viruses from Pawestri

Clade	Abbreviation	Full name
2.1.3.2	05/05	A/ Indonesia/5/2005
2.1.3.2	7393/2008	A/Indonesia/NIHRD7393/2008
2.1.3.3	7781/2008	A/Indonesia/NIHRD7781/2008
2.1.3.2	7802/2008	A/Indonesia/NIHRD7802/2008
2.1.3.2b	8987/2008	A/Indonesia/NIHRD8987/2008
2.1.3.2a	9160/2009	A/Indonesia/NIHRD9160/2009
2.1.3.2b	9340/2009	A/Indonesia/NIHRD9340/2009
2.1.3.2b	9653/2009	A/Indonesia/NIHRD9653/2009
2.1.3.2b	9665/2009	A/Indonesia/NIHRD9665/2009
2.1.3.2	10364/2010	A/Indonesia/NIHRD10364/2010
2.1.3.2b	10459/2010	A/Indonesia/NIHRD10459/2010
2.1.3.2b	10529/2010	A/Indonesia/NIHRD10529/2010
2.1.3.2b	10728/2010	A/Indonesia/NIHRD10728/2010
2.1.3.2b	11046/2011	A/Indonesia/NIHRD11046/2011
2.1.3.2b	11073/2011	A/Indonesia/NIHRD11073/2011
2.1.3.2b	11198/2011	A/Indonesia/NIHRD11198/2011
2.1.3.2	11454/2011	A/Indonesia/NIHRD11454/2011
2.1.3.2a	12078/2012	A/Indonesia/NIHRD12078/2012
2.1.3.2a	12130/2012	A/Indonesia/NIHRD12130/2012
2.1.3.2b	12162/2012	A/Indonesia/NIHRD12162/2012
2.1.3.2a	12452/2012	A/Indonesia/NIHRD12452/2012
2.1.3.2a	13157/2013	A/Indonesia/NIHRD13157/2013
2.1.3.2b	13233/2013	A/Indonesia/NIHRD13233/2013
2.1.3.2b	13269/2013	A/Indonesia/NIHRD13269/2013
2.1.3.2a	14122/2014	A/Indonesia/NIHRD14122/2014
2.1.3.2b	14157/2014	A/Indonesia/NIHRD14157/2014

The third dataset was obtained from PT. Medion Farma Jaya Indonesia. Farmers sent samples to the diagnostic lab for further analysis when a disease occurred. This newly isolated virus was tested for cross-immunity compared to the current vaccine isolate. Nine viruses were collected and proceeded for the antigenic distance calculation. Among these, eight viruses belonged to the same clade and one from a distinct clade. M166 is the current strain used for the existing vaccine. However, only three-panel antisera were used for the HI assay. The information on the virus's name, abbreviation, and clade for the third dataset used in this study were listed in Table 2. The antisera were generated by vaccinating Specific Pathogen-Free (SPF) chickens with 0.5 ml of inactivated Avian Influenza water in oil (w/o) vaccine by intramuscular inoculation. This situation represents a realistic condition to evaluate whether the current vaccine needs updates.

Table 3. List of viruses from PT. Medion Farma Jaya

Clade	Abbreviation	Full name
2.3.2.1c	M166	A/quail/Jawa Barat/12-M166/2019
2.3.2.1c	M250	A/duck/Sumatera Barat/12-M250/2021
2.3.2.1c	M254	A/duck/Sumatera Barat/12-M254/2021
2.3.2.1c	M290	A/duck/Jawa Barat/12-M290/2022
2.3.2.1c	M291	A/duck/Banten/12-M291/2022
2.3.2.1c	M296	A/chicken/Banten/12-M296/2022
2.3.2.1c	M310	A/chicken/Jawa Barat/12-M310/2022
2.3.2.1c	M320	A/duck/Sulawesi Selatan/12-M320/2022
2.3.4.4b	M294	A/duck/Hulu Sungai Utara/A1222294-1/2022

Antigenic distance calculations

The serological data is analyzed in Rstudio. HI titer values are shown in log2 values. All the dataset is analyzed using Peeters et al., (2017) script that will be called the real distance and R_acmac package (Wilks S. 2023) that will be called R_acmac distance throughout this research. In both scripts, HI titers are converted to a distance matrix where the distance between a specific antigen and antiserum pair is represented by the differences between the log2 value of the maximum observed titer to that antiserum from any antigen and the titer of the antigen to the same antiserum.

The real distances are calculated based on the R_script written by Peeters et al., (2017) which is available in the Supplemental material (). The titer HI matrix was standardized to ensure all differences were ≥ 0. This means that negative values were treated as 0. The antigenic distance is then calculated based on Euclidean distance between two points calculation according to Sitaras et al. 2014 with the following equation:

$$\text{dist}_{ij} = \sqrt{\sum_{k=1}^{n}(m_{i,k} - m_{j,k})},$$

where *i, j* represents the antigen and antisera in an n-dimensional space. $m_{i,k}$ and $m_{j,k}$ are the coordinates of point in the *i* and *j* in the *k*-th dimension. Thus, the distance is calculated by summing the squared differences between coordinates of point and taking the square root of the result.

Another way to calculate distances is using R_acmac packages based on the following website: https://acorg.github.io/Racmacs/. HI titer, in log2 values, is used to determine the table distances. First, the maximal titer value against each serum is considered to have a zero distance. This means the antigen with maximal titer is the closest to the sera. The maximal titer value against each serum is called the column base. Then, the distance can be obtained by subtracting each serum's column base and log titer. This is called the table distance. Further, MDS converts the table distance into a dimensional representation to construct the map as the acmac object. One hundred optimization numbers are used throughout the experiment. The script to obtain distances using R_acmac package is written in the Supplement material ().

R_acmac package validation

The R_acmac package in R-studio is the alternative to the website https://acmacs-web.antigenic-cartography.org/. With the guidance of GitHub, R-script was made and used for all data antigenic distance calculation. The same data was analyzed in websites and R-studio to validate the R-script. The results (map and antigenic distance) from the R-studio and website sites were compared.

Distortion analysis

The absolute result from subtracting the real distance and R_acmac distance defines the distortion in this research. Distortion was then visualized using the corrplot package in R-studio.

Vaccine strain candidate selection

The three data listed above are categorized in the variability of the clade as high (Peeters et al., 2017), moderate (Pawestri et al., 2020) and low (Medion). A threshold of 4 AU is used to eliminate a virus as a vaccine candidate for all types of data. This means a virus will be eliminated as a vaccine strain candidate when the antigenic distance of this strain to any other circulating strain is more than 4 AU.

Out of the three datasets, only the Medion dataset particularly aimed to evaluate the current vaccine strain versus the other new isolates. The two other datasets were not aimed at selecting a strain as the vaccine candidate in the paper. Therefore, two scenarios were applied to choose a vaccine strain from the high and moderate datasets. Firstly, the biological interpretation would not be considered in the analysis. Biological interpretations here mean the viruses clade. The first scenario is applied to assume that there is no previous vaccine, and we need to choose one or several candidates for a vaccine strain despite of any clade circulating. Second, biological interpretation will then be considered. One virus strain is selected and assumed as the current vaccine strain, and some strain of interest is selected as the newly circulating strain. This scenario mimics the reality of vaccine evaluation and determining whether the vaccine strain needs to be updated. Clades of the current vaccine will be considered when analysis and suggestions will be made based on biological interpretation.

RESULTS

A. Distortion Between R_acmac and Real Distance

There were three datasets analyzed in this study, which were Peeters et al., (2017) (representing high variability of AI clade), Pawestri et al., (2020)(moderately variable in AI clade) and Medion's data (low variability of AI clade). Antigenic distances in each dataset were compared pairwise for real distance calculation and R_acmac by the absolute value to acquire the distortion.

Peeters et al., (2017) data set

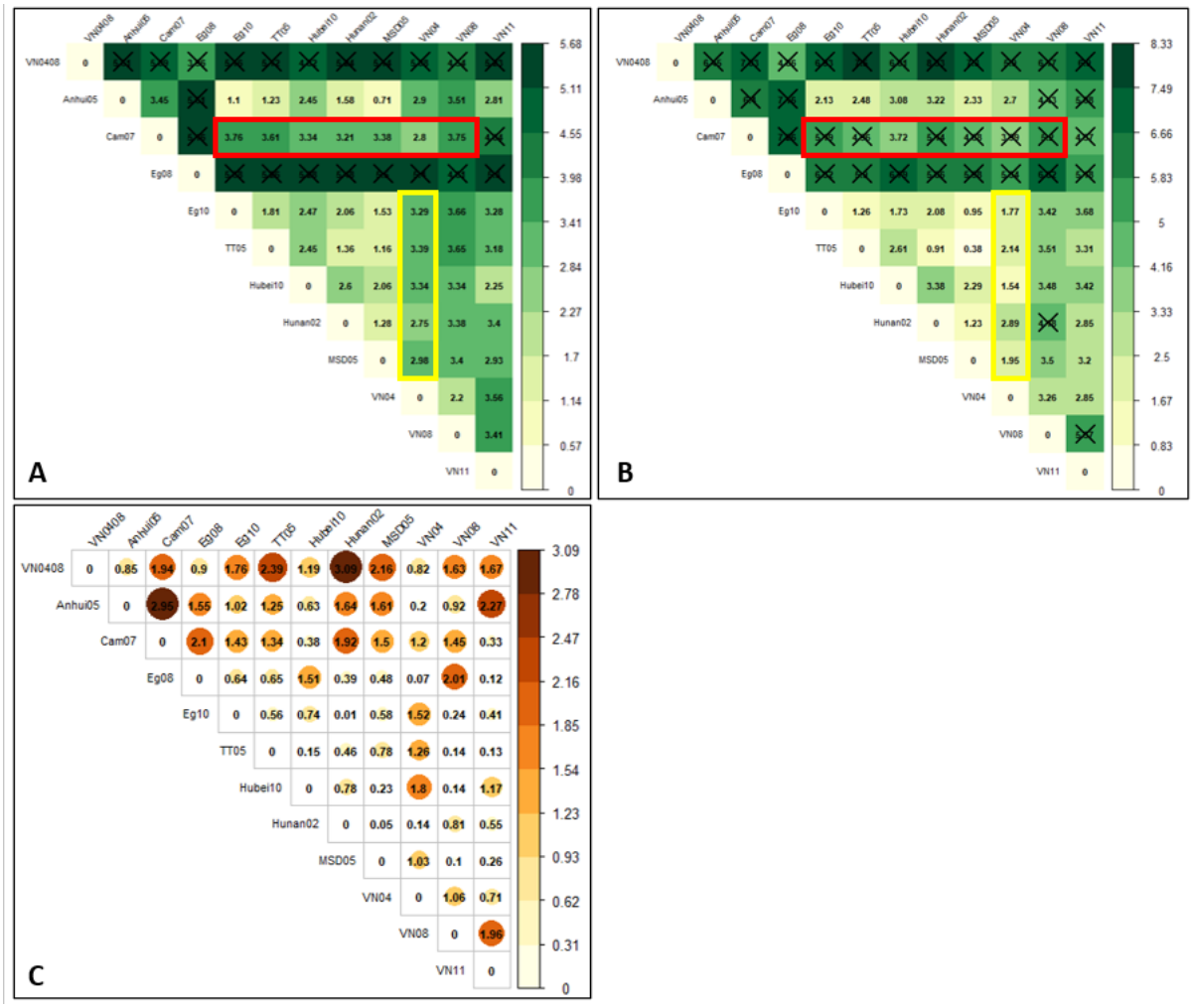

Figure 2. Antigenic distance of Peeters dataset with A. real distance, B. R-acmac and C. Antigenic distance distortion. On the x-axis is the virus, and the number within the cell gives the antigenic distance between the virus and an antiserum derived from exposure against the virus listed on the y-axis with a color-coded scheme. The higher the antigenic distance, the darker the colour becomes. The X symbol indicates an antigenic distance value above 4 where the serum no longer is deemed protective against the virus listed on the x-axis. . Yellow and red boxes show a distortion by smaller and larger values, respectively. Figure 2.C. show distortion by the colour-coded scheme and circle representing the distortion's width. More distortions are observed at the widest circle with the darkest colour.

Twelve strains, consisting of eight AI clades, were analyzed from HI and VN assays from Peeters data. Distortion in antigenic distance was observed between the first and R-acmac distances. The real distance refers to the real distance, and the second refers to the R_acmac method (Figure 2A and 2B). The crossed cells indicate that the antigenic distance between viruses exceeds 4 AU. The antigenic

distance from the R-acmac distance generates an overall higher distance than the first (red boxes). However, the value is not always higher. Some antigenic distances are shown smaller values than the real distance (yellow box). Three virus strains show a high range of distortion over other strains: VN0408, Cam07 and Eg08 (Figure 2C). The highest distortion is observed at VN0408 vs Hunan02 (3.09).

Pawestri et al., (2020) dataset

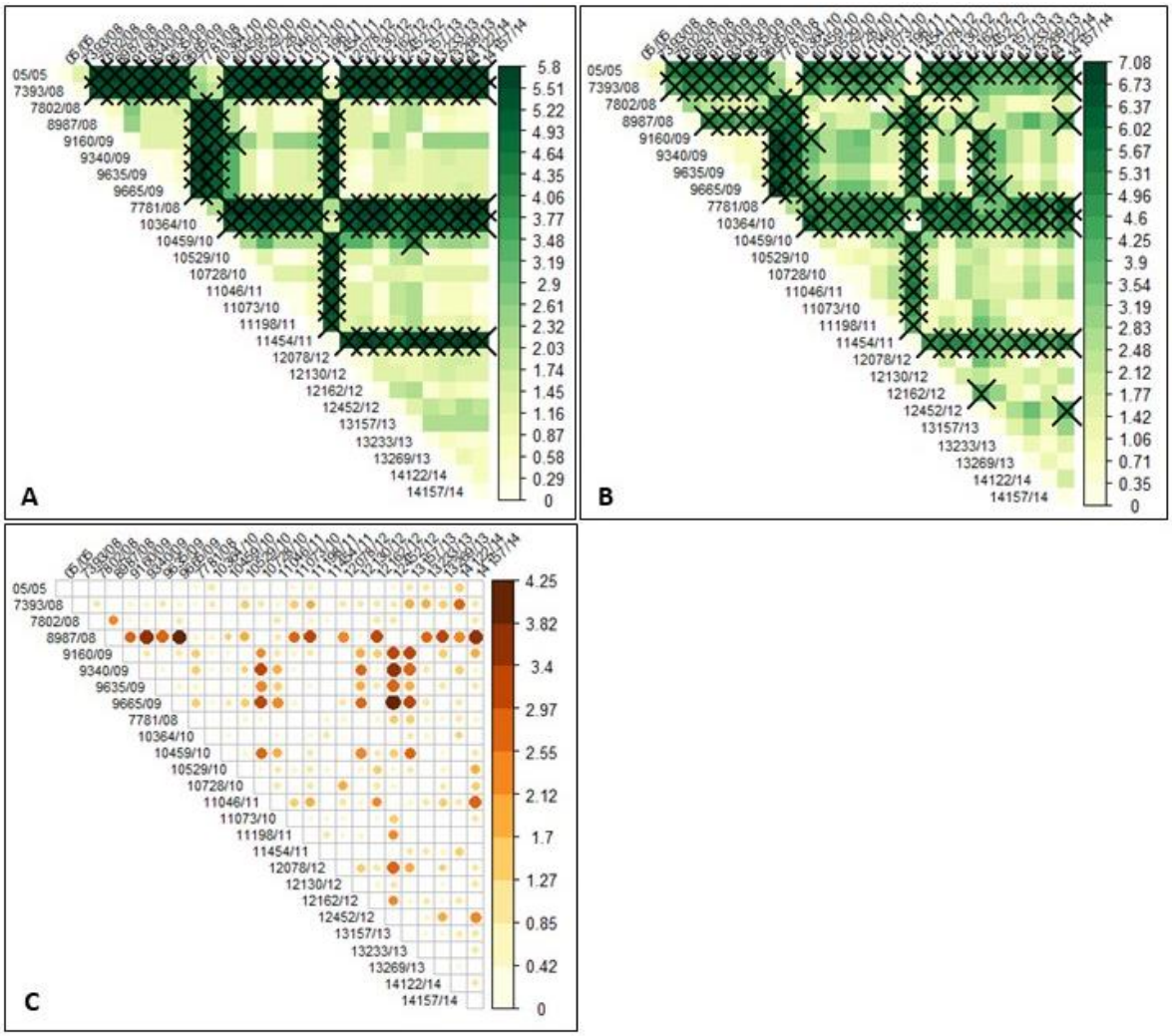

Figure 3. Antigenic distance of Pawestri et al., (2020) datasetset with A. real distance, B. R-acmac and C. Antigenic distance distortion. On the x-axis is the virus, and the number within the cell gives the antigenic distance between the virus and an antiserum derived from exposure against the virus listed on the y-axis with a color-coded scheme. The higher the antigenic distance, the darker the colour becomes. The X symbol indicates an antigenic distance value above 4 where the serum no longer is deemed protective against the virus listed on the x-axis. Figure 3.C. show distortion by the colour-coded scheme and circle representing the distortion's width. More distortions are observed at the widest circle with the darkest colour.

More strains were observed within Pawestri et al., (2020) dataset. Twenty-six strains, consisting of four different AI clades, were analyzed from HI assays of ferret antisera. More antigenic distances between viruses exceeds 4 AU in the R-acmac distance, marked by crosses in Figures 2A and 2B. A similar trend was observed in Pawestri et al., (2020) dataset: the antigenic distance from the R-acmac distance generates higher distances than the real distance. The 8987/08 strain was observed to have a high range of distortion over other strain, followed by 9160/09, 9340/09, 9635/09, 9665/09. The

highest distortion was observed at strain 9665/09 vs 12452/12 (4.25). Meanwhile, the 05/05 strain was observed to have an overall low distortion ranging from 0.14 to 1.62.

Medion dataset

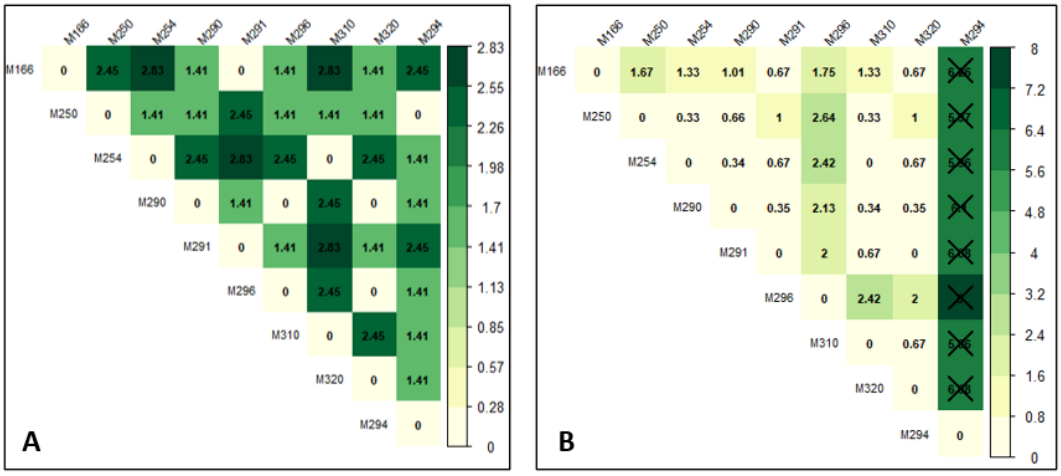

Figure 4 Antigenic distance of A. Real Distance, B. R_acmac. On the x-axis is the virus, and on the y-axis is the antigenic distance with a colour-coded scheme. The higher the antigenic distance, the darker the colour becomes. The X symbol indicates an antigenic distance value above 4 AU.

Medion data represents how vaccine effectivity towards new strains will be evaluated. Both methods agree that no values above 4 AU exist except for M294. An obvious difference observed is that no other viruses can cover strain M294 in the R_acmac distance. Further, the distance of M294 is the highest towards M296. However, In the real distance calculation, a surprising result was observed. M294 distance was all lower than four with an exact distance value as M250.

B. Consequences of vaccine selection

Both Peeters et al., (2017) and Pawestri et al., (2020) paper was not aimed at selecting a vaccine strain. Hence, two scenarios are used to help with the analysis: Not considering biological interpretation and considering biological interpretation.

1. Not considering the biological interpretation

The situation is assumed that there are no previous vaccines, multiple clades are circulating, and we need to pick up a candidate/s as the vaccine strain. This is done by creating a threshold of 4AU. The virus with an antigenic distance of more than 4 AU is considered ineffective. Viruses with broader protectivity (more value than 4AU) are preferable in this case. Viruses with more coverage towards other strains are selected as the vaccine candidate. Figures 2 and 3 (A and B) show that viruses with antigenic distances of more than 4AU are crossed and will be eliminated as vaccine candidates.

It was observed that different numbers of viruses can be candidates using Real distance and R_acmac. The evaluation is then by quantifying the strain with the most antigenic distance value of less than 4 AU over other strains. The number shown in Figure 5. is the number of strains with antigenic distances less than four towards the virus. The highest numbers of protectivity are used to select the candidate

of vaccine strain (10). It was shown that eight viruses can be selected as vaccine candidates from real distance calculation because they can protect towards 10 other viruses. Meanwhile, only two candidates can be selected from the R-acmac distance calculation while keeping the same standard. However, both methods agree that VN0408 and Eg08 are the most distant strains over other viruses (low number).

	VN0408	Anhui05	Cam07	Eg08	Eg10	TT05	Hubei10	Hunan02	MSD05	VN04	VN08	VN11
Real distance	2	10	9	2	10	10	10	10	10	10	10	9
Racmac	1	7	3	1	9	9	10	8	9	10	6	7

Figure 5. Real distance vs R_acmac for Peeters data. The figure shows the table of strains with antigenic distances less than four towards the virus between Real distance and R_acmac.

A similar observation is also seen from Pawestry data that different numbers of viruses can be candidates using Real distance and R_acmac distance (Figure 6). Firstly, the highest coverage for the real distance is towards 21 strains, while R-acmac is towards 22 strains. Thus, we also included 21 for the candidate of R_acmac. It is shown that there are supposedly 18 viruses that have more coverage in the real distance calculation. Meanwhile, only 12 viruses have more coverage towards other strains in the R_acmac. The summary of selected vs non-selected candidate number is represented in Figure 7.

	05/05	7393/08	7802/08	8987/08	9160/09	9340/09	9635/09	9665/09	7781/08	10364/10	10459/10	10529/10	10728/10
Real distance	5	5	21	21	20	21	21	21	5	5	19	21	21
R_acmac	5	14	21	14	18	19	19	17	7	5	19	22	21

	11046/11	11073/10	11198/11	11454/11	12078/12	12130/12	12162/12	12452/12	13157/13	13233/13	13269/13	14122/14	14157/14
Real distance	21	21	21	5	21	21	21	21	20	21	21	21	21
R_acmac	21	22	21	5	22	21	19	17	22	22	22	22	19

Figure 6 Pawestri data of Real distance vs R_acmac. The figure shows the table of strains with antigenic distances less than four towards the virus between Real distance and R_acmac.

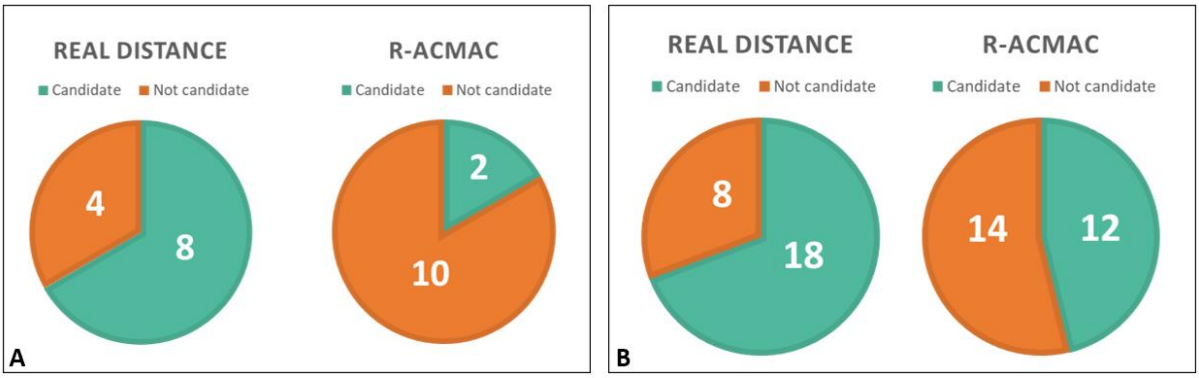

Figure 7. Summary of the number of selected candidate (teal) and not candidate (orange) in real distance and R-acmac distances. A. Peeters, et al. (2017) and B. Pawestri, et al. (2020)

2. Considering the biological interpretation

The second scenario applies to the highly and moderately variable data to further analyse the consequence of the vaccine candidate selection. In this scenario, the biological interpretation (clade of the current vaccine vs circulating strain matters) is included in the analysis. The situation is where there is an existing vaccine and an introduction of new or several strains within or out of the same clade.

MSD05 was selected as the vaccine strain because this strain is a consensus sequence of clade 2.2 AI H5N1 from the Peeters et al. (2017) dataset. Using the real distance, MSD05 cannot protect the other two strains, VN0408 (clade 7.1) and Eg08 (clade 2.2.1.1). Meanwhile, using the R-acmac distance, there is one more strain that MSD05 cannot protect: the Cam07 (1.1). This can be seen as the crossed cells in Figure 2A and 2B. It is observed that there is a different conclusion when using the antigenic distance from the first and R-acmac distances.

For the Pawestri et al. (2020) dataset, the 05/05 virus (2.1.3.2) was selected because it was the virus isolated from early human cases in Indonesia. WHO has listed this strain as a candidate for the vaccine virus. The Pawestri et al., (2020) dataset contains a more homogeneous clade, which was the monophyletic lineage of 2.1.3.X. Both methods agree that isolate 05/05 is not protective against other viruses except for 7781/08 (2.1.3.3) and 10364/10 (2.1.3.2) which cells are uncrossed in Figure 3A and 3B. Thus, the conclusion remains the same for both the real distance and R-acmac distance.

DISCUSSION

Understanding the antigenic variations by carefully observing the antigenic distance between circulating strains is essential to determine vaccine strain selection in avian influenza vaccine strategy development. It is a critical initial step in the selection process to decide when and with which strain to update an AI vaccine in poultry. The analysis of antigenic distances between strains in various datasets using real distance and R_acmac methods reveals distortion that has implications for vaccine selection. In this study, two methods of antigenic distance calculation are compared using two previous serological data. The first data was acquired from Peeters et al. (2017), which aimed to see the correlation between genetic and antigenic differences among H5N1. The second data was acquired from Pawestri et al. (2020), which aimed to characterise the genetic and antigenic properties of H5N1 Influenza from human patients in Indonesia. A surprising result was observed from the Medion data, which will be explained later. Thus, Medion data is excluded for distortion analysis. Therefore, this study requires two scenarios to help analyse the differences between the two methods in real-life situations.

This research confirms that calculating distance using R-acmac distorts the real distance and affects vaccine strain selection. Distortion is the absolute subtraction of the real distance and R-acmac distance. The distortion observed within the datasets of Peeters et al. (2017) and Pawestri et al. (2020) varies from very small to high (0.05 to 4.25). This distortion indicates a consequence of the vaccine strain selection, further confirmed by the scenario made up for analysis.

In the first scenario, no biological interpretations are considered. This situation represents multiple clades circulating; a strain with wide coverage is preferable because we want a vaccine covering various clades. From Peeters et al. (2017) dataset, eight candidates can be selected using the real distance, while only two candidates appeared to be able to be selected using R-acmac distance. The Pawestri et al. (2020) dataset also observed a different conclusion. For example, the 8987/07 strain. The real distance implies the potential selection of this strain as a vaccine candidate, whereas the R_acmac distance suggests otherwise, indicating it may not be a suitable candidate. This conclusion presents a potential limitation in vaccine strain choices. The antigenic distance between candidate vaccine strains and the circulating strains is just one of several factors to consider when choosing a vaccine strain. A wider range of choice for potential vaccine strains, indicated by the antigenic distances, simplifies several other constraints in the actual production of the vaccine. A practical implication is to look at the virus's yield. For instance, if the 8987/07 strain provides a higher yield than another virus candidate but is excluded as a vaccine candidate by the R-acmac method, an alternative candidate with a lower yield might be selected. High virus yield is an important parameter for vaccine production (Aslam et al., 2023; Lee et al., 2017). Optimal virus yield could still be achieved through serial passages in embryonated eggs (Lee et al., 2017). However, this may extend production time and lead to mutation in the HA gene, affecting viral antigenicity (Ping et al., 2016).

The second scenario considers biological interpretation. A virus is selected as the existing vaccine strain, while new genetically different virus strains are circulating. Thus, we will decide whether our current vaccine needs to be updated. The results confirm that different methods lead to different conclusions for the Peeters et al. (2017) dataset. MSD05, clade 2.2, is shown to have an antigenic distance of more than 4 AU to VN0408 (clade 7.1) and Eg08 (clade 2.2.1.1). This outcome indicates the need to update the MSD05 vaccine strain. This outcome aligns with the in vivo cross-protection findings in the Peeters et al. (2017) paper. MSD05, VN11 and Eg08 vaccines were challenged with TT05, VN11, Eg08 and VN0408 virus. The paper reported that complete protection against mortality was achieved when the antigenic distance between the vaccine and challenge strain was less than 4AU. Since they are both different clades from the current vaccine and no other strain have an antigenic distance below 4 towards these strains, the MSD05 (2.2) strain does not need to be replaced. However, constructing a combination vaccine with multiple clades, including VN0408 (7.1) and Eg08 (2.2.1.1) with MSD05 is suggested as this development vaccine strategy may increase vaccine efficacy towards the strain with high antigenic distance. This strategy has been proven to be protective towards the heterologous strain and reducing viral shedding (Hu et al., 2021; Kandeil et al., 2018; Kim et al., 2023). However, R-acmac distance shown that MSD05 is also not protective towards Cam07 (1.1). This result shows a different conclusion. While the real distance analysis suggests excluding Cam07 from the vaccine combination, this finding suggests that including Cam07 may be necessary. A classical alternative is to find another strain that has an antigenic distance below 4 towards all of these strains, which may take longer time and effort. However, with the current vaccine development, alternatives such as subunit vaccines could be used as another strategy (Dey et al., 2023; Nielsen et al., 2023; Oli et al., 2020).

A different conclusion was obtained by looking at the R-acmac distance outcome. In the Pawestri et al. (2020) dataset, both antigenic distance calculation methods agreed that 05/05 vaccine strains have more than 4 AU across most strains and that this vaccine must be updated. The 05/05 virus strain was Indonesia's early isolated virus from human-infected H5N1 HPAI. These results indicate that this strain might not offer protection against the emerging sublineages HPAI 2.1.3.2a and 2.1.3.2b. Both methods for calculating antigenic distance indicate no variation, likely due to the small range of distortion observed from both methods. Consequently, conclusions drawn from the Pawestri et al. (2020) dataset are the same.

The R_acmac method uses Metric MDS developed by Smith et al., (2004). This algorithm generates a random initial vector for each virus or antiserum in either 2 or 3 dimensions. The process involves iteratively adjusting the vector coordinates of each virus and antiserum to minimize the error function, ensuring that the calculated distances closely match the measured real distances. The method requires a pre-defined dimension L as input representing the virus and antiserum in the L-dimensional vector (Cai et al. 2010). Therefore, distortion is evident and observed in this research. The distortion, however, may lead to a different or similar conclusion as the real distance calculation. The extent of distortion can potentially be reduced by adjusting the optimization number in R-acmac. Optimization finds the optimal arrangement of antigens and antisera to reflect their antigenic similarity (Wilks S. 2023). However, determining the appropriate number of optimizations needs to be done carefully. In the ideal situation, conducting sufficient runs would result in the lowest stress that will end up with similar solutions (Wilks S. 2023). In general, a higher number of optimizations is preferable. An interesting outcome was observed when performing antigenic distance calculation in both the website and R-studio versions. With the same 100 optimizations, the web version consistently produced identical results in each run. However, different results were obtained in the R-studio versions after multiple runs of antigenic distance calculations. This outcome indicates that the web version has a fixed random seed, whereas the R-studio does not. Thus, careful attention needs to be given due to the potential variability in the results of antigenic distance using the R-studio version.

The R-acmac distance offers the advantage of visualizing on a map rather than raw serologic titer numbers. Antigenic cartography also aids in the visualization of the classification of antigens into groups/ clusters. Wang et al. (2023) analyzed antigenic variation in AI H9N2 in China. He reported that antigenic cartography categorized strains into three antigenic groups: Group 1, Group 2, and Group 3. Clustering revealed that Group 3 had underwent antigenic drift, attributed to 11 potential antigenic amino acid residue differences between the old and new clades in the HA sequence, including H48Q, G72E, S127D, D135G, Q146R, N149G, E163G, T182R, N183D, T202I, N238D. This major shifted due to antigenic drift may result in immune escape causing low immunity efficacy of the used vaccine strain (Sitaras et al., 2014). Pawestri et al. (2020) used antigenic cartography and identified two clusters of distinct antigenic groups. This finding leads to the recommendation that creating a bivalent vaccine representing both groups was required to optimize protection against the two antigenic groups. Kim et al. (2023) reported that using one vaccine containing two different strains (bivalent) induces high level of immunogenicity and protective efficacy.

Calculating the real distance is advantageous as it represents the distance without distortion in the multidimensional space. Inaccurate distance measurements can significantly influence vaccine update decisions, potentially resulting in either unnecessary updates or the retention of a vaccine with health and economic consequences (Sitaras, et al. 2017). While R-acmacs claims an increase in accuracy of the distance measurements by reducing the number of dimensions, The distance accuracy can also be increased by raising sera in more than one animal, without the drawbacks of the reduction in dimensions. According to Sitaras et. al. (2017), inter-animal variation in immune titers after vaccination can significantly vary the calculated distances between strains. Even using sera from two animals results in significant improvement in the accuracy of the antigenic map. This way, preserving the original distance matrix maintains the geometry and avoids distortion from reducing multidimensionality.

When calculating real distance, it is important to utilize a complete cross-immunity table or at least consists of a panel of antisera representing various clades to calculate the antigenic distance. This observation is evident in Medion dataset, particularly with the genetically distinct M294 strain belonging to the new clade 2.3.4.4b. While it was expected that the antigenic distance of M294 would be higher than another strains, the actual result revealed that the antigenic distance is consistently below 4 and very similar to M250. This outcome might be attributed to the absence of panel antisera specific to the newclade itself. During the calculation process, the centered matrix is derived by subtracting the raw HI titer from the mean titer. Consequently, the centred matrix of M294 and M250 were identical, influencing the subsequent calculation step and resulting in the exact antigenic distance for both strains. Conversely, the R-acmac distance results aligned more closely with the hypothesis, indicating that all antigenic distance from M294 are above 4. This alignment is expected because the raw HI data shows consistently lower titers compared to the others. However, further studies are required to validate these findings.

The real distance calculation is limited in terms of visualization, as it merely provides numerical values. To address this limitation, (Sitaras et al., 2014) proposed an alternative map that represents distances in a 2D space without reducing dimensionality. However, these maps may not effectively reveal clustering or grouping patterns. In response, alternative methods such as an antigenic dendrogram combined with a heat map can serve as an option to visualized antigenic similarity as suggested by A. C. K. Lai et al., (2012). Additionally, clustering algorithms including k-means, clara, hierarchical, EM, hcmodel, Spectral, subspace, optics, dbscan (Rodriguez et al., 2019) could be explored for further development to analyze clustering, even though it falls outside the scope of this research.

Researchers commonly use a four-fold increase in antigenic units in antigenic distance calculation as a benchmark to identify influenza antigenic drift (Anderson et al., 2018; Barnett et al., 2012; Cai et al., n.d., 2010, 2012a; M Fouchier, 2010; Y. Wang et al., 2016). These findings are taken further to confirmed the relation of antigenic distance vs vaccine efficacy in vivo in chickens by Nguyen et al., (2018). The findings revealed an inverse correlation between percent survival of chickens post-vaccination challenge and antigenic distance. Specifically, when It was found that percent survival of chickens post-vaccine challenge demonstrated an inverse relationship between survival rate and antigenic distance. the antigenic distance ranged from 3.7 to >4.9 HAU, less than 50% of chickens survived. Overall, comparing antigenic distances and percent survival suggested a trend towards lower survival rates as the antigenic distance between the challenge virus and vaccine virus increased.

Further, Sitaras et al. (2016) suggest evaluating the vaccine efficiency in stopping transmission of HPAI H5N1 by studying the vaccine dose (low or high) and the implication on HI titer. (Karo-karo et al., 2022)

Malek & Hoque, (2022) conducted a mathematical modeling study on bird flu in poultry farms, incorporating vaccination and treatment. A Susceptible-Exposed-Infectious-Treatment-Recovered (SEITR) model has been built, integrating the vaccine effective rate (1-ε) into the model. The findigs revealed that basic reproductive numbers (R0) are influenced by the transmission rate in the host individual and the vaccine effectivity rate. In summary, using vaccines that have efficiency of more than 70% will eradicate the disease along with proper diagnosis, the use of antiviral, hygienic farming system and density of the problem. without proper implementation, surveillance, robust vaccine strategies, and stringent bio-security measures, vaccination might contribute to the rapid evolution and antigenic change of H5N1 viruses, allowing them to evade vaccine protection. This underlined the urgent need of an effective animal health control and prevention measures, such as sustainable and efficacious vaccination strategies for poultry in controlling Avian Influenza (Capua & Marangon, 2007). Thus, highlighting the role of calculating accurate antigenic distance during vaccine strain selection.

It is essential to acknowledge the limitations of this research. This research uses three complexities in the variability in clades of viruses. To gain a better understanding of the distortion between real antigenic distance and R-acmac distance, additional data will be necessary. Collecting more serological datasets, either from the HI or VN test, is suggested for further study. Additionally, data simulation will help to create more datasets. Unfortunately, simulated HI data could not be generated during the study. Initially, deriving a biological interpretation from random data proved challenging. Cai et al. (2012) addressed this challenge by randomly designating 5% of the HI values as low or missing reactors (indicating weak immunological reactions). The study also accounted for variations induced by noises, such as inconsistencies in virus titration or experimental conditions. This method may be used to create HI simulation data that helps validate the outcome of this research further. In addition, the clustering algorithm may be beneficial to study as an alternative to visualizing the antigenic distances in future studies.

Following the use of this distance in a company view. Real antigenic distance calculation is suggested as it gives direct information about the distance between viruses. One consideration must be considered, which is using full cross-immunity data, or at least there should be all represented antisera for each virus clade. As shown in this research outcome, a similar HI titer pattern may give the same result of antigenic distance. Furthermore, conducting a trial using more chicken to create serum has proven to give more precise results of antigenic distance. The company can perform and validate this to enhance the precision of antigenic distance to help select the vaccine strain. However, the use of R-acmac still has advantages in visualization, especially when many strains are isolated. Using R-acmac might be inevitable as it is easier to look at the map. However, R-acmac distance result needs to be carefully optimized. It has been shown that there is a potential variability in the results of antigenic distance using the R-studio version. Validation, for example, by running the code 10 times and averaging the outcome may be used to minimize distortion. Given its advantage and validation, the use of R-acmac still possesses a distance distortion that always needs to be taken into consideration while judging the map and concluding the result. Further research about the different outcomes may also be an interesting analysis for further study to understand more about how this algorithm works.

CONCLUSION

This research has shown that distortion is evident in R-acmac distance, potentially impacting vaccine strain selection. This shows the significance of accurately calculating the antigenic distance to enhance precision in vaccine strain selection and measuring a few chickens per virus-serum combination to reduce error. The practical implications of this outcome can be utilized for vaccine development and avian influenza control strategies to formulate a broad protection vaccine that aligns effectively with circulating strains.

REFERENCES

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

Aslam, S., Rajendran, M., Kriti, D., Kurland, A., Johnson, J., van Bakel, H., Krammer, F., García-Sastre, A., & Ayllon, J. (2023). Generation of a high yield vaccine backbone for influenza B virus in embryonated chicken eggs. *Npj Vaccines*, *8*(1). https://doi.org/10.1038/s41541-023-00603-3

Barnett, J. L., Yang, J., Cai, Z., Zhang, T., & Wan, X. F. (2012). AntigenMap 3D: An online antigenic cartography resource. *Bioinformatics*, *28*(9), 1292–1293. https://doi.org/10.1093/bioinformatics/bts105

Basuno, E., Yusdja, Y., & Ilham, N. (2010). Socio-economic impacts of avian influenza outbreaks on small-scale producers in Indonesia. *Transboundary and Emerging Diseases*, *57*(1–2), 7–10. https://doi.org/10.1111/j.1865-1682.2010.01121.x

Cai, Z., Zhang, T., & Wan, X. F. (2010). A computational framework for influenza antigenic cartography. *PLoS Computational Biology*, *6*(10). https://doi.org/10.1371/journal.pcbi.1000949

Cai, Z., Zhang, T., & Wan, X. F. (2012a). Antigenic distance measurements for seasonal influenza vaccine selection. *Vaccine*, *30*(2), 448–453. https://doi.org/10.1016/j.vaccine.2011.10.051

Cai, Z., Zhang, T., & Wan, X. F. (2012b). Antigenic distance measurements for seasonal influenza vaccine selection. *Vaccine*, *30*(2), 448–453. https://doi.org/10.1016/j.vaccine.2011.10.051

Cai, Z., Zhang, T., & Wan, X.-F. (n.d.). *Concepts and applications for influenza antigenic cartography*. http://sysbio.cvm.msstate.edu/AntigenMap.

Capua, I., & Marangon, S. (2007). Control and prevention of avian influenza in an evolving scenario. *Vaccine*, *25*(30 SPEC. ISS.), 5645–5652. https://doi.org/10.1016/j.vaccine.2006.10.053

Capua, I., Terregino, C., Cattoli, G., Mutinelli, F., & Rodriguez, J. F. (2003). Development of a DIVA (Differentiating Infected from Vaccinated Animals) strategy using a vaccine containing a heterologous neuraminidase for the control of avian influenza. *Avian Pathology*, *32*(1), 47–55. https://doi.org/10.1080/0307945021000070714

Dey, P., Ahuja, A., Panwar, J., Choudhary, P., Rani, S., Kaur, M., Sharma, A., Kaur, J., Yadav, A. K., Sood, V., Suresh Babu, A. R., Bhadada, S. K., Singh, G., & Barnwal, R. P. (2023). Immune Control of Avian Influenza Virus Infection and Its Vaccine Development. In *Vaccines* (Vol. 11, Issue 3). MDPI. https://doi.org/10.3390/vaccines11030593

Eggert, D., & Swayne, D. E. (2010). Single vaccination provides limited protection to ducks and geese against H5N1 high pathogenicity avian influenza virus. *Avian Diseases*, *54*(4), 1224–1229. https://doi.org/10.1637/9410-052810-Reg.1

Escorcia, M., Vázquez, L., Méndez, S. T., Rodríguez-Ropón, A., Lucio, E., & Nava, G. M. (2008). Avian influenza: Genetic evolution under vaccination pressure. *Virology Journal*, *5*.

https://doi.org/10.1186/1743-422X-5-15
Fonville, J. M., Fraaij, P. L. A., De Mutsert, G., Wilks, S. H., Van Beek, R., Fouchier, R. A. M., & Rimmelzwaan, G. F. (2016). 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
Gaide, N., Lucas, M. N., Delpont, M., Croville, G., Bouwman, K. M., Papanikolaou, A., van der Woude, R., Gagarinov, I. A., Boons, G. J., De Vries, R. P., Volmer, R., Teillaud, A., Vergne, T., Bleuart, C., Le Loc'h, G., Delverdier, M., & Guérin, J. L. (2022). Pathobiology of highly pathogenic H5 avian influenza viruses in naturally infected Galliformes and Anseriformes in France during winter 2015-2016. *Veterinary Research*, *53*(1), 11. https://doi.org/10.1186/s13567-022-01028-x
Grund, C., Abdelwhab, E. S. M., Arafa, A. S., Ziller, M., Hassan, M. K., Aly, M. M., Hafez, H. M., Harder, T. C., & Beer, M. (2011). Highly pathogenic avian influenza virus H5N1 from Egypt escapes vaccine-induced immunity but confers clinical protection against a heterologous clade 2.2.1 Egyptian isolate. *Vaccine*, *29*(33), 5567–5573. https://doi.org/10.1016/j.vaccine.2011.01.006
Harder, T., de Wit, S., Gonzales, J. L., Ho, J. H. P., Mulatti, P., Prajitno, T. Y., & Stegeman, A. (2023). Epidemiology-driven approaches to surveillance in HPAI-vaccinated poultry flocks aiming to demonstrate freedom from circulating HPAIV. In *Biologicals* (Vol. 83). Academic Press. https://doi.org/10.1016/j.biologicals.2023.101694
Hu, J., Peng, P., Li, J., Zhang, Q., Li, R., Wang, X., Gu, M., Hu, Z., Hu, S., Liu, X., Jiao, X., Peng, D., & Liu, X. (2021). Single Dose of Bivalent H5 and H7 Influenza Virus-Like Particle Protects Chickens Against Highly Pathogenic H5N1 and H7N9 Avian Influenza Viruses. *Frontiers in Veterinary Science*, *8*. https://doi.org/10.3389/fvets.2021.774630
ICTV. (2022, July). *Orthomyxoviridae*. https://ictv.global/taxonomy/taxondetails?taxnode_id=202203956&taxon_name=Alphainfluenzavirus%20influenzae
Kandeil, A., Sabir, J. S. M., Abdelaal, A., Mattar, E. H., El-Taweel, A. N., Sabir, M. J., Khalil, A. A., Webby, R., Kayali, G., & Ali, M. A. (2018). Efficacy of commercial vaccines against newly emerging avian influenza H5N8 virus in Egypt. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-28057-x
Karo-karo, D., Bodewes, R., Restuadi, R., Bossers, A., Agustiningsih, A., Stegeman, J. A., Koch, G., & Muljono, D. H. (2022). Phylodynamics of Highly Pathogenic Avian Influenza A(H5N1) Virus Circulating in Indonesian Poultry. *Viruses*, *14*(10). https://doi.org/10.3390/v14102216
Kim, H., Cho, H. K., Kang, Y. M., Sagong, M., An, S., Kim, S., Lee, Y. J., & Kang, H. M. (2023). Protective efficacy of a bivalent H5 influenza vaccine candidate against both clades 2.3.2.1 and 2.3.4.4 high pathogenic avian influenza viruses in SPF chickens. *Vaccine*, *41*(17), 2816–2823. https://doi.org/10.1016/j.vaccine.2023.03.028
Lai, A. C. K., Wu, W. lan, Lau, S. ying, Guan, Y., & Chen, H. (2012). Two-dimensional antigenic dendrogram and phylogenetic tree of avian influenza virus H5N1. *FEMS Immunology and Medical Microbiology*, *64*(2), 205–211. https://doi.org/10.1111/j.1574-695X.2011.00884.x
Lai, S., Qin, Y., Cowling, B. J., Ren, X., Wardrop, N. A., Gilbert, M., Tsang, T. K., Wu, P., Feng, L., Jiang, H., Peng, Z., Zheng, J., Liao, Q., Li, S., Horby, P. W., Farrar, J. J., Gao, G. F., Tatem, A. J., & Yu, H. (2016). Global epidemiology of avian influenza A H5N1 virus infection in humans, 1997-2015: A systematic review of individual case data. In *The Lancet Infectious Diseases* (Vol. 16, Issue 7, pp. e108–e118). Lancet Publishing Group. https://doi.org/10.1016/S1473-3099(16)00153-5
Lapedes, A., & Farber, R. (2001). The geometry of shape space: Application to influenza. *Journal of Theoretical Biology*, *212*(1), 57–69. https://doi.org/10.1006/jtbi.2001.2347
Lee, I., Il Kim, J., Park, S., Bae, J. Y., Yoo, K., Yun, S. H., Lee, J. Y., Kim, K., Kang, C., & Park, M. S. (2017). Single PA mutation as a high yield determinant of avian influenza vaccines. *Scientific Reports*, *7*. https://doi.org/10.1038/srep40675
M Fouchier, R. A. (2010). *Use of Antigenic Cartography in Vaccine Seed Strain Selection* (Vol. 54, Issue 1).

Malek, A., & Hoque, A. (2022). Mathematical modeling of bird flu with vaccination and treatment for the poultry farms. *Comparative Immunology, Microbiology and Infectious Diseases*, *80*. https://doi.org/10.1016/j.cimid.2021.101721

Nguyen, D. T., Shepard, S. S., Burke, D. F., Jones, J., Thor, S., Van Nguyen, L., Nguyen, T. D., Balish, A., Hoang, D. N., To, T. L., Iqbal, M., Wentworth, D. E., Spackman, E., Van Doorn, H. R., Davis, C. T., & Bryant, J. E. (2018). Antigenic characterization of highly pathogenic avian influenza A(H5N1) viruses with chicken and ferret antisera reveals clade-dependent variation in hemagglutination inhibition profiles. *Emerging Microbes and Infections*, *7*(1). https://doi.org/10.1038/s41426-018-0100-7

Nielsen, S. S., Alvarez, J., Bicout, D. J., Calistri, P., Canali, E., Drewe, J. A., Garin-Bastuji, B., Gonzales Rojas, J. L., Gortázar, C., Herskin, M., Michel, V., Miranda Chueca, M. Á., Padalino, B., Roberts, H. C., Spoolder, H., Stahl, K., Velarde, A., Winckler, C., Bastino, E., … Viltrop, A. (2023). Vaccination of poultry against highly pathogenic avian influenza – part 1. Available vaccines and vaccination strategies. *EFSA Journal*, *21*(10). https://doi.org/10.2903/j.efsa.2023.8271

of the, F. U. (2012). *H5N1 Highly Pathogenic Avian Influenza Global overview ::: January-March 2012 ::: Issue no. 31*.

Oli, A. N., Obialor, W. O., Ositadimma, M., Ifeanyichukwu, Odimegwu, D. C., Okoyeh, J. N., Emechebe, G. O., Adejumo, S. A., & Ibeanu, G. C. (2020). Immunoinformatics and vaccine development: An overview. In *ImmunoTargets and Therapy* (Vol. 9, pp. 13–30). Dove Medical Press Ltd. https://doi.org/10.2147/ITT.S241064

Pawestri, H. A., Nugraha, A. A., Han, A. X., Pratiwi, E., Parker, E., Richard, M., van der Vliet, S., Fouchier, R. A. M., Muljono, D. H., de Jong, M. D., Setiawaty, V., & Eggink, D. (2020). Genetic and antigenic characterization of influenza A/H5N1 viruses isolated from patients in Indonesia, 2008–2015. *Virus Genes*, *56*(4), 417–429. https://doi.org/10.1007/s11262-020-01765-1

Peeters, B., Reemers, S., Dortmans, J., de Vries, E., de Jong, M., van de Zande, S., Rottier, P. J. M., & de Haan, C. A. M. (2017). Genetic versus antigenic differences among highly pathogenic H5N1 avian influenza A viruses: Consequences for vaccine strain selection. *Virology*, *503*, 83–93. https://doi.org/10.1016/j.virol.2017.01.012

Pfeiffer, J., Suarez, D. L., Sarmento, L., To, T. L., Nguyen, T., & Pantin-Jackwood, M. J. (2010). Efficacy of commercial vaccines in protecting chickens and ducks against H5N1 highly pathogenic avian influenza viruses from Vietnam. *Avian Diseases*, *54*(SUPPL. 1), 262–271. https://doi.org/10.1637/8715-031909-Reg.1

Ping, J., Lopes, T. J. S., Neumann, G., & Kawaoka, Y. (2016). Development of high-yield influenza B virus vaccine viruses. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(51), E8296–E8305. https://doi.org/10.1073/pnas.1616530113

Pramuwidyatama, M. G., Indrawan, D., Boeters, M., Poetri, O. N., Saatkamp, H. W., & Hogeveen, H. (2023). Economic impact of highly pathogenic avian influenza outbreaks in Western Java smallholder broiler farms. *Preventive Veterinary Medicine*, *212*. https://doi.org/10.1016/j.prevetmed.2022.105833

Rodriguez, M. Z., Comin, C. H., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. da F., & Rodrigues, F. A. (2019). Clustering algorithms: A comparative approach. *PLoS ONE*, *14*(1). https://doi.org/10.1371/journal.pone.0210236

Sandie, R., & Aris-Brosou, S. (2014). Predicting the emergence of H3N2 influenza viruses reveals contrasted modes of evolution of HA and NA antigens. *Journal of Molecular Evolution*, *78*(1), 1–12. https://doi.org/10.1007/s00239-013-9608-6

Sitaras, I., Kalthoff, D., Beer, M., Peeters, B., & De Jong, M. C. M. (2014). Immune escape mutants of highly pathogenic avian influenza H5N1 selected using polyclonal sera: Identification of key amino acids in the HA protein. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0084628

Sitaras, I., Rousou, X., Kalthoff, D., Beer, M., Peeters, B., & De Jong, M. C. M. (2016). Role of vaccination-induced immunity and antigenic distance in the transmission dynamics of highly pathogenic avian influenza H5N1. *Journal of the Royal Society Interface*, *13*(114).

https://doi.org/10.1098/rsif.2015.0976
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. In *New Series* (Vol. 305, Issue 5682).
Swayne, D. E., Spackman, E., & Pantin-Jackwood, M. (2014). Success factors for avian influenza vaccine use in poultry and potential impact at the wild bird-agricultural interface. In *EcoHealth* (Vol. 11, Issue 1, pp. 94–108). Springer New York LLC. https://doi.org/10.1007/s10393-013-0861-3
Swayne, D. E., & Suarez, D. L. (2000). Highly pathogenic avian influenza. In *Rev. sci. tech. Off. int. Epiz* (Vol. 19, Issue 2).
Theary, R., San, S., Davun, H., Allal, L., & Lu, H. (2012). New outbreaks of H5N1 highly pathogenic avian influenza in domestic poultry and wild birds in Cambodia in 2011. *Avian Diseases*, *56*(4 SUPPL.1), 861–864. https://doi.org/10.1637/10195-041012-ResNote.1
Wan, X. F. (2012). Lessons from Emergence of A/Goose/Guangdong/1996-Like H5N1 Highly Pathogenic Avian Influenza Viruses and Recent Influenza Surveillance Efforts in Southern China. In *Zoonoses and Public Health* (Vol. 59, Issue SUPPL.2, pp. 32–42). https://doi.org/10.1111/j.1863-2378.2012.01497.x
Wang, X., Liu, K., Guo, Y., Pei, Y., Chen, X., Lu, X., Gao, R., Chen, Y., Gu, M., Hu, J., Liu, X., Hu, S., Jiao, X. an, Liu, X., & Wang, X. (2023). Emergence of a new designated clade 16 with significant antigenic drift in hemagglutinin gene of H9N2 subtype avian influenza virus in eastern China. *Emerging Microbes and Infections*, *12*(2). https://doi.org/10.1080/22221751.2023.2249558
Wang, Y., Davidson, I., Fouchier, R., & Spackman, E. (2016). Antigenic Cartography of H9 Avian Influenza Virus and Its Application to Vaccine. In *Source: Avian Diseases* (Vol. 60, Issue 1).
WHO. (2007). *A Description of the Process of Seasonal and H5N1 Influenza Vaccine Virus Selection and Development* .
WHO. (2023). *Cumulative number of confirmed human cases for avian influenza A(H5N1) reported to WHO*. https://www.who.int/publications/m/item/cumulative-number-of-confirmed-human-cases-for-avian-influenza-a(h5n1)-reported-to-who--2003-2023-21-december-2023
WOAH. (2023). Avian Influenza (Including Infection with High Pathogenicity Avian Influenza Viruses). In *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals* (12th ed.). https://www.woah.org/fileadmin/Home/eng/Health_standards/tahm/3.03.04_AI.pdf
Zheng, Y., Guo, Y., Li, Y., Liang, B., Sun, X., Li, S., Xia, H., & Ping, J. (2022). The molecular determinants of antigenic drift in a novel avian influenza A (H9N2) variant virus. *Virology Journal*, *19*(1). https://doi.org/10.1186/s12985-022-01755-9