IMMUNOLOGY

Single-cell–level digital twins for preterm birth prevention strategies

Jakob Einhaus[1]†, Peter Neidlinger[1]†, Olivier Fondeur[1]†, Masaki Sato[1,2]†, Sasha Anronikov[1], Kotaro Miyazaki[1], Antoine Maechler[1], Jonas N. Amar[1], Kazuo Ando[1], Valentin Badea[1], Dyani K. Gaudilliere[3], Maximilian Sabayev[1], Dorien Feyaerts[1], Maïgane Diop[1], Amy S. Tsai[1,3], Amelie Cambriel[1,4], Edward A. Ganio[1], Romain Lagarde[1], Emmeline O'Kelly[1], Ina A. Stelzer[1,5], Julien Hedou[1,6], Ronald J. Wong[7], Yair J. Blumenfeld[8], Deirdre J. Lyell[8], Gary M. Shaw[7], Tomiko T. Oskotsky[9,10], Marina Sirota[10], Linda C. Giudice[11], David K. Stevenson[7], Nima Aghaeepour[1], Brice Gaudilliere[1]*

Copyright © 2026 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

Digital twin models can accelerate therapeutic development by enabling low-risk testing of candidate interventions. In preterm labor (PTL), a major pregnancy complication where clinical trials face ethical and financial barriers, digital twins are especially valuable for evaluating new therapies targeting immune dysfunctions driving PTL. Yet, current models lack single-cell resolution, limiting detection of cell type–specific mechanisms, off-target effects, and the design of personalized interventions. We present Simulated Immunome Modeling of Clinical Outcomes (SIMCO), a single-cell–level digital twin framework that models immunomodulatory treatment effects on the timing of labor using immunome-wide, single-cell simulations. SIMCO's digital twins are trained and validated on a newly generated mass cytometry atlas of the pregnant immunome exposed to nine candidate drugs preselected for PTL prevention. Applying SIMCO to an independent cohort of pregnant individuals, we simulate treatment effects on gestational length, screening for candidate drugs that delay labor timing and providing system-level mechanistic insight for each drug candidate. Tetrahydrofolate and the combination of aspirin and lansoprazole emerged as top candidates for PTL prevention, delaying labor onset primarily through enhanced mTOR signaling in innate immune cells and attenuated JAK/STAT signaling in naïve CD4$^+$ T cells.

INTRODUCTION

Preterm birth, defined as birth before 37 completed weeks of gestation, remains the leading cause of child mortality under age five worldwide (*1*). One in 10 pregnancies ends preterm, leading to substantial neonatal complications. Most cases (60 to 70%) are spontaneous preterm labor (sPTL) and can occur without warning signs or known risk factors (*2*). While multiple factors contribute to sPTL, increasing evidence points to dysregulation of innate and adaptive immune adaptations that maintain pregnancy from conception to labor (*3*–*11*). Early pregnancy is dominated by tolerogenic mechanisms, such as the expansion of regulatory T cells (T_{reg} cells), which promote maternal-fetal tolerance (*12*–*15*). Late pregnancy involves activation of pro-inflammatory pathways in neutrophils, monocytes, macrophages, and natural killer cell subsets (*16*–*18*). These temporal dynamics culminate in an inflammatory cascade of parturition (*19*–*21*). In sPTL, anticipated immune adaptations of healthy pregnancies occur prematurely or inappropriately (*22*).

Understanding the differences in gestational immune dynamics between pregnancies resulting in sPTL and those delivering at term is critical for identifying immunomodulatory interventions that could influence timing of labor onset (*23*, *24*). Current sPTL preventive strategies, including low-dose aspirin (*25*) and progesterone (*26*, *27*), have shown modest and inconsistent results in clinical trials, likely benefiting some but not all patients (*28*). Given the ethical and financial constraints on clinical trials during pregnancy, immune digital twins offer a transformative approach to accelerate the development of immunomodulatory drug candidates for sPTL (*29*, *30*). Incorporating single-cell resolution can equip digital twin models to capture immune system–wide drug effects, reveal cell type–specific mechanisms, identify off-target responses, and guide the rational design of personalized interventions.

Mass cytometry enables single-cell immune profiling with high-dimensional functional readouts including cell-surface receptors, cytokines, and phosphorylation states (*31*). Beyond characterizing immune states, the rich datasets generated by mass cytometry can be harnessed by machine learning to extract biomarkers of clinical outcomes. For example, the recently developed Stabl algorithm uses artificial noise to select informative features and build sparse, interpretable prediction models (*32*). Applied to peripheral blood samples across gestation, mass cytometry combined with Stabl revealed conserved temporal trajectories of immune cell signaling activity, enabling the prediction of labor timing in both term and preterm pregnancies (*5*, *33*). Besides biomarker discovery, mass cytometry can also serve as a high-throughput screening platform to study cell

[1]Department of Anesthesiology, Perioperative, and Pain Medicine, Stanford University School of Medicine, Stanford, CA, USA. [2]Department of Critical Care and Anesthesia, National Center for Child Health and Development, Tokyo, Japan. [3]Department of Surgery, Stanford University School of Medicine, Stanford, CA, USA. [4]APHP, Hopital Saint Antoine & Tenon, DMU DREAM, Department of Anesthesiology and Critical Care, Paris, France. [5]Department of Pathology, University of California San Diego, La Jolla, CA, USA. [6]Sorbonne Université, INSERM, Centre de Recherche Saint-Antoine, CRSA, F-75012 Paris, France. [7]Department of Pediatrics, Division of Neonatal and Developmental Medicine, Stanford University School of Medicine, Stanford, CA, USA. [8]Department of Obstetrics and Gynecology, Stanford University School of Medicine, Stanford, CA, USA. [9]Division of Clinical Informatics and Digital Transformation, University of California, San Francisco, San Francisco, CA, USA. [10]Bakar Computational Health Sciences Institute, University of California San Francisco, San Francisco, CA, USA. [11]Center for Reproductive Sciences, Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California San Francisco School of Medicine, San Francisco, CA, USA.
*Corresponding author. Email: gbrice@stanford.edu
†These authors contributed equally to this work.

type and pathway specificity of immunomodulatory therapeutics (*34*). With recent advances in computational modeling, a vast scale of drug effects can then be parsed at the single-cell level (*35*). In particular, optimal transport (OT) has emerged as a powerful mathematical framework to model single-cell immune perturbation responses (*36*). For mapping unpaired distributions with and without perturbation, CellOT by Bunne *et al.* (*37*) leverages a neural network for efficient OT solving. Its robust performance on unseen data enables personalized treatment response modeling beyond the original perturbation dataset.

In this study, we present Simulated Immunome Modeling of Clinical Outcomes (SIMCO), a single-cell–level digital twin framework that models immunomodulatory treatment effects on clinical outcomes by integrating simulated immune responses with multivariable modeling. Trained on a new single-cell mass cytometry atlas of the pregnant immunome, SIMCO's digital twins predict patient-specific immune modulation and resulting effects on gestational length for nine US Food and Drug Administration (FDA)–approved, low-risk (category A/B) drug candidates—eight single agents and one combination—preselected through a published computational drug repurposing pipeline for their potential to prevent preterm birth (*23*).

RESULTS

Simulating treatment effects from single-cell–level digital twins

Simulating the therapeutic potential of immunomodulatory treatment requires digital twins of the immunome that not only predict clinical outcomes but also forecast how specific treatments would alter these outcomes in individual patients. SIMCO addresses this need by integrating simulation of single-cell responses with clinical outcome modeling (Fig. 1, Materials and Methods, and fig. S1):

1) First, SIMCO is trained on high-dimensional single-cell perturbation data (e.g., flow and mass cytometry and single-cell transcriptomics) to encode drug-perturbed immune responses at single-cell resolution using OT learning with CellOT. CellOT leverages an input-convex neural network that is highly accurate in predicting heterogeneous single-cell perturbation responses (*37*).

2) Second, SIMCO projects the learned single-cell behavior onto a separate immunome dataset, comprising only baseline samples (unstimulated and untreated) with linked clinical outcomes. From these baseline samples, SIMCO simulates single-cell responses to perturbations (e.g., ex vivo stimulation or drug treatment). This simulation generates digital twins of each patient's immunome with and without exposure to the candidate drugs.

3) Last, SIMCO builds a multivariable predictive model of the clinical outcome with Stabl, a sparse machine learning algorithm, on a training set of patients using the simulated immune data. In a held-out test set of patients, SIMCO compares the predictions for each patient's untreated and treated digital twin. This difference quantifies the predicted drug effect on the outcome, enabling SIMCO to model the immunomodulatory potential of candidate drugs and to personalize treatment strategies.

A single-cell perturbation atlas capturing pregnancy-relevant immune profiles

To generate a mass cytometry single-cell perturbation atlas, we enrolled eight nonpregnant and 11 pregnant volunteers for a multiplexed drug screen assay measuring immunomodulatory effects of nine candidate drugs for the prevention of sPTL (Fig. 2A and Table 1). A second trimester sampling timepoint was chosen to capture the influence of candidate drugs on the pregnant immunome during the maintenance phase of gestation and before proinflammatory pathways of parturition become increasingly activated. Nonpregnant controls were sampled in the early follicular phase of their menstrual cycle

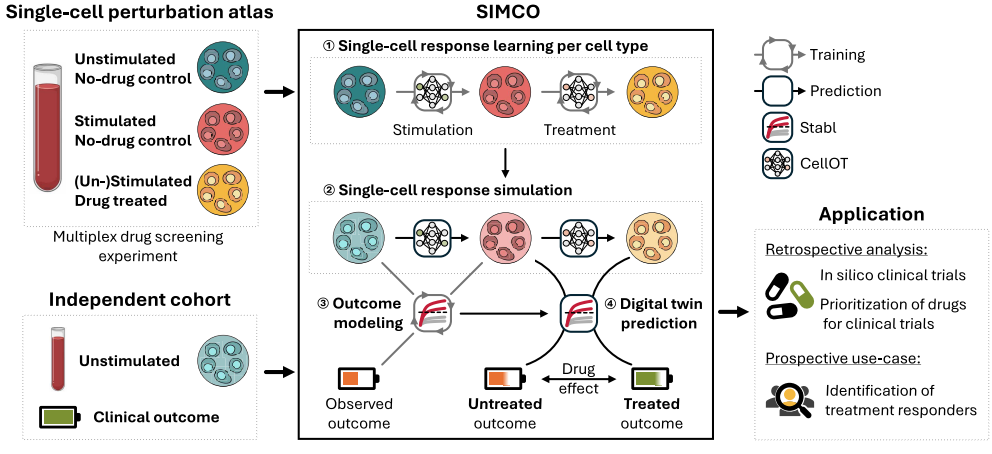


Fig. 1. Overview over the single-cell–level digital twin framework SIMCO. SIMCO is a single-cell–level digital twin framework that combines simulated immune responses with multivariable modeling of clinical outcomes to evaluate the immunomodulatory potential of candidate drugs. In vitro perturbation experiments (i.e., analysis of ex vivo stimulation responses to inflammatory ligands in the presence or absence of drug candidates) are performed to generate a single-cell perturbation atlas of immunomodulatory treatment effects. SIMCO learns OT maps (using CellOT) that encode the observed single-cell perturbation responses. For baseline samples from an independent clinical cohort linked to a clinical outcome, SIMCO simulates single-cell responses by applying the learned OT maps. A sparse machine learning model (Stabl) is trained on the simulated single-cell data before treatment (light turquoise, light red) to model the observed outcome. Applied to an unseen test set of patients, SIMCO provides a readout of simulated drug effects on the outcome as the difference in outcome prediction between the treated and untreated digital twin (light red, light yellow). SIMCO-generated immune digital twins can be used for in silico clinical trials and prioritization of candidate drugs for clinical trial testing, as well as for personalization of treatment strategies. Parts of the figure were created in BioRender. E. Ganio (2026) (https://BioRender.com/9k4hdoc).

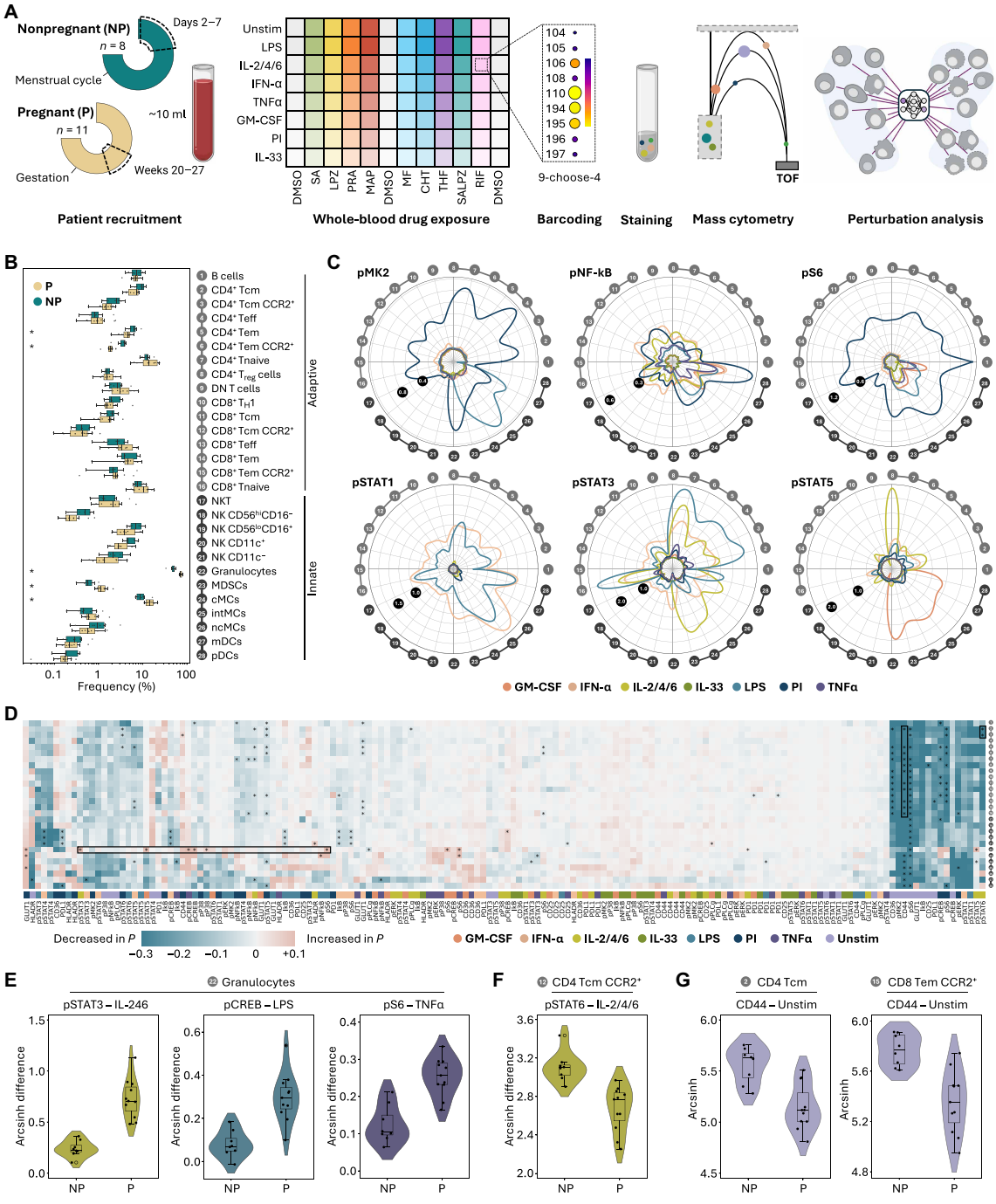


Fig. 2. A single-cell immune perturbation atlas of the second-trimester pregnancy immunome. (**A**) In 19 individuals, including eight nonpregnant (NP) between days 2 to 7 of the menstrual cycle and 11 pregnant (P) between 20 to 27 weeks of gestation, immunomodulatory effects of nine candidate drugs (eight single drugs and one drug combination) were screened in a 96-plex ex vivo whole blood immunoassay. (**B**) Proportional abundance of 28 cell populations in pregnant and nonpregnant individuals. Percent of all live leukocytes (granulocytes) and percent of mononuclear cells (all other cell types) are depicted. Significance ($P < 0.05$, Mann-Whitney U test, Benjamini-Hochberg corrected) is indicated. (**C**) Radial plots show absolute arcsinh difference between stimulated and unstimulated samples across 28 immune cell populations for selected functional markers. Cell populations are numbered as defined in (B). (**D**) Heatmap of differences in average immune responses to stimulation between pregnant and nonpregnant individuals. Stars indicate significant differences ($P < 0.05$, Wilcoxon rank-sum test, Benjamini-Hochberg corrected across all features with a group difference > 0.05 arcsinh). Feature groups with significant differences (individual features plotted in (E) to (G) are highlighted with boxes. Cell populations are numbered as defined in (B). (**E** to **G**) Example features of significant differences between pregnant and nonpregnant individuals. Pregnant individuals showed up-regulated immune responses in innate immunity (granulocytes, E), down-regulated adaptive immune responses ($CD4^+$ central memory T cells, F), and down-regulated CD44 expression (G). Depicted is arcsinh difference to unstimulated (for stimulation responses) or arcsinh-transformed signal (for unstimulated). Candidate drugs are abbreviated as follows: chlorthalidone (CHT), maprotiline (MAP), metformin (MF), lansoprazole (LPZ), pravastatin (PRA), rifabutin (RIF), SA, SALPZ, and THF. DMSO was used as no-drug control. Parts of the figure were created in BioRender. E. Ganio (2026) (https://BioRender.com/4bi6cci). TOF, time-of-flight.

Table 1. Clinical information and patient demographics for pregnant and nonpregnant volunteers. Percentage and interquartile range (IQR) within each cohort are indicated.

	Nonpregnant individuals ($n = 8$)	Pregnant individuals ($n = 11$)
Age (years)	32 (29–35, IQR)	30 (29–32, IQR)
Race		
White	5 (62.5%)	4 (36.4%)
Asian	2 (25.0%)	4 (36.4%)
Black	1 (12.5%)	0
Other	0	3 (27.3%)
Ethnicity		
Non-Hispanic	8	8
Hispanic	0	0
Other	0	3
Pregnancy information		
GA at sampling time		22.4 (20.9–22.8, IQR)
Prepregnancy BMI		24.6 (22.4–27.0, IQR)
GA at delivery		40.0 (39.5–40.05, IQR)
Medication		
Prenatal vitamin		7 (63.6%)
Omeprazole prenatal		1 (9.1%)
Contraception		
Oral contraception	1 (9.1%)	
Copper IUD	1 (9.1%)	
Hormonal IUD	2 (18.2%)	
Cycle day	3 (3–4, IQR)	
Medical history		
Hyperlipidemia		1 (9.1%)
Factor V Leiden		1 (9.1%)
PCOS	1 (12.5%)	1 (9.1%)
IUI		1 (9.1%)

to capture a low-hormone state. The candidate drugs—chlorthalidone, maprotiline, metformin, lansoprazole, pravastatin, rifabutin, salicylic acid (SA), SA + lansoprazole (SALPZ), and 5-methyltetrahydrofolate (THF)—were preselected using a previously published computational drug repurposing pipeline based on their potential to positively influence sPTL-associated gene signatures (*23*). Only candidate drugs considered relatively safe for use in pregnancy (FDA category A/B) that can be administered orally were included in the assay. For each candidate, we used the compound form predominantly present in the bloodstream after oral administration, which in some cases corresponds to an active metabolite (e.g., THF and SA). Experimental concentrations were selected to reflect in vivo plasma levels following standard oral dosing (Table 2).

Each screening experiment consisted of 96 conditions requiring only 100 μl of whole blood per condition. A two-hour, whole blood stimulation allowed us to examine physiologically relevant cell type– and stimulation-specific responses. The functional responsiveness of 28 immune cell populations was assessed as expression changes in 12 intracellular phospho-proteins and eight activation markers (fig. S2 and table S1). All conditions were analyzed simultaneously using mass cytometry by leveraging a newly developed nine-choose-four doublet-removing mass-tag barcoding scheme (see fig. S3 and Materials and Methods). In total, our single-cell perturbation atlas comprised 218 million leukocytes across 80 perturbation scenarios.

In the absence of drug treatment (0.5% dimethyl sulfoxide, DMSO-control), pregnant and nonpregnant individuals exhibited distinct immune profiles. Several innate immune populations—such as granulocytes, classical monocytes (cMCs), and myeloid-derived suppressor cells (MDSCs)—were proportionally expanded during mid-pregnancy, whereas $CD4^+$ effector memory T cells were reduced (Fig. 2B). Functional responses to seven stimuli were quantified for each cell type, revealing highly specific response profiles for different signaling pathways induced to each stimulation condition (Fig. 2C). Of the 4480 immune activation features, 163 differed significantly between the two groups (Fig. 2D).

Specifically, in samples from pregnant compared to nonpregnant participants, we observed increased innate immune signaling in granulocytes in response to interleukin-2/4/6 (IL-2/4/6), lipopolysaccharide (LPS), and tumor necrosis factor–α (TNFα) (Fig. 2E), suppressed adaptive immune responses, e.g., in pSTAT6 in $CD4^+$ central memory T cells in response to IL-2/4/6 (Fig. 2F), and reduced the expression of CD44 on innate and adaptive immune cells in unstimulated samples (Fig. 2G). These findings highlight the global adaptations of the maternal peripheral immune system during pregnancy and

Table 2. Candidate drugs for the prevention of sPTL. The screened compound, the used assay concentration, and the abbreviation (Abb.) of the compound in the figures are listed. $T_{1/2}$, half-life.

Drug candidate	Clinical reference dose	Cmax (ng/ml)	$T_{1/2}$ (hours)	Relevant metabolites	Assay concentration (ng/ml)	Screened compound	Abb.	References
Chlorthalidone	12.5 mg	75	32	None	75	Chlorthalidone	CHT	(*74*)
Maprotiline	75 mg	24	40	Desmethyl-maprotiline	10	Maprotiline	MAP	(*75*)
Metformin	500 mg	1100	4	None	800	Metformin	MF	(*76*)
Lansoprazole	30 mg	1000	2	5′-hydroxy-lansoprazole	400	Lansoprazole	LPZ	(*77*)
Pravastatin	10 mg	15	2.5	3′α-isopravastatin	7.5	Pravastatin	PRA	(*78*)
Rifabutin	150 mg	200	36	25–*O*-deacetyl-rifabutin	10	Rifabutin	RIF	(*79*)
Aspirin	81 mg	5700	2.5	Salicylic acid	2500	Salicylic acid	SA	(*80*)
Folic acid	800 μg	15	5.5	5-Methyl-THF	75	5-Methyl-THF	THF	(*81*)

recapitulate important innate and adaptive cell responses previously implicated in the maintenance of healthy pregnancies (*4*, *33*, *38*).

Candidate drugs show differential immunomodulatory properties

From the single-cell perturbation atlas, we characterized each candidate drug's immunomodulatory profile when administered in the context of the pregnant immunome. For each candidate drug, we analyzed 4480 immune cell features comparing the drug-treated measurements to their respective DMSO controls. This provided an in-depth analysis of cell type– and pathway-specific concerted effects that each candidate drug exerted on the immune cells' responses and interactions occurring over the 2-hour stimulation period (Fig. 3A and fig. S4). For example, SA globally enhanced Janus kinase–signal transducers and activators of transcription (JAK-STAT) responses in response to LPS stimulation while differentially regulating these pathways under polymethyl methacrylate/ionomycin (PI) stimulation in a highly cell type–specific manner. Hence, SA increased the pSTAT5 activation in $CD4^+$ T_{reg} cells after PI stimulation but decreased pSTAT5 in $CD8^+$ T cell subsets.

Across all treatments, we identified 1872 significant drug effects (paired *t* test, $P < 0.05$, Benjamini-Hochberg corrected; Fig. 3B). Most drug effects (94.8%) were inhibitory. Naïve $CD4^+$ and $CD8^+$ cells and central memory $CD4^+$ T cells were the most frequently targeted immune cell populations, followed by granulocytes and cMCs among innate cell types (Fig. 3C). PI and granulocyte-macrophage colony-stimulating factor (GM-CSF) stimulations were most frequently modulated.

We categorized drug effects in the following groups: inhibitory/neutralizing or stimulatory/amplifying treatment effects affecting (i) baseline immune cell activity, (ii) strong stimulation responses (arcsinh difference > 0.1 between unstimulated and stimulated), or (iii) weak stimulation responses (arcsinh difference < 0.1; Fig. 3D). Most effects amplified weak or inhibited strong stimulation responses. Overall, all tested treatments showed cell type– and pathway-specific immunomodulatory profiles that may be leveraged to counteract disease-associated immune dysregulation.

We next investigated whether the identified treatment effects differed between pregnant and nonpregnant individuals, given the distinct immune activation states of the two cohorts. For each protein marker, stimulation, and cell type, we correlated cohort-level average drug effects between the pregnant and nonpregnant cohorts. This cross-cohort correlation analysis revealed conserved treatment effects on certain protein markers (e.g., pSTAT1, pSTAT3, and CD25), stimulation conditions (e.g., PI), and cell types (e.g., granulocytes and plasmacytoid dendritic cells) across both cohorts (Fig. 3E). In contrast, effects involving PD-L1, CD44, or GM-CSF stimulation differed between pregnant and nonpregnant individuals. Correlating individual-level drug effects between individuals showed that the achieved drug effects were more consistent among pregnant individuals (Pearson R range = 0.3 to 0.5) than among nonpregnant individuals (Pearson R range = 0 to 0.5), reflecting lower variability in treatment responses in the pregnant group (Fig. 3F). We further evaluated whether variability within the pregnant cohort could be explained by maternal age, gestational age at sampling, fetal sex, or prepregnancy BMI. Across all covariates, mean absolute Spearman correlations were weak (0.24 to 0.29) and did not exceed values expected under permutation-based null models (all empirical $P > 0.05$), indicating no systematic confounding by these variables. While the cumulative effect size per drug was generally comparable between cohorts, lansoprazole and SALPZ exhibited stronger cumulative effects in pregnant individuals, whereas pravastatin showed a reduced effect (Fig. 3G). To identify drugs that preferentially modulated immune function in pregnant individuals, we calculated a pregnancy-enrichment index for each drug-feature combination, ranging from 0 (equal effect in both groups) to 1 (effect present only in pregnant individuals; see Materials and Methods). While this metric reflects relative enrichment rather than definitive pregnancy-exclusive behavior, enrichment was highest for effects mediated by SALPZ and chlorthalidone (Fig. 3H).

Notably, the combination therapy SALPZ achieved a distinct immunomodulatory profile compared to its individual components, SA and lansoprazole (fig. S5). SA alone increased JAK-STAT signaling in response to LPS, whereas lansoprazole and SALPZ exerted inhibitory effects on the same pathway. Lansoprazole strongly suppressed calcium signaling in response to PI stimulation, an effect that was less pronounced with SALPZ and absent with SA. In response to IL-2/4/6, lansoprazole enhanced pSTAT3 and pSTAT6 activation, while SALPZ attenuated these signals. To assess the interaction between the two drugs, we compared the observed effects of SALPZ to the summed effects of SA and lansoprazole, classifying them as synergistic, additive, or antagonistic effects. Across all features, SALPZ

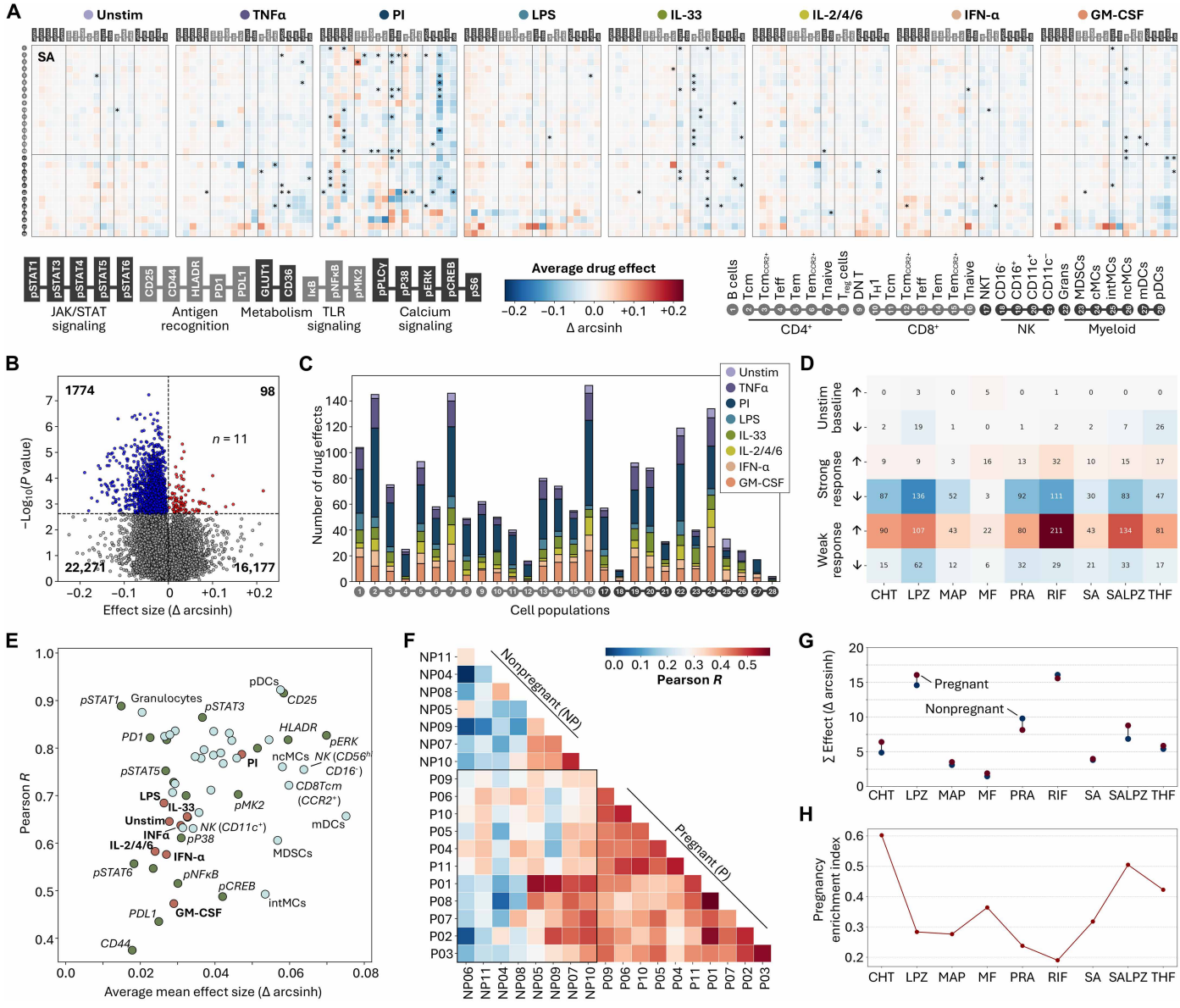


Fig. 3. Candidate drugs show differential immunomodulatory properties. (**A**) Heatmap of drug effects exerted by SA on the pregnant immunome ($n = 11$). The arcsinh difference between treated and untreated is plotted and significant effects (see B) are labeled by stars. (**B**) Volcano plot of all 40,320 drug effects for pregnant individuals ($n = 11$). Mean drug effect size and univariate P value (paired t test) are plotted. Significant drug effects with a P value < 0.05 after Benjamini-Hochberg correction are colored in blue (decreasing) or red (increasing). (**C**) Number of significant drug effects (see B) per cell population, colored by stimulation condition. (**D**) Number of significant drug effects (see B) by candidate drug, categorized as stimulatory/ amplifying (↑) or inhibiting/neutralizing (↓) effects affecting either unstimulated (baseline) immune activity, strong stimulation responses (>0.1 arcsinh difference between unstimulated and stimulated), or weak stimulation responses (<0.1 arcsinh difference). (**E**) Cross-cohort correlation analysis between cohort-level averages of pregnant (P) and nonpregnant (NP) for significant drug effects (see B) grouped by protein (green), cell type (light blue), or stimulation (orange) targeted. The average Pearson R is plotted against the average effect size. (**F**) Correlation matrix of all significant drug effects (see B) between individuals. (**G**) Average cumulative effect size for all significant drug effects (see B) comparing pregnant and nonpregnant. (**H**) Average pregnancy-enrichment index of all significant drug effects (see B).

effects were synergistic in 3%, additive in 79%, and antagonistic in 18%. Most synergistic interactions involved JAK-STAT signaling following LPS or interferon-α (IFN-α) stimulation.

SIMCO accurately simulates cell type-specific perturbation responses

Perturbation modeling approaches at the single-cell level can leverage millions of data points to provide accurate predictions across heterogeneous individuals. SIMCO uses the OT algorithm CellOT to encode the observed perturbation responses as generalizable transport maps, under the hypothesis that the multivariable marker profile of each cell before stimulation or treatment is predictive of its response. For each cell type, SIMCO was trained to model (i) stimulation responses from DMSO-control unstimulated samples and (ii) drug-treatment responses within each stimulation condition from DMSO-control samples. The trained models included only the 13

functional markers that overlapped between the perturbation dataset and the clinical dataset from Stelzer *et al.* (*33*) (Fig. 4A).

To validate SIMCO's digital twin models, we compared simulated immune responses with the observed, real-world responses in the perturbation atlas. Models were evaluated based on three metrics using a fourfold cross-validation strategy: the mean absolute error (MAE) and the coefficient of determination (R^2) between simulated and observed population-level medians as well as the maximum mean discrepancy (MMD) between simulated and observed marker distributions. Example evaluations are shown for the responses to IL-2/4/6 stimulation (Fig. 4B). Performance summary metrics for all stimulation conditions are provided in table S2. For benchmarking, we included the identity baseline (simply returning untreated distributions) and the observed training baseline (simply returning the perturbed distributions of the training patients). Across all metrics, SIMCO consistently outperformed the training baseline, validating that it captures patient heterogeneity and enables accurate individualized perturbation modeling (Fig. 4B).

Comparing simulated and observed marker distributions for example features with strong IL-2/4/6-induced stimulation response (0.1 arcsinh difference) highlighted the accuracy of SIMCO's digital twin models for heterogeneous single-cell behavior (Fig. 4C). The MAE of single-cell simulations generally increased with stronger signaling response (Fig. 4D). The simulation error correlated with interpatient variability in the observed signaling responses (Spearman's $R = 0.93$; Fig. 4D). However, SIMCO simulations still contained relative differences between patients even for the most variable responses (Fig. 4E).

SIMCO predicts treatment effects on time to labor

The main purpose of the SIMCO framework is to leverage single-cell–level digital twins to predict how immunomodulatory treatments influence clinical outcomes. Using an independent mass cytometry dataset collected by Stelzer *et al.* (*33*) over the past 120 days of pregnancy, SIMCO generated digital twins for 177 blood samples collected in term (153 samples, 58 patients) and preterm (14 samples, 5 patients) pregnancies (Fig. 5A). We first established that a sparse multivariable model built on the untreated single-cell digital twin dataset accurately predicted the clinical outcome time to labor [Pearson $R = 0.64$, root mean squared error (RMSE) = 24.1 days, term: 24.0 days, preterm of 25.4 days; Fig. 5B]. Furthermore, after estimating the expected pregnancy duration by combining the time to labor prediction with the sampling timepoint, SIMCO accurately differentiated term from preterm digital twins (AUROC = 0.75 [95% confidence interval = 0.59 to 0.88], AUPRC = 0.29 [0.12 to 0.54]; Fig. 5C and table S3). Notably, when comparing the performances of

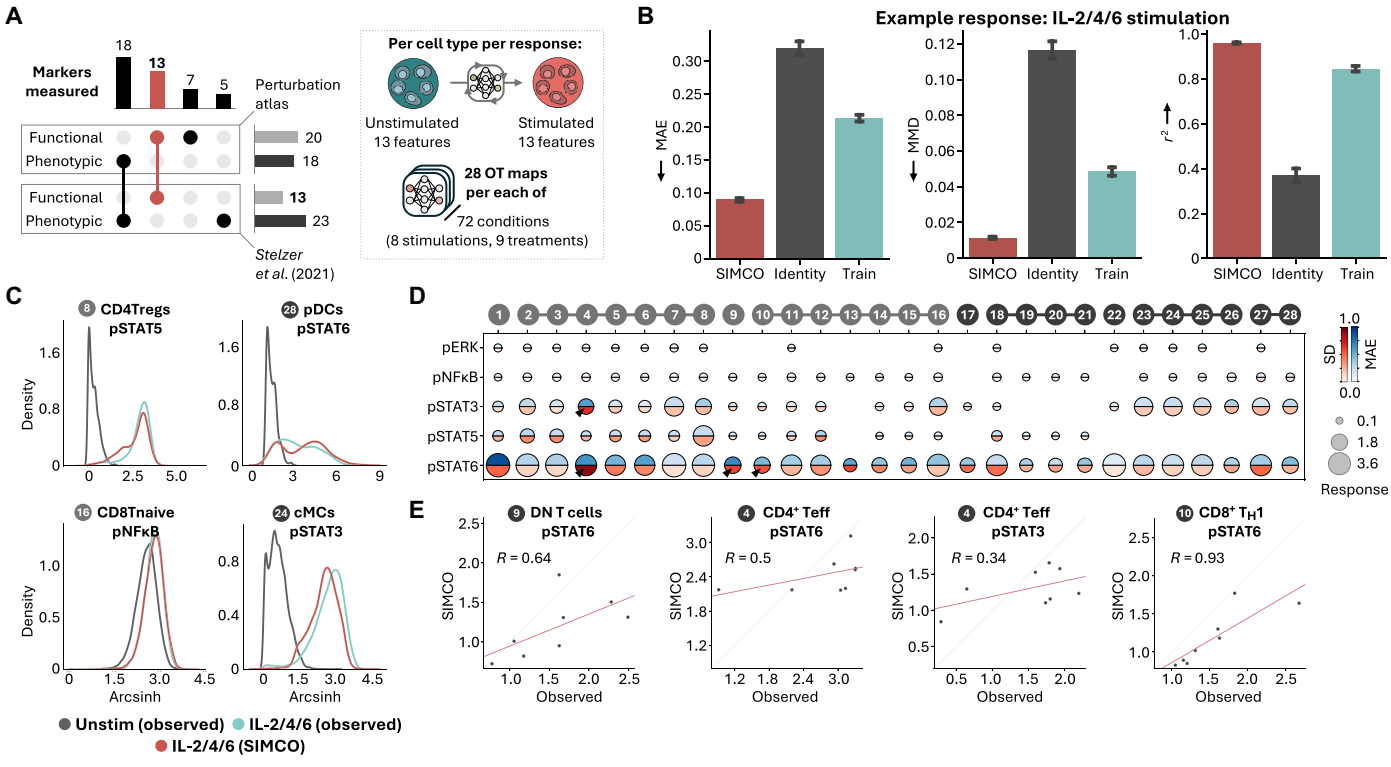


Fig. 4. Comparing observed and simulated immune responses to validate SIMCO's single-cell simulation performance. (**A**) Panel alignment between the generated single-cell perturbation atlas and the independent mass cytometry data from Stelzer *et al.* (*33*). In total, 13 functional protein measurements overlapped between the two data sets. Separate OT maps were learned per cell type and response, each predicting 13 functional protein marker changes. The cell types of the two studies were matched using the same gating strategy based on the 18 overlapping phenotypic markers. (**B**) To evaluate SIMCO's single-cell simulation performance, we compared it to identity and training mean performance in fourfold cross-validation within the single-cell perturbation atlas. MAE, MMD, and R^2 for the predicted population medians are shown for IL-2/4/6 stimulation responses across all cell types (364 features). (**C**) Distributions of protein expression at baseline and in response to IL-2/4/6, comparing observed and SIMCO-simulated single-cell distributions for selected features. (**D**) Heatmap of the MAE and the observed SD for features with an IL-2/4/6 response greater than 0.1 arcsinh difference. (**E**) Correlation (Spearman *R*) between SIMCO-simulated and observed feature values for the features with the highest ground truth standard deviation across $n = 8$ nonpregnant controls (indicated by arrows in D).

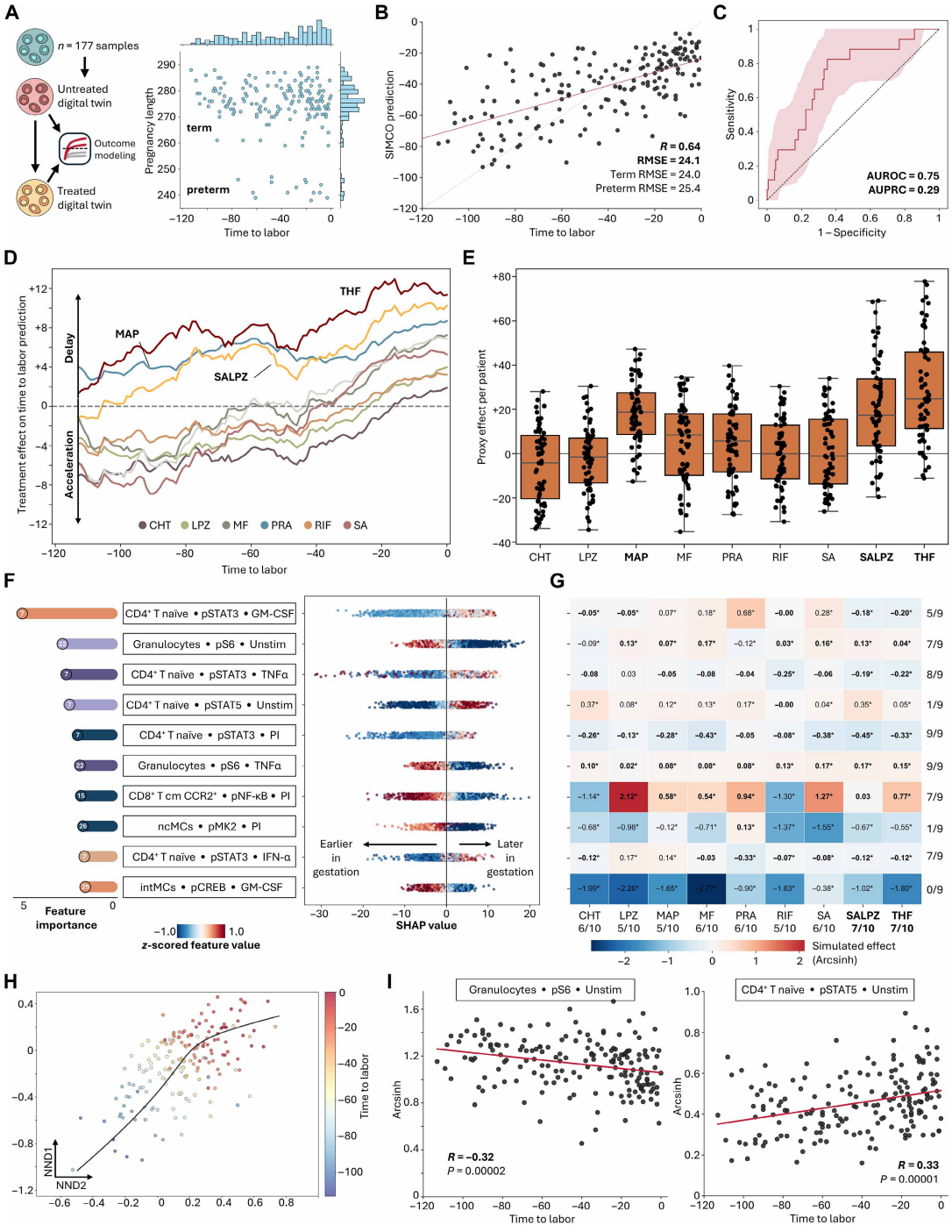


Fig. 5. SIMCO predicts treatment effects on time to labor. (**A**) SIMCO generated treated and untreated digital twins for the independent study from Stelzer *et al.* (*33*). The study cohort included 177 samples from 63 patients collected during the second and third trimesters of pregnancy (including 14 samples from 5 patients with sPTL). The primary outcome was the time to spontaneous labor onset. (**B**) SIMCO's performance to predict time to labor. Pearson *R* and root mean squared error (RMSE) are shown. (**C**) Area under the receiver operating characteristic curve (AUROC) plot for identifying preterm samples via the expected pregnancy length, calculated from SIMCO's time to labor prediction and the sampling timepoint. (**D**) Time-dependency (rolling average of 30 days) of SIMCO-predicted treatment effects across the last 120 days of pregnancy. (**E**) Proxy effect per patient ($n = 63$) on the gestational length for nine candidate drugs. (**F**) Top 10 model features (of 21 [18 to 30] features total) in descending order of mean Shapley Additive exPlanations (SHAP) value as a measure of feature importance in the SIMCO simulation. The SHAP summary plot reveals temporal association of *z*-scored feature values with gestational age. (**G**) Heatmap of average drug effects ($n = 177$) on the top 10 model features, as simulated by SIMCO. Statistical power (indicated by asterisks) was estimated empirically with Monte Carlo simulations using the observed SIMCO modeling error of each feature. Bold face indicates desired drug effects (i.e., opposing the trend toward onset of labor). Proportions of desired effects are given per candidate drugs and model feature. (**H**) Pseudotime representation of untreated samples ($n = 177$). (**I**) Trajectory of baseline example model features throughout gestation. Correlation was calculated using Spearman R. Parts of the figure were created in BioRender. E. Ganio (2026) (https://BioRender.com/qi8a8np).

SIMCO to the ground truth feature set of unstimulated data used to generate the digital twins, the addition of CellOT-simulated features yielded substantial gains in predictive performance (e.g., Pearson *R* increase from 0.34 to 0.61; AUROC of 0.69 to 0.75; AUPRC of 0.24 to 0.34; table S3). Even compared to a multivariate model built on the full ground truth feature set from Stelzer *et al.* (*33*), the addition of CellOT-simulated features modestly but consistently improved predictive performance (e.g., Pearson *R* increase of 0.79 to 0.82). Predicted drug effects and overall model performance were robust to alternative cross-validation strategies, including a stratified shuffle-groups-out cross-validation designed to leave one preterm patient out per fold (table S3).

SIMCO modeled the effect of the candidate drugs on the clinical outcome "time to labor" (see Materials and Methods): For each simulated sample, SIMCO calculated the difference in the predicted time to labor between candidate drug-treated and untreated digital twins (Fig. 5D). Maprotiline, THF, and SALPZ consistently delayed the predicted time to labor throughout all measured gestational time points. In contrast, other candidate drugs displayed a biphasic effect, first accelerating then decelerating the progression of gestation. To capture the cumulative effect of treatment over time, we aggregated data from the same patient across all three sampling time points (Fig. 5E). The strongest proxy effect on gestational length was achieved by THF [24.7 days, interquartile range (IQR) = 11.4 to 45.9], maprotiline (18.8 days, IQR = 8.6 to 27.5), and SALPZ (17.3 days, IQR = 3.5 to 33.8). Reflecting clinical reality, not all patients were modeled to benefit from a given treatment: THF, SALPZ, and maprotiline were predicted to prolong gestational length in 90.5, 87.3, and 82.5% of patients. Other candidate drugs, such as SA alone, only profited 46% of the patients.

The sparse prediction model for time to labor used 21 key informative immune features. Of the top 10 model features, five represented $CD4^+$ T cell signaling responses: Notably, reduced pSTAT3 levels in response to GM-CSF, TNFα, PI, and IFN-α was associated with earlier gestational time points, i.e., STAT3 activation increased with approaching labor (Fig. 5F). In the innate immune compartment, pS6 responses in granulocytes showed dynamic modulation over time, with both baseline and TNFα-induced pS6 levels decreasing as pregnancies approached labor.

SIMCO also enabled mechanistic insight into individual drug effects: To delay the predicted onset to labor, candidate drugs shifted key immune features of the predictive model toward patterns associated with earlier pregnancy stages. THF and SALPZ did so by consistently suppressing pSTAT3 signaling in $CD4^+$ naïve T cells and enhancing pS6 responses in granulocytes, thereby counteracting immunological signatures linked to pregnancy progression (Fig. 5G). Overall, THF and SALPZ influenced seven of the top 10 features in the desired direction. Maprotiline and THF also boosted phosphorylated NF-κB (pNF-κB) signaling after PI stimulation in $CD8^+$ $CCR2^+$ central memory T cells, which was higher in earlier gestation. These modeled effects targeted immune dynamics that showed consistent and predictable evolutions throughout gestation. Multivariate, supervised pseudotime analysis using a neural network to encode time to labor organized samples coherently by gestational age in the latent space (Fig. 5H), further supporting the interpretable temporal structure of these model features. The temporal structure was further corroborated by individual feature trajectories that illustrated the consistent evolution of signaling activity throughout gestation (Fig. 5I).

However, some drug effects paradoxically shifted features toward patterns associated with later pregnancy stages. For instance, THF and SALPZ, strongly inhibited pMK2 activation in non-CMs (ncMCs) following PI stimulation and pCREB activation in intermediate monocytes (intMCs) following GM-CSF stimulation, whereas decreased activation in these features would normally be expected closer to labor onset. Similarly, both chlorthalidone and SALPZ induced baseline activation of pSTAT5 in $CD4^+$ naïve T cells, a feature the model associated with later stages of pregnancy. However, SIMCO intrinsically integrates and prioritizes these multiple effects when generating outcome predictions for the treated digital twin. By integrating many immunomodulatory effects across cell types, stimuli, and signaling pathways, SIMCO provides a quantitative, system-level evaluation of each candidate drug's net impact on gestational timing.

DISCUSSION

SIMCO is a single-cell–level digital twin framework developed to model clinical outcome effects of immunomodulatory candidate drugs. In a first step, SIMCO learns immune mechanisms based on a high-dimensional single-cell perturbation dataset. In a second step, the trained model is parametrized with single-cell data of unseen individuals. By contrasting treated and untreated digital twin models, SIMCO quantifies how interventions alter clinical outcomes while offering single-cell–level mechanistic explainability. We showcase the application of SIMCO for sPTL prevention, a setting where in silico clinical trials are especially valuable given the constraints of interventional studies in pregnancy. Simulating treatment with nine preselected candidate drugs for the prevention of sPTL, SIMCO identified THF and SALPZ as strong candidates to delay labor onset in preterm pregnancies.

We used SIMCO with the premise that altering labor-associated immune signatures could delay labor onset and therefore prevent sPTL. While the immune system continuously adapts to the progression of gestation, an immunological intervention should sustain mid-gestational maternal immunotolerance, such as pSTAT5-driven regulatory T cell programs and immunosuppressive polarization in innate cells, while counteracting late-gestational trends of heightened NF-κB responsiveness, proinflammatory cytokine production, and innate leukocyte activation. By quantitatively evaluating drug-induced shifts of simulated immune states, SIMCO identified candidate drugs that stabilize or reverse immunological changes associated with the gestational progression toward labor. Trajectory analyses in the independent longitudinal cohort confirmed that the key signaling axes, including pSTAT5 in $CD4^+$ naïve T cells and pS6 in granulocytes, exhibited robust, monotonic trends across gestation, providing a rational basis for extrapolating drug-induced perturbations toward later gestational states. Notably, the effects were quantified based on a single-dose exposure; in practice, repeated administration (e.g., daily) would likely produce accumulating effects. Therefore, drug effects were aggregated on the patient level (across up to three samples) as a way to approximate the cumulative effect of continuous treatment over time. The three most effective drugs or drug combinations in delaying predicted time to labor were THF, maprotiline, and SALPZ. Of these, SALPZ, the combination of low-dose aspirin (81 mg) and lansoprazole (30 mg), is of particular interest: Aspirin treatment alone, when initiated between 6 and 12 weeks of gestation, has demonstrated modest reductions in the incidence of sPTL in large clinical trials (*25*, *39*, *40*). However, the small effect size has prompted ongoing

debate around optimal dosing, timing, and the potential for improved efficacy through combination with other agents. Lansoprazole, a proton pump inhibitor with known anti-inflammatory properties, inhibits LPS-induced, Toll-like receptor (TLR) 4–mediated NF-κB activation (*41*), a pathway implicated in inflammation-driven preterm labor (*42*). In a murine model of LPS-induced fetal wastage, lansoprazole treatment significantly improved fetal survival (*23*). Our findings further suggest that SA and lansoprazole act synergistically, particularly in the modulation of Janus kinase/signal transducer and activator of transcription (JAK-STAT) signaling. Although true synergy was observed only in a small proportion of features (3%), and most effects were additive (79%), the immunomodulatory profile of this combination provides a rationale for its clinical evaluation as a potential strategy to reduce sPTL incidence.

THF and maprotiline also delayed the simulated time to labor. While current guidelines recommend folic acid supplementation primarily for the prevention of neural tube defects (*43*), results from large retrospective studies show that folic acid supplementation is associated with a reduced risk of preterm delivery (*44*). Our digital twins suggest that folic acid/THF supplementation may benefit individuals at high risk for sPTL by delaying their immunological progression toward labor. Maprotiline is a discontinued tetracyclic antidepressant formerly used for treatment of major depressive disorder and anxiety. Per FDA guidelines, maprotiline was a category B drug that could be continued throughout gestation with no evidence of increased risk of sPTL (*45*, *46*). However, maprotiline use often comes with substantial side effects such as fatigue, weight gain, and anticholinergic effects. While it seems to mediate diverse immunomodulatory effects, its appeal as a preventive strategy for reducing sPTL risk is low due to its side effect profile and the general controversy around the use of antidepressants during pregnancy.

A key strength of the SIMCO framework is the ability to monitor simulated drug effects across the entire immune system and comprehensively weigh desirable and undesirable effects on the clinical outcome. Drugs that delayed the predicted time to labor predominantly modulated signaling in $CD4^+$ T cells and granulocytes. These modulated pathways demonstrated stable temporal trajectories across gestation and have established roles in the immune adaptations to pregnancy: Although naïve $CD4^+$ T cells are not classical effector populations during gestation, their signaling state critically shapes downstream immune differentiation, cytokine responsiveness, and the balance between regulatory and proinflammatory lineages. Notably, the suppression of pSTAT3 responsiveness in naïve $CD4^+$ T cells can promote immune tolerance and limit proinflammatory T cell differentiation (*47*, *48*). A sustained reduction in pSTAT3 activation might therefore constrain T helpe 17 (T_H17) or T_H1 skweing and protect against premature fading of maternal immunotolerance resulting in sPTL (*49*–*51*). Similarly, enhanced pSTAT5 responsiveness, which is central to regulatory T cells development and maintenance (*52*, *53*), may support maternal-fetal tolerance. Together, these subtle shifts in $CD4^+$ naïve T cell homeostasis observed in samples treated with THF, SALPZ, or maprotiline plausibly modulate the downstream inflammatory tone and support immunotolerant maintenance in mid-gestation.

Innate immune cells also show immunosuppressive adaptations to sustain pregnancy before taking part in the proinflammatory initiation of labor (*54*–*56*). Macrophages show an anti-inflammatory M2 polarization that helps maintain gestation to term (*57*) and granulocytic MDSCs expand, aiding directly with attenuating T cell activation (*58*). In this context, drugs that bolstered pS6 activity in granulocytes, e.g., metformin, SA, and SALPZ, beneficially influenced time to labor predictions in our simulations. While mechanistic target of rapamycin (mTOR) signaling is complex, activation of mTOR/pS6 activity has been shown to enhance IL-10 production and anti-inflammatory (M2-like) polarization in myeloid immune cells (*59*, *60*). Higher mTOR/pS6 signaling also protects from potentially placenta-damaging neutrophil extracellular traps at the maternal-fetal interface (*61*, *62*). Drugs that maintain higher mTOR pathway activity mid-pregnancy may thus help preserve immune tolerance and suppress labor-associated inflammation.

SIMCO enables meaningful data augmentation by extrapolating from unstimulated samples to infer responses under unmeasured conditions, thereby generating digital twins of the immunome. This approach assumes that the perturbation atlas is informative for the clinical dataset, as mismatched datasets from different patient populations could invalidate these extrapolations (*36*). By training CellOT models on the combined data from pregnant and nonpregnant individuals, we aimed to increase cellular diversity for improved generalizability and reduce overfitting to a small number of patients. Future pregnancy-specific and cross-gestational CellOT training could further refine the accuracy of learned single-cell behavior. While OT methods have seen growing use in basic science applications such as trajectory inference (*63*–*65*) and perturbation modeling (*66*–*68*), their adoption in translational research remains limited. Here, we showcase the utility of OT-based modeling within SIMCO to simulate gestational immune states and model outcome-level treatment effects. The strong predictive performance of the generated digital twins, approaching that of the original mass cytometry dataset from Stelzer *et al.* (*33*), indicates that SIMCO accurately models immune mechanisms underlying labor timing and bases its prediction of treatment effects on relevant immune responses. In addition, the robust performances across cross-validation folds demonstrate how CellOT likely does not overfit to individual patients but learns generalizable immune responses. While uncertainty from CellOT-based reconstruction can propagate through the SIMCO pipeline, most drug-induced feature shifts remained detectable above the noise level, although smaller effects should be interpreted cautiously, and SIMCO is best suited for comparative and mechanistic inference rather than precise quantitative prediction.

This study has several limitations. The limited number of preterm cases restricted the model's ability to capture preterm-specific immunological adaptations, and future studies with more samples from preterm pregnancies could enable targeted modeling. Although our immune perturbation atlas was profiled at the single-cell level, it included a relatively small number of individuals, captured only a restricted gestational window during the second trimester, and focused exclusively on peripheral blood, rather than tissue-resident immune populations. The experimental design reflects our goal of developing a minimally invasive, blood-based assay that can be scaled and incorporated in biomarker-driven clinical trial designs; however, it may overlook key tissue-specific immune dynamics involved in the pathophysiology of sPTL. Circulating immune signatures may reflect downstream correlates of underlying placental processes rather than upstream drivers of sPTL pathophysiology. Our work assumes that peripheral immune changes either mirror or approximate local immune effects within the same pathways (*10*, *61*, *62*, *69*, *70*). Yet, it is critical to recognize that candidate drugs may exert additional or alternative effects on local and systemic immunity beyond the model

features, thereby altering therapeutic benefits and constraining physiological immune defense mechanisms. Furthermore, our assay captures acute phospho-signaling responses but no functional readouts. While prior works have established some of these functional downstream effects, such as pSTAT3 enhancing IL-6 release while reducing IL-10 production in T_{reg} cells (*70*) or NF-κB–mediated release of TNFα from monocytes (*71*), further experiments are necessary to fully capture the drug-induced changes to the gestational immune landscape.

In addition, our in vitro experiments required selecting specific compounds concentrations for testing. Several of the investigated drugs are metabolized in the liver to active and inactive metabolites, a process that cannot be replicated in vitro. To approximate physiological relevance, we selected the metabolites with the highest plasma concentrations following oral administration (see Materials and Methods); for instance, for aspirin and folic acid, we tested SA and THF, respectively. While we cannot exclude the possibility that other metabolites also exert immunomodulatory effects, our experimental design prioritized those most likely to reflect systemic exposure. Effective concentrations strictly followed empirically observed plasma levels after oral administration of clinical standard doses. For THF, we chose a high dose (4 mg) as many patients reported taking prenatal vitamins, including THF at the recommended daily value.

Several key steps will be important to validate SIMCO as a workflow for drug screening. First, application to cross-gestational, larger cohorts as external validation will help assess the robustness and generalizability of inferred treatment effects across different populations, sampling protocols, and disease severities. Second, complementary in vivo studies in appropriate animal models could provide mechanistic validation of predicted immunomodulatory effects and their impact on gestational length and inflammatory pathways associated with preterm labor. Last, prospective clinical studies will enable direct comparison of digital twin–based predictions of individual treatment responsiveness with observed drug responses and clinical outcomes, providing an opportunity to evaluate SIMCO's predictive utility at the patient level. In addition, these clinical studies will confirm the safety profile and exclude any increased risk of infection-driven PTL resulting from immunomodulation. Together, these efforts will be critical to determine whether digital twin–predicted immune modulation translates into meaningful therapeutic benefit.

In conclusion, we present SIMCO as a proof-of-concept framework for single-cell–level digital twins that simulate immune responses and model outcome-level treatment effects. With the ability to generate digital twins from baseline immune profiles of unseen individuals, SIMCO offers a scalable and minimally invasive approach for early-stage compound prioritization and mechanistic hypothesis generation. Our work represents a critical step toward data-driven, immune-guided therapeutic strategies. Following validation in independent cohorts and prospective trials, SIMCO has the potential to transform precision obstetrics by unifying immune profiling, computational modeling, and personalized therapeutic strategies for the prevention of sPTL.

MATERIALS AND METHODS

Study participant details and experimental setup

Study design

The aim of this in vitro drug screen was to identify reproducible immunomodulatory effects of select candidate drugs for the prevention of sPTL. The study was conducted at Stanford Hospital (Stanford, CA, USA) and approved by the Institutional Review Board (protocol number 68417). All participants provided informed consent. Nonpregnant volunteers were recruited during the first week of their menstrual cycle. Inclusion criteria included age 18 to 45 years, BMI < 40, and no ongoing immunomodulatory treatment. Pregnant individuals were recruited between 20 + 0 and 27 + 7 weeks of gestation if they were 18 to 45 years old, had a prepregnancy BMI < 40, were carrying a singleton pregnancy, and without prior pregnancies (G1P0). Exclusion criteria included fertility treatment, chronic conditions outside of pregnancy (e.g., hypertension and diabetes), concurrent immunomodulatory medication, diagnosed pregnancy disorders and fetal anomalies, as well as autoimmune diseases (e.g., lupus, rheumatoid arthritis, Sjogren's, or Hashimoto's). A single blood draw of 12 ml was collected in heparinized tubes. In total, 8 nonpregnant controls and 11 pregnant individuals participated in the study. Researchers conducting the analyses were not blinded. Basic demographics and pregnancy characteristics including EGA at recruitment are summarized in Table 1.

Ex vivo drug screening assay

Whole blood was collected from study subjects in heparinized tubes and processed between 30 and 60 min after blood draw. The drug screen assessed 96 conditions per sample and was automated using an Opentrons Flex pipetting robot. Drug aliquots were prepared as 200× stocks from powder dissolved in DMSO (except 5-THF, 100× stock in 1:1 water/DMSO). In a 96-well deep-well block, RPMI with 2× drug concentrations was added and warmed to 37°C. Heparinized whole blood was added 1:1 to the RPMI/drug dilution in each well, achieving the final 1× concentration for each drug (0.5% DMSO). The final drug concentrations per drug can be found in Table 2. The whole blood pretreated with each drug was incubated for 1 hour at 37°C and 5% CO_2 in a humidified incubator on an orbital shaker. After 1 hour of preincubation with the drug to allow sufficient time for cellular uptake and target engagement, the samples were stimulated with LPS (50 ng/ml; InvivoGen, San Diego, CA); IFN-α (50 ng/ml; PBL Assay Science, Piscataway, NJ); IL-33 (50 ng/ml; R&D Systems, Minneapolis, MN); GM-CSF (50 ng/ml; R&D Systems); TNFα (50 ng/ml; R&D Systems); a cocktail of IL-2, IL-4, and IL-6 (each 50 ng/ml; R&D Systems); and PI (0.5×; Thermo Fisher Scientific, Waltham, MA) or left unstimulated for 2 hours. The 2-hour stimulation period was chosen to capture not only early signaling responses but also posttranscriptional and paracrine effects, including cytokine-mediated secondary signaling. After this 2-hour stimulation period at 37°C and 5% CO_2 in a humidified incubator on an orbital shaker, the whole-blood culture was fixed with proteomic stabilizer (SMART TUBE Inc., San Carlos, CA) for 20 min. Samples were treated with blood cell lysis buffer twice for 5 min, then washed, and stored at −80°C until further processing.

Drug metabolites and concentrations

All candidate drugs—chlorthalidone, maprotiline, metformin, lansoprazole, pravastatin, rifabutin, SA, and THF—were preselected with a computational drug repurposing pipeline (*23*) and categorized as FDA category A/B. For each candidate, we used the standard clinical oral dosing as anchor point (see Table 2). We identified in vivo studies that described plasma concentrations over 24 hours after oral administration of this reference dose and determined the empirically observed maximum concentration sustained for 3 hours, corresponding to the duration of our assay. From literature review, we selected the compound with the highest plasma concentration (either the drug itself or a metabolite) for testing. For the drug tested,

active metabolites were used for folic acid and aspirin administration (THF and SA, respectively).

96-plex layout and barcoding

All 96 screening conditions for each donor were organized in a standardized 12 × 8 format in a 2-ml deep-well block. Stimulation conditions were organized as rows, treatment conditions as columns. Three DMSO controls were included in column 1, 6, and 12, the candidate drugs were organized as depicted in Fig. 2A. A 96-plex doublet-removing barcoding scheme was optimized to allow pooling of all conditions together for mass cytometry staining and acquisition to avoid batch effects (*72*). Mass-tag barcodes were designed using the isotopes Pd104, Pd105, Pd106, Pd108, Pd110, Pt194, Pt195, Pt196, and Au197. Isotope concentrations were titrated to 10 nM (Au, Pt isotopes) and 100 nM (Pd isotopes) to achieve optimal barcode separation. Of the 126 barcodes available in a choose-4-out-of-9 scheme, the 96 barcodes with the highest purity (based on the purity of the underlying isotopes) were selected, while the remaining 30 were used to monitor the rate of wrongly assigned cells (fig. S3). Barcode reagents were prepared in DMSO and used after transient permeabilization with 0.02% saponin/phosphate-buffered saline (PBS) (*73*). Across all 19 blood donors, the average barcode assignment rate with a threshold of 0.1 was 71.8% (±4.2%). For each condition, 100 μl of blood yielded on average 138.0×10^3 (118.2×10^3 to 162.1×10^3) live leukocytes for pregnant donors and 94.6×10^3 (76.9×10^3 to 111.4×10^3) for nonpregnant donors.

Opentrons Flex operation

Protocol operation with the Opentrons Flex requires an eight-channel and a single-channel pipette. The protocol for the in vitro drug exposure and whole blood lysis required around 5 hours of operation. A custom tube rack and a pipette basin holder were three-dimensional printed to allow for the used deck configuration. The protocol for 96-well barcoding required around 2 hours of operation. The protocols and the .stl files for the custom parts are available on the GitHub repository.

Antibody staining and mass cytometry analysis

Pooled barcoded samples from all 96 conditions (40×10^6 cells) were incubated with Fc-block (Human TruStain FcX, BioLegend) for 10 min and then stained with a master mix of surface antibodies for 30 min in cell staining medium (PBS with 0.5% bovine serum albumin and 0.02% sodium azide). After washing, the cells were permeabilized in ice-cold 100% methanol, washed, and stained for 30 min with a master mix of intracellular antibodies (table S1). Following intracellular staining, cells were washed and resuspended in an iridium intercalator (Standard Biotools) solution containing 2% paraformaldehyde. Last, samples were washed with ddH_2O and pelleted. The data were acquired using a mass cytometry XT instrument (Standard Biotools). The resulting mass cytometry data were normalized across each run using 6-EQ normalization beads. The debarcoding of the data was achieved using the Single Cell Debarcoder v0.2 in MATLAB at a barcode separation threshold of 0.1 (*72*).

Stimulation responses and drug-induced effects

Using CellEngine (Primity Bio), cell populations were defined with manual gating (see fig. S2). Expression levels of intracellular phosphosignaling markers (pSTAT1, pSTAT3, pSTAT4, pSTAT5, pSTAT6, pMAPKAPK2, pCREB, pPLCγ2, pS6, pERK1/2, pP38, pNF-κB, and total IκB) and cell surface functional markers (PD-1, PD-L1, CD25, CD36, CD44, GLUT1, and HLA-DR) were exported as population median per stimulation condition, drug treatment, and study participant. Per individual, the combination of eight stimulation × 28 immune cell subsets × 20 functional proteins produced a total of 4480 features for each of the nine drug treatments and the three DMSO controls. The extracted dataset was arcsinh transformed with a cofactor of 5. The DMSO control triplicate was used to determine the intra-assay variability. Per patient per feature, a smooth spline was fitted across all columns ($k = 2$, smoothing = 1.5, weight = 1 for drug-columns, weight = 2 for DMSO columns) to account for time-dependent effects in signaling response (see fig. S7). Drug effects were calculated as difference to the spline-modeled column-specific corrected DMSO control within the same stimulation condition. Feature stimulation effects were classified as strong response if the arcsinh difference to unstimulated was >0.1 (0.01 for IL-33). For 168 of the 196 cell type/stimulation combination, at least one strong stimulation response was observed.

SIMCO workflow

Overview

We developed SIMCO to rapidly evaluate treatment effects of immunomodulatory drugs on a clinical outcome via single-cell simulations. The prediction ensemble uses CellOT to simulate single-cell immune states and Stabl for feature selection and outcome modeling. A dataset of high-dimensional perturbation experiments (e.g., from mass cytometry) should be generated in a relevant patient population, evaluating immune cell responses when challenged with s stimulatory ligands in the presence or absence of t candidate drugs (fig. S1A). After identifying cell types (e.g., using manual gating), cell type–specific CellOT models are trained for each stimulation scenarios s; separately, for each stimulation condition, CellOT models for each drug treatment scenarios *t* (fig. S1B). The learned single-cell behavior is then projected onto a separate immunome dataset with documented outcomes and matching single-cell readouts (e.g., antibody panel). For n unstimulated samples with p unstimulated features, responses are simulated to s stimulation scenarios, creating a feature matrix of $p \times (s + 1)$ features. This expanded feature matrix is then also transformed into t "treated" feature spaces (fig. S1C). Outcome modeling using the Stabl algorithm is performed on the untreated, simulation-expanded dataset in cross-validation: Per cross-validation fold, a predictive model is built on the training cohort to then predict the outcome on a hold-out test set of patients as well as their "treated" twins (fig. S1D). The global model performance is evaluated across folds; treatment effects were evaluated per patient per cross-validation fold as the difference between the predictions on untreated in treated (fig. S1E). The final model built on the entire dataset can be used to identify patients at risk in an unseen cohort. By using a specific model prediction as treatment target (continuous or categorical), the model can also steer personalized treatment strategies (fig. S1F).

Single-cell simulation using CellOT

CellOT was introduced by Bunne *et al.* (*37*) as an implementation of OT learning using an input convex neural network with four fully connected hidden layers. We adapted the network architecture to include a four-layer bottleneck structure (128, 128, 64, and 32). We reduced the latent-space dimensions from 50 to 42 to prevent overfitting and stopped training at 50,000 epochs (instead of 100,000). CellOT models were trained on single-cell data from individual cell types, identified by manual gating, as CellOT models trained per cell type achieved better predictive accuracy than CellOT models built on all leukocytes together.

Single-cell simulation performance evaluation

We evaluated CellOT's single-cell simulation performance using four-fold cross-validation with patient-level splits, excluding all cells of the held-out test patients from training. For each fold, CellOT was trained on a subset of patients, and predictions were generated for the unseen, held-out patients. Predicted single-cell data were summarized by calculating population medians per marker. To contextualize the performance, we compared CellOT against two simple baselines. In the identity baseline, predictions were set equal to the untreated (baseline) distribution of the same patient, representing the trivial case of assuming no perturbation effect was learned by CellOT. In the training baseline, predictions were set equal to the observed perturbed distributions from the training patients, representing the case of directly reusing available training data without modeling new responses. Performance metrics were then computed consistently across models: MAE between predicted and observed population medians; MMD between predicted and observed marker distributions per cell type, marker, and patient; and R^2 values quantifying the correlation between predicted and observed population medians across markers for each cell type and patient.

Final single-cell simulation model

For the final analysis, CellOT models were trained using all 19 patients (pregnant and nonpregnant) from the perturbation dataset. Our rationale for combining both groups was to increase cellular and phenotypic diversity in the single-cell training data, thereby improving robustness and reducing the risk of overfitting to pregnancy status–specific and sampling-time point–specific cell behavior rather than learning more general immune state transitions. For stimulation response models, we trained a separate CellOT model for each of the 168 cell-type stimulation combinations that showed at least one strong stimulation response. For treatment effect models, we extended the training to predict drug effects on the same 168 cell-type stimulation combinations and drug effects on all 28 cell types at unstimulated baseline.

Independent dataset with time to labor outcome

We used a longitudinal mass cytometry dataset recently published by our group as clinical reference dataset for simulation experiments (*33*). The dataset comprised 177 samples from 63 participants enrolled during the last 120 days of an uncomplicated, singleton pregnancy. Between one and three blood samples were collected throughout this period until patients went into spontaneous labor. For simulation experiments, only the 177 unstimulated samples (15 min, 37°C) were used (original stimulation conditions included unstimulated, IL-2/4/6, LPS, IFN-α, and GM-CSF). The antibody panel largely overlapped, and manual gating was adjusted to match the perturbation dataset completely. Population medians were exported and arcsinh transformed with a cofactor of 5.

Multivariable outcome modeling using Stabl

The Stabl algorithm allows for the selection of reliable biomarkers from multiomic datasets (*32*). As input data, population medians of baseline expression levels (unstimulated) or stimulation responses (arcsinh difference between stimulated and unstimulated) were used. To account for the longitudinal nature of the dataset, we used patient shuffle split to assess the robustness of our model. We applied a penalization matrix to remove weak stimulation responses (median response across all samples <0.1 arcsinh difference) from the input feature set. This penalization resulted in a dataset size of 364 unstimulated features (28 cell types × 13 markers) and 691 stimulation response features (of the 7 stimulation conditions × 28 × 13 = 2548 features). Using a grid search across multiple hyperparameters (table S4; all other hyperparameters used as default), we selected the Stabl parameters providing the best feature set based on the RMSE. For the final model fit, we then iterated over a grid search for XGBoost hyperparameters optimizing the AUROC (table S4).

Evaluation of simulated drug effects

Within the cross-validation strategy used to assess model performance, we calculated the time to labor predictions for both the held-out test patients and their "treated" counterparts. To reduce the effect of modeling error on the treatment evaluation, we compared the model prediction on the "treated" patient data to the same patient's "no drug" model prediction.

Computational resources

CellOT was executed using the Python package "CellOT" (version 0.1). CellOT training was performed on the Stanford-provided cloud computing platform Sherlock with maximum RAM request of 32 GB, maximum job time of 20 hours, maximum 16 CPU per job. Stabl was run using the Python package Stabl lw. Basic data analysis steps, including time-dependent variability correction and dataset penalization, were executed using the Python packages SciPy (version 1.10.4), scikit-learn (version 1.5.2), pandas (version 2.2.3), and numpy (version 1.26.4). For visualization, we used seaborn (version 0.13.2) and matplotlib (version 3.10.3).

Univariate analysis

Statistical comparisons between nonpregnant and pregnant individuals were calculated using a Mann-Whitney *U* test with false discovery rate (FDR) correction (FDR < 0.05, Benjamini-Hochberg). Comparisons between treatment and DMSO-control were calculated as paired Student's *t* test with Benjamini-Hochberg correction (FDR < 0.05). Pregnancy enrichment of drug effects was calculated as the absolute difference in effect size between pregnant and nonpregnant donors, normalized by the total absolute effect magnitude across both groups.

Supplementary Materials

This PDF file includes:

Figs. S1 to S7
Tables S1 to S4

REFERENCES

1. L. Liu, S. Oza, D. Hogan, Y. Chu, J. Perin, J. Zhu, J. E. Lawn, S. Cousens, C. Mathers, R. E. Black, Global, regional, and national causes of under-5 mortality in 2000–15: An updated systematic analysis with implications for the Sustainable Development Goals. *Lancet* **388**, 3027–3035 (2016).
2. R. L. Goldenberg, J. F. Culhane, J. D. Iams, R. Romero, Epidemiology and causes of preterm birth. *Lancet* **371**, 75–84 (2008).
3. R. Romero, S. K. Dey, S. J. Fisher, Preterm labor: One syndrome, many causes. *Science* **345**, 760–765 (2014).
4. X. Han, M. S. Ghaemi, K. Ando, L. S. Peterson, E. A. Ganio, A. S. Tsai, D. K. Gaudilliere, I. A. Stelzer, J. Einhaus, B. Bertrand, N. Stanley, A. Culos, A. Tanada, J. Hedou, E. S. Tsai, R. Fallahzadeh, R. J. Wong, A. E. Judy, V. D. Winn, M. L. Druzin, Y. J. Blumenfeld, M. A. Hlatky, C. C. Quaintance, R. S. Gibbs, B. Carvalho, G. M. Shaw, D. K. Stevenson, M. S. Angst, N. Aghaeepour, B. Gaudilliere, Differential dynamics of the maternal immune system in healthy pregnancy and preeclampsia. *Front. Immunol.* **10**, 1305 (2019).
5. N. Aghaeepour, E. A. Ganio, D. Mcilwain, A. S. Tsai, M. Tingle, S. Van Gassen, D. K. Gaudilliere, Q. Baca, L. McNeil, R. Okada, M. S. Ghaemi, D. Furman, R. J. Wong, V. D. Winn, M. L. Druzin, Y. Y. El-Sayed, C. Quaintance, R. Gibbs, G. L. Darmstadt, G. M. Shaw, D. K. Stevenson, R. Tibshirani, G. P. Nolan, D. B. Lewis, M. S. Angst, B. Gaudilliere, An immune clock of human pregnancy. *Sci. Immunol.* **2**, eaan2946 (2017).
6. A. L. Tarca, R. Romero, N. Benshalom-Tirosh, N. G. Than, D. W. Gudicha, B. Done, P. Pacora, T. Chaiworapongsa, B. Panaitescu, D. Tirosh, N. Gomez-Lopez, S. Draghici, S. S. Hassan,

O. Erez, The prediction of early preeclampsia: Results from a longitudinal proteomics study. *PLOS ONE* **14**, e0217273 (2019).
7. N. Gomez-Lopez, R. Romero, S. S. Hassan, G. Bhatti, S. M. Berry, J. P. Kusanovic, P. Pacora, A. L. Tarca, The cellular transcriptome in the maternal circulation during normal pregnancy: A longitudinal study. *Front. Immunol.* **10**, 2863 (2019).
8. T. T. M. Ngo, M. N. Moufarrej, M.-L. H. Rasmussen, J. Camunas-Soler, W. Pan, J. Okamoto, N. F. Neff, K. Liu, R. J. Wong, K. Downes, R. Tibshirani, G. M. Shaw, L. Skotte, D. K. Stevenson, J. R. Biggio, M. A. Elovitz, M. Melbye, S. R. Quake, Noninvasive blood tests for fetal development predict gestational age and preterm delivery. *Science* **360**, 1133–1136 (2018).
9. K. Racicot, J. Kwon, P. Aldo, M. Silasi, G. Mor, Understanding the complexity of the immune system during pregnancy. *Am. J. Reprod. Immunol.* **72**, 107–116 (2014).
10. R. Pique-Regi, R. Romero, A. L. Tarca, E. D. Sendler, Y. Xu, V. Garcia-Flores, Y. Leng, F. Luca, S. S. Hassan, N. Gomez-Lopez, Single cell transcriptional signatures of the human placenta in term and preterm parturition. *eLife* **8**, e52004 (2019).
11. A. L. Tarca, B. Á. Pataki, R. Romero, M. Sirota, Y. Guan, R. Kutum, N. Gomez-Lopez, B. Done, G. Bhatti, T. Yu, G. Andreoletti, T. Chaiworapongsa, DREAM Preterm Birth Prediction Challenge Consortium, S. S. Hassan, C.-D. Hsu, N. Aghaeepour, G. Stolovitzky, I. Csabai, J. C. Costello, Crowdsourcing assessment of maternal blood multi-omics for predicting gestational age and preterm birth. *Cell Rep. Med.* **2**, 100323 (2021).
12. J. H. Rowe, J. M. Ertelt, L. Xin, S. S. Way, Pregnancy imprints regulatory memory that sustains anergy to fetal antigen. *Nature* **490**, 102–106 (2012).
13. A. Erlebacher, Mechanisms of T cell tolerance towards the allogeneic fetus. *Nat. Rev. Immunol.* **13**, 23–33 (2013).
14. S. Tsuda, A. Nakashima, T. Shima, S. Saito, New paradigm in the role of regulatory T cells during pregnancy. *Front. Immunol.* **10**, 573 (2019).
15. D. A. Somerset, Y. Zheng, M. D. Kilby, D. M. Sansom, M. T. Drayson, Normal human pregnancy is associated with an elevation in the immune suppressive $CD25^+$ $CD4^+$ regulatory T-cell subset. *Immunology* **112**, 38–43 (2004).
16. N. Gomez-Lopez, S. Tanaka, Z. Zaeem, G. A. Metz, D. M. Olson, Maternal circulating leukocytes display early chemotactic responsiveness during late gestation. *BMC Pregnancy Childbirth* **13**, S8 (2013).
17. M. Farias-Jofre, R. Romero, J. Galaz, Y. Xu, L. Tao, C. Demery-Poulos, M. Arenas-Hernandez, G. Bhatti, Z. Liu, N. Kawahara, T. Kanninen, Z. Shaffer, T. Chaiworapongsa, K. R. Theis, A. L. Tarca, N. Gomez-Lopez, Pregnancy tailors endotoxin-induced monocyte and neutrophil responses in the maternal circulation. *Inflamm. Res.* **71**, 653–668 (2022).
18. G. Mor, P. Aldo, A. B. Alvero, The unique immunological and microbial aspects of pregnancy. *Nat. Rev. Immunol.* **17**, 469–482 (2017).
19. N. Gomez-Lopez, K. Motomura, D. Miller, V. Garcia-Flores, J. Galaz, R. Romero, Inflammasomes: Their role in normal and complicated pregnancies. *J. Immunol.* **203**, 2757–2769 (2019).
20. N. Gomez-Lopez, J. Galaz, D. Miller, M. Farias-Jofre, Z. Liu, M. Arenas-Hernandez, V. Garcia-Flores, Z. Shaffer, J. M. Greenberg, K. R. Theis, R. Romero, The immunobiology of preterm labor and birth: Intra-amniotic inflammation or breakdown of maternal–fetal homeostasis. *Reproduction* **164**, R11–R45 (2022).
21. D. Miller, V. Garcia-Flores, R. Romero, J. Galaz, R. Pique-Regi, N. Gomez-Lopez, Single-cell immunobiology of the maternal-fetal interface. *J. Immunol.* **209**, 1450–1464 (2022).
22. D. Feyaerts, M. Diop, J. Galaz, J. F. Einhaus, P. C. Arck, A. Diemert, V. D. Winn, M. Parast, C. Gyamfi-Bannerman, J. R. Prins, N. Gomez-Lopez, I. A. Stelzer, The single-cell immune profile throughout gestation and its potential value for identifying women at risk for spontaneous preterm birth. *Eur. J. Obstet. Gynecol. Reprod. Biol. X* **25**, 100371 (2025).
23. B. L. Le, S. Iwatani, R. J. Wong, D. K. Stevenson, M. Sirota, Computational discovery of therapeutic candidates for preventing preterm birth. *JCI Insight* **5**, e133761 (2020).
24. D. Feyaerts, I. Marić, P. C. Arck, J. R. Prins, N. Gomez-Lopez, B. Gaudillière, I. A. Stelzer, Predicting spontaneous preterm birth using the immunome. *Clin. Perinatol.* **51**, 441–459 (2024).
25. M. K. Hoffman, S. S. Goudar, B. S. Kodkany, M. Metgud, M. Somannavar, J. Okitawutshu, A. Lokangaka, A. Tshefu, C. L. Bose, A. Mwapule, M. Mwenechanya, E. Chomba, W. A. Carlo, J. Chicuy, L. Figueroa, A. Garces, N. F. Krebs, S. Jessani, F. Zehra, S. Saleem, R. L. Goldenberg, K. Kurhe, P. Das, A. Patel, P. L. Hibberd, E. Achieng, P. Nyongesa, F. Esamai, E. A. Liechty, N. Goco, J. Hemingway-Foday, J. Moore, T. L. Nolen, E. M. McClure, M. Koso-Thomas, M. Miodovnik, R. Silver, R. J. Derman, ASPIRIN Study Group, Low-dose aspirin for the prevention of preterm delivery in nulliparous women with a singleton pregnancy (ASPIRIN): A randomised, double-blind, placebo-controlled trial. *Lancet* **395**, 285–293 (2020).
26. Prediction and prevention of spontaneous preterm birth: ACOG Practice Bulletin, Number 234. *Obstet. Gynecol.* **138**, e65–e90 (2021).
27. Society for Maternal-Fetal Medicine (SMFM), SMFM Publications Committee, Society for Maternal-Fetal Medicine statement: Response to the Food and Drug Administration's withdrawal of 17-alpha hydroxyprogesterone caproate. *Am. J. Obstet. Gynecol.* **229**, B2–B6 (2023).
28. T. A. Manuck, C. Gyamfi-Bannerman, G. Saade, What now? A critical evaluation of over 20 years of clinical and research experience with 17-alpha hydroxyprogesterone caproate for recurrent preterm birth prevention. *Am. J. Obstet. Gynecol. MFM* **5**, 101108 (2023).
29. F. Pappalardo, G. Russo, F. M. Tshinanu, M. Viceconti, In silico clinical trials: Concepts and early adoptions. *Brief. Bioinform.* **20**, 1699–1708 (2019).
30. R. Laubenbacher, A. Niarakis, T. Helikar, G. An, B. Shapiro, R. S. Malik-Sheriff, T. J. Sego, A. Knapp, P. Macklin, J. A. Glazier, Building digital twins of the human immune system: Toward a roadmap. *npj Digit. Med.* **5**, 64 (2022).
31. S. C. Bendall, E. F. Simonds, P. Qiu, E. D. Amir, P. O. Krutzik, R. Finck, R. V. Bruggner, R. Melamed, A. Trejo, O. I. Ornatsky, R. S. Balderas, S. K. Plevritis, K. Sachs, D. Pe'er, S. D. Tanner, G. P. Nolan, Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696 (2011).
32. J. Hédou, I. Marić, G. Bellan, J. Einhaus, D. K. Gaudillière, F.-X. Ladant, F. Verdonk, I. A. Stelzer, D. Feyaerts, A. S. Tsai, E. A. Ganio, M. Sabayev, J. Gillard, J. Amar, A. Cambriel, T. T. Oskotsky, A. Roldan, J. L. Golob, M. Sirota, T. A. Bonham, M. Sato, M. Diop, X. Durand, M. S. Angst, D. K. Stevenson, N. Aghaeepour, A. Montanari, B. Gaudillière, Discovery of sparse, reliable omic biomarkers with Stabl. *Nat. Biotechnol.* **42**, 1581–1593 (2024).
33. I. A. Stelzer, M. S. Ghaemi, X. Han, K. Ando, J. J. Hédou, D. Feyaerts, L. S. Peterson, K. K. Rumer, E. S. Tsai, E. A. Ganio, D. K. Gaudillière, A. S. Tsai, B. Choisy, L. P. Gaigne, F. Verdonk, D. Jacobsen, S. Gavasso, G. M. Traber, M. Ellenberger, N. Stanley, M. Becker, A. Culos, R. Fallahzadeh, R. J. Wong, G. L. Darmstadt, M. L. Druzin, V. D. Winn, R. S. Gibbs, X. B. Ling, K. Sylvester, B. Carvalho, M. P. Snyder, G. M. Shaw, D. K. Stevenson, K. Contrepois, M. S. Angst, N. Aghaeepour, B. Gaudillière, Integrated trajectories of the maternal metabolome, proteome, and immunome predict labor onset. *Sci. Transl. Med.* **13**, eabd9898 (2021).
34. B. Bodenmiller, E. R. Zunder, R. Finck, T. J. Chen, E. S. Savig, R. V. Bruggner, E. F. Simonds, S. C. Bendall, K. Sachs, P. O. Krutzik, G. P. Nolan, Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat. Biotechnol.* **30**, 858–867 (2012).
35. N. Liu, W. E. Kattan, B. E. Mead, C. Kummerlowe, T. Cheng, S. Ingabire, J. H. Cheah, C. K. Soule, A. Vrcic, J. K. McInnich, S. Triana, M. Guzman, T. T. Dao, J. M. Peters, K. E. Lowder, L. Crawford, A. P. Amini, P. C. Blainey, W. C. Hahn, B. Cleary, B. Bryson, P. S. Winter, S. Raghavan, A. K. Shalek, Scalable, compressed phenotypic screening using pooled perturbations. *Nat. Biotechnol.* **43**, 1324–1336 (2025).
36. C. Bunne, G. Schiebinger, A. Krause, A. Regev, M. Cuturi, Optimal transport for single-cell and spatial omics. *Nat. Rev. Methods Primers* **4**, 58 (2024).
37. C. Bunne, S. G. Stark, G. Gut, J. S. Del Castillo, M. Levesque, K.-V. Lehmann, L. Pelkmans, A. Krause, G. Rätsch, Learning single-cell perturbation responses using neural optimal transport. *Nat. Methods* **20**, 1759–1768 (2023).
38. B. Abu-Raya, C. Michalski, M. Sadarangani, P. M. Lavoie, Maternal immunological adaptation during normal pregnancy. *Front. Immunol.* **11**, 575197 (2020).
39. A. J. E. M. C. Landman, M. A. De Boer, L. Visser, T. A. J. Nijman, M. A. C. Hemels, C. N. Naaktgeboren, M. C. Van Der Weide, B. W. Mol, J. O. E. H. Van Laar, D. N. M. Papatsonis, M. N. Bekker, J. Van Drongelen, M. G. Van Pampus, M. Sueters, D. P. Van Der Ham, J. M. Sikkema, J. J. Zwart, A. J. M. Huisjes, M. E. Van Huizen, G. Kleiverda, J. Boon, M. T. M. Franssen, W. Hermes, H. Visser, C. J. M. De Groot, M. A. Oudijk, Evaluation of low-dose aspirin in the prevention of recurrent spontaneous preterm labour (the APRIL study): A multicentre, randomised, double-blinded, placebo-controlled trial. *PLOS Med.* **19**, e1003892 (2022).
40. R. M. Silver, K. Ahrens, L. F. Wong, N. J. Perkins, N. Galai, L. L. Lesher, D. Faraggi, J. Wactawski-Wende, J. M. Townsend, A. M. Lynch, S. L. Mumford, L. Sjaarda, E. F. Schisterman, Low-dose aspirin and preterm birth: A randomized controlled trial. *Obstet. Gynecol.* **125**, 876–884 (2015).
41. T. Tanigawa, T. Watanabe, K. Higuchi, H. Machida, H. Okazaki, H. Yamagami, K. Watanabe, K. Tominaga, Y. Fujiwara, N. Oshitani, T. Arakawa, Lansoprazole, a proton pump inhibitor, suppresses production of tumor necrosis factor-α and interleukin-1β induced by lipopolysaccharide and Helicobacter pylori bacterial components in human monocytic cells via inhibition of 1 activation of nuclear factor-κB and extracellular signal-regulated kinase. *J. Clin. Biochem. Nutr.* **45**, 86–92 (2009).
42. A. Gilman-Sachs, S. Dambaeva, M. D. Salazar Garcia, Y. Hussein, J. Kwak-Kim, K. Beaman, Inflammation induced preterm labor and birth. *J. Reprod. Immunol.* **129**, 53–58 (2018).
43. US Preventive Services Task Force, Folic acid supplementation to prevent neural tube defects: US Preventive Services Task Force reaffirmation recommendation statement. *JAMA* **330**, 454–459 (2023).
44. Y. Wu, Y. Yuan, C. Kong, Q. Ma, H. Ye, W. Jing, J. Liu, M. Liu, The association between periconceptional folic acid supplementation and the risk of preterm birth: A population-based retrospective cohort study of 200,000 women in China. *Eur. J. Nutr.* **60**, 2181–2192 (2021).
45. C. Armstrong, ACOG guidelines on psychiatric medication use during pregnancy and lactation. *Am. Fam. Physician* **78**, 772–778 (2008).
46. J. Wang, D. Roy, M. Y.-F. Wang, D. Dinh, A. Lao, V. Mendoza, G. Xu, C. G. Chatterton, N. Ahmadieh, Timing of antidepressant use in pregnancy and preterm birth: A systematic review and meta-analysis. *O G Open* **1**, 22 (2024).
47. E. K. Deenick, S. J. Pelham, A. Kane, C. S. Ma, Signal transducer and activator of transcription 3 control of human T and B cell responses. *Front. Immunol.* **9**, 168 (2018).

48. L. Durant, W. T. Watford, H. L. Ramos, A. Laurence, G. Vahedi, L. Wei, H. Takahashi, H.-W. Sun, Y. Kanno, F. Powrie, J. J. O'Shea, Diverse targets of the transcription factor STAT3 contribute to T cell pathogenicity and homeostasis. *Immunity* **32**, 605–615 (2010).
49. F. Côté, E. Prairie, E. M. Sierra, C. Quiniou, T. Habelrih, W. Xu, B. Ferri, X. Hou, I. Lahaie, N. Côté, S.-E. Loiselle, L. Gobeil, K. Sawaya, A. Faucher, A. Beaulieu, S. Delisle, M.-P. Simard, M. A. Mohammad Nezhady, V. Laplante, A. Reuben, S. M. Kalaidji, E. Bajon, G. Cagnone, K. B. Leimert, J.-F. Gauchat, L. Gaudreau, S. Robertson, W. D. Lubell, D. M. Olson, S. Chemtob, A novel modulator of IL-6R prevents inflammation-induced preterm birth and improves newborn outcome. *EMBO Mol. Med.* **17**, 1950–1982 (2025).
50. L. Wang, Z. Liu, D. Huang, Y. Ran, H. Zhang, J. He, N. Yin, H. Qi, IL-37 exerts anti-inflammatory effects in fetal membranes of spontaneous preterm birth via the NF-κB and IL-6/STAT3 signaling pathway. *Mediators Inflamm.* **2020**, 1069563 (2020).
51. D. Marzioni, F. Piani, N. Di Simone, S. Giannubilo, A. Ciavattini, G. Tossetta, Importance of STAT3 signaling in preeclampsia (Review). *Int. J. Mol. Med.* **55**, 58 (2025).
52. L. Passerini, S. E. Allan, M. Battaglia, S. Di Nunzio, A. N. Alstad, M. K. Levings, M. G. Roncarolo, R. Bacchetta, STAT5-signaling cytokines regulate the expression of FOXP3 in $CD4^+CD25^+$ regulatory T cells and $CD4^+CD25^-$ effector T cells. *Int. Immunol.* **20**, 421–431 (2008).
53. M. A. Burchill, J. Yang, C. Vogtenhuber, B. R. Blazar, M. A. Farrar, IL-2 receptor β-dependent STAT5 activation is required for the development of $Foxp3^+$ regulatory T cells. *J. Immunol.* **178**, 280–290 (2007).
54. N. Gomez-Lopez, D. StLouis, M. A. Lehr, E. N. Sanchez-Rodriguez, M. Arenas-Hernandez, Immune cells in term and preterm labor. *Cell. Mol. Immunol.* **11**, 571–581 (2014).
55. I. Aneman, D. Pienaar, S. Suvakov, T. P. Simic, V. D. Garovic, L. McClements, Mechanisms of key innate immune cells in early- and late-onset preeclampsia. *Front. Immunol.* **11**, 1864 (2020).
56. S. Bert, E. J. Ward, S. Nadkarni, Neutrophils in pregnancy: New insights into innate and adaptive immune regulation. *Immunology* **164**, 665–676 (2021).
57. N. Gomez-Lopez, V. Garcia-Flores, P. Y. Chin, H. M. Groome, M. T. Bijland, K. R. Diener, R. Romero, S. A. Robertson, Macrophages exert homeostatic actions in pregnancy to protect against preterm birth and fetal inflammatory injury. *JCI Insight* **6**, e146089 (2021).
58. N. Köstlin, H. Kugel, B. Spring, A. Leiber, A. Marmé, M. Henes, N. Rieber, D. Hartl, C. F. Poets, C. Gille, Granulocytic myeloid derived suppressor cells expand in human pregnancy and modulate T-cell responses. *Eur. J. Immunol.* **44**, 2582–2591 (2014).
59. T. Weichhart, G. Costantino, M. Poglitsch, M. Rosner, M. Zeyda, K. M. Stuhlmeier, T. Kolbe, T. M. Stulnig, W. H. Hörl, M. Hengstschläger, M. Müller, M. D. Säemann, The TSC-mTOR signaling pathway regulates the innate inflammatory response. *Immunity* **29**, 565–577 (2008).
60. T. Weichhart, M. Hengstschläger, M. Linke, Regulation of innate immune cell function by mTOR. *Nat. Rev. Immunol.* **15**, 599–614 (2015).
61. A. Itakura, O. J. T. McCarty, Pivotal role for the mTOR pathway in the formation of neutrophil extracellular traps via regulation of autophagy. *Am. J. Physiol. Cell Physiol.* **305**, C348–C354 (2013).
62. H. Hu, Y. Peng, C. C. Wang, J. Chen, X. Yu, X. Chen, H. Ouyang, Q. Huang, J. Ma, Q. Yin, L. Ma, Z. Ding, M. Zhang, H. Ren, J. Zheng, W. Chen, Z. Tao, R. Liu, L. Chen, X. Wang, T. Zhang, M. Zhong, Neutrophil extracellular traps induce trophoblasts pyroptosis via enhancing NLRP3 lactylation in SLE pregnancies. *J. Autoimmun.* **153**, 103411 (2025).
63. A. Tong, J. Huang, G. Wolf, D. V. Dijk, S. Krishnaswamy, TrajectoryNet: A dynamic optimal transport network for modeling cellular dynamics, in *Proceedings of the 37th International Conference on Machine Learning* (PMLR, 2020), pp. 9526–9536.
64. S. Zhang, A. Afanassiev, L. Greenstreet, T. Matsumoto, G. Schiebinger, Optimal transport analysis reveals trajectories in steady-state systems. *PLoS Comput. Biol.* **17**, e1009466 (2021).
65. C. Bunne, L. Papaxanthos, A. Krause, M. Cuturi, Proximal Optimal Transport Modeling of Population Dynamics, in *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics* (PMLR, 2022), pp. 6511–6528.
66. J. Ryu, C. Bunne, L. Pinello, A. Regev, R. Lopez, Cross-modality matching and prediction of perturbation responses with labeled Gromov-Wasserstein optimal transport. arXiv:2405.00838 [q-bio.GN] (2024).
67. A. Demir, E. Solovyeva, J. Boylan, M. Xiao, F. Serluca, S. Hoersch, J. Jenkins, M. Devarakonda, B. Kiziltan, sc-OTGM: Single-cell perturbation modeling by solving optimal mass transport on the manifold of Gaussian mixtures. arXiv:2405.03726 [q-bio.GN] (2024).
68. M. Dong, B. Wang, J. Wei, A. H. De O, Fonseca, C. J. Perry, A. Frey, F. Ouerghi, E. F. Foxman, J. J. Ishizuka, R. M. Dhodapkar, D. Van Dijk, Causal identification of single-cell experimental perturbation effects with CINEMA-OT. *Nat. Methods* **20**, 1769–1779 (2023).
69. L. Fainboim, L. Arruvito, Mechanisms involved in the expansion of Tregs during pregnancy: Role of IL-2/STAT5 signalling. *J. Reprod. Immunol.* **88**, 93–98 (2011).
70. B. Liu, H. Wu, Q. Huang, M. Li, X. Fu, Phosphorylated STAT3 inhibited the proliferation and suppression of decidual Treg cells in unexplained recurrent spontaneous abortion. *Int. Immunopharmacol.* **82**, 106337 (2020).
71. M. Romão-Veiga, C. F. Bannwart-Castro, V. T. M. Borges, M. A. Golim, J. C. Peraçoli, M. T. S. Peraçoli, Increased TLR4 pathway activation and cytokine imbalance led to lipopolysaccharide tolerance in monocytes from preeclamptic women. *Pregnancy Hypertens.* **21**, 159–165 (2020).
72. E. R. Zunder, R. Finck, G. K. Behbehani, E.-A. D. Amir, S. Krishnaswamy, V. D. Gonzalez, C. G. Lorang, Z. Bjornson, M. H. Spitzer, B. Bodenmiller, W. J. Fantl, D. Pe'er, G. P. Nolan, Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. *Nat. Protoc.* **10**, 316–333 (2015).
73. G. K. Behbehani, C. Thom, E. R. Zunder, R. Finck, B. Gaudilliere, G. K. Fragiadakis, W. J. Fantl, G. P. Nolan, Transient partial permeabilization with saponin enables cellular barcoding prior to surface marker staining. *Cytometry A* **85**, 1011–1019 (2014).
74. J. V. Shah, D. P. Patel, P. A. Shah, M. Sanyal, P. S. Shrivastav, Simultaneous quantification of atenolol and chlorthalidone in human plasma by ultra-performance liquid chromatography–tandem mass spectrometry. *Biomed. Chromatogr.* **30**, 208–216 (2016).
75. K. P. Maguire, T. R. Norman, G. D. Burrows, B. A. Scoggins, An evaluation of maprotiline intravenous kinetics and comparison of two oral doses. *Eur. J. Clin. Pharmacol.* **18**, 249–254 (1980).
76. S. Eyal, T. R. Easterling, D. Carr, J. G. Umans, M. Miodovnik, G. D. V. Hankins, S. M. Clark, L. Risler, J. Wang, E. J. Kelly, D. D. Shen, M. F. Hebert, Pharmacokinetics of metformin during pregnancy. *Drug Metab. Dispos.* **38**, 833–840 (2010).
77. M. Song, X. Gao, T.-J. Hang, A.-D. Wen, Pharmacokinetic properties of lansoprazole (30-mg enteric-coated capsules) and its metabolites: A single-dose, open-label study in healthy Chinese male subjects. *Curr. Ther. Res. Clin. Exp.* **70**, 228–239 (2009).
78. M. M. Costantine, K. Cleary, M. F. Hebert, M. S. Ahmed, L. M. Brown, Z. Ren, T. R. Easterling, D. M. Haas, L. S. Haneline, S. N. Caritis, R. Venkataramanan, H. West, M. D'Alton, G. Hankins, Eunice Kennedy Shriver National Institute of Child Health and Human Development Obstetric-Fetal Pharmacology Research Units Network, Safety and pharmacokinetics of pravastatin used for the prevention of preeclampsia in high-risk pregnant women: A pilot randomized controlled trial. *Am. J. Obstet. Gynecol.* **214**, 720.e1–720.e17 (2016).
79. P. K. Narang, R. C. Lewis, J. R. Bianchine, Rifabutin absorption in humans: Relative bioavailability and food effect. *Clin. Pharmacol. Ther.* **52**, 335–341 (1992).
80. R. Dolores, N. Antunes, R. Moreno, P. Di Vaio, E. Magli, G. De Nucci, Comparative bioavailability study of two 81 mg coated tablet formulations of acetylsalicylic acid in fasting healthy volunteers. *J. Bioequiv. Availab.* **9**, 477–483 (2017).
81. X.-H. Zheng, L.-Y. Jiang, L.-T. Zhao, Q.-Y. Zhang, L. Ding, Simultaneous quantitation of folic acid and 5-methyltetrahydrofolic acid in human plasma by HPLC–MS/MS and its application to a pharmacokinetic study. *J. Pharm. Anal.* **5**, 269–275 (2015).

Acknowledgments: We thank all individuals for their participation in this study. We thank Y. El Sayed for his thoughtful input and guidance on the analyses. **Funding:** This work was supported by the Stanford Maternal and Child Health Research Institute; the March of Dimes Prematurity Research Center at Stanford University #22FY19343 (to B.G., D.K.S., and N.A.); the March of Dimes Prematurity Research Center at UCSF; the Center for Human Systems Immunology at Stanford (to B.G.); the Robertson Foundation (to B.G., N.A., and D.K.S.); the National Institute of Health P01HD106414 (to B.G., L.G., M.Si., and D.K.S.), K99HD115829 (to D.F.), and 1K99HD105016 (I.S.); and the Society for Reproductive Investigation Bayer Discovery/Innovation Grant (to D.F.). **Author contributions:** Conceptualization: J.E., M.Sat., A.S.T., V.B., D.K.G., K.A., I.A.S., N.A., B.G., M.Si., L.G., and D.K.S. Investigation: J.E., M.Sat., S.A., K.M., A.S.T., V.B., E.A.G., K.A., D.F., N.A., R.J.W., B.G., and M.Si. Formal analysis: J.E., P.N., O.F., J.N.A., V.B., M.Sab., E.A.G., I.A.S., R.J.W., and B.G. Methodology: J.E., A.S.T., J.H., V.B., D.K.S., G.M.S., T.T.O., M.Sab., K.A., I.A.S., N.A., and B.G. Data curation: J.E., M.Sat., A.S.T., V.B., E.A.G., K.A., I.A.S., O.F., and B.G. Validation: J.E., M.Sat., S.A., A.S.T., V.B., M.Sab., R.J.W., O.F., P.N., and B.G. Resources: M.Sat., S.A., A.S.T., A.C., E.A.G., K.A., V.B., R.J.W., G.M.S., I.A.S., M.Si., L.G., D.K.S., and B.G. Software: J.E., O.F., A.M., J.N.A., V.B., M.Sab., R.L., P.N., and B.G. Visualization: J.E., P.N., O.F., A.M., J.N.A., V.B., M.Sab., E.O.K., R.J.W., and B.G. Funding acquisition: A.S.T., G.M.S., M.Si., L.G., D.K.S., and B.G. Supervision: K.A., M.Si., L.G., D.K.S., and B.G. Project administration: J.E., M.Sat., S.A., A.S.T., L.G., and B.G. Writing—original draft: J.E., P.N., O.F., M.Sat., A.S.T., D.K.G., I.A.S., M.Si., L.G., and B.G. Writing—review and editing: J.E., P.N., O.F., M.Sat., S.A., K.M., J.N.A., K.A., V.B., D.K.G., M.Sab., D.F., M.D., A.S.T., A.C., E.A.G., R.L., E.O.K., I.A.S., J.H., R.J.W., Y.J.B., D.J.L., G.M.S., T.T.O., M.Si., L.G., D.K.S., N.A., and B.G. **Competing interests:** D.K.G., B.G., and J.H. are advisory board members at SurgeCare. J.H. is CEO of SurgeCare. J.E., D.K.S., M.Si., T.T.O., and B.G. are inventors on a patent application related to this work filed by Stanford University (no. PCT/US2026/017168, filed 27 February 2026). The authors declare that they have no other competing interests. **Data, code, and materials availability:** All data and code needed to evaluate and reproduce the results in the paper are present in the paper and/or the Supplementary Materials. All materials used in this study are commercially available, as specified in the supplementary tables. This study did not generate new materials. The preprocessed data of the perturbation dataset can be found as pre-gated FCS files on dryad.com (https://doi.org/10.5061/dryad.sqv9s4ngm). The source code required to reproduce analysis steps is publicly available on Github (https://github.com/ofondeur/SIMCO/) and zenodo.com (10.5281/zenodo.20481641). Further information and requests for resources and reagents should be directed to and will be fulfilled by B.G. (gbrice@stanford.edu).

Submitted 22 October 2025
Accepted 27 July 2026
Published 21 October 2026
10.1126/sciadv.aed2741