

[4] Deconvolution Analysis as a Hormone Pulse-Detection Algorithm

By MICHAEL L. JOHNSON, AMELIA VIROSTKO, JOHANNES D. VELDHUIS, and WILLIAM S. EVANS

Introduction

Simultaneous estimation of admixed basal and pulsatile secretory activity from serial plasma hormone concentration measurements presents a formidable analytical challenge due to the strong correlations among the key (unobserved) secretion and elimination parameters. However, quantifying nonpulsatile and pulsatile contributions to neuroendocrine signaling is critical to a better understanding of specific regulatory mechanisms. As an initial step toward addressing this issue, we have formulated an automated parametric deconvolution strategy. The recursive algorithm (PULSE4) applies relevant statistical tests to successive estimates of secretory burst mass and basal hormone secretion conditional on iteratively assessed peak positions. Monte Carlo simulations of luteinizing hormone- and insulin-like pulse trains superimposed upon variable basal release disclosed true-positive and false-positive pulse identification rates of 95–100% and approximately 2%, respectively, under optimal conditions. Further analysis showed that the optimal sampling frequency ranges from three to five measurements per hormone half-life and the de facto interpulse intervals exceeds the hormone half-life by 1- to 1.25-fold. A priori knowledge of the imbedded hormone half-life enhanced, but was not required for adequate estimates of unobserved basal secretion. We conclude that iterative least-squares deconvolution analysis of combined basal and pulsatile hormone secretion patterns can be formulated usefully as a peak-identification methodology.

The serum concentrations of many hormones change by orders of magnitude multiple times within each day. Examples of such hormones include luteinizing hormone (LH), growth hormone (GH), prolactin (PRL), thyrotropin (TSH), and adrenocorticotrophic hormone (ACTH). Temporal variations in hormone concentrations provide a critical means by which endocrine glands communicate with their remote target organs.¹⁻⁴ Thus, it is important to be able to identify and quantify the pulsatile nature

¹ C. Desjardins, *Biol. Reprod.* **24**, 1 (1981).

² R. J. Urban, W. S. Evans, A. D. Rogol, D. L. Kaiser, M. L. Johnson, and J. D. Veldhuis, *Endocr. Rev.* **9**, 3 (1988).

³ R. J. Urban, D. L. Kaiser, E. van Cauter, M. L. Johnson, and J. D. Veldhuis, *Am. J. Physiol.* **254**, E113 (1988).

of such time series of serum concentrations. Many algorithms have been developed for this analysis.²⁻⁵

In 1987 we noted that the temporal shape of hormone pulses is a convolution integral of secretion into and elimination from the plasma:

$$C(t) = C_0 + \int_0^t S(z)E(t-z)dz \quad (1)$$

where $C(t)$ is the concentration of serum hormone at any positive time, $t \geq 0$; C_0 is the concentration of hormone at the first time point, $t = 0$, $S(z)$ is the amount of hormone secreted at time z per unit time and unit distribution volume, and $E(t-z)$ is the amount of hormone elimination that has occurred in the time interval $t-z$.⁶ Consideration of the temporal shape of hormone pulses has provided a significant and different perspective with which to identify and to quantify episodic hormone glandular secretory activity. A relevantly chosen parametric deconvolution process allows one to evaluate the time course of the secretion of glandular products into circulation. The temporal nature of the inferred secretion profiles can subsequently be used to identify hormone pulses and thereby characterize their regulation.

The analysis of data from immunologically based assays, utilized to evaluate hormone concentration time series, poses some unique issues driven by the properties of the experimental uncertainties inherent in the data. For example, hormone assays and attendant sampling procedures are expensive, and thus provide relatively few data points. For various experimental reasons missing data points may limit analysis further by yielding incomplete data sets. In addition, current sampling and assay techniques yield considerable random variability, i.e., experimental noise. Furthermore, the magnitude of such experimental uncertainties is not constant but a complex function of the hormone concentration. Consequently, any method of hormone pulse analysis must explicitly consider (1) sparse time series, (2) missing data, (3) large experimental uncertainties, and (4) variable experimental uncertainties.

Most deconvolution methods were developed for spectroscopic and/or engineering applications where the signal-to-noise ratio is high (i.e., low experimental uncertainties), data series are ample, and observations are equally spaced (i.e., no data points are missing).⁷ For physical science

⁴ W. S. Evans, M. J. Sollenberger, R. A. Booth, A. D. Rogol, R. J. Urban, E. C. Carlsen, M. L. Johnson, and J. D. Veldhuis, *Endocr. Rev.* **13**, 81 (1992).

⁵ J. D. Veldhuis and M. L. Johnson, *Am. J. Physiol.* **250**, E486 (1986).

⁶ J. D. Veldhuis, M. L. Carlson, and M. L. Johnson, *Proc. Natl. Acad. Sci. USA* **84**, 7686 (1987).

⁷ P. A. Jansson, in "Deconvolution with Applications in Spectroscopy." Academic Press, New York, 1984.

applications, the desired result is a continuous spectrum, and, consequently, common deconvolution methodologies are nonparametric. In contrast, for biological applications, we have implemented a parametric approach, wherein one models the secretion, $S(t)$ in Eq. (1), and the elimination, $E(t)$ in Eq. (1), with distinct and physiologically relevant mathematical forms. One then performs a variably weighted least-squares fit of the hormone concentration time series in terms of the parameters for the secretion and elimination equations, $S(t)$ and $E(t)$. Given the nature (i.e., magnitudes and distributions) of the experimental uncertainties contained within the observations, this approach is a maximum likelihood method. An analogous iterative reconvolution technique is utilized in the evaluation of fluorescence lifetimes.⁸

Specifically we assume that the secretion profile can be described as a series of distinct Gaussian-shaped pulses, i.e., secretion events:

$$S(z) = S_0 + \sum_{i=1}^n e^{\log H_i} - \frac{1}{2} \left(\frac{z - PP_i}{Burst\ SD} \right)^2 \quad (2)$$

where S_0 is the basal secretion, $\log H_i$ is the logarithm of the height of the i th secretion event, PP_i is the position of the i th secretion event, and $Burst\ SD$ is the standard deviation describing the width of the secretion events. The logarithm is used to constrain the analysis to produce only positive values for the amplitudes of the secretion events, H_i . The full-width at half-height of the secretion events is approximately 2.354 times the $Burst\ SD$. The elimination function is empirically based on both one [Eq. (3)] and two [Eq. (4)] compartment pharmacokinetic models of elimination:

$$E(t-z) = \begin{pmatrix} e^{-\left[\frac{\ln 2 (t-z)}{HL}\right]}, & \text{if } t-z \geq 0 \\ 0, & \text{if } t-z < 0 \end{pmatrix} \quad (3)$$

where HL is the one compartment elimination half-life and

$$E(t-z) = \begin{pmatrix} (1-f_2)e^{-\left[\frac{\ln 2 (t-z)}{HL_1}\right]} + f_2e^{-\left[\frac{\ln 2 (t-z)}{HL_2}\right]}, & \text{if } t-z \geq 0 \\ 0, & \text{if } t-z < 0 \end{pmatrix} \quad (4)$$

where HL_1 and HL_2 are the two half-lives of the two compartment model and f_2 is the fractional amplitude corresponding to HL_2 . The zeros in the second half of these definitions indicate that elimination does not take place before secretion occurs. In mathematical terms, the elimination is multiplied by a Heaviside function.

⁸ J. R. Lakowicz, in "Principles of Fluorescence Spectroscopy." Plenum, New York, 1983.

In our 1987 construction,⁶ putative pulse locations are first nominally identified by discrete peak-detection methods, such as CLUSTER analysis.⁵ Peaks are manually removed based on failed statistical confidence intervals for the fitted pulse amplitudes. Pulses are manually added based on evident regions of consecutive positive residuals (differences between the observed and predicted reconvolution curves), thus allowing for potential subjectivity.

The present work outlines and tests an expressly automated deconvolution algorithm to analyze hormone concentration time series. We specifically examine the robustness of the operating characteristics of this algorithm in relation to the identification of (true-positive) hormone pulses under expected variance constraints and in relation to a priori known versus unknown hormone half-lives.

Methods

Algorithmic Outline

By way of overview, the automated algorithm, PULSE4, initially generates an expanding series of presumptive peak locations, and then successively removes those that do not meet statistical significance criteria. The primary objective is to evaluate the secretion event positions, PP_i s, and amplitudes, $\log H_i$ s. To create testable peaks, a natural logarithmic transformation is performed on the data values minus the “Initial Basal Concentration” (later). Data values less than the “Initial Basal Concentration” are discarded at this first stage, and only this stage, of the algorithm.

The logarithmically transformed data can be approximated to some degree as a series of discontinuous straight lines. The locations of the discontinuities correspond to the approximate locations of the secretion events, and the slopes of the predicted lines are related to the apparent single component elimination half-life. Consequently, the initial portion of the algorithm finds the most likely places to put these discontinuities by performing a series of weighted linear least-squares fits, first assuming that there are no secretion events, i.e., that the data are described by an initial concentration at time zero and a decay time.

The algorithm now locates the best location to insert the first presumptive secretion peak by evaluating the statistical consequences of inserting a single peak at every data point location (i.e., time value). For each possible location of a secretion event a weighted linear least-squares fit is performed. The resulting variance of fits are tabulated for each tested secretion-event location. The first presumptive secretion-event location is

assigned based on the location yielding the lowest variance of fit associated with a positive secretion amplitude. If there are N data points, the first step thus requires N separate weighted linear least-squares parameter estimations.

The foregoing process is repeated to evaluate the locations of the subsequent presumptive peak locations. This process is repeated for three, four, . . . presumptive peaks until the residuals appear random by three statistical tests; (1) the Runs Z score ($Z = 2.0$) for too few runs is not statistically different from zero, (2) the insertion of an additional peak does not significantly ($p < 0.05$) lower the variance of fit, and (3) the lag = 1 autocorrelation of the residuals is not significantly positive ($p < 0.05$).

At this stage, the algorithm determines whether the first data point corresponds to a new secretion event location or simply represents the tail of a previous secretion event. If the timing of the first presumptive secretion event lies within three times the “minimum secretion SD” (later), then the concentration at time zero is assigned a constant value of zero. Otherwise, the concentration at time zero is assumed to be a variable to be estimated by the subsequent weighted nonlinear least-squares (WNLLS) procedures.

To test whether the amplitudes of the individual presumptive secretion events are significant, the algorithm calculates the apparent secretion curve at intervals of one-third the minimum time between the data points using the “minimum secretion event *Burst SD*” (discussed later). Secretion events are tested by scanning the calculated secretion curve for peaks (maxima) separated by nadirs (minima). The actual *Burst SD* of the secretion event is then calculated as the average full-width at half-height of the set of identified secretion events divided by 2.354.

The final phase of the PULSE4 algorithm is the parameter refinement section wherein the statistical significance of the remaining presumptive secretion events is evaluated. For the Monte Carlo simulations presented in this report, the refinement stage proceeds by assuming either (1) that the elimination process is a single elimination function as given by the initial value of the elimination half-life determined from the slope of the line segments fitted to the logarithmic transformed data, or (2) by using the one- and two-component elimination half-lives that were used to simulate the time series. Subsequent weighted nonlinear least-squares fitting during the refinement stage is done by a procedure that is mathematically identical to our original multiparameter DECONV program.⁶ In the parameter refinement stage, all of the original non-logarithmic transformed data is utilized for the iterative parameter estimations.

For each stage in the refinement process, the algorithm performs an additional series of variably weighted least-squares parameter estimations.

If an error is detected in any of these steps, then the corresponding event is removed and the refinement process is repeated from the beginning. The specific iterative sequence entails the following:

1. Estimation of secretion-event positions only (i.e., the PP_i s) while holding the other parameters constant.
2. Estimation of secretion-event amplitudes only (i.e., S_0 and the $\log H_i$ s) while holding the other parameters at the values found in step 1.
3. Testing whether including the smallest secretion event provides a statistically significant decrease in the overall variance of fit.
4. Estimation of only the secretion-event amplitudes and the single component elimination half-life given the results of step 2, i.e., S_0 , $\log H_i$ s, and the HL ; and,
5. Retesting whether the smallest secretion event provides a statistically significant decrease in the overall variance of fit.

The foregoing sequence of parameter estimations is arbitrary but empirically satisfactory. The following error tests are imposed at one or more of the five refinement steps. If a peak position falls outside the actual range of the data in step 1, then it is removed as a presumptive secretion event and the refinement procedure restarts from the beginning. Second, if peaks are coincident (i.e., during the peak position evaluation steps one or more of the peaks has moved to within 2.5 times the secretion event SD of another peak), they are combined into a single secretion event and the refinement procedure is restarted from the beginning. And, third, to determine if any of the secretion events are actually minor shoulders on other secretion events, PULSE4 calculates the current secretion profile at each of the data points. If a nadir (minima) does not occur between consecutive presumed secretion events then the pair of events is replaced with a single event corresponding to the maximum location and the refinement procedure restarts from the beginning.

Steps 3 and 5 determine if the secretion event with the smallest amplitude peak is statistically significant. This is done by temporarily removing it and repeating the parameter estimation to determine the apparent variance of fit. An F test for an additional term is performed to determine if inclusion of the peak provided a statistically significant ($p < 0.05$) decrease in the variance. Since each secretion event adds two parameters (the amplitude and the position of the peak), the number of degrees of freedom for this F test is 2 and the number of data points minus the total number of parameters.

To accept the final solution, as outlined previously, one of two conditions must be met. If either (1) a runs test for too few runs within the residuals ($p < 0.05$) is not significant or (2) the $lag = 1$ autocorrelation of

the residuals is not significantly positive ($p < 0.05$) then the solution is acceptable. Ideally the solution should pass both of these tests and the algorithm performs even better when both conditions are met. If both of these tests fail then the entire algorithm is automatically restarted with the Initial Basal Concentration increased by 10%. Finally, (3) this last operation is performed a maximum of six times to eliminate the possibility of an infinite loop.

Monte Carlo Simulations

The algorithm was tested by a series of Monte Carlo simulations. One hundred synthetic data sets were generated for each of the simulations, based upon reasonable estimates for the various typical parameters for mid-luteal phase LH and insulin secretion, as shown in Table I. In these simulations, the Interpulse Intervals and $\log H_i$ were assumed to be Gaussian distributed with the means and SD values given. However, the interval from one secretion event to the next was assumed to be positive and as a consequence negative values in the simulation of the Interpulse Intervals were truncated. Once a series of secretion events and amplitudes was simulated, the concentration as a function of time was determined from the convolution integral [Eq. (1)].

TABLE I
TYPICAL VALUES OF THE PARAMETERS USED IN THE SIMULATIONS

	Mid-luteal LH	Insulin
Number of data points	144	150
Time between data points	10 min	1 min
Number of replicates	2	2
Interpulse interval	91 ± 18 min	12 ± 2 min
$\log H_i^a$	0.4 ± 0.2	6.0 ± 0.8
Burst SD^b	4.5 min	1.0 min
S_0^c	0.04	6.0
Single component HL	55 min	3.1 min
Two component HL_1	18 min	2.8 min
Two component f_2	0.37	0.28
Two component HL_2	90 min	5.0 min
Noise SD (minimum)	0.1	0.2
Noise CV	5.0%	3.6%

^a $\log H_i$ is the base 10 logarithm of the secretory pulse amplitude (units of concentration per minute).

^b Burst SD represents the secretory pulse full-width at half-height in minutes divided by 2.354.

^c S_0 denotes the basal secretory rate (units of concentration per minute).

For each simulation, a realistic level of Gaussian distributed pseudo-random noise was added to the data sets. The variance model for this noise as a function of the simulated hormone concentration is given by

$$\begin{aligned} \text{Variance of noise (concentration)} = & [\text{noise } SD]^2 \\ & + \left[\frac{\text{concentration} [\text{noise } CV\%]}{100} \right]^2 \end{aligned} \quad (5)$$

The *noise SD* term is included so that the simulated experimental uncertainties approach a realistic nonzero minimum value as the simulated concentration approaches zero. Whereas these simulations will not exactly duplicate the actual observed mid-luteal LH and insulin concentration time series, they provide a set of realistic profiles to test algorithmic performance under a variety of conditions. Since the “answers” are known for each of these simulations they can be utilized to test the performance characteristics of the algorithm.

Data sets were simulated based upon typical values for the secretion event interpulse intervals, secretion event amplitudes, basal secretion, elimination half-lives, and with realistic pseudo-random “experimental observational errors” added. It is clearly impossible to present all possible permutations of the interpulse intervals, amplitudes, basal secretion, elimination half-lives, and “observation errors.” Thus, the approach taken here is to present analyses across a plausible range of perturbations.

Results

Figure 1 illustrates a typical simulation of a mid-luteal LH concentration (± 1 SEM) time series based upon the values presented in Table 1. The corresponding secretion profile, as evaluated by the PULSE4 algorithm with an unknown single elimination half-life, is shown in Fig. 2. Figure 1 also shows the calculated LH concentration profile (reconvolution curve) based upon the inferred secretion profile.

The computational output of the PULSE4 algorithm is not actually a continuous curve as shown in Fig. 2, but rather a table of secretion event locations, PP_i s, and the logarithms of the corresponding secretion event heights, $\log H_i$ s. Figure 2 is simply a calculated secretion curve based upon the final table of pulse-parameters values estimated by the PULSE4 algorithm.

The particular example shown in Figs. 1 and 2 was chosen to demonstrate some difficult to locate secretion events that were located by the PULSE4 algorithm. The four diamonds in Fig. 1 correspond to the time locations of

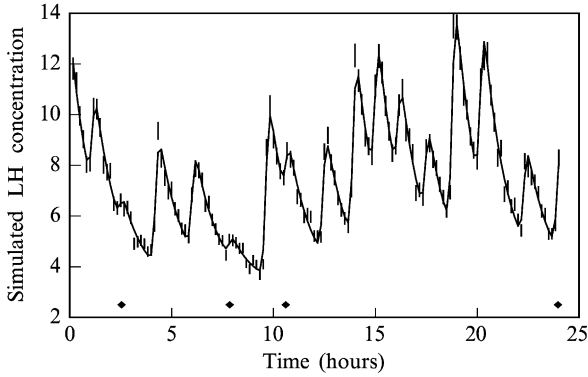


FIG. 1. The vertical lines represent ± 1 SEM error bars for a simulated human mid-luteal phase LH concentration time series based upon the typical parameters presented in Table I. The smooth curve in this figure corresponds to the estimated concentration time series based upon the calculated secretion profile shown in Fig. 2, as evaluated by the PULSE4 algorithm assuming an unknown single elimination half-life. The diamonds correspond to the time locations of difficult to locate simulated secretion events illustrated in this analysis.

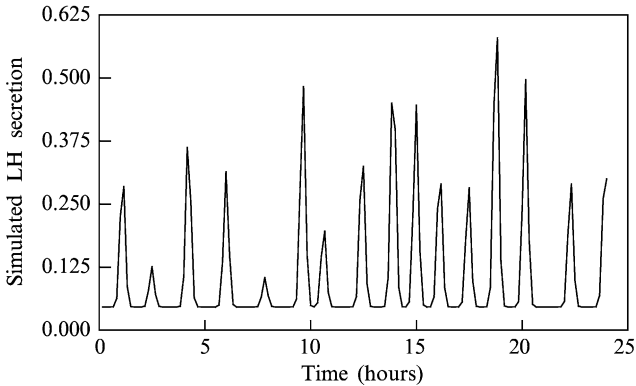


FIG. 2. LH secretion profile inferred from the simulated data shown in Fig. 1 by the PULSE4 algorithm.

four difficult to locate simulated secretion events located by the PULSE4 algorithm in this particular set of data. For example, the fourth diamond marks a simulated peak at 1437.6 min, only 2.4 min before the last simulated data point. The other three marked secretion events are on the trailing edge of a previous secretion event. Sixteen peaks were simulated in this example, and all were located accurately by the PULSE4 algorithm.

TABLE II
 CLUSTER AND PULSE4 ALGORITHMIC ANALYSES OF SIMULATED *LH* DATA SETS^a

CLUSTER algorithm	Simulated single <i>HL</i> data		Simulated dual <i>HL</i> data	
True-positive rate (%) ^b	91.0		92.4	
False-negative rate (%) ^c	9.0		7.6	
False-positive rate (%) ^d	0.6		0.4	
Peak position error (min)	10.1 ± 4.4		8.7 ± 3.5	
PULSE4 algorithm	Unknown <i>HL</i>	Known <i>HL</i>	Unknown <i>HL</i>	Known <i>HLs</i>
True-positive rate (%)	95.5	99.3	96.7	99.2
False-negative rate (%)	4.5	0.7	3.3	0.8
False-positive rate (%)	0.4	0.8	1.2	1.4
Basal secretion (S_0)	0.054 ± 0.025	0.033 ± 0.003	0.064 ± 0.027	0.035 ± 0.026
<i>HL</i> (min)	48 ± 12		38 ± 10	
Peak position error (min)	0.7 ± 3.3	0.5 ± 1.7	-0.6 ± 2.5	0.4 ± 1.6
Peak mass (observed-true)	0.082 ± 0.105	0.092 ± 0.081	0.002 ± 0.094	0.081 ± 0.101

^a Values are the mean for 100 simulated series ± SD. See Table I for simulation parameters. *HL* denotes the apparent half-life embodied in the $E(t - z)$ kinetic function [Eqs. (3) and (4)].

^b The true-positive rate is the percentage of simulated peaks that was located to within two data points by the algorithms.

^c The false-negative rate is the percentage of simulated peaks that was not located to within two data points by the algorithms.

^d The false-positive rate is the percentage of peaks located by the algorithms that did not correspond to a simulated peak.

To appraise peak-detection performance, we tabulated true-positive peak identification and false-positive errors as discrimination indices on simulated data series. For example, 100 data sets were simulated using the nominal single elimination LH values in Table I and a second set of 100 data sets was simulated based upon the nominal two component elimination half-lives. The results of the analysis of each of these data set with known and unknown half-lives are summarized in Table II. For example, when the PULSE4 algorithm was applied to the single elimination simulated data under the assumption that the elimination function corresponds to a single unknown half-life, the true-positive rate was 95.5%, the false-negative rate was 4.5%, and the false-positive rate was 0.4%. For comparison, when the same simulated data sets were analyzed with the CLUSTER algorithm,⁵ the resulting true-positive rate was 91.0%. The CLUSTER algorithm provides no information about the apparent elimination half-life or basal secretion. The simulated basal secretion rate was 0.04 and the

PULSE4 estimated value was 0.054 ± 0.025 (SD). Uncertainties are represented as SDs that provide a better description of the expected variability for an individual data set than do SEMs. The simulated elimination half-life was 55 min and the estimated value by the PULSE4 algorithm was 48 ± 12 (SD). The mean error in locating a secretion event position was 0.7 ± 3.3 (SD) mins. Corresponding results for the CLUSTER method averaged 10.1 ± 4.4 mins. In judging the error in locating a secretion event remember that the data were simulated at 10-min intervals.

The PULSE4 algorithm can also evaluate the apparent mass of hormone released within each secretion event. The area under a Gaussian-shaped secretion event is

$$\text{Burst Area}_i = \frac{SD \ 10^{\log H_i}}{\sqrt{2\pi}} \quad (6)$$

For the present simulation, the average simulated burst mass was 1.96 units. In this setting, the average error (i.e., the difference between the apparent and the simulated burst masses) in burst mass of true-positive peaks was 0.08 ± 0.10 (data were simulated as a single elimination half-life and analyzed by the PULSE4 assuming a single unknown half-life).

One question of particular interest is how often experimental observations should be collected. This question is of immediate practical significance to the experimentalist. For example, being able to utilize a 20-min instead of 10-min sampling frequency can represent substantial research savings. [Figure 3](#) presents the true-positive and false-positive rates as a function of sampling interval for simulated mid-luteal LH data analyzed by the PULSE4 algorithm. It is interesting to note that for long sampling intervals the true-positive rate decreases and for short sampling intervals the false-positive rate increases. For the particular choice of the parameters used in the mid-luteal LH simulation model, the maximum adequate sampling interval appears to be approximately 15 min while the minimum sampling interval appears to be approximately 10 min. However, the important general concept here is that the minimum number of data points per elimination half-life is approximately four, when secretory-pulse detection is the primary outcome of interest.

[Figure 3](#) indicates that at least for this simulation, the true-positive rates decrease as the sampling interval increases and the false-positive rates increase as the sampling interval decreases. However, these changes do not occur as a uniform change. For example, for the 20-min sampling with two known half-lives a total of 223 false-negatives were identified for the 100 simulated data sets. Forty-five of the simulations found zero false-negatives giving a mode of zero. Nineteen of the simulations found

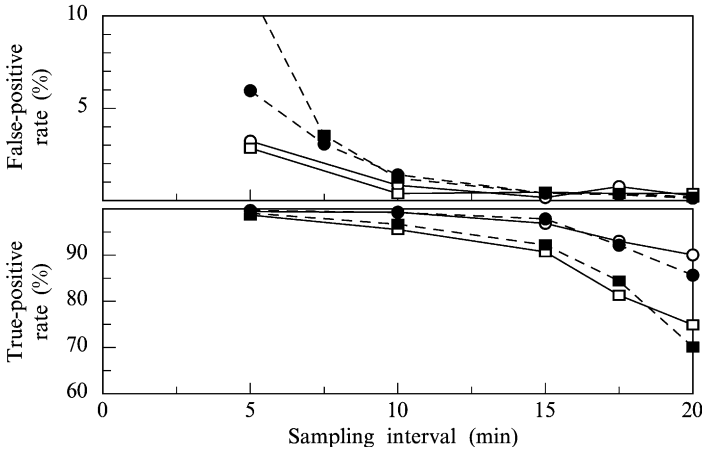


FIG. 3. Apparent true-positive and false-positive rates for the PULSE4 algorithm when applied to mid-luteal phase LH concentration data simulated as described in Table I. The open symbols and solid lines correspond to data simulated for a single-elimination model. The closed symbols and the dashed lines correspond to a two-exponential model. The circles denote PULSE4 analyses based on known elimination half-lives and the squares are for analyses with an unknown single elimination half-life.

only a single false-negative. More than 50% of the false-negatives were from only 11% of the simulations.

The general simulation and testing procedure outlined previously can be utilized to address a number of questions about the sensitivity of the PULSE4 algorithm that cannot be tested directly *in vivo*. For example, Fig. 4 presents an analysis of the sensitivity of the PULSE4 algorithm to a varying interpulse interval of hormone release. For each of the points in Fig. 4, the mid-luteal phase LH parameters shown in Table I were utilized except that the interpulse interval was varied from 40 to 120 min (i.e., from approximately 50% of a half-life to approximately two half-lives). Simulation results indicated that the lower bound of optimal pulse detection is an interpulse interval a little more than one half-life. It is also interesting to note that the PULSE4 algorithm with known half-lives always performs better than PULSE4 with an unknown half-life and both perform better than the CLUSTER method.

The operating characteristics of the PULSE4 algorithm were also explored in relation to other pulsatile hormones. For example, Table III presents the performance data based on 100 simulated insulin data sets reflecting the typical insulin secretory parameters shown in Table I (i.e., the average secretion event mass for this insulin simulation was 2.4).

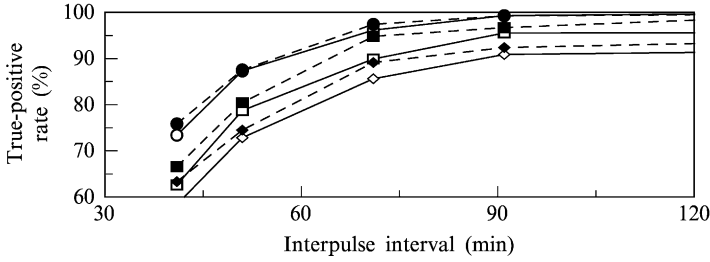


FIG. 4. Apparent true-positive rates for the PULSE4 and CLUSTER algorithms applied to simulated mid-luteal LH phase data of varying mean interpulse intervals. The open symbols and solid lines correspond to data simulated for a single-exponential elimination model. The closed symbols and the dashed lines correspond to data simulated for a two-exponential model. The circles reflect results of PULSE4 analysis with known elimination half-lives, the squares PULSE4 analysis with an unknown single elimination half-life, and the diamonds the CLUSTER algorithm.

TABLE III
COMPARISON OF THE CLUSTER AND PULSE4 ALGORITHM APPLIED TO THE ANALYSIS
OF SIMULATED INSULIN DATA SETS^a

CLUSTER algorithm	Simulated single <i>HL</i> data		Simulated dual <i>HL</i> data	
True-positive rate (%)	92.3		91.5	
False-negative rate (%)	7.7		8.5	
False-positive rate (%)	0.2		0.7	
Peak position error (min)	1.44 ± 0.51		1.46 ± 0.55	
PULSE4 algorithm	Unknown <i>HL</i>	Known <i>HL</i>	Unknown <i>HL</i>	Known <i>HL</i>
True-positive rate (%)	99.1	99.5	99.7	99.3
False-negative rate (%)	0.9	0.5	0.3	0.72
False-positive rate (%)	1.0	1.0	0.4	1.4
Basal secretion (S_0)	5.21 ± 0.75	5.85 ± 0.076	5.33 ± 0.80	5.95 ± 0.11
<i>HL</i> (min)	3.47 ± 0.35		3.82 ± 0.42	
Peak position error (min)	-0.097 ± 0.217	-0.002 ± 0.174	-0.11 ± 0.22	-0.01 ± 0.19
Peak mass (observed-true)	-0.124 ± 0.189	-0.072 ± 0.177	-0.155 ± 0.176	-0.076 ± 0.192

^aSimulation parameter choices are given in Table I.

A comparison of [Tables II and III](#) indicates that the PULSE4 algorithm performs better for insulin than for mid-luteal LH. The origin of this improvement is in the distribution of peak sizes. The distribution of the insulin secretion event amplitudes, $\log H_i$, is such that small secretion events do not occur in the insulin simulations. However, for LH a significant number of very small peaks is simulated, which imposes lower true-positive rates.

Conclusions

The present analyses describe and implement an iterative parametric deconvolution technique to identify and quantitate basal and pulsatile hormone release, while concurrently estimating endogenous hormone half-lives. Simulated examples of the application of this methodology indicated that at least for plausible synthetic time series, this parametric deconvolution approach is superior to the CLUSTER method (i.e., one of the commonly accepted procedures) in locating hormone secretion events. PULSE4 also provides crucial information about secretion-event burst mass, elimination half-lives, and basal secretion.

In general, it is very difficult to separate (i.e., deconvolve) the 3-fold contributions to hormone concentration data of pulse amplitudes, basal secretion, and elimination half-lives. This is, in essence, the quintessential ill-posed problem in numerical analysis. However, the present simulations illustrate that a parametric deconvolution approach conditional on iterative reappraisal of peak positions can provide a good estimate of basal secretion without any prior knowledge of the elimination half-lives, and offers an excellent approximation of the amount of basal secretion when the half-lives are known.

The accompanying simulations also offer guidelines about the features of the experimental data that are required for an accurate characterization of the pulsatile nature of hormone release. Specifically, the sampling frequency should be such that there are approximately four or more data points per elimination half-life. In the simulated LH case with a single component elimination half-life of about 55 min the true-positive rates dropped rapidly when data points were simulated less often than every 15 min and the false-positive rates increased when the data were sampled more often than every 10 min ([Fig. 3](#)). For insulin with a single component elimination half-life of about 3.1 min, the minimum sampling rates appeared to be about 1 min (data not shown).

Parametric deconvolution-based hormone pulse identification methods also have specific limitations in the maximum detectable hormone pulse frequency. After each secretion event, the concentration of the hormone begins to decrease according to the elimination kinetics. To accurately

parameterize the elimination process and basal secretion, a significant amount of elimination must take place before the next secretion event. In general, the minimum time between secretion events needs to equal or exceed 1.25 elimination half-lives (Fig. 4).

Although the present analyses have not explored operating characteristics of pulse detection algorithms for all possible hormones, the principles implied by investigating LH and insulin time-series should be applicable to other neuroendocrine systems. In addition, the foregoing results imply that accurate prior knowledge of hormone-specific kinetics aids significantly in true-positive pulse detection. Likewise, iterative peak analysis conditional on recurrently tested pulse positions avoids subjective peak addition and deletion, thus achieving a refined parameter estimation set via automated computation followed by statistical confirmation.

Software is available from the author (M.L.J.) upon written request.

Acknowledgments

The authors acknowledge the support of the National Science Foundation Science and Technology Center for Biological Timing at the University of Virginia (NSF DIR-8920162), the General Clinical Research Center at the University of Virginia (NIH RR-00847), and the University of Maryland at Baltimore Center for Fluorescence Spectroscopy (NIH RR-08119 and NIH R01 AG14799).