

Deconvolution of Serum Cortisol Levels by Using Compressed Sensing

Rose T. Faghih^{1,2,3*}, Munther A. Dahleh^{1,2}, Gail K. Adler^{5,6}, Elizabeth B. Klerman^{5,7}, Emery N. Brown^{3,4,5,8}

1 Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **2** Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **3** Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, Massachusetts, United States of America, **4** Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **5** Harvard Medical School, Boston, Massachusetts, United States of America, **6** Division of Endocrinology, Diabetes and Hypertension, Brigham and Women's Hospital, Boston, Massachusetts, United States of America, **7** Division of Sleep Medicine, Brigham and Women's Hospital, Boston, Massachusetts, United States of America, **8** Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America

Abstract

The pulsatile release of cortisol from the adrenal glands is controlled by a hierarchical system that involves corticotropin releasing hormone (CRH) from the hypothalamus, adrenocorticotropin hormone (ACTH) from the pituitary, and cortisol from the adrenal glands. Determining the number, timing, and amplitude of the cortisol secretory events and recovering the infusion and clearance rates from serial measurements of serum cortisol levels is a challenging problem. Despite many years of work on this problem, a complete satisfactory solution has been elusive. We formulate this question as a non-convex optimization problem, and solve it using a coordinate descent algorithm that has a principled combination of (i) compressed sensing for recovering the amplitude and timing of the secretory events, and (ii) generalized cross validation for choosing the regularization parameter. Using only the observed serum cortisol levels, we model cortisol secretion from the adrenal glands using a second-order linear differential equation with pulsatile inputs that represent cortisol pulses released in response to pulses of ACTH. Using our algorithm and the assumption that the number of pulses is between 15 to 22 pulses over 24 hours, we successfully deconvolve both simulated datasets and actual 24-hr serum cortisol datasets sampled every 10 minutes from 10 healthy women. Assuming a one-minute resolution for the secretory events, we obtain physiologically plausible timings and amplitudes of each cortisol secretory event with R^2 above 0.92. Identification of the amplitude and timing of pulsatile hormone release allows (i) quantifying of normal and abnormal secretion patterns towards the goal of understanding pathological neuroendocrine states, and (ii) potentially designing optimal approaches for treating hormonal disorders.

Citation: Faghih RT, Dahleh MA, Adler GK, Klerman EB, Brown EN (2014) Deconvolution of Serum Cortisol Levels by Using Compressed Sensing. PLoS ONE 9(1): e85204. doi:10.1371/journal.pone.0085204

Editor: Andrew Wolfe, John Hopkins University School of Medicine, United States of America

Received: September 11, 2013; **Accepted:** November 22, 2013; **Published:** January 28, 2014

Copyright: © 2014 Faghih et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: RTF's work was supported in part by the NSF Graduate Fellowship. For this work ENB is supported in part by NIH DP1 OD003646 and NSF 0836720. MAD is supported in part by EFRI-0735956 while EBK is supported in part by NIH P01-AG09975, K24-HL105664, RC2-HL101340, R01-HL-11408, and HFP01603 and HFP02802 from the National Space Biomedical Research Institute through NASA NCC 9-58. Finally, GKA is supported in part by NIH K24 HL103845. Clinical studies were performed on the Brigham and Women's Hospital General Clinical Research Center. Grants supporting the original data collection include NIH R01AR43130 (GKA), M01RR20635 (BWH GCRC), K01AG00661 (EBK), R01GM 53559 (ENB), and NASA NCC9-58 with the NSBRI. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: EBK received an unrestricted gift to the Brigham and Women's Hospital from Sony Corporation in 2011. The rest of the authors have declared that no competing interests exist. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: rfaghih@mit.edu

Introduction

Since many endocrine hormones such as cortisol are released in pulses instead of having slowly varying concentrations, understanding normal and abnormal endocrine functioning requires detection and quantification of both timing and amplitude of hormone pulses [1]. Determining the number, timing, and amplitude of hormone pulses is a challenging problem. For hormones that are released under control of the hypothalamus and anterior pituitary, release is regulated hierarchically; hypothalamic hormone release induces pituitary hormone release [2]. Then, either the pituitary hormone induces the release of another hormone from an endocrine gland (e.g. cortisol from the adrenal glands or thyroid from the thyroid gland) or the hormone of interest is directly released from the pituitary (e.g. growth

hormone). These hormones are absorbed from the blood stream, and implement regulatory functions throughout the body. Some hormones exert negative feedback on release of their hypothalamic and pituitary regulatory hormones, and consequently their further release [2].

Current data analysis methods for pulsatile hormone secretion either assume that the timing of the impulses belongs to a certain class of stochastic processes (e.g. birth-death process) [3] or use pulse detection algorithms [4]. The problem of recovering the model parameters as well as the number, timing, and amplitude of hormone pulses from a limited number of observations is ill-posed (i.e., there could be multiple solutions). However, by taking advantage of the sparse nature of hormone pulses and adding more constraints, the problem becomes more tractable. Veldhuis et al. [5] and Keenan et al. [6] have recently reviewed various

methods used for analyzing pulsatile hormone secretion. One method used for analyzing hormones is to assume point process models for the secretory events and employ methods such as the Markov Chain Monte Carlo (MCMC) algorithm [3]. For example, Johnson et al. find hormone pulses by embedding a birth-death process in an MCMC algorithm [3]. Keenan et al. propose an algorithm that uses a Bayesian approach to identify the pulses, and the secretion and clearance rates [6]. These methods assume that the occurrence of the secretory events is stochastic. In a recent work, Vidal et al. [4] use a pulse detection algorithm and remove peaks whose heights are small compared to the other detected pulses or some threshold. This pulse detection algorithm is implemented on luteinizing hormone (LH) data from ewes. In the LH data from ewes, one pulse decays significantly before the next pulse occurs. GH time series patterns also consist of hormone pulses that decay significantly before the next pulse occurs. However, one cortisol pulse can occur before the previous one has significantly decayed to near zero. Therefore, the extraction of the pattern of pulses may be less clear in cortisol than in LH and GH data, and analyzing cortisol data is more challenging.

As a first step in the study of novel methods of quantifying pulsatile activity of hormones, we investigate cortisol secretion in the hypothalamic-pituitary-adrenal (HPA) axis. Glucocorticoids, including cortisol, are crucial in neurogenesis, glucose homeostasis, metabolism, stress response, cognition, and response to inflammation [7]. Diseases that are linked to abnormalities in the HPA axis include diabetes, visceral obesity and osteoporosis, life-threatening adrenal crises and disturbed memory formation [8], [9]. Cortisol secretion is initiated by the release of corticotropin releasing hormone (CRH) from the hypothalamus; CRH induces pulsatile release of adrenocorticotropin hormone (ACTH) from the anterior pituitary, and through their stimulation by ACTH, the adrenal glands produce cortisol. Cortisol exerts negative feedback effect on the release of CRH and ACTH, and consequently future cortisol release. Normal physiology includes variation of the cortisol level over a 24-hour period; these variations are due to the circadian modulation of the amplitude of the secretory events and ultradian modulation of the timing of the secretory events. Thus, there is a complex physiology required to produce the cortisol pulses and in particular this physiology depends critically on the release of CRH from the hypothalamus and pulsatile release of ACTH from the anterior pituitary. In this study, we are only concerned with the problem of estimating the pharmacokinetics properties of cortisol as well as the number, timing and amplitude of the cortisol pulses from the time-series of the serum cortisol levels. Over 24 hours, there are 15 to 22 secretory events [10,11] with an average of 18 secretory events in both males and females [12]. Hence, since there are a small number of secretory events that are significant, these hormone pulses can be considered sparse and specific analytic techniques can be applied.

In this paper, we use the characteristic of the sparsity of hormone pulses (i.e., there are a small number of secretory events that are important) and recover the timing and amplitude of individual hormone pulses using compressed sensing techniques. Compressed sensing is a technique for perfect reconstruction of sparse and compressible signals using fewer measurements than required by the Shannon/Nyquist sampling theorem [13]. For compressible signals, where only a small number of coefficients are large (i.e., most coefficients are small or zero) and small coefficients can be discarded, the signal can be approximated by a sparse representation and recovered using optimization or greedy algorithms [13]. We describe a coordinate descent approach to recover cortisol secretory events and model parameters. To demonstrate the performance of the algorithm, we apply it to

cortisol data. Although we know the sparsity range of the input, the number of pulses for each participant is an open question. In finding the number of pulses, there is a trade-off between capturing the residual error and the sparsity. We use generalized cross-validation to find the number of pulses such that there is a balance between the residual error and the sparsity. This algorithm potentially can be applied to other pulsatile endocrine hormones such as GH, thyroid hormone, LH, and testosterone.

Methods

Experiment

Blood from 10 healthy women collected every 10 minutes for 24 hours was assayed for cortisol in duplicate. The 24 hours began with 8 hours of scheduled sleep followed by 16 hours of wake. The participants were recruited via advertisements in local newspapers to serve as healthy controls for a study on women with fibromyalgia; participants with abnormal laboratory test results or current medical problems were excluded [14]. None of the participants had received glucocorticoids or estrogen/progesterone within the year or 4 months before the study, respectively [14]. For 3 consecutive nights, at the General Clinical Research Center of the Brigham and Womens Hospital, participants had 8 hours of scheduled sleep in the dark at their habitual sleep-wake times and three meals and two snacks; on day 4, all participants started a constant routine protocol that included continuous wake, constant posture, and small meals given hourly [14]. Blood drawing for hormones analyzed in this study was performed during the third night of sleep and during the first 16 hours of the constant routine. A detailed description of the experiment is in [14]; clinical characteristics of the participants are given in Table 1. The experimental protocol was designed to minimize the effects of ambulatory temperature, activity, eating, and stress on the participants. Therefore, this dataset can be used to quantify cortisol variations as a function of the time of the day, as controlled by the circadian and the ultradian patterns.

Ethics Statement

This project has been reviewed and approved by the Brigham and Women's Hospital Institutional Review Board (IRB). During the review of this project, the IRB specifically considered (i) the risks and anticipated benefits, if any, to subjects; (ii) the selection of

Table 1. Clinical Characteristics of the Participants.

AGE	BMI ($\frac{kg}{m^2}$)	Systolic BP (mmHg)	Diastolic BP (mmHg)
28	20.8	100	70
41	23.6	120	80
41	22.5	130	70
24	20.7	108	78
23	27.4	108	68
26	25.2	108	68
23	29.3	110	74
37	29.6	138	80
44	29.9	135	82
42	22.9	108	62

BMI and BP refer to Body Mass Index and Blood Pressure, respectively. None of the participants had a current diagnosis of depression
doi:10.1371/journal.pone.0085204.t001

subjects; (iii) the procedures for securing and documenting informed consent; (iv) the safety of subjects; and (v) the privacy of subjects and confidentiality of the data. All subjects provided written informed consent that was approved by the ethics committees.

Model Formulation

We build a model based on the stochastic differential equation model of diurnal cortisol patterns [10]. This model is based on the first-order kinetics for cortisol synthesis in the adrenal glands, cortisol infusion to the blood, and cortisol clearance by the liver, while considering a doubly stochastic pulsatile input in the adrenal glands that has Gaussian amplitudes and gamma distributed interarrival times [10]. This input can be considered as an abstraction of hormone pulses and marks the timing and amplitude of the secretory events leading to cortisol secretion. We assume that there are between 15 and 22 secretory events over a 24-hour period that result in the observed cortisol profile [10,11,15,16]. This model is represented as follows:

$$\frac{dx_1(t)}{dt} = -\theta_1 x_1(t) + u(t) \text{ (Adrenal Glands)} \tag{1}$$

$$\frac{dx_2(t)}{dt} = \theta_1 x_1(t) - \theta_2 x_2(t) \text{ (Serum)} \tag{2}$$

where x_1 is the cortisol concentration in the adrenal glands and x_2 is the serum cortisol concentration. θ_1 and θ_2 , respectively, represent the infusion rate of cortisol from the adrenal glands into the blood and the clearance rate of cortisol by the liver. $u(t) = \sum_{i=1}^N q_i \delta(t - \tau_i)$ is an abstraction of the hormone pulses that result in cortisol secretion where q_i represents the amount of the hormone pulse initiated at time τ_i (q_i is zero if a hormone pulse did not occur at time τ_i), and we assume that impulses occur at integer minute values. N corresponds to the length of the input ($N = 1440$).

Table 2. The Estimated Model Parameters and the Squares of the Multiple Correlation Coefficients (R^2) for the Fits of the Experimental Cortisol Time Series.

Participant	$\theta_1(\text{min}^{-1})$	$\theta_2(\text{min}^{-1})$	N	R^2
1	0.0739	0.0067	18	0.97
2	0.0762	0.0057	17	0.93
3	0.0921	0.0082	16	0.96
4	0.1248	0.0061	17	0.93
5	0.0585	0.0122	18	0.95
6	0.0726	0.0095	20	0.96
7	0.0799	0.0107	16	0.97
8	0.0365	0.0091	16	0.93
9	0.0361	0.0090	16	0.92
10	0.0864	0.0073	20	0.94
Median	0.0751	0.0086	17	0.95

The parameters θ_1 and θ_2 are, respectively, the estimated infusion rate of cortisol into the circulation from the adrenal glands and the estimated clearance rate of cortisol by the liver. N is the estimated number of hormone pulses and R^2 is the square of the multiple correlation coefficient.
doi:10.1371/journal.pone.0085204.t002

Blood was collected, beginning at y_0 and then, with a sampling interval of 10 minutes, for M samples ($M = 144$). All samples were assayed for cortisol. Let $y_{t_{10}}, y_{t_{20}}, \dots, y_{t_{10M}}$

$$y_{t_k} = x_2(t_k) + v_{t_k} \tag{3}$$

where y_{t_k} and v_{t_k} represent the observed serum cortisol level and the measurement error, respectively, and missing data points can be interpolated. For each participant, blood samples were assayed in duplicate; hence, for each participant, we could obtain the standard deviation of noise and model the corresponding measurement error. A Gaussian density is a reasonably good approximation of the probability density of the immunoassay error [10], and, by using a least squares approach in our estimation algorithm, we model the noise as a Gaussian random variable. Using the serum cortisol level (x_2) with a sampling interval of 10 minutes, we would like to estimate θ_1 and θ_2 , and obtain the number of pulses, their timing, and their amplitude.

Considering the known physiology of *de novo* cortisol synthesis (i.e., no cortisol is stored in the adrenal glands) [10], we assume that the initial condition of the cortisol level in the adrenal glands is zero ($x_1(0) = 0$) [3], and solve for y_{t_k} . Considering that we have discrete data points sampled every 10 minutes, assuming that the input occurs at integer minutes, and y_0 is the initial condition of the serum cortisol concentration, every observed data point y_{t_k} can be represented as follows:

$$y_{t_k} = a_{t_k} y_0 + b'_{t_k} u + v_{t_k} \tag{4}$$

where

$$b_{t_k} = \left[\frac{\theta_1}{\theta_1 - \theta_2} (e^{-\theta_2 k} - e^{-\theta_1 k}) \quad \frac{\theta_1}{\theta_1 - \theta_2} (e^{-\theta_2(k-1)} - e^{-\theta_1(k-1)}) \dots \frac{\theta_1}{\theta_1 - \theta_2} (e^{-\theta_2} - e^{-\theta_1}) \underbrace{0 \dots 0}_{N-k} \right]' a_{t_k} = e^{-\theta_2 k}, \text{ and } u \text{ represents the entire input over 24 hours (elements of } u \text{ take values } q_i \text{ for } i = 1, \dots, 1440).$$

Let $y = [y_{t_{10}} \quad y_{t_{20}} \quad \dots \quad y_{t_{10M}}]'$, $\theta = [\theta_1 \quad \theta_2]'$, $A_\theta = [a_{t_{10}} \quad a_{t_{20}} \quad \dots \quad a_{t_{10M}}]'$, $B_\theta = [b_{t_{10}} \quad b_{t_{20}} \quad \dots \quad b_{t_{10M}}]'$, and $v = [v_{t_{10}} \quad v_{t_{20}} \quad \dots \quad v_{t_{10M}}]'$.

We can represent this system as:

$$y = A_\theta y_0 + B_\theta u + v. \tag{5}$$

Estimation

In modeling cortisol secretion over 24 hours, results from previous studies suggest that there are 15 to 22 secretory events [10,11,15,16], and on average, there are 18 secretory events [12]. Hence, we assume u contains 15 to 22 nonzero elements out of 1440 possibilities and all these nonzero elements are nonnegative ($15 \leq \|u\|_0 \leq 22, u \geq 0$). To estimate the model parameters, we follow [10] and assume that the infusion rate of cortisol from the adrenal glands to the circulation is at least four times the clearance rate of cortisol by the liver (i.e., $4\theta_2 \leq \theta_1$). Because the hormone infusion and clearance rates cannot be negative, we further assume in the optimization algorithm that $\theta \geq 0$. We can formulate this problem as an optimization problem:

$$\min \|y - A_\theta y_0 - B_\theta u\|_2^2 \tag{6}$$

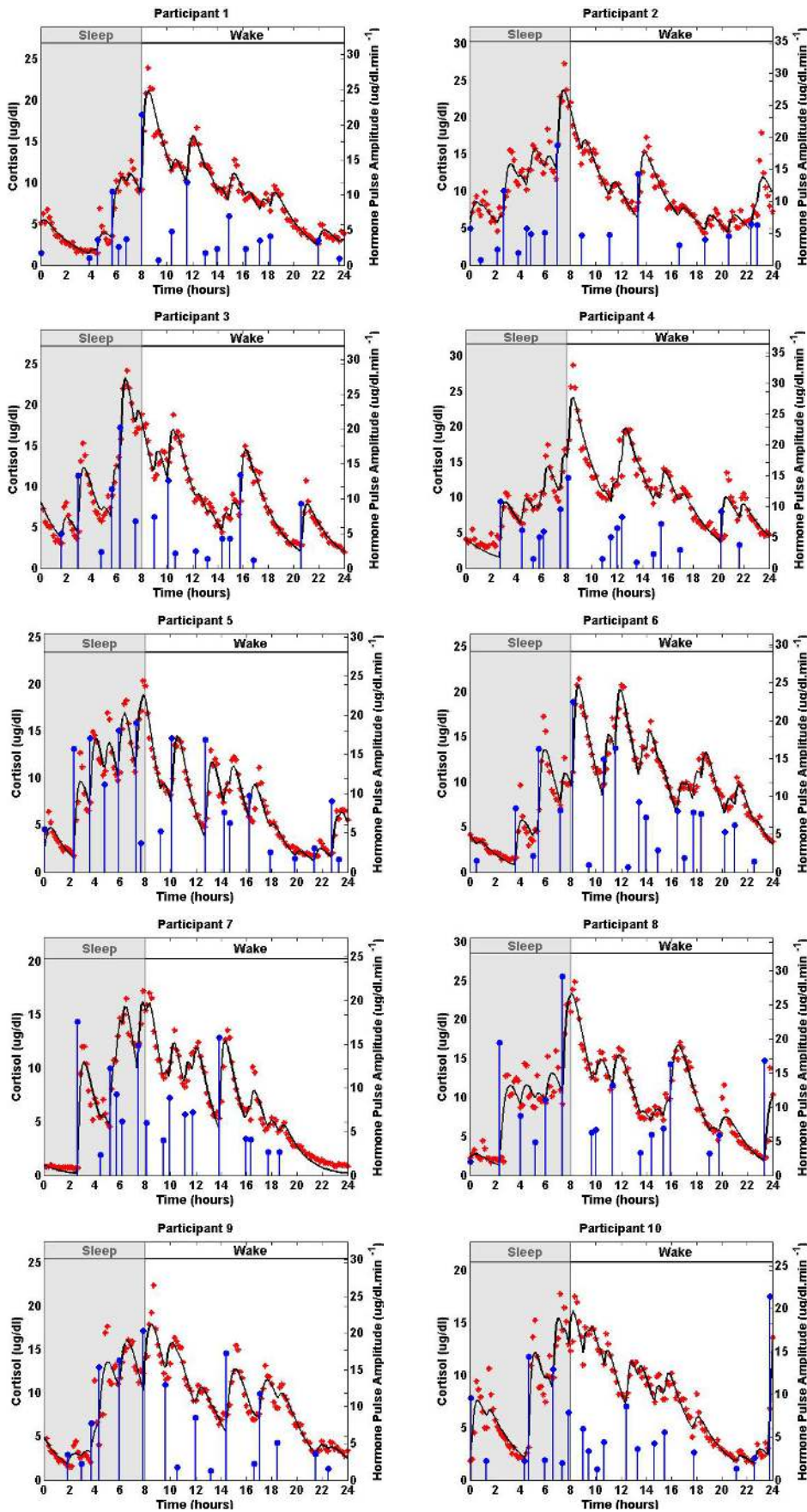


Figure 1. Estimated Deconvolution of the Experimental Twenty-Four-Hour Cortisol Levels in 10 Women. Each panel shows the measured 24-hour cortisol time series (red stars), the estimated cortisol levels (black curve), the estimated pulse timing and amplitudes (blue vertical lines with dots) for one of the participants. The estimated model parameters are given in Table 2.
doi:10.1371/journal.pone.0085204.g001

$$\begin{aligned} \text{s.t.} \\ 15 \leq \|u\|_0 \leq 22 \\ u \geq 0 \\ C\theta \leq b \end{aligned}$$

where $C = \begin{bmatrix} -1 & -1 & 0 \\ 4 & 0 & -1 \end{bmatrix}'$, $b = [0 \ 0 \ 0]'$.

This optimization problem is generally considered as NP-hard. It is possible to use an ℓ_1 -norm relaxation and solve this problem using different computational strategies such as the basis pursuit, greedy algorithms, iterative-thresholding algorithms, or the FOCUSS algorithm and its extensions [17]. We solve this problem with an extension of the FOCUSS algorithm by first casting this optimization problem as:

$$\min_{\substack{u \geq 0 \\ C\theta \leq b}} J_\lambda(\theta, u) = \frac{1}{2} \|y - A_\theta y_0 - B_\theta u\|_2^2 + \lambda \|u\|_p^p \quad (7)$$

where the ℓ_p -norm is an approximation to the ℓ_0 -norm ($0 < p \leq 2$) and λ is chosen such that the sparsity of u is between 15 to 22. Then, by using a coordinate descent approach, this optimization problem can be solved iteratively through the following steps (for $l=0,1,2,\dots$) until convergence is achieved:

1.

$$u^{(l+1)} = \underset{u \geq 0}{\operatorname{argmin}} J_\lambda(\theta^{(l)}, u) \quad (8)$$

2.

$$\theta^{(l+1)} = \underset{C\theta \leq b}{\operatorname{argmin}} J_\lambda(\theta, u^{(l+1)}) \quad (9)$$

The optimization problem in (8) now can be solved using the FOCUSS algorithm, which uses a re-weighted norm minimization approach. The solution at each iteration is found by minimizing the ℓ_2 -norm and the iteration refines the initial estimate to the final localized energy solution [18]. In the FOCUSS algorithm, assuming that a gradient factorization exists, the stationary points of (8) satisfy $u = P_u B_\theta^T (B_\theta P_u B_\theta^T + \lambda I)^{-1} y_\theta$ [18], where $P_u = \operatorname{diag}(|u_i|^{2-p})$, and $y_\theta = y - A_\theta y_0$. By iteratively updating λ and u until convergence, we can solve for the sparse vector u . In the optimization problem in (8), λ balances between the sparsity of u and the residual error $\|y_\theta - B_\theta u\|_2$. The sparsity of u increases with λ .

A version of the FOCUSS algorithm called FOCUSS+ proposed by Murray [19] allows for solving for u such that the maximum sparsity of u is n ($n=22$ for our current problem) and u is nonnegative. This algorithm uses a heuristic approach for updating λ , which tunes the trade-off between the sparsity and the residual error by increasing λ to a maximum regularization λ_{max} as the residual error decreases. FOCUSS+ works as follows:

1. $P_u^{(r)} = \operatorname{diag}(|u_i^{(r)}|^{2-p})$
2. $\lambda^{(r)} = \left(1 - \frac{\|y_\theta - B_\theta u^{(r)}\|_2}{\|y_\theta\|_2} \right) \lambda_{max}, \lambda > 0$

Table 3. The Estimated Model Parameters and the Squares of the Multiple Correlation Coefficients (R^2) for the Fits of the Simulated Cortisol Time Series.

Participant	$\hat{\theta}_1$ (min ⁻¹)	$\hat{\theta}_2$ (min ⁻¹)	N	R^2	σ_n ($\frac{\text{ug}}{\text{dl}}$)	$\frac{ \theta_1 - \hat{\theta}_1 }{\theta_1}$ (%)	$\frac{ \theta_2 - \hat{\theta}_2 }{\theta_2}$ (%)
1	0.0628	0.0068	16	0.99	0.38	15.02%	1.49%
2	0.0569	0.0056	15	0.97	0.75	25.33%	1.75%
3	0.0747	0.0078	15	0.98	0.67	18.89%	4.88%
4	0.0918	0.0071	16	0.94	1.44	26.44%	16.39%
5	0.0492	0.0123	18	0.99	0.52	15.90%	0.82%
6	0.0490	0.0122	19	0.99	0.29	32.51%	28.42%
7	0.0770	0.0098	16	0.97	0.98	3.63%	8.41%
8	0.0375	0.0094	18	0.99	0.33	2.74%	3.30%
9	0.0365	0.0091	16	0.99	0.35	1.11%	1.11%
10	0.0654	0.0076	19	0.99	0.31	24.30%	4.11%
Median	0.0599	0.0084	16	0.99	0.45	17.39%	3.70%

The parameters $\hat{\theta}_1$ and $\hat{\theta}_2$ are, respectively, the estimated infusion rate of cortisol into the circulation from the adrenal glands and the estimated clearance rate of cortisol by the liver. N is the estimated number of hormone pulses and R^2 is the square of the multiple correlation coefficient. σ_n is the standard deviation of the zero mean Gaussian noise added to each simulated data point. For each participant, blood samples were assayed in duplicate, and the corresponding standard deviation of noise was obtained. The parameters θ_1 and θ_2 are, respectively, the infusion rate of cortisol into the circulation from the adrenal glands and the clearance rate of cortisol by the liver used in simulating each dataset. The values of θ_1 and θ_2 are given in Table 2.

doi:10.1371/journal.pone.0085204.t003

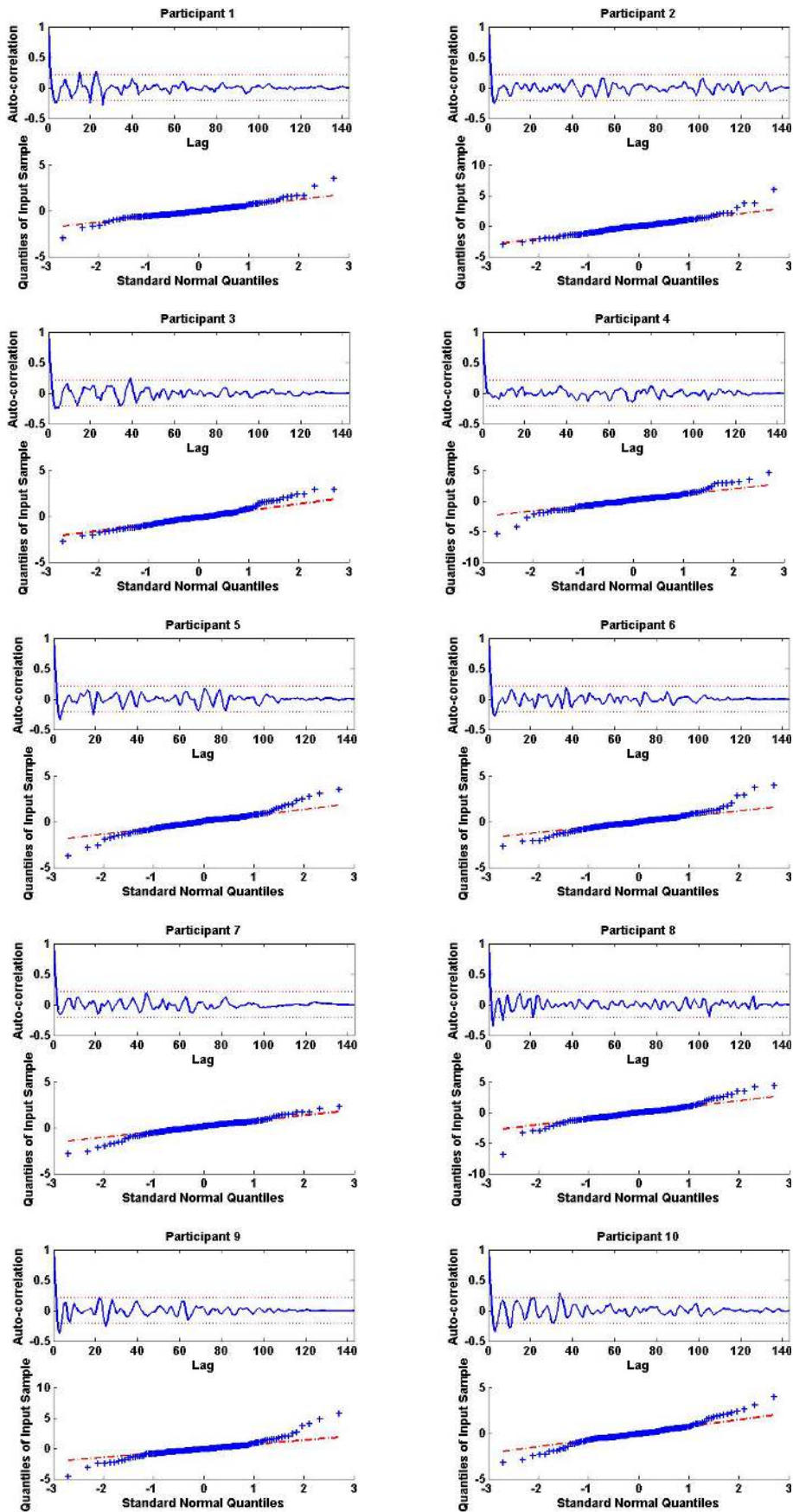


Figure 2. White Gaussian Structure in the Model Residuals of 10 Women. In each panel, (i) the top sub-panel displays the autocorrelation function of the model residuals in one of the 10 participants; the graph shows that the model captures the dynamics and that residuals are white; (ii) the bottom sub-panel displays the quantile-quantile plot of the model residuals for that participant; the graph shows that the residuals are Gaussian. doi:10.1371/journal.pone.0085204.g002

3. $\mathbf{u}^{(r+1)} = \mathbf{P}_u^{(r)} \mathbf{B}_\theta^T (\mathbf{B}_\theta \mathbf{P}_u^{(r)} \mathbf{B}_\theta^T + \lambda^{(r)} \mathbf{I})^{-1} \mathbf{y}_0$
4. $\mathbf{u}_i^{(r+1)} \leq 0 \rightarrow \mathbf{u}_i^{(r+1)} = 0$
5. After more than half of the selected number of iterations, if $\|\mathbf{u}^{(r+1)}\|_0 > n$, select the largest n elements of $\mathbf{u}^{(r+1)}$ and set the rest to zero.
6. Iterate

FOCUSS+ usually converges in 10 to 50 iterations [19]. Although an estimate of the unknown quantities can be obtained after iteratively solving for \mathbf{u} and θ (8–9) using FOCUSS+, θ and \mathbf{u} should be updated by finding an optimal choice of λ such that enough noise is filtered out and the estimated \mathbf{u} is not capturing residual error by finding a less sparse solution. We use the Generalized Cross-Validation (GCV) technique [20] for estimating the regularization parameter such that there is a balance in filtering out the noise and the sparsity of \mathbf{u} . The GCV function is defined as:

$$G(\lambda) = \frac{L\|(I - H_\lambda)\mathbf{y}_0\|^2}{(\text{trace}(I - H_\lambda))^2}, \quad (10)$$

where L is the number of data points, and H_λ is the influence matrix. For the FOCUSS algorithm, $H_\lambda = \mathbf{B}_\theta \mathbf{P}_u \mathbf{B}_\theta^T (\mathbf{B}_\theta \mathbf{P}_u \mathbf{B}_\theta^T + \lambda \mathbf{I})^{-1}$. Zhaoshui et al. [17] employ the GCV technique for estimating the regularization parameter for the FOCUSS algorithm through singular value decomposition:

$$G(\lambda) = \frac{L \sum_{i=1}^L \zeta_i^2 \left(\frac{\lambda}{\sigma_i^2 + \lambda} \right)^2}{\left(\sum_{i=1}^L \frac{\lambda}{\sigma_i^2 + \lambda} \right)^2}, \quad (11)$$

where $\zeta = \mathbf{R}^T \mathbf{y}_0 = [\zeta_1 \ \zeta_2 \ \dots \ \zeta_L]^T$ and $\mathbf{B}_\theta \mathbf{P}_u^{\frac{1}{2}} = \mathbf{R} \Sigma \mathbf{Q}^T$ with $\Sigma = \text{diag}\{\sigma_i\}$; \mathbf{R} and \mathbf{Q} are unitary matrices and σ_i 's are the singular values of $\mathbf{B}_\theta \mathbf{P}_u^{\frac{1}{2}}$. [17]. Zhaoshui et al. [17] propose minimizing $G(\lambda)$ such that λ is bounded between some minimum and maximum values (λ_{\min} and λ_{\max}) using the MATLAB function `f_minbnd` (MATLAB R2011b), which is an implementation of the golden section (GS) search. Although the GS search only finds a local extremum, considering that $G(\lambda)$ is unimodal, the GS search always finds the desired solution given a large range for λ [17]. We used a range of zero to 10 for λ .

Considering the ill-posedness of deconvolution problems, small variations in the data can result in large changes in the solution, and a balanced choice of regularization is required to filter out the effect of noise. Tikhonov regularization, truncated singular value decomposition, and the method of L-curve are well-known methods used when dealing with such problems [21]; among these methods, the L-curve method appears to be the most commonly used. Automatically searching for the minimum of the GCV function is easier than finding the corner of the L-curve as the L-curve method is computationally expensive, requiring computation of the solution for several samples of the parameter [17]. Moreover, Zhaoshui et al. point out that GCV is usually more accurate in estimating the regularization parameter than the L-curve method [17]. In the GCV technique, the optimal choice of regularization minimizes the predictive mean-squared error.

Hence, one can update \mathbf{u} by modifying FOCUSS+ using the GCV method. For $r=0,1,2,\dots$ GCV - FOCUSS+ works as follows:

1. $\mathbf{P}_u^{(r)} = \text{diag}(|\mathbf{u}_i^{(r)}|^{2-p})$
2. $\mathbf{u}^{(r+1)} = \mathbf{P}_u^{(r)} \mathbf{B}_\theta^T (\mathbf{B}_\theta \mathbf{P}_u^{(r)} \mathbf{B}_\theta^T + \lambda^{(r)} \mathbf{I})^{-1} \mathbf{y}_0$
3. $\mathbf{u}_i^{(r+1)} \leq 0 \rightarrow \mathbf{u}_i^{(r+1)} = 0$
4. $\lambda^{(r+1)} = \underset{0 \leq \lambda \leq 10}{\text{argmin}} G(\lambda)$
5. Iterate until convergence

By combining the GCV method with FOCUSS+, one can find an optimal choice of λ at each iteration such that enough noise is filtered out when solving for \mathbf{u} , and iterate between solving (8) and (9) until convergence is achieved. The following is the algorithm that we propose for deconvolution of cortisol data:

1. Initialize $\tilde{\theta}^0$ by sampling a uniform random variable w on $[0,1]$, and let $\tilde{\theta}^0 = [w \ \frac{w}{4}]^T$ and $j=1$
2. Set $\tilde{\theta}$ equal to $\tilde{\theta}^{j-1}$; using FOCUSS+, solve for $\tilde{\mathbf{u}}^j$ by initializing the optimization problem in (8) at a vector of all ones
3. Set $\tilde{\mathbf{u}}$ equal to $\tilde{\mathbf{u}}^j$; using the interior point method, solve for $\tilde{\theta}^j$ by initializing the optimization problem in (9) at $\tilde{\theta}^{j-1}$
4. Iterate between steps 2–3 for 30 iterations
5. Initialize $\hat{\theta}^0$ and $\hat{\mathbf{u}}^0$ by setting them equal to the $\tilde{\theta}^j$ and $\tilde{\mathbf{u}}^j$ that minimize $J_\lambda(\theta, \mathbf{u})$ in (7), and let $i=1$
6. Set $\hat{\theta}$ equal to $\hat{\theta}^{i-1}$; using GCV - FOCUSS+, solve for $\hat{\mathbf{u}}^i$ by initializing the optimization problem in (8) at $\hat{\mathbf{u}}^{i-1}$
7. Set $\hat{\mathbf{u}}$ equal to $\hat{\mathbf{u}}^i$; using the interior point method, solve for $\hat{\theta}^i$ by initializing the optimization problem in (9) at $\hat{\theta}^{i-1}$
8. Iterate between steps 6–7 until convergence
9. Repeat steps 1–8 for various initializations
10. Set the estimated model parameters $\hat{\theta}$ and input $\hat{\mathbf{u}}$ equal to the values that minimize $J_\lambda(\theta, \mathbf{u})$ in (7). Since this optimization problem is non-convex, there are multiple local minima, and a reasonable procedure to choose among the local minima is to select the one with the best goodness of fit.

Step 1 initializes the algorithm randomly. Steps 2–4 use the random initialization to find a good initialization for the unknowns while using FOCUSS+ for sparse recovery and interior point method for finding the model parameters. Step 5 finds a good initial condition by comparing the estimates obtained in Steps 2–4, and selecting the θ and \mathbf{u} values that minimize the cost function. The values found at Step 5 are used for initializing the main algorithm. Steps 6–8 use a coordinate descent approach to estimate the unknowns until convergence is achieved. Sparse recovery is achieved using GCV - FOCUSS+ which uses generalized cross-validation for finding the regularization parameter; GCV - FOCUSS+ selects the regularization parameter such that there is a balance between capturing the sparsity and the noise, and finds different sparsity levels for different individuals. Step 9 repeats the initialization and estimation steps for various

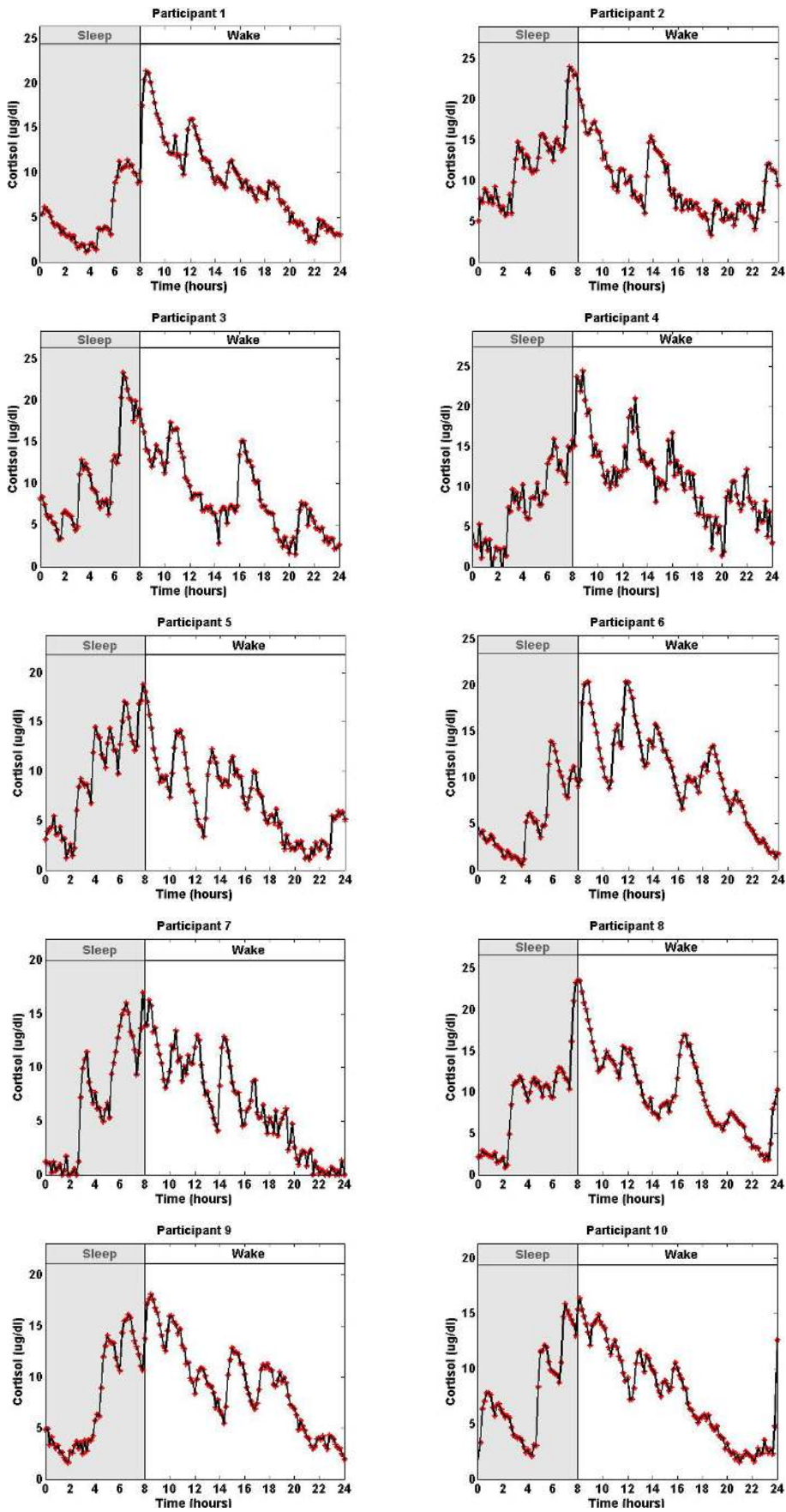


Figure 3. Simulated Twenty-Four-Hour Cortisol Levels with Measurement Errors Corresponding to Datasets from 10 Women. Each panel displays the simulated serum cortisol levels based on pulse patterns in Figure 1 and estimated model parameters θ_1 and θ_2 in Table 2 in one of the 10 participants, assuming a zero mean Gaussian measurement error with standard deviation σ_n in Table 3. In all simulations the initial conditions are $x_1(0)=0$, $x_2(0)$ equals the initial cortisol level of the corresponding participant, and the cortisol levels are recorded every 10 minutes. doi:10.1371/journal.pone.0085204.g003

initializations. Step 10 selects the θ and u values that minimize the cost function.

The main idea behind the algorithm is to solve the non-convex problem in (7) using a coordinate descent approach to converge to a local minimum for different initializations, and choose the local minimum that minimizes the problem in (7). Convergence properties of coordinate descent algorithms are well-studied and a discussion can be found in [22].

We have implemented the algorithm by assuming that hormone pulses occur at integer minutes. We ran the proposed algorithm using 10 random initializations for each dataset. According to [23], the FOCUSS algorithm converges faster for $0 < \rho < 1$ compared to $1 < \rho < 2$; however, it should be noted that for $0 < \rho < 1$, the optimization problem is not convex and if ρ is too small (e.g., $\rho = 0.1$), it is possible to stagnate into a local minimum; hence, they suggest selecting a value slightly smaller than 1 but not too small [23]. When running GCV - FOCUSS+, we let $\rho = 0.5$ to solve for u . Data analysis, estimation, and simulations were performed in MATLAB R2011b.

Using the cortisol secretion model (1–2), we simulated ten 24-hour cortisol datasets using parameters θ_1 and θ_2 in Table 2 and impulse trains in Figure 1. Then, using the observation model (3), we added zero mean Gaussian noise with a standard deviation of σ_n to the simulated cortisol levels; σ_n in Table 3 models the immunoassay error for each participant and was obtained using blood samples assayed in duplicate.

Results

Figure 1 shows experimental data from each participant. For most participants, the cortisol level is low at the beginning of the scheduled sleep; the cortisol level increases rapidly around the wake time and gradually decreases throughout the day.

Figure 1 shows the estimated amplitude and timing of hormone pulses, experimental cortisol data, and model-predicted cortisol estimates for each participant. There are variations in the timing and amplitudes of the detected hormone pulses. The circadian amplitudes of the recovered pulses demonstrate the known circadian variation of cortisol time series [10]; for most participants, the recovered pulses are small at the beginning of the scheduled sleep, and there is a large pulse towards the end of the sleep period or beginning of the wake period. There are multiple small and medium sized pulses during the wake period. The number of detected pulses for all participants are within their corresponding physiologically plausible ranges [10,11,15,16] with a square of the multiple correlation coefficient (R^2) above 0.92 (Table 2). The R^2 is a statistical measure of the goodness of fit obtained by model-predicted estimates of data; it measures the fraction of the sample variance of data that is predicted by the model: R^2 values close to 1.0 suggest that the model is good at estimating the data. Figure 2 shows the autocorrelation function and the quantile-quantile plots of the model residuals for the 10 participants suggesting that the model captures the dynamics, and that the residuals have a Gaussian structure and are white.

We simulated 10 datasets, each corresponding to one of the 10 experimental datasets (Figure 3). These datasets were simulated by using the recovered pulses of the 10 experimental datasets shown in Figure 1 as well as the estimated model parameters for each of

the corresponding datasets as shown in Table 2. Then, we added zero mean Gaussian noise with standard deviations shown in Table 3 to the simulated data. These 10 simulated datasets were sampled every 10 minutes.

Table 3 shows the estimated model parameters, number of pulses, the square of the multiple correlation coefficient (R^2), the percentage error in estimating the model parameters, and the standard deviation of zero mean Gaussian noise used in simulating the 10 datasets. These simulations were performed by adding various values of zero mean Gaussian noise with standard deviations σ_n ranging from 0.29 to 1.44; the level of noise was based on the signal-to-noise ratio for data collection for each of the participants. Errors in estimating θ_1 and θ_2 range from small values (1.11% and 0.82%) to high values (32.51% and 28.42%), respectively. The maximum error in finding the number of pulses is two. There is zero error in finding the number of pulses for participant 7 where the maximum error in the detected support is 14 minutes. The overall performance of the estimation for the simulated dataset is best for the data that corresponds to participant 7 with $\sigma_n = 0.98$, and 3.63% and 8.41% error in estimating θ_1 and θ_2 , respectively. These estimates were obtained by 10 random initializations, and considering that the optimization problem solved for this estimation is non-convex, there are multiple local minima and the estimation can be improved by using more initializations. Since the noise added to the simulated data is comparable in amplitude to small pulses of cortisol, the pulsatile patterns of the simulated cortisol time series differ from their corresponding experimental time series. This explains the higher error in the estimated model parameters θ_1 and θ_2 , and the error in the estimated timing and amplitudes of pulses. This choice of noise is based on the immunoassay error for each experimental time series, so that the simulations are based on multiple metrics of the experimental data. The algorithm performs better for lower levels of noise that do not significantly affect the pulsatile patterns of the simulated time series.

Figure 4 shows the actual sparse input, the recovered input, the simulated cortisol data, and the estimated cortisol levels. Table 4 shows the maximum error in the detected timing of pulses, the number of small pulses that are not detected, and the number of extra pulses that are detected for each simulated dataset. The recovered and the actual inputs are in good agreement for a variety of noise levels when detecting significant pulses; however, a few of the small pulses cannot be detected for some cases or in some cases noise is captured as a small pulse. The maximum error in detecting the support is 26 minutes.

Discussion

Understanding the cortisol secretion process and modeling the underlying system is a challenging problem for several reasons. (I) Due to the simultaneous release and clearance of hormones and the unknown timing and amplitudes of the secretory events, identifying the pulsatile input to the system and the infusion and clearance rates is challenging. (II) Due to data collection difficulties and cost, the sampling interval is usually relatively large (10–60 minutes) compared to the expected inter-pulse intervals as well as secretion and clearance rates of cortisol. This low resolution of the data makes identifying the delays in the system and potential

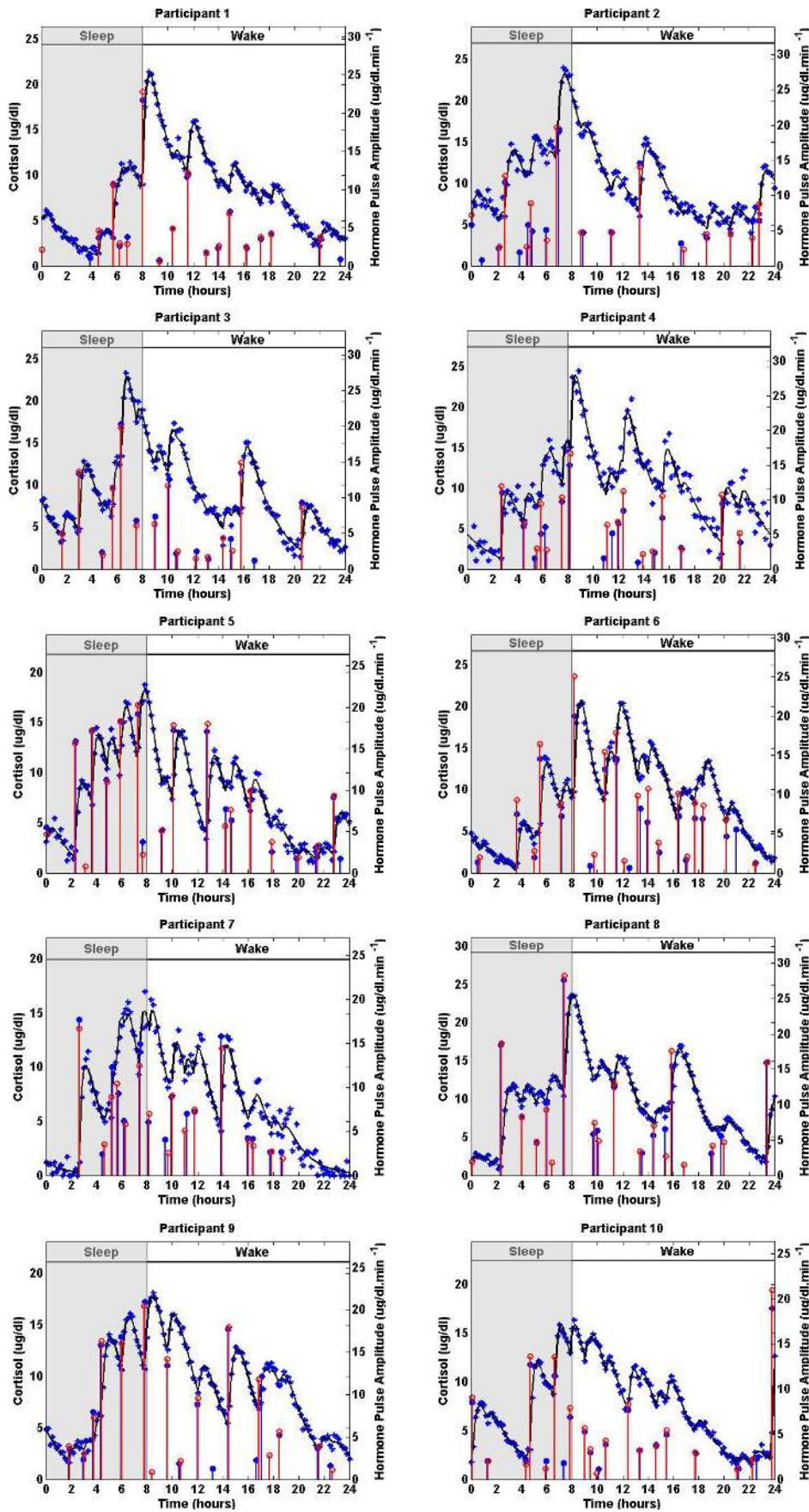


Figure 4. Estimated Deconvolution of Simulated Twenty-Four-Hour Cortisol Levels with Different Measurement Errors Corresponding to Datasets from 10 Women. Each panel shows the simulated 24-hour cortisol time series (blue stars), the estimated cortisol levels (black curve), the simulated pulse timing and amplitudes (blue vertical lines with dots) and the estimated pulse timing and amplitudes (red vertical lines with empty circles) for one of the simulated datasets that each correspond to a participant. The estimated parameters are given in Table 3.

doi:10.1371/journal.pone.0085204.g004

consecutive pulses that occur over one sampling interval challenging. (III) Cortisol secretion differs in sleep and wake states and at different circadian times; therefore, the model parameters might be time-varying. (IV) There is inter-individual variation, even among healthy individuals. (V) The properties of noise in the system are not known.

In this paper, we modeled secretory events that result in cortisol time series, and proposed a coordinate descent approach to estimate the model parameters and recover the sparse time-varying secretory input. Considering the sparsity of the input, we recovered the impulses using compressed sensing. While a range for the sparsity of the hormone pulses is known, the exact number of pulses varies from one individual to another and is unknown. To recover the accurate number of hormone pulses, the regularized problem should be solved such that there is a balance between capturing the sparsity and the residual error. We used generalized cross-validation for choosing the regularization parameter and finding the number of pulses for each individual. The algorithm described in this paper provides a general framework that can also be implemented on other hormones. For the case of cortisol data, the high R^2 values (found to be greater than 0.92 for all 10 participants) suggest that our proposed algorithm can successfully uncover physiologically plausible hormone pulse information underlying cortisol secretion. There are variations in the timing and amplitudes of cortisol secretory events. The amplitude variations throughout the day occur as a result of the circadian rhythm underlying cortisol release; the variations in the timing of impulses reflect the ultradian rhythm underlying cortisol release. Our algorithm makes it possible to capture the circadian and ultradian features of hormone pulses as well as the parameters underlying the first-order kinetics of cortisol release. Furthermore, our algorithm recovered the impulse train input from simulated data with various noise levels and detected the significant pulses; however, depending on the dataset, it misses one or two of the

insignificant pulses or captures noise as one or two small pulses. Our approach can be applied to GH, thyroid hormone, and gonadal hormones in a similar fashion. Future directions for this research include modeling simultaneous measurements of ACTH and cortisol data. In addition, as we postulate that the human body controls hormone secretion by solving an optimization problem, we can use these methods to begin to understand the underlying physiology. Correlation analysis of the residuals suggests that the model captures the dynamics, and residuals are Gaussian and white.

Many data analysis methods for modeling hormone pulsatility either assume the timing of the impulses belongs to a certain class of stochastic processes or use pulse detection algorithms [3,4]. These procedures work well when the pulses are readily identifiable and are more challenging when the pulses are more difficult to discern by visual inspection as in the case of cortisol. Johnson et al. [3] and Vidal et al. [4], respectively, analyze LH data using a birth-death process in an MCMC algorithm and a pulse detection algorithm. While our algorithm also can be applied to LH, we analyzed cortisol data, in which the timing of the hormone pulses is not as clearly defined as in LH data. In summary, our proposed algorithm works well even when the pulses are not easily identifiable while still being applicable to cases in which pulses are identifiable by visual inspection. Furthermore, our algorithm does not require assumptions about the inter-arrival times of the pulses, and timings of pulses can be recovered for different classes of distributions of inter-arrival times.

Although our proposed algorithm runs on average in less than half an hour, it can be accelerated. In the data that we analyzed, the likelihood of two pulses occurring during small intervals is low because the experiment was conducted when the participants were under relatively low stress, were fed regularly, had low constant activity levels, and the ambient temperature was held constant. We have not tested our algorithm on data from settings in which the likelihood of two impulses occurring in short intervals is high due to external factors such as stress. Moreover, in our analysis, we make the assumption that the pulses occur at any minute; hence, the detected pulses could have happened within a minute before or after the pulse was detected. Using an approach similar to the one proposed here, with an appropriate number of data points, one could bin the data assuming the pulses occur at any second, or even millisecond and detect the pulses with a higher resolution.

Author Contributions

Conceived and designed the experiments: EBK GKA. Performed the experiments: EBK GKA. Analyzed the data: RTF. Contributed reagents/materials/analysis tools: RTF MAD ENB. Wrote the manuscript: RTF.

References

1. Vis DJ, Westerhuis JA, Hoefsloot HCJ, Pijl H, Roelfsema F, et al. (2010) Endocrine pulse identification using penalized methods and a minimum set of assumptions. *American Journal of Physiology - Endocrinology And Metabolism* 298: E146–E155.
2. Kettyl WM, Arky RA (1998) *Endocrine Pathophysiology*. Philadelphia: Lippincott-Raven.
3. Johnson TD (2003) Bayesian deconvolution analysis of pulsatile hormone concentration profiles. *Biometrics* 59: 650–660.

Table 4. The Error in Estimated Pulses of the Simulated Cortisol Time Series.

Participant	M_e (min)	N_u	N_d
1	4	2	0
2	14	2	0
3	6	1	0
4	26	1	0
5	10	1	1
6	22	1	0
7	14	0	0
8	22	0	2
9	13	2	2
10	22	1	0

M_e is the maximum error in the detected timing of pulses. N_u and N_d are, respectively, the number of pulses that are not detected and the number of extra pulses that are detected for each simulated dataset.

doi:10.1371/journal.pone.0085204.t004

4. Vidal A, Zhang Q, Medigue C, Fabre S, Clement F (2012) Dynpeak: An algorithm for pulse detection and frequency analysis in hormonal time series. *PLoS ONE* 7: e39001.
5. Veldhuis JD, Keenan DM, Pincus SM (2008) Motivations and methods for analyzing pulsatile hormone secretion. *Endocrine Reviews* 29: 823–864.
6. Keenan DM, Chattopadhyay S, Veldhuis JD (2005) Composite model of time-varying appearance and disappearance of neurohormone pulse signals in blood. *Journal of Theoretical Biology* 236: 242–255.
7. Sarabdjitsingh R, Jols M, de Kloet E (2012) Glucocorticoid pulsatility and rapid corticosteroid actions in the central stress response. *Physiology & Behavior* 106: 73–80.
8. Vinther F, Andersen M, Ottesen JT (2011) The minimal model of the hypothalamic pituitary adrenal axis. *Journal of Mathematical Biology* 63: 663–690.
9. Conrad M, Hubold C, Fischer B, Peters A (2009) Modeling the hypothalamus pituitary adrenal system: homeostasis by interacting positive and negative feedback. *Journal of Biological Physics* 35: 149–162.
10. Brown EN, Meehan PM, Dempster AP (2001) A stochastic differential equation model of diurnal cortisol patterns. *American Journal of Physiology - Endocrinology And Metabolism* 280: E450–E461.
11. Veldhuis JD, Iranmanesh A, Lizarralde G, Johnson ML (1989) Amplitude modulation of a burstlike mode of cortisol secretion subserves the circadian glucocorticoid rhythm. *American Journal of Physiology-Endocrinology And Metabolism* 257: E6–E14.
12. Roelfsema F, van den Berg G, Frlich M, Veldhuis JD, van Eijk A, et al. (1993) Sex-dependent alteration in cortisol response to endogenous adrenocorticotropin. *Journal of Clinical Endocrinology & Metabolism* 77: 234–40.
13. Boufounos P, Duarte M, Baraniuk R (2007) Sparse signal reconstruction from noisy compressive measurements using cross validation. In: *IEEE Workshop on Statistical Signal Processing*. pp. 299–303.
14. Klerman EB, Goldenberg DL, Brown EN, Maliszewski AM, Adler GK (2001) Circadian rhythms of women with fibromyalgia. *Journal of Clinical Endocrinology & Metabolism* 86: 1034–1039.
15. Faghieh RT, Savla K, Dahleh MA, Brown EN (2011) A feedback control model for cortisol secretion. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. pp. 716–719. doi:10.1109/IEMBS.2011.6090162.
16. Faghieh RT (2010) The FitzHugh-Nagumo model dynamics with an application to the hypothalamic pituitary adrenal axis. SM thesis, Massachusetts Institute of Technology.
17. Zdunek R, Cichocki A, Member S (2008) Improved m-focuss algorithm with overlapping blocks for locally smooth sparse signals. *IEEE Transactions on Signal Process* 56: 4752–4761.
18. Gorodnitsky I, Rao B (1997) Sparse signal reconstruction from limited data using focuss: a reweighted minimum norm algorithm. *IEEE Transactions on Signal Processing* 45: 600–616.
19. Murray JF (2005) Visual recognition, inference and coding using learned sparse overcomplete representations. PhD thesis, University of California, San Diego.
20. Golub GH, Heath M, Wahba G (1979) Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics* 21: 215–223.
21. Hansen PC (2000) The l-curve and its use in the numerical treatment of inverse problems. In: *Computational Inverse Problems in Electrocardiology*, ed. P. Johnston. *Advances in Computational Bioengineering*. WIT Press, pp. 119–142.
22. Attouch H, Bolte J, Redont P, Soubeyran A (2010) Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the kurdyka-lojasiewicz inequality. *Mathematics of Operations Research* 35: 438–457.
23. He Z, Xie S, Cichocki A (2012) On the convergence of focuss algorithm for sparse representation. *arXiv preprint arXiv 1202.5470*.