Virologica Sinica 40 (2025) 288–291

Contents lists available at ScienceDirect

Virologica Sinica

journal homepage: www.keaipublishing.com/en/journals/virologica-sinica
www.virosin.org

Letter

Evaluating the performance of the PREDAC method in flu vaccine recommendations over the past decade (2013–2023)

Yousong Peng [a,1], Lei Yang [b,1], Weijuan Huang [b], Mi Liu [c], Xiao Ding [d], Xiangjun Du [e], Yuelong Shu [f,*], Taijiao Jiang [g,*], Dayan Wang [b,*]

[a] Bioinformatics Center, College of Biology, Hunan Provincial Key Laboratory of Medical Virology, Hunan University, Changsha, 410082, China
[b] National Institute for Viral Disease Control and Prevention, China CDC, Beijing, 102206, China
[c] Jiangsu Institute of Clinical Immunology, The First Affiliated Hospital of Soochow University, Suzhou, 215006, China
[d] Suzhou Institute of Systems Medicine, Suzhou, 215123, China
[e] Department of Biostatistics and Systems Biology, School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, Guangdong, 510275, China
[f] Key Laboratory of Pathogen Infection Prevention and Control (MOE), State Key Laboratory of Respiratory Health and Multimorbidity, National Institute of Pathogen Biology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 102629, China
[g] Guangzhou National Laboratory, Guangzhou, 510005, China

Dear Editor,

Influenza viruses cause significant mortality and morbidity in humans. Vaccination is currently the most effective way to combat the virus (Perofsky and Nelson, 2020). Unfortunately, the influenza virus frequently changes its antigenicity through rapid mutations, leading to decreased vaccine efficacy or even failure. To improve vaccine effectiveness, it is necessary to monitor antigenic variation and update vaccine strains when significant antigenic variation occurs (Perofsky and Nelson, 2020; Malik et al., 2024). To this end, the World Health Organization (WHO) convenes technical consultations with global flu experts every February and September to recommend flu vaccines for the upcoming flu seasons in the Northern and Southern Hemispheres. Despite the great efforts of WHO, mismatches between recommended vaccine strains and circulating strains occur frequently (Du et al., 2012).

With the rapid development of DNA sequencing technology, several sequence-based methods have been developed to model the antigenic evolution of influenza viruses, showing potentials of improving flu vaccine recommendations (Du et al., 2012; Hayati et al., 2023; Huddleston and Bedford, 2024; Shah et al., 2024). We previously developed a computational method called PREDAC (PREDict Antigenic Cluster) to predict antigenic clusters—groups of viruses with similar antigenicity—of influenza A(H3N2) viruses based on HA protein sequences (Du et al., 2012; Peng et al., 2020). PREDAC not only achieves high accuracy in predicting antigenic variations but also models the dynamics of antigenic patterns of influenza viruses in the form of antigenic clusters (Du et al., 2012; Liu et al., 2015; Peng et al., 2017; Zhai et al., 2024).

The PREDAC method consists of three steps (Fig. 1A). First, the antigenic relationship between viruses is predicted with PREDAV (PREDict Antigenic Variant), a naive Bayes classifier that uses 12 physicochemical features extracted from the HA1 (the immunodominant part of the HA protein) protein sequence. Second, an antigenic correlation network (ACnet) is constructed by connecting viruses with predicted similar antigenicity. Finally, antigenic clusters are identified through network clustering. Based on PREDAC, we further proposed a novel framework for flu vaccine strain recommendations (Du et al., 2012). Briefly, the method works as follows (Fig. 1B): First, all available HA protein sequences of influenza A (H3N2) viruses from public databases like GISAID are collected before the time of vaccine recommendations by WHO (usually January or August of each year) and classified into antigenic clusters using PREDAC. Then, the monthly antigenic dynamics are analyzed. If a novel antigenic cluster emerges and meets two conditions: i) it continues increasing over the past three months and ii) it represents more than 15% of the virus population in the current month, the representative strain of the novel antigenic cluster is recommended as a new vaccine strain. Otherwise, the previous vaccine strain is recommended. It has been demonstrated that coupling PREDAC with large-scale HA sequencing of influenza A(H3N2) viruses in Chinese mainland could help improve vaccine recommendations in retrospective testing (Du et al., 2012).

Since 2013, PREDAC has been used to recommend vaccine strains for influenza A (H3N2) viruses. Here, we summarize its performance over the past ten years (2013–2023). A total of 19 vaccine recommendations have been made using the PREDAC method. These recommendations

* Corresponding authors.
E-mail addresses: dayanwang@cnic.org.cn (D. Wang), jiang_taijiao@gzlab.ac.cn (T. Jiang), yshu@cnic.org.cn (Y. Shu).
[1] Yousong Peng and Lei Yang contributed equally to this work.

https://doi.org/10.1016/j.virs.2025.01.004
Received 13 November 2024; Accepted 24 January 2025
Available online 28 January 2025

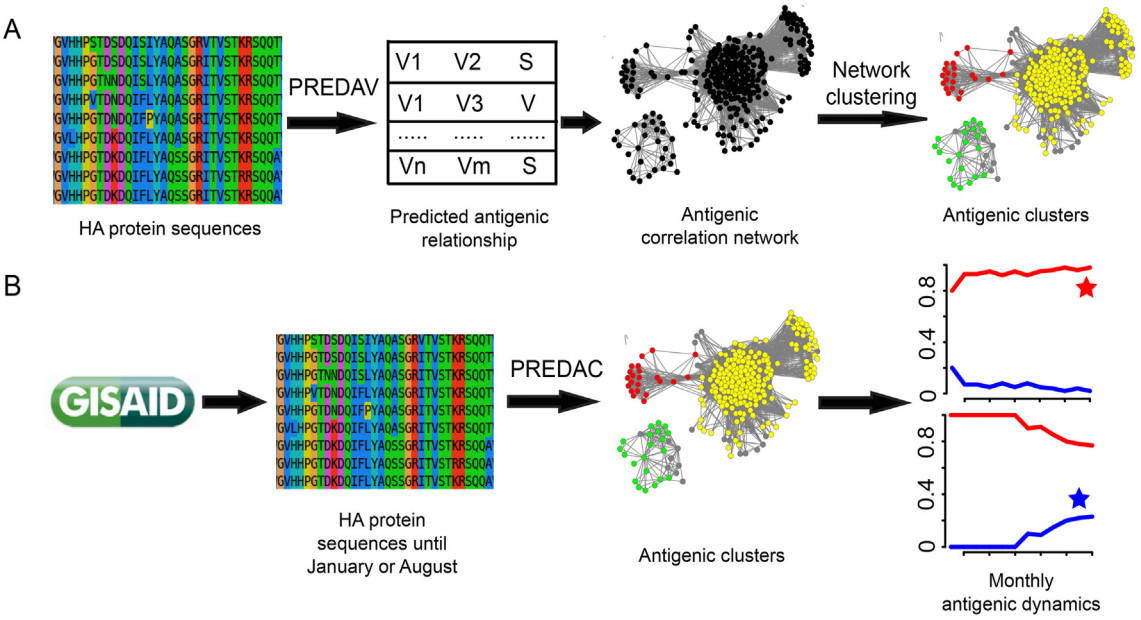

Fig. 1. Overview of PREDAC workflow (**A**) and its applications in influenza vaccine recommendations (**B**).

were validated using WHO-reported antigens circulating in the subsequent flu season after the recommendation date. As shown in Fig. 2, in six cases—including February 2013, September 2014, both February and September of 2015, February 2016, and February 2022—the recommended vaccine strains had the same antigenicity as that of the dominant circulating strains (marked by ticks). For example, in February 2015,

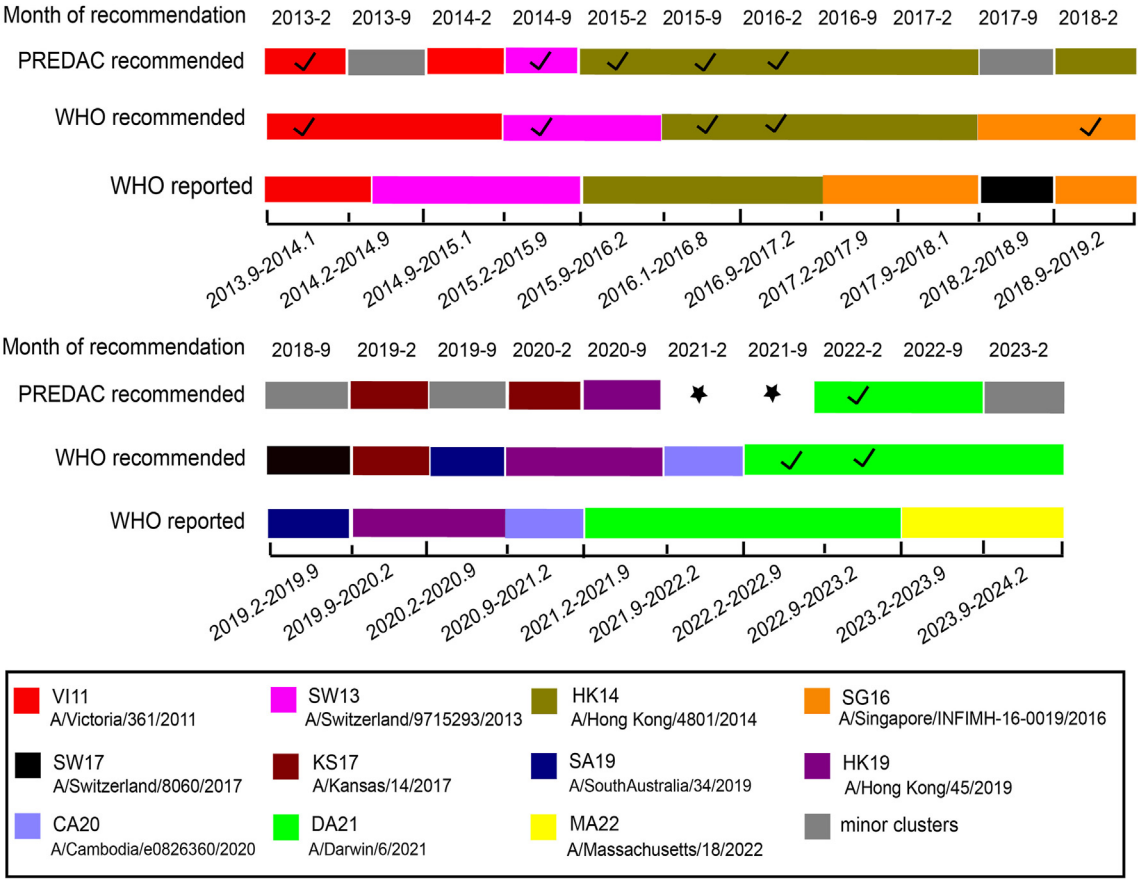

Fig. 2. Validation of vaccine recommendations of influenza A(H3N2) viruses by PREDAC method and WHO. Antigens were colored as indicated by the figure legend in the bottom. The gray refers to minor antigenic clusters which didn't predominate in any influenza season. The ticks refer to that the recommended vaccine strains had the same antigenicity as the dominant strains; the stars refer to that the recommended vaccine strains were antigenic similar to the dominant strains; the gray circles refer to no recommendations by PREDAC in 2021 due to the COVID-19. The WHO reported antigens were compiled from the vaccine recommendation reports by WHO.

PREDAC recommended the A/Hong Kong/4801/2014-like strain as the vaccine for use in the 2015–2016 Northern Hemisphere influenza season (Figs. 2 and 3A), which aligned with the reported dominant strain for that season. In another six cases (marked by stars), such as in February 2014 and September 2016, the recommended vaccine strains were antigenically similar to the dominant strains, although some antigenic drift had occurred. For instance, in February 2014, PREDAC recommended A/Victoria/361/2011-like strains for use in the 2014–2015 Northern Hemisphere influenza season, which were antigenically similar to the dominant strain A/Switzerland/9715293/2013 for that season (Supplementary Table S1). We also evaluated WHO's performance, and similar to PREDAC, in 12 instances out of 21 influenza seasons, WHO's recommended strains had the same or similar antigenicity as the dominant strains (marked by ticks or stars).

We next analyzed the reasons for mismatched vaccine recommendations made by PREDAC. The primary reason was the lack of sequences or epidemic records of the future prevalent strains at the time of recommendation, which accounted for most of mismatched cases (6/7), including September 2018, February 2019, February and September 2020, September 2022, and February 2023 (Fig. 3, Supplementary Fig. S1). For example, in September 2022, the antigenic cluster DA21 (A/Darwin/6/2021-like) had dominated for the past half-year (Fig. 3B), while the antigenic cluster MA22 (A/Massachusetts/18/2022-like) began circulating in February 2023 (Fig. 2). It was difficult to recommend MA22-like strains as vaccine candidates in September 2022.

The second reason was the inaccurate antigenic clustering by PREDAC, which often led to errors in identifying novel antigenic clusters or a failure to recommend new clusters, such as in the vaccine recommendations of September 2017, 2018, and February 2023. However, in some cases, the recommended vaccine strains were still antigenically similar to the dominant strains, such as those from September 2013 and 2019. For example, in February 2023, a minor antigenic cluster named NL22 (A/Netherlands/00034/2022-like viruses) was identified through antigenic clustering (Fig. 3C). It began circulating in early 2022, with its prevalence steadily increasing throughout the year (Fig. 3C), peaking in October and November. Although its prevalence decreased in December 2022 and January 2023, NL22 was still recommended as a vaccine candidate for the 2023–2024 Northern Hemisphere influenza season, based on the assumption that DA21 had been circulating for over two years and was unlikely to persist. Unfortunately, further analysis revealed that NL22 was not a novel cluster but a part of DA21, as the representative strain of NL22, i.e., A/South Australia/389/2022 (belonging to the clade 3C.2a1b.2a.2b), was antigenically similar to A/Darwin/6/2021 (the representative strain of DA21) (WHO, 2023).

The third reason was the significant antigenic heterogeneity of the virus, observed during the period from 2017 to 2019. As reported by the WHO, the viruses during that time primarily belonged to the 3C.2a1, 3C.2a2, and 3C.3a clades, which were antigenically distinct from each other. Additionally, the antigenic characterization of 3C.2a viruses was technically challenging because most of these viruses did not agglutinate red blood cells, making it difficult to determine the antigenic diversity within the clade (WHO, 2018; WHO, 2019). PREDAC did not predict any major antigenic clusters during this period (2017–2019). Instead, it identified multiple minor antigenic clusters (Fig. 3D, Supplementary Fig. S1). Making vaccine strain recommendations based on antigenic clustering was challenging. In the case of September 2017, the minor antigenic cluster, which had been steadily increasing over the past three months and reached a proportion of over 10%, was selected as the vaccine candidate (Fig. 3D).

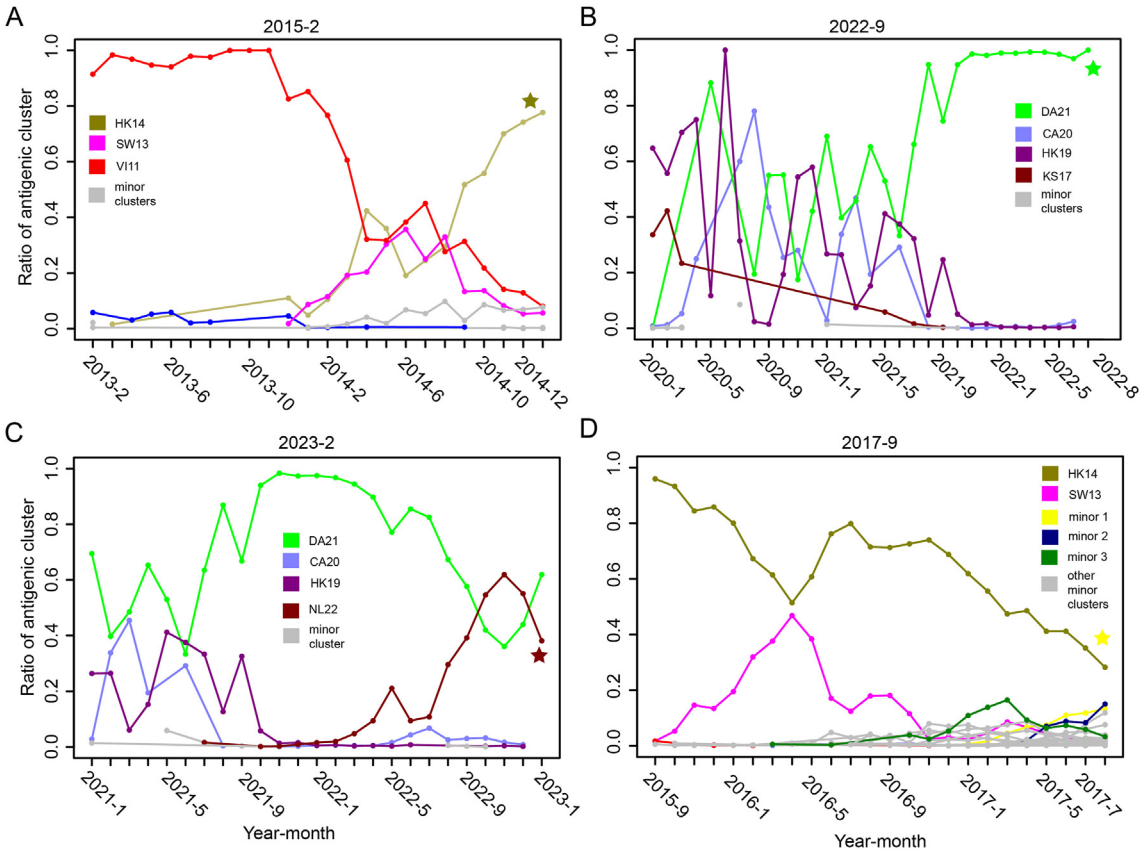

Fig. 3. Examples of vaccine recommendations by PREDAC method. **A** February 2015, **B** September 2022, **C** February 2023, D September 2017. The major antigenic clusters were colored according to the legend. The minor antigenic clusters were colored in gray, except some minor antigenic clusters displayed in panel **D**. No HA protein sequences in the month of vaccine recommendations were available in public databases due to the delay of sequence submission (Supplementary Fig. S2).

Based on the above mentioned analyses, the vaccine recommendations by PREDAC could be improved in several ways. First, new methods are needed to predict the fitness of antigenic clusters and to forecast their dynamics for the following year. This is particularly important because most inaccurate vaccine recommendations result from the absence or limited viral sequences or epidemic records of novel antigenic clusters at the time of recommendation. Second, timely sequence submission to public databases may aid in analyzing novel antigenic clusters. Regrettably, the submission of sequences to public databases significantly lags behind virus isolation (Supplementary Figs. S2 and S3). A median delay of 90 days was observed between virus isolation and submission to the GISAID database. However, this lag has slightly improved in recent years (Supplementary Fig. S2). The HA protein sequences from the three months preceding the vaccine recommendation period are crucial for inferring the future evolutionary dynamics of the virus and are key for accurate vaccine recommendations. However, the proportion of these sequences relative to all sequences from the flu season is less than 20% (Supplementary Fig. S3), which hinders accurate flu vaccine recommendations. A recent report showed that enhanced global solidarity in genomic surveillance of respiratory viruses would significantly strengthen the ability to rapidly detect novel viruses, even with a substantially reduced number of sequenced viruses (Jong et al., 2023). It is suggested to strengthen the flu surveillance and sequencing infrastructure globally, particularly in developing countries.

Third, the antigenic clustering method can be enhanced to more accurately identify novel antigenic clusters, particularly during the early circulation of novel clusters. At this stage, only a few viruses are available, and can be easily grouped with older antigenic clusters in the clustering process. Recently, a novel method called PREDAC-CNN, based on deep learning, was developed to predict antigenic clusters and has been shown to capture novel clusters more accurately (Meng et al., 2024). Applying PREDAC-CNN may improve flu vaccine recommendations.

Fourth, additional data such as genetic evolution analysis results (key amino acid mutations, phylogenetic tree analysis, selective pressure analysis, etc.) and the epidemic history of antigenic clusters should be incorporated to assist in vaccine recommendations. In addition, experimental methods for antigenicity characterization such as hemagglutination inhibition (HI) assay and microneutralization assay are important for improving the accuracy of recommended vaccine strains. Finally, novel methods are needed to select the best vaccine candidate from an antigenic cluster that contains a large number of viruses.

In summary, PREDAC has performed comparably to the WHO in recommending vaccines for influenza A(H3N2) viruses over the past ten years. It is a semi-automated and user-friendly method compared to the more complex and time-consuming traditional methods used by the WHO. PREDAC can serve as a useful supplement to the traditional approach for flu vaccine recommendations. However, accurately recommending the correct flu vaccines remains a challenge for both PREDAC and the WHO. More intensive surveillance of antigenic variations in the influenza virus, along with advanced computational methods, are needed to improve the accuracy of vaccine recommendations.

FOOTNOTES

This work was supported by the Major Project of Guangzhou National Laboratory (GZNL2024A01002), National Key Plan for Scientific Research and Development of China (2022YFC2303802), National Natural Science Foundation of China (32170651 & 32370700), and Hunan Provincial Natural Science Foundation of China (2024JJ2015). Prof. Yongsong Peng is an editorial board member for *Virologica Sinica* and was not involved in the editorial review or the decision to publish this article. The authors declare that they have no competing interests. We thank members in Peng Lab for helpful discussions on the manuscript.

All data used in the study are available in the GISAID database. Supplementary data to this article can be found online at https://doi.org/10.1016/j.virs.2025.01.004.

REFERENCES

Du, X., Dong, L., Lan, Y., Peng, Y., Wu, A., Zhang, Y., Huang, W., Wang, D., Wang, M., Guo, Y., Shu, Y., Jiang, T., 2012. Mapping of H3N2 influenza antigenic evolution in China reveals a strategy for vaccine strain recommendation. Nat. Commun. 3, 709.

Hayati, M., Sobkowiak, B., Stockdale, J.E., Colijn, C., 2023. Phylogenetic identification of influenza virus candidates for seasonal vaccines. Sci. Adv. 9, eabp9185.

Huddleston, J., Bedford, T., 2024. Timely vaccine strain selection and genomic surveillance improves evolutionary forecast accuracy of seasonal influenza A/H3N2. medRxiv. https://doi.org/10.1101/2024.09.11.24313489.

Jong, S.P.J., E, Nichols, B., Anniek, R., Parker, E., Mitesser, V., Happi, C., Jong, M.D.,X., Han, A., Russell, C., 2023. Global solidarity in genomic surveillance improves early detection of respiratory virus threats. medRxiv. https://doi.org/10.1101/2023.11.01.23297901.

Liu, M., Zhao, X., Hua, S., Du, X.J., Peng, Y.S., Li, X.Y., Lan, Y., Wang, D.Y., Wu, A.P., Shu, Y.L., Jiang, T.J., 2015. Antigenic patterns and evolution of the human influenza a (H1N1) virus. Sci. Rep. 5, 14171.

Malik, S., Asghar, M., Waheed, Y., 2024. Outlining recent updates on influenza therapeutics and vaccines: a comprehensive review. Vaccine X 17, 100452.

Meng, J., Liu, J., Song, W., Li, H., Wang, J., Zhang, L., Peng, Y., Wu, A., Jiang, T., 2024. PREDAC-CNN: predicting antigenic clusters of seasonal influenza A viruses with convolutional neural network. Briefings Bioinf. 25, bbae033.

Peng, Y., Wu, A., Meng, J., Yang, L., Wang, D., Shu, Y., Jiang, T., 2020. Automated recommendation of the seasonal influenza vaccine strain with PREDAC. Biosafety and Health 2, 117–119.

Peng, Y., Li, X., Zhou, H., Wu, A., Dong, L., Zhang, Y., Gao, R., Bo, H., Yang, L., Wang, D., Lin, X., Jin, M., Shu, Y., Jiang, T., 2017. Continual antigenic diversification in China leads to global antigenic complexity of avian influenza H5N1 viruses. Sci. Rep. 7, 43566.

Perofsky, A.C., Nelson, M.I., 2020. The challenges of vaccine strain selection. Elife 9, e62955.

Shah, S.A.W., Palomar, D.P., Barr, I., Poon, L.L.M., Quadeer, A.A., McKay, M.R., 2024. Seasonal antigenic prediction of influenza A H3N2 using machine learning. Nat. Commun. 15, 3833.

WHO, 2018. Recommended composition of influenza virus vaccines for use in the 2019 southern hemisphere influenza season. https://www.who.int/publications/m/item/recommended-composition-of-influenza-virus-vaccines-for-use-in-the-2019-southern-hemisphere-influenza-season. (Accessed 13 November 2024).

WHO., 2019. Recommended composition of influenza virus vaccines for use in the 2019–2020 northern hemisphere influenza season. https://www.who.int/publications/m/item/recommended-composition-of-influenza-virus-vaccines-for-use-in-the-2019-2020-northern-hemisphere-influenza-season. (Accessed 13 November 2024).

WHO., 2023. Recommended composition of influenza virus vaccines for use in the 2024 southern hemisphere influenza season. https://www.who.int/publications/m/item/recommended-composition-of-influenza-virus-vaccines-for-use-in-the-2024-southern-hemisphere-influenza-season. (Accessed 13 November 2024).

Zhai, K., Dong, J., Zeng, J., Cheng, P., Wu, X., Han, W., Chen, Y., Qiu, Z., Zhou, Y., Pu, J., Jiang, T., Du, X., 2024. Global antigenic landscape and vaccine recommendation strategy for low pathogenic avian influenza A (H9N2) viruses. J. Infect. 89, 106199.