

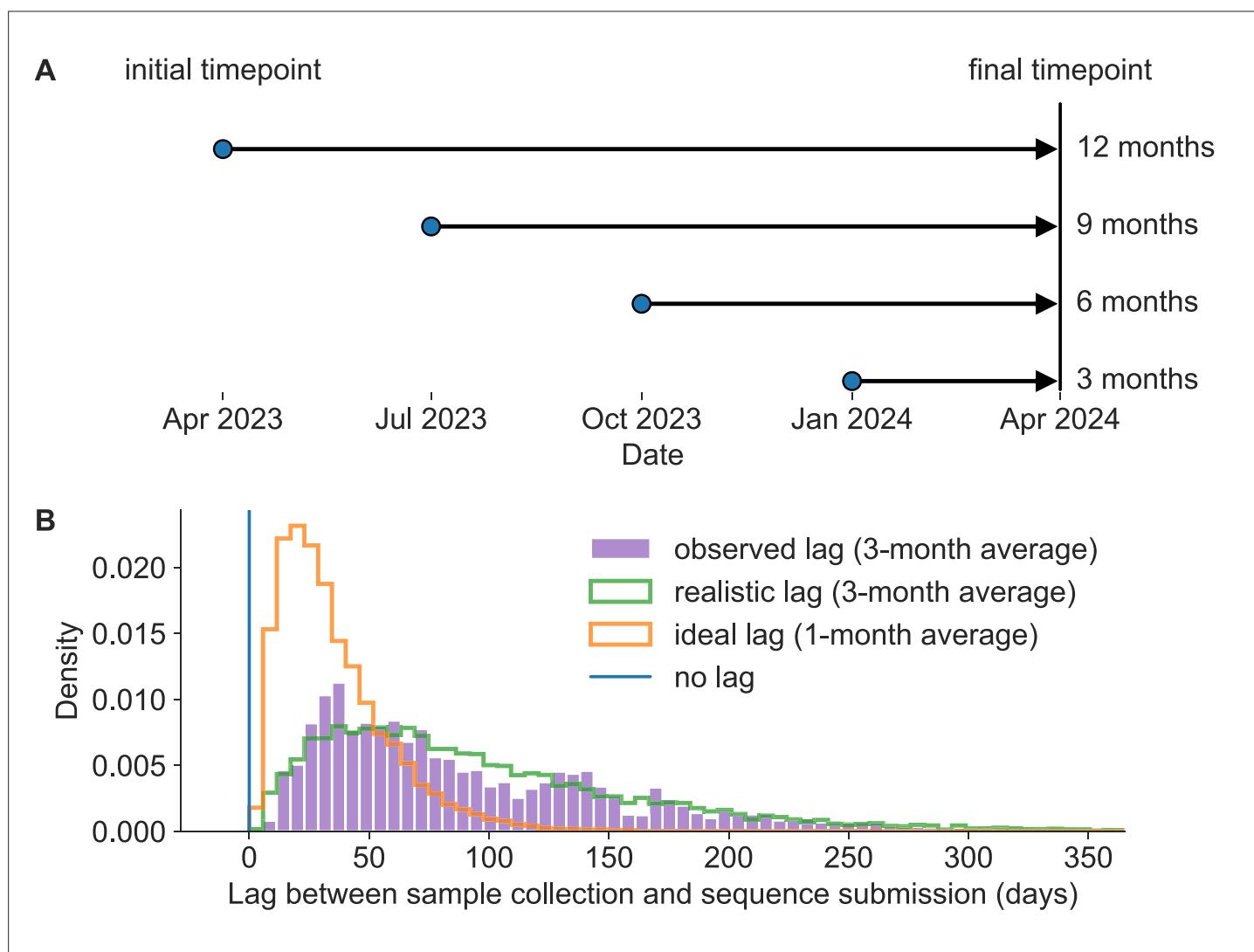


---

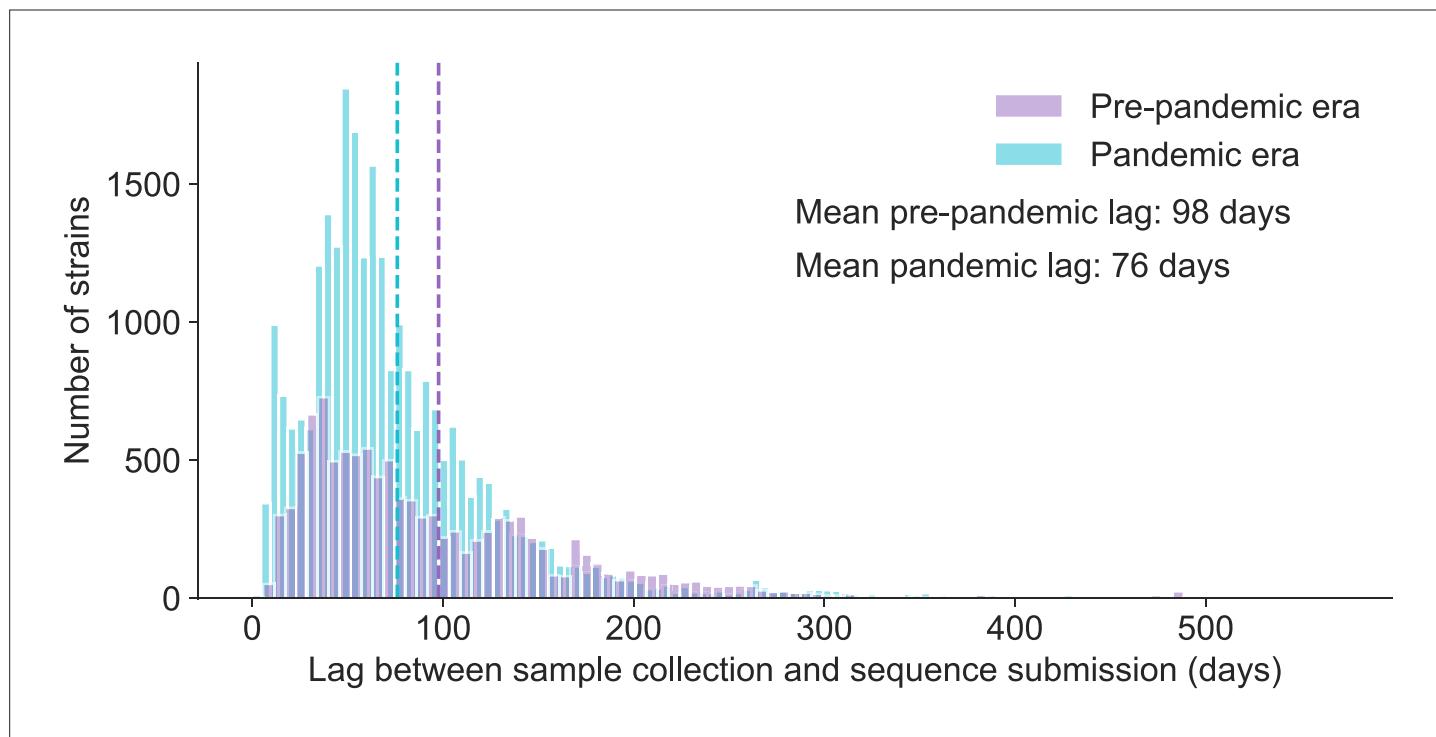
## Figures and figure supplements

Timely vaccine strain selection and genomic surveillance improve evolutionary forecast accuracy of seasonal influenza A/H3N2

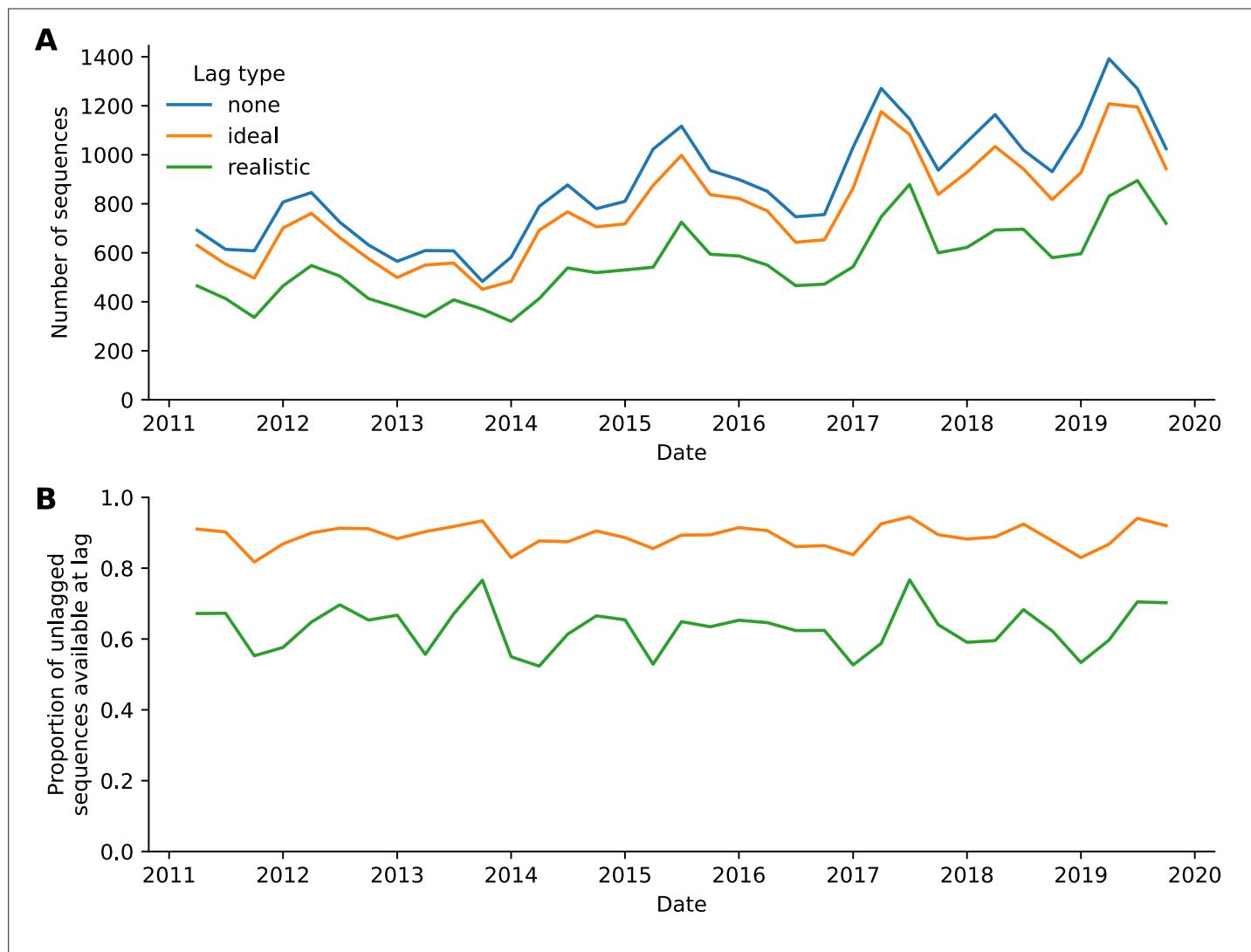
**John Huddleston and Trevor Bedford.**



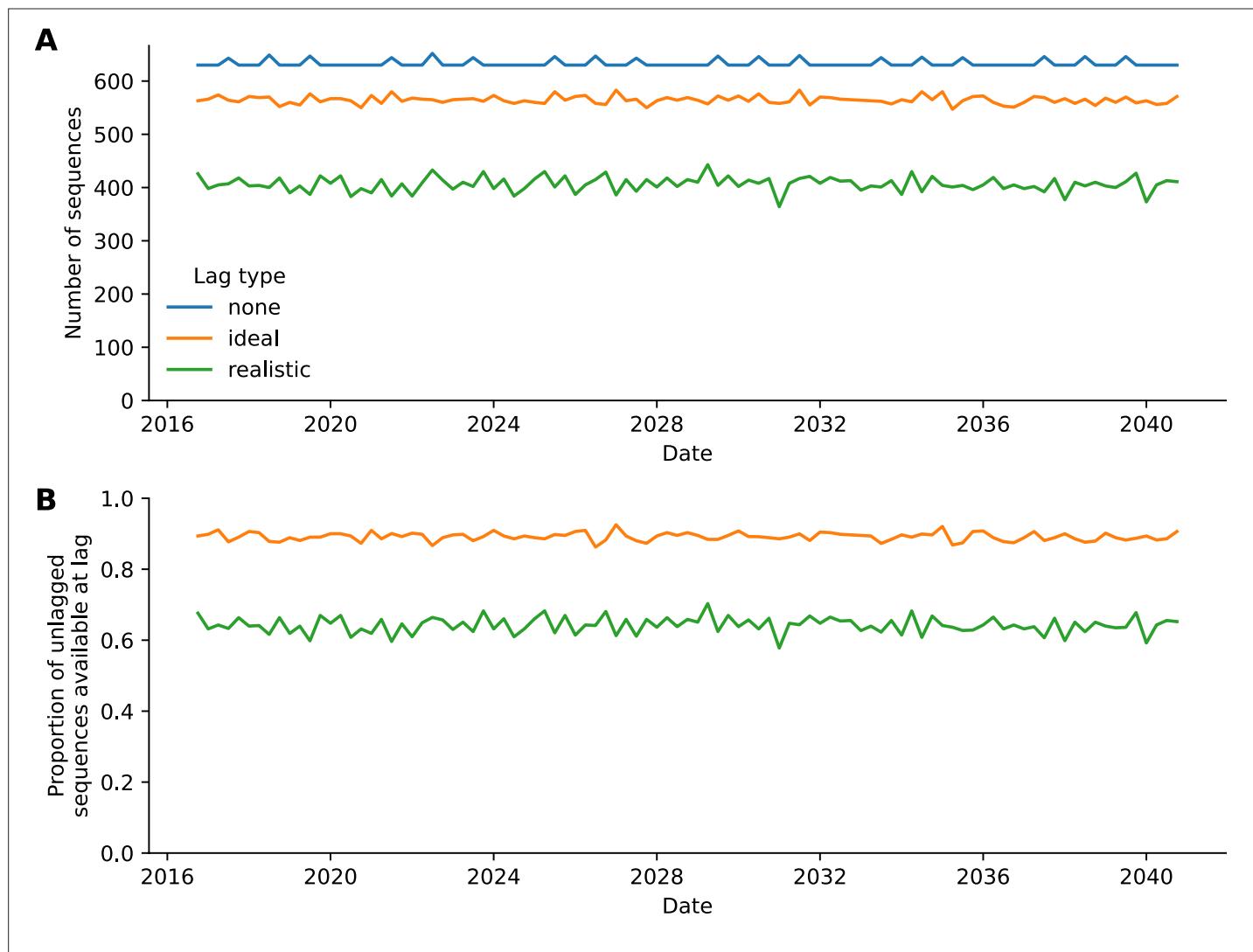
**Figure 1.** Model of forecast horizons and submission lags. **(A)** Long-term forecasting models historically predicted 12 months into the future from April and October because of the time required to develop and distribute a new vaccine (*Luksza and Lässig, 2014*). We tested three additional shorter forecast horizons in 3-month intervals of 9, 6, and 3 months prior to the same time in the future season. For each forecast horizon, we calculated the accuracy of forecasts under each of the three submission lags described below including no lag, realistic lag, and ideal lag. **(B)** Observed lags in days between collection of viral samples and submission of corresponding hemagglutinin (HA) sequences to Global Initiative on Sharing All Influenza Data (GISaid) (purple) for samples collected in 2019 have a mean of 98 days (approximately 3 months). A gamma distribution fit to the observed lag distribution with a similar mean and shape (green) represents a realistic submission lag that we sampled from to assign “submission dates” to simulated and natural A/H3N2 populations. A gamma distribution with a mean that is one-third of the realistic distribution (orange) represents an ideal submission lag analogous to the 1-month average observed lags for SARS-CoV-2 genomes. Retrospective analyses including fitting of forecasting models typically filter HA sequences by collection date instead of submission dates in which case there is no lag (blue).



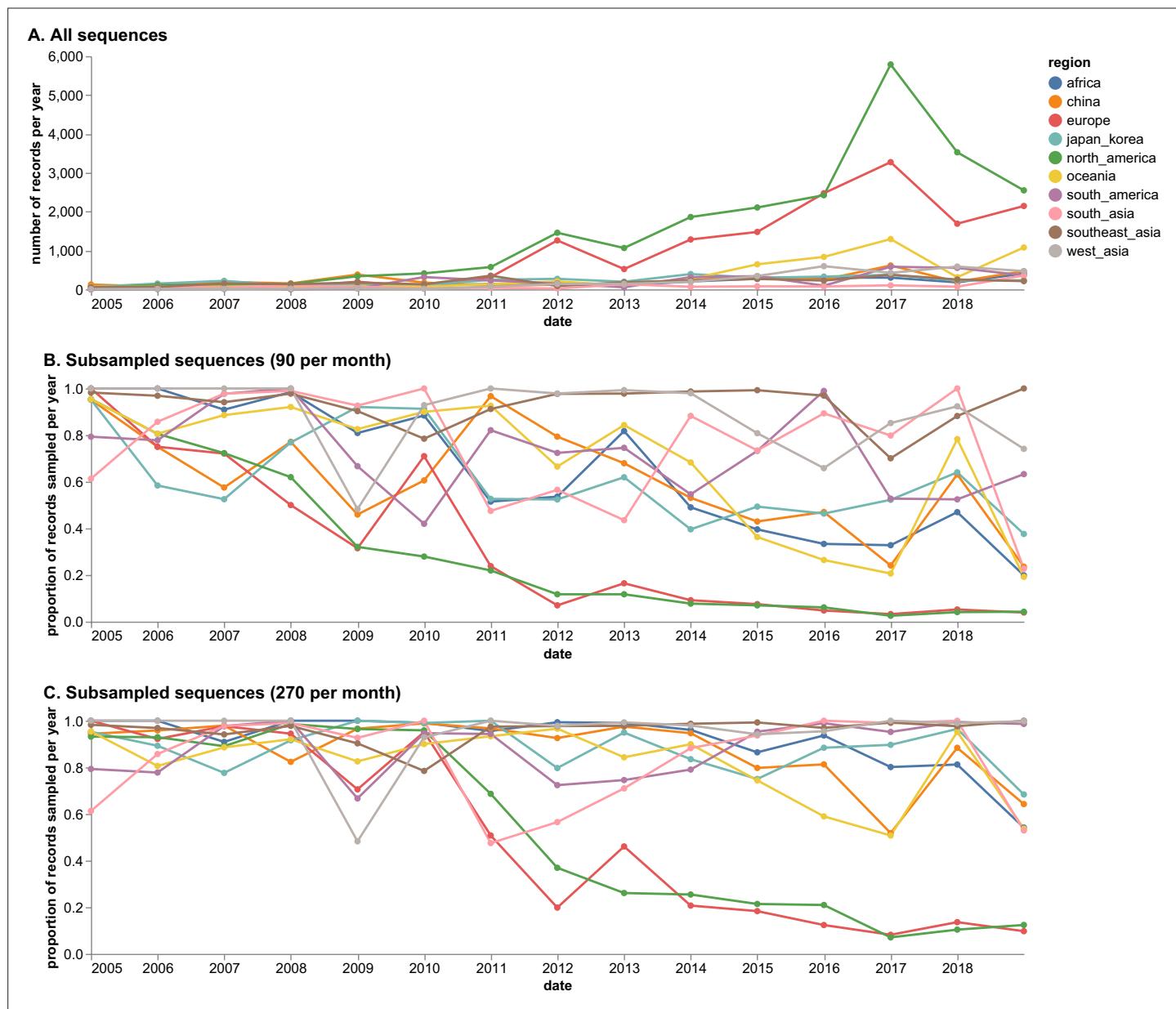
**Figure 1—figure supplement 1.** Distribution of submission lags in days for the pre-pandemic era (2019–2020) and pandemic era (2022–2023 in orange). Vertical dashed lines represent mean lags for each distribution.



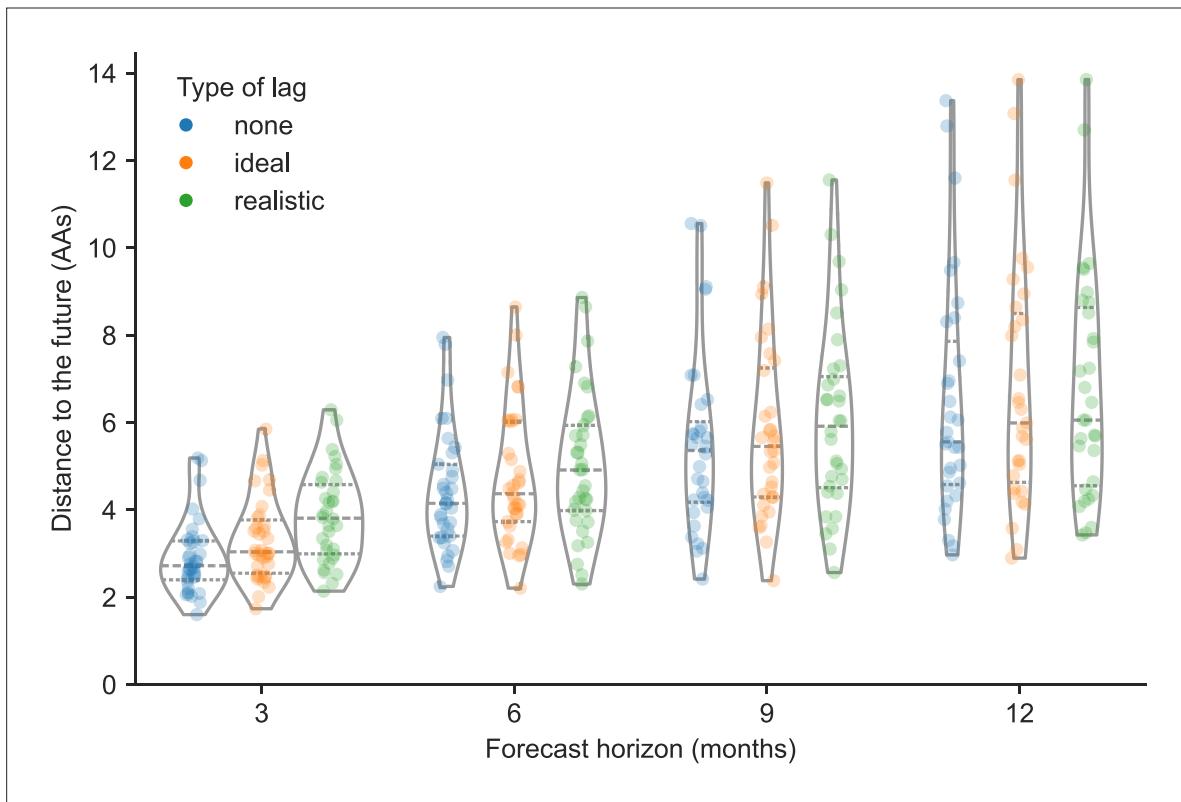
**Figure 1—figure supplement 2.** Number and proportion of A/H3N2 sequences available per timepoint and lag type. **(A)** Number of A/H3N2 sequences available per timepoint and lag type. **(B)** Proportion of all A/H3N2 sequences without lag per timepoint and lag type.



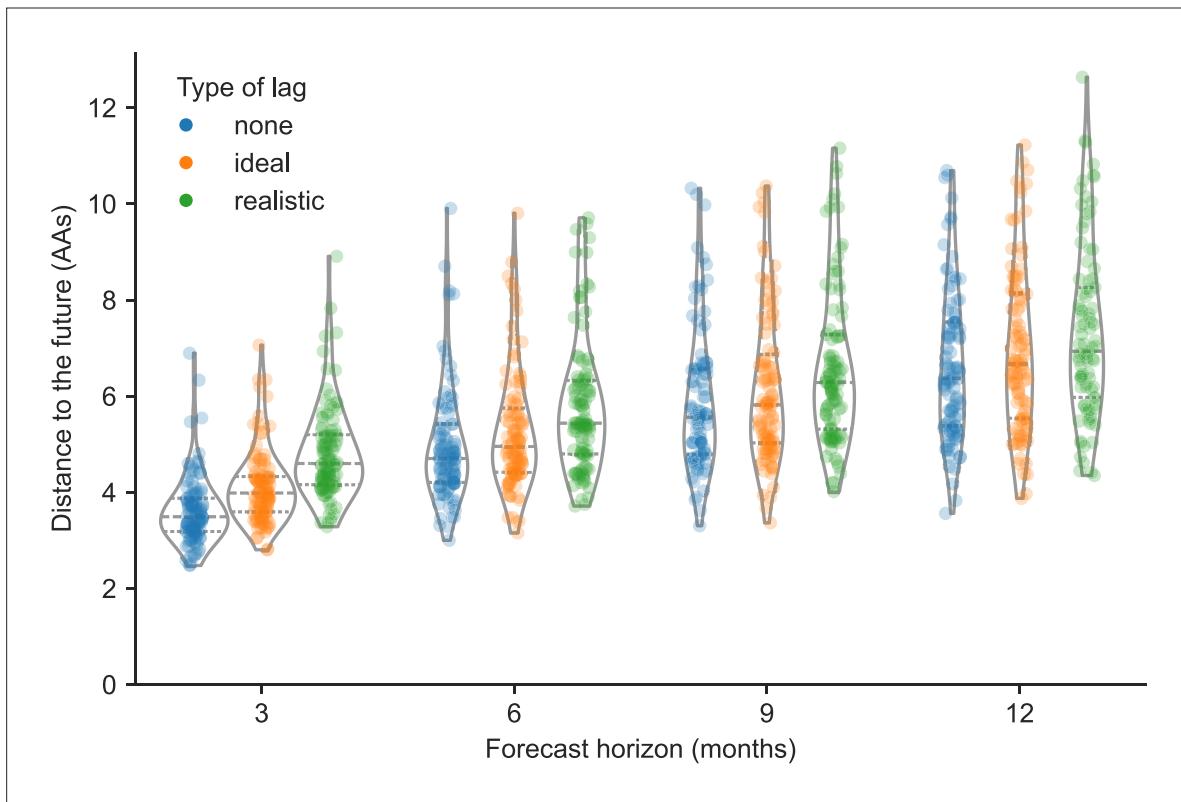
**Figure 1—figure supplement 3.** Number and proportion of simulated A/H3N2-like sequences available per timepoint and lag type. **(A)** Number of simulated A/H3N2-like sequences available per timepoint and lag type. **(B)** Proportion of all simulated A/H3N2-like sequences without lag per time point and lag type.



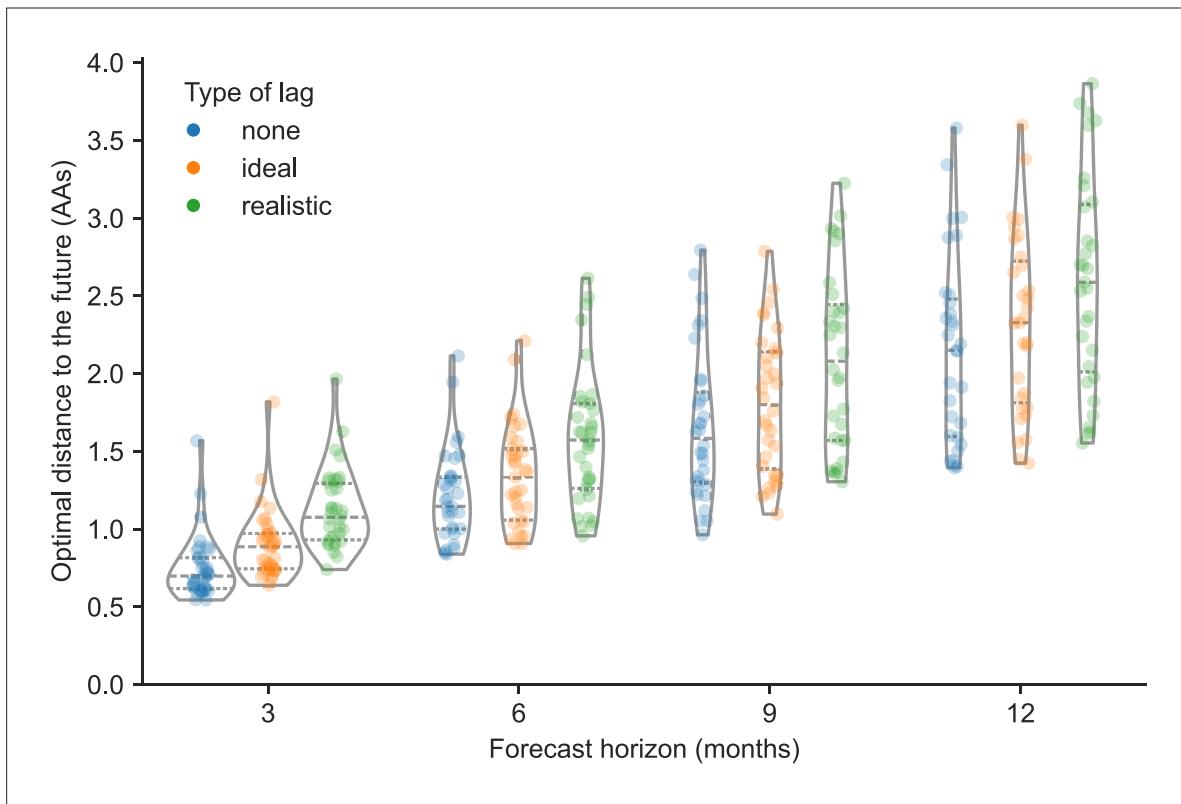
**Figure 1—figure supplement 4.** Number of all available sequences per region and year and proportion of sequences sampled by two different subsampling methods. **(A)** Number of all available sequences per region and year in the Global Initiative on Sharing All Influenza Data (GISaid) EpiFlu database for the study period between April 1, 2005, and October 1, 2019. **(B)** Proportion of sequences sampled per region and year with even subsampling across regions and year/month combinations at 90 viruses per month. **(C)** Proportion of sequences sampled per region and year with even subsampling across regions and year/month combinations at 270 viruses per month.



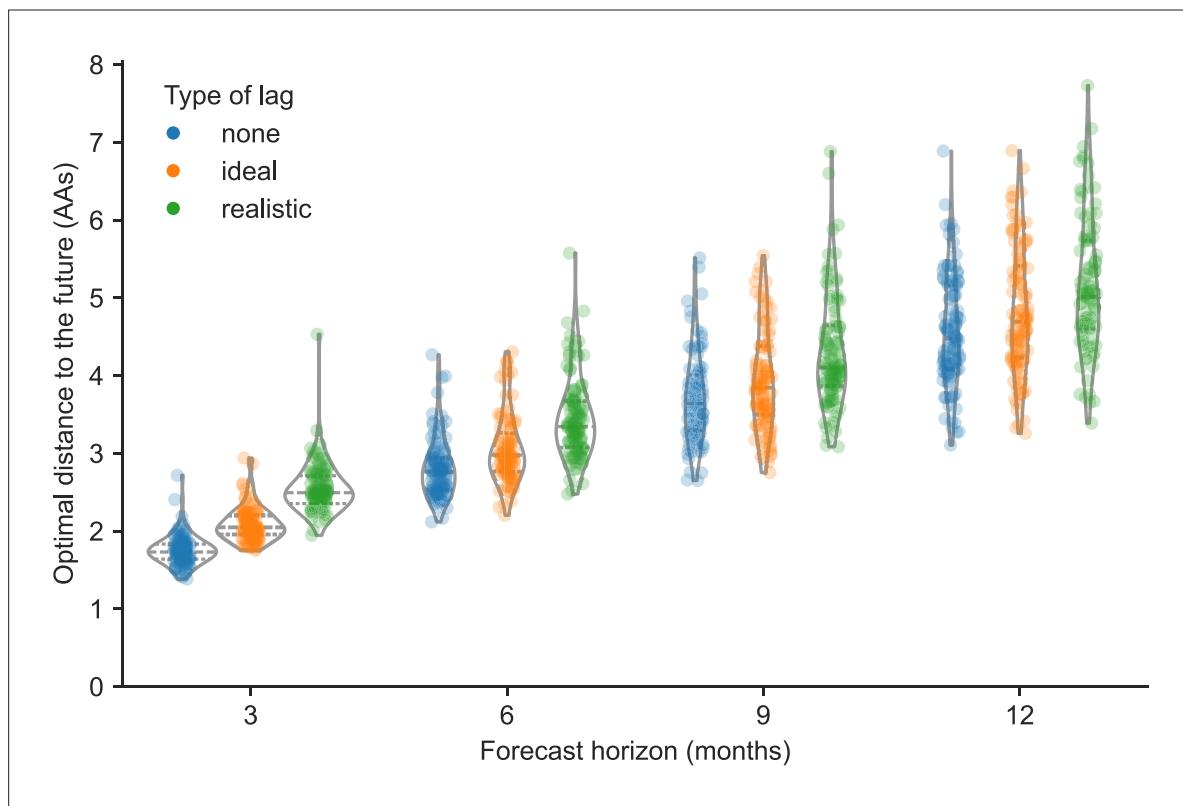
**Figure 2.** Distance to the future per timepoint (AAs) for natural A/H3N2 populations by forecast horizon and submission lag type based on forecasts from the local branching index (LBI) and mutational load model. Each point represents a future timepoint whose population was predicted from the number of months earlier corresponding to the forecast horizon. Points are colored by submission lag type including forecasts made with no lag (blue), an ideal lag (orange), and a realistic lag (green).



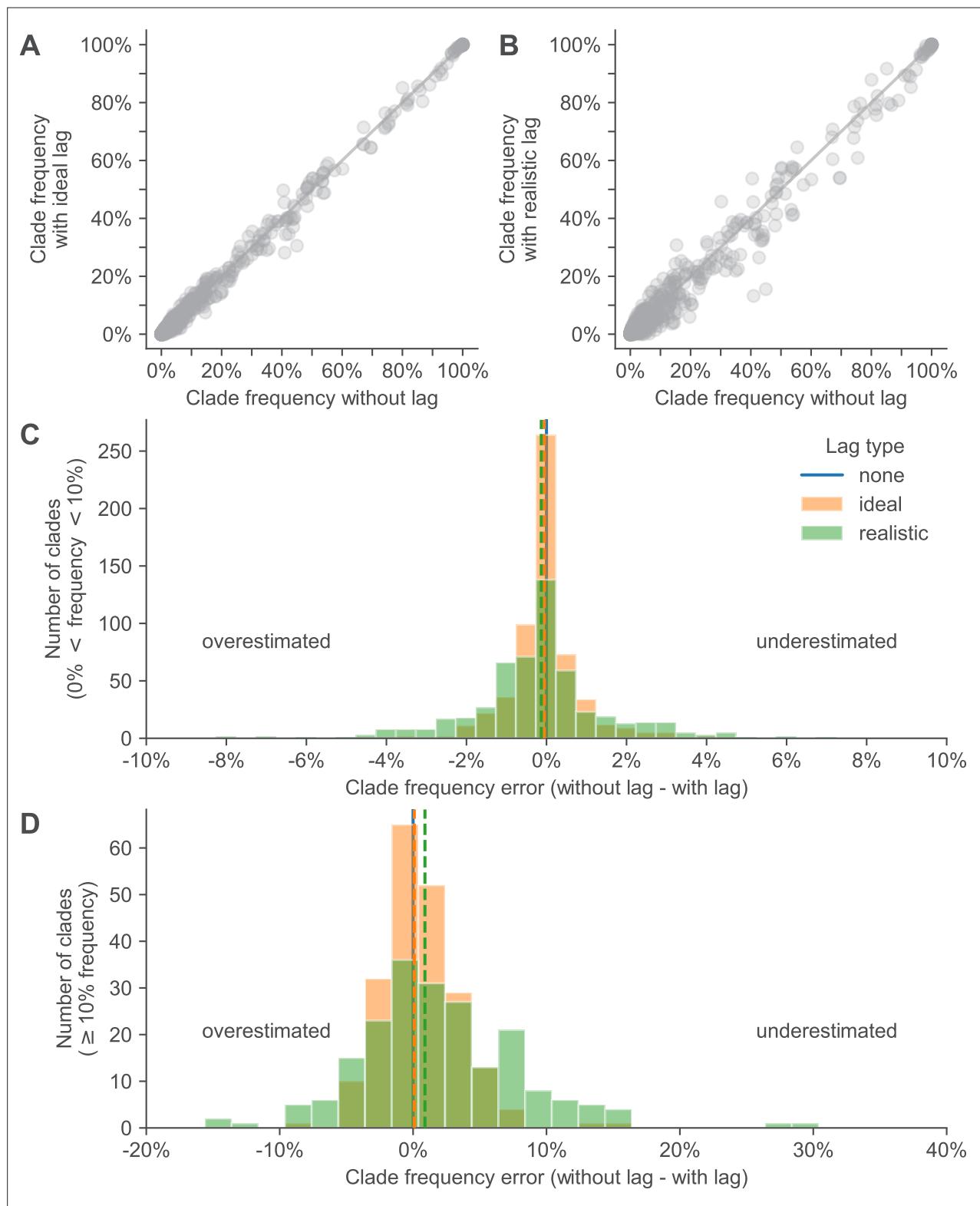
**Figure 2—figure supplement 1.** Distance to the future for simulated A/H3N2-like populations by forecast horizon and submission lag type based on forecasts from the “true fitness” model.



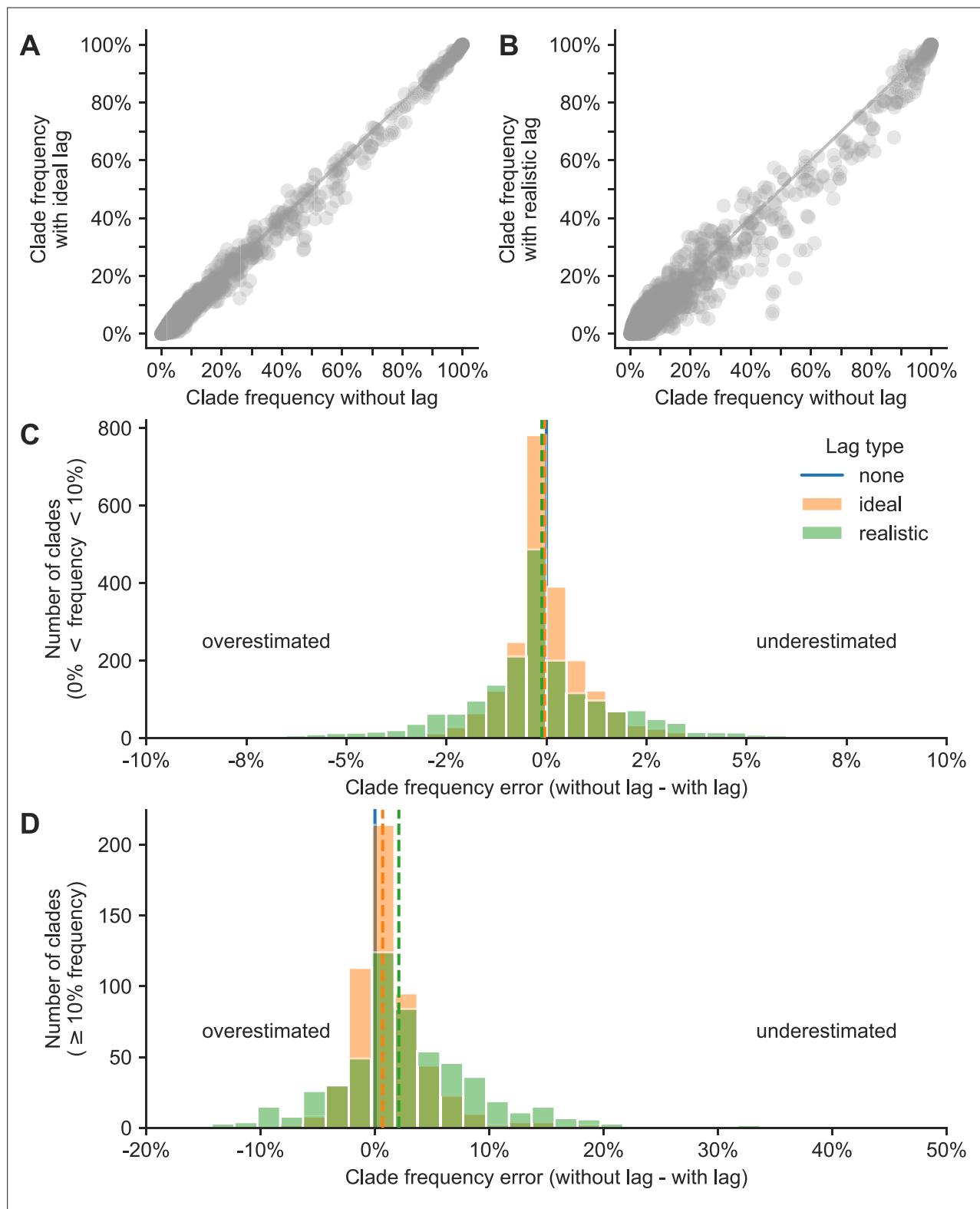
**Figure 2—figure supplement 2.** Optimal distance to the future for natural A/H3N2 populations by forecast horizon and submission lag type based on post hoc empirical fitness of the initial population.



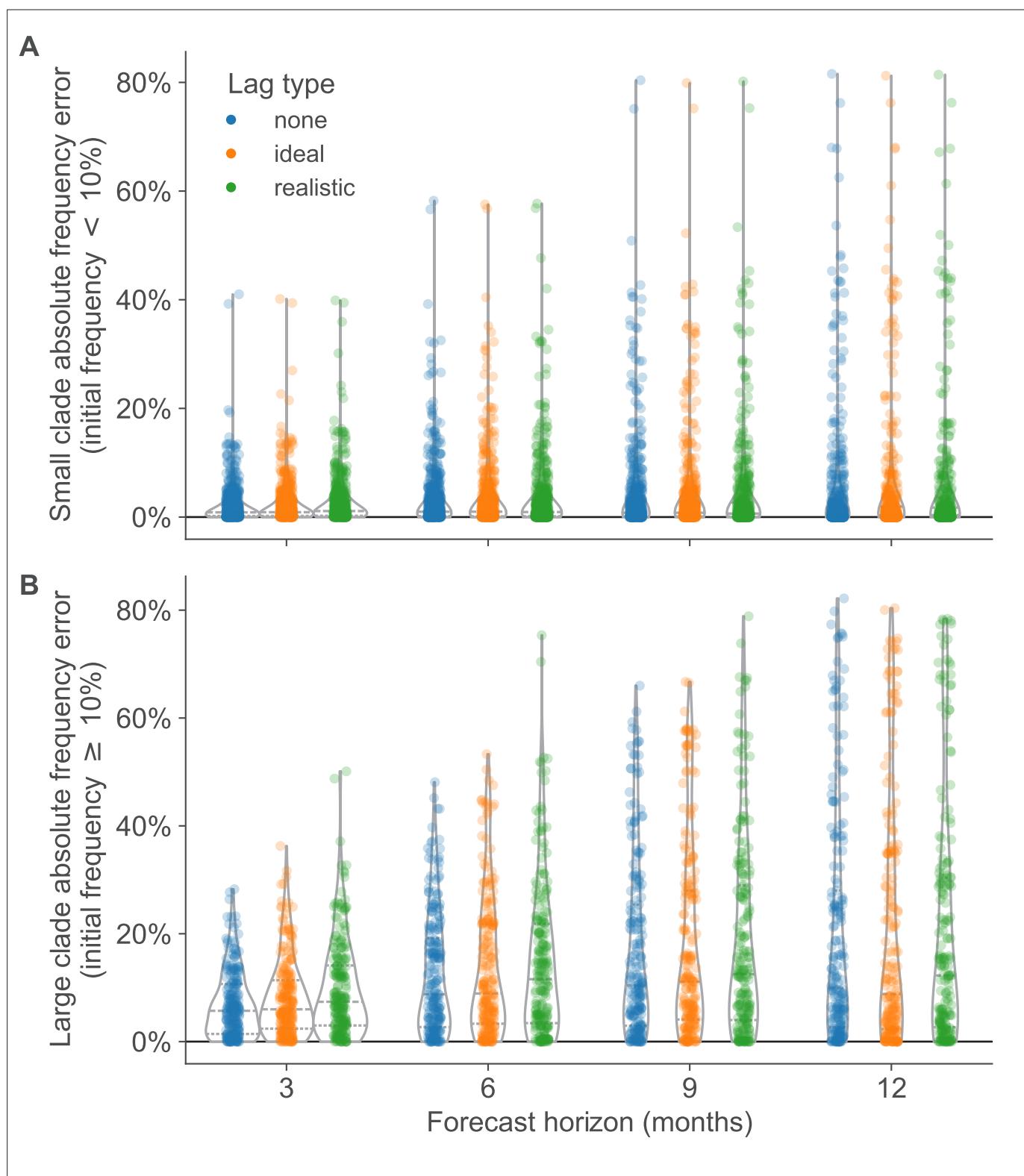
**Figure 2—figure supplement 3.** Optimal distance to the future for simulated A/H3N2-like populations by forecast horizon and submission lag type based on post hoc empirical fitness of the initial population.



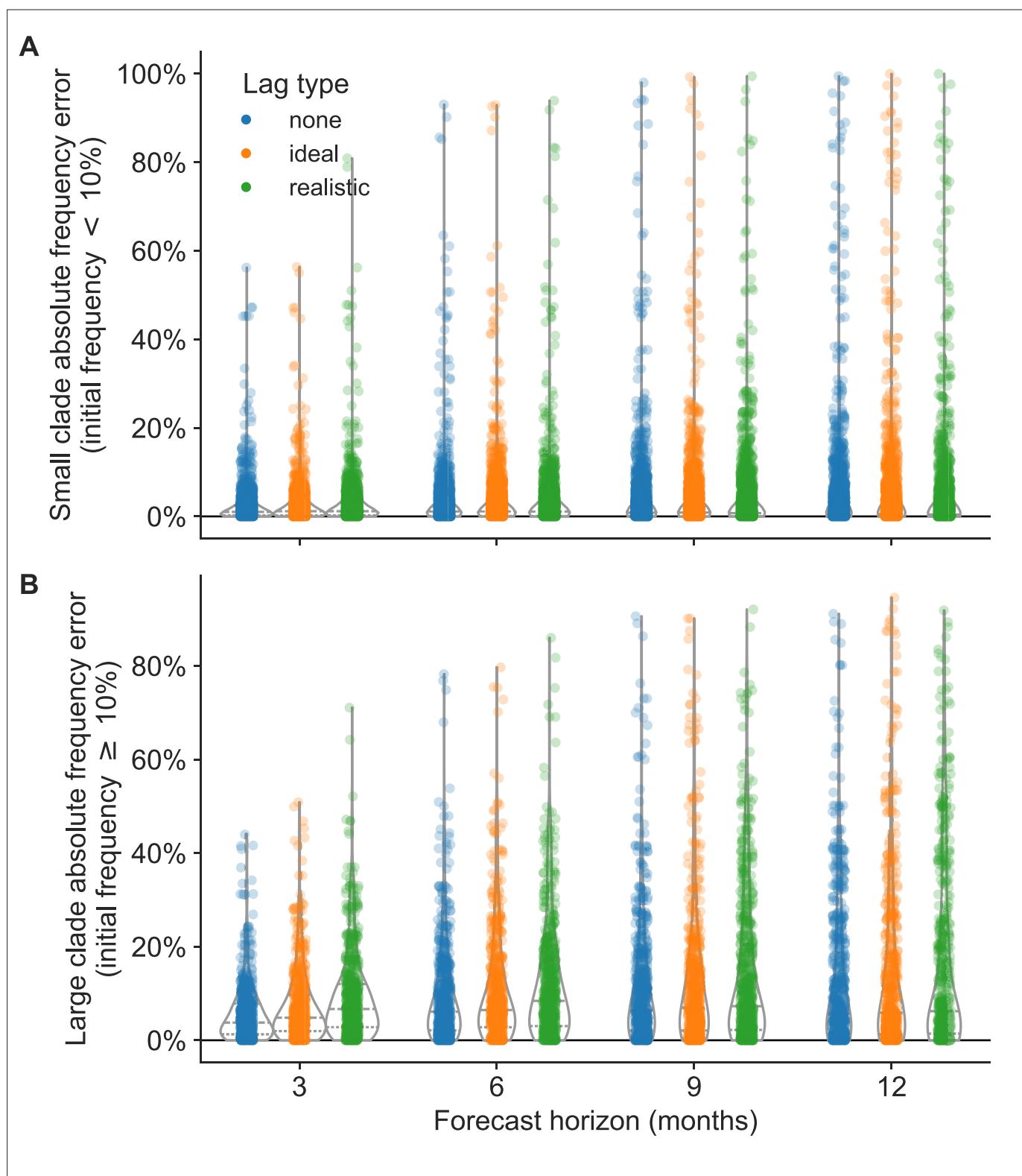
**Figure 3.** Clade frequency errors for natural A/H3N2 clades. Clade frequency errors for natural A/H3N2 clades at the same timepoint calculated as the difference between clade frequencies without submission lag and corresponding frequencies with either (A) ideal or (B) realistic submission lags. Distributions of frequency errors appear normally distributed in both lag scenarios for both (C) small clades ( $>0\%$  and  $<10\%$  frequency) and (D) large clades ( $\geq 10\%$ ). Dashed lines indicate the median error from the distribution of the lag type with the same color.



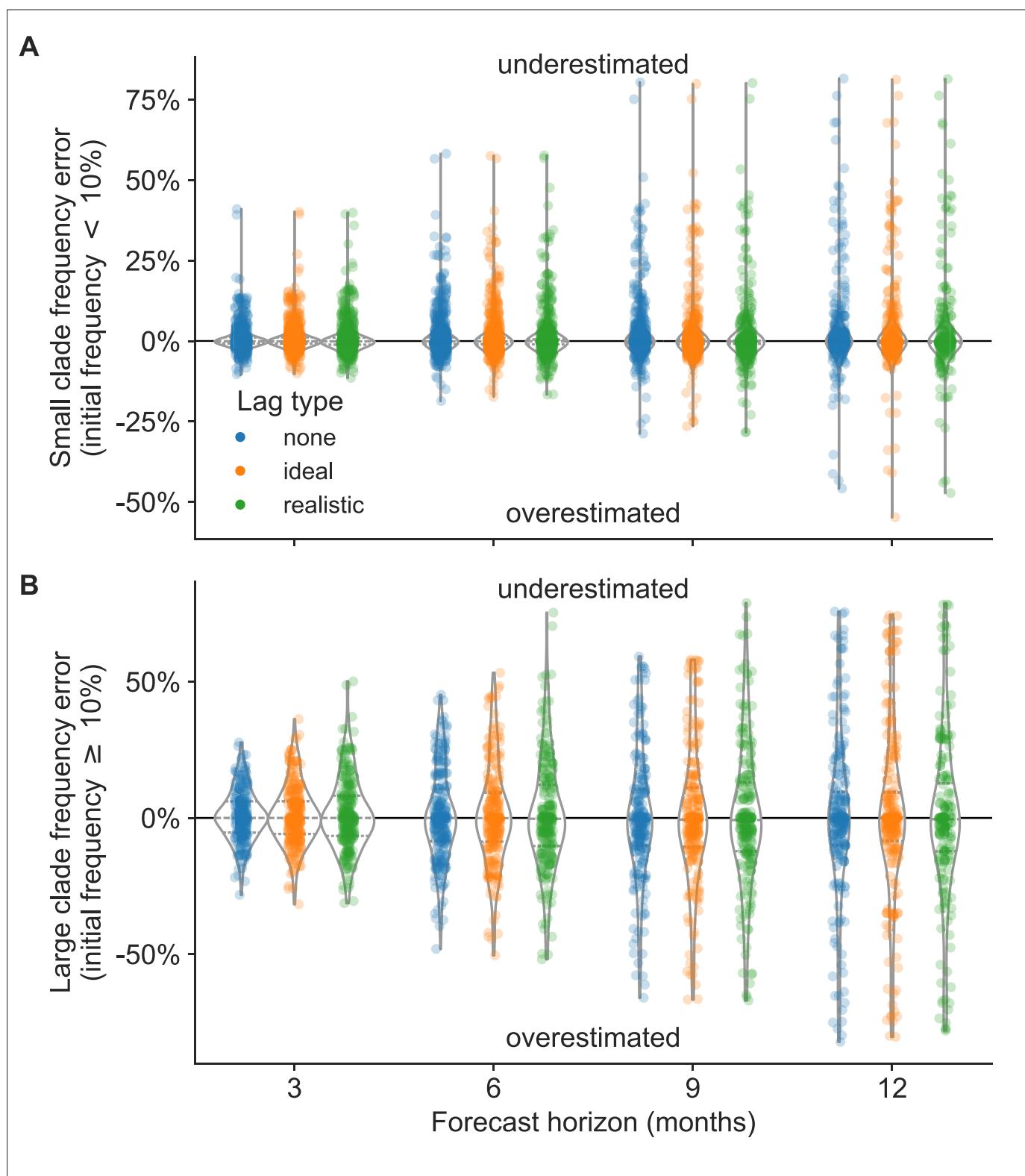
**Figure 3—figure supplement 1.** Clade frequency errors between simulated A/H3N2-like HA populations with ideal or realistic submission lags and populations without any submission lag.



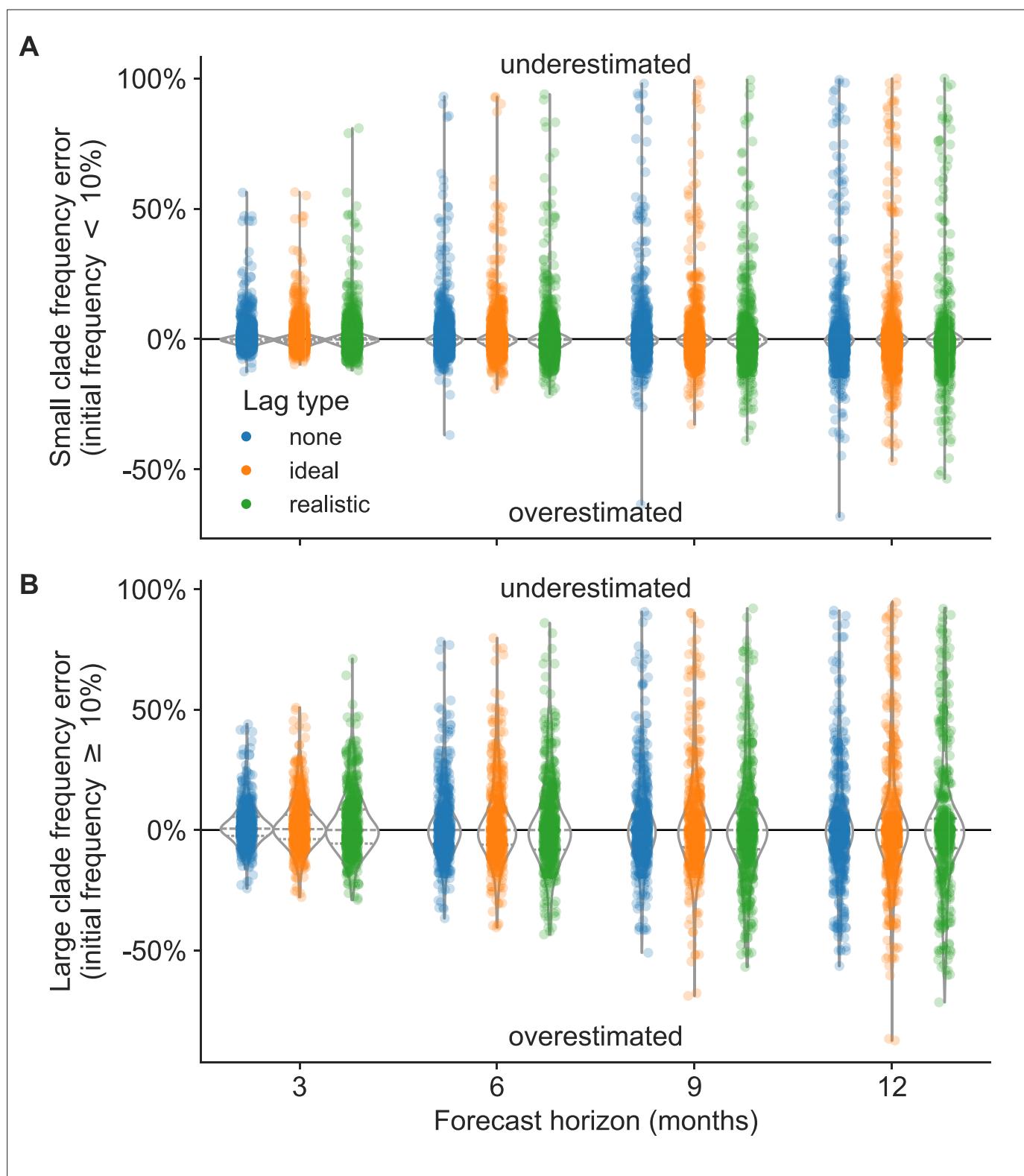
**Figure 4.** Absolute forecast clade frequency errors for natural A/H3N2 populations by forecast horizon in months and submission lag type (none, ideal, or observed) for (A) small clades ( $<10\%$  initial frequency) and (B) large clades ( $\geq 10\%$  initial frequency).



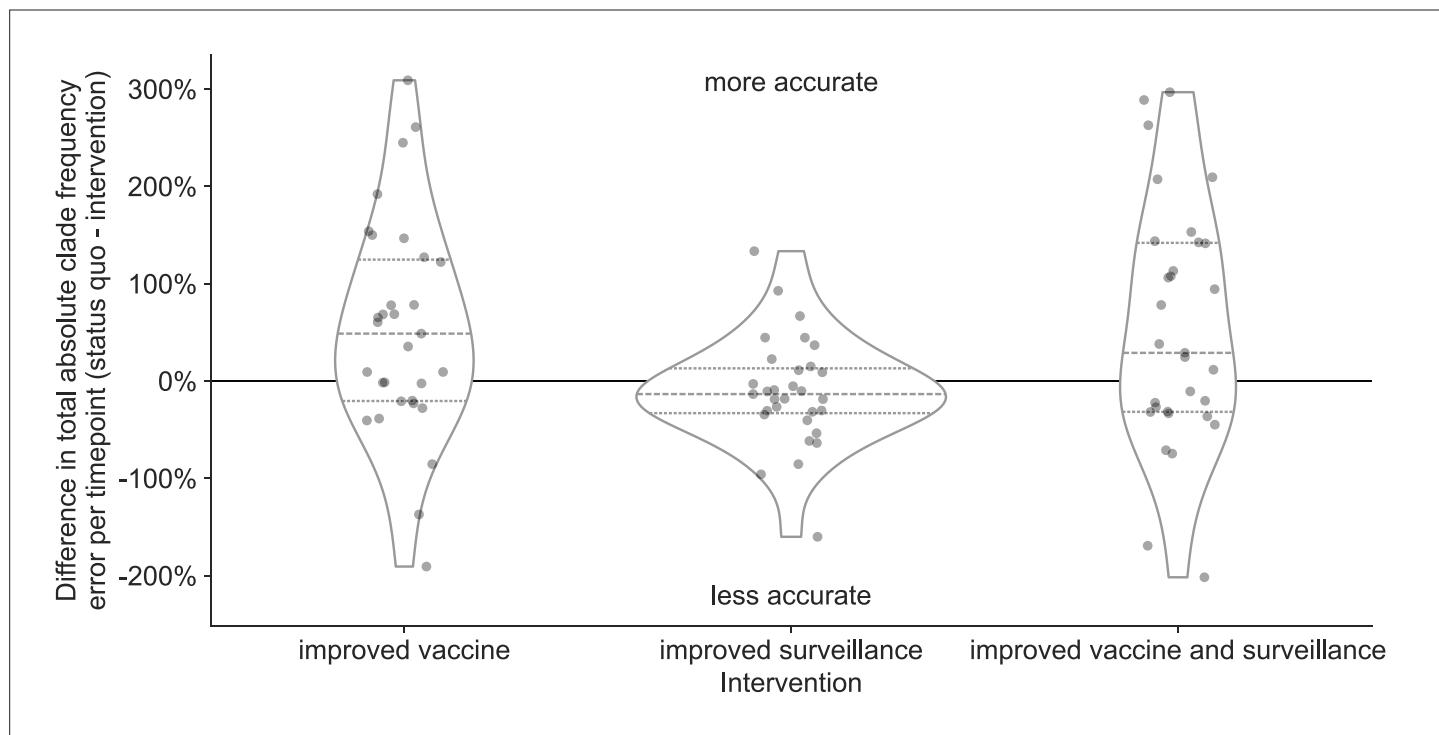
**Figure 4—figure supplement 1.** Absolute forecast clade frequency errors for simulated A/H3N2-like HA populations by forecast horizon in months and submission lag type (none, ideal, or realistic) for (A) small clades (<10% initial frequency) and (B) large clades (≥10% initial frequency).



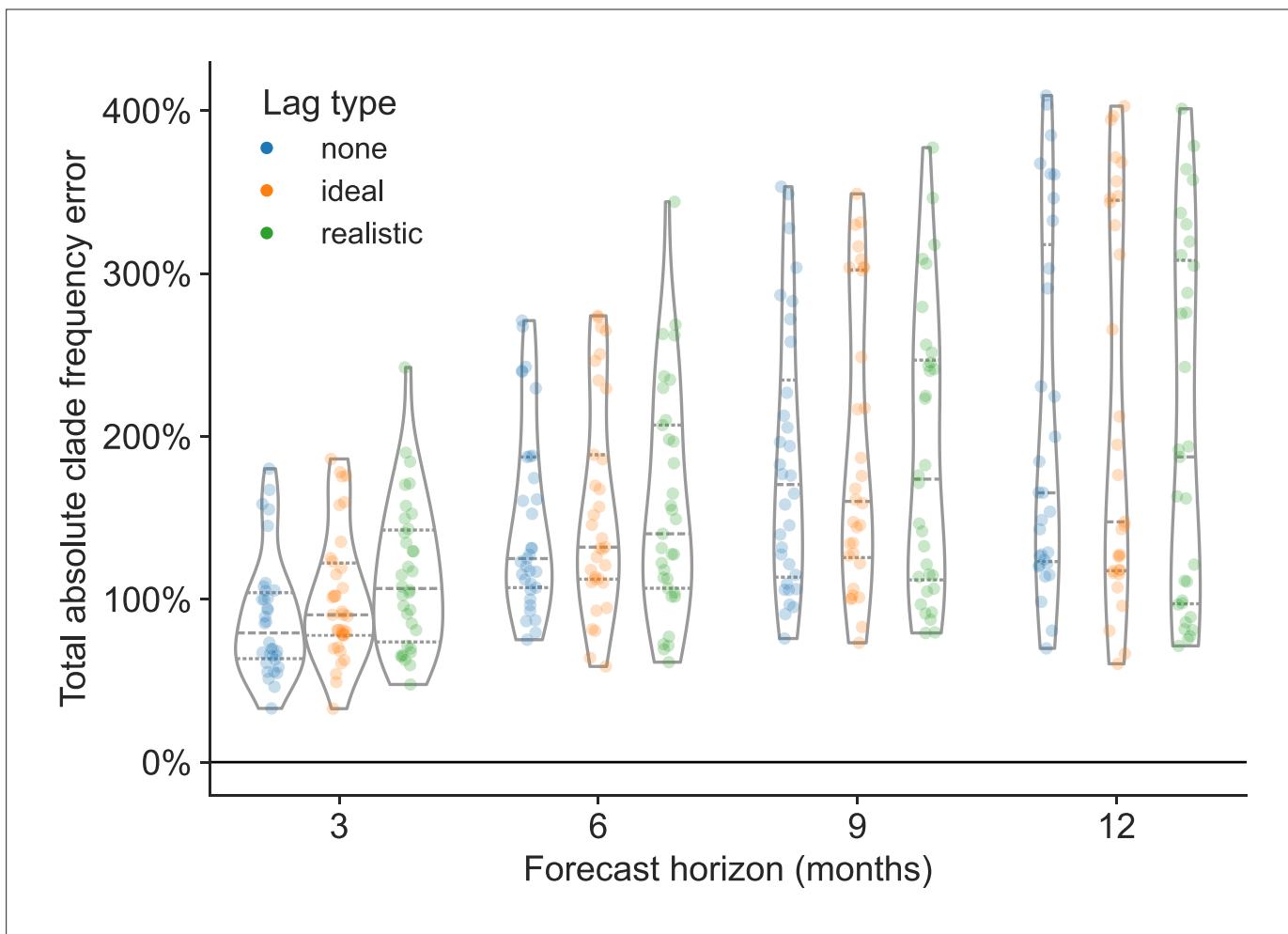
**Figure 4—figure supplement 2.** Forecast clade frequency errors for natural A/H3N2 HA populations by forecast horizon in months and submission lag type (none, ideal, or realistic) for (A) small clades (10% initial frequency) and (B) large clades ( $\geq 10\%$  initial frequency).



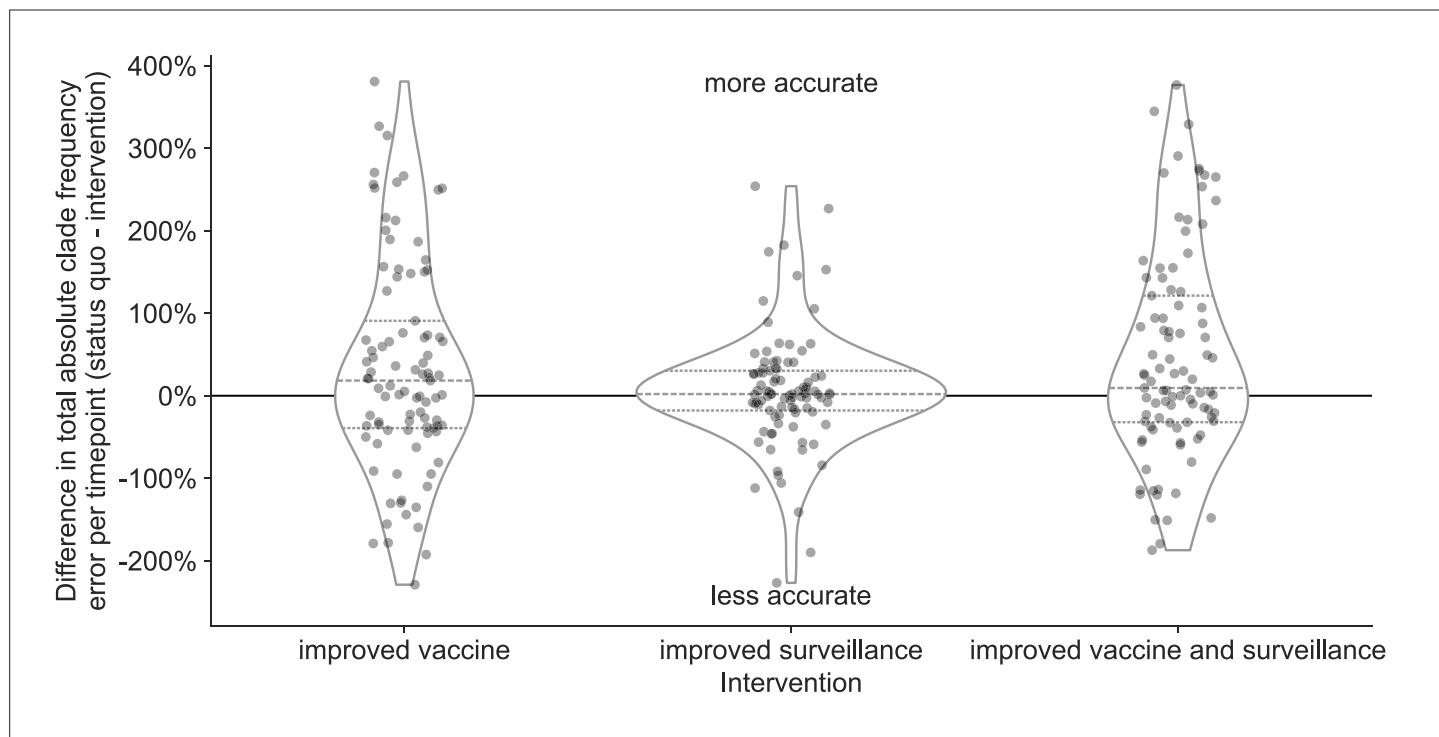
**Figure 4—figure supplement 3.** Forecast clade frequency errors for simulated A/H3N2-like HA populations by forecast horizon in months and submission lag type (none, ideal, or realistic) for (A) small clades (<10% initial frequency) and (B) large clades ( $\geq 10\%$  initial frequency).



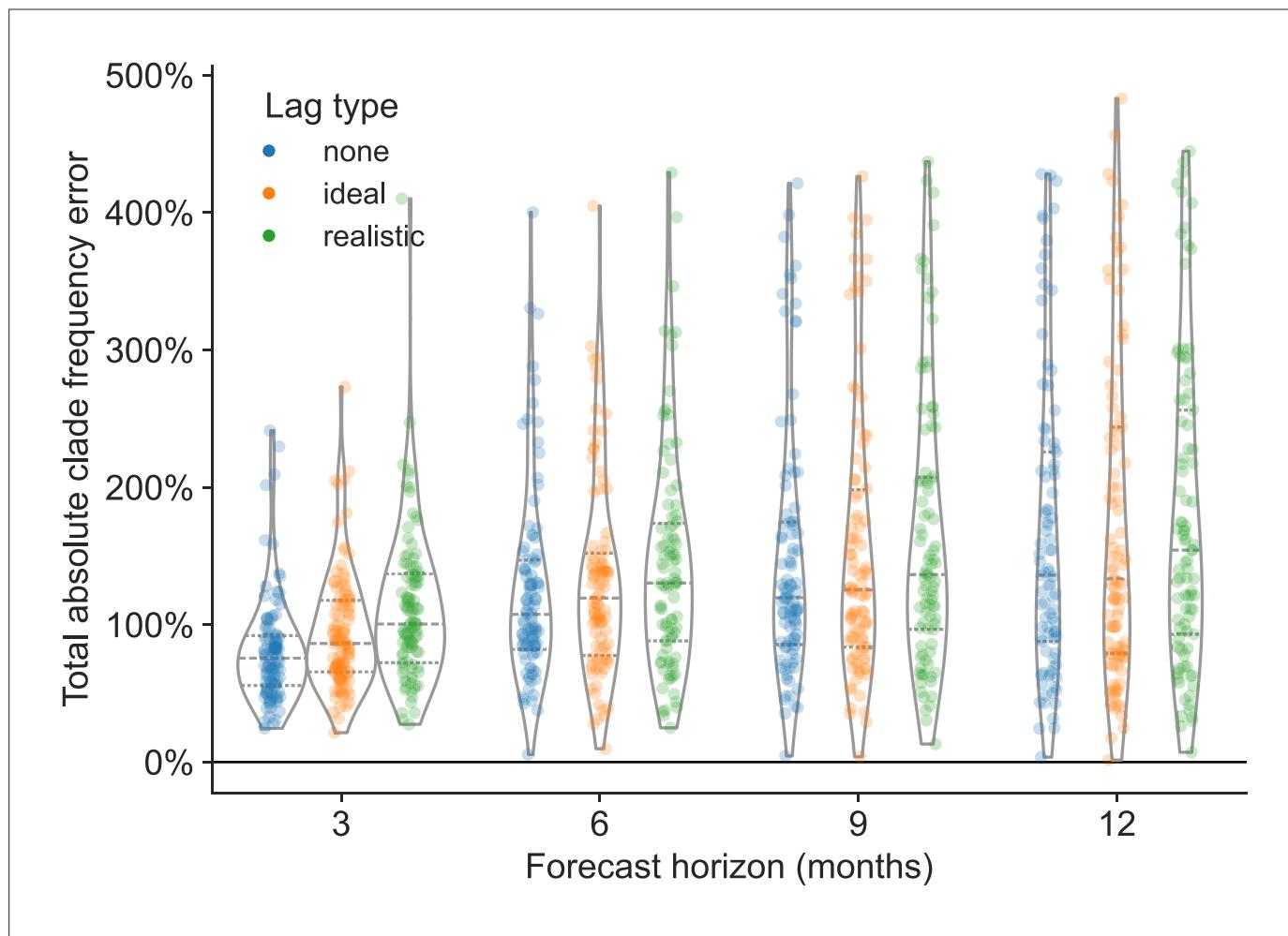
**Figure 5.** Improvement of clade frequency errors for A/H3N2 populations between the status quo (12-month forecast horizon and realistic submission lags) and realistic interventions of improved vaccine development (reducing 12-month to 6-month forecast horizon), improved surveillance (reducing submission lags from 3 months on average to 1 month), or a combination of both interventions. We measured improvements from the status quo as the difference in total absolute clade frequency error per future timepoint. Positive values indicate increased forecast accuracy, while negative values indicate decreased accuracy. Each point represents the improvement of forecasts for a specific future timepoint under the given intervention. Horizontal dashed lines indicate median improvements. Horizontal dotted lines indicate upper and lower quartiles of improvements.



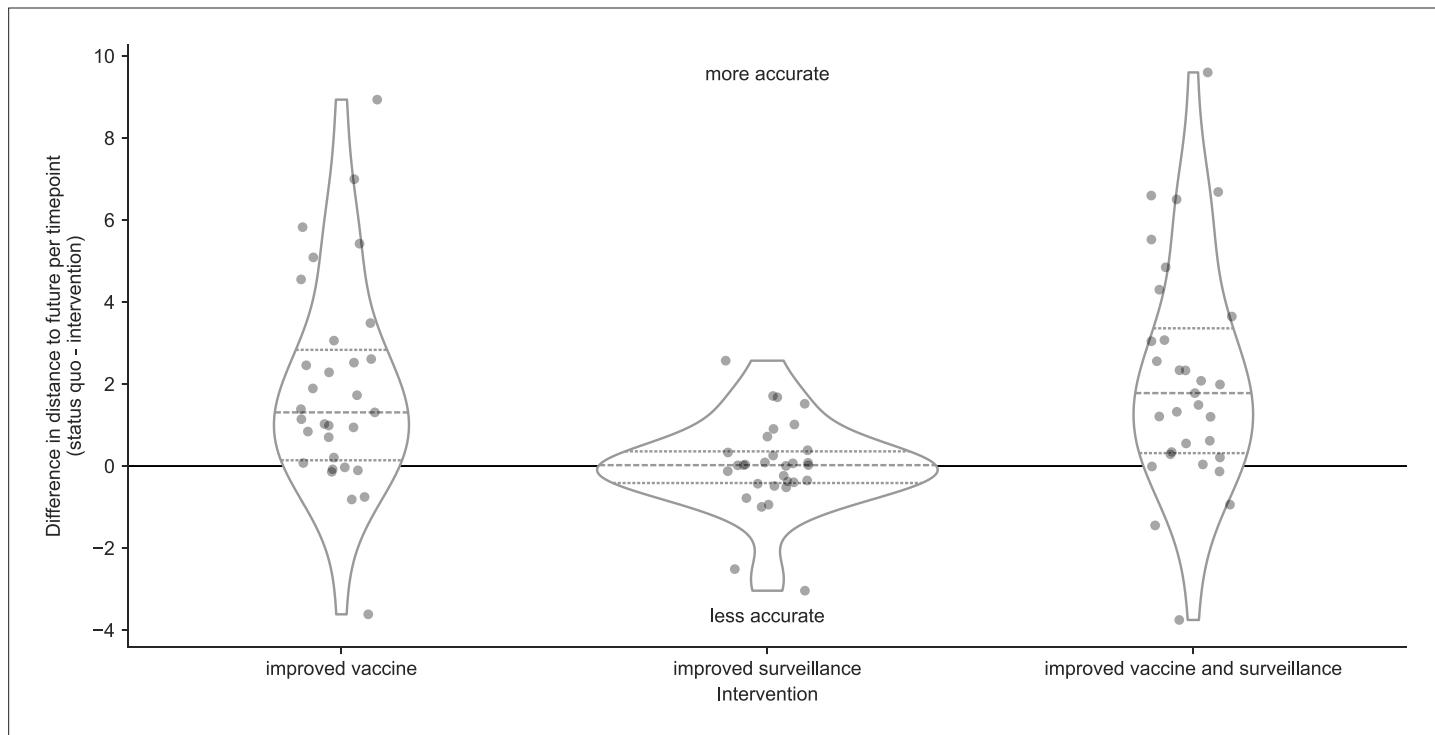
**Figure 5—figure supplement 1.** Distribution of total absolute clade frequency errors summed across clades per future timepoint for A/H3N2 populations. We calculated the effects of interventions as the difference between these values per future timepoint under the status quo (12-month forecast horizon and realistic submission lag) and specific interventions.



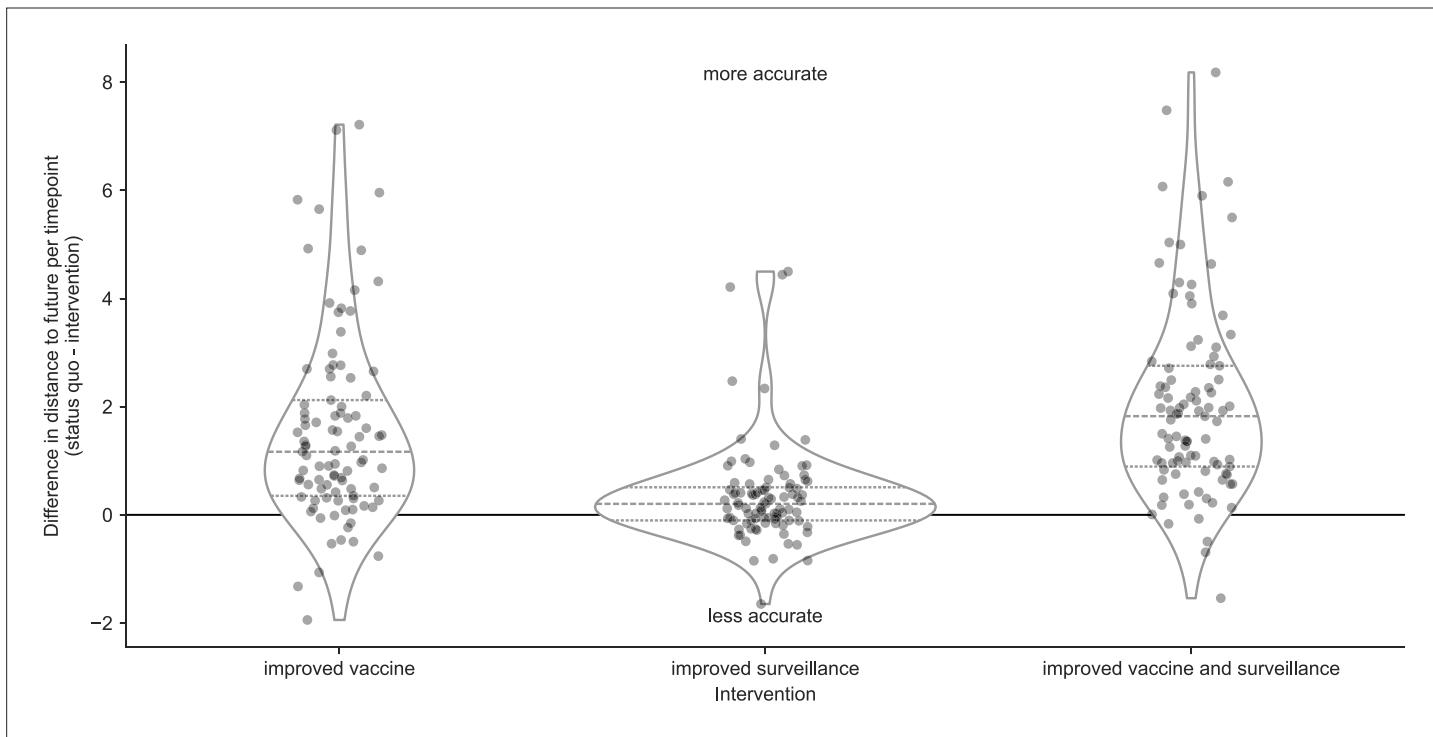
**Figure 5—figure supplement 2.** Improvement of clade frequency errors for simulated A/H3N2-like populations between the status quo and realistic interventions.



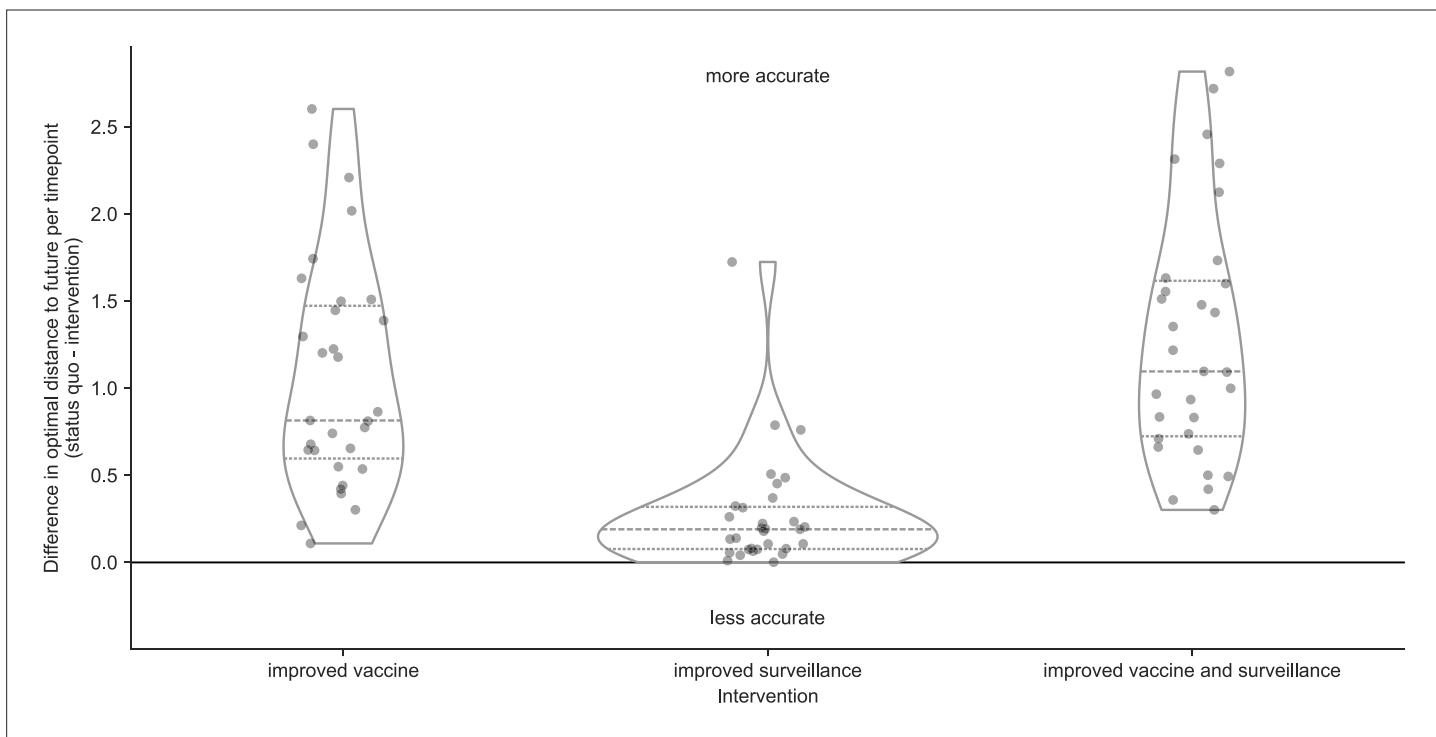
**Figure 5—figure supplement 3.** Distribution of total absolute clade frequency errors summed across clades per future timepoint for simulated A/H3N2-like populations.



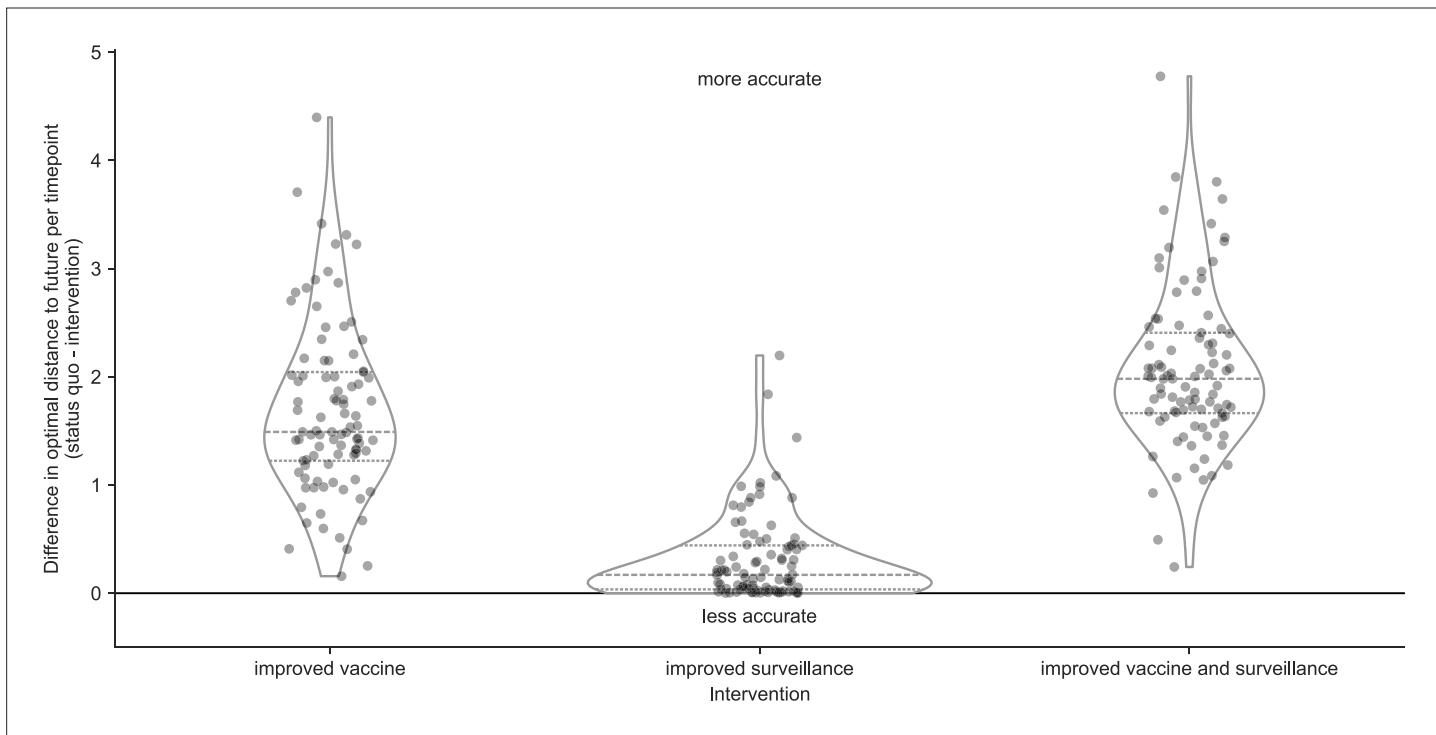
**Figure 5—figure supplement 4.** Improvement of distances to the future (AAs) for A/H3N2 populations between the status quo (12-month forecast horizon and realistic submission lags) and realistic interventions. The effects of interventions are the differences between distances to the future per future timepoint under the status quo and specific interventions.



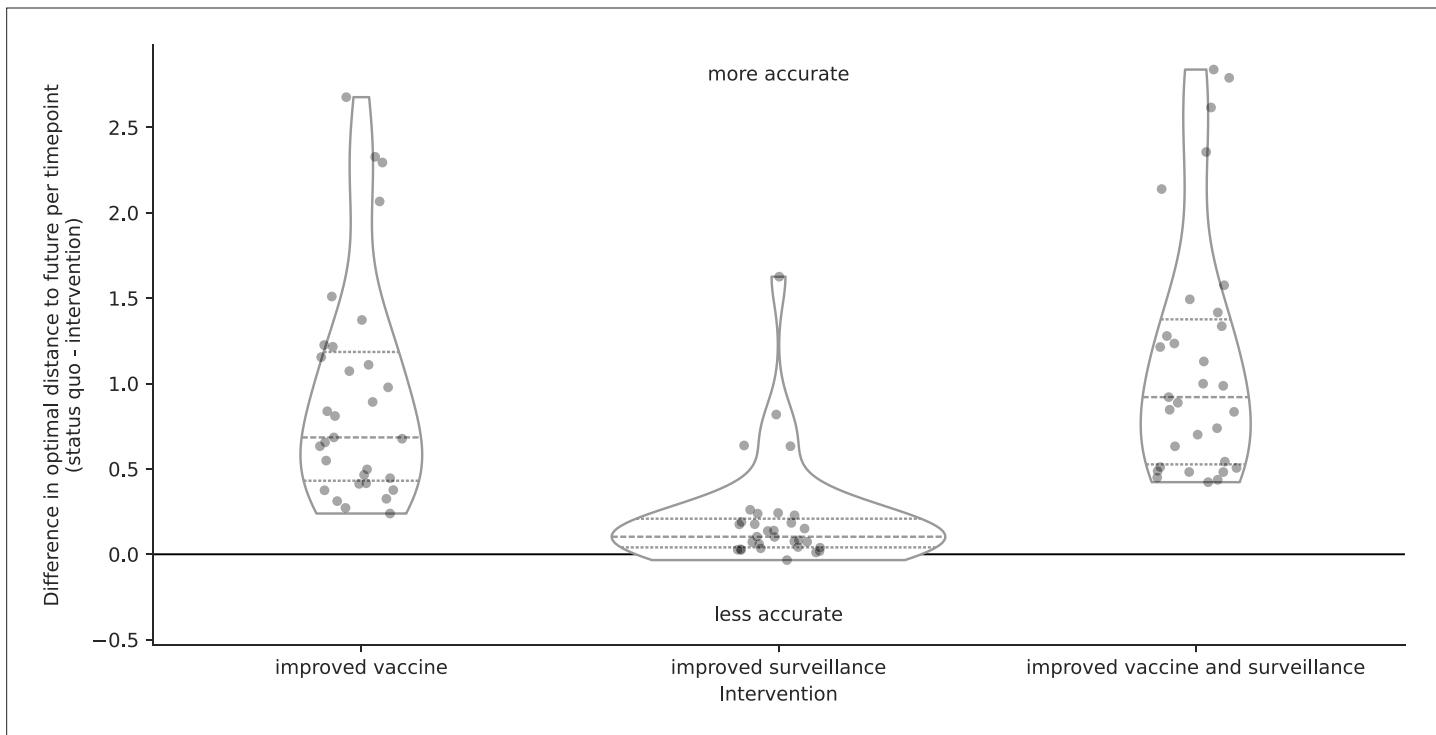
**Figure 5—figure supplement 5.** Improvement of distances to the future (AAs) for simulated A/H3N2-like populations between the status quo (12-month forecast horizon and realistic submission lags) and realistic interventions. The effects of interventions are the differences between distances to the future per future timepoint under the status quo and specific interventions.



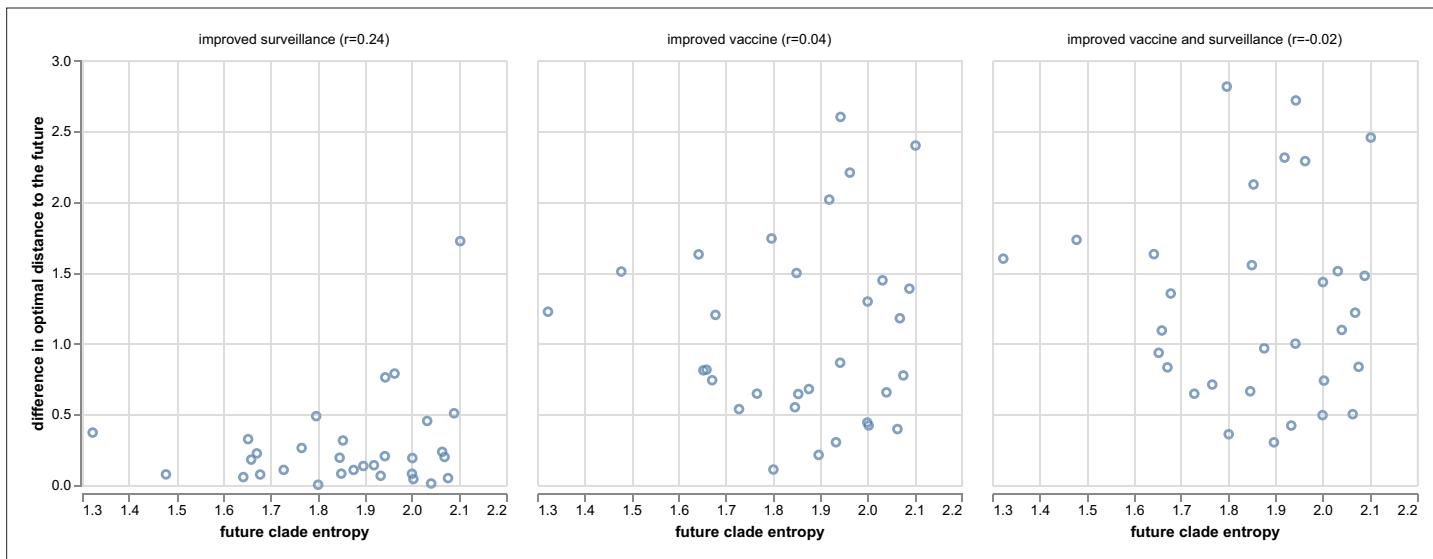
**Figure 6.** Improvement of optimal distances to the future (AAs) for A/H3N2 populations between the status quo (12-month forecast horizon and realistic submission lags) and realistic interventions of improved vaccine development (reducing 12-month to 6-month forecast horizon), improved surveillance (reducing submission lags from 3 months on average to 1 month), or a combination of both interventions. We measured improvements from the status quo as the difference in optimal distances to the future per future timepoint. Positive values indicate increased forecast accuracy, while negative values indicate decreased accuracy. Each point represents the improvement of forecasts for a specific future timepoint under the given intervention. Horizontal dashed lines indicate median improvements. Horizontal dotted lines indicate upper and lower quartiles of improvements.



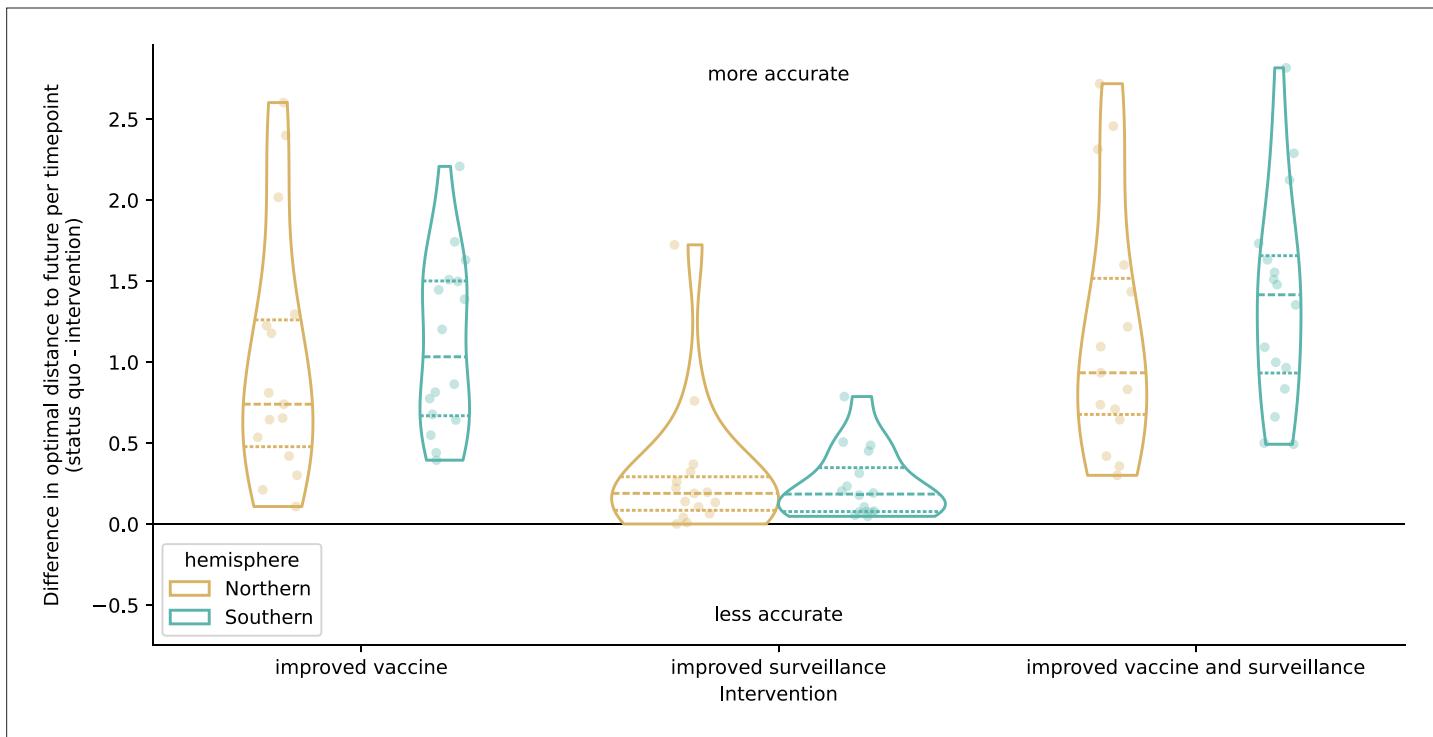
**Figure 6—figure supplement 1.** Improvement of optimal distances to the future (AAs) for simulated A/H3N2-like populations between the status quo (12-month forecast horizon and realistic submission lags) and realistic interventions.



**Figure 6—figure supplement 2.** Improvement of optimal distances to the future (AAs) for A/H3N2 populations between the status quo (12-month forecast horizon and realistic submission lags) and realistic interventions using forecasts based on sampling 270 viruses per month instead of the 90 viruses-per-month sampling used in the main results.



**Figure 6—figure supplement 3.** Improvement of optimal distances to the future (AAs) for A/H3N2 populations compared to the Shannon entropy of clade frequencies (estimated without submission lags) at the future timepoint being forecast to. Panel titles include the Pearson  $r$  value between the improvement in distance and the future clade entropy.



**Figure 6—figure supplement 4.** Improvement of optimal distances to the future (AAs) for A/H3N2 populations by intervention and the hemisphere with an active season during the future timepoint being predicted. We labeled future timepoints that occurred in October or January as "Northern" and "Southern" if the dates were in April or July.