Pregnancy and postpartum dynamics revealed by an atlas of millions of lab tests

Alon Bar¹, Netta Mendelsohn-Cohen², Yael Korem Kohanim¹, Avi Mayo¹, Yoel Toledano³, Uri Alon¹

¹ Dept. Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel
² Dept. of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel
³ Division of Maternal Fetal Medicine, Helen Schneider Women’s Hospital, Rabin Medical Center, Petah Tikva, Israel

Abstract
Pregnancy involves dynamic alterations in most physiological systems. Delivery initiates a transition to postpartum recovery, which similarly exhibits widespread dynamic changes. Physiological dynamics during pregnancy and after delivery have not been systematically analyzed at high temporal resolution in a large human population. Here, we present a dynamic atlas of 78 laboratory tests with weekly time resolution, spanning the period from 20 weeks preconception until 80 weeks postpartum, totaling 45 million measurements from over half a million pregnancies. Tests exhibited discontinuities and rebound effects after delivery, and showed extended recovery times of up to a year postpartum. The accuracy of the data revealed effects of preconception supplements and intricate temporal responses to changes in blood volume and renal filtration rate. The dynamical data allowed discrimination between distinct mechanistic models of adaptation in the thyroid axis. These results provide a comprehensive dynamic portrait of the systems physiology of pregnancy.

Introduction
During pregnancy, the mother undergoes physiological changes that support fetal growth and development. The cardiovascular, respiratory, renal, gastrointestinal, skeletal, metabolic, endocrine and immune systems are all affected by fetal demand and massive endocrine secretion by the placenta¹⁻⁴. Increased demand for oxygen and nutrients cause an increase in cardiac output and up to 50% growth in blood volume ¹. The kidney increases glomerular filtration rate by 50%, leading to increased urine production ¹. The immune system is modulated to prevent rejection of the fetus, and coagulation and red blood cells show marked changes ¹,²,⁵. Metabolism shifts to increased insulin resistance and lipid production in order to supply energy for fetal growth ⁶.

Delivery marks a profound change, as the fetus and placenta exit the body and abruptly cease their metabolic and endocrine effects. The mother accordingly undergoes a series of adaptations in which various physiological systems recover with different timescales - from hours to months ⁷. Pregnancy and postpartum periods have increased risk of pathologies including gestational diabetes, thyroid disorders, anemia, depression and eclampsia ⁸. Understanding healthy physiology is essential for both advancing basic science and as a baseline for studying pathology. To understand the physiological changes during pregnancy and postpartum requires accurate temporal data on numerous physiological parameters. However, almost all studies have a limited number of participants, consider only a few parameters, and have low temporal resolution,
typically of one time point per trimester. This is even more salient for the postpartum period in which a single time point is usually measured. Meta-analyses have collected these smaller studies to construct normal ranges for tests in each trimester\(^9,10\). All together, our knowledge of the physiological time course is thus limited to low temporal resolution.

Here we harness a large national health record database\(^11\) to study over half a million pregnancies in terms of 78 major lab tests, totaling over 45 million tests, and present this information at weekly resolution as a resource. We identify global dynamical trends and show how this accurate data can illuminate physiological mechanisms of compensation.

**Results**

**Lab tests at weekly resolution throughout pregnancy and postpartum**

To obtain an atlas of lab tests in pregnancy and postpartum, we utilized a Clalit healthcare database\(^11\) that includes about half of the pregnancies in Israel in the period 2003-2020 with broad socioeconomic representation. We considered the first pregnancies of 588,964 women aged 25 to 35 and analyzed 78 common lab tests starting from 60 weeks before delivery to 80 weeks after delivery (Methods) (Fig 1). Each test had between 7.5K and 1.5M data points, totalling 45.5M data points (Fig 1A).

We binned the tests into weekly bins; each weekly bin included 20 to 2900 data points. As reference values we used tests from an age-matched female cohort from the same Clalit dataset\(^12\) (Methods).

As in all electronic health record studies, it is crucial to control for ascertainment bias, because many of the tests were done for specific medical reasons. To do so, we compared the present data to a previous meta-analysis of lab tests in healthy women done at six time points during pregnancy with one time point postpartum\(^9\). Most tests showed strong and significant correlations with this study (16/21 of tests in this comparison had R^2>0.5, p<0.05). We also developed an ascertainment bias score based on the number of tests, participant age and BMI (Methods). We study the 78 tests (out of an initial 112 tests) which we consider to be not strongly biased (Methods). The data on all 112 tests is provided online and in the SI (SI section S2). To avoid outlier values and subpopulations with extreme test results, we performed quantile analysis\(^12\) and transformed the mean quantile scores into test values (FIG 1BC) (Methods).
All 78 tests showed significant temporal variation across pregnancy and postpartum (Methods), Fig 1C. The variation over time in mean test value ranged between a few percent to hundreds of percent for different tests. This variation was proportional to the variation of each test within the reference population - tests that vary widely between individuals also varied strongly over time in pregnancy and postpartum (Fig 2A) (Methods). Thus, homeostatic processes seem to scale the variation in pregnancy so that tests remain within their physiological range. Mean test value varied over time (peak to trough) by 5 to 52 percentile points relative to reference, with a mean variation of 29 percentiles (Fig 2B).
Fig 2. Main trends of lab tests during pregnancy and postpartum. A) Distribution of maximal quantile variation over time of the tests. B) Maximal percent change of a test during pregnancy is approximately proportional to the variation of the absolute test value in the reference population, CV is standard deviation divided by mean in percentages. C-F) Four clusters of normalized test dynamical profiles, gray are individual tests, colored lines are their mean. G-H) Theoretical profiles for first order response I-J) Theoretical profiles for a system with a slow compensatory mechanism. K) (1) Circuit in which the load of pregnancy affects test X as a first order system. (2) Circuit in which that load affects test Y with X as a compensatory system. L) Pregnancy load in the theoretical model rises in pregnancy and drops abruptly at delivery. M) First order system X responds continuously, N) Compensated system Y shows a discontinuity and rebound effect.

Test dynamics can show jumps and rebounds after delivery
In order to study the types of dynamics, we clustered tests according to their normalized temporal profiles (Z-scored in the pregnancy and postpartum periods separately to emphasize the dynamics in both periods, Methods) (Fig 2C-F). We found four clear clusters, which define four mean profiles. Two of the profiles are continuous: Profile 1 rises during pregnancy and drops
postpartum in a continuous rise and fall, and prolife 2 is its mirror image, declining in pregnancy and rising postpartum. Profiles 3 and 4 show a discontinuous rebound effect. Profile 3 rises during pregnancy but jumps to low values after delivery and then recovers, and prolife 4 is its mirror image.

To understand the origin of continuous versus rebound dynamics, we explore canonical physiological mechanisms ([Methods](#)). Pregnancy exerts a load on physiological variables in order to meet the needs of the developing fetus. This load pushes physiological variables away from their normal set points or adjusts new set points that reflect the changing physiological priorities. Upon delivery this load is suddenly relaxed ([Fig 2K,L](#)). Variables with continuous dynamics can be explained by first-order recovery to baseline ([Fig 2K,M](#)). The system is pushed by the load away from steady state during pregnancy, and then recovers exponentially with a typical timescale. There is no rebound.

In contrast, a rebound can occur when there exists an additional, slowly varying compensation mechanism ([Fig 2K,N](#)). During pregnancy this mechanism strengthens to keep the variable from moving too far from its setpoint. Upon delivery the load is suddenly reduced but the compensation mechanism is still strong - causing overcompensation that induces a sharp rebound. Return to baseline is governed by the return of the compensation mechanism. A mathematical description of these two possibilities fits the dynamic curves well ([Fig 2G-J](#)). Below, we discuss in more detail a specific example of such compensation mechanisms in the thyroid axis.

**Physiological changes showed slow postpartum recovery**

The typical time courses thus show an asymmetry in which postpartum dynamics can have a different shape than the dynamics during pregnancy. To quantify this, we consider the data as a vector of 78 mean lab test values at each of 140 weekly timepoints. To study the global temporal trajectory, we reduced dimensionality using principal component analysis ([SI section 5](#)). The trajectory shows hysteresis - tests change during pregnancy but return to baseline via a different trajectory postpartum ([Fig 3A](#)).

Postpartum adaptation has two main phases - a rapid return 10 weeks after delivery followed by prolonged recovery. Many tests take months to return to baseline after delivery. To quantify the timing of return we calculated the settling time for each test. The settling time is defined by time after which the test remains within 0.2 SDS of its postpartum baseline ([Fig 3B](#)) ([Methods](#)).

A third (26/78) of the tests have long settling times that exceed 10 weeks ([Fig 3C,D](#)). Among these tests are liver functions aspartate transaminase (AST) and alanine transaminase (ALT) that take about half a year to recover, metabolic factors such as cholesterol, and alkaline phosphatase that settles only after about a year ([Fig 3D,E](#)). About a half (41/78) of the tests settle rapidly within a month to 10 weeks after delivery ([Fig 3C,D](#)).

Slow settling times can arise from several factors. Metabolism may follow BMI that settles over months. Breastfeeding may also affect some tests, such as calcium, phosphate, PTH and prolactin.
Fig 3 Postpartum recovery times of tests range between days and a year. A) Dimensionality reduction of test mean values as a function of time using PCA shows that trajectory during pregnancy differs from the postpartum trajectory (PCA 1 and 3 were chosen for graphical reasons, SI). B) Settling time is defined as the time after which the test remains within 0.2 SDS of its postpartum baseline. C) Distribution of settling time for the tests in the atlas. D) Tests ordered by settling time with confidence intervals estimated by bootstrapping. E) Settling time for the tests arranged by the physiological system.
Compensatory thyroid dynamics explained by changes in thyroid mass

To test the potential of this high-resolution dataset to address physiological mechanisms, we choose a clinically relevant system with a calibrated mechanistic mathematical model - the thyroid axis. The thyroid gland produces thyroid hormone T4 which regulates metabolic rate and is crucial for fetal development 17,18 (Fig 4A, blue).

The thyroid axis faces strong challenges during pregnancy 17,19,20. The fetus takes up maternal T4 during the first trimester and takes up maternal iodine in order to produce T4 in the second and third trimesters. Maternal thyroid hormones are degraded in the placenta, diluted by hemodilution, removed by increased GFR and sequestered by increased levels of thyroid binding globulin (TBG). These factors increase the demands on the maternal thyroid axis. To partially compensate, the hormone hCG from the placenta acts as a weak analogue of thyroid stimulating hormone (TSH) to stimulate thyroid hormone secretion 21.

The physiological understanding of the thyroid axis is based on a three-hormone cascade, the hypothalamic-pituitary-thyroid axis (HPT axis, Fig 4A, blue)22,23. The hypothalamus secretes thyrotropin releasing hormone (TRH) that induces the pituitary thyrotroph cells to secrete TSH, which induces the thyroid to secrete T4. Thyroid hormone acts in a negative feedback loop on the secretion rate of the two upstream hormones TSH and TRH.

This classical model was recently updated to include changes in the thyroid and pituitary gland functional mass 24 (Fig 4B, green, Methods). This model can account for example for growth of the thyroid when iodine is low, a phenomenon called goiter 25,26. The changes in gland mass are caused by the hormones in the axis that act as growth factors: TSH stimulates thyroid growth whereas T4 inhibits pituitary thyrotroph growth 27,28. Since gland mass grows slowly on the timescale of months, these growth interactions add a slow timescale to the classical model.

To compare these two models, we use the present high resolution data in the postpartum period. The thyroid hormones settle to within 0.2SDS of their normal range within a few weeks. Their detailed dynamics, however, has a slow time scale of months. Mean Free T4 (FT4) - the biologically active T4 unbound to TBG - is high after delivery and then undershoots for about 50 weeks, before rising back to baseline (Fig 4C). TSH begins low and adapts over 30 weeks (Fig 4D). The classical model predicts much faster adaptation, with a time scale determined by the halflife of the hormones, about 7 days for FT4 and 1h for TSH 29,30. Furthermore, the classical model predicts a unique value of FT4 for each level of TSH, whereas the data shows a hysteresis-like curve where a given TSH corresponds to different FT4 levels (Fig 4E) at different times.

The gland-mass model, in contrast, captures the postpartum dynamics well, including the FT4 undershoot (Fig 4CDE) (R=0.9, p-value<10^-3) (Methods). It predicts that upon delivery the thyroid is about 10% larger than normal (11.4% CI [10.1%-12.8%]), in agreement with clinical measurements 31–33 (Fig 4F). The thyroid mass normalizes with a timescale given by its cellular turnover time 24,28,29 of about 3 months (85 days CI [76 days - 95 days]). The pituitary thyrotroph functional mass in the model is roughly normal postpartum (Fig 4F). Thus, growth of the thyroid gland acts as a compensation mechanism that contributes to the rebound dynamics of thyroid hormones and can explain their slow adaptation timescale.

We conclude that the high resolution data can help to distinguish between physiological mechanisms, and provides a testing ground for mechanistic models that form the basis of systems physiology.
Fig 4. Thyroid dynamics can reject classical mechanisms and support compensation by gland mass growth. A) Classical thyroid axis model shows the hormone cascade and negative feedback by thyroid hormones on upstream hormone secretion. B) Gland mass model adds to the classical model the effects of hormones as growth factors (TSH on thyroid) or growth inhibitors (T4 on pituitary thyrotropes) that change the functional mass of the glands. C) Postpartum dynamics of free T4 and D) TSH is not captured by classical model and is captured by gland-mass model. E) T4 versus TSH dynamics, arrows from gland-mass model. F) Relative change in thyroid gland and pituitary thyrotroph mass in gland-mass model.

Tests affected by pregnancy supplements show preconception dynamics
We noticed that about a third of the tests (24/78) show significant dynamical trends before conception, in the period of 60 to 38 weeks before delivery ([Fig 5A](#)) ([Methods]). One of the strongest changes is a rise in folic acid ([Fig 5B](#)). Folic acid supplements are taken in the months before conception by about half of the relevant population 34–36.

We reasoned that supplements such as folic acid, vitamin B-12 and iron might exert physiological effects on other test values. Indeed, most of the preconception variation is explicable based on known effects of folic acid and other preconception supplements. These changes include positive effects on anemia, anticoagulative effects and lowering of lipids 37–44.

Some of the changes seen in preconception are not easily attributed to known effects of supplements. This includes changes in some immune cell counts, ALT, AST, Na, Urea and Urine pH. One possibility is that these tests are affected by yet unknown mechanisms by supplements, or that they are affected by other preconception health behaviors such as reduced rates of smoking, alcohol consumption and improved diet 46.

We conclude that the accuracy of the present atlas allows the detection of preconception changes that may correlate with health behaviors.
Fig 5. Tests affected by preconception health behaviors show preconception dynamics (effect size>0.1, p<0.05 FDR) A) Preconception slope of lab tests arranged by physiological system showing nonzero preconception slope (effect size>0.1, p<0.05 FDR). B) Folic acid test values show strong preconception rise. Conception is indicated by a dashed line. The regression line between test values and time preconception is shown. The inset highlights the preconception period.

System by system overview of pregnancy and postpartum dynamics

We next provide a system-by-system overview of the physiological dynamics observed in the atlas (Methods) and relate them to previous knowledge. The atlas captures the major known trends and reveals previously unknown temporal changes. See SI (Section S2) for the full dataset including absolute test value distributions at each timepoint.

Kidney

Renal physiology is affected by pregnancy in multiple ways. There is an increase in renal volume, blood flow and glomerular filtration rate (GFR), as well as an increase in reabsorption of nutrients and electrolytes. After delivery, diuresis causes elimination of excess fluids within days and GFR returns to normal within 6-8 weeks. Kidney volume might take up to 6 months to decrease to prepregnant volume.

The kidney removes waste products from the blood including creatinine, urea and uric acid. The atlas shows that the corresponding test values drop sharply in the first trimester, consistent with the temporal profile of increased GFR. Urea and creatinine remain low until delivery, whereas uric acid rises in the last trimester, peaking at delivery, and normalizes over months postpartum. Blood electrolytes (Na, Cl, K, Mg) show a drop at the first trimester. They remain low until delivery, consistent with the GFR profile, and show a mild overshoot postpartum, with the exception of chloride, which normalizes in the second trimester. The observed patterns of electrolytes are in line with the known decrease in serum osmolarity.

Urine composition, controlled by the kidneys, is also subject to multiple changes. The atlas shows that urinary pH becomes more alkaline during pregnancy. It shows a pattern that is inverse to other renal blood tests such as creatinine, urea, and sodium, including an undershoot postpartum. Specific gravity, which reflects urine particle concentration, shows a gradual decline during pregnancy, which is resolved within several weeks postpartum. (Fig 6)
Liver
The liver changes to support the growing fetus and the metabolic demands of the mother's body. Placental hormones, such as estrogen and progesterone, affect hepatic metabolic, synthetic and excretory functions. The liver produces most of the blood proteins including albumin and globulin. The atlas shows that albumin and total protein concentration tests drop gradually during the first trimester, and stay low until delivery and then recover over a few weeks postpartum. This follows the temporal profile of blood volume growth, and may be attributed to uncompensated hemodilution. In contrast, globulin stays relatively constant. Bilirubin tests report on the ability of the liver to clear bilirubin. They show a previously unreported U-shape curve, with a nadir around mid-pregnancy.
Liver enzyme tests GGT, ALT and AST are used to detect liver damage. In the atlas, their dynamics differ - GGT shows a drop in pregnancy similar to bilirubin and rapid recovery, whereas ALT and AST do not drop as much during pregnancy and show a pronounced overshoot after delivery. This is surprising since AST and ALT are usually considered unchanged in pregnancy. The overshoot in ALT and AST levels might reflect a period of postpartum liver remodeling due to residual liver stiffness, or might be related to muscle remodeling since these two enzymes are also found in muscle. (Fig 7)
Musculoskeletal system

Weight gain and the growing fetus impose forces and biochemical stress on the maternal skeleton, muscles and ligaments. Demand for calcium and phosphorus rises during pregnancy and lactation, and bone metabolism is affected by estrogen deficiency during lactation. Bone mass recovery can take up to a year after delivery.

The atlas shows that indicators of muscle and bone damage or turnover, including alkaline phosphatase (ALP), lactate dehydrogenase (LDH) and creatinine kinase (CK), vary strongly over pregnancy and postpartum. ALP and LDH drop in the first trimester and then rise to peak around delivery. Postpartum, the tests return to baseline, with LDH normalizing faster than ALP. The rise in ALP is thought to be due to the production of placental and bone isoenzymes, and breastfeeding may contribute to its prolonged postpartum dynamics of up to 18 months. CK tests also drop in the first trimester and rise again towards delivery. Postpartum CK begins normal, undershoots and adapts to baseline within weeks.

The atlas shows that calcium and phosphorus, major bone components, differ in their dynamics. Calcium drops during pregnancy, with a nadir around week 20, due to increased fetal demand. It shows discontinuity following delivery and rises to normal range within weeks. Calcium then slowly decays postpartum. In contrast to the strong changes in calcium, phosphorus dynamics are complex but generally close to the normal range, with postpartum overshoot that slowly adapts. Negative calcium balance several weeks postpartum and the prolonged increase in serum
phosphorus might be a result of adaptation mechanisms that support the calcium demand of lactation \(^\text{66}\), as is the rise in the calcium hormone PTH (see SI section S2). (Fig 8)

![Musculoskeletal test dynamics, pregnancy is in gray.](image)

**Red blood cells**

During pregnancy, red blood cell concentration (RBC), volume fraction (HCT) and hemoglobin (Hgb), tests which reflect oxygen carrying capacity, all show the same U-shape decline in the atlas that approaches the lower bound of the normal range. This is due to hemodilution that is only partially compensated by increased RBC production leading to dilutional anemia \(^\text{67}\). These tests show sharp discontinuity around delivery, possibly due to hemorrhage \(^\text{68}\), and return to prepregnancy levels within weeks. (Fig 9A) Mean cell volume and MCV and hemoglobin MCH rise during pregnancy \(^\text{69}\) and show a postpartum undershoot. Red cell size distribution width RDW rises and then adapts to normal levels with an additional peak postpartum. These effects are consistent with increased RBC production \(^\text{70}\) and iron-deficient anemia \(^\text{71,72}\). (Fig 9A) Ferritin, which reflects the body’s iron stores, shows a gradual decline, with a nadir during the third trimester. Postpartum it overshoots for about 20 weeks, parallel to an undershoot in MCHC. Blood iron shows a mild rise during the first trimester, but then declines until delivery. Postpartum, iron falls significantly, and rises over few weeks. Hgb, ferritin and iron show a sequence of nadirs during late pregnancy, potentially reflecting a hierarchy in iron metabolism. Transferrin, a protein that transports iron in the circulation, increases dramatically during pregnancy \(^\text{73}\). It shows discontinuity as it decreases and undershoots postpartum. Taken together, these test dynamics
reflect the increased demand for iron and the adaptation processes that increase its bioavailability \(^{69}\) (Fig 9B). Additional RBC related tests are shown in Fig 9C.

**Fig 9. Red blood cell test dynamics, pregnancy is in gray.**

**Immune system**

The maternal immune system changes to protect the mother and fetus from pathogens and to provide tolerance to the allogeneic fetus \(^{74}\) (Fig 10). Lymphocytes in the atlas show U-shaped dynamics during pregnancy, with a nadir around mid-pregnancy \(^{5}\) and an overshoot postpartum, in a pattern that reflects the incidence of T-cell mediated autoimmune diseases \(^{75,76}\). Neutrophils, the most abundant white blood cells, and total white blood cell counts (WBC) both rise and plateau in the third trimester \(^{74}\). They return to normal range within a few weeks postpartum, but then undershoot and take about 30 weeks to reach stable value. Monocytes show a mild rise during pregnancy \(^{74}\) and return to baseline shortly after delivery. Eosinophil and basophil counts, which are generally thought to be unchanged in pregnancy \(^{74}\) are shown in the atlas to decrease slightly during pregnancy and overshoot postpartum. Fig 10A shows the blood counts, and Fig 10B the relative counts.

The inflammatory marker CRP (C-reactive protein) increases during pregnancy \(^{74}\), peaking shortly after delivery. The atlas shows that CRP remains slightly higher than pre-pregnancy baseline for up to 80 weeks. The atlas shows that procalcitonin, a marker of bacterial infection, decreases and plateaus around the second trimester and normalizes rapidly after delivery. A previous study showed an opposite trend \(^{77}\). Large unstained cells, a marker for viral infection, show a mild decrease during pregnancy. Fig 10C.
Fig 10. Immune system test dynamics, pregnancy is in gray.

**Coagulation**

Coagulation undergoes multiple changes and shifts towards an hypercoagulable state\(^78\), **Fig 11**. In the atlas fibrinogen, a procoagulant factor, rises during pregnancy \(^78\), and undershoots postpartum. It reaches normal range within a few weeks but takes about 20 weeks to stabilize, reflecting the increased risk of hemorrhage\(^79\). During pregnancy clotting becomes more rapid,\(^78,80\) as seen in the decline in clotting tests APTT-R, APTT-sec and PT-INR, which resolves within a few weeks postpartum. Platelet count in the atlas shows a decrease due to hemodilution\(^81\), with a slight overshoot postpartum. We also observe changes in mean platelet distribution width (PDW), which increases during pregnancy, indicating changes in production and removal that counter hemodilution.

![Coagulation Test Dynamics](image)

Fig 11. Coagulation test dynamics, pregnancy is in gray.

**Metabolism**

Metabolism changes to ensure nutrient supply for the fetus. These changes are regulated by human placental lactogen (hPL) secreted by the placenta \(^67\). The atlas shows that glucose drops in the first trimester and then rises, mirroring the decline and rise of insulin resistance \(^82,83\). Postpartum, the hPL-dependent insulin resistance that developed during pregnancy is relieved rapidly, but compensation by insulin-secreting beta cells is still strong and lasts for weeks \(^82-84\), explaining the drop in glucose seen in the atlas in the first weeks postpartum.
In early pregnancy, the maternal body begins to store fat, and serum lipids increase, as seen in the atlas by a rise in cholesterol and triglycerides that slows down towards delivery. Postpartum, these tests show recovery over months. The prolonged recovery of these tests is possibly due to the mother’s adipose tissue stores that decrease over weeks, and breastfeeding that affects lipids metabolism. (Fig 12)

Fig 12. Metabolism test dynamics, pregnancy is in gray.

Discussion
We present an atlas of lab tests as a function of time in preconception, pregnancy and postpartum and analyze its global trends and specific mechanisms. The atlas is unprecedented in terms of number of participants and temporal resolution, and covers all major laboratory tests. The atlas reveals slow postpartum recovery times of up to a year for about a third of the tests. The recovery trajectories can show discontinuity upon delivery with postpartum rebound effects. We demonstrate how this data can help to test physiological mechanisms using the thyroid axis as an example - showing how changes in gland size can explain elaborate variations in thyroid axis hormones. The accuracy of the atlas allows detection of novel dynamical changes, including the impact of preconception supplements. This study thus provides a resource for understanding pregnancy and the postpartum period, and demonstrates how it may be used to understand mechanisms in systems physiology.

The 78 lab tests show two stereotypical profiles of change - either continuous, where delivery redirects the direction of change back to baseline, or discontinuous, where delivery causes a sharp reversal and an overshoot-like rebound effect. Rebounds and discontinuities have not been systematically characterized previously because studying them requires high temporal resolution which was lacking in most previous studies.

These dynamics can be rationalized based on general physiological principles. Systems that show continuous change can be described as a first order process in which the load of pregnancy...
pushes physiological parameters away from baseline. When this load is relieved, the system adapts with a typical timescale. Rebound dynamics, in contrast, can not be explained by such a first order process. Instead, they are consistent with a compensatory mechanism that grows during pregnancy, and remains high after delivery causing overcompensation and overshoots.

An example of such compensation occurs in the thyroid axis, where thyroid functional mass grows during pregnancy under control of TSH and hCG, increasing the capacity to produce thyroid hormones. This extra mass takes many weeks to recover postpartum given the slow turnover of thyroid cells, causing overshoot dynamics in thyroid hormones. The ability of endocrine glands to change mass has important beneficial functions. They have been shown to provide powerful dynamic compensation to endocrine systems, to contribute to hormone seasonality, to explain subclinical endocrine pathological stages, and to explain extended dysregulation after chronic stress is relieved.

The present atlas greatly expands our knowledge about the postpartum period, since most postpartum studies considered only one or a few timepoints. Upon delivery, the fetus and placenta exit the body, ceasing their profound metabolic and endocrine effects. Lactation begins in some mothers, along with behavioral changes. We accordingly find that the return of the tests to baseline postpartum occurs by a trajectory that differs from the trajectory of change during pregnancy - a phenomenon called hysteresis. Postpartum adaptation is a distinct physiological process and not merely the reverse of pregnancy dynamics.

The first 10 weeks postpartum show large restoration in most tests, and about two thirds of the tests return to baseline, with some returning very quickly such as coagulation tests. However, a third of the tests take on the order of 4 months to a year to adapt to their post pregnancy baseline. Examples of such slow-adapting tests are alkaline phosphatase, albumin, AST and ALT as well as sodium and uric acid. These findings reveal two sensitive periods postpartum, where deviation from baseline is evident in most tests in the first months and persists in a fraction of the tests up to a year after delivery.

Interestingly, some of the changes seen in pathologies of pregnancy are exaggerated forms of the milder changes found here that lie mostly in the normal range of each test. For example, TSH and T4 changes postpartum are mild versions of postpartum hyperthyroidism and hypothyroidism. Mild dilutional anemia is seen in blood counts and iron related tests. Insulin resistance in the second and third trimesters counters the first trimester drop in glucose and resets it, echoing gestational diabetes in which insulin resistance causes glucose to exceed its normal range in the last trimester. Future studies can use similar high resolution approaches to study pathologies, by focusing on relevant subpopulations; the present atlas can serve as a baseline control.

This study has limitations associated with use of medical datasets, including the effects of any ascertainment bias which bypassed the controls used here. This study considered first pregnancies in a single country; future work can consider the difference between first and later pregnancies and the effects of different locales.
In conclusion, this study shows how large medical datasets can be used to produce detailed temporal trajectories of pregnancy and postpartum physiology. These trajectories reveal compensation processes that allow the mother’s physiology to adapt to the multi-systemic load of pregnancy and to navigate the abrupt effects of delivery and extended effects of postpartum lactation and other changes. A similar data-driven approach might be useful for understanding other temporal transitions such as growth and development in childhood and puberty, menopause and the course of specific diseases and their recovery processes. Such accurate data can serve to develop and calibrate mechanistic models of physiological adaptation. We hope that the present atlas will lead to a better understanding of pregnancy and postpartum biology, and inspire similar studies of other crucial physiological processes that unfold over time.

Data Availability
Anonymized data and the source code used to perform the analysis is available at the GitHub repository: https://github.com/alonbar110/Atlas-of-the-physiology-of-pregnancy-based-on-millions-of-lab-tests/

Acknowledgments
We thank members of our labs for fruitful discussions. We thank Gabi Barabash and Ran Balicer for the Clalit–Weizmann collaboration. Data acquisition was approved by the Clalit Helsinki Committee RMC-1059-20. Figure 1A and figures 4A, B were made with BioRender (BioRender.com). This work was supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant Agreement No 856487).

Declaration of interests
None declared by authors.
Methods

Clalit lab tests at weekly resolution over pregnancy and postpartum
We analyzed laboratory test data from the Clalit medical database. Clalit is Israel’s largest health service organization which included about 4.6 million active members in Clalit health services as of 2020. We considered the first pregnancy of 588964 women aged 25 to 35 from 2003 to 2020, and extracted the 112 most common lab tests starting from 60 weeks before delivery to 80 weeks after delivery. To exclude second pregnancies, we removed tests that were obtained from women that had another delivery up to 40 weeks following the measurement in the postpartum period. Data were anonymized by hashing of personal identifiers and addresses and randomization of dates by a random number of weeks uniformly sampled between 0 and 13 weeks for each patient and adding it to all dates in the patient diagnoses, laboratory, and medication records. This approach maintains differential data analysis per patient, and thus the timing relative to delivery is accurate.

For each test, we then transformed the raw measurement into quantiles per age. We then binned the tests into bins of 1 week. The top and bottom 5% were removed from each bin to remove outliers. For each bin, we calculated various statistics - Number of tests in bin, Mean, median, standard variation, 25th percentile, 75th percentile of raw test value, Mean, median, 25th percentile, 75th percentile of quantile scores, Mean, 25th percentile, 75th percentile of binned patient’s age and Mean, 25th percentile, 75th percentile of binned patient's most proximal BMI measurement.

Reference values
To establish reference ranges for the raw lab tests analyzed in our study, we used a cohort of age-matched female individuals from the same Clalit healthcare. Specifically, we used the LabNorm R package for analyzing laboratory measurements. LabNorm provides functions for normalizing standard laboratory measurements according to age and sex as described in Cohen et al and are based on a cohort from the same Clalit database. This reference cohort is of women who were not pregnant and did not have any known medical conditions that could impact the lab test results. This cohort is matched to the study population in terms of age and demographic factors. However, because the cohort is based on non-pregnant women, it differs by some degree from the pregnant population analyzed in this paper due to differences in birth rates in the Israeli population.

Selection of lab tests for analysis
Three tests out of the 112 that were duplicates of other tests. These are lab tests that appeared in the database under two different IDs. These tests are for blood potassium, blood sodium, and red blood cell distribution width. We omitted the duplicates that had fewer measurements, resulting in a total of 109 tests. We then considered only tests that had reference ranges given by LabNorm, resulting in a total of 89 tests. Endocrine tests were mostly excluded at this stage. Next, we developed an ascertainment bias score (See below) to exclude tests that may have been influenced by other factors. We excluded tests that had a bias score lower than 3. After applying
these filters, we remained with 78 tests that are included in the final analysis. Excluded tests are available in the SI (Section S2). The list of tests included in the final analysis can be found in Table S1 (SI Section S1).

**Ascertainment bias score**

The bias score is composed of a sum of five factors, F1 to F5. We reasoned that very common tests, such as blood counts, are less likely to be biased. We thus scored the median test number across bins as F1=0 for n<250, 0.5 for 250<n<5000, and 1 for n>5000. We reasoned that tests that are taken with a temporal profile similar to blood counts are less likely to be biased. We therefore clustered the temporal profiles of the number of tests per bin, using hierarchical clustering over the pairwise distances matrix of the Z-scored log of number of tests, with a criteria of 45% of the max distance in the matrix. We scored tests that fall in the same cluster as blood counts with F2=1 and F2=0 otherwise. We reasoned that biased tests would be more likely in participants with outlier age or BMI. Thus factor 3 scores F3=0 on tests whose age per bin profiles are outliers and F3=1 otherwise, and factor 4 does the same for BMI. Outliers were defined using PCA 1 and 2. Finally, for the 21 tests for which ‘ground truth’ values at several times-points were available from a previous study, we scored low on tests which did not correlate - r<0.5 F5=0, and F5=1 otherwise. The total bias score is the sum of F1 to F5.

**Data processing**

To provide a robust measure for the mean test value at each time point, we consider the 'value at mean quantile' rather than the mean test value. To calculate the 'value at mean quantile' for each lab test, we utilized LabNorm, which allows for the transformation of quantile scores to absolute test values. We developed a transfer function for each lab test based on its distribution in the reference population. Specifically, for each test, we obtained the test values at the 1st, 3rd, 10th, 15th, 25th, 35th, 50th, 65th, 75th, 85th, 90th, 97th, and 99th percentiles of the reference population. We then linearly interpolated test values for intermediate percentiles. Since the distribution of test values varies across different age groups, we used the median age across the 140 weeks as a reference age for each test. The quantile-to-value transfer function for each test is provided in the supplementary information (SI Section S3).

To ensure comparable accuracy across different lab tests, we merged consecutive bins for tests with smaller sample sizes. Consecutive bins were merged such that the error of test value per bin is at least 15% of the standard variation of test values across time. To ensure that we did not merge too many bins and lose important information, we set the maximum number of bins we allowed to merge to 7.

Thyroid data in figure 4 was processed by merging every two consecutive bins for both TSH and T4. Test values from the prepartum and postpartum periods were normalized by the corresponding baseline to represent percent change. In Fig 6-12 we present percentile change relative to preconception baseline defined by the mean of weeks -60 to -50.
Analysis of variation across pregnancy
We assessed the magnitude of variation across pregnancy by comparing the bins with the minimum and maximum values for each test. To determine if this variation was statistically significant, we performed a two-sample t-test for each test between these bins and adjusted for multiple comparisons using the false discovery rate (FDR) method with a threshold of 0.05. We assessed this variation both in quantile scores and in value at mean quantile.

To measure the variation of test values within the reference population, we calculated the coefficient of variation (CV) for each test as the ratio of the standard deviation to the mean test value. We investigated the relationship between the variation across pregnancy in test values and the variation of each test within the population using a linear regression analysis. For Fig 3A we used the Python library scikit-learn to perform PCA.

Normalized temporal profiles
In order to cluster the dynamical profiles of the tests, we sought to perform a normalization process that maintains the dynamical properties of the tests both in pregnancy and postpartum. Usually, z-score transformation of the data can achieve this - it allows for comparison of time series that ranges over different scales. We adapted this approach to our dataset by two additional steps. First, we removed time points that are 2 weeks prior to delivery, and up to 3 weeks after delivery to exclude spike-like dynamics proximal to delivery in some tests. Next, we divided each time series into two periods - preterm and postpartum and performed z-score transformation on each period independently. This process provided a similar weight for each period in the clustering analysis. To cluster the dynamic profiles of the lab tests, we used t-SNE, to reduce the dimensionality of the normalized lab test data, with the scikit-learn package in Python, with 2 components. Next, we used K-means to cluster the lab tests based on the t-SNE representation. We used the elbow method in order to choose the optimal number of clusters.

Minimal mathematical model of dynamical profiles
We consider a simple mathematical model based on the incoherent feed-forward loop (IFFL) circuit. In this model, factor $X$, a first order system, is produced at rate $u(t) b_X$ and removed at rate $a_X$ giving rise to the differential equation $\frac{dX}{dt} = u(t) b_X - a_X X$. At steady state with $u(t) = \text{const.}$ we solve $\frac{dX}{dt} = 0$ to obtain the steady state $X_{st} = u(t) b_X / a_X$. The second factor in this model, $Y$, represents a test for which there is a compensation process, namely $X$. Thus $Y$ has an additional layer of regulation where $X$ reduces $Y$ production rate $\frac{dY}{dt} = u(t) b_Y / X - a_Y Y$. Pregnancy exerts a time-varying load on production. We simulated the model over 140 weeks starting with a constant preconception production rate for 22 weeks, followed by a pregnancy load of 38 weeks of a exponential increase in $u(t)$. The 80 week postpartum period was simulated by resetting production rate to baseline. Removal rates $a_X$, $a_Y$ where chosen to be $1/20$ weeks for $X$ and $1/\text{weeks}$ $Y$. Other parameters were chosen to give a steady state of 1. During pregnancy $u(t)$ rises linearly from 1 with a slope of 0.01/week. An analytical solution of this model is provided in the SI (Section S4).
Settling time
We define the postpartum baseline using the mean test values in the last 10 weeks of the atlas, weeks 70-80 after delivery, and define SDS using the average of the standard deviations of test values in these 10 bins. We next smoothed the mean test value data using a Gaussian rolling average over 7 weeks to reduce the effect of outlier timepoints, and removed the first 3 weeks in the postpartum period. The settling time is the time after which at least 90% of the smoothed timepoints remain within 0.2SDS of the baseline values (SI section S6). We estimated the error bars of settling time by bootstrapping. Settling time for tests that settled immediately (within the first time point) was adjusted to “up to 4 weeks”. For RBC, Calcium, Transferrin and Microcytic% we used a cutoff of 0.18 SDS to avoid spurious fluctuations (SI section S6).

Mechanistic models of the thyroid axis
We use the model of Korem et al. The three hormone concentrations TRH, TSH and T4 are denoted $x_1, x_2, x_3$. The functional mass of the pituitary thyrotrophs is $P$ and of the thyroid is $T$. The classic cascade model is

$$\begin{align*}
    \frac{dx_1}{dt} &= \frac{b_1}{x_3} x_1 - a_1 x_1 \\
    \frac{dx_2}{dt} &= \frac{b_2 P x_1}{x_3} - a_2 x_2 \\
    \frac{dx_3}{dt} &= \frac{b_3 T x_2}{x_3} - a_3 x_3
\end{align*}$$

where hormone lifetimes are given by the removal rate inverses $1/a_1 = 6$ minutes, $1/a_2 = 1$ hour, $1/a_3 = 7$ days. In the classic model, $P$ and $T$ are constant functional masses. The model of Korem et al adds to this classic model equations for $P$ and $T$ as time changing functional masses under control of the axis hormones

$$\begin{align*}
    \frac{dP}{dt} &= P \left( \frac{b_p}{x_3} - a_p \right) \\
    \frac{dT}{dt} &= T (b_T x_2 - a_T)
\end{align*}$$

The steady-state hormone levels (averaged over months) in this model is robust to almost all model parameters:

$$\begin{align*}
    x_{1st} &= \frac{b_1 a_p}{a_1 b_p} x_{2st} = \frac{a_T}{b_T}, \\
    x_{3st} &= \frac{b_p}{a_p}, \\
    P_{st} &= \frac{a_1 a_2 b_p^2}{b_1 b_2 a_p^2 b_T}, \\
    T_{st} &= \left( \frac{a_3 b_p b_T}{b_3 a_p a_T} \right)^{1/3}
\end{align*}$$

Model parameters were chosen as $b_1 = a_1, b_2 = a_2, b = a_3, b_p = a_p, b_T = a_T$ to give a steady-state equal to 1. Full parameter values are listed at table 1. See SI (Section S7) for model simulation throughout pregnancy.
Equations for the model on a timescale of months can be derived by using quasi-steady-state approximation for the hormones. The quasi steady-state of the fast equations is:

\[ x_{1\text{qst}} = P^{-1/3}T^{-1/3} \]  
\[ x_{2\text{qst}} = P^{1/3}T^{-2/3} \]  
\[ x_{3\text{qst}} = P^{1/3}T^{1/3} \]

Substituting the quasi steady-state (Eq 8,9) into the slow gland mass equations (Eq 4,5) results in two coupled ODEs for the functional masses \( P \) and \( T \):

\[ \frac{dP}{dt} = a_P P(P^{-1/3}T^{-1/3} - 1) \]
\[ \frac{dT}{dt} = a_T T(P^{1/3}T^{-2/3} - 1) \]

Postpartum \((t = 0)\), we assume that gland masses are not at steady state, and that all model parameters are constant. The solution of these equations, namely \( P_{pp}(t) \) and \( T_{pp}(t) \), is a function of the initial conditions - the functional masses of the glands at delivery, and their turnover rates.

\[ P_{pp}(t) = f(P(0), T(0), a_p, a_T) \]
\[ T_{pp}(t) = g(P(0), T(0), a_p, a_T) \]

To solve the hormones postpartum dynamics we numerically solve the glands dynamics for a given set of parameters \((P(0), T(0), a_p, a_T)\). Next, the postpartum dynamics of the hormones is then given by the quasi steady-state (Eq 8,9):

\[ x_{2pp}(t) = P_{pp}(t)^{1/3}T_{pp}(t)^{-2/3} \]
\[ x_{3pp}(t) = P_{pp}(t)^{1/3}T_{pp}(t)^{1/3} \]

Using Scipy’s ‘curve_fit’ we fit TSH and FT4 data, merging every two consecutive bins, at week 5 to 80 postpartum and estimated thyroid and pituitary initial sizes and turnover times. (See SI (Section S8) for approximate analytical solution of the model)
Table 1. HPT model parameter values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>$0.17/\text{min} \ 29.30$</td>
</tr>
<tr>
<td>$b_1$</td>
<td>$0.17/\text{min} \ (\text{steady state normalization to 1})$</td>
</tr>
<tr>
<td>$a_2$</td>
<td>$1/\text{hour} \ 29.30$</td>
</tr>
<tr>
<td>$b_2$</td>
<td>$1/\text{hour} \ (\text{steady state normalization to 1})$</td>
</tr>
<tr>
<td>$a_3$</td>
<td>$1/\text{week} \ 29.30$</td>
</tr>
<tr>
<td>$b_3$</td>
<td>$1/\text{week} \ (\text{steady state normalization to 1})$</td>
</tr>
<tr>
<td>$a_P$</td>
<td>$1/30 \text{ days} \ (\text{model fit})$</td>
</tr>
<tr>
<td>$b_P$</td>
<td>$1/30 \text{ days} \ (\text{steady state normalization to 1})$</td>
</tr>
<tr>
<td>$a_T$</td>
<td>$1/85 \text{ days} \ (\text{model fit})$</td>
</tr>
<tr>
<td>$b_T$</td>
<td>$1/85 \text{ days} \ (\text{steady state normalization to 1})$</td>
</tr>
</tbody>
</table>

Preconception dynamics

To study the dynamic trends of lab test values before conception, we used the mean quantile test values. We used linear regression to model the relationship between the test mean quantile score and time (in weeks) in the preconception period (60 to 38 weeks before delivery). We considered a test to have a significant dynamical trend if the absolute value of the regression line slope was greater than 0.1 and the p-value was less than 0.05 after controlling for false discovery rate (FDR).

Bibliography


regulation of thyroid cell proliferation and differentiation by thyrotropin and other factors. Physiological Reviews 72, 667–697. 10.1152/physrev.1992.72.3.667.


73. Lebel, Y., Milo, T., Bar, A., Mayo, A., and Alon, U. (2023). Excitable dynamics of flares and relapses in autoimmune diseases. 2023.05.03.539265. 10.1101/2023.05.03.539265.


Trimester-specific coagulation and anticoagulation reference intervals for healthy pregnancy. Thrombosis Research 156, 82–86. 10.1016/j.thromres.2017.05.021.