



OPEN ACCESS

EDITED BY

Julie Olson,
University of Minnesota Twin Cities,
United States

REVIEWED BY

Aruna Pal,
West Bengal University of Animal and
Fishery Sciences, India
LaToya Richards Montgomery Richards,
Beckman Coulter Life Sciences,
United States

*CORRESPONDENCE

David R. Mcllwain
✉ davidmcllwain@unr.edu

[†]These authors have contributed equally
to this work

[‡]These authors share senior authorship

RECEIVED 13 January 2026
REVISED 13 February 2026
ACCEPTED 23 February 2026
PUBLISHED 19 March 2026

CITATION

Rizzo Nervo K, Hajiakhoond Bidoki N,
Chen H, Rahil Z, Bjornson Z, Kim K,
Bock B, Affrime M, Bauerle L,
Huynh PBT, Liebowitz D, Tucker S,
Gherardini PF, Nolan GP, Aghaeepour N
and Mcllwain DR (2026) Estimating time
since influenza virus exposure using
single-cell proteomic data.
Front. Immunol. 17:1787198.
doi: 10.3389/fimmu.2026.1787198

COPYRIGHT

© 2026 Rizzo Nervo, Hajiakhoond Bidoki,
Chen, Rahil, Bjornson, Kim, Bock, Affrime,
Bauerle, Huynh, Liebowitz, Tucker,
Gherardini, Nolan, Aghaeepour and
Mcllwain. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance
with accepted academic practice. No
use, distribution or reproduction is
permitted which does not comply with
these terms.

Estimating time since influenza virus exposure using single-cell proteomic data

Klodiana Rizzo Nervo^{1†}, Neda Hajiakhoond Bidoki^{2†}, Han Chen³,
Zainab Rahil³, Zach Bjornson³, Kenneth Kim⁴, Bonnie Bock⁵,
Melton Affrime⁵, Logan Bauerle¹, Pham Bao Tran Huynh¹,
David Liebowitz⁶, Sean Tucker⁶, Pier Federico Gherardini⁷,
Garry P. Nolan⁸, Nima Aghaeepour^{2‡} and David R. Mcllwain^{1*‡}

¹Department of Microbiology and Immunology, School of Medicine, University of Nevada, Reno, Reno, NV, United States, ²Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University School of Medicine, Stanford, CA, United States, ³Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA, United States, ⁴Ark Clinical Research, LLC, Long Beach, CA, United States, ⁵WCCT Global, Cypress, CA, United States, ⁶Vaxart, Inc., South San Francisco, CA, United States, ⁷Department of Biology, University of Rome "Tor Vergata", Rome, Italy, ⁸Department of Pathology, Stanford University School of Medicine, Stanford, CA, United States

Introduction: Determining when the onset of a respiratory infection occurred is important for effective clinical management and can aid in mapping transmission events. However, current diagnostic assays report only pathogen detection status and do not provide any information about the timing of infection, in part because of the lack of biomarkers that inform time since exposure.

Methods: To address this gap, we developed immune-based predictive models of infection timing and shedding status using data from a controlled human challenge with influenza A/California/2009 (H1N1), in which major immune cell subsets were longitudinally profiled across multiple time points before and after viral challenge using 42-marker mass cytometry panels. Random forest machine learning models were trained to address two predictive objectives: (1) distinguishing virus shedders from non-shedders and (2) estimating days post-infection challenge (DPC) from immune profiles. Model performance was evaluated within the primary challenge cohort and independently validated using data from a separate controlled human influenza challenge study using the same virus.

Results: Our analysis revealed that single-cell immune population dynamics alone encode a robust and reproducible temporal structure following influenza infection, enabling accurate estimation of virus exposure timing.

Discussion: These findings provide foundational insight into host immune responses during influenza infection and represent an early step toward a future class of immune-based diagnostics that could extend beyond pathogen detection to inform infection timing and the duration of periods associated with contagiousness.

KEYWORDS

controlled human virus challenge, influenza A (H1N1), machine learning, mass cytometry (CyTOF), random forest models

Introduction

Influenza viruses pose a continual threat to public health by causing significant seasonal and pandemic respiratory diseases. In the United States alone, annual epidemics are estimated to cause 9–49 million infections, 140,000–710,000 hospitalizations, and 12,000–79,000 deaths each year (1). Worldwide, the catastrophic 1918 outbreak underscores the virus's ability for rapid spread, severe morbidity, and large-scale mortality (2). Despite widespread vaccination efforts and the availability of antiviral therapies, influenza continues to exert a substantial and recurring health and economic burden.

Existing diagnostics for influenza and other respiratory infections rely primarily on rapid antigen detection assays and nucleic acid amplification tests (NAATs), such as reverse transcription-polymerase chain reaction (RT-PCR). Rapid antigen tests are inexpensive and offer near-immediate results but suffer from limited sensitivity. NAATs provide superior sensitivity and specificity but typically require more time, cost, and laboratory infrastructure (3). Clinically, both testing modalities generally offer only a binary readout of viral presence but reveal no information on when the infection began or how long an individual may remain infectious (4). As a result, clinicians and public authorities lack actionable tools to estimate infectious duration or guide time-sensitive decisions related to treatment, isolation, and contact tracing.

Controlled human challenge studies offer an ideal context to identify factors capable of estimating time since infection (TSI) because exposure timing is precisely defined. Prior efforts to estimate TSI have primarily focused on host transcriptional responses measured in controlled human challenge settings. Longitudinal gene expression signatures derived from influenza A (H1N1, H3N2), respiratory syncytial virus (RSV), and human rhinovirus (HRV) challenge studies have been shown to accurately infer timing of influenza exposure (5). Whole-blood transcriptomic analyses from individuals challenged with HRV, RSV, H1N1, and H3N2 have also identified early biomarkers that discriminate uninfected controls from individuals who subsequently become viral shedders within the first 32 hours post-infection (6). Additional studies have demonstrated distinct temporal transcriptional dynamics associated with asymptomatic versus symptomatic influenza A infection (7). Similar transcriptomic approaches applied to COVID-19 cohorts have successfully predicted both viral shedding status and stage of infection (8). Serological approaches have likewise been used to estimate infection recency beyond the acute phase (9).

While prior studies make it clear that peripheral blood contains temporal information, there remains a lack of studies leveraging single-cell data to estimate TSI for respiratory infections. Circulating cellular immune responses can capture coordinated changes across multiple immune compartments during infection and may provide a more direct path for identifying a minimal feature set needed for robust time since infection inference.

High-dimensional single-cell technologies, such as mass cytometry, also known as cytometry by time-of-flight (CyTOF), enable detailed characterization of immune system dynamics across

infection stages by simultaneously measuring dozens of protein markers at the single-cell level. Mass cytometry has become a powerful tool to study infectious disease and vaccine immunology, offering the resolution needed to capture nuanced shifts in cellular immunity across multiple infection and vaccination contexts, including severe acute respiratory syndrome SARS-CoV-2, influenza A, dengue, Ebola, and others (10–16).

In this study, we analyzed peripheral blood samples from healthy adults challenged with influenza A/California/2009 (H1N1) using 42-marker mass cytometry panels to profile immune cell subsets before and after infection. We applied random forest machine learning models to classify subjects as virus shedders or non-shedders and to predict days post-infection challenge (DPC) from immune profiles, with model generalizability evaluated in an independent challenge cohort receiving the same influenza challenge virus. We demonstrate that high-dimensional single-cell immune population dynamics alone encode a robust and reproducible temporal signal following influenza infection, enabling accurate inference of infection timing across independent controlled human challenge studies. These results lay the groundwork for potential future diagnostic approaches capable of detecting infection and estimating when it began, and by extension, how long infectiousness may persist.

Materials and methods

Mass cytometry and clinical cohorts

Mass cytometry data were retrospectively analyzed from two independent controlled human influenza challenge studies, referred to as Study A and Study B, both of which used the same A/California/2009 (H1N1) challenge strain (17).

For Study A (11), data were analyzed from 35 volunteers with samples collected at baseline and on days 1–7 post-challenge (DPC). Volunteers were classified as non-shedders ($n = 19$) or viral shedders ($n = 16$), indicating having two or more positive nasopharyngeal swab (qRT-PCR) measurements of viral RNA.

For Study B (18), data were analyzed from 31 volunteers selected to provide a cohort comparable to Study A with respect to sampling timepoints and viral shedding status. This cohort included 15 non-shedders and 16 shedders and, with representation across the three study treatment arms. Viral shedding status was determined in the original study using nasopharyngeal swab qRT-PCR data and was used as reported here. Samples were collected between days 1 and 7 DPC; however, the study design did not include collection at DPC4 or DPC6, and DPC7 samples were not collected for non-shedders or for four shedders.

Mass cytometry was performed on venipuncture whole blood samples that were fixed and cryopreserved using Smart Tube Proteomic Stabilizer. A consistent workflow for staining, data acquisition, and data processing was applied across both studies. Detailed descriptions of the experimental workflow and antibody staining panels used for each study have been published previously (11, 12).

Input data for machine learning consisted of mass cytometry measurements from the Study A and Study B cohorts. Collection day variables were standardized by converting study-specific day numbering schemes into a unified scale relative to viral challenge (DPC).

Normalization

Cell population abundances from mass cytometry were calculated as frequencies of total mononuclear cells, as previously described (11, 12). To reduce noise associated with inter-individual variability, immune features were baseline-normalized by subtracting each participant's pre-challenge values from post-challenge measurements. Two datasets were generated: baseline-normalized and not-normalized (raw). Both were analyzed using identical modeling pipelines, and model performance for viral shedding classification and time-since-infection prediction was compared to assess the impact of normalization.

Classification model

For the binary classification task of distinguishing virus shedders from non-shedders, we implemented a random forest classifier (19). Within Study A, models were trained and validated both at individual DPC and across all days combined. Performance was evaluated using a 5-fold cross-validation strategy, in which folds were stratified by subject ID to ensure that all samples from a given participant were assigned exclusively to either the training or validation set. This design prevented information leakage across datasets and avoided artificially inflated performance estimates. The choice for five-folds was empirical and reflected a balance between computational efficiency and effective control over model bias and variance. To further strengthen robustness, the full cross-validation procedure was repeated one hundred times, and performance metrics were averaged across repetitions. Classification performance was summarized by the mean and standard deviation of key evaluation indicators, including receiver operating characteristics (ROC) and corresponding area under the curve (AUC) values. To ensure computational reproducibility, the random state of the random forest models was controlled using a fixed seed.

For external validation, models trained using Study A data were applied directly to the independent Study B cohort using the same feature set and preprocessing pipeline. No cross-validation was performed within Study B. Classification performance metrics were computed from predictions on Study B samples to enable direct comparison of model generalizability across cohorts.

Receiver operating characteristics

Receiver Operating Characteristic (ROC) curves were generated to evaluate the performance of the random forest classifier in distinguishing virus shedders from non-shedders. ROC curves were constructed by plotting the true positive rate (sensitivity) against the false positive rate ($1 - \text{specificity}$) at varying classification thresholds. Sensitivity was calculated as $TP/(TP + FN)$, and specificity as $TN/(TN + FP)$. The area under the ROC

curve (AUC) was computed as a scalar measure of the classifier's discriminatory performance, ranging from 0 (no discrimination) to 1 (perfect discrimination).

ROC curves and AUC values were calculated using the scikit-learn Python library. Analyses were performed for each individual day post-infection as well as across all days combined. ROC/AUC metrics were also used to compare model performance with and without baseline normalization, allowing assessment of the contribution of normalization and immune feature dynamics to shedding classification.

Regression model

To estimate the DPC at which each blood sample was collected, we trained a random forest regression model using immune cell population features from Study A. Model performance within Study A was evaluated using a leave-one-subject-out (LOSO) cross-validation strategy. In each fold, all samples from one participant across all available time points were held out, and predictions were generated for that participant's samples using a model trained on all remaining participants. This approach ensured that each evaluation was performed on entirely unseen individuals, preventing data leakage and providing a realistic assessment of subject-level generalizability. Predictions from all LOSO folds were pooled to quantify overall model performance.

Model generalizability was then assessed using an independent external validation cohort (Study B). In this setting, the random forest model was trained once using all available samples from Study A and applied directly to Study B samples to generate predicted DPC values. No cross-validation was performed in Study B. Predicted DPC values were summarized as median trends with 95% confidence intervals. Model accuracy was quantified using Pearson correlation coefficients (r) and associated p-values to assess agreement between predicted and actual sampling days.

Regression model prediction correlation

Pearson's correlation coefficient (r) was used to assess the linear association between the regression model's predicted values and the observed post-challenge time points. Correlation analyses were performed in Python using the `pearsonr()` function from the `scipy.stats` module, which provides both the correlation coefficient and a two-tailed p-value testing the null hypothesis of no correlation. Pearson correlation was used as a descriptive measure of agreement between predicted and observed DPC values to assess recovery of temporal structure along the infection timeline, rather than parametric assumption testing about the underlying data distributions.

Root mean square error

Root Mean Square Error (rMSE) was used to quantify prediction accuracy in the random forest regression models. rMSE measures the average magnitude of prediction errors between the model's predicted time post-infection and the actual time points of sample collection. rMSE was defined as follows,

where y_i are observed values and $\hat{y}_i$ are predicted values:

$$rMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

In this study, rMSE was calculated in two ways. First, rMSE was computed separately for each day post-challenge (DPC) using only samples collected at that time point, allowing prediction accuracy to be compared across stages of infection.

Second, an overall rMSE was computed by pooling all predictions across DPCs and calculating a single rMSE, equivalent to averaging the squared prediction errors across all samples before taking the square root. rMSE calculations were implemented in Python using NumPy by computing the square root of the mean squared difference between predicted and observed values, excluding non-finite entries.

Correlation network analysis

Correlation networks were constructed to visualize univariate immune cell population feature importance and to contextualize these features based on cell population identity and similarity to other cell populations. Nodes represent gated immune cell populations, with node diameter encoding scaled univariate feature importance. Edges connect populations with similar phenotypic marker expression (Spearman $\rho \geq 0.7$). Spearman correlation was used as a non-parametric measure to reduce sensitivity to non-normal distributions and outliers.

Nodes were embedded in two dimensions using t-distributed stochastic neighbor embedding (t-SNE) to cluster populations with similar phenotypic marker expression profiles. Nodes were also colored by cell population category. Univariate feature importance was quantified by the correlation between normalized population abundance and sampling day, summarized as $-\log_{10}(p)$ or R^2 values, which were normalized for visualization and used to scale node diameter. Larger diameter nodes therefore indicate immune populations with relatively stronger infection- or time-associated patterns.

Trajectory visualization by UMAP

Longitudinal immune cell trajectories were visualized using uniform manifold approximation and projection (UMAP) applied to immune cell frequency data across all samples and participants. Prior to dimensionality reduction, immune cell frequencies were normalized across samples to place features on a comparable scale. UMAP was then performed on the resulting feature matrix to embed samples into two dimensions, with UMAP1-UMAP2 coordinates reflecting similarity in immune cell frequency composition. Each sample was represented as a single point. The position of points in 2D space is dictated solely by immune cell frequency features and is independent of days post-challenge (DPC). After UMAP embedding, samples were colored by either actual or model-predicted DPC for visualization. The trajectory line traces the central tendency of samples across DPC for visualization purposes, with bold numbers marking each DPC.

Software

All analyses were performed in Python using the pandas, NumPy, SciPy, Matplotlib, Seaborn, and scikit-learn libraries (20, 21), as commonly adopted in other single-cell studies (22).

Results

Study design

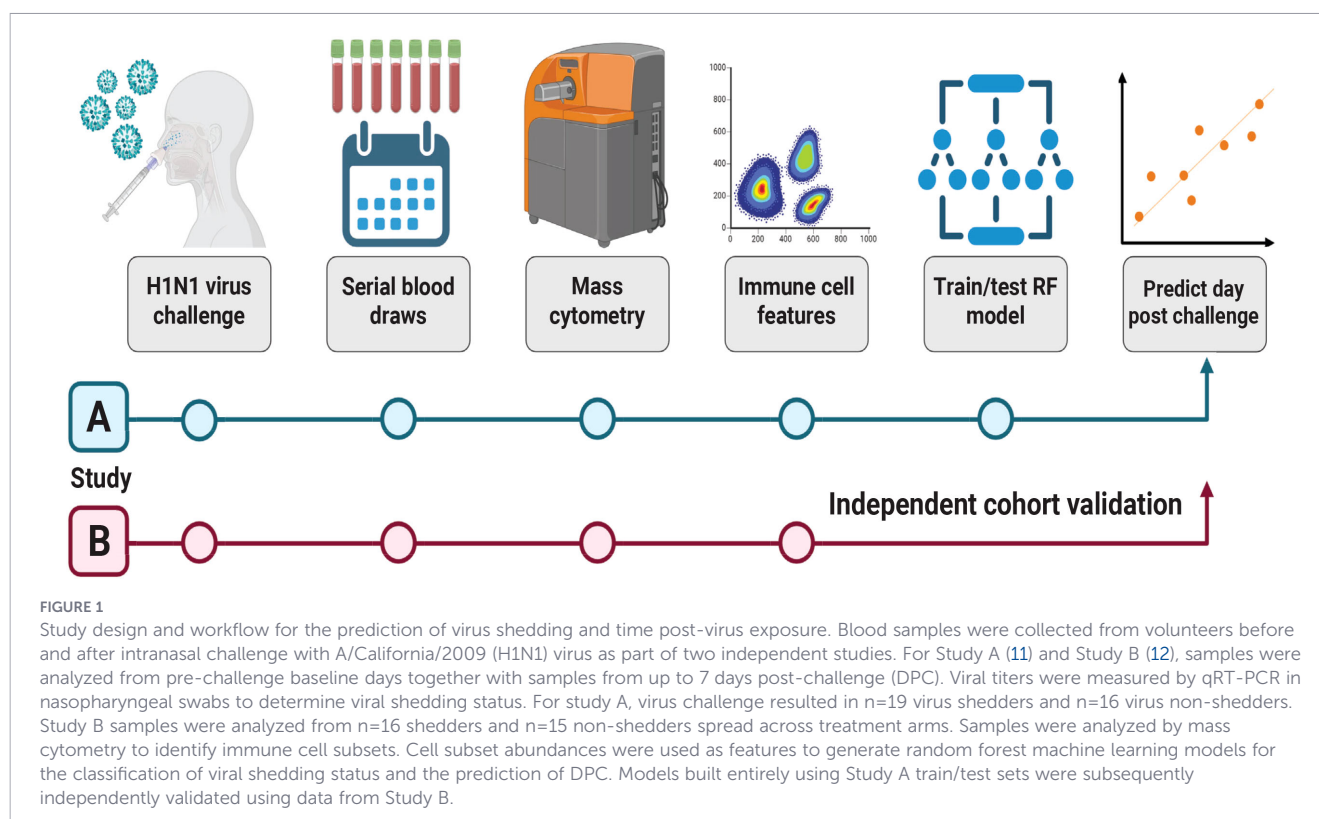
The objective of this study was to determine whether integrated immune cell responses quantified by mass cytometry could be used in a machine learning framework to predict the timing of influenza virus infection. To achieve this, we analyzed data from two controlled human challenge studies using the same influenza A/California/2009 (H1N1) strain, described below. Use of an identical challenge strain and closely aligned experimental protocols across studies provided a consistent experimental context with respect to viral genetics and exposure conditions for evaluating temporal immune dynamics. Together, these cohorts provided complementary datasets for training and validating predictive models of infection dynamics (Figure 1).

Study A (11) enrolled 35 healthy adults who were intranasally inoculated with a standardized dose of A/California/2009 (H1N1), delivered by atomizer, and were intensively monitored during screening, confinement, and the post-challenge follow-up. Peripheral blood samples were collected daily until the seventh day post challenge (DPC), including two pre-challenge baseline samples. Participants were classified as virus shedders or non-shedders based on quantitative RT-PCR results from nasal swabs and supporting clinical assessments.

Study B (12, 18), was a clinical trial assessing the investigational oral influenza vaccine from Vaxart (VXA-A1.1) alongside a marketed vaccine and placebo control. These individuals were challenged with the same A/California/2009 (H1N1) strain under conditions aligned with study A. Peripheral blood samples were similarly collected across the course of infection, and shedder status was assigned using equivalent criteria applied in Study A.

Immune profiling by mass cytometry

Mass cytometry was performed using 42-antibody panels to identify and quantify immune cell populations in peripheral blood samples collected from both studies. Manual gating resolved 37 immune cell populations (Supplementary Table 1), that reliably matched across both studies at consistent DPC and were used for modeling (Supplementary Table 2). These included major subsets of monocytes, granulocytes, B-cell, T-cell, and NK-cells. The full mass cytometry datasets from each study used as model inputs are provided in Supplementary Table 3 and Supplementary Table 4. Prior analyses of the Study A dataset showed that these immune profiles resolved distinct immune response patterns between virus shedders and non-shedders (11).



Several immune populations displayed distinct population-specific temporal patterns following viral challenge (Figure 2). Classical monocytes (cMCs) and intermediate monocytes (intMCs) peaked around DPC 3 in shedders, with intMCs reaching maximum abundance slightly later on DPC 4. Nonclassical monocytes (ncMCs) increased until DPC 5 before declining. Basophils, defined as CD123⁺HLA-DR⁻ cells, decreased sharply in shedders. Activated and proliferating lymphocytes, including CD4⁺ T cells (CD38⁺Ki67⁺), CD8⁺ memory T cells (CD38⁺Ki67⁺), and NK cells (CD56^{lo}CD16⁺ CD38⁺Ki67⁺), increased steadily in shedders through DPC 7. Class-switched memory B cells declined for shedders reaching nadirs at day 4 or 5. In contrast, non-shedders maintained relatively stable immune cell frequencies across most populations.

These coordinated and resolved changes in immune cell abundances over time reveal a clear trajectory of immune cell mobilization and contraction during influenza infection. The strong temporal structure of these immune responses suggests that peripheral immune profiles may encode sufficient information to infer the presence of infection and infection timing.

Predicting viral shedding status with machine learning

To determine whether immune profiles could reliably distinguish infection outcomes, we trained random forest classifiers to differentiate virus shedders from non-shedders using baseline-normalized data. The models were trained and validated independently at each study day in both Study A and Study B. Furthermore, we used a bootstrapping strategy combined with five-

fold cross-validation to account for the modest sample size and ensure robust performance estimates.

The classifiers achieved consistent predictive accuracy across multiple time points, with area-under-the-curve (AUC) values summarized in Figures 3A, B. Models trained on Study A data generalized to the independent Study B cohort with cross-cohort AUC values ranging from 0.50 to 0.81 across time points, indicating that peripheral immune signatures contain sufficient discriminatory power to classify viral shedding status following influenza challenge. In the absence of baseline normalization, models exhibited reduced overall predictive performance, indicating that anchoring immune features to each participant's pre-challenge baseline is critical for amplifying infection-related immune signals (Supplementary Figures 1A, B).

To investigate the contribution of individual immune cell populations to model performance, we constructed network visualizations that highlight relationships among cell populations based on phenotypic marker expression (Figures 3C, D). Feature importance analyses indicated contributions from both innate and adaptive immune cell populations across studies, with NK cell and granulocyte subsets among the most consistently informative.

These results demonstrate that integrated immune cell population data provides a robust framework for distinguishing likely infection outcomes across independent human challenge studies.

Predicting time since infection onset with machine learning

We next evaluated whether immune cell dynamics could be leveraged to infer the time since infection onset in viral shedders.

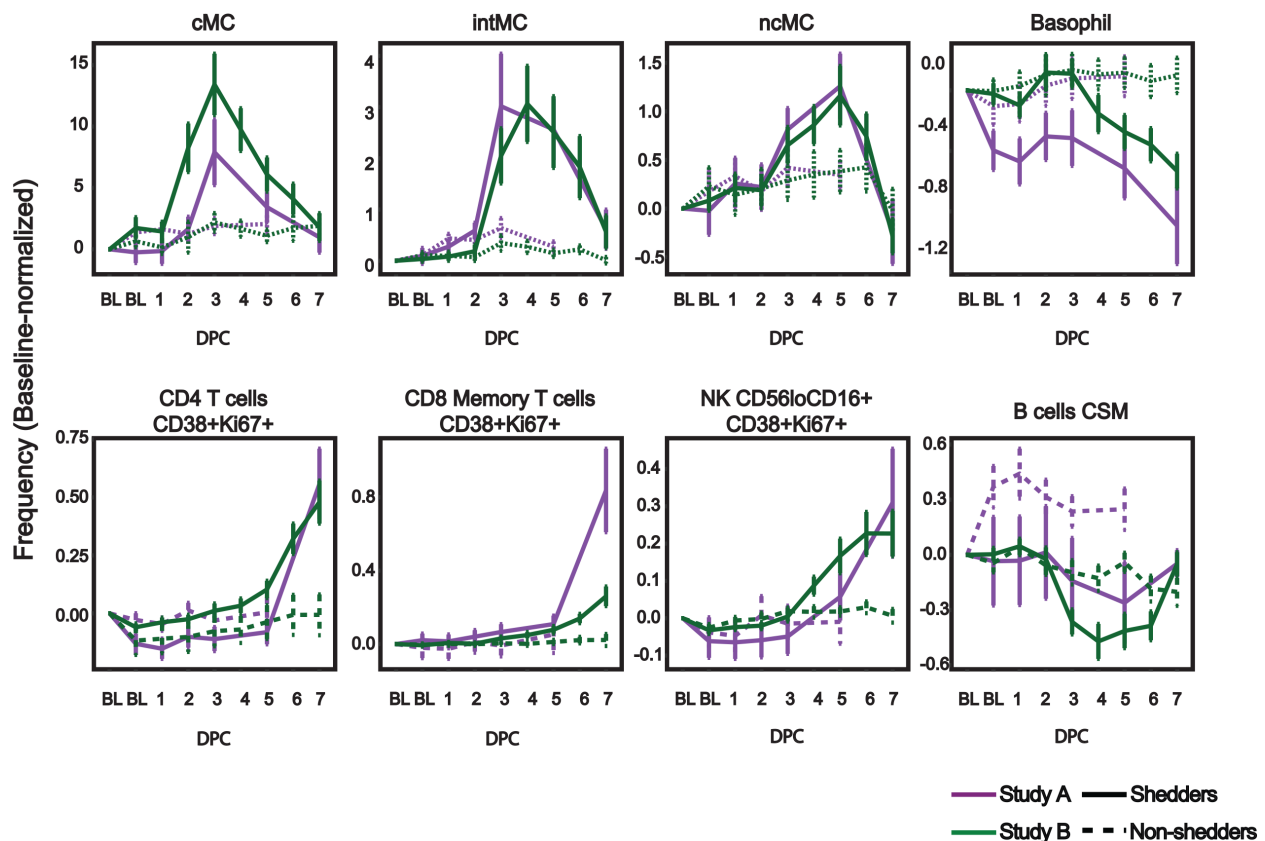


FIGURE 2

Post-challenge immune population kinetics in viral shedders align across studies. Relative abundances of indicated immune cell populations across baseline (BL) and day post-infection challenge (DPC). Classical monocytes (cMCs)(CD14+CD16-), intermediate monocytes (intMC)(CD14+CD16+), non-classical monocytes (ncMC)(CD16+CD14-), basophil (CD123+HLA-DR-), activated proliferating CD4 T cells (CD4+CD38+Ki67+), activated proliferating CD8+ Memory T cells (CD8+ CD45RA-CD38+ Ki67+), activated proliferating NK cells (CD56 loCD16+ CD38+ Ki67+), class-switched memory B cells (B cells CSM) (IgM-CD27+). For gating, see Rahil et al. and McIlwain and Chen et al. (11, 12). Values show relative abundance (% of mononuclear CD66- cells normalized by subtraction of baseline); filled squares and solid lines indicate virus shedders, and open squares and dashed lines indicate virus non-shedders. Green lines show samples from Study A (n=19 shedders, n=16 non-shedders), purple lines show samples from Study B (n=16 shedders, n=15 non-shedders). Plotted values are mean frequency $\pm$ SEM.

Using Study A data, random forest regression models with one hundred estimators were trained using repeated stratified five-fold cross-validation (100 iterations), in which each iteration excluded one participant for testing while the model was trained on the remaining subjects. Both baseline-normalized and raw abundance data were analyzed to determine the effect of normalization on model performance, based on the expectation that incorporating each individual's baseline immune profile would reduce biological variability and improve predictive accuracy.

For Study A, predicted DPC showed a strong correlation with actual DPC for viral shedders ($r = 0.92$, $p\text{-value} = 5.87 \times 10^{-53}$ for normalized data; $r = 0.70$, $p\text{-value} = 4.19 \times 10^{-20}$ for unnormalized data), as illustrated in Figure 4A and in Supplementary Figures 2A, B. Overall prediction error (rMSE, in days) was 1.060 for the Study A normalized model and 1.490 for the Study A unnormalized model. Predictions also generalized effectively by applying the model to the independent Study B cohort, yielding a robust correlation between predicted and actual DPC for viral shedders ($r = 0.78$, $p\text{-value} = 5.04 \times 10^{-17}$ for normalized data; $r = 0.57$, $p\text{-value} = 8.61 \times 10^{-8}$ for unnormalized data), as displayed in Figure 4B and in Supplementary Figures 2C, D. Corresponding overall rMSE values in Study B were 1.483 (normalized) and 1.490 (unnormalized).

Attempts to model temporal trajectories in non-shedders did not produce meaningful temporal predictions, consistent with the absence of sustained infection-driven immune perturbation in these individuals (Supplementary Figure 3). Taken together, these results indicate that immune cell dynamics contain a measurable temporal signal that can be decoded to estimate time since infection, and that anchoring each participant's data to their own baseline substantially enhances the stability and the accuracy of these estimates.

To examine the evolution of immune cell trajectories over the course of infection, we visualized all samples from all participants using uniform manifold approximation and projection (UMAP) built based on immune cell frequency data (Figures 4C, D). Here each point represents a sample, with UMAP1-UMAP2 coordinates reflecting global similarity in immune cell frequency composition. In both Studies A and B, samples followed an elliptical trajectory in UMAP space, with peak viral load timepoints located furthest from baseline. This pattern is consistent with a coordinated progression from infection onset through maximal immune activation and subsequent recovery. Points were colored by either actual or model-predicted days post-challenge (DPC). In both cases, samples with similar DPC clustered together along the UMAP

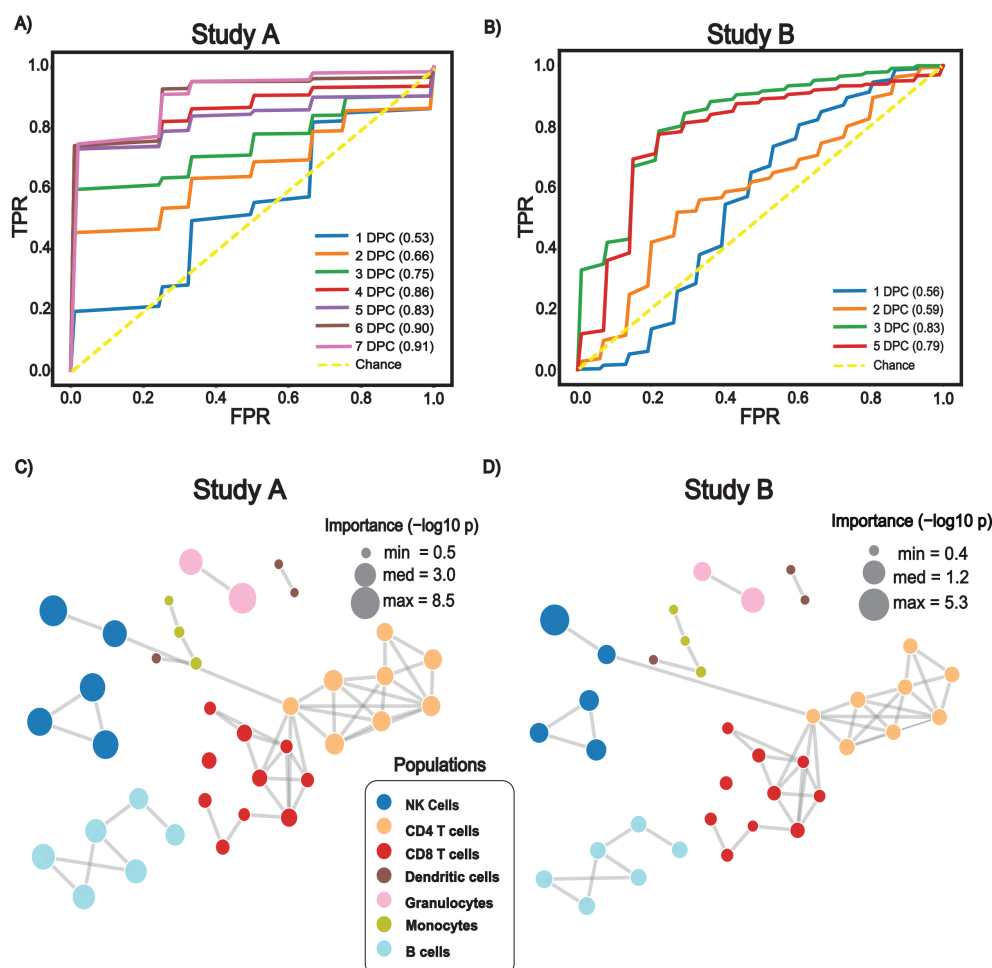


FIGURE 3

Random forest model effectively classifies virus shedding across studies. (A, B) Receiver operating characteristic (ROC) curves evaluating the performance of random forest models built using baseline-normalized data to classify virus shedders from non-shedders. Lines and AUC values show model performance at the indicated DPCs. (A) Model trained and validated on Study A, (B) Model trained on Study A and applied to the independent Study B cohort. (C, D) Correlation networks used to visualize and contextualize univariate immune population feature importance for Study A (C) and Study B (D), where nodes represent gated immune cell populations, and edges indicate similarity in phenotypic marker expression between populations. Univariate importance scores ($-\log_{10}(p)$) for distinguishing virus shedders from non-shedders are projected onto the network and scale node size, with larger nodes indicating stronger predictive importance. Networks contextualize the identity of immune cell populations and their relative importance to the classification models.

trajectory, demonstrating close agreement between predicted timing and observed immune cell dynamics.

Model performance was further examined by comparing predicted versus actual DPC for each participant across each study (Figures 4E, F). The resulting heatmaps display accurate temporal ranking across individuals and timepoints. Equivalent heatmaps generated using unnormalized data showed reduced consistency, reinforcing the advantage of baseline-normalization (Supplementary Figure 4).

Lastly, to explore cellular features driving temporal predictions, we applied the same network-based approach used for the classifier analyses. Networks depicting relationships among immune cell populations based on phenotypic marker expression were overlaid with univariate feature importance scores for days post-challenge (DPC) prediction (Figures 5A, B). Node size was scaled by the strength of each population's association with time, as measured by Spearman correlation values, highlighting features with the strongest temporal relationships. To directly compare feature importance

across studies, a scatter plot was used to visualize head-to-head rankings between models (Figure 5C). Immune populations consistently ranked among the most informative in both studies included activated and proliferating T cell and NK cell subsets (Figure 5C), consistent with their well-characterized expansion and activation dynamics during acute respiratory infection.

Collectively, these results suggest that temporal variations in immune cell populations encode a measurable and reproducible signal of infection timing, enabling reliable estimation of time since exposure across independent controlled human influenza challenge studies.

Discussion

We applied machine learning to high-dimensional immune profiling data from two controlled human influenza challenge

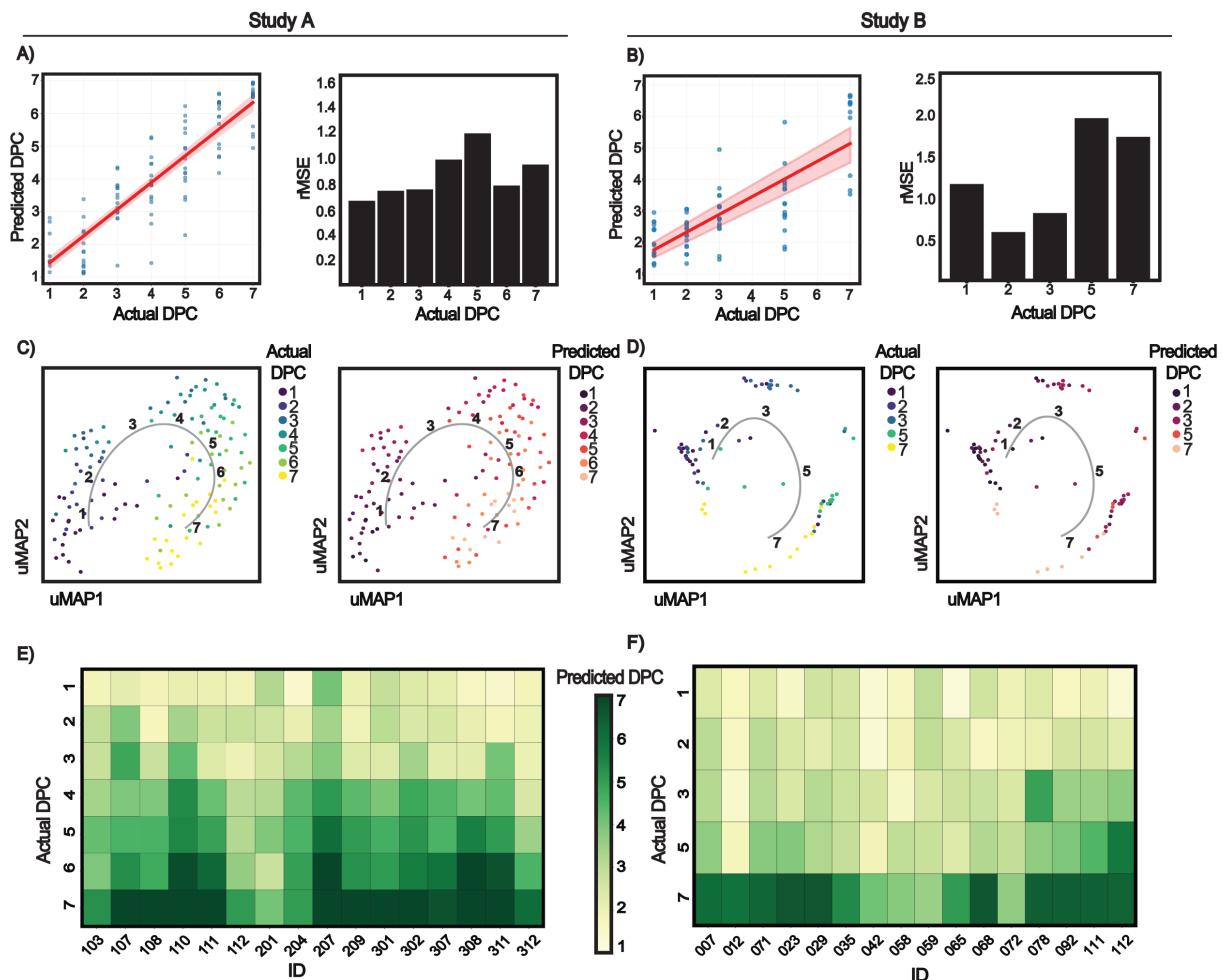


FIGURE 4

Machine learning-based prediction of time post virus exposure. (A, B) Random forest models predicting time post virus exposure in virus shedders using baseline-normalized data. (A) Leave-one-out cross-validation in Study A (Pearson $r = 0.92$, $p = 5.87 \times 10^{-53}$, $n = 19$) with corresponding root mean squared error (rMSE). (B) Model trained on Study A and applied to the independent Study B cohort (Pearson $r = 0.78$, $p = 5.04 \times 10^{-17}$, $n = 16$) with rMSE shown. (C, D) UMAP projections visualizing immune state trajectories over time and colored by actual or predicted DPC. UMAP projections of longitudinal immune cell frequencies from Study A (C) and Study B (D), illustrating trajectories from baseline through peak infection and toward resolution; legends indicate whether color scale displays actual or model-predicted DPC, which show strong concordance. Each point represents a single sample embedded based on immune cell frequency features only (independent of DPC). Point colors are applied after UMAP embedding and do not influence the UMAP coordinates. The trajectory line traces the central tendency of samples across DPC, with bold numbers marking each DPC. (E, F) Heatmaps comparing predicted versus actual DPC by subject for Study A (E) and Study B (F), demonstrating accurate temporal inference and generalization to an independent cohort.

studies to demonstrate that circulating immune cell dynamics contain predictive information about infection stage. Immune population trajectories not only distinguished viral shedders from non-shedders but also enabled accurate prediction of time since infection challenge. Random forest models trained in one cohort generalized to an independent challenge study, underscoring the robustness and reproducibility of the temporal immune signatures identified under controlled experimental conditions. Together, these findings establish single-cell immune population dynamics as an independent and previously underexplored modality for infection staging.

Previous work has shown that respiratory viral infection stage can be inferred from blood-based biomarker data through transcriptional and serological profiling (5–9). Our results complement this body of work by demonstrating that single-cell immune phenotyping alone, without transcriptomic or serological

data, captures temporal patterns of infection at high resolution. These findings indicate that cellular immune dynamics represent an underexplored source of biological information for infection staging. Future translation of these insights toward practical diagnostics will likely depend on identifying a minimal set of temporally informative cellular features that can be measured using scalable, lower-complexity platforms more suitable than mass cytometry.

Integrating measurement of such features with existing molecular diagnostic modalities could substantially improve clinical decision-making by providing estimates of infection status and timing, enabling more precise guidance on treatment interventions. Because the temporal relationship between time since exposure and viral shedding has been characterized for influenza, accurate estimation of time since infection alone may provide an indirect framework to improve estimates of expected

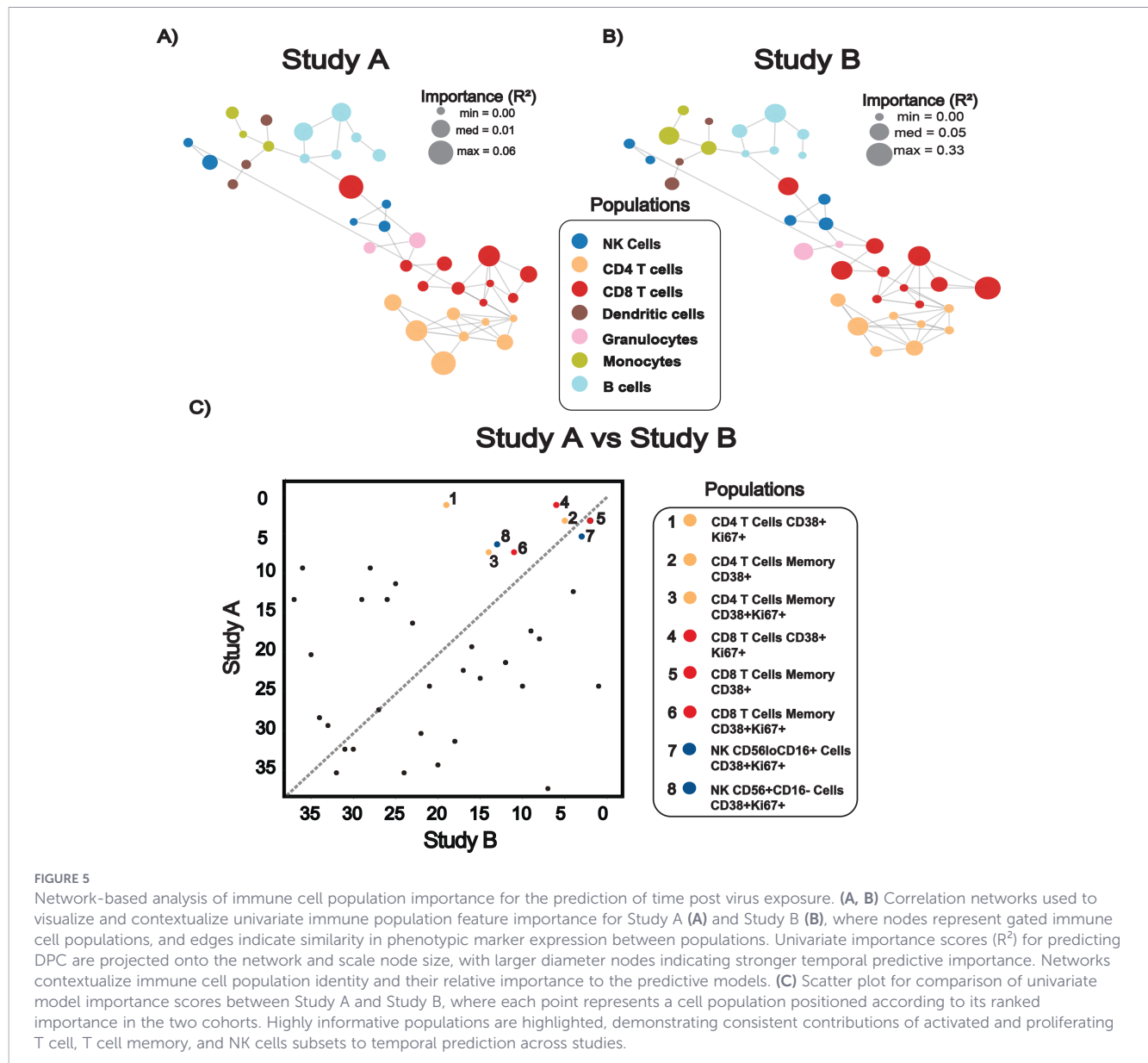


FIGURE 5

Network-based analysis of immune cell population importance for the prediction of time post virus exposure. (A, B) Correlation networks used to visualize and contextualize univariate immune population feature importance for Study A (A) and Study B (B), where nodes represent gated immune cell populations, and edges indicate similarity in phenotypic marker expression between populations. Univariate importance scores (R^2) for predicting DPC are projected onto the network and scale node size, with larger diameter nodes indicating stronger temporal predictive importance. Networks contextualize immune cell population identity and their relative importance to the predictive models. (C) Scatter plot for comparison of univariate model importance scores between Study A and Study B, where each point represents a cell population positioned according to its ranked importance in the two cohorts. Highly informative populations are highlighted, demonstrating consistent contributions of activated and proliferating T cell, T cell memory, and NK cells subsets to temporal prediction across studies.

shedding windows. Whether similar immune-cell-based models could be developed in studies with larger cohorts and more precisely resolved longitudinal shedding measurements to directly predict the duration of viral shedding warrants investigation in future work. Identifying circulating immune cell responses predictive of atypical prolonged viral shedding in community-acquired respiratory viral infection would be particularly valuable.

In this study, model performance benefited from normalization to individual pre-challenge immune baselines, highlighting a dependency on baseline measurements for optimal temporal inference. Although this requirement could pose practical challenges for deployment, it also underscores the broader value of establishing individual immune baselines for interpreting immune perturbations, which may prove relevant across multiple healthcare contexts, including infection, vaccination, and other inflammatory or physiological stressors (23). As longitudinal

immune profiling becomes more feasible, baseline-referenced models may support more precise interpretation of immune state changes.

Influenza A viruses evolve through antigenic drift and, less frequently, antigenic shift, with prominent variation in the surface glycoproteins hemagglutinin (HA) and neuraminidase (NA). Viral genetic differences can alter replication kinetics, host immunity, and the timing and magnitude of inflammatory and cellular responses (2). An important limitation of our study is that models were trained and validated using two studies conducted with a single A/California/2009 (H1N1) challenge strain; the extent to which the same immune features and learned mappings generalize to antigenically distinct strains remains uncertain. Contemporary influenza A controlled human infection models have largely focused on H1N1 and H3N2 challenge strains (24), reflecting practical and ethical constraints on human challenge-strain

development and use. An important next step is to evaluate performance across additional controlled influenza challenge strains; matched multi-strain challenge cohorts would enable a direct test of model generalizability.

Furthermore, because this analysis was performed in controlled human influenza challenge studies with precisely defined exposure timing and standardized infection conditions, additional work would be required to evaluate how immune-cell-based temporal inference models perform in community-acquired influenza settings. In natural infection contexts, greater variability in viral genetics, infectious dose, and host population characteristics may influence immune response trajectories, while less uniform clinical sampling schedules may introduce additional noise. Cohort sizes in the present study were modest, constraining model complexity and limiting power to detect subtler cellular features. Future studies in larger and more diverse populations that span broader influenza strain diversity and reflect real-world, less uniform community sampling will be essential to evaluate generalizability and to develop more comprehensive temporal prediction models that remain robust to evolving influenza viruses.

In summary, this study demonstrates that coordinated changes in circulating immune cell populations encode measurable information about infection timing that can be extracted using high-dimensional phenotyping coupled with machine learning. Across two independent controlled human influenza challenge studies, these immune dynamics supported accurate inference of infection stage and generalized across cohorts challenged with the same viral strain. Together, these findings establish a benchmark for immune-based temporal inference under well-defined experimental conditions and provide a foundation for future studies evaluating the performance of such approaches in more heterogeneous, community-acquired infection settings.

Data availability statement

The datasets used for model building in this study are available in [Supplementary Tables 3 and 4](#). Code used in this study is available from Github (<https://github.com/davemcilwain/Flu-Day-Prediction>).

Ethics statement

This study involves retrospective analysis of data derived from previously reported clinical trials and does not itself constitute human subjects research. The source studies have been described previously by Rahil et al. (11) and Liebowitz et al. (18), with trial details available on ClinicalTrials.gov under identifiers NCT04106817 and NCT02918006. Both studies were conducted with written informed consent from participants, were approved by institutional review boards with data and safety monitoring board oversight.

Author contributions

KR: Conceptualization, Methodology, Supervision, Investigation, Writing – review & editing, Writing – original draft. NH: Methodology, Supervision, Conceptualization, Writing – original draft, Writing – review & editing, Investigation. HC: Investigation, Writing – review & editing. ZR: Investigation, Writing – review & editing. ZB: Writing – review & editing, Investigation. KK: Methodology, Writing – review & editing, Investigation. BB: Writing – review & editing, Investigation. MA: Writing – review & editing, Investigation. LB: Writing – review & editing. PH: Writing – review & editing. DL: Writing – review & editing. ST: Investigation, Methodology, Writing – review & editing. PG: Methodology, Writing – review & editing. GN: Methodology, Supervision, Writing – review & editing, Investigation. NA: Writing – original draft, Conceptualization, Methodology, Writing – review & editing, Investigation, Supervision. DM: Methodology, Supervision, Investigation, Writing – review & editing, Conceptualization, Writing – original draft.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported in part by US Department of Health and Human Services (HHS), Office of the Assistant Secretary for Preparedness and Response, and Biomedical Advanced Research and Development Authority (contract number HHSO100201500034C); US NIH Cooperative Centers for Translational Research in Human Immunology and Biodefense Opportunity Fund seed grant; US Food and Drug Administration Medical Countermeasures Initiative contracts 75F40120C00176 and HHSF223201610018C. This article reflects the views of the authors and should not be construed as representing the views or policies of HHS agencies or other groups and companies providing funding or affiliated with the authors.

Conflict of interest

KK was employed by company Ark Clinical Research, LLC. DL, ST are current or former employees of Vaxart Inc. MA, BB and KK are former employees of WCCT Global.

The remaining author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declare that this work received funding from Vaxart Inc (Vaxart/Stanford sponsored research agreement #137364). The funder was involved in investigation, methodology, and writing/review and editing.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2026.1787198/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Random forest model classifying virus shedding across studies without baseline normalization. ROC curves evaluating the performance of random forest models built using raw, not-normalized data to classify virus shedders from non-shedders. Lines and AUC values show model performance at the indicated DPCs. (A) Model trained and validated on Study A. (B) Model trained on Study A and applied to the independent Study B cohort. Results illustrate reduced overall classification accuracy and generalization compared to baseline-normalized models shown in Figure 3.

SUPPLEMENTARY FIGURE 2

Machine learning-based prediction of time post virus exposure without baseline normalization. (A–D) Random forest models predicting time post virus exposure in virus shedders using raw, not-normalized data. (A, B) Leave-one-out cross-validation in Study A (Pearson $r = 0.70$, $p = 4.19 \times 10^{-20}$, $n = 19$) with corresponding root mean squared error (rMSE). (C, D) Model trained on Study A and applied to the independent Study B cohort (Pearson $r = 0.57$, $p = 8.61 \times 10^{-9}$, $n = 16$) with rMSE shown. Results illustrate reduced temporal accuracy and generalization compared to baseline-normalized models shown in Figure 4.

SUPPLEMENTARY FIGURE 3

Expected absence of predictive temporal immune trajectories in non-shedders. Random forest models predicting time post virus exposure in non-shedders. (A–D) Models built using baseline-normalized data. (A) Leave-one-out cross-validation in Study A (Pearson $r = 0.45$, $p = 1.26 \times 10^{-7}$, $n = 16$) with corresponding root mean squared error (rMSE) shown in (B). (C) Model trained on Study A and applied to the independent Study B cohort (Pearson $r = 0.17$, $p = 1.93 \times 10^{-1}$, $n = 15$) with rMSE shown in (D). (E–H) Models built using raw, not-normalized data. (E) Leave-one-out cross-validation in Study A (Pearson $r = 0.27$, $p = 2.13 \times 10^{-3}$, $n = 16$) with rMSE shown in (F). (G) Model trained on Study A and applied to the independent Study B cohort (Pearson $r = 0.01$, $p = 9.14 \times 10^{-1}$, $n = 15$) with rMSE shown in (H).

SUPPLEMENTARY FIGURE 4

Predicted versus actual DPC without baseline normalization. (A, B) Heatmaps comparing predicted versus actual DPC by subject for Study A (A) and Study B (B) using raw, non-normalized data. Each tile represents the median predicted DPC for a given subject relative to the actual DPC, illustrating reduced temporal accuracy and generalization compared to baseline-normalized models shown in Figure 4.

SUPPLEMENTARY TABLE 1

Mapping of immune cell populations between study A and study B. Columns list cell populations from each study dataset, with corresponding or equivalent subpopulations aligned within the same row.

SUPPLEMENTARY TABLE 2

Alignment of sampling timepoints between study A and study B. The DPC (days post-challenge) column indicates the unified timepoint alignment used in this study. The Study A and Study B columns list the original study-specific day numbering schemes reported in (11) and (12) respectively. For Study A, viral challenge occurred on study day 1 (original numbering) following blood draw sampling. For Study B, viral challenge occurred on study day 90 (original numbering) following blood draw sampling. For both studies, the DPC 1 blood draw corresponds to approximately 24 hours post-challenge.

SUPPLEMENTARY TABLE 3

Study A immune cell abundance data used for model features. Columns include volunteer identifier (VOLUNTEER), day post-challenge of blood draw (DPC), and virus shedding status (SHEDDER), encoded as 1 for shedders and 0 for non-shedders, for participants in Study A. Remaining columns report the relative abundances of immune cell populations measured by mass cytometry and used as model features.

SUPPLEMENTARY TABLE 4

Study B immune cell abundance data used for model features. Columns include volunteer identifier (VOLUNTEER), day post-challenge of blood draw (DPC), and virus shedding status (SHEDDER), encoded as 1 for shedders and 0 for non-shedders, for participants in Study B. Remaining columns report the relative abundances of immune cell populations measured by mass cytometry and used as model features.

References

1. CDC. *Flu Burden* (2024). Available online at: <https://www.cdc.gov/flu-burden/php/about/index.html>. About Estimated Flu Burden (Accessed December 10, 2025).
2. Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, et al. Influenza. *Nat Rev Dis Primers*. (2018) 4:3. doi: 10.1038/s41572-018-0002-y
3. Uyeki TM, Bernstein HH, Bradley JS, Englund JA, File TM, Fry AM, et al. Clinical practice guidelines by the infectious diseases society of america: 2018 update on diagnosis, treatment, chemoprophylaxis, and institutional outbreak management of seasonal influenza. *Clin Infect Dis*. (2019) 68:e1–47. doi: 10.1093/cid/ciy866
4. McIlwain DR, Chen H, Apkarian M, Affrime M, Bock B, Kim K, et al. Performance of BioFire array or QuickVue influenza A + B test versus a validation qPCR assay for detection of influenza A during a volunteer A/California/2009/H1N1 challenge study. *Viral J*. (2021) 18:45. doi: 10.1186/s12985-021-01516-0
5. Sharma K, Aminian M, Ghosh T, Liu X, Kirby M. Using machine learning to determine the time of exposure to infection by a respiratory pathogen. *Sci Rep*. (2023) 13:5340–0. doi: 10.1038/s41598-023-30306-7
6. Aminian M, Ghosh T, Peterson A, Rasmussen AL, Stiverson S, Sharma K, et al. Early prognosis of respiratory virus shedding in humans. *Sci Rep*. (2021) 11:17193. doi: 10.1038/s41598-021-95293-z
7. Huang Y, Zaas AK, Rao A, Dobigeon N, Woolf PJ, Veldman T, et al. Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza A infection. *PLoS Genet*. (2011) 7:e1002234. doi: 10.1371/journal.pgen.1002234
8. Jeyanthan P. Prolonged viral shedding prediction on non-hospitalized, uncomplicated SARS-CoV-2 patients using their transcriptome data. *Comput Methods Programs BioMed Update*. (2022) 2:100070. doi: 10.1016/j.cmpbup.2022.100070
9. Pelleau S, Woudenberg T, Rosado J, Donnadieu F, Garcia L, Obadia T, et al. Kinetics of the severe acute respiratory syndrome coronavirus 2 antibody response and serological estimation of time since infection. *J Infect Dis*. (2021) 224:1489–99. doi: 10.1093/infdis/jiab375
10. Simoni Y, Chng MHY, Li S, Fehlings M, Newell EW. Mass cytometry: a powerful tool for dissecting the immune landscape. *Curr Opin Immunol*. (2018) 51:187–96. doi: 10.1016/j.coi.2018.03.023
11. Rahil Z, Leylek R, Schürch CM, Chen H, Bjornson-Hooper Z, Christensen SR, et al. Landscape of coordinated immune responses to H1N1 challenge in humans. *J Clin Invest*. (2020) 130:5800–16. doi: 10.1172/JCI137265
12. McIlwain DR, Chen H, Rahil Z, Bidoki NH, Jiang S, Bjornson Z, et al. Human influenza virus challenge identifies cellular correlates of protection for oral vaccination. *Cell Host Microbe*. (2021) 29:1828–37. doi: 10.1016/j.chom.2021.10.009

13. Reeves PM, Sluder AE, Paul SR, Scholzen A, Kashiwagi S, Poznansky MC. Application and utility of mass cytometry in vaccine development. *FASEB J.* (2018) 32:5–15. doi: 10.1096/fj.201700325r
14. Mukherjee R, Eggesbø LM, Wolf AS, Kjønsstad IF, Solum G, Ravussin A, et al. Mass cytometry reveals cellular correlates of immune response heterogeneity to SARS-CoV-2 vaccination in the elderly. *NPJ Vaccines.* (2024) 9:238. doi: 10.1038/s41541-024-01028-2
15. Hamlin RE, Rahman A, Pak TR, Maringer K, Mena I, Bernal-Rubio D, et al. High-dimensional CyTOF analysis of dengue virus-infected human DCs reveals distinct viral signatures. *JCI Insight.* (2017) 2:e92424. doi: 10.1172/jci.insight.92424
16. Kotliar D, Lin AE, Logue J, Hughes TK, Khoury NM, Raju SS, et al. Single-cell profiling of ebola virus disease *in vivo* reveals viral and host dynamics. *Cell.* (2020) 183:1383–1401.e19. doi: 10.1016/j.cell.2020.10.002
17. Watson JM, Francis JN, Mesens S, Faiman GA, Makin J, Patriarca P, et al. Characterisation of a wild-type influenza (A/H1N1) virus strain as an experimental challenge agent in humans. *Virology.* (2015) 12:13. doi: 10.1186/s12985-015-0240-5
18. Liebowitz D, Gottlieb K, Kolhatkar NS, Garg SJ, Asher JM, Nazareno J, et al. Efficacy, immunogenicity, and safety of an oral influenza vaccine: a placebo-controlled and active-controlled phase 2 human challenge study. *Lancet Infect Dis.* (2020) 20:435–44. doi: 10.1016/S1473-3099(19)30584-5
19. Breiman L. Random forests. *Mach Learning.* (2001) 45:5–32. doi: 10.1023/A:1010933404324
20. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods.* (2020) 17:261–72. doi: 10.1038/s41592-019-0686-2
21. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. *J Mach Learn Res.* (2011) 12:2825–30.
22. Weinreb C, Wolock S, Tusi BK, Socolovsky M, Klein AM. Fundamental limits on dynamic inference from single-cell snapshots. *Proc Natl Acad Sci U S A.* (2018) 115: E2467–76. doi: 10.1073/pnas.1714723115
23. Tsang JS, Schwartzberg PL, Kotliarov Y, Biancotto A, Xie Z, Germain RN, et al. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell.* (2014) 157:499–513. doi: 10.1016/j.cell.2014.03.031
24. Balasingam S, Wilder-Smith A. Randomized controlled trials for influenza drugs and vaccines: a review of controlled human infection studies. *Int J Infect Diseases.* (2016) 49:18–29. doi: 10.1016/j.ijid.2016.05.013