Bryn Mawr College
Scholarship, Research, and Creative Work at Bryn Mawr College

Mathematics Faculty Research and Scholarship | Mathematics

2023

Reduced model for female endocrine dynamics: Validation and functional variations

Erica Graham
Bryn Mawr College, ejgraham@brynmawr.edu

Noémie Elhadad
Columbia University

David Albers
University of Colorado Denver

Follow this and additional works at: https://repository.brynmawr.edu/math_pubs

 Part of the Mathematics Commons

Let us know how access to this document benefits you.

Citation

Graham, E. J., Elhadad, N. & D. Albers. 2023. "Reduced model for female endocrine dynamics: Validation and functional variations." *Mathematical Biosciences* 358: 108979.

This paper is posted at Scholarship, Research, and Creative Work at Bryn Mawr College.
https://repository.brynmawr.edu/math_pubs/31

For more information, please contact repository@brynmawr.edu.

Reduced model for female endocrine dynamics: Validation and functional variations

Erica J. Graham[a,*], Noémie Elhadad[b], David Albers[c]

[a]*Mathematics Department, Bryn Mawr College, Bryn Mawr, PA 19010, USA*
[b]*Department of Biomedical Informatics, Columbia University, New York, NY 10032, USA*
[c]*Pediatrics Department, University of Colorado Denver–Anschutz Medical Campus, Aurora, CO 80045, USA*

Abstract

A normally functioning menstrual cycle requires significant crosstalk between hormones originating in ovarian and brain tissues. Reproductive hormone dysregulation may cause abnormal function and sometimes infertility. The inherent complexity in this endocrine system is a challenge to identifying mechanisms of cycle disruption, particularly given the large number of unknown parameters in existing mathematical models. We develop a new endocrine model to limit model complexity and use simulated distributions of unknown parameters for model analysis. By employing a comprehensive model evaluation, we identify a collection of mechanisms that differentiate normal and abnormal phenotypes. We also discover an intermediate phenotype–displaying relatively normal hormone levels and cycle dynamics–that is grouped statistically with the irregular phenotype. Results provide insight into how clinical symptoms associated with ovulatory disruption may not be detected through hormone measurements alone.

Keywords: Ovulation, endocrinology, polycystic ovary syndrome

1. Introduction

Female endocrine physiology is an incompletely understood system, particularly as it pertains to reproductive health and disease. Metabolic and mental health problems are also associated with dysfunctional female reproductive endocrinology [1, 2]. Two relatively common disorders, which independently affect between 5 and 20% of individuals of reproductive age, are polycystic ovary syndrome (PCOS) [2] and endometriosis (EM) [3]. PCOS is often characterized by infrequent or absent ovulation and excess ovarian production of *androgens* (male sex hormones), especially testosterone [4]. EM is characterized by lesions of endometrial-like tissue throughout the body, significant pain, and infertility [3]. The systems physiologic understanding of PCOS and EM is both incomplete and unavoidably complex. Given the prevalence of these disorders, comprehensive reproductive hormone data is relatively limited. Aside from their relatively high prevalence, both conditions have unknown etiologies, are often associated with infertility, and can lead to additional complications and a significant reduction in quality of life. As a result, PCOS and EM are good targets for mathematical physiology, which has the ability to link endocrine function/dysfunction with its physiological underpinnings in a quantitative manner.

Although the importance of these disorders and their clinical impact cannot be overstated, of particular importance in quantitative approaches is an understanding of baseline ovulatory function, which depends on the interactions of several reproductive hormones and their targets tissues. Collectively, this system is known as the hypothalamic-pituitary-ovarian (HPO) axis. Briefly we describe the essential processes involved in developing a basic quantitative description of ovulation and HPO function, for which there are

*Corresponding author. Address: Mathematics Department, 101 N. Merion Ave., Bryn Mawr, PA 19010, USA.
Email address: ejgraham@brynmawr.edu (Erica J. Graham)

three primary components. First is the signal generated in the hypothalamus which, through gonadotropin-releasing hormone (GnRH) neurons, triggers the production of follicle-stimulating hormone (FSH) and luteinizing hormone (LH) from the pituitary gland. Second, there are the FSH- and LH-dependent processes of (1) follicle growth and maturation, (2) ovulatory follicle selection, and (3) ovum release. Third, there is the feedback from ovarian steroid hormones–e.g., estrogens, progesterones, and inhibins–which regulate production and release of FSH and LH from the brain.

Relative to typical ovulatory function, in many cases of PCOS the HPO axis is altered by elevated androgen, e.g. testosterone, production resulting from insulin resistance, increased LH synthesis due to disrupted GnRH pulse frequencies, and an increased LH:FSH ratio [4]. The consequence of these disruptions is reduced or complete absence of ovulation–which can be observed by very low P_4 levels. In EM, the story is less clear, as pituitary and ovarian hormone levels tend to be relatively similar between individuals with and without EM [1]. In fact, alleviation of severe pain in EM can be achieved through intentional disruption of ovulation, which prevents estrogen production and the resulting pain-causing inflammation [5]. Reducing other sources of circulating estradiol, e.g. from fat, skin, or endometrial tissue, could also be therapeutic in EM [4]. While PCOS can result from dysregulation of reproductive hormones, EM instead reflects pathological responses to those hormones.

A number of mathematical models of ovulation have been developed over the past few decades, each focusing on one or more aspects of the HPO axis. For example, GnRH neuron pulse generation has been modeled in [6, 7], whereas the downstream effect on FSH and LH is modeled in [8]. Follicle dynamics have been described mathematically in [9, 10, 11, 12], and ovarian steroid biosynthesis in [13, 14], with a semi-mechanistic approach taken in [15]. Some models focus on a two-part pituitary-ovarian axis [16, 17, 18, 19], whereas others describe the multi-scale behavior of the HPO-axis, including GnRH pulse regulation [20, 21, 22, 23, 24]. Of these models, [19, 25, 15] incorporate testosterone and are capable of producing PCOS-like dynamics in the form of ovulatory dysfunction and anovulation. Although a recent model in [26] describes a phenomenological-based model of endometrial dynamics, there appears to be no mathematical models that address endometriosis in the context of the ovulatory cycle.

Whereas existing mathematical models may highlight one specific disease or another, we aim initially to reduce the viable parameter space of fundamental models using a 'bottom-up' approach: physiology first, pathophysiology second. In particular, we aim to improve tractability while still maintaining a reasonable quantitative representation of the underlying physiology.

In this paper, and toward building a unified, physiologically-based framework to shed light on diseases like PCOS and EM, we develop a new endocrine model with reduced complexity, both in biological description and in the number of unknown parameters. We hypothesize that structural reduction of the model may provide greater insight to relevant and essential processes governing the typical ovulatory cycle and aid in identifiability. Although PCOS and EM are important disorders with many open questions regarding their etiologies, we presently aim to study *generalized* ovulatory dysfunction, which may or may not stem from existing clinical phenotypes, including hyperandrogenic PCOS and EM. The goal is to fit the new model to available data (described in Section 2.2) and follow up with a comprehensive model evaluation that provides insight into relevant parameter-informed mechanisms that can distinguish ovulatory phenotypes. The approach taken in this work is useful because it allows us to examine dysfunction based on limited yet easily identifiable clinical information, such as an individual's 'time since last period.' A typical approach to mathematically study female endocrine physiology often relies on analyzing hormone levels and long-term behavior in response to some stimulus, which is mathematically informative but often not clinically feasible due to cost and invasiveness of the data collection process. We aim to push the boundaries of what we can study in female reproductive endocrinology based on intentionally minimal information.

In Section 2, we begin with an existing mathematical model of ovulation [15] that has been fit to data from the literature [27], and perform a semi-mechanistic model reduction. In Section 3, we construct a systematic methodology for model evaluation and use it to examine the new model's ability to capture essential biology and represent data. We also examine emergent phenotypes [28] through analysis of the parameter space. In Section 4, we present simulation and evaluation results, followed by a discussion of their implications in Section 5.

2. Methods

We develop a new endocrine model to describe essential processes in ovulation. Following [8, 18, 19, 15], we use a compartmental framework to model the ovulatory system. This new model is a reduction of a model developed by Graham and Selgrade that uses ordinary differential equations to describe the ovulatory cycle under androgen, i.e. testosterone (T), influence [15]. In addition to explicitly incorporating T-mediated feedback, the model explores the effects of normal and premature luteinization on a T-dependent ovulatory cycle and provides mechanistic insight into the different PCOS phenotypes that might emerge in a high-androgen high-insulin state. Effectively, we show that we can exclude T–which is absent from many existing models that accurately describe the ovulatory cycle–from the model and still capture important features of physiological, but not necessarily pathological, hormone dynamics. Further, we choose to begin with the Graham-Selgrade model because unlike its predecessors, it (1) successfully captures the dynamics of clinical data without delay differential equations, and (2) describes ovarian follicle dynamics using only three distinct stages (described in Section 2.1), compared to the nine [18] or 12 stages used previously [19]. Thus, our starting point for model reduction is an already simplified framework. For comparison, the original model equations are listed in Appendix A.

2.1. Model Development

Here we describe our reduced model, which comprises three major subsystems and describes changes in the pituitary-ovarian axis without an explicit role for androgens. For each subsystem, we derive the model equations and highlight major modifications made to reduce the number of unknown model parameters used in [15].

I. Pituitary regulation.

LH and FSH are the primary hormones produced by the pituitary gland, and we assume their synthesis and release are regulated by the ovarian steroid hormones E_2 and P_4 (see Figure 1(a)). In subsystem (1)–(4), we track FSH and LH, which are split between releasable (denoted FSH_ρ and LH_ρ) and serum (denoted FSH and LH) compartments. The equations governing FSH_ρ and LH_ρ represent the balance between synthesis and release into serum, subject to both stimulatory and inhibitory feedback via estradiol (E_2) and progesterone (P_4). Serum hormones are assumed to undergo first-order decay. Equations (1)–(4) are almost identical to the ones presented in [15], except we eliminate the effects of testosterone in the basal rate of LH synthesis (v_{0L}) and in LH inhibition via P_4 ($K_{iL,P}$).

$$\text{Releasable FSH:} \quad \frac{dFSH_\rho}{dt} = \frac{v_F}{1 + c_{F,I}\frac{S\Lambda}{K_{iF,I}+S\Lambda}} - k_F \frac{1 + c_{F,P}P_4}{1 + c_{F,E}E_2^2} FSH_\rho \tag{1}$$

$$\text{Serum FSH:} \quad \frac{dFSH}{dt} = \frac{1}{V} \cdot k_F \frac{1 + c_{F,P}P_4}{1 + c_{F,E}E_2^2} FSH_\rho - \delta_F FSH \tag{2}$$

$$\text{Releasable LH:} \quad \frac{dLH_\rho}{dt} = \left[v_{0L} + \frac{v_{1L}E_2^n}{K_{mL}^n + E_2^n} \right] \cdot \frac{1}{1 + P_4/K_{iL,P}} - k_L \frac{1 + c_{L,P}P_4}{1 + c_{L,E}E_2} LH_\rho \tag{3}$$

$$\text{Serum LH:} \quad \frac{dLH}{dt} = \frac{1}{V} \cdot k_L \frac{1 + c_{L,P}P_4}{1 + c_{L,E}E_2} LH_\rho - \delta_L LH \tag{4}$$

II. Follicle dynamics.

Follicle growth, maturation, and differentiation are assumed to occur in a series of three sequential stages: (1) *follicular* (Φ), (2) *ovulatory* (Ω), and (3) *luteal* (Λ), as illustrated in Figure 1(b). The follicular phase is characterized by recruitment and growth of stimulated follicles. The ovulatory phase is characterized by ovum release from a designated follicle in response to a mid-cycle surge in LH. Finally, the luteal phase is

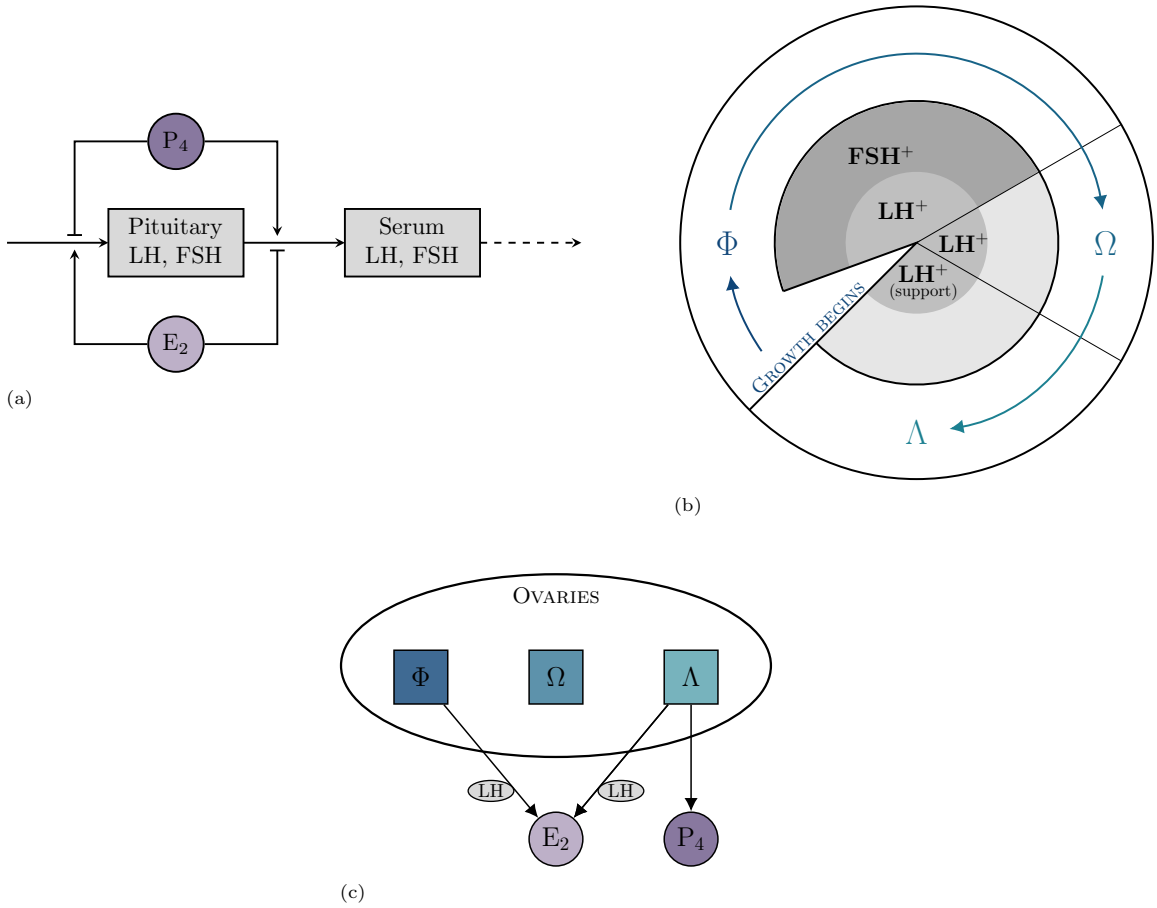

Figure 1. Model schematic for three subsystems: (a) pituitary regulation, (b) ovarian follicle dynamics, and (c) ovarian steroid production.

characterized by the formation and, in the absence of fertilization, regression of the corpus luteum. An LH support variable, S, reflects the dependence of the corpus luteum growth and regression on LH.

$$\text{Follicular phase:} \quad \frac{d\Phi}{dt} = \left(\frac{f_1 FSH^2}{h_1^2 + FSH^2} - \frac{f_2 LH^2}{h_2^2 + LH^2} \right) \cdot \Phi \tag{5}$$

$$\text{Ovulatory phase:} \quad \frac{d\Omega}{dt} = \frac{f_2 LH^2}{h_2^2 + LH^2} \cdot \Phi - wS\Omega \tag{6}$$

$$\text{Luteal phase:} \quad \frac{d\Lambda}{dt} = wS\Omega - l(1-S)\Lambda \tag{7}$$

$$\text{LH support:} \quad \frac{dS}{dt} = \hat{s}\frac{LH^4}{LH^4 + h_s^4}(1-S) - \delta_S S \tag{8}$$

Compared to the model in [15], we simplify two follicular processes in Equations (5)–(8). First, we eliminate the effects of testosterone on follicle sensitivity to FSH (h_1). Second, we simplify LH sensitivity by omitting FSH-dependent upregulation of LH receptors (h_2).

III. Ovarian steroidogenesis.

Throughout the ovulatory cycle, follicles may produce E_2 and P_4. Intracellular steroid production is primarily FSH- and LH-dependent during a typical cycle and is subject to functional maturation of individual

follicles, as illustrated in Figure 1(c). This subsystem exploits the *two-cell two-gonadotropin* theory of ovarian steroid production, which describes the differential functionality of theca cells and granulosa cells within ovarian follicles [4]. In Equations (9) and (10) we assume that intracellular E_2 is immediately converted from its testosterone precursor without explicit dependence on FSH. However, we do retain the LH dependence required for precursor synthesis. Once released into the serum during the follicular and luteal stages only, E_2 dynamics are subject to peripheral production and first-order decay. P_4 conversion within theca and granulosa cells requires enzymes that are also regulated by LH. However, given the very low concentration of LH required for this process, we assume P_4 production occurs with rate constant p during the luteal stage only.

$$\text{Serum } E_2: \quad \frac{dE_2}{dt} = e_0 - \delta_E E_2 + t_{g1}\frac{LH}{LH+\kappa_2}\cdot(\Phi + \eta\Lambda S) \tag{9}$$

$$\text{Serum } P_4: \quad \frac{dP_4}{dt} = -\delta_P P_4 + p\Lambda S \tag{10}$$

Compared to the Graham-Selgrade model, we have simplified the semi-mechanistic steroid production terms by eliminating both testosterone state variables (T and T_γ) altogether, assuming relatively constant dynamics, and by removing the role of basal LH in P_4 production. The impacted parameters are t_{g1} and p.

The new model is given by Equations (1)–(10), with 10 state variables, compared to 12 previously. With 27 unknown parameters, compared to the original model's 41, we have reduced the parameter space by more than a third.

2.2. Data

Complete data sets that track the pituitary hormones, LH and FSH, as well as ovarian steroids E_2 and P_4 during the course of an entire cycle are uncommon; for example, see [27] and [29]. What is missing is a complete hormone profile that also includes androgen levels through the course of a normal ovulatory cycle. Arguably, androgens may not have a substantial impact on regular ovulatory function. Still, in [15], averaged T data from [30] are used to inform T dynamics.

To revert back to a lack of androgen data in the present work, we use two data sets, one synthetic and one real. The first data set is generated synthetically using the numerical solution of the Graham-Selgrade model. We use this to show that the reduced model can capture the qualitative dynamics of the original model. The second data set is the hormone data available in [27]. These data contain average daily measurements for 33 normally cycling women during the course of one complete ovulatory cycle for FSH, LH, E_2, and P_4. This second data set is used to demonstrate both the ability of the model to estimate data well and how to use the model to better understand physiology for given data.

3. Comprehensive Model Evaluation

Given the complex cross-talk in the reproductive endocrine system, analysis of hormone concentrations alone likely provides insufficient insight into the subtleties of ovulatory dysfunction. To address this issue, we develop and implement a five-step algorithm (summarized in Algorithm 1) that allows us to carry out a comprehensive evaluation of the reduced model, with a focus on primarily *clinically relevant phenotypes* and secondarily mathematically relevant phenotypes. Then we discuss the simulations and statistical methods used to analyze the results of Algorithm 1.

3.1. Terminology: physiological vs. mathematical cycles

To discuss model evaluation and results, we explicitly distinguish between physiological and mathematical notions of a 'cycle'. For properties of mathematical ovulation, we explicitly refer to the inter-ovulatory interval (IOI), which denotes the length of time between consecutive simulated LH surges. Physiologically, the IOI is equivalent to the time between two *ovulatory cycles*; however, multiple IOIs may be required before the solution completes a single mathematical (limit) cycle. For clarity and consistency, we restrict our generalized use of 'cycle' to refer to physiological ovulation and IOI to the calculated times between these cycles.

3.2. Algorithm for comprehensive model evaluation

Step 1: Synthetic data. We generate the synthetic data by numerically solving the system (A.11)–(A.22), using the parameters in [15], for a sufficiently long time to approach a stable limit cycle for normal ovulation. We then align the trajectories so that the LH surge occurs at the end of day 15 of the first cycle. Finally, we extract daily data between days 0 and 30 and then, to avoid propagated numerical inaccuracies, repeat these same data twice more for each variable. We also expand the set of data to include Φ, Ω, and Λ, under the assumption that follicular dynamics should follow a similar pattern to the original model. Note that we do not make this assumption for the releasable pools of pituitary hormones (FSH_ρ and LH_ρ). With the exclusion of testosterone from the model, we have a total of 7 state variables (namely FSH, LH, E_2, P_4, Φ, Ω, and Λ) with $n = 93$ data points each.

Step 2: Optimization. To determine how well the new model compares to the original model, we first estimate the 27 parameters of the reduced model by fitting output to the synthetic data. Then we verify the model behavior under the influence of testosterone-mediated feedback. Because this latter step is model-specific, we postpone a detailed discussion of setup and implementation to Section 4.

To capture essential ovulatory behavior, we optimize parameters using a weighted least squares approach, described as follows. For a given variable $X_i(t)$, where $i \in \{FSH, LH, E_2, P_4, \Phi, \Omega, \Lambda\}$, we first assign default weights $w_i = 1/\mathrm{Var}(X_i(t))$ to each data point at $t_j = 0, 1, \ldots, 92$. Because we cannot guarantee the expected behavior of follicular dynamics, we do not incorporate additional time-dependent weights for $i = \Phi, \Omega, \Lambda$. However, for the hormones we increase weights by variable factors at important peaks, troughs, and plateaus within the data. These weights are adjusted to acquire the best qualitative fit to the data, with the understanding that local minimization of the cost function may be sensitive to variation in weights and may not produce a globally optimal solution.

Let $\mathbf{y}_i$ represent the vector of measurements corresponding to reduced model output variable $\mathbf{x}_i(\boldsymbol{\varphi})$, defined by parameters $\boldsymbol{\varphi}$. We define the optimization problem that minimizes the sum of the squared error as

$$\min_{\boldsymbol{\varphi}} \frac{1}{|V| \cdot n} \sum_i w_i ||\mathbf{y}_i - \mathbf{x}_i(\boldsymbol{\varphi})||^2, \tag{11}$$

where $V = \{FSH, LH, E_2, P_4, \Phi, \Omega, \Lambda\}$ and denote the optimal parameter vector satisfying Equation (11) by $\boldsymbol{\varphi}^*$. We use MATLAB's `fminsearch`, which implements the Nelder-Mead simplex method, to determine the optimal $\boldsymbol{\varphi}$. In most cases, initial parameter guesses are taken from the original model. In others, they are derived from the adjustments made in the reduction process.

Step 3: Monte Carlo simulations. We determine the distributions of the 27 model parameters using a Monte Carlo approach to generate a collection of best-fit parameters using various initial guesses in the estimation scheme described in Step 2 and by comparing the output to data. We first assume that the values in $\boldsymbol{\varphi}^*$ represent mean quantities and that initial guesses, $\boldsymbol{\varphi}^{(0)}$, are uniformly distributed within

Algorithm 1 Comprehensive model evaluation.

Step 1. Generate synthetic data set using Equations (A.11)–(A.22).

Step 2. Optimize reduced model parameters using weighted least squares and synthetic data.

Step 3. Run N Monte Carlo simulations, initialized with perturbed best-fit parameters from Step 2 and refit to clinical data.

Step 4. Compute numerical solutions for each parameter profile generated in Step 3, and store resulting hormone data over multiple cycles.

Step 5. Use results of Steps 3 and 4 to define distributions for each of the 27 reduced model parameters.

$\pm 10\%$ of the mean. That is, $\varphi_k^{(0)} \sim U(0.9\varphi_k^*, 1.1\varphi_k^*)$ for $k = 1, 2, \ldots, 27$. To ensure a representative sampling of $N = 500$ parameter combinations from each individual subinterval of length $1/N$ ranging from $0.9\varphi_k^*$ to $1.1\varphi_k^*$, we use Latin hypercube sampling (LHS)–a type of Monte Carlo sampling–to randomly generate initial parameter guesses (see [31] for further discussion on LHS). We assume a uniform distribution when generating these parameter values because we have insufficient biological information to infer a more confined distribution and because each parameter spans few, if any, orders of magnitude, as discussed in [32]. For each initial parameterization we then minimize a cost function similar to Equation (11), fit to average daily clinical data $\mathbf{y}_i,\ i \in \{FSH, LH, E_2, P_4\}$, as reported in [27].

Step 4: Range of simulated model output. We numerically solve the reduced model over 186 days using the estimated parameters and generate an ensemble of these solutions. We align each LH surge (assuming one exists) to day 15 and determine the length of each IOI. The LH surge is defined to be a peak LH concentration that is followed by an apparent luteal phase; any other local maxima in LH failing to meet this criterion are ignored. Because we have restricted our sampling scheme in Step 3, we guarantee that the model does not approach a stable equilibrium. Although this limitation does not capture complete ovarian failure (i.e., the absence of a cycle at all), it does allow for reasonable comparisons in the presence of oscillatory dynamics.

Step 5: Parameter distributions.

We calculate empirical parameter and cumulative distributions based on the optimized parameter sets obtained from Step 3.

4. Computational Results

In this section, we (1) describe the simulations carried out for the new endocrine model, and (2) examine additional emergent features arising from the model and its analysis. In particular, we are interested how clinical phenotypes present themselves in the new model, and how these results uncover important mechanisms associated with ovulatory function and dysfunction.

4.1. Simulations

Following Step 1 of the algorithm, we numerically solve the original model to obtain synthetic data, which include four hormones and three ovarian stages. In Step 2, we use the synthetic data to estimate the 27 parameters of the reduced model (Section 4.2). With the set of best-fit reduced model parameters, we compare the model behavior under the influence of testosterone-via-insulin to that of the original model; in effect, we redefine relevant parameters $(t_{g1}, v_{0L}, K_{iL,P}, h_1)$ in response to insulin influence parameter α (Section 4.2). In Step 3, we use a collection of perturbed parameter regimes ($N = 500$) to initialize re-fitting of the reduced model, this time to clinical data in [27]. In Step 4, we use the resulting sets of optimized parameters to generate numerical solutions, and compile model output for the hormones FSH, LH, E_2, and P_4 (Section 4.3). Finally, in Step 5, we compute the empirical frequency and cumulative distributions for each parameter (Section 4.4).

4.2. Model calibration

The best-fit parameters of the current model compared to original data from the model in [15] are presented in Table 1. In Figure 2 we numerically solve the reduced model using these parameters and compare the results to output from the original model. For frame of reference, we also include the clinical data from [27]. Notably, for all numerical solutions that display hormone trajectories, we solve the model for a sufficiently long time to overcome initial transient behavior in an effort to only capture the solution once it has approached a limit cycle. It is only then that we determine the location of the LH surge, which is always assumed to occur on day 15 of the displayed ovulatory cycles. All subsequent LH surges occur relative to this initial surge.

Table 1. Estimated pituitary and ovarian parameters generated from fitting the reduced model to original model output. Other fixed parameters appearing in the model remain unchanged from [15].

Pituitary Parameters

Parameter	**Units**	**Value**
v_F	μg/d	3219.9
$K_{iF,I}$	μg	149.76
k_F	d^{-1}	3.0212
$c_{F,P}$	$(\mu\mathrm{g/L})^{-1}$	65.229
$c_{F,E}$	$(\mathrm{ng/L})^{-2}$	0.0024047
$c_{F,I}$	#	3.0188
v_{0L}	μg/d	308.35
v_{1L}	μg/d	44700
Km_L	μg/L	226.37
$K_{iL,P}$	μg/L	3.2279
k_L	d^{-1}	0.67146
$c_{L,P}$	$(\mu\mathrm{g/L})^{-1}$	0.015844
$c_{L,E}$	$(\mu\mathrm{g/L})^{-1}$	0.00068867

Ovarian Parameters

Parameter	**Units**	**Value**
f_1	d^{-1}	1.0958
f_2	d^{-1}	46.225
h_1	μg/L	146.31
h_2	μg/L	798.39
w	d^{-1}	0.23497
l	d^{-1}	0.64178
$\hat{s}$	d^{-1}	2.6338
δ_S	d^{-1}	0.38256
η	#	0.81426
κ_2	μg/L	8.276
h_s	μg/L	11.691
t_{g1}	ng/(L·μg·d)	6.3594
e_0	ng/L·d	9.6377
p	1/L·d	0.22851

Fixed Parameters

Parameter	**Units**	**Value**
δ_F	d^{-1}	8.21
V	L	2.5
δ_L	d^{-1}	14
δ_E	d^{-1}	1.1
δ_P	d^{-1}	0.5

The qualitative dynamics are well captured, with two primary quantitative discrepancies. First, early follicular phase E_2 (days 0 through 10 of the ovulatory cycle) undershoots both synthetic and clinical data. However, the biological impact of E_2 during this stage of the ovulatory cycle is minimal, and E_2—unlike LH and P_4–is not used to determine whether ovulation has been successful. The second discrepancy is in the peak P_4 concentration. This arises due to an overshoot of the data in the middle of luteal stage Λ. Since P_4 levels are known to peak clinically around this time, we consider this behavior to be within a physiologically relevant and normal range for the hormone. Further, because we assume that the ovarian stages are crude approximations to actual follicular dynamics, there may be substantial variability in the trajectories that may nevertheless yield normal ovulatory function, as illustrated in [30].

Testosterone-mediated dysfunction

A fundamental change in the reduced framework is the omission of testosterone, T. Although absent from the model, we may still examine how T might influence pathological ovulation. This approach also serves as proof of concept when using the reduced model in lieu of the original one. To re-incorporate T into the present framework, we modify relevant parameters. Following [15], we let α denote the degree of insulin influence, where $\alpha = 0$ reflects a normal state with basal insulin (and hence T) levels. Assuming testosterone remains constant over time, we define its concentration using a linear function in α, denoted T_α:

$$T_\alpha = T_0 \cdot [1 + (\delta_T - 1) \cdot (1 + \alpha)]/\delta_T, \tag{12}$$

where T_0 is the initial T concentration in the absence of hyperinsulinemia and δ_T is the first-order clearance rate constant for T, as defined originally. The parameters to be altered by T in the reduced model are t_{g1}, v_{0L}, $K_{L,P}$, and h_1. We only consider the case of normal luteinization (see [15] for details) because we have omitted FSH-dependent upregulation of follicle LH receptors, which would impact parameter h_2. We first redefine $t_{g1} \to t_{g1}(1 + \alpha)$, based on the assumption made for precursor E_2. To incorporate the remaining modifications, we also redefine the parameters $v_{0L} \to v_{0L}\xi_1$, $K_{L,P} \to K_{L,P}\xi_2$, and $h_1 \to h_1\xi_3$ for $\alpha > 0$,

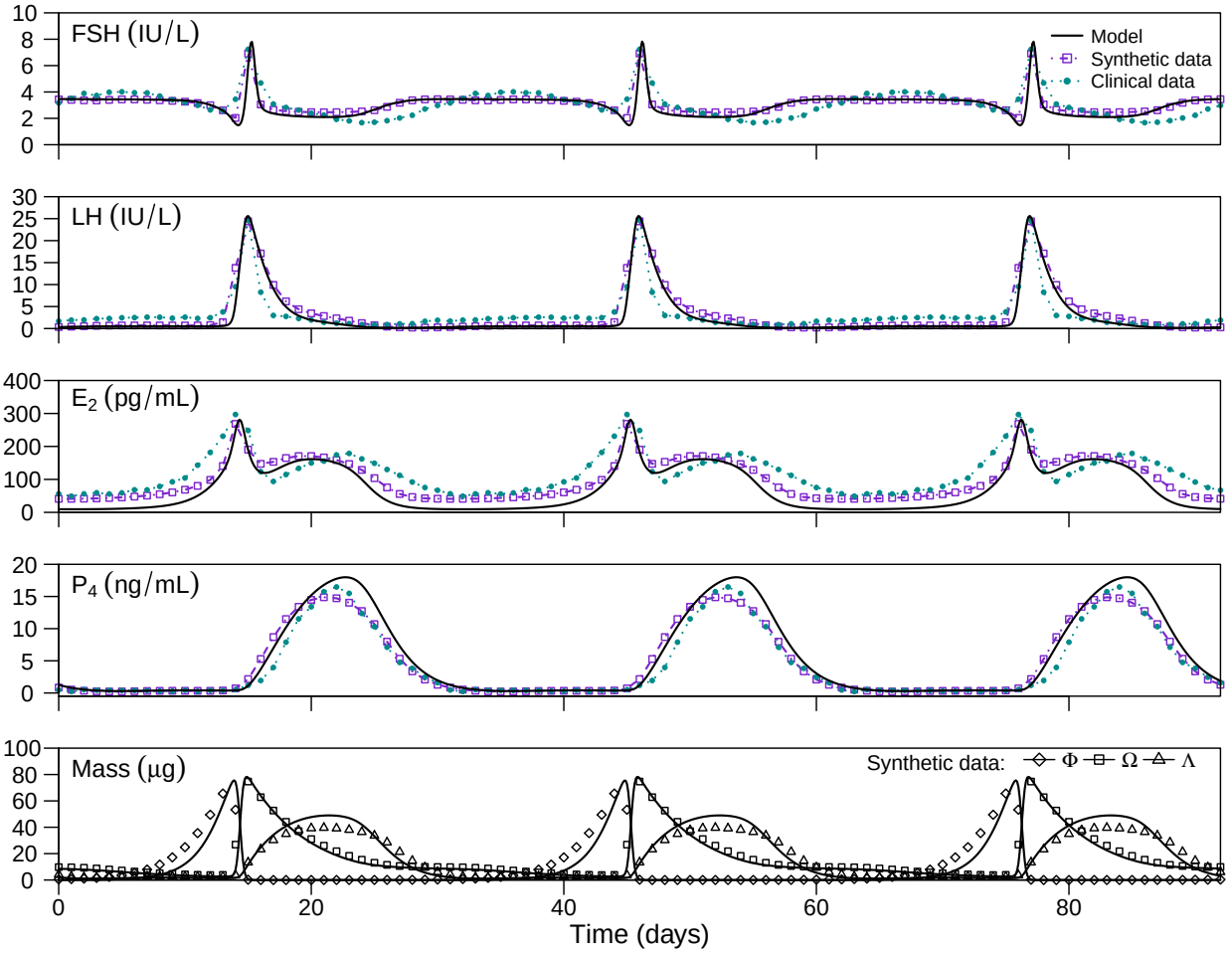

Figure 2. Fit of reduced model to Graham-Selgrade model [15] over 92 days. FSH and LH are displayed in standard international units according to the 2nd international reference preparation, where 1 IU FSH = 45 μg and 1 IU LH = 15 μg [33]. Conversion factors are based on the NIH preparation used in [27], which is also the source of the clinical data shown for FSH, LH, E_2, and P_4. Ovarian stages reflect synthetic data from the original model only.

where

$$\xi_1 = \frac{(\beta_1 + T_0) \cdot T_\alpha}{(\beta_1 + T_\alpha) \cdot T_0}, \tag{13a}$$

$$\xi_2 = \frac{1 + \beta_2 T_\alpha}{1 + \beta_2 T_0}, \text{ and} \tag{13b}$$

$$\xi_3 = \frac{1 + \beta_3}{1 + \beta_3 T_\alpha / T_0}. \tag{13c}$$

The ξ_i in Equations (13) determine the scaling of the model parameters as insulin influence increases and are plotted in Figure 3a. The constants β_i are defined according to the original model, with the caveat that bifurcation values of α may be shifted based on the values of these parameters. The derivation of the ξ_i are given in Appendix B.

In Figure 3b, we plot the long-term local maximum and minimum values corresponding to the LH surge for various α. At first glance, the model displays considerable sensitivity to the magnitude of α, such that periodic behavior is sustained for roughly $\alpha < 0.6$, followed by a Hopf bifurcation, characterized by

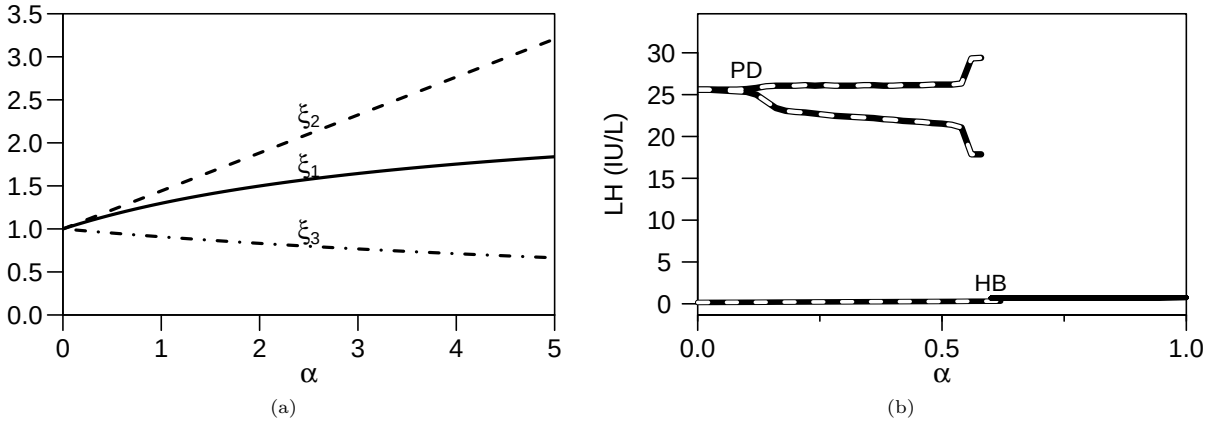

Figure 3. Insulin influence and testosterone in reduced model development and results. (a) Dimensionless functional forms used to incorporate T into reduced model, as in Equations (12) and (13). Each ξ_i contributes a T-dependent change (percent increase or decrease) in relevant parameters from the original model [15]. ξ_1 increases LH synthesis parameter v_{0L}, ξ_2 increases P_4-mediated LH inhibition parameter $K_{iL,P}$, and ξ_3 decreases FSH sensitivity parameter h_1. α: degree of insulin influence. (b) Simulated bifurcation diagram depicting adjusted role for T and insulin influence (α). Maximal and minimal LH concentrations are shown for various values of $\alpha \geqslant 0$. For $\alpha < 0.2$, LH oscillates between two values, suggesting a stable limit cycle. LH peaks alternate between consecutive IOIs for $\alpha \geqslant 0.2$, suggesting a period-doubling bifurcation (PD) with stable oscillations.

a loss of periodicity and a stable equilibrium. A stable limit cycle is roughly evident for $\alpha < 0.1$, with an apparent period doubling bifurcation giving rise to alternating LH surge amplitudes. Minimal LH levels remain relatively constant. This suggests that the reduced framework responds to elevated T by altering the amplitudes and timing of LH surges. Although the dynamic mechanisms governing ultimate dysfunction may differ from the original model, we are able to capture disruptive behavior, which takes the form of sustained oscillations under normal luteinization, with slightly shorter limit cycle lengths, as seen in [15]. The primary discrepancy is that the original model, under normal luteinization, maintains limit-cycle behavior for a wider range of α, i.e. for $0 < \alpha < 5$. However, under premature luteinization, the Graham-Selgrade model does undergo a Hopf bifurcation near $\alpha = 4.5$. Collectively, these results suggest that the reduced model with testosterone-mediated feedback illustrates a more severe level of dysfunction given the right trigger. Interestingly, should t_{g1} *α-independent*, the reduced model exhibits sustained limit cycle behavior for a much wider range of α (results not shown). This suggests that one mechanism of dysfunction might depend more on the presence of increased E_2 rather than an androgen-driven response.

4.3. Emergent behavior, phenotypes, and clinical relevance

From Steps 3 and 4 of Algorithm 1, we obtain an ensemble of trajectories from numerical simulations. By observation of these results, we find that we can use the values of the IOIs to ensure that pathological trajectories are reflected by the presence of abnormally long or short IOIs at any time. Considering that IOI is often the first step in recognizing a problem in ovulation, we wish to study the characteristics of individuals–each with their own parameter regime in the new model–who might be considered 'abnormal' from a clinical office visit. This approach is useful as it allows us to *study mechanisms of dysfunction based on limited information*, such as the time since the last period. Therefore, the criterion we use to categorize individuals is based solely on IOIs calculated throughout one's ovulatory trajectory. Specifically, we assign each trajectory (representing one person) to one of two phenotypes. The *regular* phenotype describes simulations in which both minimal and maximal IOIs fall between 25 and 35 days, which is the textbook standard range for normal ovulatory cycles [4]. The *irregular* phenotype describes those simulations failing to satisfy this criterion, i.e. those containing at least one IOI outside of the standard range.

Figure 4 shows hormone trajectories over 186 days for two representative solutions, one regular and one irregular. For reference, the timing of the LH surge for the regular phenotype is indicated with a vertical line. Stable limit cycle behavior is exhibited for the regular cycle with a characteristic length of 30.9 days.

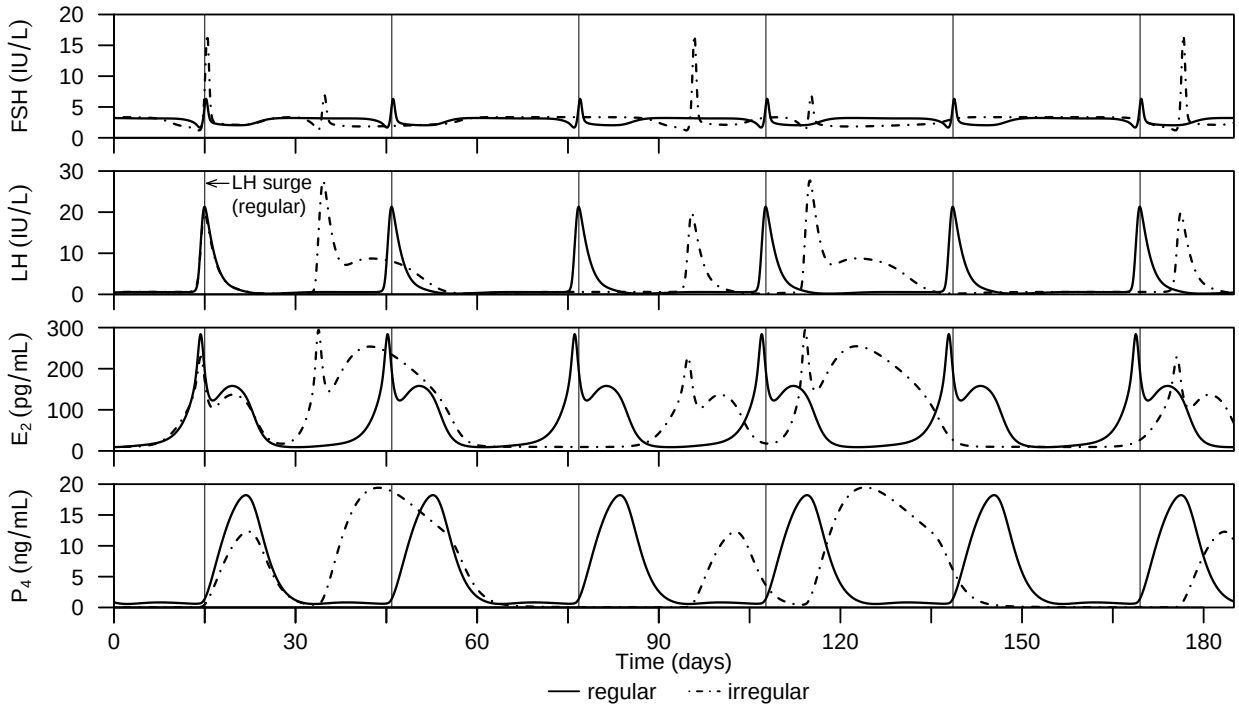

Figure 4. Comparison of representative regular and irregular trajectories simulated by the reduced model. The regular cycle displays a characteristic length of 30.9 days. The irregular cycle has a total length of 80.7 days, with IOIs of 19.5 and 61.2 days.

The irregular phenotype, however, consists of nonuniform behavior of the major hormones, indicating a certain degree of intra-individual variation. Specifically, the irregular limit cycle has a length of 80.7 days, with 19.5 and 61.2 days passing between consecutive LH surges. Although hormone levels are relatively normal through the course of the irregular cycle, there are marked differences in hormone patterns that could suggest ovulatory dysfunction.

Phenotype refinement

To examine how the important parameters and the accuracy of their accompanying numerical solutions when fit to clinical data vary, we calculate the mean squared error (MSE) between the model output (variables LH, FSH, E_2, and P_4) and the averaged data in [27]. We do observe a threshold MSE value—estimated from the MC output—above which all irregular phenotype results lie and below which roughly 85% of regular results lie. We use this threshold to assign an additional subcategory to simulations belonging to the regular phenotype. Specifically, regular solutions that yield MSE values below the computed threshold, and hence fit hormone data relatively well, are denoted $regular^+$. Regular solutions that yield above-threshold MSE values, and hence fit hormone data less well, are denoted $regular^-$. Qualitatively, we consider the regular$^+$ phenotype to reflect 'regular IOI-regular dynamics' and regular$^-$ to reflect 'regular IOI-irregular dynamics'. Notably that there does exist a subset of parameters for which the IOI varies by 50%, where both regular and irregular IOIs are observed yet the limit cycle length is fixed. Because of this, regular$^+$ implies both low intra-cycle hormone variability compared with data and also low IOI variability.

In Figure 5, we compute 95% confidence intervals of simulated hormone concentrations over four months to examine how hormone profiles influence these refined phenotypes. Briefly, to compute the confidence intervals for each phenotype, we calculate the upper 95% and lower 5% quantiles of the simulated trajectory data over time and then shade region in between the two boundaries. The result is an aesthetically improved, yet still representative, illustration of the trends in individual trajectories. As before, the simulated LH surge of the first cycle is forced to occur on day 15. Regular$^+$ simulations exhibit the least variation across all

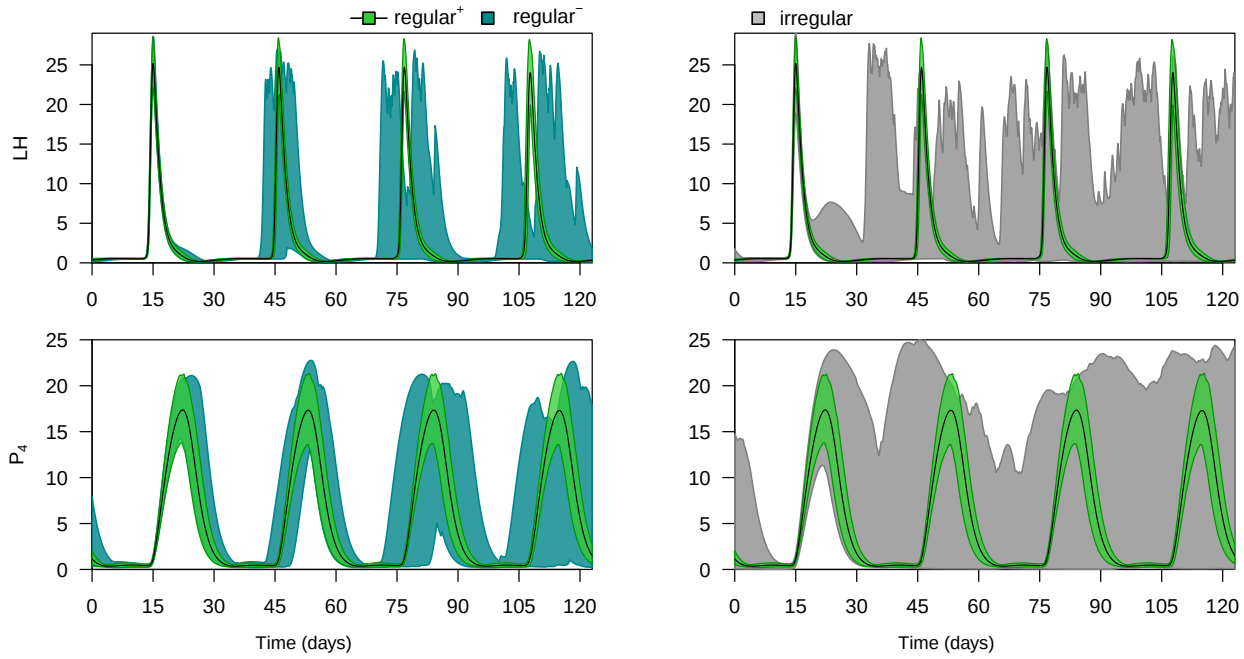

Figure 5. 95% confidence intervals of reduced model output over four regular cycles. Regular^{+} (green) compared to (*left*,teal) regular^{-} and (*right*,gray) irregular phenotypes. Time-dependent regular^{+} means are indicated with black curves.

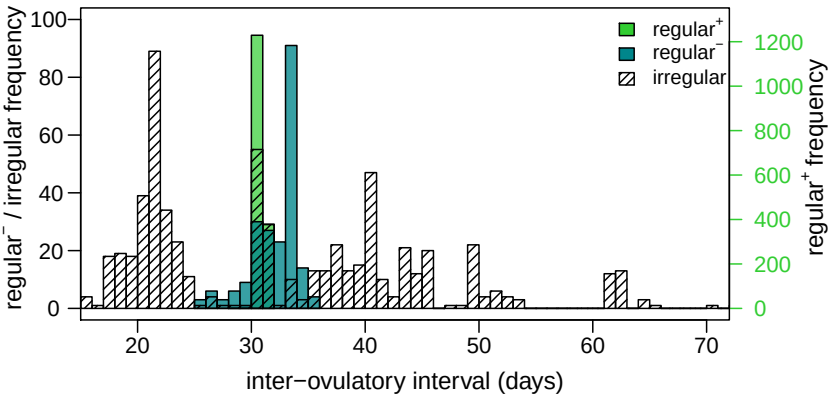

Figure 6. Distribution of inter-ovulatory intervals (IOIs) across phenotypes. Histogram computes the range of frequencies based on individual IOIs, rather than the set of IOIs belonging to independent trajectories. Irregular phenotypes exhibit significantly more variation in IOI than regular phenotypes.

cycles (green regions). Beyond the first LH surge, regular$^-$ phenotypes (left panel, teal regions) have more variation in the timing of characteristic ovulatory events (e.g. LH surge and luteal formation) than regular$^+$, but considerably less variation than the irregular phenotypes (right panel, gray region). As a result, we have reduced predictability of ovulation when we refine phenotypes according to data fitting. In addition, no level of observed irregularity can produce a complete absence of ovulation, either through loss of oscillations or subthreshold hormone concentrations in LH or P_4.

In Figure 6, we examine the distribution of IOIs for each phenotype. Frequencies are determined by the collection of all IOIs, rather than a statistic describing generalized behavior. This is especially useful for the irregular case, which displays much wider variability than either of the regular phenotypes. Further, there appear to be multiple modes in the distribution of IOIs for irregular trajectories, observed at IOIs of 20, 30, and 40 days.

4.4. Important parameters: Identification and distributions

We use the results from Step 5 of Algorithm 1 to calculate empirical parameter distributions, which we can now examine in a phenotype-specific manner. In Figure 7 we illustrate resulting distributions for eight of the reduced model parameters: the complete distribution (white boxes), along with the distributions for the regular phenotype (slanted line boxes) and the irregular phenotype (solid gray boxes). In addition, we compute the empirical cumulative distribution functions for all parameters distinguished by phenotype (see Appendix C). These distribution results form the basis of our remaining model analysis and computational results.

Statistical significance

To assess whether each parameter distribution differs from its counterpart in the opposing phenotype, we use the Kolmogorov-Smirnov (KS) test, which determines whether two samples are drawn from the same distribution [34, 35]. The test uses the Kolmogorov-Smirnov statistic, which is defined as the L_∞ norm of the distance between two cumulative probability distribution functions. For each parameter, we apply `ks.test`, the R implementation of the two-sampled KS test, to analyze the phenotype-specific empirical distributions generated from our simulations.

KS test results for the parameter distributions are illustrated in Figure 8. Each box is shaded according to the minimal level of significance that allows us to accept the alternative hypothesis, i.e. that regular and irregular distributions are statistically different. Darker shaded squares correspond to higher levels of significance. Of the 27 parameters remaining in the reduced model, we identify eight that have significantly different distributions between regular and irregular phenotypes, with $p < 0.01$ (indicated by *). These parameters are given in Table 2, along with their associated p-values from the KS test. These are the same eight parameters shown in Figure 7. Our remaining analysis focuses on these eight important parameters.

4.5. Dimensional reduction of phenotypes

Beyond the structure manually imposed on the Monte Carlo dataset, we are interested in determining whether distinct phenotypes can be identified in another way. Patterns in the generated data may depend

Table 2. Eight parameters identified as most important based on the Kolmogorov-Smirnov test. Parameters are ranked in order from most (1) to least (8) significant, according to the p-value obtained.

Rank	p-value	Parameter	Description
1	2.44×10^{-15}	η	luteal E_2 production;
2	5.31×10^{-5}	v_F	maximal FSH synthesis rate;
3	4.44×10^{-4}	h_1	follicle sensitivity to FSH;
4	5.11×10^{-4}	$\hat{s}$	LH support maximal growth rate;
5	5.12×10^{-4}	K_{mL}	half-maximal E_2 stimulation level;
6	8.93×10^{-4}	δ_s	LH support decay rate;
7	2.51×10^{-3}	l	maximal luteolysis rate;
8	4.98×10^{-3}	f_1	maximal follicle growth rate.

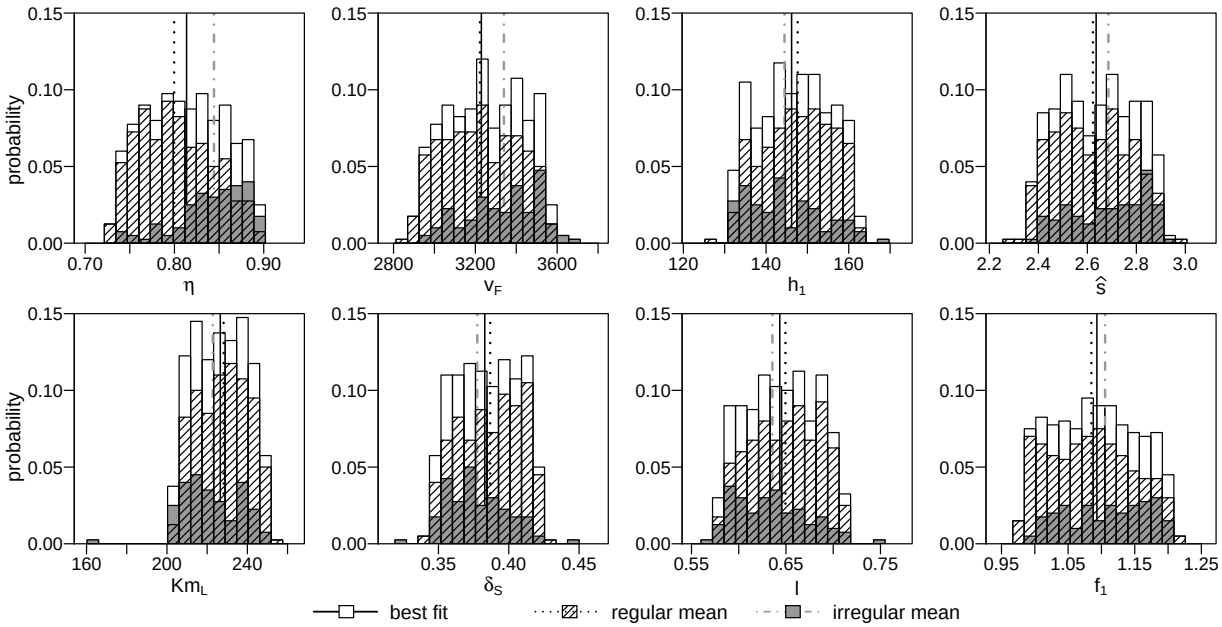

Figure 7. Parameter distributions from Step 5 of the algorithm.

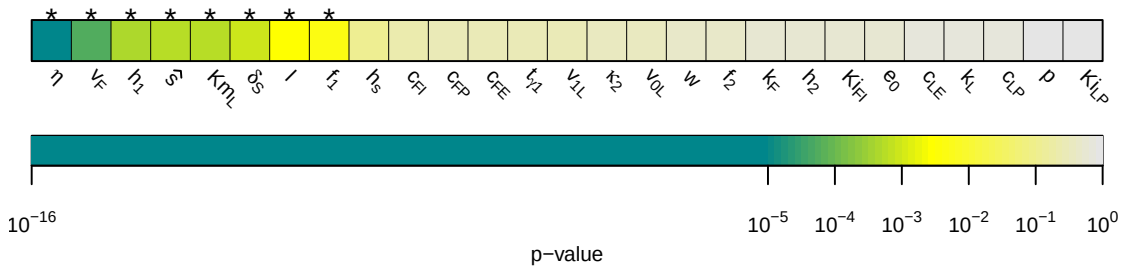

Figure 8. Two-sample Kolmogorov-Smirnov test. Shaded according to p-value, in increasing order from left to right. $^*p < 0.01$.

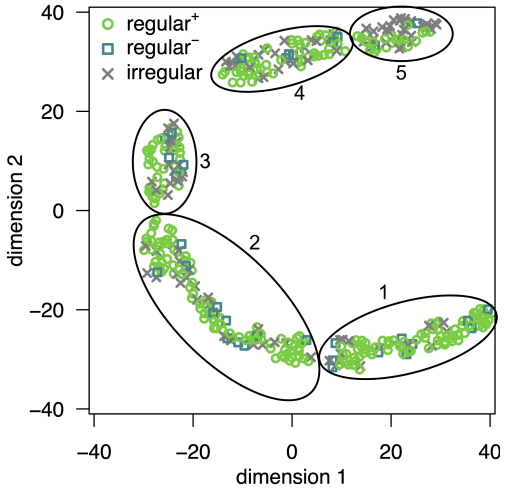

Figure 9. t-Distributed Stochastic Neighbor Embedding of model results. Dimensional reduction of identified phenotypes based on the eight significant parameters $\eta,\ v_F,\ h_1,\ \hat{s},\ K_{mL},\ \delta_s,\ l,\ f_1$ gives a two-dimensional embedding of model output.

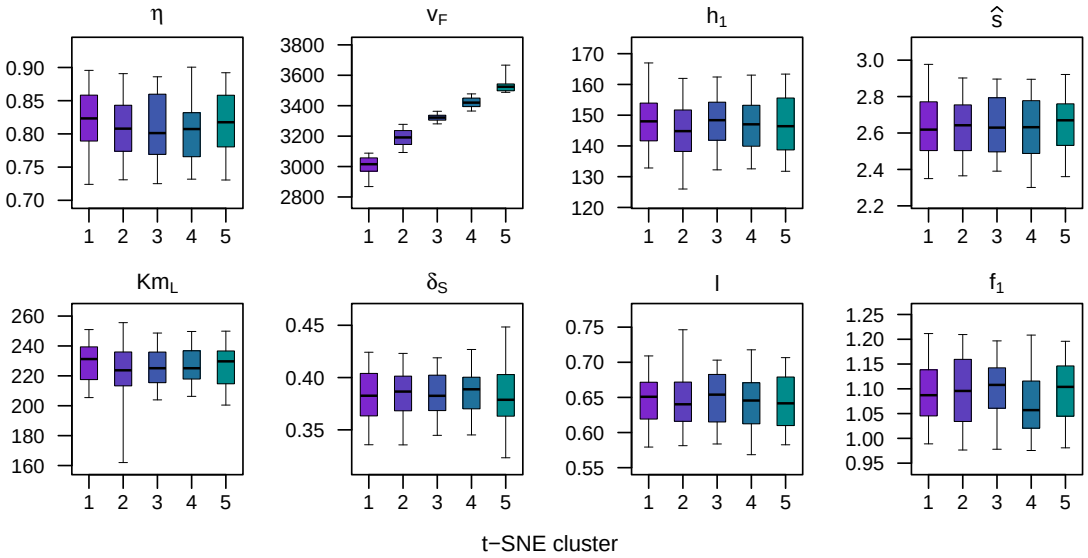

Figure 10. Significant parameter estimates for t-SNE clusters. Cluster-specific behavior is evident for parameter v_F, which corresponds to the maximal rate of FSH synthesis in the brain.

	t-SNE Cluster				
	1	2	3	4	5
$\text{regular}^{\pm}_{(100)}$	$0.86_{(28)}$	$0.81_{(33)}$	$0.70_{(13)}$	$0.71_{(18)}$	$0.48_{(8)}$
regular$^+$	*0.76*	*0.72*	*0.61*	*0.62*	*0.42*
regular$^-$	*0.10*	*0.09*	*0.09*	*0.09*	*0.06*
$\text{irregular}_{(100)}$	$0.14_{(14)}$	$0.19_{(22)}$	$0.30_{(16)}$	$0.29_{(22)}$	$0.52_{(26)}$
mean v_F	3008.0	3189.1	3321.8	3421.8	3530.7

Table 3. Proportion of t-SNE-clustered trajectories that belong to a particular phenotype. Subscripts in parentheses give the percentage distribution of all phenotype-specific trajectories (regular or irregular) among the five t-SNE clusters. Note: *regular*$^{\pm}$ values are the sum of regular$^+$ and regular$^-$ proportions.

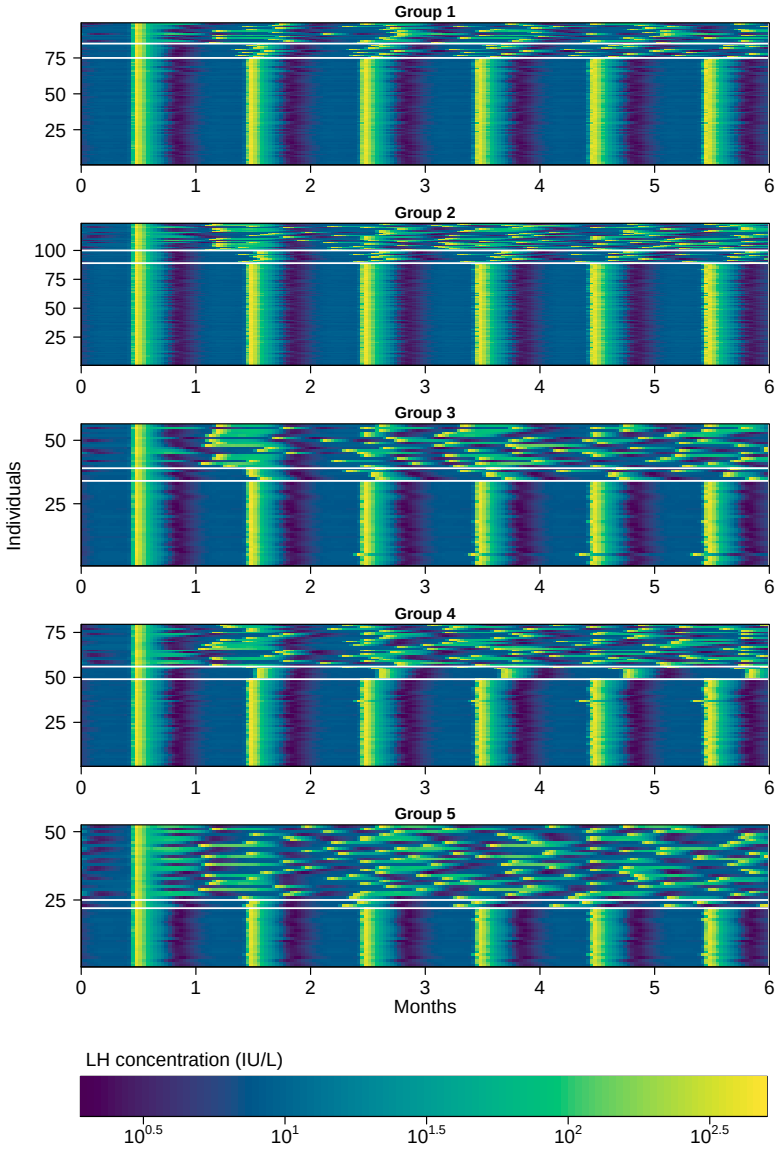

Figure 11. The collection of LH trajectories belonging to each t-SNE-identified group. Each row within a group panel is a single trajectory over 6 months, color-coded according the the magnitude of LH. From bottom to top, individual trajectories are plotted beginning with the regular^{+}, then regular^{-}, then irregular phenotypes. White lines indicate the transition between phenotypes.

on any of 93 data points for each of four hormones, or any of the 27 parameter estimates. Without a comprehensive understanding of the interplay between each of these elements, we seek a methodology that will answer the binary question of whether there are inherent differences (seen or unseen) between regular and irregular phenotypes. t-Distributed stochastic neighbor embedding (t-SNE) is a machine learning tool for reduction of high-dimensional data to lower dimensions [36]. We wish to determine whether phenotypes can be clearly clustered by a profile of selected model parameters.

To examine refined phenotypes based on the eight important parameter estimates, we implement a t-SNE of the parameter profiles, with points distinguished according to the assigned primary (regular or irregular) and secondary (+ or -) phenotypes. We use the *Rtsne* package in R to apply the t-SNE. In a two-dimensional reduction of the eight-dimensional parameter space, we find no discernible differences between phenotypes. Instead, five clusters do emerge from the two-dimensional t-SNE, which have been arbitrarily numbered one through five in Figure 9. These results indicate that the set of significant parameters cannot alone isolate reproductive phenotypes. In other words, although we can use phenotypes to identify reproductive parameter regimes, we cannot use the regimes themselves to decode their respective phenotypes. This is perhaps unsurprising given what little information has gone into our phenotyping approach. The resulting clusters tell us which characteristics are more closely related when considering our eight-dimensional parameter space.

Given these results, we can explore the characteristics of the five t-SNE clusters further by plotting the individual parameters according to cluster (see Figure 10). Of the eight important parameters we have identified, v_F—representing the maximal rate of FSH synthesis—is the only one that exhibits clear cluster-specific behavior. The other parameters vary by group, but not in any clearly discernible way. In Table 3, we calculate the distribution of regular and irregular phenotypes present in each cluster, accompanied by the mean v_F attained within each grouping. We also include the percentage of an overall phenotype belonging to each group. We find that v_F is positively correlated with the frequency of irregular phenotypes, to the extent that lower values of v_F occur in greater frequency with regular ovulatory cycles.

In Figure 11, we provide a visual representation of the results in Table 3, while also exploiting our time-dependent information at our disposal. In particular, we provide a two-dimensional representation of the simulations for LH over a period of six months, separated by t-SNE cluster. Each individual row corresponds to a Monte Carlo trajectory, with all regular$^+$ at the bottom, followed by regular$^-$, and irregular phenotypes closest to the top of each panel. White lines are added to provide a visible boundary between phenotypes. In all groups regular$^+$ individuals demonstrate predictable ovulatory function and relatively constant IOIs. However, regular $^-$ trajectories appear to become *more regular*-looking as we increase v_F toward Group 5. That is, there appears to be more uniformity in the LH concentrations, to the extent that trajectories 'line up' better with each other as we examine the groups in order of increasing v_F. Finally, and as expected, there is no immediately discernible pattern in the output for irregular trajectories. However, it does appear that even though LH trajectories are less uniform over time, there are nevertheless a relatively standard number of ovulatory cycles within the six-month timespan, as indicated by the LH surge concentrations in yellow. Collectively, these results suggest that the reduced model introduced herein displays ovulatory irregularity as a by-product of elevated FSH production.

5. Discussion

In this study we introduce a new, reduced endocrine model that inherently demonstrates both regular and irregular phenotypes, which we classify based on the timing of ovulation. The model produces distinct phenotypes as a result of altered time-independent parameter regimes and in the absence of disease-specific factors, e.g. testosterone-mediated dysfunction in PCOS. Through a comprehensive model evaluation algorithm, we identify a subset of model parameters that provide insight into physiological mechanisms of dysfunction. Further, the reduced framework provides a testable hypothesis of model prediction: consistently similar inter-ovulatory intervals (IOIs) between individuals likely reflect similar reproductive hormone dynamics. These results also imply that there is potentially a many-to-one relationship between endocrine states and observable endocrine dynamics and dysfunction, e.g., between physiological parameters and hormone dynamics. This fuzzy causation is not uncommon in physiologic systems or in biomedicine broadly; but

to develop better clinical treatment, it is critical to minimize the number of potential causes of an observable problem while maximizing the understanding of the physiologic mechanics driving endocrine dynamics.

Based on the most significant parameters identified by the present work, the model highlights key mechanisms associated with pituitary hormone synthesis (v_F, K_{mL}), follicle growth (h_1, f_1), luteal dynamics ($\hat{s}$, δ_s, l), and ovarian E_2 production (η). However, the redundancy in the biological processes associated with these parameters allows us to more succinctly characterize sources of dysfunction based on two major processes: (1) altered follicular growth and (2) feedback associated with E_2 concentrations. Both of these biological processes are relevant to our discussion of PCOS and EM [37, 3], to the extent that we can adapt the current model to circumstances specific to these disorders, especially where downstream signals–beyond the typical reproductive hormone profiles–are concerned.

Altered follicular growth. In vitro experiments suggest that granulosa cells may be more sensitive to FSH in PCOS, affecting follicle growth [4]. Follicular growth is stimulated by FSH, and the model's maximal FSH synthesis rate parameter modulates pituitary stores of FSH. In the irregular phenotype, there is a tendency toward increased mid-cycle FSH levels, which are considered elevated for physiological FSH concentrations (roughly 20 IU/L). In addition, increased v_F—identified as a distinguishing parameter in our t-SNE analysis—accompanies increased peak FSH levels, regardless of phenotype. This suggests that the reduced model accounts for ovulatory disruption through changes in FSH, which is also consistent with the current literature, wherein elevated FSH is a determining factor in premature ovarian insufficiency (POI) [37, 38]. Although the maximal FSH levels produced by the model are relatively lower than those expected from a confirmed POI individual, these levels also occur in the face of residual ovulatory function, albeit irregular.

E_2-mediated feedback. Variations in E_2 are implicated in multiple manifestations of ovulatory dysfunction. For example, decreased E_2 is characteristic of menopausal women. Prolonged exposure to elevated E_2 has been associated with ovulatory disruption in previous mathematical models [18, 39], and elevated E_2 formation has been found in *in vitro* PCOS models [4]. Further, E_2 acting via the estrogen receptor-β is a primary trigger for inflammation leading to severe pain in EM [4, 1]. As such, increased physiological, but not necessarily pathological, E_2 levels can contribute to dysfunction downstream of the ovulatory processes discussed herein. In the current work, parameters associated with luteal stage dynamics are altered in the irregular phenotype, such that appearance and disappearance rates of LH support are increased and decreased, respectively. This supports greater ovarian mass during the luteal phase, which contributes to significantly elevated E_2 during this period. Simulated irregular cycles are also associated with higher E_2 production rates from functional luteal cells and increased pituitary sensitivity to E_2, which can prematurely trigger the LH surge. Elevated subthreshold E_2 prolongs suppression of FSH and LH release into the serum, thereby inhibiting follicle growth. In extreme cases, this results in two ovulation events close together, followed by an increased period of ovulatory suppression. This is exhibited in Figure 4, with a two-month lapse between ovulation events in the representative irregular phenotype.

The reduced framework is amenable to modifications allowing us to explore testosterone-mediated ovulatory dysfunction, as in [15]. Clinically, it remains unclear how disruptions propagate in the face of hyperandrogenism. We find that when we alter pituitary-specific processes—particularly with respect to LH production—and follicle growth processes with linearly increasing levels of T, cyclic behavior ceases. Further, the steady state approached for sufficiently large insulin influence includes a clinically low level of LH. In contrast, LH is often found to be elevated in PCOS populations, but with high interindividual variability. These results suggest that we may not associate the T-mediated disruptions within the reduced framework with specific PCOS symptoms, but rather as part of a more generalized manifestation of ovulatory dysfunction due to abnormal responses in the HPO axis.

Without testosterone as an explicit driver of dysfunction, all phenotypes in the new endocrine model exhibit successful ovulatory events, with some variations in frequency. Hormone concentrations arising from irregular cycles lie within their respective physiological ranges, and interestingly, the range of IOI for irregular phenotypes is consistent with the ranges reported for individuals near menarche or approaching menopause [4]. The model cannot, nor is it designed to, produce an increase in small ovarian cysts that can accompany PCOS. Yet, it does capture observable information—such as cycle length and the absence of androgen excess—that could indicate a less severe phenotype of PCOS, which would be characterized

mathematically by oligo-ovulation. It also appears that our ability to identify defects via reproductive hormones depends on the sampling frequency of data.

5.1. Limitations

A number of limitations are evident in process of mathematically modeling female endocrine physiology, especially in the realm of reproductive hormone regulation. Although the model reduction introduced here allows us to further refine our study of parameter-mediated dysfunction, there are some challenges that require further analyses to overcome. We discuss a few of these here.

Global sensitivity analysis and the parameter space

The model evaluation algorithm, especially in Steps 1 and 2, provides a clear procedure to bridge the gap between the original and the reduced model. In particular, we use synthetic data fit the model initially. In doing so, we are able to use ovarian stage data—which is unavailable clinically—to aid in model fitting. Most of the parameters obtained from this approach are similar, with respect to orders of magnitude, to their counterparts in the original model. The primary differences in parameter values are due to the removal of testosterone. Because we deem the reduced model as a surrogate for the original model, the similarities between parameter sets is neither unanticipated nor undesirable. Further, a preliminary attempt to fit the reduced model to the averaged clinical data in [27] over a 3-month period rather than synthetic data yields equally similar parameters (results not shown). That is, either approach results in a parameter regime that remains close to that of the original model. This suggests a local minimum in the parameter space, which may be explored with an in depth global sensitivity analysis (GSA).

A natural course of action in determining salient model behavior is global sensitivity analysis (GSA) of parameters. This allows us to determine the relative sensitivity of model output to changes in the parameters. There are multiple challenges associated with the Graham-Selgrade model that make GSA a suboptimal next step in model analysis. First, the model contains considerably more parameters than the data available for estimation. Second, coupling between state variables is highly nonlinear. Third, stable limit cycles are not guaranteed for all parameter combinations. Collectively, standard GSA approaches provide limited insight. In particular, a partial rank correlation coefficient (PRCC)-based approach would be inappropriate, as simulations do not yield monotonic hormone responses that can be interpreted in any meaningful way (preliminary work, not shown). Alternatives such as the extended Fourier amplitude sensitivity test (eFAST) [40] may also prove more useful, as discussed in [32]. GSA can only be as good as the signal being measured in response to variations in the parameters. The challenge with models of ovulation is the periodicity of model solutions, coupled with a reasonably stiff system of differential equations. As a result, appropriate selection of model output remains a challenge, but an alternative approach in future work could focus on the rates of change in numerical solutions, as in [24].

Complexities of data and analytical challenges

Data for primary reproductive hormone measurements are useful for delineating broadly defined clinical abnormalities and quantifying generalized ovulatory states. Two prototypical data sets reported in the literature include pituitary and ovarian hormones collected daily over the course of a typical cycle [27, 29]; we use the data in [27] in this paper. However, these data provide only a partial view to more subtle abnormalities. For example, PCOS can result in the complete absence of, or sporadic, ovulation. But, distinguishing between mechanisms governing these two observable clinical manifestations is difficult because clinically feasible diagnostic tools rely on measurements taken either at a single time point or over the course of a few hours [4]. A similar challenge lies in the diagnosis of EM–in which a collection of symptoms and isolated hormone measurements rarely point to a single cause [41]. In the worst cases, diagnosis itself is a months- or years-long process that can reduce quality of life of those affected [42, 3]. In essence, we would require data spanning multiple months in order to build a comprehensive hormone profile with any hope of revealing important reproductive features, especially in the absence of clearly identifiable ovulatory states.

In the present context, it is important to note that a high-fidelity, data-driven, robust and expansive definition of normal ovulatory function does not currently exist. This makes defining 'normal' and 'dysfunctional' a complex task, as dysfunction is usually defined as a deviation from normal. Because of this,

we adopt a narrow definition of normal and consequently limit our ability to discover different-from-normal phenotypes. This limitation is due to the lack of data; with more data, the methodology here could provide more phenotypic fidelity. Ideally, we seek an alternative to patterns in hormone dynamics to distinguish between ovulatory phenotypes, with the hope that identifying underlying mechanisms of dysfunction lies in our ability to connect clinical symptoms with mechanisms that may not be apparent in hormone measurements alone.

The available data have three primary limitations that influence our work. First, recall that normal is generally poorly defined, where 'normal' means no known pathophysiologic cycle features. Second, it is known that there is substantial variation in IOIs even for an individual. For example, it is not uncommon for the same person to have IOIs that vary from 20 to 40 days; these data obscure such intraindividual variability by taking an average. And third, because the data are an average, they induce three potential issues whose presence we may not be able to detect: (i) an average can fail to represent anyone if the mean is not representative of the population; (ii) an average smooths individualized daily variability, which can be substantial, is not present in data, and will not be explicitly estimated by the models; and (iii) variability of cycle length and dynamics coupled to cycle length for both 'normal' and 'abnormal' cycle lengths is entirely missing.

5.2. Conclusions

The over-arching goal is to use models for predictive decision support and to deepen our understanding of physiology. We wish to not only understand mechanisms of function but also the factors that differentiate those mechanisms. Endometriosis and polycystic ovary syndrome are two high-impact disorders governed by physiology, both with incompletely understood etiologies. We wish to shed insight on these disorders to better inform intervention and treatment decisions. The current model and evaluation process allows us to delineate dysfunction based on physiology, which can then be applied to these disorders of interest in future work. As constructed, the model is flexible enough to allow us to (1) highlight important—generalizable or disorder-specific—mechanisms of dysfunction; (2) determine the clinical span of the model compared to other models and alternative data sets; (3) identify how and when clinical intervention is feasible, necessary, or effective; and (4) reverse-engineer parameter profiles to differentiate physiological from pathological outcomes.

Declarations of interest

None.

Funding

E.J.G. reports funding from the Simons Foundation [MPS 585858]. D.A. reports funding from the NLM R01 [LM012734].

References

[1] S. Vannuccini, S. Clemenza, M. Rossi, F. Petraglia, Hormonal treatments for endometriosis: The endocrine background, Reviews in Endocrine and Metabolic Disorders 23 (3) (2022) 333–355.

[2] L. Pal, D. B. Seifer, Polycystic Ovary Syndrome, 2nd Edition, Springer Cham, 2022.

[3] T. M. Gruber, S. Mechsner, Pathogenesis of endometriosis: the origin of pain and subfertility, Cells 10 (6) (2021) 1381.

[4] J. F. Strauss, R. L. Barbieri, Yen & Jaffe's Reproductive Endocrinology E-Book: Physiology, Pathophysiology, and Clinical Management, 7th Edition, Elsevier Health Sciences, 2013.

[5] T. Tanbo, P. Fedorcsak, Endometriosis-associated infertility: aspects of pathophysiological mechanisms and treatment options, Acta obstetricia et gynecologica Scandinavica 96 (6) (2017) 659–667.

[6] J. Evans, T. Wilkinson, D. Wall, A two-pathway mathematical model of the lh response to gnrh that predicts self-priming, International journal of endocrinology 2013.

[7] A. Pratap, K. L. Garner, M. Voliotis, K. Tsaneva-Atanasova, C. A. McArdle, Mathematical modeling of gonadotropin-releasing hormone signaling, Molecular and cellular endocrinology 449 (2017) 42–55.

[8] P. M. Schlosser, J. F. Selgrade, A model of gonadotropin regulation during the menstrual cycle in women: Qualitative features, Environmental health perspectives (2000) 873–881.
[9] H. Lacker, Regulation of ovulation number in mammals. a follicle interaction law that controls maturation, Biophysical journal 35 (2) (1981) 433–454.
[10] F. Clément, D. Monniaux, Multiscale modelling of ovarian follicular selection, Progress in biophysics and molecular biology 113 (3) (2013) 398–408.
[11] N. M. Panza, A. A. Wright, J. F. Selgrade, A delay differential equation model of follicle waves in women, Journal of biological dynamics 10 (1) (2016) 200–221.
[12] F. Clément, D. Monniaux, Mathematical modeling of ovarian follicle development: A population dynamics viewpoint, Current Opinion in Endocrine and Metabolic Research 18 (2021) 54–61.
[13] M. S. Breen, D. L. Villeneuve, M. Breen, G. T. Ankley, R. B. Conolly, Mechanistic computational model of ovarian steroidogenesis to predict biochemical responses to endocrine active compounds, Annals of biomedical engineering 35 (6) (2007) 970–981.
[14] C. Louw, Computational modelling of steroid hormone biosynthesis and metabolism, Ph.D. thesis, Stellenbosch: Stellenbosch University (2020).
[15] E. J. Graham, J. F. Selgrade, A model of ovulatory regulation examining the effects of insulin-mediated testosterone production on ovulatory function, Journal of theoretical biology 416 (2017) 149–160.
[16] R. Bogumil, M. Ferin, J. Rootenberg, L. Speroff, R. V. Wiele, Mathematical studies of the human menstrual cycle. i. formulation of a mathematical model, The Journal of Clinical Endocrinology & Metabolism 35 (1) (1972) 126–143.
[17] R. Bogumil, M. Ferin, R. V. Wiele, Mathematical studies of the human menstrual cycle. ii. simulation performance of a model of the human menstrual cycle, The Journal of Clinical Endocrinology & Metabolism 35 (1) (1972) 144–156.
[18] L. H. Clark, P. M. Schlosser, J. F. Selgrade, Multiple stable periodic solutions in a model for hormonal control of the menstrual cycle, Bulletin of mathematical biology 65 (1) (2003) 157–173.
[19] A. O. Hendrix, C. L. Hughes, J. F. Selgrade, Modeling endocrine control of the pituitary–ovarian axis: Androgenic influence and chaotic dynamics, Bulletin of mathematical biology 76 (1) (2014) 136–156.
[20] I. Reinecke, P. Deuflhard, A complex mathematical model of the human menstrual cycle, Journal of Theoretical Biology 247 (2) (2007) 303–330.
[21] S. Röblitz, C. Stötzel, P. Deuflhard, H. M. Jones, D.-O. Azulay, P. H. van der Graaf, S. W. Martin, A mathematical model of the human menstrual cycle for the administration of gnrh analogues, Journal of theoretical biology 321 (2013) 8–27.
[22] C. Chen, J. P. Ward, A mathematical model for the human menstrual cycle, Mathematical Medicine and Biology 31 (2014) 65–86.
[23] F. Clément, Multiscale mathematical modeling of the hypothalamo-pituitary-gonadal axis, Theriogenology 86 (1) (2016) 11–21.
[24] S. Fischer-Holzhausen, S. Röblitz, Hormonal regulation of ovarian follicle growth in humans: Model-based exploration of cycle variability and parameter sensitivities, Journal of Theoretical Biology (2022) 111150.
[25] A. O. Hendrix, J. F. Selgrade, Bifurcation analysis of a menstrual cycle model reveals multiple mechanisms linking testosterone and classical pcos, Journal of theoretical biology 361 (2014) 31–40.
[26] D. Arbeláez-Gómez, S. Benavides-López, M. P. Giraldo-Agudelo, J. P. Guzmán-Álvarez, C. Ramirez-Mazo, L. M. Gómez-Echavarría, A phenomenological-based model of the endometrial growth and shedding during the menstrual cycle, Journal of Theoretical Biology 532 (2022) 110922.
[27] R. I. McLachlan, N. L. Cohen, K. D. Dahl, W. J. Bremner, M. R. Soules, Serum inhibin levels during the periovulatory interval in normal women: relationships with sex steroid and gonadotrophin levels, Clinical endocrinology 32 (1) (1990) 39–48.
[28] G. Hripcsak, D. Albers, High-fidelity phenotyping: richness and freedom from bias, J Am Med Inform Assoc.
[29] C. K. Welt, D. J. McNicholl, A. E. Taylor, J. E. Hall, Female reproductive aging is marked by decreased secretion of dimeric inhibin, The Journal of Clinical Endocrinology & Metabolism 84 (1) (1999) 105–111.
[30] C. C. Keefe, M. M. Goldman, K. Zhang, N. Clarke, R. E. Reitz, C. K. Welt, Simultaneous measurement of thirteen steroid hormones in women with polycystic ovary syndrome and control women using liquid chromatography–tandem mass spectrometry, PloS one 9 (4) (2014) e93805.
[31] S. M. Blower, H. Dowlatabadi, Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example, International Statistical Review/Revue Internationale de Statistique (1994) 229–243.
[32] S. Marino, I. B. Hogue, C. J. Ray, D. E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology, Journal of theoretical biology 254 (1) (2008) 178–196.
[33] A. Labhart, Clinical endocrinology: theory and practice, Springer Science & Business Media, 2012.
[34] V. Rohatgi, A. Saleh, Wiley series in probability and statistics, Hoboken: John Wiley & Sons, Inc.
[35] L. Mora-López, J. Mora, An adaptive algorithm for clustering cumulative probability distribution functions using the Kolmogorov–Smirnov two-sample test, Expert Systems with Applications 42 (8) (2015) 4016–4021.
[36] L. van der Maaten, G. Hinton, Visualizing data using t-sne, Journal of machine learning research 9 (Nov) (2008) 2579–2605.
[37] S. Mikhael, A. Punjala-Patel, L. Gavrilova-Jordan, Hypothalamic-pituitary-ovarian axis disorders impacting female fertility, Biomedicines 7 (1) (2019) 5.
[38] X. Jiao, T. Meng, Y. Zhai, L. Zhao, W. Luo, P. Liu, Y. Qin, Ovarian reserve markers in premature ovarian insufficiency: Within different clinical stages and different etiologies, Frontiers in endocrinology 12.
[39] L. A. Harris, J. F. Selgrade, Modeling endocrine regulation of the menstrual cycle using delay differential equations, Mathematical biosciences 257 (2014) 11–22.
[40] A. Saltelli, S. Tarantola, K.-S. Chan, A quantitative model-independent method for global sensitivity analysis of model

output, Technometrics 41 (1) (1999) 39–56.
[41] L. Lode, M. Often Sveen, M. Rudnicki, Abnormal pathways in endometriosis in relation to progesterone resistance: a review, Journal of Endometriosis and Pelvic Pain Disorders 9 (4) (2017) 245–251.
[42] I. Kaur, V. Suri, S. V. Rana, A. Singh, Treatment pathways traversed by polycystic ovary syndrome (pcos) patients: A mixed-method study, PloS one 16 (8) (2021) e0255830.

Appendix A. Graham-Selgrade Model Description and Equations

The Graham-Selgrade model [15] uses a compartmental framework to examine changes in ovulation due to increased androgens. The model follows the approaches of [8, 18, 19] and comprises three major subsystems, which describe changes in the pituitary-ovarian axis with mechanisms of steroidogenesis: pituitary regulation, follicle dynamics, and ovarian steroidogenesis. Collectively, the model consists of 12 state variables, tracking serum concentrations of five important reproductive hormones, follicle stimulating hormone (FSH), luteinizing hormone (LH), estradiol (E_2), progesterone (P_4), and testosterone (T), along with precursors/intermediaries of LH, FSH, and T. It also describes the dynamics of three follicular stages and of the follicle response to LH, termed *LH sensitivity*. The final model contains 41[1] unknown parameters which are estimated—to a locally minimizing set—by fitting the model to data from the literature [27, 30].

The complete list of equations for the original Graham-Selgrade model may be found in Appendix A.

I. *Pituitary regulation.* LH and FSH are the primary hormones produced by the pituitary gland. Synthesis and release of these hormones are regulated by ovarian steroid hormones, including E_2, P_4, and T. The equations governing changes in FSH and LH are split between releasable (denoted FSH_ρ and LH_ρ) and serum (denoted FSH and LH) pools of the hormones and incorporate stimulatory and inhibitory feedback by ovarian steroids. Using this compartmental approach, we can differentiate feedback processes governing pituitary hormone synthesis versus release.

Here we provide a generalized description of pituitary dynamics. Let $H(t)$ denote the serum concentration of a pituitary hormone (either FSH or LH) and $H_\rho(t)$ its releasable amount at time t. For $H = FSH,\ LH$, the differential equations governing releasable and serum quantities have the form

$$\frac{dH_\rho}{dt} = k_{\text{synthesis}}(\cdot) - k_{\text{release}}(E_2, P_4)H_\rho, \tag{A.1}$$

$$\frac{dH}{dt} = k_{\text{release}}(E_2, P_4)H_\rho/V - \delta_H H. \tag{A.2}$$

Each $k(\cdot)$ term denotes a function of state variables and describes the change in hormone levels due to the process indicated. Synthesis of FSH and LH is determined by different processes—with precise arguments to $k_{\text{synthesis}}$ omitted to reflect this—whereas their release is mediated solely by E_2 and P_4. Release into the serum is scaled by the blood volume, V, and clearance of the hormones is assumed to be a first-order process, with rate constant δ_H. Regardless of the highly nonlinear form of ovarian feedback, the subsystem remains linear in H_ρ and H. Collectively, the pituitary subsystem comprises four differential equations, with Equations (A.1) and (A.2) defined explicitly for both FSH and LH.

II. *Follicle dynamics.* Follicle growth, maturation, and differentiation are assumed to occur in a series of three sequential stages: (1) *follicular*, (2) *ovulatory*, and (3) *luteal*. We denote these using variables $\Phi(t)$, $\Omega(t)$, and $\Lambda(t)$, respectively. The follicular phase is characterized by recruitment and growth of stimulated follicles. The ovulatory phase is characterized by ovum release from a designated follicle in response to a mid-cycle surge in LH. Finally, the luteal phase is characterized by the formation and, in the absence of fertilization, regression of the corpus luteum. The three follicular stages are modeled as follows:

$$\frac{d\Phi}{dt} = k_{\text{recruitment}}(T) + k_{\text{growth}}(FSH, T)\Phi - k_{\text{ovulation}}(FSH, LH)\Phi, \tag{A.3}$$

[1]The model presented in [15] contains a typographical error in one of the equations, which omits one parameter ($c_{\Phi,T}$) from the total parameter count cited.

$$\frac{d\Omega}{dt} = k_{\text{ovulation}}(FSH, LH)\Phi - k_{\text{luteal}}(S)\Omega, \tag{A.4}$$

$$\frac{d\Lambda}{dt} = k_{\text{luteal}}(S)\Omega - k_{\text{regression}}(S)\Lambda. \tag{A.5}$$

Transitions to subsequent stages are unidirectional and depend on pituitary hormone levels. The model also incorporates a role for T in follicle recruitment and growth. Graham and Selgrade further define a new LH support variable, $S(t)$, to model the tonic LH-dependence of growth and premature regression of the corpus luteum. Specifically, S decays exponentially (with rate δ_S) to 0 in the absence of LH and approaches a maximal level of 1 for sufficiently large LH:

$$\frac{dS}{dt} = k_{\text{activation}}(LH)(1 - S) - \delta_S S. \tag{A.6}$$

III. *Ovarian steroidogenesis.* Throughout the ovulatory cycle, follicles may produce E_2, P_4, and T. Intracellular steroid production is primarily FSH- and LH-dependent during a typical cycle and is subject to functional maturation of individual follicles. This subsystem exploits the *two-cell two-gonadotropin* theory of ovarian steroid production, which describes the differential functionality of theca cells and granulosa cells within ovarian follicles [4]. The Graham-Selgrade model also introduces a semi-mechanistic description of testosterone production for examining a role for insulin in promoting hyperandrogenism. For $T_\gamma(t)$ denoting the 'intermediate' concentration of T destined to be converted into E_2, we write

$$\frac{dT_\gamma}{dt} = k_{\text{entry}}(LH, \alpha) - k_{\text{aromatization}}(FSH)T_\gamma. \tag{A.7}$$

In a growing follicle, theca cells compose the outermost layers of cells surrounding the ovum and granulosa cells the innermost layers. Importantly, theca cells possess androgen (i.e. T) production machinery and are stimulated by LH alone, whereas only neighboring granulosa cells can convert these androgens into estrogens, in an FSH-dependent process called *aromatization.* Therefore, we consider T_γ to reflect the average concentration of T that enters granulosa cells from theca cells.

Finally, we model the major ovarian outputs of the model: serum concentrations of E_2, T, and P_4:

$$\frac{dE_2}{dt} = k_{\text{basal,E}} - \delta_E E_2 + k_{\text{aromatization}}(FSH)T_\gamma \cdot f_E(\Phi, \Omega, \Lambda), \tag{A.8}$$

$$\frac{dT}{dt} = k_{\text{basal,T}} - \delta_T T + \left[k_{\substack{\text{ovarian} \\ \text{production}}}(LH, \alpha) + k_{\substack{\text{peripheral} \\ \text{production}}}(LH, \alpha)\right] \cdot f_T(\Phi, \Omega, \Lambda), \tag{A.9}$$

$$\frac{dP_4}{dt} = k_{\text{basal,P}} - \delta_P P_4 + k_{\text{secretion}}(LH) \cdot f_P(\Phi, \Omega, \Lambda). \tag{A.10}$$

The first two terms in Equations (A.8)–(A.10) represent basal secretion by the adrenal gland and first-order clearance of individual steroids, defined by rate constants $k_{\text{basal,I}}$ and δ_I, respectively, where $I = E, T, P$. The last term in each equation defines secretion of steroid hormones into the circulation, which is assumed to occur immediately upon production. The average production rate per follicle is multiplied by a function $f_I(\Phi, \Omega, \Lambda)$, $I = E, T, P$, that describes the relative contribution of each follicular stage to the production of a given steroid.

Importantly, steroidogenesis is altered through feedback from FSH and LH, according to the two cell-two gonadotropin theory. Whereas LH is required almost exclusively for T (theca only) and P_4 (theca and granulosa) production, FSH is entirely responsible for E_2 (granulosa only). Because P_4 is an androgen precursor in the theca, it is assumed that circulating P_4 is produced primarily by granulosa cells for modeling purposes. To address insulin's influence in ovulatory dysfunction, the Graham-Selgrade model contains a detailed formulation of T production, wherein ovarian and peripheral conversion of T from its precursors are treated as two distinct processes. In Equations

(A.7) and (A.9), the parameter α represents the relative degree to which insulin may increase T production.

Releasable FSH:
$$\frac{dFSH_\rho}{dt} = \frac{v_F}{1 + c_{F,I}\frac{S\Lambda}{K_{iF,I}+S\Lambda}} - k_F\frac{1 + c_{F,P}P_4}{1 + c_{F,E}E_2^2}FSH_\rho \tag{A.11}$$

Serum FSH:
$$\frac{dFSH}{dt} = \frac{1}{V}\cdot k_F\frac{1 + c_{F,P}P_4}{1 + c_{F,E}E_2^2}FSH_\rho - \delta_F FSH \tag{A.12}$$

Releasable LH:
$$\frac{dLH_\rho}{dt} = \left[\frac{v_{0L}T}{K_{L,T}+T} + \frac{v_{1L}E_2^n}{K_{mL}^n + E_2^n}\right]\cdot\frac{1}{1 + \frac{P_4}{K_{iL,P}\left(1+c_{L,T}T\right)}} - k_L\frac{1 + c_{L,P}P_4}{1 + c_{L,E}E_2}LH_\rho \tag{A.13}$$

Serum LH:
$$\frac{dLH}{dt} = \frac{1}{V}\cdot k_L\frac{1 + c_{L,P}P_4}{1 + c_{L,E}E_2}LH_\rho - \delta_L LH \tag{A.14}$$

Follicular phase:
$$\frac{d\Phi}{dt} = f_0\cdot\frac{T}{T_0} + \left[\frac{f_1FSH^2}{\left(\frac{h_1}{1+c_{\Phi,T}T/T_0}\right)^2 + FSH^2} - \frac{f_2LH^2}{\left(\frac{h_2}{1+c_{\Phi,F}FSH}\right)^2 + LH^2}\right]\cdot\Phi \tag{A.15}$$

Ovulatory phase:
$$\frac{d\Omega}{dt} = \frac{f_2LH^2}{\left(\frac{h_2}{1+c_{\Phi,F}FSH}\right)^2 + LH^2}\cdot\Phi - wS\Omega \tag{A.16}$$

Luteal phase:
$$\frac{d\Lambda}{dt} = wS\Omega - l(1-S)\Lambda \tag{A.17}$$

LH Support:
$$\frac{dS}{dt} = \frac{\hat{s}LH^m}{h_s^m + LH^m}\cdot(1-S) - \delta_s S \tag{A.18}$$

Serum T:
$$\frac{dT}{dt} = t_0 - \delta_T T + \left[t_1\mathcal{G}_1\left(F_1 + c_{T,F_2}F_2\right) + t_2\mathcal{G}_1\mathcal{G}_2F_1\right]\cdot \cdot\left[\Phi + \tau_1\Omega + \tau_2S\Lambda + \tau_3\left(1 - \frac{\Phi+\Omega+\Lambda}{\Psi}\right)\right] \tag{A.19}$$

Intermediate T:
$$\frac{dT_\gamma}{dt} = t_{g1}\mathcal{G}_1\mathcal{G}_2F_1 - \frac{t_{g2}FSH}{h_3 + FSH}T_\gamma \tag{A.20}$$

Serum E_2:
$$\frac{dE_2}{dt} = e_0 - \delta_E E_2 + \frac{t_{g2}FSH}{h_3 + FSH}T_\gamma\cdot[\Phi + \eta\Lambda S] \tag{A.21}$$

Serum P_4:
$$\frac{dP_4}{dt} = -\delta_P P_4 + \frac{pLH}{LH + h_p}\cdot\Lambda S \tag{A.22}$$

Functional Forms.

- **Insulin-stimulated conditions** ($\alpha > 0$)
  $\mathcal{G}_1 = \mathcal{G}_1(\alpha)$
  $\mathcal{G}_2 = \mathcal{G}_2(\alpha)$
  $\mathcal{D}(\alpha) = LH^2\left[\mathcal{G}_2 + A\right] + LH\left[\mathcal{G}_2B + A\cdot(B+C)\right] + A\cdot B\cdot C$
  $F_1(LH,\alpha) = LH^2/\mathcal{D}(\alpha)$
  $F_2(LH,\alpha) = LH/\mathcal{D}(\alpha)$

- **Basal conditions** ($\alpha = 0$)
  $\mathcal{G}_1 = \mathcal{G}_2 = 1$
  $\kappa_1 = 1 + A$
  $\kappa_2 = B + A(B+C)$
  $\kappa_3 = ABC$
  $\mathcal{D} = \kappa_1LH^2 + \kappa_2LH + \kappa_3$
  $F_1(LH) = LH^2/\mathcal{D}$
  $F_2(LH) = LH/\mathcal{D}$

Appendix B. Derivation of Testosterone-Dependent Terms

To incorporate testosterone implicitly in the reduced model, we need to modify parameters v_{0L}, $K_{iL,P}$, and h_1. We will use $\tilde{p}$ to denote parameters used in the original Graham-Selgrade model, which we will then redefine to incorporate into the reduced framework.

Derivation of ξ_1. In the original model, basal LH synthesis occurs at rate $\tilde{v}_{0L}T/(T+\beta_1)$, where $\beta_1 = K_{L,T} = 420$. We assume for the reduced model that

$$v_{0L}\xi_1 = \tilde{v}_{0L}\frac{T_\alpha}{T_\alpha+\beta_1},$$

where $\tilde{v}_{0L}$ is redefined so that $\xi_1 = 1$ when $T_\alpha = T_0$. That is, we define $\tilde{v}_{0L} = v_{0L}(T_0+\beta_1)/T_0$. It follows that

$$v_{0L}\xi_1 = v_{0L}\frac{T_0+\beta_1}{T_0}\frac{T_\alpha}{T_\alpha+\beta_1} = v_{0L}\frac{(\beta_1+T_0)\cdot T_\alpha}{(\beta_1+T_\alpha)\cdot T_0}.$$

Derivation of ξ_2. In the original model, P_4 inhibition of LH synthesis is scaled by the factor $\tilde{K}_{iL,P}(1+\beta_2 T)$, where $\beta_2 = c_{L,T} = 0.00959$. Similar to the derivation of ξ_1, we assume

$$K_{iL,P}\xi_2 = \tilde{K}_{iL,P}(1+\beta_2 T_\alpha),$$

so that

$$\tilde{K}_{iL,P} = \frac{K_{iL,P}}{1+\beta_2 T_0} \quad \text{and} \quad K_{iL,P}\xi_2 = K_{iL,P}\frac{1+\beta_2 T_\alpha}{1+\beta_2 T_0}.$$

Derivation of ξ_3. In the original model, follicle sensitivity to FSH has the form $h_1/[1+\beta_3 T/T_0]$, where $\beta_3 = c_{\Phi,T} = 0.19878$. We assume

$$h_1\xi_3 = \frac{\tilde{h}_1}{1+\beta_3 T_\alpha/T_0},$$

so that

$$\tilde{h}_1 = h_1(1+\beta_3),$$

which implies

$$h_1\xi_3 = h_1\frac{1+\beta_3}{1+\beta_3 T_\alpha/T_0}.$$

Appendix C. Empirical Distributions by Phenotype

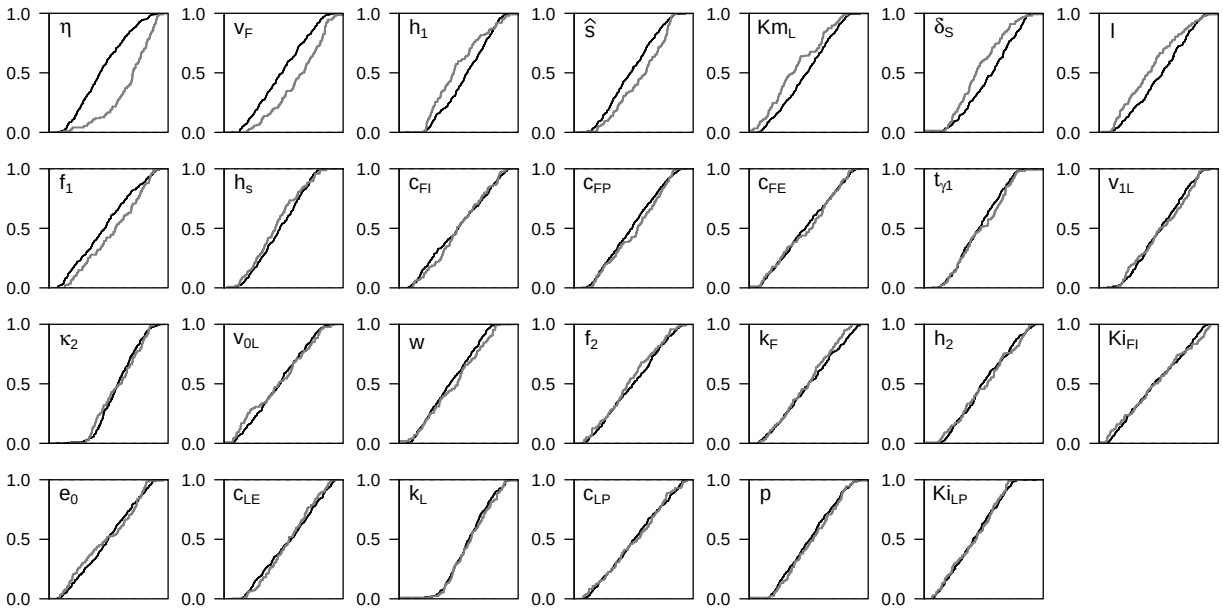

Figure Appendix C.1. Empirical cumulative distribution functions for reduced model parameters, separated by regular (black) and irregular (gray) phenotypes. Parameters are listed, beginning from the top row, in order of decreasing significance.