Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Digital Twins of Pregnancy from Public Single-Cell Data: An In-Silico Framework for Predicting Maternal Immune Responses to Infection and Vaccination

Tingting Zhai

zhaitingting@cpu.edu.cn

China Pharmaceutical University

Research Article

Keywords: digital twin, pregnancy, maternal immunity, single-cell RNA sequencing, interferon signaling, COVID-19, maternal vaccination, in-silico intervention

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11799717/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Background

Pregnant women face elevated risks of severe respiratory viral infection, yet WHO reports only 4% of trials in the past decade allowed their enrollment, leaving prevention and treatment unguided. We tested whether a digital twin built from limited public data could evaluate maternal immune responses to infection and vaccination.

Methods

From public single-cell RNA sequencing of 181,112 peripheral blood mononuclear cells from 26 donors with severe, asymptomatic, or no COVID-19, we built a Harmony-integrated atlas of 14 immune cell types and donor-level response maps (composition, differential expression, pathway modules). An L1-regularized logistic model read out pregnancy status and severity (leave-one-out cross-validation). The COVID-derived interferon signature was tested in independent influenza A infection and vaccination cohorts; mother-to-child transfer was quantified from published maternal–cord pairs, and the vaccine-target epitope was mapped onto the viral receptor structure. In-silico restoration experiments asked which immune programs normalize the pregnant twin.

Results

The model detected pregnancy from immune cell states alone (cross-validated AUROC 0.87), driven by dampening of cytotoxic, interferon-stimulated, and antigen-presentation programs with regulatory T-cell expansion; all 15 interferon-stimulated genes were down-regulated in both severe and asymptomatic infection (adjusted $p < 0.05$). Severity prediction failed (AUROC 0.21) owing to donor-batch confounding, a negative result reported deliberately. The signature transferred to influenza A infection and rose one day after vaccination in 44 pregnant women (+ 0.91 SD). In 129 mother–newborn dyads, cord-blood antibodies tracked maternal titers (Spearman $\rho = 0.80$), with higher transfer efficiency after second- than third-trimester vaccination, and the platform's vaccine-target epitope mapped onto the viral receptor footprint. Restoring the full cell-state program normalized all six pregnant twins; no single program did.

Conclusions

Public data support a digital twin of maternal immunity that functions as a computational alternative to animal models, defines what future studies should measure, and charts a reusable path to first-in-pregnancy vaccine evaluation.

Introduction

Pregnancy reprograms maternal immunity to tolerate the fetus, and the reprogramming comes at a cost: pregnant women with COVID-19 face higher risks of severe disease than matched non-pregnant women. Yet the cell-state mechanisms distinguishing a pregnant immune response remain incompletely characterized, because the trials and preclinical models that would resolve them exclude pregnant women: WHO now calls their inclusion in clinical research "no longer optional", targeting routine ethical inclusion by 2030[1], yet only 4% of trials registered in the past decade allowed their enrollment and pregnancy-specific evidence remains scarce. The practical path, until inclusion becomes routine, is to extract maximal insight from the public single-cell atlases that already exist — which requires a framework for turning them into predictive, testable models.

Our own bench program illustrates the impasse. Echovirus 18 is a leading cause of aseptic meningitis in infants, who can deteriorate within days of birth; with no approved vaccine, the realistic prevention strategy is maternal immunization — vaccinate the pregnant woman so that transplacental IgG protects the neonate. We developed a needle-free, floss-mediated gingival immunization platform presenting an HBc-E18-3 virus-like particle, which induces durable, FcRn-dependent intestinal IgG in mice[2]. Translating the platform to maternal immunization requires predicting how the pregnant immune system responds to a mucosal vaccine — a question mice cannot answer: mice are non-permissive to Echovirus 18, which uses the human neonatal Fc receptor as its dual attachment-and-uncoating receptor[3], maternal antiviral responses are poorly modeled in animals, and lymphoid-organoid systems cannot assess systemic immunity. The U.S. FDA's April 2025 plan to phase out animal-testing requirements for monoclonal antibodies and other drugs, replacing them with "New Approach Methodologies" that include AI-based computational models, makes this gap regulatory as well as scientific[4].

Scott recently articulated a blueprint for "building digital pregnancies from the single cell up": construct a single-cell perturbation atlas, transfer it to patient cells, and fit a model linking cell states to clinical outcomes[5]. The observational ingredients already exist. Maternal autoantibody signatures associated with preterm birth[6], compartment-specific soluble immune profiles of adverse birth outcomes[7], maternal metabolomic risk scores for preterm birth[8], dynamic prediction of preterm birth from longitudinal electronic health records[9], longitudinal destabilization of the maternal gut virome[10], and immunological maladaptation detectable weeks before spontaneous preterm birth[11] together establish that maternal multi-omic states precede and predict clinical outcomes. What has not been demonstrated is the infection-immunology half of the blueprint: a digital twin that reads out a maternal immune response to acute infection, validates externally across pathogens, and supports in-silico intervention.

Here we show that this half of the blueprint can be executed end-to-end from existing public data. We built a re-analyzed, Harmony-integrated cell-state atlas of 181,112 peripheral blood mononuclear cells from 26 donors spanning pregnancy and COVID-19 severity; derived donor-level response maps showing that pregnancy blunts the blood interferon response and reshapes immune composition; fit a sparse, interpretable model that detects pregnancy status from immune cell states (cross-validated AUROC 0.87) and honestly reports a negative result for severity; validated the COVID-derived signature in an

independent influenza A cohort; and ran a multi-arm “treated twin” restoration experiment identifying which immune programs must change to normalize the pregnant twin. The pipeline is fully scripted and reusable.

Methods

Data sources

All data are publicly available. The discovery cohort, GSE239450/GSE239452, comprises CITE-seq of peripheral blood mononuclear cells (PBMCs) from 26 donors: 9 pregnant (3 severe/critical, 3 asymptomatic, 3 healthy control) and 17 non-pregnant (11 severe, 6 asymptomatic) women with SARS-CoV-2 infection[12]. These 26 donors are the complete single-cell subset of the source study’s 226-participant cohort (152 pregnant, 74 non-pregnant); the remaining participants contributed flow-cytometry, plasma cytokine, and stool metagenomic measurements not part of the public single-cell deposit. We used the author-processed object (log-normalized expression, raw counts, pre-computed principal components, Harmony-corrected principal components, and UMAP embeddings) and re-processed raw CellRanger outputs for 11 samples. The external validation cohort, GSE243629, is a single-cell atlas of the peripheral immune response to influenza A virus infection: 102,892 cells from 11 donors in five groups — healthy controls (n = 2), infected adults (n = 2), pregnant controls (n = 2), pregnant infected (n = 3), infected children (n = 2)[13].

Atlas construction and annotation

For the discovery cohort we re-clustered cells per lineage (T/NK, myeloid, B/plasma) on Harmony-corrected components using Leiden clustering (resolution 1.2), then merged and refined clusters into 14 cell types by canonical expression (Supplementary Methods). Donor was the batch unit throughout, and all downstream modeling is donor-level. For the influenza cohort, raw CellRanger matrices were quality-controlled (200–5,000 genes per cell; mitochondrial fraction ≤ 25%), log-normalized, reduced to highly variable genes, embedded by principal-component analysis, clustered, and annotated by maximum marker score across canonical marker sets.

Response maps

Cellular composition. Per-donor cell-type proportions were compared across the five clinical groups by Kruskal–Wallis tests; abundance changes for pregnant versus non-pregnant donors with severe COVID-19 are reported as log2 fold changes. *Pseudobulk differential expression.* Per-donor, per-cell-type counts were aggregated and tested with PyDESeq2, contrasting pregnant versus non-pregnant donors within severe COVID-19 and within asymptomatic COVID-19; Benjamini–Hochberg-adjusted $p < 0.05$ was called significant. *Pathway module scores.* Six gene modules (IFN-stimulated, IL27 axis, monocyte activation, cytotoxicity, antigen presentation, exhaustion checkpoints) were scored per cell and averaged per donor per cell type (gene sets in Supplementary Methods).

Outcome models

Donor-level features combined cell-type proportions (14) and module scores (14 cell types × 6 modules; 98). We fit L1-regularized logistic regression (C = 0.5, liblinear, standardized features) for two binary readouts — pregnancy status (23 infected donors, 6 pregnant) and severity — with leave-one-out cross-validation and area under the receiver-operating-characteristic curve (AUROC) as the metric. Feature-block variants (composition only, modules only, all) localized the signal.

External validation

The 15-gene pregnancy-dampened interferon signature was scored in the influenza cohort after filtering to detected genes. Donor-mean scores were compared between infected and control groups (Mann–Whitney U), and NK-cell fractions were computed per donor. For vaccine responsiveness, the same signature was scored in a public whole-blood microarray cohort of 44 pregnant women sampled before and 1 and 7 days after inactivated influenza vaccination (GSE166545)[14]: probe-level expression was z-scored across samples, gene scores averaged over probes, and paired changes tested with Wilcoxon signed-rank tests.

Mother-to-child transfer and structural mapping

Mother-to-child antibody transfer was quantified from the published paired maternal–cord plasma measurements accompanying Atyeo et al.[15] (source data file 2): dyads with one maternal and one cord sample and maternal optical density ≥ 0.2 (detectable seropositivity; n = 129) were analyzed. The transfer ratio was defined as cord divided by maternal Spike-specific IgG; Spearman correlation quantified concordance; ratios were grouped by vaccination trimester and platform. For structural mapping, the E18-3 epitope (37 VP1 residues) was located in the cryo-EM structure of the echovirus 18 protomer in complex with the neonatal Fc receptor (PDB 9TF0[16]) by sequence alignment; per-residue minimum distances to receptor atoms were computed, and receptor-contact residues were taken from the structure-based analysis of Chen et al.[3].

Treated-twin simulation

For each pregnant donor, feature blocks were linearly interpolated toward the non-pregnant class mean in original units — interferon modules only, all immune modules, cellular composition only, or the full feature program — at restoration fractions of 0, 0.25, 0.5, 0.75, and 1.0, and the fitted pregnancy model re-scored each simulated donor. This in-sample, hypothesis-generating experiment interrogates which feature blocks carry the pregnancy signal, not the effect of any real treatment.

Statistical analysis

All tests were two-sided. Differential expression used Benjamini–Hochberg adjustment at 0.05. Composition and validation analyses are exploratory: we report nominal p values and point estimates

without multiplicity adjustment. Analyses used Python (scanpy, anndata, PyDESeq2, scikit-learn); full software versions and quality-control thresholds are in the Supplementary Methods.

Results

A cell-state atlas spanning pregnancy and COVID-19 severity

We assembled 181,112 PBMCs from 26 donors in five clinical groups and re-clustered them into 14 immune cell types (Fig. 1). Harmony correction removed donor-level batch structure while preserving biology, and author annotations were independently reproduced. Because every non-pregnant donor was processed as its own batch while pregnant donors were hashtag-pooled, all downstream analyses operate at the donor level (n = 26; 23 infected).

Pregnancy blunts the blood interferon response and reshapes composition

Pseudobulk differential expression contrasting pregnant versus non-pregnant donors revealed a coherent signature: all 15 interferon-stimulated genes queried (ISG15, IFIT1/2/3, MX1/2, OAS1/2/3, IFI6, IFI44L, RSAD2, IFI27, CXCL10, IFITM3) were significantly down-regulated in pregnant donors, in severe COVID-19 (log2 fold change − 2.65 [IFITM3] to − 5.78 [IFIT1]) and again in asymptomatic infection (− 1.43 [OAS2] to − 7.30 [IFI27]; all adjusted $p < 0.05$) (Fig. 2A, B). The blunted interferon state is thus a property of pregnancy, not of disease severity.

Cellular composition shifted in parallel (Fig. 2C, D): across the five groups, γδ T cells ($p = 0.020$), NK cells ($p = 0.028$), proliferating T/NK cells ($p = 0.041$), and plasmacytoid dendritic cells ($p = 0.043$) differed by Kruskal–Wallis, with pregnant severe donors showing depletion of NK (log2 fold change − 1.21), γδ T (− 1.18), and pDC (− 2.50) and expansion of regulatory T cells (+ 1.65, $p = 0.062$) relative to non-pregnant severe donors. These abundance changes do not survive correction across cell types (adjusted $p = 0.149$); we therefore treat composition as supporting, not primary, evidence.

Pregnancy status is readable from immune cell states — severity is not

A sparse L1 logistic model read out pregnancy status from donor-level immune features with cross-validated AUROC 0.873 (Fig. 3A). The signal is carried by cell states, not composition: restricting features to module scores preserved AUROC 0.873, whereas composition alone gave 0.814. The full fit used only 5 of 98 features — lower cytotoxicity in CD4 + T cells (weight − 0.72) and γδ T cells (− 0.24), lower interferon-stimulated activity in dendritic cells (− 0.51), lower antigen-presentation activity in CD4 + T cells (− 0.26), and higher regulatory T-cell fraction (+ 0.18) (Fig. 3B) — a coherent picture of coordinated dampening of effector and interferon programs with regulatory expansion.

In contrast, severity prediction failed outright (AUROC 0.214; all feature-block and regularization variants 0.21–0.54) — a negative result we report deliberately. With 23 infected donors, every non-pregnant donor constitutes its own processing batch, and severity is confounded with batch; a model that appears to predict severity on such data would more likely encode cohort structure than biology.

The signature transfers across pathogens

If the pregnancy-dampened interferon state is real rather than a cohort artifact, it should appear in an independent cohort infected with a different virus. Scoring the 15-gene signature in the influenza A atlas, infected donors separated from controls in direction: pregnant infected versus pregnant control differed by + 2.03 and infected adults versus adult controls by + 2.23 donor-mean standard-deviation units (Mann–Whitney $p = 0.20$ and 0.33; 2–3 donors per group) (Fig. 4). NK-cell fractions were 3.0-fold lower in pregnant infected than in pregnant control donors (3.9% vs 11.9%), concordant with the COVID-cohort NK depletion. Every effect replicates in direction; formal significance awaits larger cohorts.

The interferon program is also mobilizable by vaccination: in a public microarray cohort of 44 pregnant women who received inactivated influenza vaccination, the signature rose by a mean of + 0.91 standard deviations one day after vaccination (paired Wilcoxon $p < 0.001$) and returned to baseline by day 7 (Supplementary Figure S1) — a transient, controlled response showing that the pregnancy-dampened program is a dynamic, vaccine-responsive readout rather than a fixed state.

Vaccinating the mother matters only if the resulting antibodies reach the infant. In 129 published mother–newborn dyads with paired measurements after COVID-19 vaccination[15], cord-blood Spike-specific IgG tracked maternal titers closely (Spearman $\rho = 0.80$, $p < 0.001$), with a median cord-to-maternal transfer ratio of 1.15 (76% of dyads at or above 1); transfer was more efficient after second-trimester than third-trimester vaccination (median 1.23 vs 0.99; 92% vs 49% of dyads at or above 1) (Supplementary Figure S3; Table S8). Maternal vaccination therefore delivers vaccine-elicited antibodies to the neonate, with timing a modifiable efficiency factor.

The twin also speaks to target choice. Mapping the E18-3 epitope of our bench VLP platform onto the cryo-EM structure of the echovirus 18 protomer bound to FcRn (PDB 9TF0[16]), the 37-residue epitope (VP1 residues 133–169) contains two of the four VP1 receptor-contact residues (I148, P149) and approaches FcRn to 3.1 Å — the epitope directly overlaps the receptor footprint that primes viral uncoating[3,17] (Supplementary Figure S2). Antibodies raised against this epitope are therefore structurally positioned to block receptor attachment; the twin defines the pregnancy-side readouts a first-in-pregnancy evaluation should measure (Table 1).

The treated twin: no single program is sufficient

Finally, we asked what would have to change for a pregnant donor's digital twin to read as non-pregnant (Fig. 5). Restoring interferon modules alone moved the mean predicted probability of pregnancy only from 0.807 to 0.650 (17% of pregnant donors reclassified below 0.5); restoring cellular composition alone had almost no effect (0.773; 0%); restoring all immune modules normalized every pregnant donor (0.202), as did the full program (0.163, versus a non-pregnant reference of 0.186). The pregnancy signature is therefore an emergent property of coordinated multi-program cell states: interferon dampening is its most visible facet, but interferon restoration alone does not normalize the twin. For

intervention design — including vaccine-adjuvant choice — the twin suggests targeting the coordinated program rather than a single axis.

Discussion

We have shown that the digital-twin blueprint articulated for obstetrics[5] extends to infection immunology and can be executed entirely on public data: atlas construction, response mapping, outcome modeling, external cross-pathogen validation, and in-silico intervention — without a new cohort or a new animal experiment. The resulting twin reproduces the central empirical finding of the source study — pregnancy blunts the blood interferon response to SARS-CoV-2[12] — but packages it as a predictive, transferable, and interrogable model.

Three methodological choices were essential. First, donor-level modeling: with hashtag-pooled pregnant donors and per-donor batches, cell-level analyses would confound batch with group. Second, honest negative reporting: the severity model's failure (AUROC 0.21) is itself a finding about the limits of public-data reuse; suppressing it would paint a misleadingly rosy picture. Third, multi-arm restoration: the treated-twin experiment is informative precisely because it compares arms — it localizes the signal to coordinated cell-state programs rather than any single pathway.

For our bench program, the twin specifies which cell states a maternal mucosal vaccine must engage. The floss-mediated HBc-E18-3 VLP platform induces durable, FcRn-dependent intestinal IgG in preclinical models[2]; the twin suggests that in pregnancy the relevant readouts extend to the cytotoxic, antigen-presentation, and regulatory programs identified here, and that single-axis interferon induction is unlikely to normalize maternal responses. The influenza-vaccination microarray cohort (GSE166545[14]) shows that the interferon axis is a readable vaccine-response readout in pregnancy; single-cell profiling of mRNA vaccination[18] would extend this to cell-state resolution when access-restricted data become available.

The same logic scales to a paradigm: for a new maternal vaccine — mRNA, VLP, or protein — the twin functions as a computational alternative to the animal models that cannot be built: it specifies which samples to collect, when, and what to measure, with decision criteria for safety, immunogenicity, and neonatal delivery (Table 1; Fig. 6). For our own platform, the CpG (TLR9) adjuvant's innate interferon readout is directly measurable by the twin's interferon axis — which pregnancy dampens and vaccination transiently mobilizes — making adjuvant dose and timing empirical questions rather than assumptions. As regulators begin accepting computational evidence in place of animal studies[4], we suggest this readout-first protocol as the default first step in maternal vaccine development: the twin generates the evaluation framework before the first participant is enrolled.

Table 1
Twin-defined readout panel for a first-in-pregnancy vaccine evaluation.

Component	Specification	Rationale from the twin
Samples	Maternal PBMC, plasma, stool; cord blood at delivery	Donor-level blood cell states; mucosal compartment for enteric targets; neonatal transfer
Time points	Days 0, 1, 7 after vaccination; delivery	Interferon mobilization peaks at day 1 and returns to baseline by day 7 (GSE166545)
Safety readouts	Interferon-signature dynamics; composition shifts	Pregnancy dampens blood interferon; a controlled transient response, not sustained inflammation, is the expected pattern
Immunogenicity readouts	Neutralizing IgG in maternal serum; the five non-zero model features (cytotoxicity in CD4 + and γδ T cells, interferon activity in dendritic cells, antigen presentation in CD4 + T cells, Treg fraction)	These features carry the pregnancy signal the twin detected (AUROC 0.87)
Transfer readout	Cord-to-maternal IgG transfer ratio at delivery	Median 1.15 after COVID-19 vaccination; higher efficiency with second- than third-trimester vaccination
Decision criteria	Transient interferon mobilization without sustained shift; maternal seroconversion; transfer ratio ≥ 1	Combines the twin's state model with the transfer bridge to the neonate

This study has four principal limitations. First, sample size and batch structure: the discovery cohort comprises 26 donors, every non-pregnant donor is its own processing batch, and pregnant donors were hashtag-pooled; all models are donor-level and cross-validated, but confidence intervals are wide and severity prediction is confounded beyond rescue. Second, in-silico restoration is not treatment prediction: the treated-twin experiment linearly interpolates features toward a class mean on an in-sample classifier; it is hypothesis-generating, not causal, and claim nothing about any real intervention. Third, validation granularity: the influenza cohort's annotation is dominated by CD14 + monocytes, so external validation is limited to donor-level module scores and coarse composition. Fourth, restricted data: the planned single-cell vaccination-arm validation (Zenodo record 13969283) is access-restricted; the vaccine analysis is therefore limited to bulk microarray dynamics (GSE166545), which capture interferon mobilization but not cell-state resolution, and the single-cell cohort should be substituted when access is available.

Public single-cell data are sufficient to build a digital twin of the maternal immune response to infection and vaccination: a model that detects pregnancy status from immune cell states, replicates across pathogens, quantifies the antibody bridge from mother to neonate, and supports in-silico intervention

experiments that localize the pregnancy program to coordinated cell states rather than single axes. Because the twin defines what to measure next, every new pregnancy cohort directly improves its accuracy and transferability — turning the repositories we already have into the experiments we cannot run, and the field toward data-guided maternal medication and vaccine development.

Declarations

Funding. No external funding was received for this work.

Conflicts of interest. The author declares no conflicts of interest. (Disclosure forms are submitted via the journal's Convey system.)

Ethics. This study is a secondary analysis of publicly available, de-identified human data. The original cohort studies obtained ethical approval and informed consent; no new human-subjects data were collected or generated, and no data were redistributed. The work conforms to the Declaration of Helsinki.

Data sharing. All input data are publicly available: GSE239450/GSE239452 (discovery cohort)[12], GSE243629 (influenza A validation cohort)[13], and GSE166545 (influenza vaccination dynamics cohort)[14] from the Gene Expression Omnibus, and the published mother–newborn dyad titers from the source data of Atyeo et al.[15]. Donor-level output tables are provided in the Supplementary Appendix.

Code availability. The code supporting the findings of this study is available from the corresponding author upon reasonable request.

Author contributions. T.Z. conceived the study, performed all analyses, interpreted the results, wrote the manuscript, and approved the final version. W.Y. critically revised the manuscript with specific input on the epitope-structural mapping section.

Acknowledgements We thank Professor Jie Wu for expert guidance on infection-immunology insights for this work.

Use of artificial intelligence. The author used a large language model (LLM) assistant to support code development, figure preparation, and manuscript drafting and editing. All analyses, interpretations, and conclusions were generated, verified, and approved by the author, who takes full responsibility for the integrity and accuracy of the work. Consistent with journal policy, LLM tools do not qualify for authorship.

References

1. World Health Organization (2025) Equity and health: the inclusion of pregnant and breastfeeding women in clinical trials. June 18, (https://www.who.int/news/item/18-06-2025-equity-and-health--the-inclusion-of-pregnant-and-breastfeeding-women-in-clinical-trials.)

2. Zhai T, Jiang S Floss-mediated gingival mucosal immunization with HBc-E18-3 VLPs induces long-lasting intestinal IgG and provides a candidate strategy for intervention of FcRn-related autoimmune injury. bioRxiv 2026. 10.64898/2026.08.10.743934
3. Chen X, Qu X, Liu C et al (2022) Human FcRn is a two-in-one attachment-uncoating receptor for echovirus 18. MBio. ;13:e01166-22. 10.1128/mbio.01166-22
4. Food US, Administration D (2025) FDA announces plan to phase out animal testing requirement for monoclonal antibodies and other drugs. April 10, (https://www.fda.gov/news-events/press-announcements/fda-announces-plan-phase-out-animal-testing-requirement-monoclonal-antibodies-and-other-drugs.)
5. Scott AK (2026) Building digital pregnancies from the single cell up. Sci Adv 12:eaen0211. 10.1126/sciadv.aen0211
6. Rackaityte E, Borges B, Kortbawi HM et al (2026) Evidence for maternal autoantibodies in the pathogenesis of preterm birth. Nat Immunol 2280–2292. 10.1038/s41590-026-02689-y
7. Corry J, Zotova N, Tabala M et al (2026) Compartment-specific soluble immune profiles associated with preterm birth, perinatal death, and low birthweight in pregnant individuals living with HIV. Sci Rep 16:34647. 10.1038/s41598-026-75643-5
8. Tan Y, Liang D, Eick SM et al (2026) Maternal metabolomic signatures for risk of preterm and early term birth in African American women. Nat Health 1:1106–1120. 10.1038/s44360-026-00150-w
9. Gu L, Li W, Zhu Y et al (2026) Dynamic prediction of preterm birth and gestational age at delivery based on longitudinal electronic health records. npj Digit Med. 10.1038/s41746-026-03071-w
10. Jiao X, Yang Y, Li F et al (2026) Longitudinal dynamics of the maternal gut virome associate with metabolic features of preterm birth. Nat Commun 17:9294. 10.1038/s41467-026-76220-0
11. Stelzer IA, Gillard J, Urbschat C et al (2026) Immunological maladaptation preceding spontaneous preterm birth in human pregnancies. Nat Commun 17:7121. 10.1038/s41467-026-75605-5
12. Oh DS, Kim E, Normand R et al (2024) SARS-CoV-2 infection elucidates features of pregnancy-specific immunity. Cell Rep 43:114933. 10.1016/j.celrep.2024.114933
13. Zhang Y, Zong L, Zheng Y et al (2023) A single-cell atlas of the peripheral immune response in patients with influenza A virus infection. iScience 26:108507. 10.1016/j.isci.2023.108507
14. NCBI Gene Expression Omnibus GSE166545: Blood gene expression profiles following influenza vaccination in pregnant women. (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE166545.)
15. Atyeo CG, Shook LL, Brigida S et al (2022) Maternal immune response and placental antibody transfer after COVID-19 vaccination across trimester and platforms. Nat Commun 13:3571. 10.1038/s41467-022-31169-8
16. RCSB Protein Data Bank 9TF0: Structure of echovirus 18 in complex with neonatal Fc receptor. (https://www.rcsb.org/structure/9TF0.)

17. Mukhamedova L, Buchta D, Trebichalská Z et al (2026) Particles of echovirus 18 open to release their genomes in vivo. Proc Natl Acad Sci U S A 123:e2601182123. 10.1073/pnas.2601182123
18. Kang YT, Lu TJ, Cheng PL et al (2026) Single-cell mapping of peripheral immune dynamics in pregnant women after Moderna mRNA COVID-19 vaccination. Genes Immun 27:374–383. 10.1038/s41435-026-00394-2

Figures

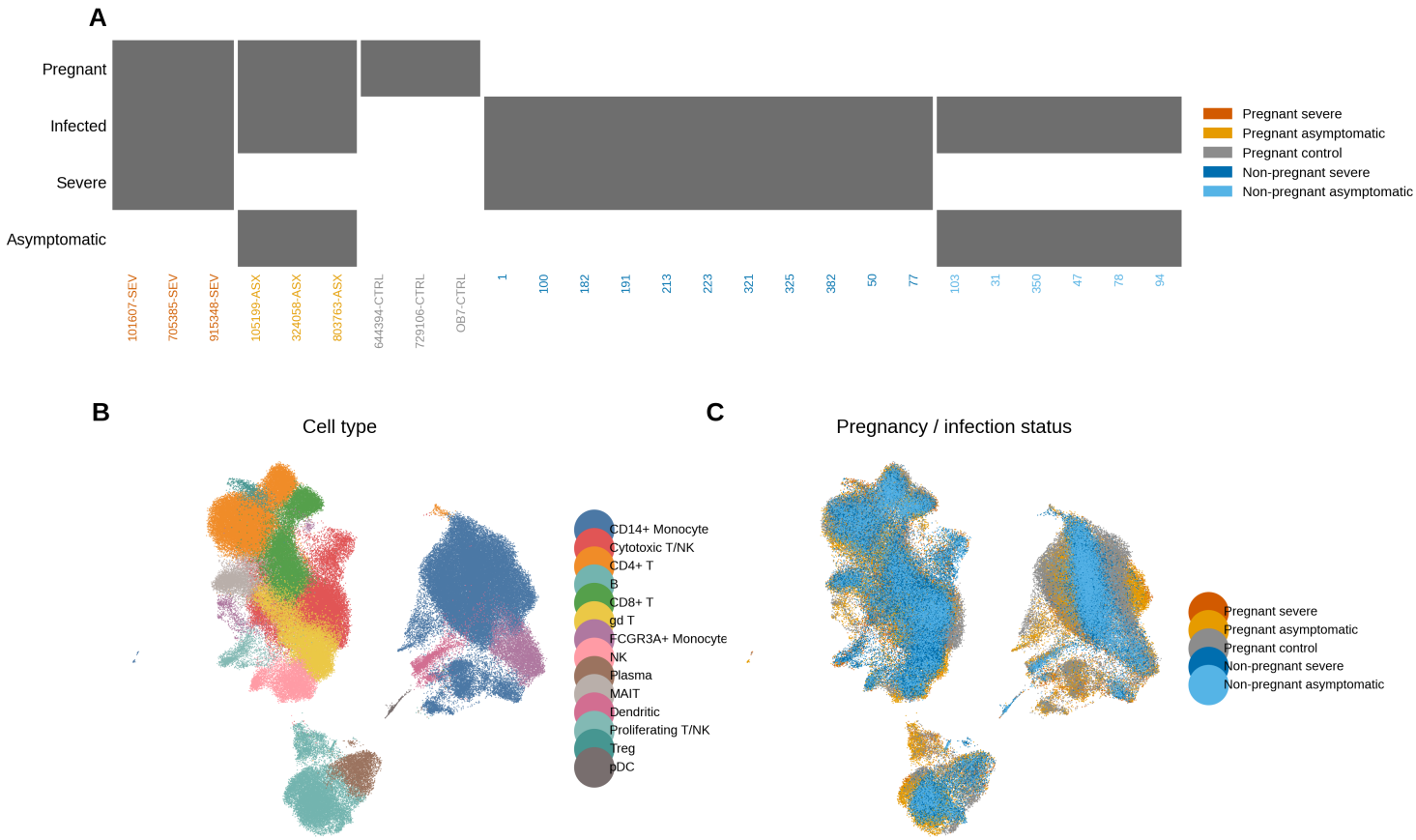


Figure 1

Cohort and cell-state atlas. (A) Donor-by-attribute matrix for the 26 donors (181,112 PBMCs) of the public SARS-CoV-2 pregnancy CITE-seq cohort (GSE239450/GSE239452), grouped by pregnancy and disease status; filled cells mark attributes, and donor IDs are colored by group. **(B)** UMAP of the Harmony-integrated atlas colored by the 14 annotated immune cell types. **(C)**UMAP colored by pregnancy/infection group.

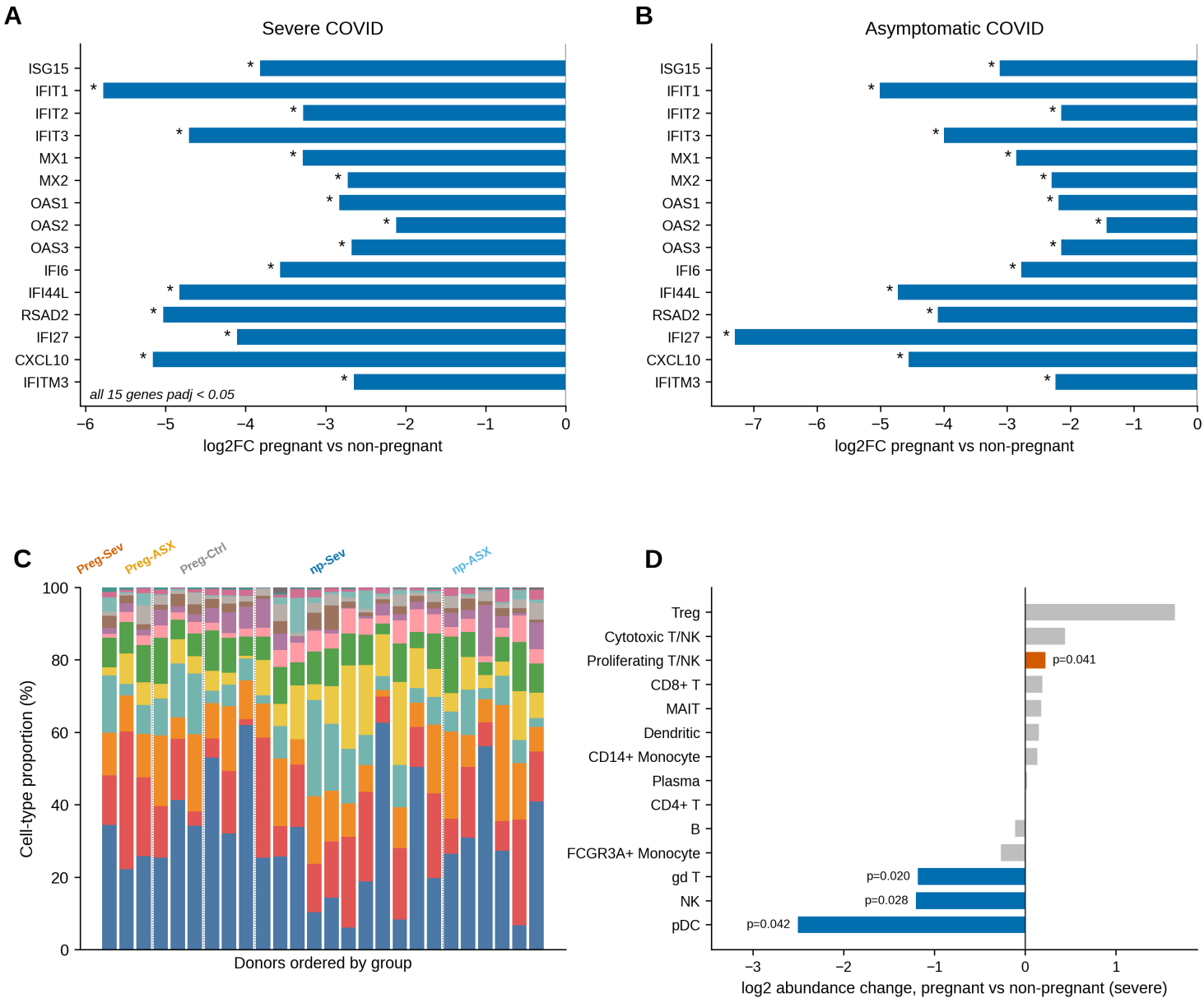


Figure 2

Pregnancy blunts the interferon response and reshapes immune composition. (A) Pseudobulk log2 fold change (pregnant vs non-pregnant donors) for 15 interferon-stimulated genes in severe COVID-19; all genes adjusted p < 0.05. **(B)** The same in asymptomatic COVID-19. **(C)** Per-donor cell-type proportions ordered by clinical group. **(D)** Log2 abundance change, pregnant vs non-pregnant donors with severe COVID-19; colored bars are nominally significant (Kruskal–Wallis p < 0.05), grey bars are not.

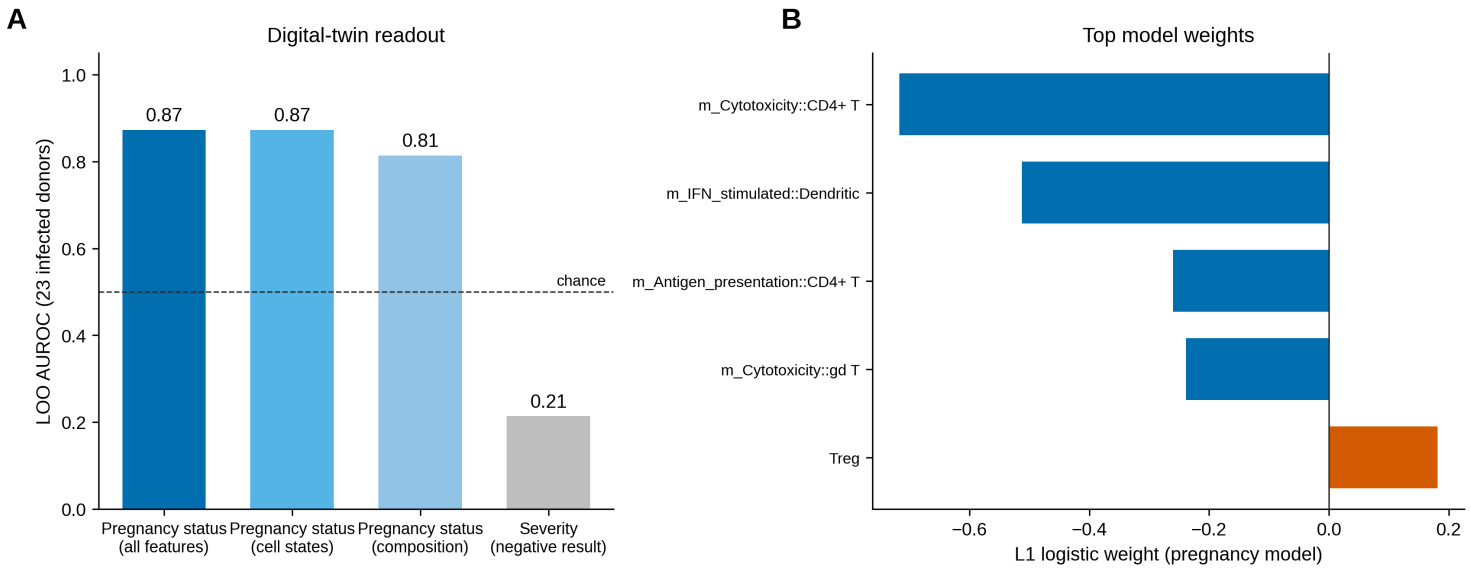


Figure 3

Pregnancy-status readout from immune cell states. (A) Leave-one-out cross-validated AUROC for pregnancy status (all features, cell states only, composition only) and for severity (negative result); dashed line marks chance. **(B)** The five non-zero L1 logistic weights of the pregnancy model.

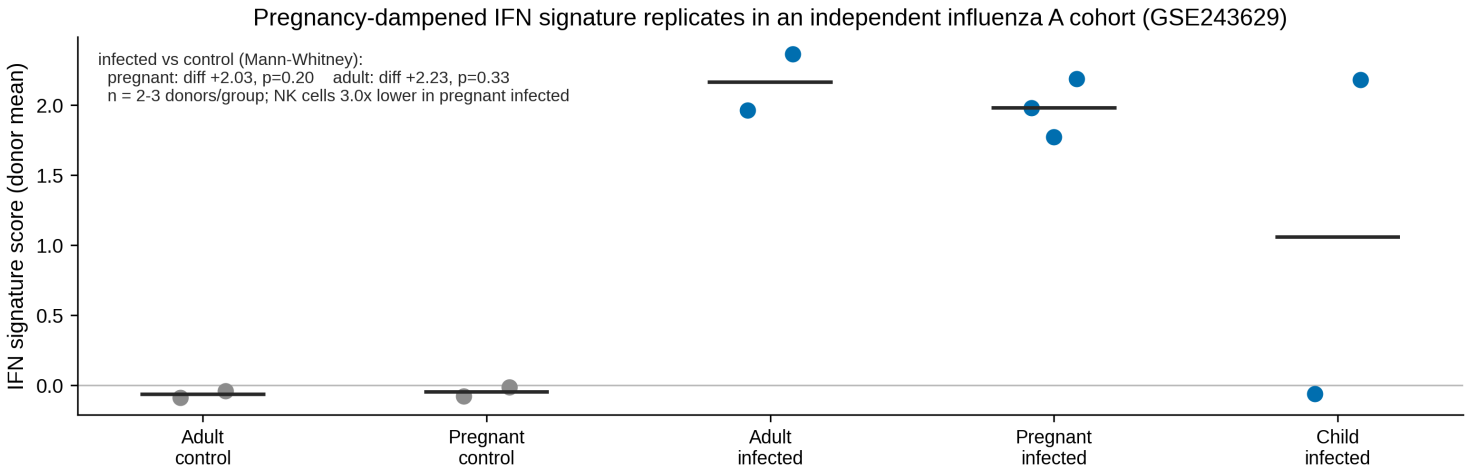


Figure 4

External validation in an independent influenza A cohort (GSE243629). Donor-mean scores of the COVID-derived pregnancy-dampened interferon signature across the five donor groups (AC, Adult control; PC, Pregnant control; AI, Adult infected; PI, Pregnant infected; CI, Child infected); NK cells are 3.0-fold lower in pregnant infected vs pregnant control donors.

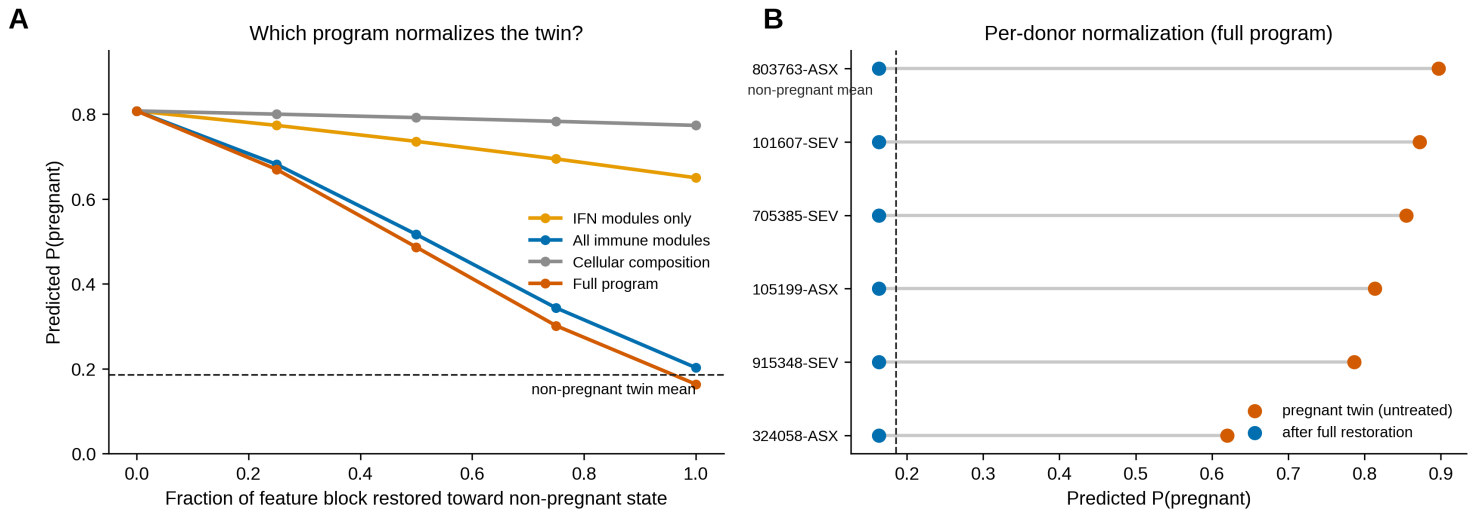


Figure 5

The treated twin. (A) Mean predicted P(pregnant) for pregnant donors as feature blocks are restored toward the non-pregnant state (dashed line = non-pregnant twin mean). **(B)**Per-donor P(pregnant) before vs after full-program restoration.

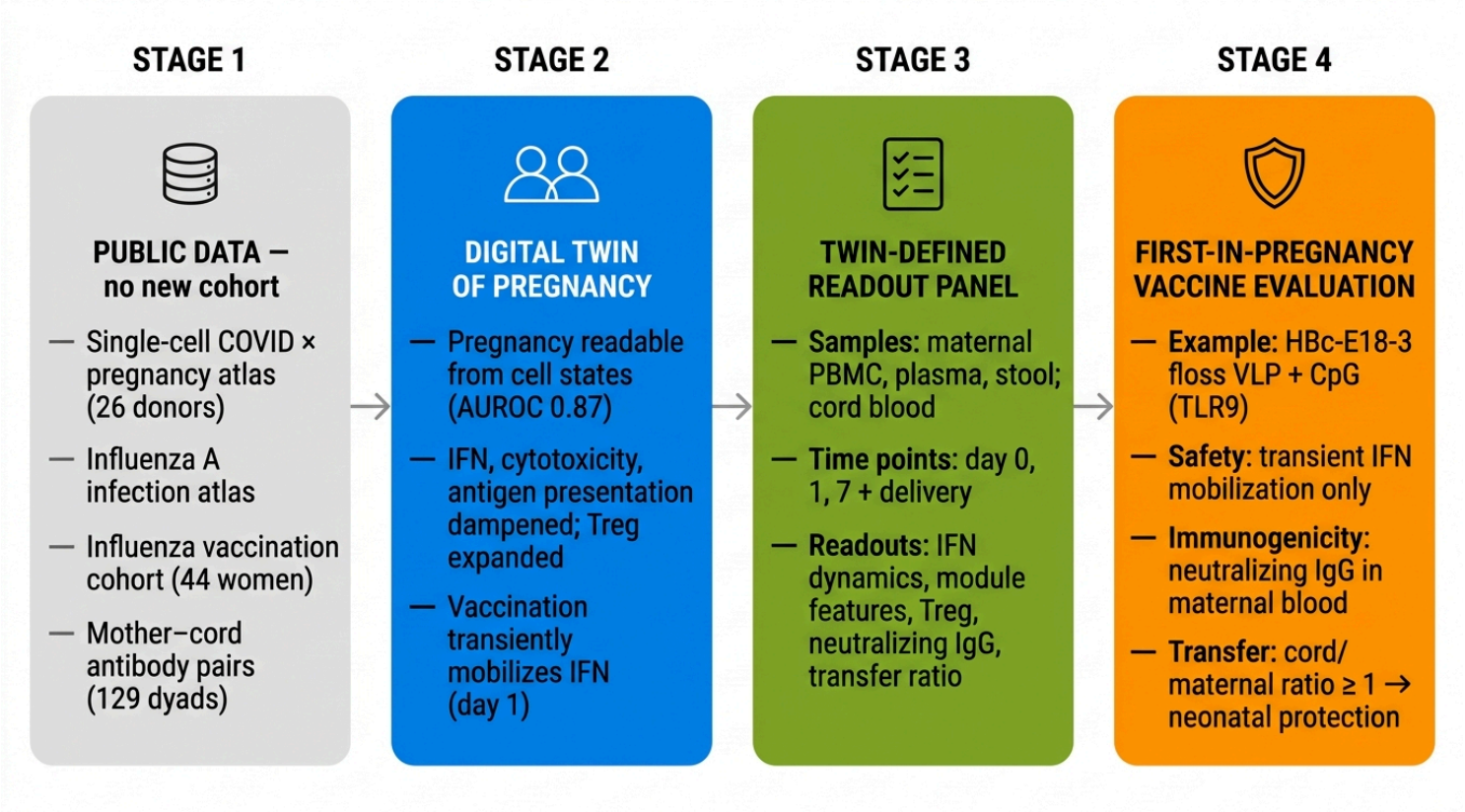


Figure 6

The digital twin as a computational animal model for maternal vaccine evaluation. Public data build the twin; the twin defines the readout panel (samples, time points, measurements); the panel structures first-

in-pregnancy evaluation of new vaccine platforms, with the bench HBc-E18-3 floss VLP as the worked example and mother-to-child antibody transfer completing the protection chain.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 02supplementaryappendix.docx