nature immunology

Article

https://doi.org/10.1038/s41590-023-01590-2

Baseline innate and T cell populations are correlates of protection against symptomatic influenza virus infection independent of serology

Received: 16 December 2022

Accepted: 13 July 2023

Published online: 17 August 2023

Check for updates

Robert C. Mettelman[1,10], Aisha Souquette[1,10], Lee-Ann Van de Velde[1], Kasi Vegesana[1], E. Kaitlynn Allen[1], Christina M. Kackos[2], Sanja Trifkovic[2], Jennifer DeBeauchamp[2], Taylor L. Wilson[1,3], Deryn G. St. James[1,3], Smrithi S. Menon[1], Timothy Wood[4], Lauren Jelley[4], Richard J. Webby[2,11] ✉, Q. Sue Huang[4,11] ✉, Paul G. Thomas[1,11] ✉ & SHIVERS-II Investigation Team*

Evidence suggests that innate and adaptive cellular responses mediate resistance to the influenza virus and confer protection after vaccination. However, few studies have resolved the contribution of cellular responses within the context of preexisting antibody titers. Here, we measured the peripheral immune profiles of 206 vaccinated or unvaccinated adults to determine how baseline variations in the cellular and humoral immune compartments contribute independently or synergistically to the risk of developing symptomatic influenza. Protection correlated with diverse and polyfunctional $CD4^+$ and $CD8^+$ T, circulating T follicular helper, T helper type 17, myeloid dendritic and $CD16^+$ natural killer (NK) cell subsets. Conversely, increased susceptibility was predominantly attributed to nonspecific inflammatory populations, including γδ T cells and activated $CD16^-$ NK cells, as well as $TNF\alpha^+$ single-cytokine-producing $CD8^+$ T cells. Multivariate and predictive modeling indicated that cellular subsets (1) work synergistically with humoral immunity to confer protection, (2) improve model performance over demographic and serologic factors alone and (3) comprise the most important predictive covariates. Together, these results demonstrate that preinfection peripheral cell composition improves the prediction of symptomatic influenza susceptibility over vaccination, demographics or serology alone.

Influenza viruses are among the leading etiologic agents responsible for human respiratory illness, with annual infection estimates of 10–49 million and upward of 650,000 deaths globally[1,2]. Although seasonal influenza vaccinations are an important preventative measure and limit severe disease, their year-to-year effectiveness is inconsistent (ranging between 10% and 60%) (refs. 3–7). Some of this variability is attributable to incomplete matching between circulating and vaccine strains, but another possible factor is the reliance on these vaccines to generate protective humoral immunity alone. Although current seasonal influenza vaccine platforms can promote robust antibody

A full list of affiliations appears at the end of the paper. ✉e-mail: richard.webby@stjude.org; sue.huang@esr.cri.nz; paul.thomas@stjude.org

https://doi.org/10.1038/s41590-023-01590-2

titers and the T helper cell responses necessary to promote them, none are specifically designed to stimulate durable, virus-specific cellular immunity, thereby neglecting a potential ally in preventing and controlling influenza virus infection. In recent years, there have been several large-scale initiatives (for example, the Centers of Excellence for Influenza Research and Response (https://www.ceirr-network.org/) network and the Collaborative Influenza Vaccine Innovation Centers (https://www.niaidcivics.org/) program, both funded by the National Institute of Allergy and Infectious Diseases (NIAID)) to improve influenza vaccine design by increasing surveillance, improving strain risk assessment and prediction, identifying universal humoral or T cell antigens, evaluating new delivery platforms and considering demographic-specific formulations[8,9]. Driving this push is the need to understand the complete range of protective immune responses limiting symptomatic influenza illness—termed correlates of protection (CoP)—which, to date, have been largely defined by serologic measurements alone. Few studies have examined how efficacious serologic data are as predictors of protection against influenza within the context of additional immune measures, such as cellular and innate immunity; how individual cell populations contribute to protection in the presence of existing antibody responses; or if and how these compartments coordinate.

Further complicating the study of CoP against influenza viruses is the substantial degree of baseline immune variation that exists across humans, now recognized as a key determinant in predicting the efficacy of some vaccines and therapeutics and in modeling disease outcomes[10–15]. Differences in host genetics, environmental and demographic factors, and infectious exposure histories influence the composition of the baseline innate, cell-mediated immune (CMI; adaptive T cells) and humoral compartments, and may augment varied responses to influenza virus infection or vaccination[13,16–29]. Therefore, identifying baseline CoP necessitates the use of large, well-curated, influenza virus-seropositive human cohorts in which immune and demographic data are considered together. To address these challenges, the second iteration of the Southern Hemisphere Influenza Vaccine Effectiveness and Response Study (SHIVERS-II) was established to follow unvaccinated and vaccinated adults within a natural community setting in New Zealand. Using baseline serum and peripheral blood samples from 206 SHIVERS-II adult participants, we aimed to define baseline immune cell subsets correlated with protection against symptomatic influenza independently from or synergistically with humoral responses. Further, we investigated the quantitative and relative contributions of baseline cellular and humoral immune responses in mediating protective anti-influenza virus immunity. Our analytic approach incorporated high-dimensional flow cytometry data, serology measures, vaccination status and demographic data into statistical models, allowing the identification of individual baseline cell populations associated with increased risk of or protection against symptomatic influenza across vaccinated and unvaccinated adults while accounting for baseline immune and demographic variations. Using univariate analyses and multivariate partition and regression models, we demonstrate that the baseline composition of peripheral cells improves the prediction of influenza susceptibility over serology, vaccination or demographics alone. Our results underscore the complexity and variability of baseline cellular responses, support influenza vaccine design strategies targeting optimized cell subsets to induce long-lasting heterosubtypic immunity, and provide improved methods to compare vaccine efficacy.

Results

Study population

A total of 206 adult participants (Table 1) were selected from the 2018 SHIVERS-II cohort for inclusion in this study (Fig. 1a) and analyzed according to a predefined analysis pipeline (Extended Data Fig. 1). The selected participants were divided into four roughly age- and sex-matched comparator groups based on influenza virus vaccination and infection status: vaccinated–uninfected ($n = 75$), vaccinated–infected ($n = 33$), unvaccinated–uninfected ($n = 76$) and unvaccinated–infected ($n = 22$) (Fig. 1b). Participant age was comparable across sex and vaccination status, reducing the potential for confounding effects due to sample selection bias (Fig. 1c). Baseline serum and peripheral blood mononuclear cell (PBMC) samples were collected in the preseason period (March through May) for unvaccinated participants or 14 days after vaccination for vaccinees. Postseason serum samples were collected from September through January, with a majority collected in November. All participants exhibiting respiratory symptoms and meeting the World Health Organization (WHO)-defined criteria for influenza-like illness (ILI) were further tested by PCR. Participants with ILI and an associated influenza-positive PCR result were defined as having 'symptomatic' influenza. For the 2018 influenza season (May through September), 41 of 55 (74.5%) influenza virus infections were symptomatic, with a majority attributed to A(H1N1) viruses ($n = 29$, 70.7%), although infections with A(H3N2) ($n = 5$, 12.2%), B/Yamagata lineage ($n = 1$, 2.4%), B/Victoria lineage ($n = 1$, 2.4%) and untyped influenza

Table 1 | Participant demographics

Characteristic	*n*	%
Age (years)[a]	43.8±12.6	
BMI (kg m^{-2})[a]	27.5±5.6	
Influenza virus infection status (2018)		
Negative	151	73.3
Positive	55	26.7
Influenza virus vaccination status (2018)		
Unvaccinated	98	47.6
Vaccinated	108	52.4
Sex (assigned at birth)		
Female	124	60.2
Male	82	39.8
Ethnicity (participant-reported)		
Asian	14	6.8
European	171	83
Maori	12	5.8
Pacific	3	1.5
Other	2	1
Not reported	4	1.9
BMI group[b]		
Underweight (<18.5)	1	0.5
Normal weight (18.5–<25)	73	35.4
Overweight (25–<30)	59	28.6
Obese (≥30)	56	27.2
Not reported	17	8.3
Influenza virus strain		
A(H1N1)	29	14.1
A(H3N2)	5	2.4
A(untyped)	5	2.4
B/Victoria lineage	1	0.5
B/Yamagata lineage	1	0.5
Cryptic	14	6.8
Uninfected	151	73.3

[a]Data shown as mean and s.d. [b]Centers for Disease Control and Prevention (CDC)-defined BMI groupings for adults aged ≥20 years (kg m^{-2}).

A (n = 5, 12.2%) viruses were also recorded (Supplementary Table 1). Cryptic influenza cases, identified by seroconversion (fourfold or greater increase in hemagglutination inhibition (HAI) antibody titer ≥1:40) in the absence of a PCR-confirmed symptomatic influenza episode, comprised the other 14 of 55 (25.5%) infections. Influenza virus strains associated with infection were proportional between male and female participants (Fig. 1d). Supplementary Tables 1 and 2 detail the vaccination status and demographics of the participants according to influenza virus strain.

Baseline serology measures are associated with protection

Anti-influenza virus antibodies targeting hemagglutinin (HA) and neuraminidase (NA) are known CoP[30–35]. To determine the degree of preexisting humoral immunity to dominant influenza viruses circulating in the region as well as to those present in the 2018 quadrivalent influenza vaccine[36–38], we measured inhibiting antibody titers through inhibition assays (HAI or NA inhibition (NAI)) and total binding antibody titers through ELISA. The overall correlation between inhibiting and total anti-HA (Fig. 2a) or anti-NA (Fig. 2b) serology measures against homosubtypic influenza virus targets was positive, reflecting a large degree of concordance between the assays. Three correlation clusters were also identified across all serology measures (Extended Data Fig. 2a), suggesting an association between existing anti-influenza antibody responses to homo- and heterosubtypic targets. In 2018, the quadrivalent influenza vaccine was 38% effective at preventing influenza-associated hospitalizations[38]. In our study, this vaccine elicited HAI titers of ≥1:40, a purported cutoff for protection, in roughly half of the study participants sampled 14 days after immunization (Extended Data Fig. 2b). Vaccinated participants had significantly increased HAI and NAI titers against all targets compared to unvaccinated participants (Fig. 2c,d). The ELISA results showed that the total binding antibody titers were also significantly increased in vaccinees, except the titers of antibodies against B/Victoria lineage HA and A(N1), which were lower in vaccinees than in unvaccinated participants (Fig. 2e,f). Further, we compared the antibody levels in uninfected participants and participants with cryptic influenza to those in participants with symptomatic influenza. For these and other downstream analyses, uninfected and cryptic infection cases were grouped together (uninfected/cryptic group), as these cases did not meet the study criteria for symptomatic influenza and represent a more protected group. Nearly all NAI titers were significantly increased in the uninfected/cryptic group (Fig. 2h), whereas only the total anti-A(H1) and anti-B/Yamagata lineage NA binding antibody levels were significantly increased in uninfected/cryptic cases (Fig. 2i,j).

Previous studies have suggested that demographic factors can influence humoral responses, illness susceptibility and influenza vaccine effectiveness[17,39–41]. Therefore, we examined the correlations between age, sex, body mass index (BMI) and serology measures, using locally estimated scatterplot smoothing, to inform downstream statistical modeling. Age and BMI were not correlated with sex (Extended Data Fig. 2c); however, we observed significant relationships between age or BMI and several serology measures (Extended Data Fig. 2d–k). Although these results indicate that the effect of demographics on antibody levels is limited to specific targets, this variability may influence serology measures in downstream modeling and was therefore taken into account.

As higher antibody titers were observed in the uninfected/cryptic group, we used generalized logistic regression models (GLMs) to establish the risk of symptomatic influenza virus infection given individual serology measures while adjusting for demographics and influenza vaccination status. Elevated NAI titers against A(N1), A(N2), B/Victoria lineage NA and B/Yamagata lineage NA and total binding antibodies targeting A(H1) and B/Yamagata lineage HA were significantly associated with protection (odds ratio (OR) < 1) (Fig. 2k,l). In line with the elevated HAI titer observed in symptomatic influenza cases (Fig. 2g), increased risk (OR > 1) was associated with an elevated A(H3) HAI titer (Fig. 2k). This likely reflects the dominance of regional A(H3N2) virus infections in the previous year (2017) (ref. 42), skewing preexisting antibody levels away from an effective titer needed to neutralize the A(H1N1) viruses circulating during 2018 (ref. 38). Together, the serology results from SHIVERS-II are consistent with the results of numerous studies demonstrating protection mediated by anti-influenza antibodies. Further, these results suggest that the baseline levels of anti-NA inhibitory antibodies are particularly important in determining the risk of symptomatic influenza infection in this cohort.

Immune measures vary across vaccination and infection status

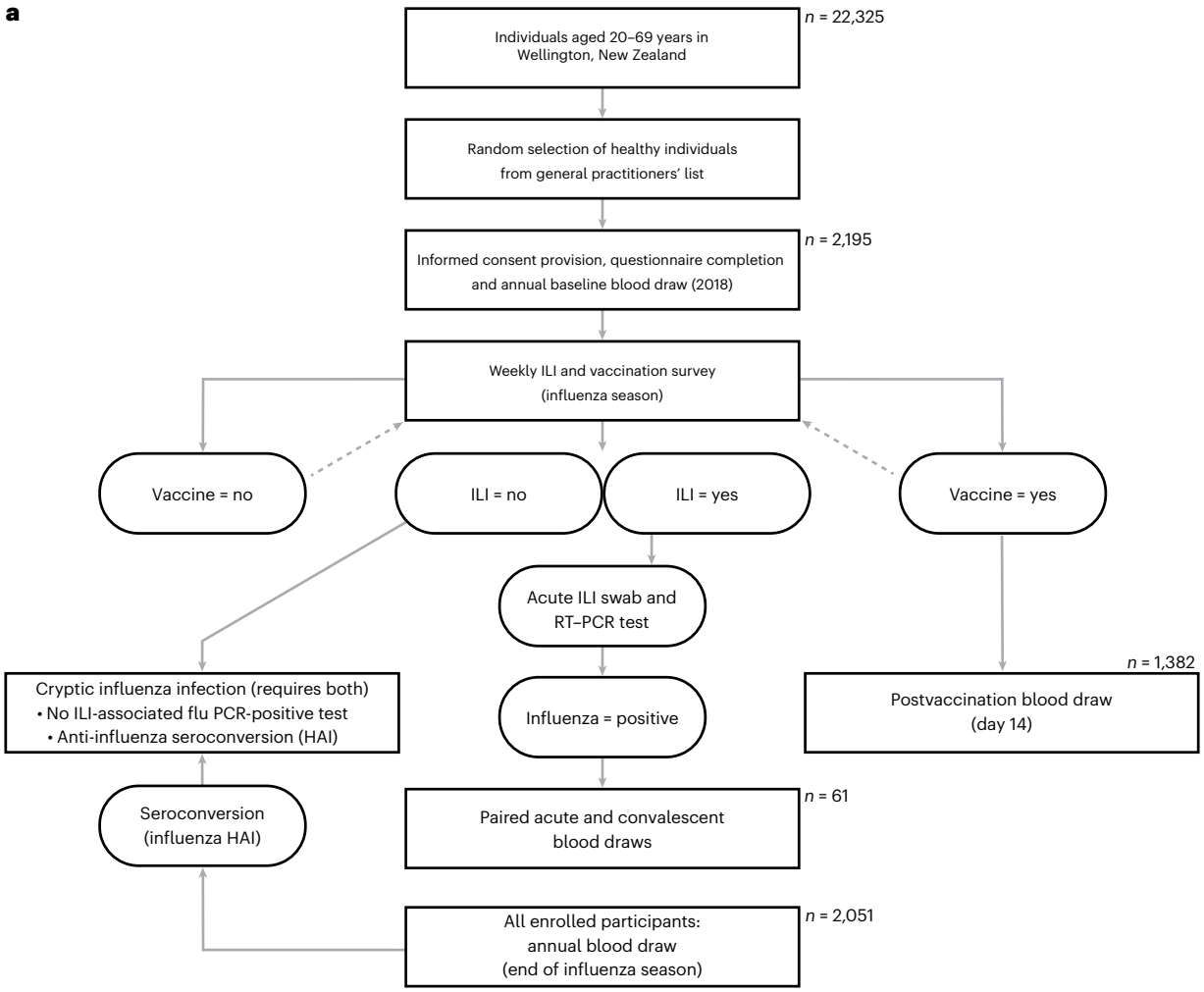
As our participants are relatively evenly divided by vaccination status, this allowed us to observe the statistical behavior of all variables separately across vaccination and infection groups. Statistical differences in individual baseline demographic, serologic and cellular covariates were evaluated between participants with and without symptomatic infection and are presented in Supplementary Table 3. Regardless of vaccination and infection status, no significant differences in demographic parameters (age, sex, BMI and ethnicity) were observed (Supplementary Table 3). Regarding serology measures, unvaccinated participants in the uninfected/cryptic infection group had higher median values of total binding antibodies to A(H1) and B/Yamagata lineage NA and NAI titers against B/Yamagata lineage, whereas the median A(H3) HAI titer was elevated in symptomatic influenza cases (Supplementary Table 3). Regardless of vaccination status, higher median NAI titers against A(N1), A(N2) and B/Victoria lineage were observed in the uninfected/cryptic group (Supplementary Table 3).

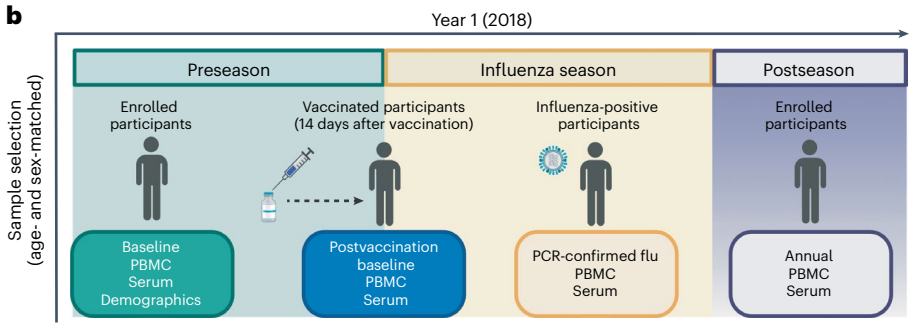
To evaluate variations in the baseline cell profile across participants, PBMCs were stained with fluorescent antibody panels distinguishing a wide array of phenotypic and functional myeloid (Extended Data Fig. 3) or lymphoid (Extended Data Fig. 4) cell subsets. To promote influenza virus-specific cytokine production for functional intracellular cytokine staining (ICS) in the lymphoid/functional panel, a portion of PBMCs from the participants were stimulated with conserved influenza virus peptide pools or live A(H1N1) or A(H3N2) viruses. Across vaccinated and unvaccinated participants, 11 cell populations had a

Fig. 1 | SHIVERS-II study design, participant enrollment, sample collection and participant demographics. a, Schematic depiction of the SHIVERS-II study design and participant numbers for year 1 (2018). RT–PCR, reverse transcription followed by PCR. **b**, Following consented enrollment, demographic information and whole blood samples were collected from all vaccinated and unvaccinated participants in the preseason period (nonvaccinated baseline) and 14 days after vaccination (vaccinated baseline). Participants meeting the WHO-defined criteria for ILI were tested for influenza viruses by PCR, and confirmed cases were sampled further during acute infection. All enrolled participants were sampled after the season. Cryptic infections were adjudicated in the postseason period from ILI- and PCR-negative participants with a fourfold or greater increase in HAI antibody titers without postvaccination HAI seroconversion. Right, 206 enrolled participants were selected for study inclusion from four baseline comparator groups (unvaccinated–uninfected, unvaccinated–infected, vaccinated–uninfected and vaccinated–infected) based on age- and sex-matching. **c**, Sex (assigned at birth) of n = 206 participants stratified by vaccination status and age (years) and compared by two-sided Wilcoxon rank-sum test (unvaccinated female (n = 58) versus unvaccinated male (n = 42), P = 0.07; vaccinated female (n = 66) versus vaccinated male (n = 42), P = 0.63). Boxes represent the median and 25th–75th percentiles; whiskers indicate the minimum (left) and maximum (right) values no further than 1.5 times the interquartile range (IQR); notches extend to 1.58 × IQR/sqrt(n), providing the 95% confidence interval (CI). P < 0.05 indicates significance. **d**, Participants' sex stratified by influenza virus infection status and strain. Influenza A (FluA) viruses include A(H1N1) and A(H3N2) strains; influenza B (FluB) viruses include the B/Victoria lineage and B/Yamagata lineage strains.

https://doi.org/10.1038/s41590-023-01590-2

significantly increased frequency in uninfected/cryptic cases (Supplementary Table 3). Conversely, 20 cell populations had a significantly increased frequency in symptomatic infection cases. Although some of these cell populations had elevated frequencies regardless of vaccination status, the overall cell profiles were distinct, suggesting an effect of immunization on protective or susceptible cellular profiles.

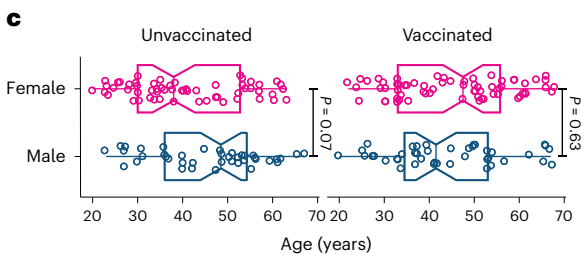


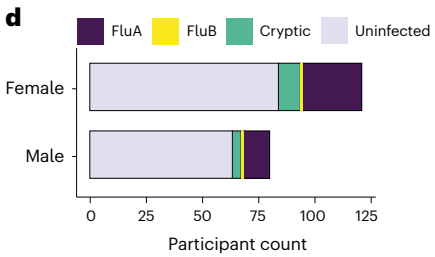


206 selected participants (this study)

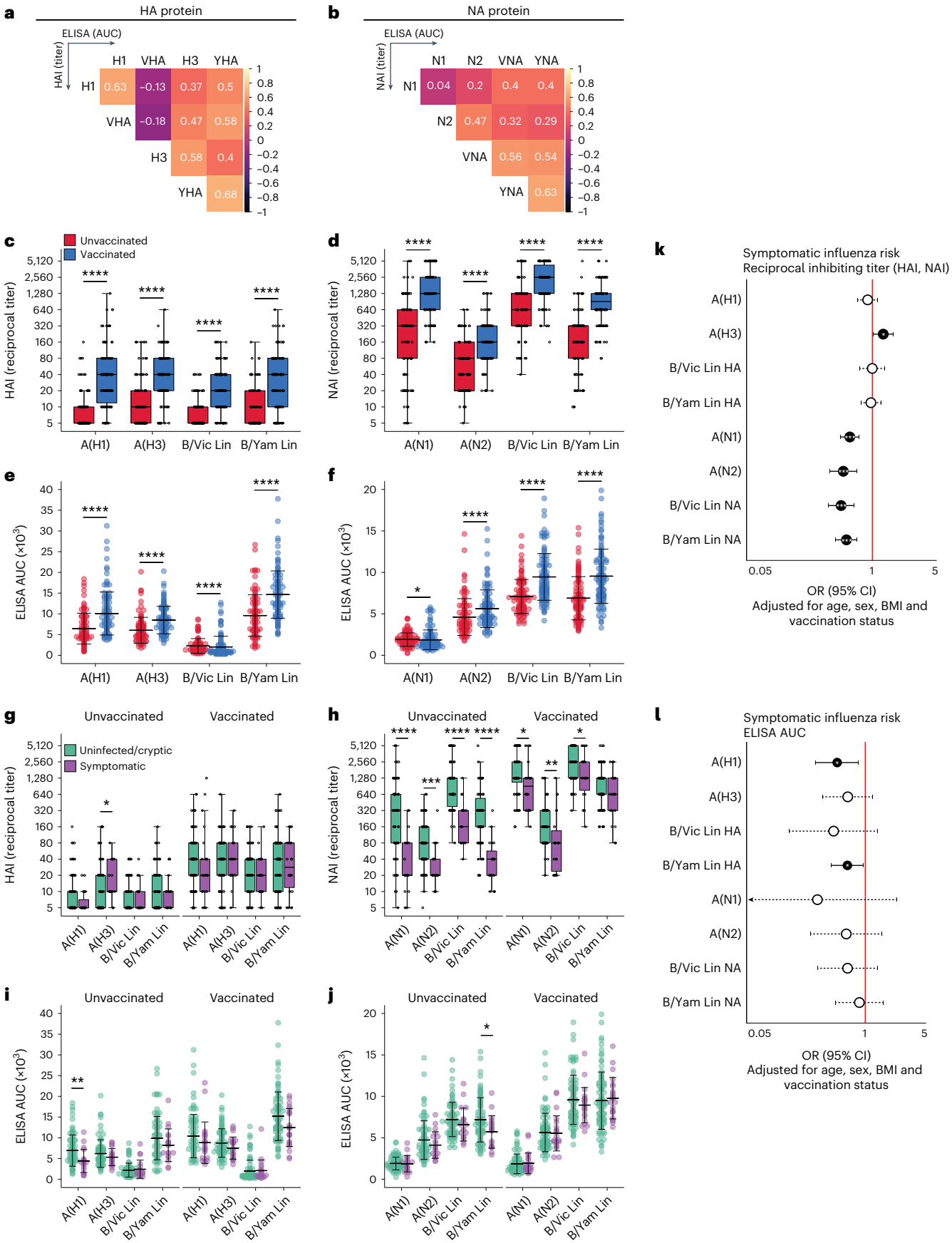
2018 influenza vaccine status		2018 influenza virus infection status		
		Negative	Positive[a]	All
Unvaccinated	*n*	76	22	98
	%	77.6	22.4	100
Vaccinated	*n*	75	33	108
	%	69.4	30.6	100
All participants	*n*	151	55	206
	%	73.3	26.7	100

[a]Includes PCR-defined (symptomatic) and serology-defined (cryptic) cases.





https://doi.org/10.1038/s41590-023-01590-2



https://doi.org/10.1038/s41590-023-01590-2

Fig. 2 | Individual serology measures correlate with protection from symptomatic influenza disease. a,**b**, Spearman rank correlations between anti-HA (HAI, ELISA) (**a**) or anti-NA (NAI, ELISA) (**b**) serology measures. VHA, B/Victoria lineage HA; YHA, B/Yamagata lineage HA; VNA, B/Victoria lineage NA; YNA, B/Yamagata lineage NA. **c**–**f**, Serology measures against HA (**c**,**e**) and NA (**d**,**f**) by influenza virus source, comparing vaccinated ($n = 108$) to unvaccinated ($n = 98$) participants by two-sided Wilcoxon rank-sum test. For **c**, A(H1) $P < 2 \times 10^{-16}$, A(H3) $P = 4.5 \times 10^{-11}$, B/Victoria lineage (B/Vic Lin) $P = 1.6 \times 10^{-15}$, B/Yamagata lineage (B/Yam Lin) $P = 4.1 \times 10^{-12}$; for **d**, A(N1) $P < 2 \times 10^{-16}$, A(N2) $P = 1.5 \times 10^{-8}$, B/Victoria lineage $P = 6 \times 10^{-12}$, B/Yamagata lineage $P = 9.4 \times 10^{-16}$; for **e**, A(H1) $P = 5.3 \times 10^{-9}$, A(H3) $P = 2 \times 10^{-9}$, B/Victoria lineage $P = 4.5 \times 10^{-5}$, B/Yamagata lineage $P = 4.8 \times 10^{-11}$; for **f**, A(N1) $P = 0.033$, A(N2) $P = 7.4 \times 10^{-5}$, B/Victoria lineage $P = 9.7 \times 10^{-11}$, B/Yamagata lineage $P = 1.1 \times 10^{-9}$. **g**–**j**, Vaccinated and unvaccinated participants' serology measures against HA (**g**,**i**) and NA (**h**,**j**) by influenza virus source, comparing uninfected/cryptic ($n = 165$) to symptomatic ($n = 41$) influenza virus infection through a two-sided Wilcoxon rank-sum test. For **g**, unvaccinated: A(H1) $P = 0.280$, A(H3) $P = 0.024$, B/Victoria lineage $P = 1.0$, B/Yamagata lineage $P = 0.922$; vaccinated: A(H1) $P = 0.32$, A(H3) $P = 0.57$, B/Victoria lineage $P = 0.33$, B/Yamagata lineage $P = 0.65$. For **h**, unvaccinated: A(N1) $P = 1.6 \times 10^{-6}$, A(N2) $P = 3.8 \times 10^{-4}$, B/Victoria lineage $P = 4.4 \times 10^{-7}$, B/Yamagata lineage $P = 2.7 \times 10^{-8}$; vaccinated: A(N1) $P = 0.012$, A(N2) $P = 0.0013$, B/Victoria lineage $P = 0.033$, B/Yamagata lineage $P = 0.212$. For **i**, unvaccinated: A(H1) $P = 0.0037$, A(H3) $P = 0.450$, B/Victoria lineage $P = 0.713$, B/Yamagata lineage $P = 0.239$; vaccinated: A(H1) $P = 0.103$, A(H3) $P = 0.155$, B/Victoria lineage $P = 0.646$, B/Yamagata lineage $P = 0.053$. For **j**, unvaccinated: A(N1) $P = 0.781$, A(N2) $P = 0.393$, B/Victoria lineage $P = 0.161$, B/Yamagata lineage $P = 0.028$; vaccinated: A(N1) $P = 0.90$, A(N2) $P = 0.95$, B/Victoria lineage $P = 0.40$, B/Yamagata lineage $P = 0.48$. Boxes represent the median and 25th–75th percentiles; whiskers indicate the minimum and maximum values no further than 1.5 times the IQR; dot plots are presented as mean ± s.d. Inhibiting antibody titers are presented as reciprocal endpoint dilutions calculated from HAI or NAI assays using A(H1N1), A(H3N2), B/Victoria lineage and B/Yamagata lineage viruses. Total binding antibody titers are reported as AUC values calculated from the results of ELISA against purified, full-length HA or NA proteins derived from influenza A(H1N1), A(H3N2), B/Victoria lineage and B/Yamagata lineage viruses. Means were compared using a two-sided Wilcoxon rank-sum test. **k**,**l**, Relative risk of symptomatic infection among all participants ($n = 206$) given individual anti-HA and anti-NA serology measures by HAI and NAI assays (**k**) or ELISA (**l**), adjusted for participant age (years), sex, BMI (kg m^{-2}) and influenza vaccination status (2018), from individual GLMs. Data are shown as ORs (circles) with 95% CIs (bars) derived from the exponential transformation of GLM estimate (logit) values. HAI and NAI reciprocal inhibition endpoint values are depicted at a 1 = $\log_2$ interval; ELISA AUC values are depicted at a 1 = 5,000 interval. Positive effects on symptomatic influenza observed with OR >1, negative effects observed with OR <1 and no effect observed at or spanning OR = 1 (red line). Individual HAI or NAI models were generated from 185 degrees of freedom (d.f.) (symptomatic $n = 35$, uninfected/cryptic $n = 151$), except A(H3) HAI and B/Victoria lineage HAI (d.f. = 184; symptomatic $n = 35$, uninfected/cryptic $n = 150$). Significance was determined from the GLM Pr(>|z|) output, where z = estimate/SE. Resulting two-tailed P values: A(H1) HAI $P = 0.368$, A(H3) HAI $P = 0.033$, B/Victoria lineage HAI $P = 0.983$, B/Yamagata lineage HAI $P = 0.775$, A(N1) NAI $P = 2.4 \times 10^{-6}$, A(N2) NAI $P = 3.0 \times 10^{-5}$, B/Victoria lineage NAI $P = 1.9 \times 10^{-6}$, B/Yamagata lineage NAI $P = 4.1 \times 10^{-6}$. Individual HA and NA ELISA models were generated from 187 d.f. (symptomatic $n = 35$, uninfected/cryptic $n = 153$). Significance was determined from the GLM Pr(>|z|) output, where z = estimate/SE. Resulting two-tailed P values: A(H1) AUC $P = 0.011$, A(H3) AUC $P = 0.172$, B/Victoria lineage HA AUC $P = 0.163$, B/Yamagata lineage HA AUC $P = 0.033$, A(N1) AUC $P = 0.232$, A(N2) AUC $P = 0.3$, B/Victoria lineage NA AUC $P = 0.252$, B/Yamagata lineage NA AUC $P = 0.63$. Not significant (blank); *$P \le 0.05$, **$P \le 0.01$, ***$P \le 0.001$, ****$P \le 0.0001$.

Predictive baseline cell profiles differ by vaccination status

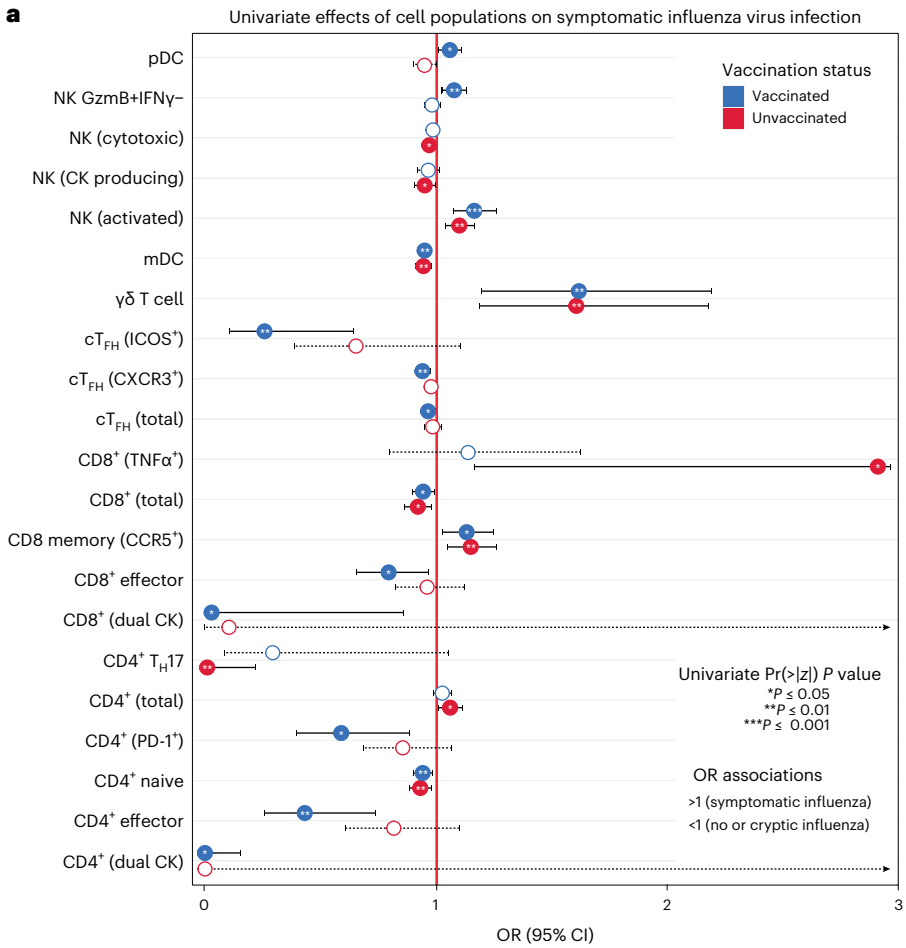
To define cellular CoP against symptomatic influenza, cell populations were then evaluated using univariate GLMs, with respect to vaccination status, to determine whether baseline cell population frequencies were associated with a change in the relative risk of acquiring symptomatic influenza infection during the 2018 season (Fig. 3a and Supplementary Table 4). We identified 14 cell populations with significant protective associations (OR < 1) within vaccinated (myeloid dendritic cells (mDCs), naïve $CD4^+$ T cells, effector $CD4^+$ and $CD8^+$ T cells, dual-cytokine (interleukin-2 (IL-2)/tumor necrosis factor-&aplha; (TNF&aplha;)/interferon-γ (IFNγ))-producing $CD8^+$ and $CD4^+$ T cells, programmed cell death protein 1 (PD-1)$^+CD4^+$ T cells, total circulating T follicular helper (cT_{FH}) cells, inducible costimulator (ICOS)$^+$ cT_{FH} cells, CXC chemokine receptor 3 (CXCR3)$^+$ cT_{FH} cells, total $CD8^+$ T cells) and unvaccinated (mDCs, cytotoxic natural killer (NK) cells, cytokine-producing NK cells, naïve $CD4^+$ T cells, $CD4^+$ T helper type 17 (T_H17) cells, total $CD8^+$ T cells) participants (Fig. 3a). We also identified seven populations that increased the risk of symptomatic influenza (OR > 1) among vaccinated (plasmacytoid dendritic cells (pDCs), granzyme B (GzmB)$^+$IFNγ$^-$ NK cells, activated NK cells, γδ T cells, CC chemokine receptor 5 (CCR5)$^+CD8^+$ memory T cells) or unvaccinated (activated NK cells, γδ T cells, CCR5$^+CD8^+$ memory T cells, total $CD4^+$ T cells, TNF&aplha;$^+CD8^+$ T cells) participants (Fig. 3a). The overall cellular CoP profile defined by the univariate analyses indicates a diversity of polyfunctional responses to influenza viruses, engaging both the adaptive and innate compartments (Fig. 3b, left). In contrast, the susceptibility-associated profile skewed toward nonspecific inflammatory populations and single-cytokine producers (Fig. 3b, right). Comparison of cell profiles from unvaccinated participants across protective and nonprotective status also indicates that the baseline composition of innate populations is an important determinant of protection from symptomatic influenza in the absence of recent vaccine-induced immunity.

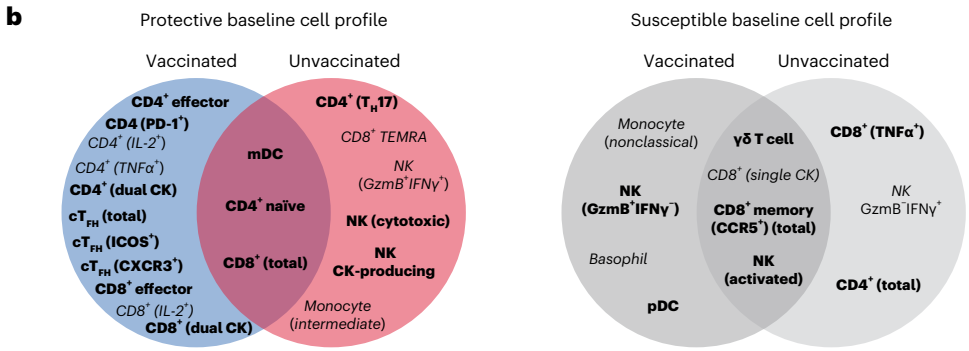
ROC thresholds define protection and susceptibility cutoffs

From the univariate analyses, we identified individual immune measures associated with protection from or increased risk of developing symptomatic influenza. To provide a quantitative value denoting

Fig. 3 | Univariate effects of cell populations on symptomatic influenza by vaccination status. a, Univariate GLM constructed to determine the individual effect of cell population frequency (% parent) on symptomatic influenza by vaccination status. Individual GLMs of unvaccinated participants were generated from 98 d.f. (uninfected/cryptic $n = 79$, symptomatic $n = 19$); GLMs of vaccinated participants were generated from 108 d.f. (uninfected/cryptic $n = 86$, symptomatic $n = 22$). GLM estimate values (logit) were transformed to ORs (circles) using the e-function and are presented with the associated 95% CIs (bars). Cell populations with positive effects on symptomatic influenza have OR >1, cell populations with negative effects have OR <1 and cell populations with no effect observed at or spanning OR = 1 (red line). Univariate significance was determined from the GLM Pr(>|z|) output (closed, significant; open, not significant), where z = estimate/SE. Significance was defined as follows: *$P \le 0.05$, **$P \le 0.01$ and ***$P \le 0.001$. Exact Pr(>|z|) P values and false discovery rate (FDR)-adjusted q values are reported in Supplementary Table 4. **b**, Baseline cell profiles associated with protection from (left) and susceptibility to (right) developing symptomatic influenza based on the 2018 influenza virus vaccination status. Participants with significant cell frequency outliers were determined by Grubbs' test ($P < 0.05$) and removed from this analysis. All cell populations presented significance by one-sided Kruskal–Wallis (italicized) or both Kruskal–Wallis and univariate GLM (bold) evaluations. Exact P values are reported in Supplementary Tables 3 and 4. **c**, Representative thresholds are depicted as cell frequency (% parent) with their associated ROC curves. A threshold defines the cell frequency at which the ROC curve sensitivity (true positive rate) equals 0.5 and represents the cutoff above which an individual factor correctly associates 50% of cases as protected or susceptible. AUC values indicate the overall quality (true versus false positives) of the individual measure in discerning protection or susceptibility. $CD4^+$, $CD4^+$ T cell; $CD8^+$, $CD8^+$ T cell; CK, cytokine; single CK, single-cytokine-producing; dual CK, dual-cytokine-producing; TEMRA, terminally differentiated effector memory.

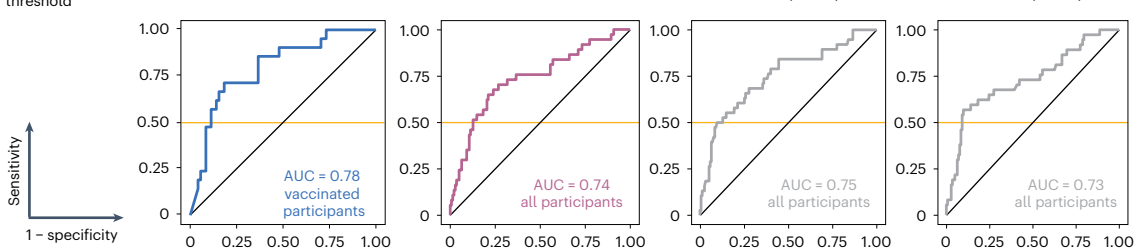
https://doi.org/10.1038/s41590-023-01590-2





c

	CD4$^+$ (dual CK)	CD4$^+$ naive	NK (activated)	γδ TCR$^+$ T cell
Threshold (% parent)	0.022%	10.65%	12.7%	2.75%
Parent population	CD3$^+$CD4$^+$CD8$^-$	CD3$^+$CD4$^+$CD8$^-$	CD56$^+$CD16$^+$CD56^{low}	CD3$^+$
Association above threshold	Protection	Protection	Susceptibility	Susceptibility



https://doi.org/10.1038/s41590-023-01590-2

protection or susceptibility, we derived 'threshold' values for each cell population (Supplementary Table 5) or serology measure (Supplementary Table 6) through receiver-operating characteristic (ROC) analysis. Thresholds define the value at which an individual parameter accurately identifies ~50% of true positive cases (sensitivity) while minimizing false positives (specificity). Therefore, a threshold is a quantifiable measure at and above which a factor accurately associates with protection or susceptibility. The area under the ROC curve (AUC) defines the quality of the immune measure as a classifier. For example, a baseline frequency of $CD4^+$ dual-cytokine producers above a threshold of 0.022% identifies at least 50% of vaccinated individuals as protected from symptomatic influenza (Fig. 3c). Conversely, a threshold frequency of ≥12.7% activated NK cells correctly identifies at least 50% of individuals as susceptible to symptomatic influenza (Fig. 3c). Comparing the protection-associated factors with the highest AUC values (Supplementary Tables 5 and 6), we observed that serology measures were more accurate in classifying cases in unvaccinated participants (for example, NAI B/Yamagata lineage AUC = 0.9, NAI B/Victoria lineage AUC = 0.86 and NAI A(N1) AUC = 0.85), whereas cellular measures were more accurate in vaccinated participants (for example, dual-cytokine-producing $CD4^+$ T cell AUC = 0.78, IL-2^+CD4^+ T cell AUC = 0.77 and $ICOS^+$ cT_{FH} cell AUC = 0.77). This is likely because the overall antibody levels are elevated in vaccinated individuals, so differences in cell populations provide improved resolution; in unvaccinated individuals, increased baseline antibody levels are more variable and therefore more accurate classifiers likely reflecting a more recent exposure to influenza antigens. However, as these are single-parameter measures, we cannot determine from the AUC values alone whether other confounding factors affect these classifications.

Cryptic influenza cases are associated with unique cell profiles

Susceptibility to symptomatic influenza encompasses risks associated with (1) virus exposure and (2) symptom development following exposure. Outside of regional prevalence estimates, it was difficult to determine whether uninfected study participants were indeed exposed to influenza viruses during the 2018 season. Therefore, we analyzed a subset of participants with confirmed exposures: symptomatic influenza cases (confirmed by PCR, $n = 41$) and cryptic infection cases (confirmed by seroconversion, $n = 14$). Univariate regression analyses demonstrated that elevated baseline levels of several anti-HA or anti-NA inhibitory or total binding antibody measures were significantly associated with reduced risk of symptomatic influenza, whereas demographic factors and recent influenza vaccination history were not (Fig. 4a). We then compared cell frequencies between symptomatic and cryptic cases. In concordance with our univariate results, the frequencies of both activated and $GzmB^-IFN\gamma^+$ NK cells were increased at baseline in symptomatic cases, further suggesting a role in susceptibility (Fig. 4b). Conversely, we identified 15 cell populations with increased baseline frequencies in cryptic infection cases, comprising a unique set of protective responses, including several cell populations (conventional DC type 2 (cDC2) and IL-21^+ cT_{FH} cells) not identified in the prior univariate analysis (Fig. 4c). Similar to the univariate protective profile, the baseline cell populations with increased frequencies in cryptic cases were diverse and polyfunctional, reflecting adaptive T cell activation and cytokine production ($CD4^+$ effector T cells, IL-17^+CD4^+ T cells, IL-2^+CD4^+ T cells, IL-2^+CD8^+ T cells, $IFN\gamma^+CD8^+$ T cells, dual-cytokine-producing $CD8^+$ T cells), humoral engagement (cT_{FH} subsets) and innate immune activities (cytotoxic and cytokine-producing NK cells, mDCs).

Immune cell populations cluster into co-regulated modules

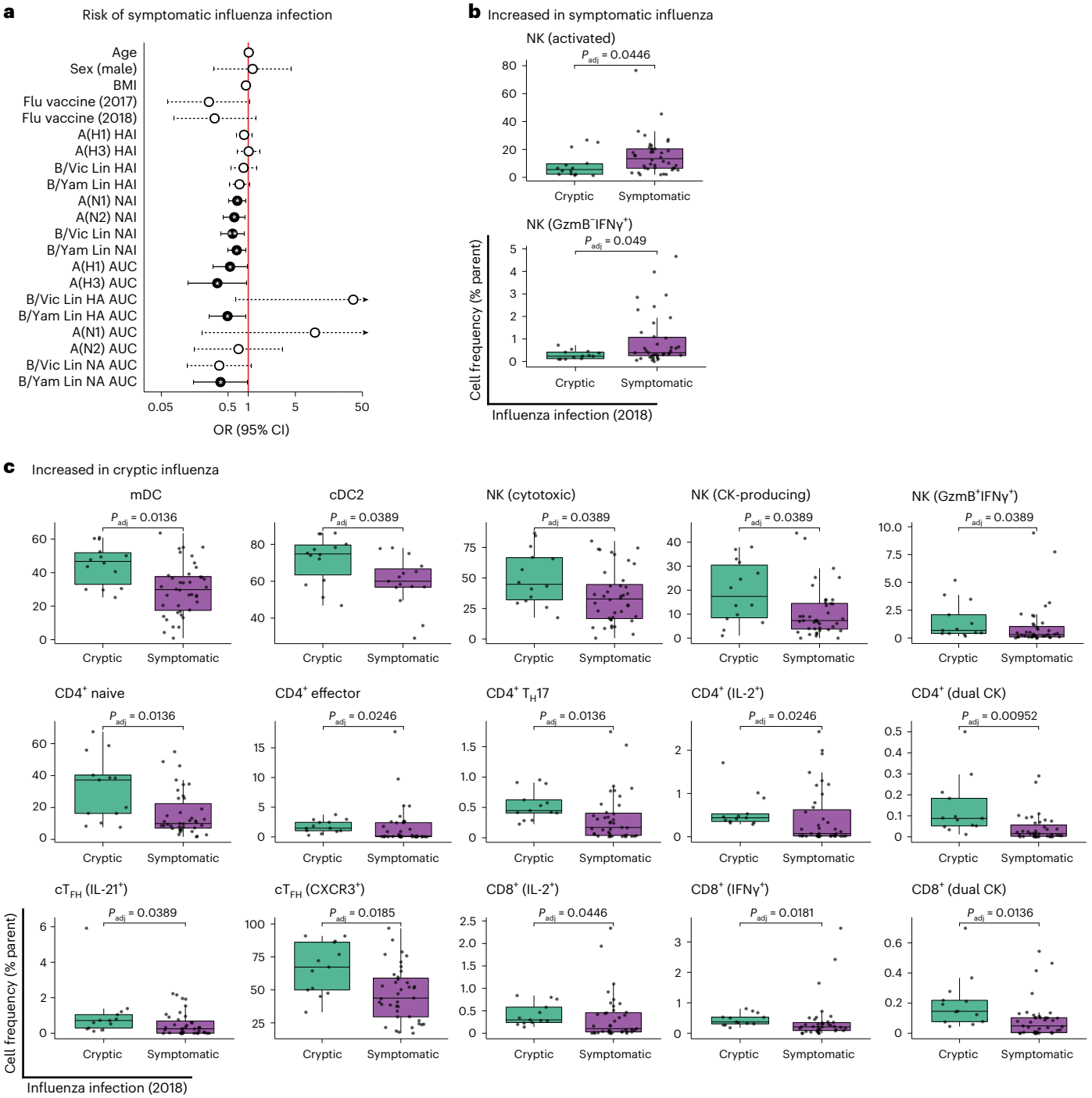
The above univariate analyses consider cell populations independently and do not account for coordinated responses often observed across the adaptive and innate immune compartments. We sought to understand the cell-to-cell relationships within our dataset and determine which populations represent co-regulated immune cell modules, defined here as cell populations with a strong positive correlation in cell frequencies suggestive of parallel responses. Using Pearson's bivariate correlation, we identified 9 myeloid (Fig. 5a) and 13 lymphoid/functional (Fig. 5b) cell modules across all study participants. We also evaluated immune cell correlations separately in vaccinated and unvaccinated participants (Extended Data Fig. 5) and found them to largely reflect the same correlation groups identified from the complete set of participants. Correlation results from the complete set of participants were used for downstream analyses. The cell frequencies from individual populations within a given module were averaged to provide the final 'cluster' frequency values, which were compared according to influenza virus infection status. Clusters with increased average cell frequencies in both influenza-negative and cryptic infection cases included myeloid clusters 1 (mDCs, $CD14^+CD16^+$ intermediate monocytes), 3 (cytotoxic and cytokine-producing NK cells, neutrophils) and 8 (cDC2) (Fig. 5c), as well as lymphoid/functional clusters 4 (TNF&aplha;$^+$, $IFN\gamma^+$, IL-17^+ and dual-cytokine (IL-2/TNF&aplha;/IFNγ)-producing $CD4^+$ T cells; $IFN\gamma^+$ and dual-cytokine (IL-2/TNF&aplha;/IFNγ)-producing $CD8^+$ T cells), 8 ($CD4^+$ and $CD8^+$ effector T cells), 9 (PD-1^+CD4^+ and PD-1^+CD8^+ T cells, PD-1^+ and $CXCR3^+$ cT_{FH} cells) and 13 (total $CD8^+$ T cells, $GzmB^+CD4^+$ and $GzmB^+CD8^+$ T cells, γδ T cells) (Fig. 5d). Conversely, the cell frequency of myeloid cluster 7 (activated NK cells) was found to be increased in symptomatic cases (Fig. 5c), supporting the previous observations made in the univariate GLM analysis. Together, these results reflect the involvement of coordinated cellular responses in determining influenza virus susceptibility in cases of confirmed exposure.

Fig. 4 | Cryptic infections are associated with unique cellular responses. **a**, Relative risk of symptomatic infection among influenza virus-infected participants given demographic, serologic or vaccination history covariates. GLM estimate values (logit) were transformed to ORs (circles) using the e-function and are presented with the associated 95% CIs (bars). Positive effects on symptomatic influenza observed with OR >1, negative effects observed with OR <1, and no effect observed at or spanning OR = 1 (red line). Univariate GLMs were generated from 54 d.f. (symptomatic $n = 41$, cryptic $n = 14$), except BMI (d.f. = 46; symptomatic $n = 35$, cryptic $n = 12$). Significance was determined from the GLM Pr(>|z|) output, where z = estimate/SE. Resulting two-tailed P values: age $P = 0.528$, sex (male) $P = 0.827$, BMI $P = 0.24$, flu vaccine (2017) $P = 0.062$, flu vaccine (2018) $P = 0.111$, A(H1) HAI $P = 0.321$, A(H3) HAI $P = 0.945$, B/Victoria lineage HAI $P = 0.495$, B/Yamagata lineage HAI $P = 0.091$, A(N1) NAI $P = 0.014$, A(N2) NAI $P = 0.015$, B/Victoria lineage NAI $P = 0.01$, B/Yamagata lineage NAI $P = 0.014$, A(H1) AUC $P = 0.041$, A(H3) AUC $P = 0.041$, B/Victoria lineage HA AUC $P = 0.082$, B/Yamagata lineage HA AUC $P = 0.026$, A(N1) AUC $P = 0.249$, A(N2) AUC $P = 0.666$, B/Victoria lineage NA AUC $P = 0.076$, B/Yamagata lineage NA AUC $P = 0.045$. Not significant (blank); $^*P \le 0.05$, $^{**}P \le 0.01$. **b**,**c**, Comparison of individual cell population frequencies (% parent) across participants with symptomatic (**b**) and cryptic (**c**) influenza. Cell populations are grouped based on increased frequencies in symptomatic or cryptic influenza cases. In **b**, NK (activated) (cryptic $n = 14$ versus symptomatic $n = 40$); NK ($GzmB^-IFN\gamma^+$) (cryptic $n = 13$ versus symptomatic $n = 39$). In **c**, mDC, NK (cytotoxic), NK (CK-producing) (cryptic $n = 14$ versus symptomatic $n = 40$); cDC2 (cryptic $n = 14$ versus symptomatic $n = 15$); NK ($GzmB^+IFN\gamma^+$), $CD4^+$ naïve, $CD4^+$ effector, $CD4^+$ T_H17, $CD4^+$ (IL-2^+), $CD4^+$ (dual CK), cT_{FH} (IL-21^+), cT_{FH} ($CXCR3^+$), $CD8^+$ (IL-2^+), $CD8^+$ ($IFN\gamma^+$), $CD8^+$ (dual CK) (cryptic $n = 13$ versus symptomatic $n = 39$). Lymphoid/functional panel frequencies represent the average frequency (% parent) across the virus (multiplicity of infection (MOI) = 4, A/Michigan/45/2015 (H1N1)pdm09 or A/Singapore/INFIMH-16-019/2016 (H3N2)) and peptide (peptide pools containing matrix 1 protein (M1), nucleoprotein (NP) and polymerase basic 1 protein (PB1), 1–5 μM per peptide) stimulation groups. Boxes represent the median and 25th–75th percentiles; whiskers indicate the minimum and maximum values no further than 1.5 times the IQR. Means were compared using a two-sided Wilcoxon rank-sum test with FDR adjustment, with $q \le 0.01$ indicating significance.

https://doi.org/10.1038/s41590-023-01590-2

Immune cells most accurately categorize influenza cases

By defining the univariate cellular CoP and susceptibility through regression modeling, we have shown the individual contributions of baseline innate and adaptive T cell populations in mitigating or enhancing symptomatic influenza. We next sought to provide additional context to these cellular responses by determining which covariates, spanning the demographic, vaccine, serology and immune cell categories, would best predict future symptomatic infections during the 2018 influenza season. To do this, we built, trained and compared random forest models from an 80:20 (train to test) split of 200 participants, ensuring equivalent proportions of cases (symptomatic influenza) and controls (uninfected, cryptic influenza). Four random forest models were built: (1) base model predicting symptomatic infection from demographic, 2018 influenza vaccination status and serology covariates; (2) lymphoid model comprising lymphoid/functional panel cell frequencies and all base model variables; (3) myeloid model comprising myeloid panel cell frequencies and all base model variables; and (4) combined model comprising all available variables. Model performance was measured by sensitivity and specificity metrics and by scoring data from the testing set in ROC curves (Fig. 6a). We observed that both the base and myeloid models categorized participants with 79% accuracy (Fig. 6a, bottom right), whereas the lymphoid model increased sensitivity and specificity and substantially improved categorization accuracy to 86%, demonstrating the influence of lymphoid populations in categorizing influenza cases. As the combined model (84% accuracy) did not improve accuracy over the lymphoid model, we examined the



https://doi.org/10.1038/s41590-023-01590-2

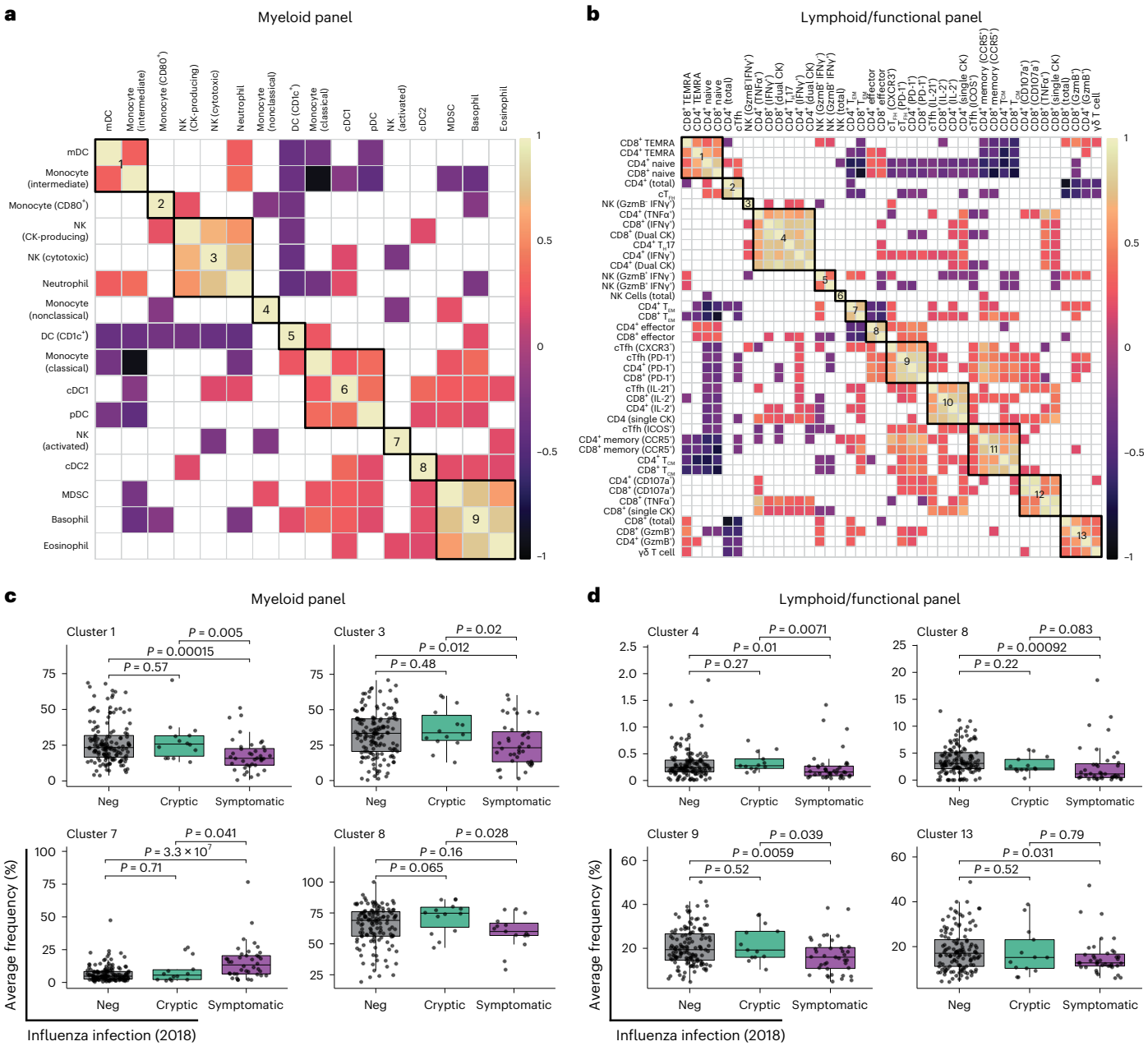


Fig. 5 | Co-regulated CMI and innate immune cell modules. a,**b**, Co-regulated cell modules ('clusters') determined by the average frequency (% parent) of individual cell populations from myeloid (**a**) or lymphoid/functional (**b**) compartments with a significant positive correlation (Pearson's bivariate correlation). Lymphoid/functional cell population frequencies were calculated as the average frequency across ICS stimulation conditions including peptide pools (M1, NP and PB1) and viruses (A(H1N1), A(H3N2)). *P* values were adjusted using FDR (*q*) correction for multiple comparisons. Significance was defined at $q \leq 0.05$ (filled squares); not significant (blank squares). **c**,**d**, Mean cell cluster frequencies from myeloid (**c**) or lymphoid/functional (**d**) correlation matrixes compared according to influenza virus infection status (myeloid clusters 1, 3 and 7: uninfected (Neg) $n = 145$ versus cryptic $n = 14$ versus symptomatic $n = 40$; myeloid cluster 8: uninfected $n = 129$ versus cryptic $n = 14$ versus symptomatic $n = 15$; lymphoid/functional clusters 4, 8, 9 and 13: uninfected $n = 133$ versus cryptic $n = 13$ versus symptomatic $n = 39$) through a two-sided Wilcoxon rank-sum test. Boxes represent the median and 25th–75th percentiles; whiskers indicate the minimum and maximum values no further than 1.5 times the IQR. T_{EM}, effector memory T cell; T_{CM}, central memory T cell.

relative contributions of lymphoid and myeloid populations in model performance. Three additional random forest models—myeloid-only, lymphoid-only and lymphoid + myeloid—were also built, trained and compared (Extended Data Fig. 6). The lymphoid-only model (79%) performed better than the myeloid-only model (72%) and was as accurate as the base model (79%) in classifying symptomatic influenza cases. The lymphoid + myeloid model also exhibited higher accuracy (81%) than the base model. These data indicate that lymphoid populations improve overall accuracy in classifying symptomatic influenza cases compared to myeloid populations, serology, vaccination status and demographics alone. Using variable importance (VIP) analysis of the combined model, we then derived 'importance' values representing the individual effect strength on the dependent variable. Strikingly, the top 4 (and 25 of the top 30) most important covariates used in categorizing symptomatic and uninfected/cryptic influenza cases were of cellular origin, with $ICOS^+$ cT_{FH} cells having the highest importance (Fig. 6b). Indeed, the frequency of $ICOS^+$ cT_{FH} cells was significantly higher in uninfected/cryptic influenza cases across both unvaccinated

https://doi.org/10.1038/s41590-023-01590-2

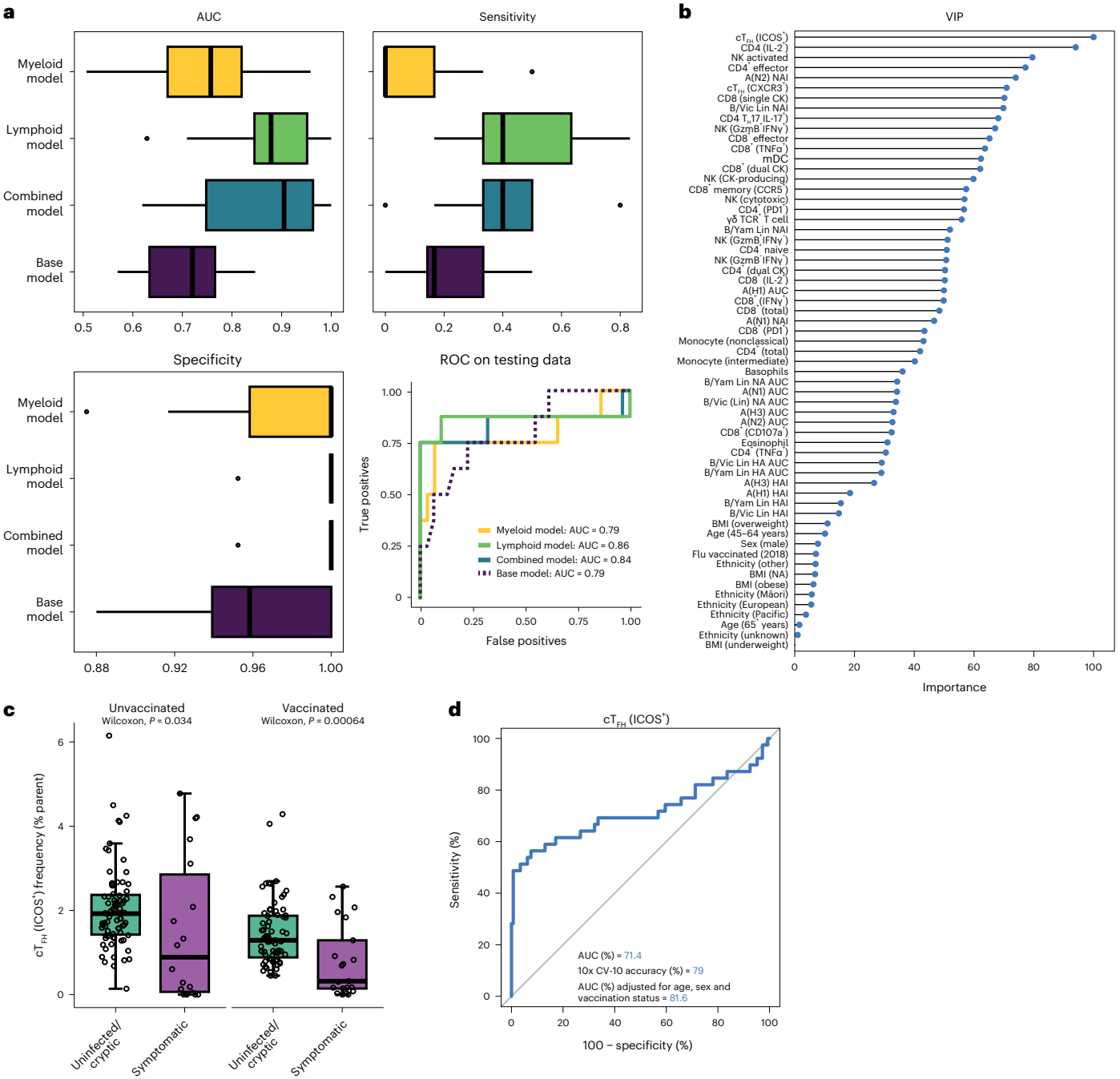
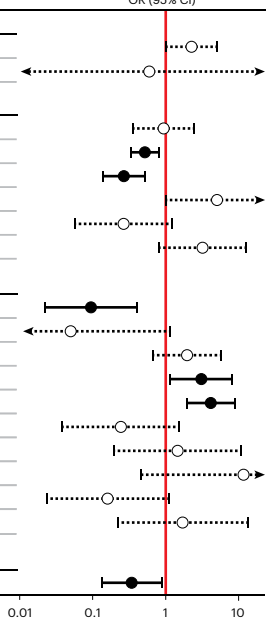


Fig. 6 | Decision tree model comparison. a, Comparison of the base (demographic factors + serology + vaccination status), myeloid (base + myeloid panel cell populations), lymphoid (base + lymphoid/functional panel cell populations) and combined (base + myeloid + lymphoid) random forest models built to categorize symptomatic and uninfected/cryptic influenza. Participants were split 80:20 into a training set (symptomatic cases $n = 31$, uninfected/cryptic controls $n = 128$) and a testing set (symptomatic cases $n = 8$, uninfected/cryptic controls $n = 33$), ensuring equal proportions of cases and controls. Models were trained, tested and cross-validated using ten times cross-validation (10× CV-10). Sensitivity (true positive rate), specificity (false positive rate), and AUC metrics and an out-of-sample evaluation of the models (bottom right) are provided. Boxes represent the median and 25th–75th percentiles; whiskers indicate the minimum and maximum values no further than 1.5 times the IQR. **b**, Relative importance of each baseline covariate in the combined random forest model; baseline covariates with high importance best categorize symptomatic and uninfected/cryptic influenza cases. Participants with significant cell frequency outliers were determined by Grubbs' test ($P < 0.05$) and removed from this analysis. **c**, Mean ICOS$^+$ cT$_{FH}$ cell frequency comparison between uninfected/cryptic and symptomatic influenza infection according to the 2018 influenza vaccination status through a two-sided Wilcoxon rank-sum test (unvaccinated: symptomatic $n = 18$ versus uninfected/cryptic $n = 75$; vaccinated: symptomatic $n = 21$ versus uninfected/cryptic $n = 71$). Boxes represent the median and 25th–75th percentiles; whiskers indicate the minimum and maximum values no further than 1.5 times the IQR. **d**, ROC curve generated from the GLM that predicts symptomatic influenza infection from the ICOS$^+$ cT$_{FH}$ cell frequency. AUC values were from the ICOS$^+$ cT$_{FH}$ GLM as a univariate model, following 10× CV-10 and after adjustment for age (years), sex and 2018 influenza vaccination status.

https://doi.org/10.1038/s41590-023-01590-2

a

	OR (95% CI)	*P* value	1 unit	Association	Cell module
Demographics and vaccination status					
Age (scaled)	2.26 (1–5.14)	0.051	12.7		
Vaccinated (2018 influenza vaccine)	0.6 (0–181.34)	0.859			
Serologic factors					
A(H3) HAI (log_2)	0.94 [0.35–2.52)	0.902			
A(N1) NAI (log_2)	0.52 (0.32–0.83)	0.007		Protection	
B/Vic lin NAI (log_2)	0.27 (0.13–0.54)	>0.001		Protection	
A(N1) total (scaled)	5.06 (0.99–25.96)	0.052	1,036.7		
B/Vic lin HA total (scaled)	0.27 (0.06–1.27)	0.097	2,245.0		
Vaccinated: A(H3) HAI (scaled)	3.2 (0.8–12.86)	0.100			
Lymphoid compartment					
$CD4^+$ ($CD107a^+$) (scaled)	0.09 (0.02–0.42)	0.002	4.2	Protection	L-2
$CD4^+$ T_H17 (scaled)	0.05 (0–1.18)	0.064	0.4		L-4
cT_{FH} ($ICOS^+$) (scaled)	1.97 (0.65–5.9)	0.229	1.1		L-12
$CD8^+$ ($TNF\alpha^+$) (scaled)	3.09 (1.13–8.45)	0.028	1.0	Susceptibility	L-3
γδ TCR^+ T cell (scaled)	4.15 (1.89–9.13)	>0.001	1.7	Susceptibility	L-13
NK ($GzmB^-IFN\gamma^+$) (scaled)	0.24 (0.04–1.58)	0.139	4.7		L-7
NK ($GzmB^+IFN\gamma^+$) (scaled)	1.46 (0.19–11.19)	0.716	1.0		L-11
Vaccinated: $CD4^+$ T_H17	11.68 (0.44–306.51)	0.141			
Vaccinated: cT_{FH} ($ICOS^+$)	0.16 (0.02–1.13)	0.066			
Vaccinated:NK ($GzmB^+IFN\gamma^+$) (scaled)	1.71 (0.21–13.58)	0.614			
Myeloid compartment					
Monocyte (intermediate) (scaled)	0.34 (0.13–0.9)	0.030	12.7	Protection	M-8



b

Lymphoid panel

Cell module (L)	Cell type
1	$CD4^+$ (single CK)
	$CD4^+$ (IL-2^+)
	$CD8^+$ (IL-2^+)
	cT_{FH} (IL-21^+)
2	$CD4^+$ ($CD107a^+$)
	$CD8^+$ ($CD107a^+$)
3	$CD8^+$ (single CK)
	$CD8^+$ ($TNF\alpha^+$)
4	$CD4^+$ ($TNF\alpha^+$)
	$CD4^+$ (dual CK)
	$CD8^+$ (dual CK)
	$CD4^+$ ($IFN\gamma^+$)
	$CD8^+$ ($IFN\gamma^+$)
	$CD4^+$ T_H17
5	$CD4^+$ effector
	$CD8^+$ effector
6	$CD4^+$ memory ($CCR5^-$)
	cT_{FH} ($CXCR3^+$)
	$CD4^+$ ($PD1^+$)
	$CD8^+$ ($PD1^+$)
	cT_{FH} ($PD1^+$)
	$CD8^+$ memory ($CCR5^-$)
7	NK ($GzmB^-IFN\gamma^+$)
8	$CD8^+$ naive
	$CD4^+$ naive
	$CD4^+$ Tcm
	$CD4^+$ Tem
	$CD4^+$ TEMRA
	$CD8^+$ T_{CM}
	$CD8^+$ T_{EM}
	$CD8^+$ TEMRA
9	NK (total)
10	NK ($GzmB^+IFN\gamma^-$)
11	NK ($GzmB+IFN\gamma^+$)
12	cT_{FH} ($ICOS^+$)
13	γδ T cell
14	$CD4^+$ ($GzmB^+$)
15	$CD8^+$ ($GzmB^+$)
16	cT_{FH} $CD4^+$ (total)
	$CD8^+$ (total)

Myeloid panel

Cell module (M)	Cell type
1	MDSC
	Basophil
	Eosinophil
2	Monocyte ($CD80^+$)
3	Monocyte (nonclassical)
4	NK (activated)
5	cDC1
6	cDC2
7	pDC
8	Monocyte (classical)
	Monocyte (intermediate)
9	mDC
10	DC ($CD1c^+$)
11	Neutrophil
	NK (CK-producing)
	NK (cytotoxic)

Fig. 7 | Baseline predictors of influenza virus infection susceptibility, accounting for demographic, vaccination, serology and cellular covariates. **a**,**b**, A multivariate GLM that predicts symptomatic influenza virus infection, accounting for baseline demographic, vaccination, serology, myeloid and lymphoid covariates, was constructed. **a**, Final logistic regression model, with covariates by category, predicting the risk of symptomatic influenza (symptomatic $n = 39$ and uninfected/cryptic $n = 161$ participants). For covariate selection in the final model, baseline covariates with strong multicollinearity by VIF assessment were excluded; single representative cell populations from each co-regulated cell cluster were included. Models comprising these covariates were compared by stepwise AIC (forward and reverse) and BMA to account for inherent model uncertainty arising from the variable selection. Interaction terms are denoted with a colon (:). OR (95% CI) values were derived from the exponential transformation of the GLM estimate (logit) value; two-tailed *P*-value significance was determined from the GLM Pr(>|*z*|) output (closed, significant; open, not significant), where z = estimate/SE. Positive effects on symptomatic influenza observed with OR >1, negative effects observed with OR <1 and no effect observed at or spanning OR = 1 (red line). Variables were scaled as indicated based on their median and s.d. values. The value for 1 unit of change is indicated for each scaled value. Units for scaled values are years (age), AUC (total antibody ELISA measures), log_2(reciprocal endpoint titer) (HAI and NAI) and % parent frequency (cell populations). For example, for every increase of 12.7 years in age, there is a 2.26 increase in the odds of influenza infection, suggesting that younger individuals are better protected. The association of significant baseline covariate predictors with symptomatic (susceptibility) or uninfected/cryptic (protection) influenza is indicated. Participants with significant cell frequency outliers were determined by Grubbs' test ($P < 0.05$) and removed from this analysis. **b**, Cell modules comprising cell populations from the lymphoid/functional or myeloid panel with strong absolute value correlations. The 'Cell module' column in **a** corresponds to strongly correlated cell modules clustered in **b** based on the absolute value of correlation. MDSC, myeloid-derived suppressor cells.

and vaccinated participants (Fig. 6c) and had a predictive accuracy of 81.6% following cross-validation and adjustment for age, sex and vaccination status (Fig. 6d). Together, the results from the partition analyses emphasize the strong contributions of both baseline innate and adaptive cellular immunity in regulating responses to influenza virus and mitigating symptomatic disease.

Baseline immune measures predict influenza susceptibility

Although partition analysis helps determine which combined or individual baseline variables best categorize symptomatic and uninfected/cryptic cases, these models are limited in risk assessment. Indeed, the random forest and VIP analyses here do not ascribe effect directionality nor associate any factor with increased or decreased risk of symptomatic influenza. Therefore, we used GLMs to assess influenza risk given all measures of immune and demographic variation present at baseline. A multivariate GLM was constructed to identify the differential effects of cell populations, anti-influenza virus antibodies, vaccination status and demographics on determining the overall risk of symptomatic influenza (Fig. 7). GLMs with highly correlated factors (collinearity) often produce unreliable coefficients with high standard error (SE) values[43]. Owing to the high dimensionality and observed correlation between the cellular and serology immune measures in our data, we took several steps to reduce variable collinearity (Pearson's correlation and variance inflation factor (VIF)) and optimized selected covariates by comparing stepwise models through Akaike information criterion (AIC) and Bayesian model averaging (BMA) (Extended Data Fig. 1). The final multivariate GLM was constructed using the set of selected covariates against a response variable denoting symptomatic infections (Fig. 7a). The model comprises 19 covariates, including 4 ($CD107a^+CD4^+$ T cells, intermediate monocytes, A(N1) NAI and B/Victoria lineage NAI) that are significantly associated with protection and 2 (TNF&aplha;$^+CD8^+$ T cells and γδ T cell receptor (TCR)$^+$ T cells) that significantly predict increased symptomatic influenza risk. Using $CD107a^+CD4^+$ T cells as an example, the GLM readout shows that for each 4.2% increase in $CD107a^+CD4^+$ T cell frequency, the odds associated with developing symptomatic infection decrease by 11.1-fold (1/0.09). Although neither age ($P = 0.051$) nor 2018 vaccination status ($P = 0.859$) had significant effects on risk, these variables were included in the GLM to account for variability across participants. Owing to the strong correlations observed among cell populations, we can further infer protection or susceptibility associations by identifying the myeloid or lymphoid modules within which these populations reside (Fig. 7b) and by referencing the univariate associations for a given cell population (Fig. 3b). Together, the final GLM results indicate that baseline serologic

and cellular factors, when considered together, synergistically predict influenza virus infection outcomes and contribute to the relative risk of developing influenza disease. Results from this study support recent calls to target adaptive T cell populations in next-generation vaccine designs[9] and strongly suggest that both cellular and serology measures are necessary to fully assess the efficacy of influenza vaccines.

Discussion

Immune responses to influenza viruses are complex and include coordinated innate, CMI and serologic responses to clear active infections; build durable immune memory; and rapidly neutralize subsequent challenges. Equally nuanced are the determinants of risk and susceptibility to symptomatic influenza, which comprise individual comorbidities and demographic risk factors, vaccination and infection histories, and variations in baseline immune profiles. In our study, we investigated which baseline immune cell populations affect the risk of developing symptomatic influenza and examined how differences in preexisting anti-influenza virus antibody titers, influenza vaccination status and participant-level demographics influence these associations.

Our univariate analyses demonstrated that protection from symptomatic influenza correlates with increased frequencies of diverse and polyfunctional influenza virus-specific $CD4^+$ and $CD8^+$ T cells; cells associated with the engagement of humoral responses, including cT_{FH} cells, mDCs and T_H17 cells; and innate immune effector $CD16^+$ cytotoxic and cytokine-producing NK cells, many of which have been observed in other studies. As expected, protection profiles were distinct between vaccinated and unvaccinated adults. The vaccine-associated protective profile favored effector and polyfunctional adaptive populations, pointing toward influenza virus-specific T cell activation. Increased frequencies of antigen-presenting mDCs and cT_{FH} populations in vaccinees are also suggestive of humoral crosstalk, discussed in detail below. In unvaccinated participants, protection was associated with increased baseline frequencies of populations exhibiting analogous, albeit antigen-agnostic, functions to those observed following vaccination. For example, increased frequencies of T_H17 cells, which were protective in our study, have been shown to be involved in anti-influenza virus responses either directly[44], through B cell engagement promoting immunoglobulin G (IgG), IgA and IgM[45,46], or by recruiting preexisting B cells into the lung without vaccine priming[46–48]. Therefore, T_H17 cells may compensate for the absence of other vaccine-induced $CD8^+$ or $CD4^+$ T helper subsets. Cytokine-producing and cytotoxic NK cells exhibiting virus-limiting potential similar to $CD8^+$ T cells were also strongly protective in unvaccinated participants, supporting studies demonstrating the active roles of NK cells in anti-influenza virus responses and in limiting disease severity[49–56]. NK cells, however, may also be a double-edged sword. One group reported a link between $IFN\gamma^+$ NK cells and extrapulmonary inflammation leading to poor outcomes in pregnant women with influenza virus infection[57]. Coinciding with these findings, we also observed increased risk of symptomatic influenza in participants with increased baseline frequencies of activated NK cells ($CD56^{low}HLA\text{-}DR^{hi}CD16^-$), suggesting that NK cells have pleiotropic roles during influenza virus infection.

A main takeaway from this study is the importance of cell–humoral crosstalk promoting increased baseline levels of anti-influenza virus antibodies. We report that elevated baseline NAI titers are the best serology predictors of symptomatic influenza and are associated with protection in this cohort, supporting other recent findings[32–34,58–60]. However, we also found that the addition of lymphoid cell population composition improves the prediction of symptomatic influenza cases, arguing that baseline antibody levels cannot be considered alone. Indeed, sustained B–T cell crosstalk within germinal centers, as well as cT_{FH} cells in peripheral blood, is associated with productive anti-influenza virus humoral response[61–65]. In our study, evidence of this communication is supported by the association of $CD4^+$ populations, including cT_{FH} subsets, with protection. In fact, $ICOS^+$ cT_{FH} cells represent the most important variable in determining the accuracy of the combined random forest model. From our analysis of co-regulated cell modules, we found that $ICOS^+$ cT_{FH} cells correlate with $CD4^+$ and $CD8^+$ memory T cell populations (lymphoid/functional cluster 11), linking increased $ICOS^+$ cT_{FH} cell frequency to long-term memory. We also identified correlations across T_H17, $IFN\gamma^+$, and polyfunctional $CD4^+$ and $CD8^+$ T cells (lymphoid/functional cluster 4), which are known to improve outcomes during influenza virus infection[66]. Further, our multivariate regression modeling identified that reduced risk of symptomatic influenza is independently associated with $CD107a^+CD4^+$ T cells, $CD14^+CD16^+$ intermediate monocytes, and anti-influenza virus NAI titers targeting A(H1N1) and B/Victoria lineage viruses. CD107a is a marker of degranulation[67] and implies that these $CD4^+$ T cells were specifically activated by influenza virus antigens during ICS and flow cytometry. That these populations are observed at baseline indicates that cT_{FH} and other influenza-responsive T cells persist following influenza antigen exposure in prior seasons and may confer 'carryover' protection in a subset of participants[68]. However, as our study was limited to immune measures present in circulation, we did not evaluate contributions from mucosa-associated cells, secretory IgA, or IgG produced by respiratory-resident memory B cells, which are likely CoP (reviewed in refs. 22,69). As $ICOS^+$ cT_{FH} and other functional lymphoid cells rank highly in our study, we hypothesize that cellular correlates might provide better insights toward mucosal immune responses (including antibody responses) compared to serum antibody measures.

A strength of our statistical modeling approach is the ability to define not only CoP but also baseline factors that increase the risk of symptomatic influenza. Among the strongest risk-associated cell populations identified across univariate and multivariate models were γδ T cells. A prior study from our group demonstrated that γδ T cells are integral for lung repair in neonatal mice[70]; however, in this adult SHIVERS-II cohort, the opposite effect was observed. This incongruence could be attributed to differences in γδ TCR repertoires, which change substantially with age, tissue localization and infection history[71,72]. Pediatric γδ T cells exhibit nonoverlapping TCR repertoires, whereas adult γδ T cells heavily favor Vγ9 and Vδ2 chain usage associated with polycytotoxic cytokine profiles[72]. Therefore, it is possible that, within the context of adult influenza virus infection, γδ T cells could contribute to inflammation-mediated pathology rather than tissue repair.

The results presented in this study emphasize the need to evaluate anti-influenza virus responses and vaccine efficacy from both the serologic and cellular standpoints, as well as argue that the composition of baseline cell populations is a better overall predictor of influenza susceptibility than serology and vaccination status alone. As baseline cellular and humoral immune landscapes significantly vary within human populations and across study cohorts, we propose that the protection-associated cell profiles defined in this study are a better metric to evaluate influenza susceptibility than any one cell population or serology measure. Although individual cellular CoP against symptomatic influenza likely differ between cohorts, their functions may contribute to a converging protective immune profile, which can be more easily compared across studies. Further, the statistical and predictive models described here will also serve as useful tools to evaluate vaccine efficacy, identify new targets for next-generation vaccines promoting both cellular and humoral responses, and identify additional at-risk human populations for vaccination. However, further questions remain. It will be important for follow-up studies to validate these risk and protection models across more diverse cohorts to determine whether there are population-specific baseline cell profiles correlated with vaccine failure or disease severity. Assessment of CMI and innate responses is especially important during pregnancy[73] and in adults over the age of 65 years, who are known to generate poor protective humoral responses following immunization[74,75].

The continuing aim of SHIVERS-II is to assess the immunological parameters of influenza virus infection and vaccination over several years. As the data presented in this study are derived from participants enrolled in the inaugural year of the SHIVERS-II study, there is great potential for future studies using this cohort to validate and expand on these results for years to come.

Study limitations

We acknowledge several limitations of our study. The first is the relatively small sample size across comparator groups and the increased circulation of the A(H1N1) virus during the study year. An over-representation of infections caused by the A(H1N1) virus means that no strong strain-specific conclusions regarding influenza B or A(H3N2) influenza viruses can be drawn. Future studies using samples collected in subsequent years of the SHIVERS-II cohort will focus on providing further validation to the models described here and on expanding the analyses to include a wider range of influenza virus types and subtypes. Second, participants enrolled in the SHIVERS-II cohort represent a limited regional, age and genetic scope; therefore, specific interpretations of these data should be made solely within this context. Further validation of our risk modeling and associated cell profiles should be compared across a more diverse ethnogeographic space to account for the large degree of baseline variability across participants. Finally, limited considerations were made regarding genetic contributions to the risk of influenza as, although ethnicity was reported, this is not a robust evaluation of genetic background[76]. Despite these limitations, data presented here are from a subset of year 1 participants, and it is our directive to continue to investigate this important population across multiple years of the SHIVERS-II study.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-023-01590-2.

References

1. Centers for Disease Control and Prevention. Estimated influenza illnesses, medical visits, hospitalizations, and deaths in the United States—2017–2018 influenza season. https://www.cdc.gov/flu/about/burden/2017-2018.htm#Table1 (2018).
2. Iuliano, A. D. et al. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet* **391**, 1285–1300 (2018).
3. Flannery, B. et al. Interim estimates of 2016–17 seasonal influenza vaccine effectiveness—United States, February 2017. *Morb. Mortal. Wkly. Rep.* **66**, 167–171 (2017).
4. Flannery, B. et al. Interim estimates of 2017–18 seasonal influenza vaccine effectiveness—United States, February 2018. *Morb. Mortal. Wkly. Rep.* **67**, 180–185 (2018).
5. Jackson, M. L. et al. Influenza vaccine effectiveness in the United States during the 2015–2016 season. *N. Engl. J. Med.* **377**, 534–543 (2017).
6. Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases. Estimated flu-related illnesses, medical visits, hospitalizations, and deaths in the United States—2019–2020 flu season. https://www.cdc.gov/flu/about/burden/2017-2018.htm (2021).
7. Zimmerman, R. K. et al. 2014–2015 Influenza vaccine effectiveness in the United States by vaccine type. *Clin. Infect. Dis.* **63**, 1564–1573 (2016).
8. Erbelding, E. J. et al. A universal influenza vaccine: the strategic plan for the National Institute of Allergy and Infectious Diseases. *J. Infect. Dis.* **218**, 347–354 (2018).
9. Morens, D. M., Taubenberger, J. K. & Fauci, A. S. Rethinking next-generation vaccines for coronaviruses, influenzaviruses, and other respiratory viruses. *Cell Host Microbe* **31**, 146–157 (2023).
10. Fourati, S. et al. Pan-vaccine analysis reveals innate immune endotypes predictive of antibody responses to vaccination. *Nat. Immunol.* https://doi.org/10.1038/s41590-022-01329-5 (2022).
11. Hagan, T. et al. Transcriptional atlas of the human immune response to 13 vaccines reveals a common predictor of vaccine-induced antibody responses. *Nat. Immunol.* https://doi.org/10.1038/s41590-022-01328-6 (2022).
12. Liston, A., Humblet-Baron, S., Duffy, D. & Goris, A. Human immune diversity: from evolution to modernity. *Nat. Immunol.* **22**, 1479–1489 (2021).
13. Tsang, J. S. et al. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell* **157**, 499–513 (2014).
14. Tsang, J. S. et al. Improving vaccine-induced immunity: can baseline predict outcome? *Trends Immunol.* **41**, 457–465 (2020).
15. Souquette, A. et al. Integrated drivers of basal and acute immunity in diverse human populations. Preprint at *bioRxiv* https://doi.org/10.1101/2023.03.25.534227 (2023).
16. Gounder, A. P. & Boon, A. C. M. Influenza pathogenesis: the effect of host factors on severity of disease. *J. Immunol.* **202**, 341–350 (2019).
17. Mettelman, R. C. & Thomas, P. G. Human susceptibility to influenza infection and severe disease. *Cold Spring Harb. Perspect. Med.* **11**, a038711 (2021).
18. Brodin, P. et al. Variation in the human immune system is largely driven by non-heritable influences. *Cell* **160**, 37–47 (2015).
19. Carr, E. J. et al. The cellular composition of the human immune system is shaped by age and cohabitation. *Nat. Immunol.* **17**, 461–468 (2016).
20. Lakshmikanth, T. et al. Human immune system variation during 1 year. *Cell Rep.* **32**, 107923 (2020).
21. Lee, M. N. et al. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science* **343**, 1246980 (2014).
22. Mettelman, R. C., Allen, E. K. & Thomas, P. G. Mucosal immune responses to infection and vaccination in the respiratory tract. *Immunity* **55**, 749–780 (2022).
23. O'Neill, M. B. et al. Single-cell and bulk RNA-sequencing reveal differences in monocyte susceptibility to influenza A virus infection between Africans and Europeans. *Front. Immunol.* **12**, 768189 (2021).
24. Orrù, V. et al. Genetic variants regulating immune cell levels in health and disease. *Cell* **155**, 242–256 (2013).
25. Patin, E. et al. Natural variation in the parameters of innate immune cells is preferentially driven by genetic factors. *Nat. Immunol.* **19**, 302–314 (2018).
26. Raj, T. et al. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science* **344**, 519–523 (2014).
27. Randolph, H. E. et al. Genetic ancestry effects on the response to viral infection are pervasive but cell type specific. *Science* **374**, 1127–1133 (2021).
28. Roederer, M. et al. The genetic architecture of the human immune system: a bioresource for autoimmunity and disease pathogenesis. *Cell* **161**, 387–403 (2015).
29. Ye, C. J. et al. Intersection of population variation and autoimmunity genetics in human T cell activation. *Science* **345**, 1254665 (2014).
30. Cox, R. J. Correlates of protection to influenza virus, where do we go from here? *Hum. Vaccin. Immunother.* **9**, 405–408 (2013).
31. Guthmiller, J. J. et al. First exposure to the pandemic H1N1 virus induced broadly neutralizing antibodies targeting hemagglutinin head epitopes. *Sci. Transl. Med.* **13**, eabg4535 (2021).

32. Memoli, M. J. et al. Evaluation of antihemagglutinin and antineuraminidase antibodies as correlates of protection in an influenza A/H1N1 virus healthy human challenge model. *mBio* **7**, e00417 (2016).
33. Monto, A. S. et al. Antibody to influenza virus neuraminidase: an independent correlate of protection. *J. Infect. Dis.* **212**, 1191–1199 (2015).
34. Ng, S. et al. Novel correlates of protection against pandemic H1N1 influenza A virus infection. *Nat. Med.* https://doi.org/10.1038/s41591-019-0463-x (2019).
35. Steel, J. et al. Influenza virus vaccine based on the conserved hemagglutinin stalk domain. *mBio* **1**, e00018 (2010).
36. Bandaranayake, D. et al. Risk factors and immunity in a nationally representative population following the 2009 influenza A(H1N1) pandemic. *PLoS ONE* **5**, e13211 (2010).
37. Huang, Q. S. et al. Risk factors and attack rates of seasonal influenza infection: results of the Southern Hemisphere Influenza and Vaccine Effectiveness Research and Surveillance (SHIVERS) seroepidemiologic cohort study. *J. Infect. Dis.* **219**, 347–357 (2019).
38. Institute of Environmental Science and Research. 2018 Annual influenza summary. https://www.esr.cri.nz/assets/Intelligence-Hub-2023/Surveillance-reports-and-dashboards/Influenza/InfluenzaAnn2018.pdf (2018).
39. Honce, R. & Schultz-Cherry, S. Impact of obesity on influenza A virus pathogenesis, immune response, and evolution. *Front. Immunol.* **10**, 1071 (2019).
40. Ursin, R. L. & Klein, S. L. Sex differences in respiratory viral pathogenesis and treatments. *Annu. Rev. Virol.* **8**, 393–414 (2021).
41. Wu, Y., Goplen, N. P. & Sun, J. Aging and respiratory viral infection: from acute morbidity to chronic sequelae. *Cell Biosci.* **11**, 112 (2021).
42. Institute of Environmental Science and Research. 2017 Annual influenza summary. https://www.esr.cri.nz/assets/Intelligence-Hub-2023/Surveillance-reports-and-dashboards/Influenza/2017-Influenza-Annual-report.pdf (2017).
43. Dormann, C. F. et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46 (2013).
44. Hamada, H. et al. Tc17, a unique subset of CD8 T cells that can protect against lethal influenza challenge. *J. Immunol.* **182**, 3469–3481 (2009).
45. Annunziato, F. et al. Phenotypic and functional features of human Th17 cells. *J. Exp. Med.* **204**, 1849–1861 (2007).
46. Takagi, R. et al. B cell chemoattractant CXCL13 is preferentially expressed by human Th17 cell clones. *J. Immunol.* **181**, 186–189 (2008).
47. Wang, X. et al. A critical role of IL-17 in modulating the B-cell response during H5N1 influenza virus infection. *Cell. Mol. Immunol.* **8**, 462–468 (2011).
48. Bao, J. et al. Decreased frequencies of Th17 and Tc17 cells in patients infected with avian influenza A (H7N9) virus. *J. Immunol. Res.* **2019**, 1418251 (2019).
49. Schultz-Cherry, S. Role of NK cells in influenza infection. *Curr. Top. Microbiol. Immunol.* **386**, 109–120 (2015).
50. Riese, P. et al. Responsiveness to influenza vaccination correlates with NKG2C-expression on NK cells. *Vaccines* **8**, 281 (2020).
51. Jost, S. et al. Changes in cytokine levels and NK cell activation associated with influenza. *PLoS ONE* **6**, e25060 (2011).
52. Björkström, N. K., Ljunggren, H.-G. & Sandberg, J. K. CD56 negative NK cells: origin, function, and role in chronic viral disease. *Trends Immunol.* **31**, 401–406 (2010).
53. Fox, A. et al. Severe pandemic H1N1 2009 infection is associated with transient NK and T deficiency and aberrant CD8 responses. *PLoS ONE* **7**, e31535 (2012).
54. Heltzer, M. L. et al. Immune dysregulation in severe influenza. *J. Leukoc. Biol.* **85**, 1036–1043 (2009).
55. Bongen, E., Vallania, F., Utz, P. J. & Khatri, P. KLRD1-expressing natural killer cells predict influenza susceptibility. *Genome Med.* **10**, 45 (2018).
56. Dou, Y. et al. Influenza vaccine induces intracellular immune memory of human NK cells. *PLoS ONE* **10**, e0121258 (2015).
57. Kay, A. W. et al. Enhanced natural killer-cell and T-cell responses to influenza A virus during pregnancy. *Proc. Natl Acad. Sci. USA* **111**, 14506–14511 (2014).
58. Giurgea, L. T., Morens, D. M., Taubenberger, J. K. & Memoli, M. J. Influenza neuraminidase: a neglected protein and its potential for a better influenza vaccine. *Vaccines* **8**, 409 (2020).
59. Rajendran, M., Krammer, F. & McMahon, M. The human antibody response to the influenza virus neuraminidase following infection or vaccination. *Vaccines* **9**, 846 (2021).
60. Wohlbold, T. J. et al. Vaccination with adjuvanted recombinant neuraminidase induces broad heterologous, but not heterosubtypic, cross-protection against influenza virus infection in mice. *mBio* **6**, e02556 (2015).
61. Turner, J. S. et al. Human germinal centres engage memory and naive B cells after influenza vaccination. *Nature* **586**, 127–132 (2020).
62. Bentebibel, S. E. et al. Induction of $ICOS^+CXCR3^+CXCR5^+$ TH cells correlates with antibody responses to influenza vaccination. *Sci. Transl. Med.* **5**, 176ra32 (2013).
63. Bentebibel, S. E. et al. $ICOS^+PD\text{-}1^+CXCR3^+$ T follicular helper cells contribute to the generation of high-avidity antibodies following influenza vaccination. *Sci. Rep.* **6**, 26494 (2016).
64. Lindgren, G. et al. Induction of robust B cell responses after influenza mRNA vaccination is accompanied by circulating hemagglutinin-specific $ICOS^+PD\text{-}1^+CXCR3^+$ T follicular helper cells. *Front. Immunol.* **8**, 1539 (2017).
65. Currenti, J. et al. Tracking of activated cTfh cells following sequential influenza vaccinations reveals transcriptional profile of clonotypes driving a vaccine-induced immune response. *Front. Immunol.* **14**, 1133781 (2023).
66. Pizzolla, A. et al. Influenza-specific lung-resident memory T cells are proliferative and polyfunctional and maintain diverse TCR profiles. *J. Clin. Invest.* **128**, 721–733 (2018).
67. Lorenzo-Herrero, S., Sordo-Bahamonde, C., Gonzalez, S. & López-Soto, A. CD107a degranulation assay to evaluate immune cell antitumor activity. in *Cancer Immunosurveillance: Methods and Protocols* (eds. López-Soto, A. & Folgueras, A. R.) 119–130 (Springer, 2019); https://doi.org/10.1007/978-1-4939-8885-3_7
68. Wilkinson, T. M. et al. Preexisting influenza-specific $CD4^+$ T cells correlate with disease protection against influenza challenge in humans. *Nat. Med.* **18**, 274–280 (2012).
69. Krammer, F. The human antibody response to influenza A virus infection and vaccination. *Nat. Rev. Immunol.* **19**, 383–397 (2019).
70. Guo, X. J. et al. Lung γδ T cells mediate protective responses during neonatal influenza infection that are associated with type 2 immunity. *Immunity* **49**, 531–544 (2018).
71. Clark, B. L. & Thomas, P. G. A cell for the ages: human γδ T cells across the lifespan. *Int. J. Mol. Sci.* **21**, 8903 (2020).
72. Sant, S. et al. Human γδ T-cell receptor repertoire is shaped by influenza viruses, age and tissue compartmentalisation. *Clin. Transl. Immunol.* **8**, e1079 (2019).
73. Savic, M. et al. Distinct T and NK cell populations may serve as immune correlates of protection against symptomatic pandemic influenza A(H1N1) virus infection during pregnancy. *PLoS ONE* **12**, e0188055 (2017).
74. McElhaney, J. E. et al. T cell responses are better correlates of vaccine protection in the elderly. *J. Immunol.* **176**, 6333–6339 (2006).

75. McElhaney, J. E. et al. Granzyme B: correlates with protection and enhanced CTL response to influenza vaccination in older adults. *Vaccine* **27**, 2418–2425 (2009).
76. Mersha, T. B. & Abebe, T. Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. *Hum. Genomics* **9**, 1 (2015).

[1]Department of Immunology, St. Jude Children's Research Hospital, Memphis, TN, USA. [2]Department of Host–Microbe Interactions, St. Jude Children's Research Hospital, Memphis, TN, USA. [3]Department of Microbiology, Immunology and Biochemistry, College of Graduate Health Sciences, University of Tennessee Health Science Center, Memphis, TN, USA. [4]Institute of Environmental Science and Research Limited (ESR), Wallaceville Science Centre, Upper Hutt, New Zealand. [10]These authors contributed equally: Robert C. Mettelman, Aisha Souquette. [11]These authors jointly supervised this work: Richard J. Webby, Q. Sue Huang, Paul G. Thomas. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: richard.webby@stjude.org; sue.huang@esr.cri.nz; paul.thomas@stjude.org

SHIVERS-II Investigation Team

Judy Bocacao[4], Jacqui Ralston[4], Jessica Danielewicz[5], Wendy Gunn[4], Nayyereh Aminisani[4], Ben Waite[4], R. Pamela Kawakami[4], Annette Nesdale[6], Michelle Balm[6], Nikki Turner[7,8] & Tony Dowell[9]

[5]Institute of Environmental Science and Research Limited (ESR), Mt Albert Science Centre, Auckland, New Zealand. [6]Te Whatu Ora Health New Zealand (Capital, Coast and Hutt Valley), Wellington, New Zealand. [7]Department of General Practice and Primary Health Care, School of Population Health, University of Auckland, Auckland, New Zealand. [8]Immunisation Advisory Centre, University of Auckland, Auckland, New Zealand. [9]Department of Primary Health Care and General Practice, University of Otago, Wellington, New Zealand.

Methods

Study ethics statement

This study received ethics approval from the New Zealand Health & Disability Ethics Committee (NTX11.11.102.AM36 & AM36 & AM49 & AM51). All participants provided written consent to participate in the study. Deidentified participant samples (PBMCs, serum) and associated demographic information were provided under the St. Jude Children's Research Hospital Center of Excellence for Influenza Research and Surveillance NIAID contract HHSN272201400006C.

SHIVERS-II study design and participant definitions

The SHIVERS-II study is a population-based, longitudinal (2018–2022) prospective cohort study in Wellington, New Zealand, initiated to evaluate cellular and serologic immune responses at baseline and during recall among adults with current influenza virus vaccination and/or infection. During year 1 (2018), more than 22,000 individuals aged 20–69 years were randomly selected for recruitment from three participating primary health organizations' healthy patient networks. Selected individuals received a study invitation and information packet by mail, and those interested provided consent through an online electronic consent form and questionnaire in which demographic and contact information, vaccination status, and health and ILI status were provided. A total of 2,195 individuals were ultimately enrolled following written informed consent (Fig. 1a). Each participant had one baseline blood draw (preseason, usually during March through May) and a postseason blood draw (usually during October through December). Consented participants received a New Zealand $30 gift card after each blood or swab sample collection to recognize their time and effort. Whole blood samples were collected in vacuum tubes containing heparin. Additional samples, including blood draw samples or nasal swabs, were obtained from participants with PCR-confirmed influenza virus infection or who received an influenza virus vaccination during the influenza season (May through September). Briefly, participants who received an influenza virus vaccination during this time provided a blood draw sample 14 days after vaccination (postvaccination baseline). All participants were monitored, through weekly surveys, for ILI, defined by the WHO as "acute respiratory illness with cough and a history of fever/measured fever of ≥38 °C and illness onset within the past 10 days" (ref. 77). If ILI presented, a respiratory specimen (nasopharyngeal or throat swab) was collected and subjected to PCR molecular testing for respiratory viruses (see methods below). Participants with both ILI and an associated influenza-positive PCR result were categorized as having symptomatic influenza virus infection, and additional acute (1–2 weeks after ILI onset) and convalescent (4–7 weeks after ILI onset) blood draws were conducted. Respiratory samples from these PCR-positive influenza cases were further processed to determine the influenza virus subtype (see methods below). Inhibiting antibody titers against HA and NA were determined from baseline and postseason serum samples by HAI assay or NAI enzyme-linked lectin assay (NAI-ELLA), respectively (see methods below). Total binding antibody titers against purified, full-length HA and NA were measured by ELISA (see methods below). HAI titers were used to detect cryptic influenza virus infection cases not identified by PCR. Briefly, cryptic influenza virus infections were defined by (1) the absence of an ILI-associated influenza-positive PCR result and (2) one of the following seroconversion criteria: unvaccinated participants with a fourfold or greater increase in the preseason to postseason HAI titer, vaccinated participants with a fourfold or greater increase in the postvaccination to postseason HAI titer, or vaccinated participants with a fourfold or greater increase in the preseason to postseason HAI titer without seroconversion following vaccination (that is, seroconversion cannot be accounted for by vaccination), with the second titer being at least 1:40 in all cases.

Participant categorization and selection

Participants were categorized based on 2018 influenza virus vaccination (vaccinated or unvaccinated) or infection (symptomatic, cryptic or uninfected) status. The symptomatic influenza virus infection group met the WHO-defined criteria for ILI and had a positive influenza virus PCR test. Cryptic influenza virus infections were defined as cases with zero to one mild symptom that did not meet the WHO-defined ILI criteria (and therefore do not have an associated PCR test) but were confirmed to have had a cryptic influenza virus infection by seroconversion. All cryptic infections in this study are, by definition, subclinical as no influenza PCR-positive-associated ILI was reported. The uninfected group did not meet the WHO-defined ILI criteria, did not undergo an associated PCR test and were negative for cryptic infection. A total of 206 participants were selected from year 1 of the SHIVERS-II study and comprised 82 men (39.8%) and 124 women (60.2%) residing in the Wellington, New Zealand, catchment area who had an average age of 43.8 ± 12.6 years (range 20–68 years), had a BMI of 27.5 ± 5.6 kg m^{-2} (range 18.3–48.5 kg m^{-2}) and predominantly self-identified as being of New Zealand–European descent (83.0%) (Table 1). Participants were selected into four comparator groups: vaccinated–uninfected ($n = 75$), vaccinated–infected ($n = 33$), unvaccinated–uninfected ($n = 76$) and unvaccinated–infected ($n = 22$). Participants' age (in years) and sex (assigned at birth) were roughly matched when selecting across comparator groups. Sample size requirements for statistical testing of the covariates were calculated using the R package WMWssp (v.0.4.0) (refs. 78–80). At a defined power of 0.8, our sample was determined to be sufficient in size for the planned comparisons.

Composition of the 2018 influenza virus vaccine

The 2018 seasonal quadrivalent influenza virus vaccine had the following four components: (1) A/Michigan/45/2015 (H1N1)pdm09-like virus, (2) A/Singapore/INFIMH-16-0019/2016 (H3N2)-like virus, (3) B/Phuket/3073/2013-like virus (B/Yamagata lineage) and (4) B/Brisbane/60/2008-like virus (B/Victoria lineage).

Molecular testing for respiratory viruses

For the detection of respiratory viruses including influenza viruses, the CDC standard RT–PCR assay was performed on RNA extracted from the participants' nasopharyngeal or throat swabs, as previously described[81–85]. All samples were tested for a standard panel including influenza A (A(H1N1)pdm09, A(H3N2)) viruses, influenza B (B/Yamagata and B/Victoria lineages) viruses, respiratory syncytial virus, rhinovirus, human metapneumovirus, parainfluenza types 1–3, adenovirus and enterovirus.

Influenza virus antigenic typing

Antigenic typing of the infecting influenza virus strain was performed on all samples identified by qPCR as influenza virus positive with low cycle threshold values. Following virus isolation by sample inoculation into Madin–Darby canine kidney cells stably expressing human α-2,6-sialyltransferase[86], antigenic typing of influenza viruses was performed using an HAI assay with standard antisera supplied by the WHO Collaborating Center in Melbourne, Australia. Any untypeable influenza A viruses were forwarded to the WHO Collaborating Center in Melbourne or the CDC in Atlanta, GA, for further characterization.

HAI assay

Anti-HA antibody titers were determined by HAI assay as previously described for influenza A (A/Michigan/45/2015 (H1N1)pdm09-like, A/Singapore/INFIMH-16-0019/2016 (H3N2)-like) and influenza B (B/Brisbane/60/2008-like (B/Victoria lineage), B/Phuket/3073/2013-like (B/Yamagata lineage)) viruses present in the 2018 Southern Hemisphere vaccine and in regional circulation[36–38]. Briefly, serum samples were treated with receptor-destroying enzyme (Denka Seiken) at 37 °C

overnight, heat-inactivated at 56 °C for 30 min and tested by HAI assay with 0.5% turkey red blood cells. Serum samples were serially diluted twofold in 96-well U-bottom plates beginning at a 1:10 dilution. Four agglutinating doses were added to each well, and the samples were incubated. Titers were read after 30 min of incubation with 50 µl of 0.5% turkey red blood cells.

NAI-ELLA assay

NAI antibody titers were determined by ELLA for influenza A and B viruses[37,87,88]. Briefly, recombinant viruses, composed of the NA from each of the influenza A viruses used for HAI and a mismatched H6 on a PR8 backbone, were generated by reverse genetics and used as antigens. Serum samples were tested at a starting dilution of 1:10. Because no recombinant antigen was available for influenza B viruses, ELLA was performed against the whole B/Brisbane/60/2008 virus (B/Victoria lineage) as previously described[89].

ELISA

Enzyme immunoassay/radioimmunoassay 96-well plates (Corning, cat. no. 3590) were individually coated with 0.5 µg purified recombinant influenza virus HA or NA protein at 0.01 mg ml^{-1}. HA antigens derived from A/Michigan/45/2015 (representative H1; Sino Biological, cat. no. 40567-V08H1), A/Singapore/INFIMH-16-0019/2016 (representative H3; Sino Biological, cat. no. 40580-V08H), B/Phuket/3073/2013 (representative B/Yamagata lineage; Sino Biological, cat. no. 40498-V08B) or B/Brisbane/60/2008 (representative B/Victoria lineage; Sino Biological, cat. no. 40016-V08H) and NA antigens derived from A/Michigan/45/2015 (representative N1; Sino Biological, cat. no. 40568-V07H), A/Singapore/INFIMH-16-0019/2016 (representative N2; Sino Biological, cat. no. 40802-V08B), B/Phuket/3073/2013 (representative B/Yamagata lineage; Sino Biological, cat. no. 40502-V07B) or B/Washington/02/2019 (representative B/Victoria lineage; Sino Biological, cat. no. 40790-V08B) were used. Coated plates were sealed and stored at 4 °C for 24 h. Serum samples were treated by adding 3 volumes of receptor-destroying enzyme (Accurate Chemical, cat. no. YCC340122) to 1 volume of serum and incubated at 37 °C overnight. Serum samples were then heat-treated at 56 °C for 1 h, after which 6 volumes of 1× PBS was added to obtain a 1:10 starting dilution of serum. Plates were washed three times with 1× PBS and coated with 100 µl of blocking buffer per well (1% BSA (Gibco, cat. no. 15260037) and 0.01% Tween-20 (ThermoFisher, cat. no. 85113) in 1× PBS) for 6 h at 25 °C. Serum samples treated with receptor-destroying enzyme were titrated at 1:160–1:20,480 in blocking buffer, and 50 µl of each serum dilution or blocking buffer (negative control) was added to replicate wells following the removal of blocking buffer. The plates were then sealed and incubated at 4 °C overnight. Plates were washed three times with wash buffer (0.05% Tween-20 in 1× PBS), after which 100 µl of anti-human IgG (Fab-specific)–horseradish peroxidase secondary antibody (Sigma, cat. no. A0293, polyclonal, lot no. 0000201676, dilution 1:3,000) was added per well and plates were incubated at 4 °C overnight. The plates were then washed three times in wash buffer, and 50 µl of tetramethylbenzidine (Sigma, cat. no. T0440) was added to each well. After 8 min, 50 µl of stop solution (1 N H_2SO_4; Sigma, cat. no. T0440) was added per well and the plates were read at 450 nm on the BioTek Synergy H1 microplate reader. To normalize the optical density (OD), average OD values from negative control wells were subtracted from the average OD of replicate sample wells per plate. AUC values were calculated from the *xy* plots of dilution by the normalized OD values in GraphPad Prism 9 (v.9.5.1 (528)), using the following parameters: baseline $y = 0$; minimum peak height, <10% of the distance from minimum to maximum *y*; all peaks above baseline; five significant digits.

Flow cytometry: surface and intracellular staining

Cryopreserved aliquots of PBMCs were thawed at 37 °C and suspended in RPMI 1640 supplemented with 10% heat-inactivated FBS (Gibco, cat. no. 16140071), 1% nonessential amino acids (Gibco, cat. no. 11140-050), 1 mM sodium pyruvate (Gibco, cat. no. 11360-070) and 1% penicillin–streptomycin (10,000 U ml^{-1} penicillin and 10,000 µg ml^{-1} streptomycin; Gibco, cat. no. 15140-122). Cells were plated at 3.0×10^5–4.0×10^5 cells per well in a 96-well U-bottom plate.

For peptide-stimulated PBMCs, pooled peptides (BEI Resources, cat. no. NR-2667) derived from influenza virus M1, NP and PB1 were added to wells at a final concentration of 5 µg ml^{-1}. For virus stimulation, PBMCs were infected with H1N1pdm09 (A/Michigan/45/2015) or H3N2 (A/Singapore/INFIMH-16-019/2016) at an MOI of 4 in virus infection medium (RPMI 1640 completed with 0.3% BSA, 0.225% $NaCO_3$, 1% penicillin–streptomycin (10,000 U/ml^{-1} penicillin and 10,000 µg ml^{-1} streptomycin; Gibco), 1% L-glutamine (Gibco) and 1% vitamin solution (Gibco)). Cell stimulation cocktail (BioLegend, cat. no. 423302; eBioscience, cat. no. 00-4970-93) was added to positive control wells. Medium alone was used in unstimulated wells. Thereafter, cells were incubated for 2 h (peptide, mock) or 12 h (virus, stimulation cocktail) at 37 °C. Brefeldin A (BioLegend, cat. no. 420601; BD, cat. no. 555029) and monensin (BD, cat no. 554715) were then added along with anti-CD107a fluorophore-conjugated antibodies where appropriate, and samples were incubated at 37 °C for an additional 6 h. Cells were washed twice in FACS buffer (Dulbecco's PBS, 2% FBS, 1 mM EDTA pH 8.0) and treated with human Fc-block (BioLegend, cat. no. 422302) for 10 min at 4 °C.

Cells were stained for 30 min at 4 °C with the following antibodies or dyes: CCR3 AF647 (BioLegend, cat. no. 310710, clone 5E8, lot no. B312457, dilution 1:50), CCR5 peridinin chlorophyll (PerCP)/eFluor 710 (ThermoFisher, cat. no. 46-1951-82, clone 7A4, lot no. 4324286, dilution 1:80), CCR7 phycoerythrin (PE)/cyanine 7 (Cy7) (BioLegend, cat. no. 353226, clone G04H7, lot no. B305236, dilution 1:20), CD1c PE/Cy7 (BioLegend, cat. no. 331516, clone L161, lot no. B256006, dilution 1:20), CD3 BV750 (BioLegend, cat. no. 344846, clone SK7, lot no. B303395, dilution 1:200), CD3 BV510 (BioLegend, cat. no. 300448, clone UCHT1, lot no. B281774, dilution 1:200), CD4 BB515 (BD, cat. no. 565996, clone SK3, lot no. 9343297, dilution 1:200), CD8 BV570 (BioLegend, cat. no. 301038, clone RPA-T8, lot no. B281322, dilution 1:50), CD11b CD785 (BioLegend, cat. no. 301346, clone ICRF44, lot no. B315938, dilution 1:100), CD11c BV650 (BioLegend, cat. no. 337238, clone Bu15, lot no. B304739, dilution 1:20), CD14 allophycocyanin (APC)/Fire 750 (BioLegend, cat. no. 367120, clone 63D3, lot no. B257669, dilution 1:100), CD16 fluorescein isothiocyanate (FITC) (BioLegend, cat. no. 360716, clone B73.1, lot no. B300038, dilution 1:100), CD19 BV510 (BioLegend, cat. no. 302242, clone HIB19, lot no. B281769, dilution 1:200), CD33 PE (BioLegend, cat. no. 303404, clone WM53, lot no. B288411, dilution 1:20), CD40 AF700 (BioLegend, cat. no. 334328, clone 5C3, lot no. B272227, dilution 1:20), CD45RA V450 (ThermoFisher, cat. no. 48-0458-42, clone HI100, lot no. 2263338, dilution 1:200) or BV570 (BioLegend, cat. no. 304132, clone HI100, lot no. B319894, dilution 1:20), CD45RO APC/Fire 750 (BioLegend, cat. no. 304250, clone UCHL1, lot no. B288981, dilution 1:20), CD56 PE/Cy5 (BioLegend, cat. no. 362516, clone 5.1H11, lot no. B317268, dilution 1:100), CD64 BV711 (BD, cat. no. 740782, clone 10.1, dilution 1:20), CD80 Super Bright 436 (ThermoFisher, cat. no. 62-0809-42, clone 2D10.4, lot no. E113289A, dilution 1:20), CD103 BV605 (BioLegend, cat. no. 350217, clone Ber-ACT8, lot no. B359323, dilution 1:20), CD107a BV711 (BioLegend, cat. no. 328640, clone H4A3, lot no. B301430, dilution 1:50) or PE/Dazzle 594 (BioLegend, cat. no. 328646, clone H4A3, lot no. B319152, dilution 1:20), CD123 PE (BioLegend, cat. no. 306006, clone 6H6, lot no. B269700, dilution 1:100) or PerCP/Cy5.5 (BioLegend, cat. no. 306016, clone 6H6, lot no. B318963, dilution 1:20), CD141 BV421 (BioLegend, cat. no. 344144, clone M80, lot no. B315574, dilution 1:100), CD177 APC (BioLegend, cat. no. 315808, clone MEM-166, lot no. B265153, dilution 1:50), CXCR3 PE/Dazzle 594 (BioLegend, cat. no. 353736, clone G025H7, lot no. B289168, dilution 1:100) or BV711 (BioLegend, cat. no. 353732, clone G025H7, lot no. B264427, dilution 1:20), CXCR5 Super Bright 436 (ThermoFisher, cat. no. 62-9185-42, clone MU5HBEE, lot no. 2279709, dilution 1:100), HLA-A2 BV650 (BioLegend,

cat. no. 343324, clone BB7.2, lot no. B290341, dilution 1:200), HLA-DR AF700 (BioLegend, cat. no. 307626, clone L243, lot no. B272227, dilution 1:100) or PE/Dazzle 594 (BioLegend, cat. no. 307654, clone L243, lot no. B320167, dilution 1:20), ICOS PerCP/Cy5.5 (BioLegend, cat. no. 313518, clone C398.4A, lot no. B297477, dilution 1:20), Ghost Dye Violet 510 viability dye (Tonbo, cat. no. 13-0870-T500, lot no. D0870061322133, dilution 1:400), PD-1 FITC (BioLegend, cat. no. 329904, clone EH12.2H7, lot no. B253784, dilution 1:100), Siglec-8 BV480 (BD, cat. no. 747874, clone 837535, lot no. 1006002, dilution 1:50) and TCRγδ AF647 (BioLegend, cat. no. 331214, clone B1, lot no. B274274, dilution 1:80).

Cells were washed twice with FACS buffer and then fixed and permeabilized using the Cytofix/CytoPerm fixation and permeabilization kit (BD, cat. no. 554714) according to the manufacturer's instructions. Cells were washed twice in Perm/Wash buffer and then stained for 30 min at 4 °C with the following intracellular antibodies: CD68 PerCP/Cy5.5 (BioLegend, cat. no. 333814, clone Y1/82A, lot no. B273670, dilution 1:50), GzmB AF700 (BioLegend, cat. no. 372222, clone QA16A02, lot no. B321316, dilution 1:25), IL-2 APC (ThermoFisher, cat. no. 17-7029-82, clone MQ1-17H12, lot no. 2172935, dilution 1:25), IL-13 BV421 (BioLegend, cat. no. 501916, clone JES10-5A2, lot no. B319799, dilution 1:25), IL-17 BV785 (BioLegend, cat. no. 512338, clone BL168, lot no. B310823, dilution 1:25), IL-21 PE (BioLegend, cat. no. 513004, clone 3A3-N2, lot no. B299502, dilution 1:25), IFNγ BV480 (BD, cat. no. 566100, clone B27, lot no. 0261295, dilution 1:10) and TNFα BV605 (BioLegend, cat. no. 502936, clone Mab11, lot no. B312493, dilution 1:25). Cells were washed twice and resuspended in FACS buffer for flow cytometric analysis. Data were acquired on a three-laser Cytek Aurora spectral flow cytometer using SpectroFlow (v.2.2, Cytek) and analyzed using FlowJo (v.10.7.1, TreeStar). Spectral cytometry panels conform to best-practice principles in spectral analysis[90,91].

Flow cytometry: selecting cell subsets for primary analysis

Cell frequencies were resolved for more than 89 CMI and innate subtypes, including populations previously shown to correlate with anti-influenza virus response. For unbiased selections, several blinded quality control measures were used to identify cell populations for inclusion in the primary analysis. These included filtering out individual samples with low cell viability (≤5%) and/or high cellular debris (>25% of events) and eliminating cell populations with limited dynamic range in cell frequency. Using these measures, we reduced the number of individual myeloid populations from 23 to 16 and the number of lymphoid/functional cell populations from 66 to 41 for the primary analysis. Refer to Supplementary Table 7 for cell population naming conventions used in this study.

Defining co-regulated immune cell modules

Co-regulated immune cell modules were independently determined for myeloid or lymphoid/functional cell populations across all vaccinated and unvaccinated participants. For the myeloid cell populations, modules ('clusters') were identified by a significant positive correlation between cell frequencies (% parent), using Pearson's bivariate correlation and an FDR adjustment cutoff of $q \leq 0.05$. To define modules within the lymphoid/functional compartment, cell population frequencies (% parent) were first averaged across virus (MOI = 4, A/Michigan/45/2015 (H1N1)pdm09 or A/Singapore/INFIMH-16-019/2016 (H3N2)) and peptide (peptide pools containing M1, NP and PB1, 1–5 µM per peptide) stimulation conditions to obtain the average stimulated cell frequency ('Stim Average'). Lymphoid/functional clusters were then identified by a significant positive correlation between Stim Average frequencies, using Pearson's bivariate correlation with an FDR adjustment cutoff of $q \leq 0.05$. Analyses were performed using the base library in R (v.3.6.0) (ref. 92).

Statistical modeling

Univariate statistical comparisons (Supplementary Table 3) were performed independently on vaccinated and unvaccinated participants and are presented as differences between cases (symptomatic influenza) and controls (uninfected and cryptic influenza) within each group. Categorical variables and associated levels are displayed as counts and percentages across the target variable, and Fisher's exact test was used to compute *P* values to test the significance of the association of the variable across cases and controls. Continuous variables are represented by their median and IQR across the symptomatic and cryptic/uninfected groups, and the significance of the difference between the two groups was computed using a single-sided Kruskal–Wallis test. Statistical analyses were performed using the base library in R (v.3.6.0) (ref. 92) and sjTabone (v.0.1.0) (ref. 93). Six participants with consistent cell frequency outliers across multiple cell populations determined by Grubbs' test ($P \leq 0.05$) were excluded, and a total of 200 participants were analyzed.

Univariate regression modeling was performed individually on all cellular covariates (separately for vaccinated and unvaccinated participants), on selected co-regulated cell modules, and on a subset of participants with cryptic and symptomatic infection (for serology measures). Binomial regression models were built from each cellular or serology measure or from the cluster frequency (independent variable) to evaluate the effects on symptomatic infection risk (dependent variable). Individual multivariate regression models describing the risk of symptomatic influenza, given individual serology measures adjusted for age (years), sex, BMI (kg m^{-2}) and 2018 vaccination status, were also generated. Logit estimates from each GLM were transformed exponentially to obtain OR values.

For the multivariate logistic regression considering demographic, vaccination status, serology and cell immune measures, a stepwise approach was used to identify appropriate variables to include in the GLM. First, we identified highly correlated cell population clusters for both lymphoid/functional and myeloid compartments (Fig. 7b) or across serology measures (Extended Data Fig. 2a). The resulting Pearson's correlation coefficients were informative in selecting representative immune measures from each cluster to be used in statistical models. Inhibiting titers for all HAI and NAI were $\log_2$ transformed. All serology measures and cell populations found to be influenced by vaccination status in the univariate analyses were considered as both individual covariates and interacting terms with the 2018 influenza vaccination status. As collinearity is a common problem in studies in which the underlying phenomenon that generates the observations is not fully specified, we also assessed multicollinearity across all covariates through VIF analysis. The results from the correlation and VIF analyses identified a 'selected' set of covariates that could be used in traditional regression methods without overfitting. A GLM was constructed with the final set of selected covariates against a response variable denoting symptomatic infection status. Stepwise regression models (both forward-selection and backward-elimination) were also constructed using AIC as the metric of model effectiveness. The final stepwise model had a lower AIC than the select GLM but resulted in higher residual deviance. BMA was also used to test the strength of the selected covariates when averaged over a wider range of models. Continuous dependent covariates were normalized using scaled mean and s.d. values (age, ELISA AUC, % parent cell frequency) or by $\log_2$ transformation (HAI, NAI). Multivariate analyses were performed using the base library in R (v.3.6.0) (ref. 92) and BMA (v.3.18.17) (ref. 94).

Univariate protection and susceptibility thresholds

ROC curves were used to test the effectiveness of each univariate immune measure—from the complete set of participants, the vaccinated subset and the unvaccinated subset—as a diagnostic indicator[95,96]. From the ROC analysis, we (1) measured trade-offs between the sensitivity and specificity of a given immune measure and (2) determined the optimal threshold achieving 50% sensitivity to generalize the behavior of the measure. The AUC was also determined for each

immune measure and indicates the overall quality of the measure in capturing true positives. Analyses were performed using the base library in R (v.3.6.0) (ref. 92).

Predictive modeling

An 80:20 (train to test) split was performed on 200 participants (after the removal of six participants with significant cell population frequency outliers) while ensuring equivalent proportions of cases (symptomatic influenza) and controls (uninfected and cryptic influenza). Four separate random forest models were built: (1) a base model that predicts symptomatic infection using demographic (age, sex, BMI category, ethnicity, 2018 influenza virus vaccination status) and serologic (HAI A(H1), HAI A(H3), HAI B/Victoria lineage, HAI B/Yamagata lineage, NAI A(N1), NAI A(N2), NAI B/Victoria lineage, NAI B/Yamagata lineage, A(H1) AUC, A(H3) AUC, B/Victoria lineage HA AUC, B/Yamagata lineage HA AUC, A(N1) AUC, A(N2) AUC, B/Victoria lineage NA AUC, B/Yamagata lineage NA AUC) covariates; (2) a lymphoid model that uses the variables from the base model and cell populations from the lymphoid/functional panel (total $CD4^+$ T cells, TNF&aplha;$^+CD4^+$ T cells, $CD4^+$ effector T cells, IL-2^+CD4^+ T cells, naïve $CD4^+$ T cells, PD-1^+CD4^+ T cells, $CD4^+$ T_H17 cells, dual-cytokine-producing $CD4^+$ T cells, total $CD8^+$ T cells, $CD107a^+CD8^+$ T cells, $CD8^+$ effector T cells, $IFN\gamma^+CD8^+$ T cells, IL-2^+CD8^+ T cells, $CCR5^+CD8^+$ memory T cells, PD-1^+CD8^+ T cells, TNF&aplha;$^+CD8^+$ T cells, single-cytokine-producing $CD8^+$ T cells, dual-cytokine-producing $CD8^+$ T cells, $CXCR3^+$ cT_{FH} cells, $ICOS^+$ cT_{FH} cells, $GzmB^-IFN\gamma^+$ NK cells, $GzmB^+IFN\gamma^+$ NK cells, $GzmB^+IFN\gamma^-$ NK cells, γδ T cells); (3) a myeloid model that uses the variables from the base model and cell populations from the myeloid panel (mDCs, cytokine-producing NK cells, activated NK cells, cytotoxic NK cells, intermediate monocytes, nonclassical monocytes, basophils, eosinophils); (4) a combined model that uses all available variables; (5) a myeloid-only model that contains populations from the myeloid panel alone; (6) a lymphoid-only model that contains populations from the lymphoid panel alone; and (7) a lymphoid + myeloid model that contains populations from both the lymphoid and myeloid panels. For cellular variables with high frequency correlation, a representative population was selected for inclusion into the models. Fivefold cross-validation was performed to avoid overfitting the data and repeated three times to generate reliable metrics of model performance. Model performance was measured by sensitivity and specificity and by scoring test-set data and plotting ROC curves. VIP analysis was performed on the combined model, and values were scaled between 0 and 100, with higher values denoting a greater impact on prediction. Random forest and VIP analyses were performed using the base library in R (v.3.6.0) (ref. 92) and caret (v.6.0-92) (ref. 97).

Data visualization

Data visualizations were performed in R (v.3.6.0) using ggplot2 (v.3.4.0) (ref. 98) and ggpubr (v.0.5.0) (ref. 99) and assembled using Adobe Illustrator 2023 software. Figure 1b was created with BioRender.com and exported under a paid subscription with an associated publication license.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The published article includes all datasets generated or analyzed as a part of this study. Individual source data are provided with associated figures (where appropriate) per the data sharing agreement stipulated under the Ruth L. Kirschstein National Research Service Award Individual Postdoctoral Fellowship (award no. F32AI157296; R.C.M.). Raw flow cytometry source files can be made available upon reasonable request. Source data are provided with this paper.

Code availability

A minimum dataset containing deidentified study participant information and biological assay results along with custom study-generated R code for analysis was uploaded to GitHub (https://github.com/kvegesan-stjude/SHIVERS2) per the data sharing agreement stipulated under the Ruth L. Kirschstein National Research Service Award Individual Postdoctoral Fellowship (award no. F32AI157296; R.C.M.). Additional basic R code can be made available upon reasonable request.

References

77. World Health Organization. *Global Epidemiological Surveillance Standards for Influenza*. https://www.who.int/publications/i/item/9789241506601 (World Health Organization, 2013).
78. Brunner, E., Bathke, A. C. & Konietschke, F. *Rank and Pseudo-Rank Procedures for Independent Observations in Factorial Designs* (Springer, 2018); https://doi.org/10.1007/978-3-030-02914-2
79. Bürkner, P. C., Doebler, P. & Holling, H. Optimal design of the Wilcoxon–Mann–Whitney-test. *Biom. J.* **59**, 25–40 (2017).
80. Happ, M., Bathke, A. C. & Brunner, E. Optimal sample size planning for the Wilcoxon–Mann–Whitney test. *Stat. Med.* **38**, 363–375 (2019).
81. Huang, Q. S. et al. Implementing hospital-based surveillance for severe acute respiratory infections caused by influenza and other respiratory pathogens in New Zealand. *Western Pac. Surveill. Response J.* **5**, 23–30 (2014).
82. Kim, C. et al. Comparison of nasopharyngeal and oropharyngeal swabs for the diagnosis of eight respiratory viruses by real-time reverse transcription-PCR assays. *PLoS ONE* **6**, e21610 (2011).
83. Kodani, M. et al. Application of TaqMan low-density arrays for simultaneous detection of multiple respiratory pathogens. *J. Clin. Microbiol.* **49**, 2175–2182 (2011).
84. Olsen, S. J. et al. Incidence of respiratory pathogens in persons hospitalized with pneumonia in two provinces in Thailand. *Epidemiol. Infect.* **138**, 1811–1822 (2010).
85. Shu, B. et al. Design and performance of the CDC real-time reverse transcriptase PCR swine flu panel for detection of 2009 A (H1N1) pandemic influenza virus. *J. Clin. Microbiol.* **49**, 2614–2619 (2011).
86. Matrosovich, M., Matrosovich, T., Carr, J., Roberts, N. A. & Klenk, H.-D. Overexpression of the α-2,6-sialyltransferase in MDCK cells increases influenza virus sensitivity to neuraminidase inhibitors. *J. Virol.* **77**, 8418–8425 (2003).
87. Sandbulte, M. R. et al. Cross-reactive neuraminidase antibodies afford partial protection against H5N1 in mice and are present in unexposed humans. *PLoS Med.* **4**, e59 (2007).
88. Wong, S. S. et al. H5N1 influenza vaccine induces a less robust neutralizing antibody response than seasonal trivalent and H7N9 influenza vaccines. *NPJ Vaccines* **2**, 16 (2017).
89. Couzens, L. et al. An optimized enzyme-linked lectin assay to measure influenza A virus neuraminidase inhibition antibody titers in human sera. *J. Virol. Methods* **210**, 7–14 (2014).
90. Ferrer-Font, L. et al. Panel design and optimization for high-dimensional immunophenotyping assays using spectral flow cytometry. *Curr. Protoc. Cytom.* **92**, e70 (2020).
91. Liechti, T. et al. An updated guide for the perplexed: cytometry in the high-dimensional era. *Nat. Immunol.* https://doi.org/10.1038/s41590-021-01006-z (2021).
92. R Core Team. *R: A Language and Environment for Statistical Computing* (R Project for Statistical Computing, 2018).
93. Vegesana, K. sjTabone. *Github* https://github.com/kvegesan-stjude/sjTabone#readme (2023).
94. Raftery, A., Hoeting, J., Volinsky, C., Painter, I. & Yeung, K. Y. BMA: Bayesian model averaging. https://cran.r-project.org/web/packages/BMA/BMA.pdf (2022).

95. Mandrekar, J. N. Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* **5**, 1315–1316 (2010).
96. Greiner, M., Pfeiffer, D. & Smith, R. D. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev. Vet. Med.* **45**, 23–41 (2000).
97. Kuhn, M. et al. caret: classification and regression training. https://cran.r-project.org/web/packages/caret/caret.pdf (2016).
98. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2009); https://doi.org/10.1007/978-0-387-98141-3
99. Kassambara, A. ggpubr: 'ggplot2' based publication ready plots. https://rpkgs.datanovia.com/ggpubr/index.html (2022).

Acknowledgements

We thank A. DeCleene, K. MacGregor, M. Gawith and M. Mitchell for their work as study nurses from Regional Public Health and all unnamed members of the SHIVERS-II team. We thank T. Hertz (Ben-Gurion University of the Negev) for insightful discussions and expert feedback on modeling methodology. We thank all consented enrollees and their families for their participation and commitment to the SHIVERS-II study. This publication was supported by the American Lebanese Syrian Associated Charities at St. Jude Children's Research Hospital (SJCRH), the SJCRH Center of Excellence for Influenza Research and Surveillance (P.G.T., R.J.W., Q.S.H.) contract HHSN272201400006C, US Department of Health and Human Services (HHS) contract 75N93021C00016 for the St. Jude Centers of Excellence for Influenza Research and Response, HHS contract 75N93019C00052 for the Center for Influenza Vaccine Research for High Risk Populations of the Collaborative Influenza Vaccine Innovation Centers, National Institute of Allergy and Infectious Diseases award 3U01AI144616-02S1 (P.G.T.), U01AI150747 (P.G.T.), R01AI154470 (P.G.T.), and Ruth L. Kirschstein National Research Service Award Individual Postdoctoral Fellowship award F32AI157296 (R.C.M.). The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

R.C.M. and A.S. contributed equally as co-first authors. L.-A.V.d.V. and K.V. contributed equally as co-second authors. Conceptualization: R.C.M., A.S., R.J.W., Q.S.H. and P.G.T. Formal analysis: R.C.M., A.S., L.-A.V.d.V. and K.V. Investigation: R.C.M., A.S., L.-A.V.d.V., K.V., E.K.A., C.M.K., S.T., J.D.B., T.L.W., D.G.S.J. and S.S.M. Methods development: R.C.M., A.S., L.-A.V.d.V., K.V., E.K.A., T.L.W., C.M.K., J.D.B. and S.T. Resources: T.W., L.J. and Q.S.H. Data and sample curation: T.W., L.J., Q.S.H. and the SHIVERS-II Investigation Team. Writing—original draft: R.C.M. and A.S. Writing—review and editing: R.C.M., A.S., L.-A.V.d.V., K.V., E.K.A., R.J.W., Q.S.H. and P.G.T. Visualization: R.C.M. and K.V. Supervision: R.J.W., Q.S.H. and P.G.T. Funding acquisition: R.C.M., R.J.W., Q.S.H. and P.G.T.

Competing interests

P.G.T. has consulted or received honoraria and/or travel support from Illumina, J&J, Pfizer and 10x Genomics. P.G.T. serves on the scientific advisory board of ImmunoScape and CytoAgents. The remaining authors declare no competing interests.

Additional information

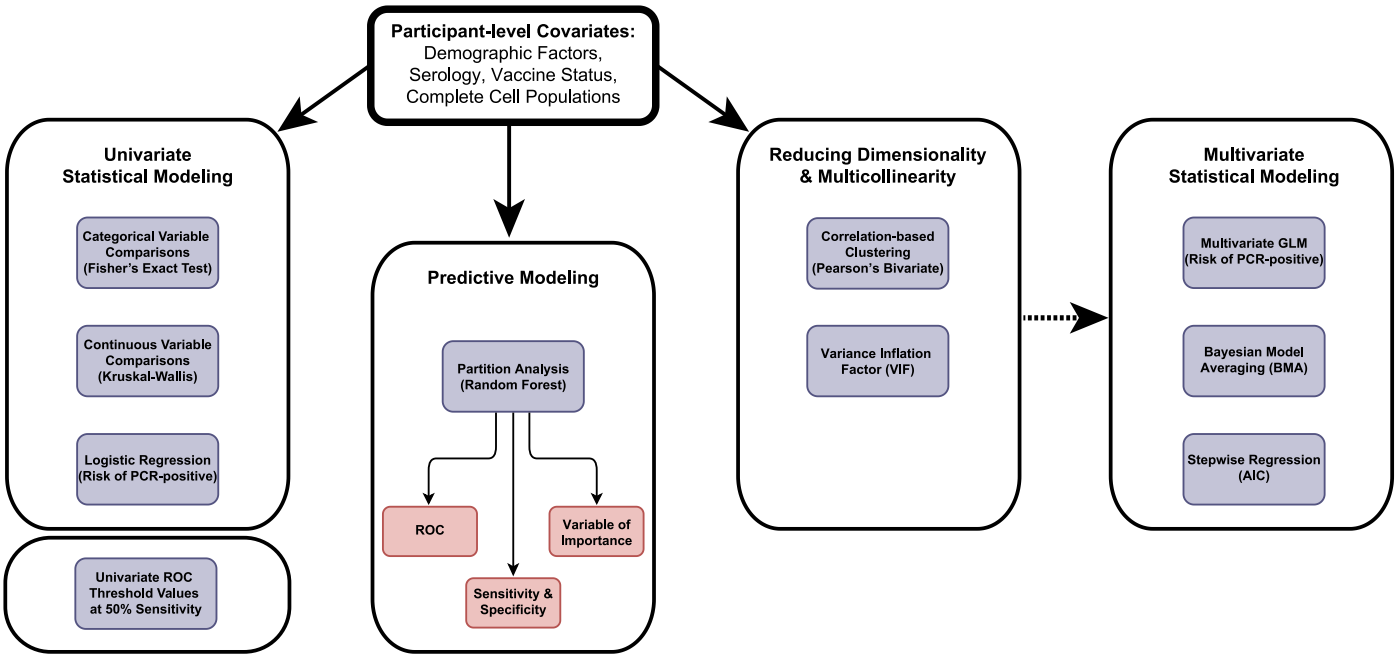
Extended data is available for this paper at https://doi.org/10.1038/s41590-023-01590-2.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-023-01590-2.

Correspondence and requests for materials should be addressed to Richard J. Webby, Q. Sue Huang or Paul G. Thomas.

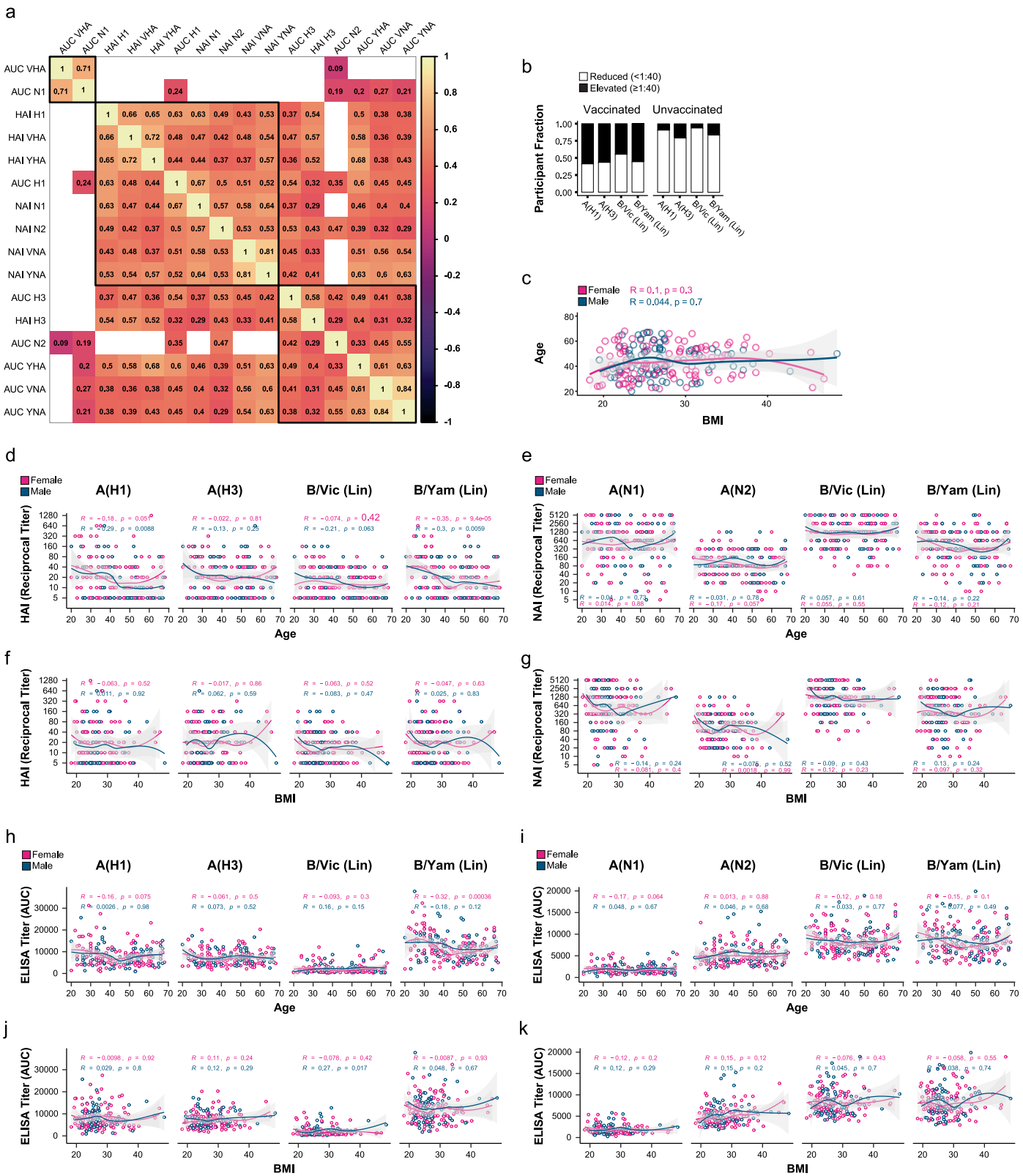
Peer review information *Nature Immunology* thanks Peter Openshaw, Sophie Valkenburg and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available. Primary Handling Editor: L.A. Dempsey, in collaboration with the *Nature Immunology* team.

https://doi.org/10.1038/s41590-023-01590-2



Extended Data Fig. 1 | Data analysis pipeline for predictive and statistical modeling. The analysis pipeline was designed to integrate participant-level demographic, serology, vaccine histories, and cellular flow cytometry data into statistical and predictive models. Univariate analyses, including Fisher's exact test, Kruskal-Wallis, logistic regression, and ROC thresholds were performed first on single, independent variables. These analyses help determine if an individual immune measure is statistically different between influenza virus infection and vaccination comparator groups (Fisher's exact test; Kruskal-Wallis), the risk of symptomatic influenza associated with an individual measure (logistic regression), and the threshold at which an individual immune measure can accurately describe 50% of symptomatic cases (ROC threshold). As univariate comparisons do not account for confounding factors, multivariate analyses were performed on combined variables including decision tree analysis (random forest) and logistic regression. The random forest allows comparison of performance (that is categorization accuracy) across models (ROC; Sensitivity & Specificity) as well as the relative importance of individual covariates within a model (VIP analyses). While random forest considers which models or individual covariates best categorize cases (symptomatic influenza) and controls (uninfected/cryptic), they do not provide information on association or risk. Multivariate generalized linear modeling (GLM) was used to determine the risk of symptomatic influenza associated with individual immune measures while accounting for the effects of others. The GLM was built on a select set of variables determined following reduction of dimensionality (correlation-based clustering) and multicollinearity (VIF) using stepwise regression (Akaike Information Criteria; AIC) and evaluated using Bayesian Model Averaging (BMA).

https://doi.org/10.1038/s41590-023-01590-2

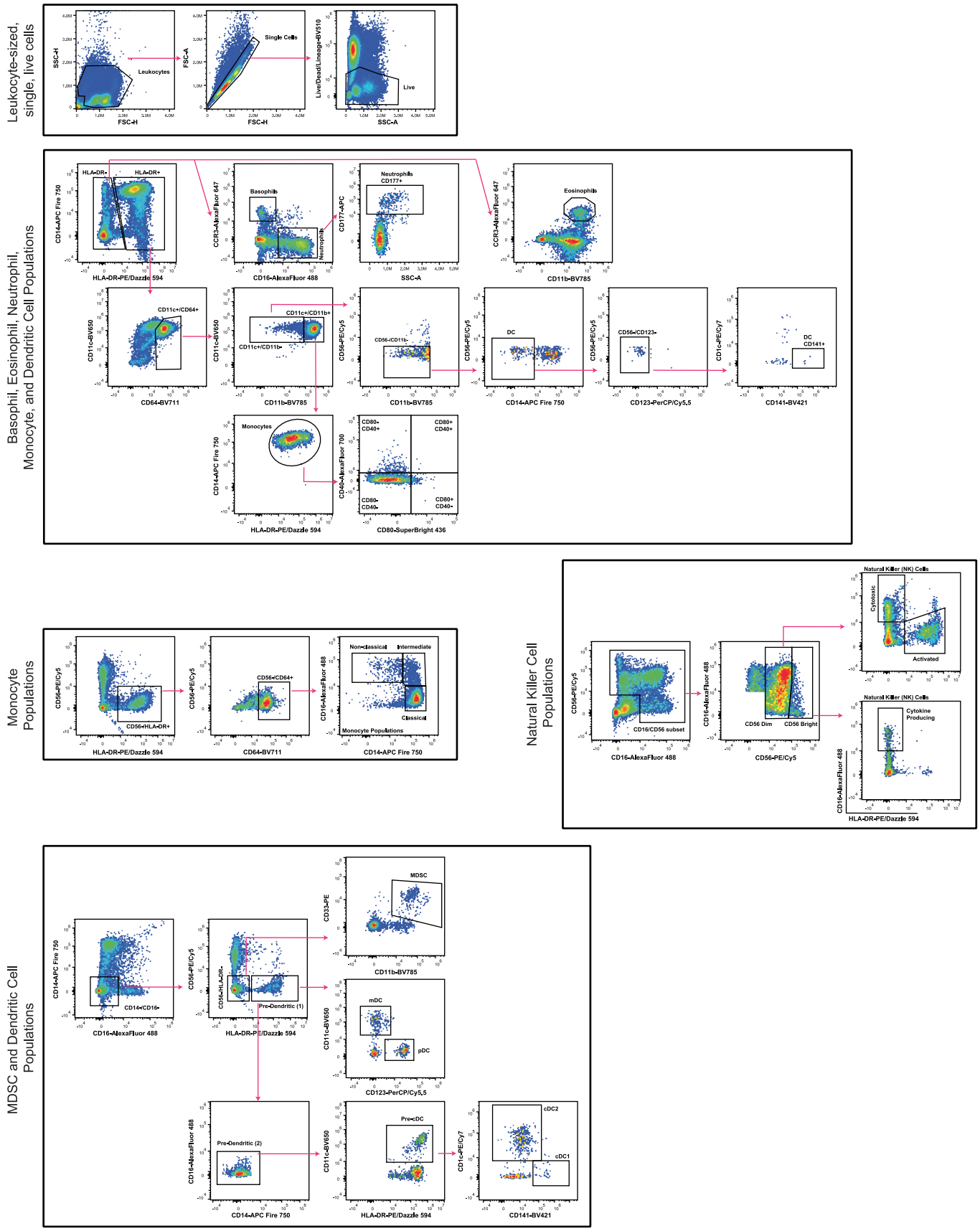


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41590-023-01590-2

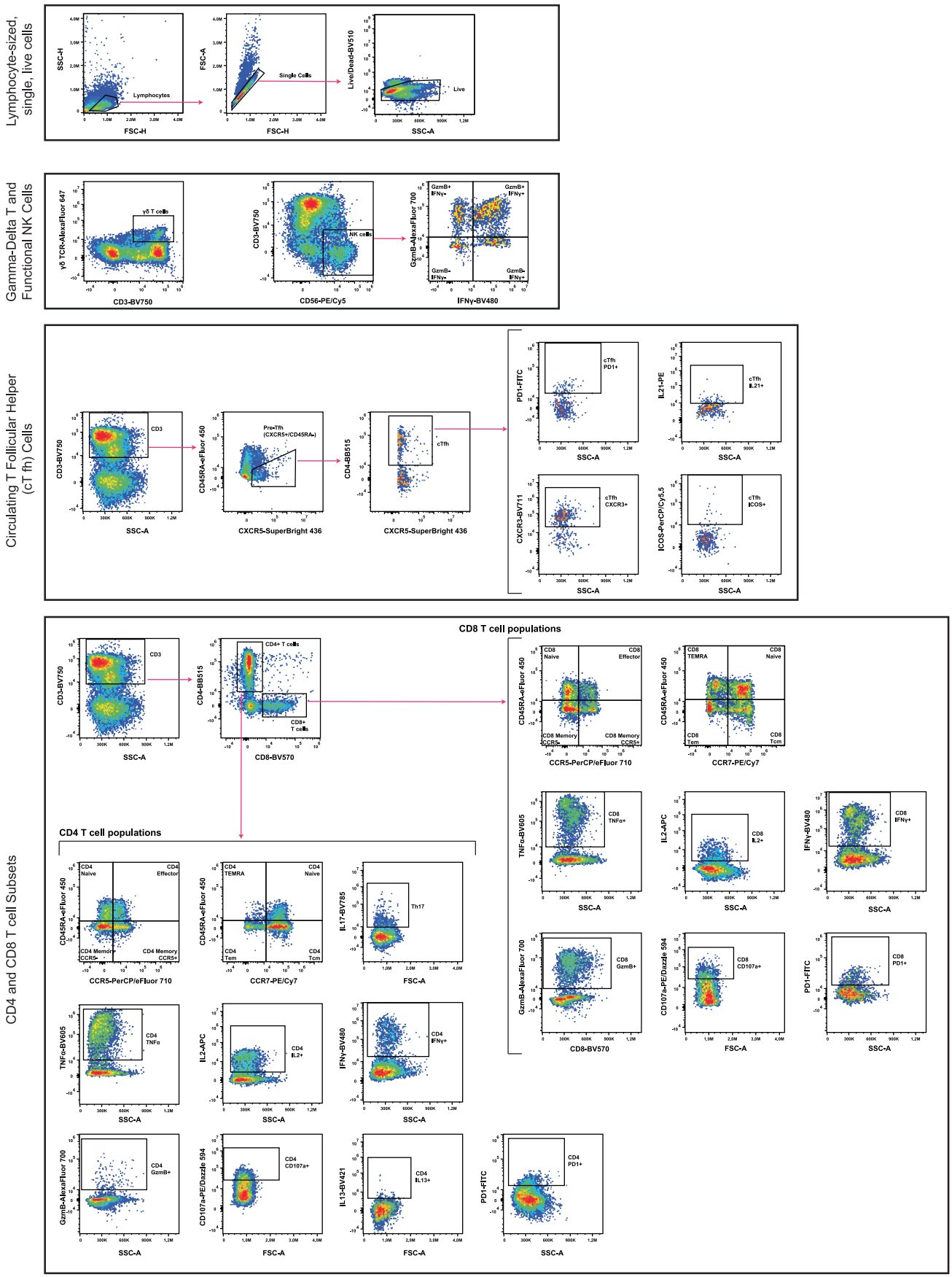
Extended Data Fig. 2 | Participant demographic and serologic correlations. **a)** Spearman Rank correlations between serology measures. Significant values (FDR-adjusted; $q \leq 0.05$) depicted with correlation coefficients and within correlation groups (black rectangles). Insignificant values blank. **b)** Frequency of unvaccinated or vaccinated study participants with baseline anti-HA and anti-NA antibody titers at elevated (≥1:40) or reduced (<1:40) levels for each influenza strain. **c)** Spearman Rank correlation (R; coefficient) between participant age (years) and BMI (kg m^{-2}) by sex. **d-k)** Spearman Rank Correlation (R; coefficient) between participant BMI (kg m^{-2}) and age (years) by baseline serologic measures stratified by sex. Reciprocal inhibiting antibody titer against **(d,f)** HA or **(e,g)** NA. Inhibiting titer calculated from HAI or NAI assays using A(H1N1), A(H3N2), B/Victoria (lineage), and B/Yamagata (lineage) viruses. Total **(h,j)** anti-HA or **(i,k)** anti-NA binding antibody titers. Total binding antibody titers reported as AUC values calculated from ELISA assay against purified, full-length HA or NA proteins derived from influenza A(H1N1), A(H3N2), B/Victoria lineage, and B/Yamagata lineage viruses. Regression analysis using locally estimated scatterplot smoothing (LOESS) method depicted with LOESS fit line (center line; smoothed local regression using least squares) and 95% CI (grey). Significant associations defined at $p \leq 0.05$.

https://doi.org/10.1038/s41590-023-01590-2



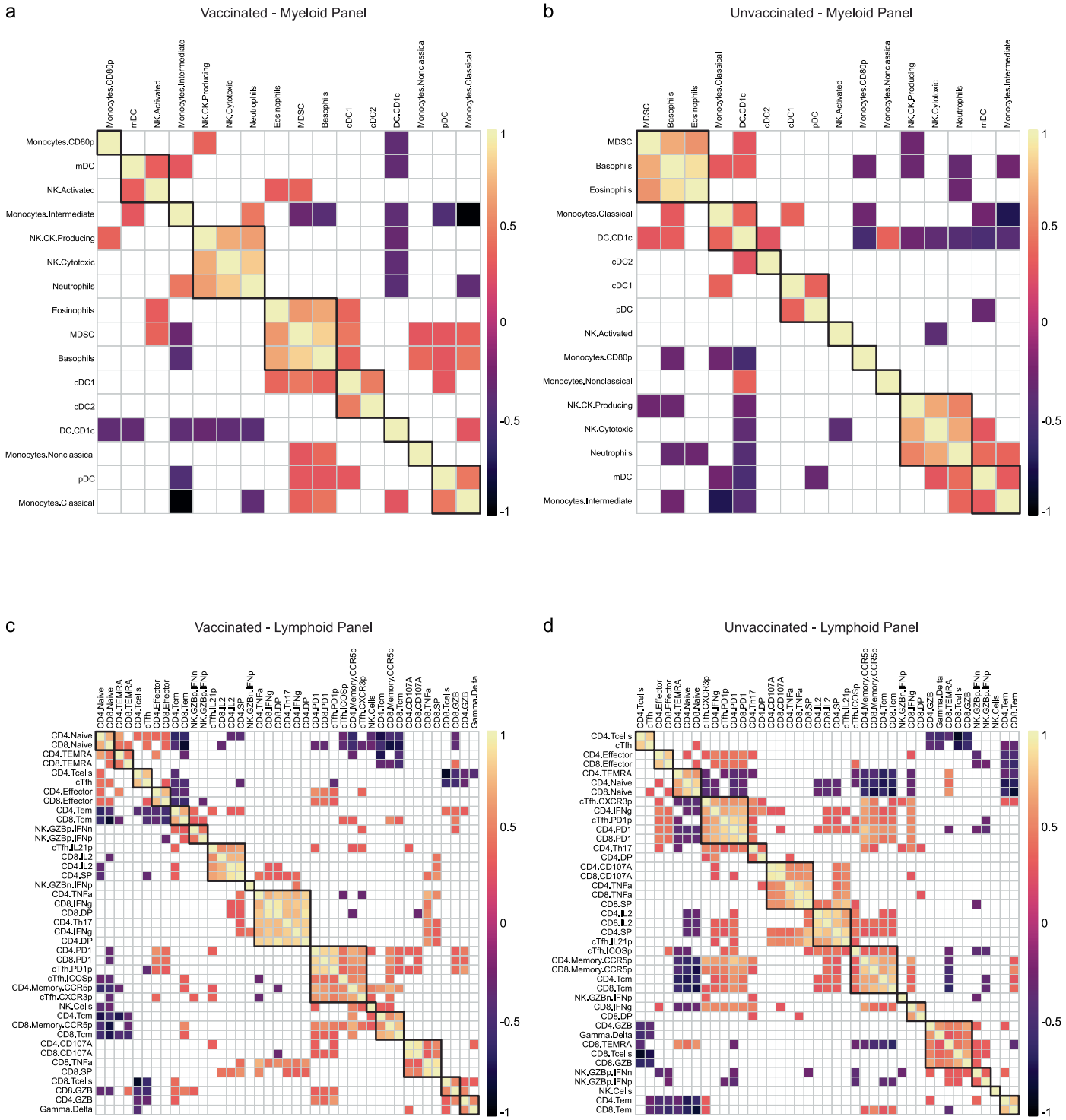
Extended Data Fig. 3 | Myeloid panel gating strategy. Flow cytometry gating strategy to resolve cell populations within the myeloid compartment. All gates applied to leukocyte-sized, single, live cells. Gates depict frequency as % of parent gate.

https://doi.org/10.1038/s41590-023-01590-2



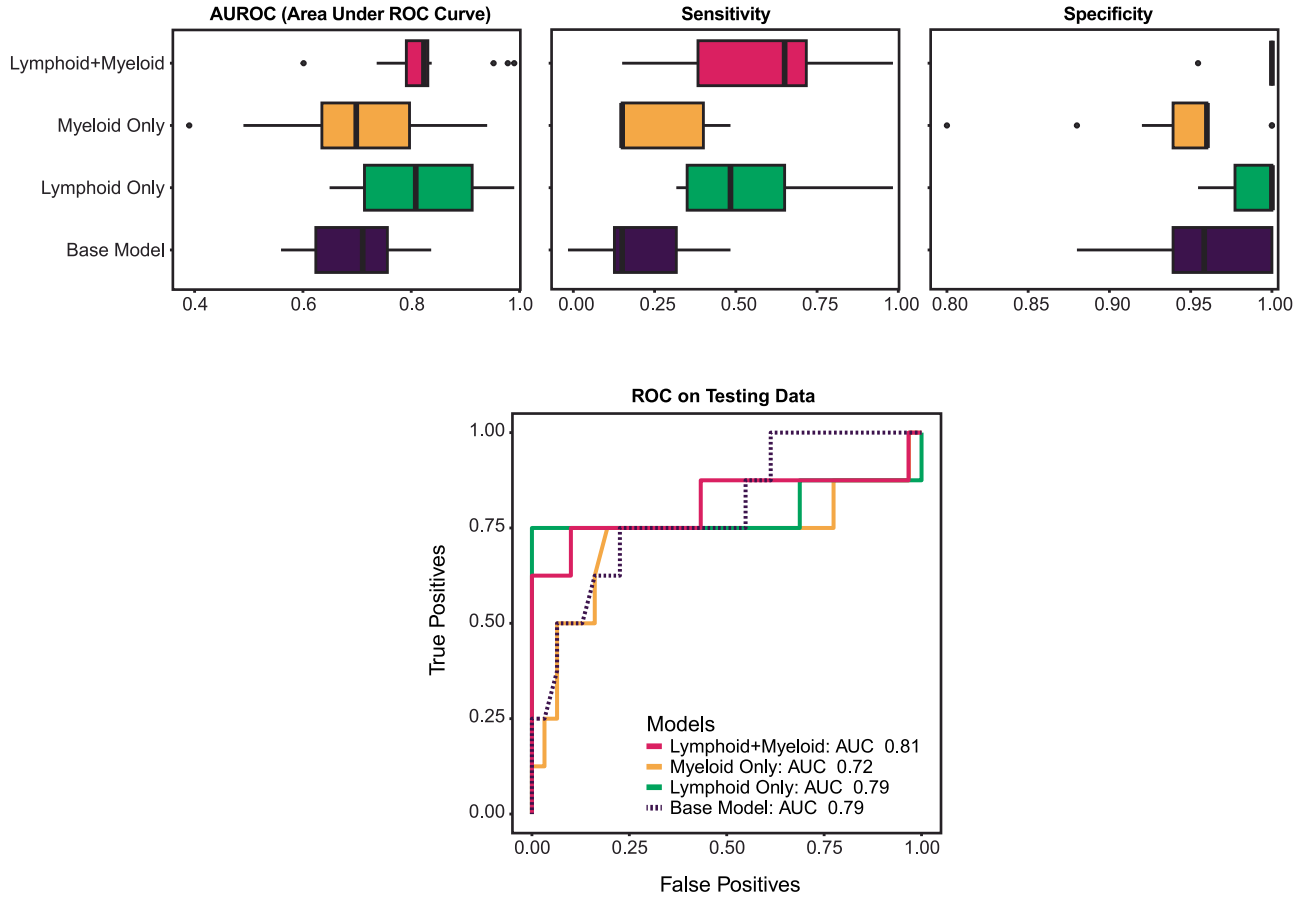
Extended Data Fig. 4 | Lymphoid and ICS panel gating strategy. Flow cytometry gating strategy to resolve cell populations within the lymphoid/functional compartment by ICS. All gates applied to lymphocyte-sized, single, live cells. Gates depict frequency as % of parent gate.

https://doi.org/10.1038/s41590-023-01590-2



Extended Data Fig. 5 | Co-regulated immune cell clusters by vaccine status. **a-d**, Co-regulated cell modules ('clusters') from Vaccinated participants' myeloid **(a)** and lymphoid/functional **(c)** cell populations, or Unvaccinated participants' myeloid **(b)** and lymphoid/functional **(d)** cell populations determined by average frequency (% parent) of individual cell populations with significant positive Pearson's bivariate correlation. Lymphoid/functional panel cell frequencies represent the average frequency (% parent) across virus (MOI = 4A/Michigan/45/2015 H1N1pdm09 or A/Singapore/INFIMH-16-019/2016 H3N2) and peptide (1–5 µM /peptide pools containing M1, NP, PB1) stimulation groups. *P* values were adjusted using false discovery rate (FDR; *q*) correction for multiple comparisons with significance $q \leq 0.05$ denoted with color; not significant (blank).

https://doi.org/10.1038/s41590-023-01590-2



Extended Data Fig. 6 | Decision tree model comparison from cellular covariates. Comparison of the Base (demographic factors + serology + vaccination status), Myeloid Only (myeloid panel cell populations), Lymphoid Only (lymphoid/functional panel cell populations), and Lymphoid+Myeloid (cell populations from the lymphoid/functional and myeloid panel) random forest models built to categorize symptomatic and uninfected/cryptic influenza cases. Participants were split 80:20 into a training set (symptomatic cases n = 31, uninfected/cryptic controls n = 128) and testing set (symptomatic cases n = 8, uninfected/cryptic controls n = 33) ensuring equal proportions of cases and controls. Models were trained, tested, and cross-validated using 10× CV-10. Sensitivity, Specificity and AUROC (area under the receiver-operating characteristic curve) provided. An out-of-sample evaluation of the models (bottom) shows a comparison of the AUC accuracy. Boxes represent the median and 25th to 75th percentiles; whiskers indicate the minimum and maximum values no further than 1.5 times the interquartile (IQR).

nature portfolio

Corresponding author(s): Paul G. Thomas, PhD
Q. Sue Huang, PhD
Richard J. Webby, PhD

Last updated by author(s): Jul 10, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Flow cytometery data were collected using SpectroFlow v2.2 software (Cytek). All other biologic data were collected without software.
Data analysis	All statistical analyses were performed in R v3.6.0, including sample size assessments (WMWssp v0.4.0), univariate statistical modeling (sjTabone 0.1.0), multivariate predictive modeling (caret v6.0-92), and Bayesian model averaging (BMA v3.18.17) except calculation of area under the curve (AUC) values for ELISA, which were determined with GraphPad Prism 9 (v9.5.1 (528)) using the following parameters: baseline y=0; minimum peak height <10% of the distance from min to max Y; all peaks above baseline; 5 significant digits. Visualization in R was performed using ggplot2 v3.4.0 and ggpubr v0.5.0. Flow cytometry data were analyzed using FlowJo v10.7.1 (TreeStar). Figures were assembled using Adobe Illustrator 2023 or created with BioRender.com (exported under a paid subscription with an associated publication license). A detailed description of each analysis can be found in the "Methods" section of the manuscript. A minimum dataset containing de-identified study participant information and biological assay results along with custom study-generated R code for analysis was uploaded to GitHub (https://github.com/kvegesan-stjude/SHIVERS2) per the data sharing agreement stipulated under the NRSA-NIAID Individual Postdoctoral Fellowship award number F32AI157296 (R.C.M). Additional basic R code can be made available upon reasonable request. The published article includes all datasets generated or analyzed as a part of this study. Individual Source Data are provided with associated figures (where appropriate) per the data sharing agreement. Raw flow cytometry source files can be made available upon reasonable request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The published article includes all datasets generated or analyzed as a part of this study. Individual Source Data are provided with associated figures (where appropriate) per the data sharing agreement stipulated under the Ruth L. Kirschstein National Research Service Award (NRSA) Individual Postdoctoral Fellowship (NRSA-NIAID) award number F32AI157296 (R.C.M). Raw flow cytometry source files can be made available upon reasonable request.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Participant sex (assigned at birth) is reported as a biologic attribute. Attention was paid to participant sex during the design of the study and sample selection so rough equivalent Male and Female participants were represented. Sex was considered in the correlation and univariate analyses, and as a covariate in the multivariate modeling. Results reported apply to all sexes, unless specifically stated. Information on comparator groups (vaccine and influenza infection status) by sex can be found in Figure 1c-1d, Table 1, and Supplementary Table 2.
Population characteristics	Participant age (in years or as age group), sex (biologic attribute; assigned at birth), body-mass index (kg/m^2), ethnicity (participant-reported), 2018 influenza virus infection status (by PCR-confirmed infecting strain), and 2018 influenza vaccine status were collected and are reported in Table 1.
Recruitment	During year 1 (2018) of the SHIVERS-II study, more than 22,000 individuals aged 20-69 years were randomly selected for recruitment from 3 participating primary health organizations' healthy patient networks in Wellington, New Zealand, with 2,195 ultimately enrolled following written informed consent. The study staff identified prospective adults aged 20-69 years through random selection from those healthy individuals listed in the management systems of selected primary care general practices. The study staff mailed the invitation and study information packet to those selected individuals. Those interested individuals provided their consent through online electronic consent form and also filled in an online questionnaire by providing demographic and contact information, vaccination and health and influenza-like illness status. For those consented participants, study activities included blood/swab collections. Each participant received a NZ$30 gift card after each blood or swab sample collection to recognize their time and effort.
Ethics oversight	This study received ethics approval from the New Zealand Health & Disability Ethics Committee (NTX11.11.102.AM36 & AM36 & AM49 & AM51). All participants provided written consent to participate in the study. De-identified participant samples (PBMC, serum) and associated demographic information were provided under the St. Jude Children's Research Hospital Center of Excellence for Influenza Research and Surveillance NIAID contract #HHSN272201400006C.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	A total of 206 participants were selected from year 1 of the SHIVERS-II study comprising vaccinated-uninfected (n=75), vaccinated-infected (n=33), unvaccinated-uninfected (n=76), and unvaccinated-infected (n=22). Subjects' age (in years) and sex (assigned at birth) were roughly matched when selecting across comparator groups. Sample size requirements for statistical testing of the covariates were calculated using the R package WMWssp v0.4.0 with a defined power of 0.8 and were determined to be sufficient in size for the planned comparisons.
Data exclusions	Following collection of flow cytometry data we applied several blinded quality control measures for unbiased selections to identify cell populations to include in our primary analysis. These included filtering out individual samples with low cell viability (5% and below) and/or high cellular debris (>25% of events), and eliminating cell populations with a limited dynamic range in cell frequency. In the downstream statistical and modeling analyses, subjects with significant cell frequency outliers consistent across multiple cell population types were determined by Grubbs' test and excluded where indicated (n=6 removed).

Replication

The PI and research team maintain practices and atmosphere that place high value on scientific rigor and reproducibility. This is accomplished by appropriate study design, by performing biological replicates as well as technical replicates when possible, implementing good practice in data analysis, and by statistical evaluation. Due to the limited volumes of blood collected from the human subjects, technical replication of the cellular analyses were not possible. However, values collected from flow cytometry or anti-influenza antibody titers were compared to other human cohorts of similar composition to verify that ranges were appropriate. Rigorous statistical methods were performed on all data and reported as such; adjustments were made for multiple comparisons where applicable. For random forest modeling, an 80:20 (train:test) split was performed on the 200 (outliers removed as indicated above) participants while ensuring equivalent proportions of cases and controls were included. Models underwent 5-fold cross-validation, which was repeated 3 times for each model. All reagents and methods are described in detail to best support reproducibility.

Randomization

Sample selections were based on rough age- and sex- matching across the infection and vaccination comparator groups to ensure similar numbers in each. Sample order was randomized prior to cellular and serology assays to limit batch effects.

Blinding

Individual sample blinding was not performed during sample selection as the comparator groups required rough age, sex, and sample number inclusion equivalents. Subject-level covariates were considered in the univariate and multivariate models and therefore were unblinded. Following flow cytometry, information regarding sample and cell population were blinded during quality control steps and selection of cell populations for primary analysis. During the multivariate modeling, computer-generated covariate selection (blinded; unbiased) was compared to covariates selected by the investigators; comparison confirmed the investigator-selected covariates appropriately fit the data.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Flow cytometry surface antibodies used: CCR3 AF647 (BioLegend cat# 310710, clone 5E8, lot# B312457, dilution 1:50), CCR5 PerCP/eFluor710 (ThermoFisher cat# 46-1951-82, clone 7A4, lot# 4324286, dilution 1:80), CCR7 PE/Cy7 (BioLegend cat# 353226, clone G04H7, lot# B305236, dilution 1:20), CD1c PE/Cy7 (BioLegend cat# 331516, clone L161, lot# B256006, dilution 1:20), CD3 BV750 (BioLegend cat# 344846, clone SK7, lot# B303395, dilution 1:200), CD3 BV510 (BioLegend cat# 300448, clone UCHT1, lot# B281774, dilution 1:200), CD4 BB515 (BD cat# 565996, clone SK3, lot# 9343297, dilution 1:200), CD8 BV570 (BioLegend cat# 301038, clone RPA-T8, lot# B281322, dilution 1:50), CD11b CD785 (BioLegend cat# 301346, clone ICRF44, lot# B315938, dilution 1:100), CD11c BV650 (BioLegend cat# 337238, clone Bu15, lot# B304739, dilution 1:20), CD14 APC/Fire750 (BioLegend cat# 367120, clone 63D3, lot# B257669, dilution 1:100), CD16 FITC (BioLegend cat# 360716, clone B73.1, lot# B300038, dilution 1:100), CD19 BV510 (BioLegend cat# 302242, clone HIB19, lot# B281769, dilution 1:200), CD33 PE (BioLegend cat# 303404, clone WM53, lot# B288411, dilution 1:20), CD40 AF700 (BioLegend cat# 334328, clone 5C3, lot# B272227, dilution 1:20), CD45RA V450 (ThermoFisher cat# 48-0458-42, clone HI100, lot# 2263338, dilution 1:200) or BV570 (BioLegend cat# 304132, clone HI100, lot# B319894, dilution 1:20), CD45RO APC/Fire 750 (BioLegend cat# 304250, clone UCHL1, lot# B288981, dilution 1:20), CD56 PE-Cy5 (BioLegend cat# 362516, clone 5.1H11, lot# B317268, dilution 1:100), CD64 BV711 (BD cat#740782, clone 10.1, dilution 1:20), CD80 SuperBright 436 (ThermoFisher cat# 62-0809-42, clone 2D10.4, lot# E113289A, dilution 1:20), CD103 BV605 (BioLegend cat# 350217, clone Ber-ACT8, lot#B359323, dilution 1:20), CD107a BV711 (BioLegend cat# 328640, clone H4A3, lot# B301430, dilution 1:50) or PE/Dazzle 594 (BioLegend cat# 328646, clone H4A3, lot# B319152, dilution 1:20), CD123 PE (BioLegend cat# 306006, clone 6H6, lot# B269700, dilution 1:100) or PerCP/Cy5.5 (BioLegend cat#306016, clone 6H6, lot# B318963, dilution 1:20), CD141 BV421 (BioLegend cat# 344144, clone M80, lot# B315574, dilution 1:100), CD177 APC (BioLegend cat# 315808, clone MEM-166, lot# B265153, dilution 1:50), CXCR3 PE/Dazzle594 (BioLegend cat# 353736, clone G025H7, lot# B289168, dilution 1:100) or BV711 (BioLegend cat# 353732, clone G025H7, lot# B264427, dilution 1:20), CXCR5 SuperBright436 (ThermoFisher cat# 62-9185-42, clone MU5HBEE, lot# 2279709, dilution 1:100), HLA-A2 BV650 (BioLegend cat# 343324, clone BB7.2, lot# B290341, dilution 1:200), HLA-DR AF700 (BioLegend cat# 307626, clone L243, lot# B272227, dilution 1:100) or PE/Dazzle 594 (BioLegend cat#307654, clone L243, lot# B320167, dilution 1:20), ICOS PerCP/Cy5.5 (BioLegend cat# 313518, clone C398.4A, lot# B297477, dilution 1:20), GhostDye Violet 510 Viability Dye (Tonbo cat# 13-0870-T500, lot# D0870061322133, dilution 1:400), PD1 FITC (BioLegend cat# 329904, clone EH12.2H7, lot# B253784, dilution 1:100), Siglec-8 BV480 (BD cat# 747874, clone 837535, lot# 1006002, dilution 1:50), and TCRγδ AF647 (BioLegend cat# 331214, clone B1, lot# B274274, dilution 1:80)

Flow cytometry intracellular staining antibodies used: CD68 PerCP/Cy5.5 (BioLegend cat# 333814, clone Y1/82A, lot# B273670, dilution 1:50), Granzyme B AF700 (BioLegend cat# 372222, clone QA16A02, lot# B321316, dilution 1:25), IL2 APC (ThermoFisher cat# 17-7029-82, clone MQ1-17H12, lot# 2172935, dilution 1:25), IL13 BV421 (BioLegend cat# 501916, clone JES10-5A2, lot# B319799, dilution 1:25), IL17 BV785 (BioLegend cat# 512338, clone BL168, lot# B310823, dilution 1:25), IL21 PE (BioLegend cat# 513004, clone 3A3-N2, lot# B299502, dilution 1:25), IFNγ BV480 (BD cat# 566100, clone B27, lot# 0261295, dilution 1:10), and TNFα BV605

(BioLegend cat# 502936, clone Mab11, lot# B312493, dilution 1:25)

ELISA secondary antibody used: anti-Human IgG (Fab specific)-HRP secondary antibody (Sigma cat #A0293, polyclonal, lot# 0000201676, dilution 1:3000)

Validation

All antibodies were purchased from commercial suppliers including BD, BioLegend, Tonbo, ThermoFisher, Sigma, and eBiosciences with validation data and applicable citations available on product listings for all antibodies (see individual catalog numbers). Antibodies that have previously been validated in the literature were preferred and used at specified dilutions or according to the manufacturer's specifications.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Cryopreserved aliquots of PBMCs were thawed at 37°C, suspended in RPMI-1640 supplemented with 10% heat-inactivated FBS (Gibco, cat# 16140071), 1% non-essential amino acids (Gibco, cat# 11140-050), 1 mM sodium pyruvate (Gibco, cat# 11360-070), and 1% penicillin-streptomycin (10,000 U/mL penicillin and 10,000 ug/mL streptomycin; Gibco, cat# 15140-122). Cells were plated at 3.0-4.0x10^5 cells/well in a 96-well U-bottom plate.

Instrument

Cytek 3-laser Aurora spectral flow cytometer

Software

SpectroFlow v2.2 software (Cytek) was used to collect data, which were analyzed using FlowJo v10.7.1 (TreeStar)

Cell population abundance

Following collection of flow cytometry data we applied several blinded quality control measures for unbiased selections to identify cell populations to include in our primary analysis. These included filtering out individual samples with low cell viability (5% and below) and/or high cellular debris (>25% of events), and eliminating cell populations with a limited dynamic range in cell frequency. Where possible, individual cell population frequencies were compared to known frequencies in published human cohorts.

Gating strategy

The flow cytometry gating strategy to resolve cell populations within the myeloid compartment is presented in Extended data figure 3 and Supplementary Table 7. All gating applied to leukocyte-sized, single, live cells and depict frequency as % of parent gate. The flow cytometry gating strategy to resolve cell populations within the lymphoid compartment by ICS is presented in Extended data figure 4 and Supplementary Table 7. All gates applied to lymphocyte-sized, single, live cells. Gates depict frequency as % of parent gate.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.