PERGAMON

Computers in Biology and Medicine 29 (1999) 229–242

Computers in Biology and Medicine

An algorithm for detecting features of the hormone profiles of the human menstrual cycle

Edward F. Krieg Jr*, James S. Kesner, Edwin A. Knecht

Division of Biomedical and Behavioral Science, National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, Ohio 45226, USA

Received 5 November 1998; accepted 25 January 1999

Abstract

An algorithm for detecting features of the cycles of the gonadotropic and ovarian hormones in women is described. The algorithm can detect hormone peaks and normal cycles defined in terms of the peaks in sequences of measurements that have an arbitrary starting point in the menstrual cycle and are of arbitrary length. The algorithm makes use of fuzzy set theory and is optimized using signal detection theory. Published by Elsevier Science Ltd.

Keywords: Algorithm; Fuzzy set; Signal detection; Menstrual cycle; Hormone

1. Introduction

Function of the human menstrual cycle is controlled and defined by a complex interaction between hormones secreted by the pituitary and ovaries. These hypophysial and ovarian hormones can be measured in blood or urine for medical or research purposes. For example, endocrine profiles can aid individuals seeking medical reproductive assistance. Researchers can use endocrine data to evaluate the reproductive health of human female populations who may have been exposed to potential hazards in the workplace or general environment.

Endocrine profiles are serial samples of hormone concentrations taken during one or more menstrual cycles. In a clinic setting, blood might be drawn on several days of a cycle. In field

* Corresponding author. Tel.: +1-513-533-8160; fax: +1-513-533-8510.
E-mail address: erk3@cdc.gov (E.F. Krieg Jr)

0010-4825/99/$ - see front matter Published by Elsevier Science Ltd.
PII: S0010-4825(99)00007-4

studies, an individual might collect urine on every day of a cycle. Certain circumstances impede the development of a general method to detect the features of hormone profiles. These include missing samples, sampling that begins mid-way into a cycle, and variations in baseline hormone levels between women, between cycles of the same woman, and between analytical methods.

This paper describes an algorithm that identifies qualitative endocrine features of the human menstrual cycle. The algorithm detects hormone peaks and normal cycles defined in terms of the peaks. The algorithm was developed and tested using urinary values for luteinizing hormone (LH, mIU/ml), follicle stimulating hormone (FSH, mIU/ml), estrone 3-glucuronide ($E_1$3G, the principal urinary metabolite of estradiol, ng/ml), and pregnanediol 3-glucuronide (Pd3G, the principal urinary metabolite of progesterone, μg/ml). These endocrine endpoints were measured using time-resolved fluoroimmunoassays [1,2] and were divided by urinary creatinine concentrations (mg/ml) to normalize for the urine dilution [3]. The algorithm also detects the peak of the $E_1$3G to Pd3G ratio. This ratio is used to identify the day of transition from follicular production of estrogen to luteal production of progesterone. The day of transition is used as an estimate of the day of ovulation [4].

The algorithm makes use of fuzzy set theory, which is an extension of set theory in which the degree of membership in a set is recognized [5]. The degree to which an element is a member of a set is indicated by a value ranging from zero to one. Zero indicates no membership and one indicates complete membership. Operations such as union, intersection, and complement can be applied to fuzzy sets. Unlike classical set theory, there are multiple definitions for these operations, and the fuzzy operators do not always satisfy the laws of classical set theory, e.g., the laws of contradiction ($A \cap \bar{A} = \phi$) and the excluded middle ($A \cup \bar{A} = X$). Here, ϕ is the null set and X is the universal set. In our algorithm, the minimum operator is used for intersection, and the maximum operator is used for union.

Fuzzy set theory is useful when the concepts one is dealing with have an inherent vagueness, like young and old, or big and small. It has been used to control machines and processes, to recognize patterns, and to perform approximate reasoning. Fuzzy solutions to problems can be faster to develop, simpler, and easier to understand and interpret than more conventional methods.

The algorithm was optimized using signal detection theory [6]. In this theory, 'hits' are true positives, 'correct rejections' are true negatives, 'false alarms' are false positives, and 'misses' are false negatives. A receiver operating characteristic (ROC) curve is a plot of the probability of a hit, $P(S \mid s)$, as a function of the probability of a false alarm, $P(S \mid n)$. $P(S \mid s)$ is the probability that a system detects a signal, S, given the presence of the signal and noise, s. $P(S \mid n)$ is the probability that a system detects a signal given the presence of noise only, n. The area under the ROC curve, $P(A)$, can be used as a measure of the performance of a signal detection system. If the area equals 0.5, the probability of a hit equals the probability of a false alarm. The area equals 1.0 if no errors are made.

The steps of the algorithm are described in detail in the following sections and summarized here. The input variables are series of hormone measurements from individual women. The hormone measurements are standardized statistically in the first step of the algorithm. In the second step, the values of the fuzzy membership functions are calculated. In the last step, hormone peaks and normal cycles are searched for. If the fuzzy membership value for a peak or a normal cycle on a given day exceeds a specified value, α, then the algorithm detects a peak or normal cycle on that day. The outputs of the algorithm are binary variables indicating whether or not a peak occurs on a given day, and whether or not a normal cycle occurs. The algorithm is optimized by selecting values of α which maximize the number of correct decisions.

2. Description of the algorithm

2.1. Step 1—standardizing measurements

The first step in the algorithm is standardizing the measurements. All the endocrine measurements were divided by urinary creatinine (mg/ml) to adjust for urine dilution. Then each series of hormone measurements for each woman, including the estrogen to progesterone ratio, was standardized separately.

First the mean ($\bar{x}_1$) and standard deviation (s_1) of a series of measurements were calculated. Values greater than $\bar{x}_1 + 3s_1$, or less than $\bar{x}_1 - 3s_1$, were considered extreme. The mean ($\bar{x}_2$) was recalculated with the extreme values removed, and the standard deviation (s_2) was recalculated by summing the squared deviations of all values less than or equal to the mean ($\bar{x}_2$). Then all the values in the series were standardized by subtracting the mean ($\bar{x}_2$) and dividing by the standard deviation (s_2).

Removing extreme values for the calculation of $\bar{x}_2$ resulted in lower means because the extremely high values of some peaks were not included. Calculating the 'standard deviation', s_2, omitting the values above the mean also eliminated high values and resulted in smaller values. The net result is that the standardized scores of the values of the peaks were accentuated relative to other values.

The standardized values were used in the next step of the algorithm.

2.2. Step 2—calculating membership functions values

Values for the membership functions were calculated in the next step of the algorithm. A membership function assigns a value indicating the degree of membership in a set to each element in the universal set. In our algorithm the membership function representing a peak, *Peak*, is defined as the intersection of two other membership functions, $Peak = Highest \cap High = \min\,(Highest,\ High)$. *Highest*, is a non-fuzzy, or crisp, function. It equals 1 if the standardized hormone value is a local maximum over $\pm d$ days, or 0 otherwise. *High* is defined as

$$High(X) = \begin{cases} 0, & \text{if } X < a \\ \dfrac{X-a}{b-a}, & \text{if } a \le X \le b\,. \\ 1, & \text{if } X > b \end{cases} \tag{1}$$

Here, X represents a set of standardized measurements.

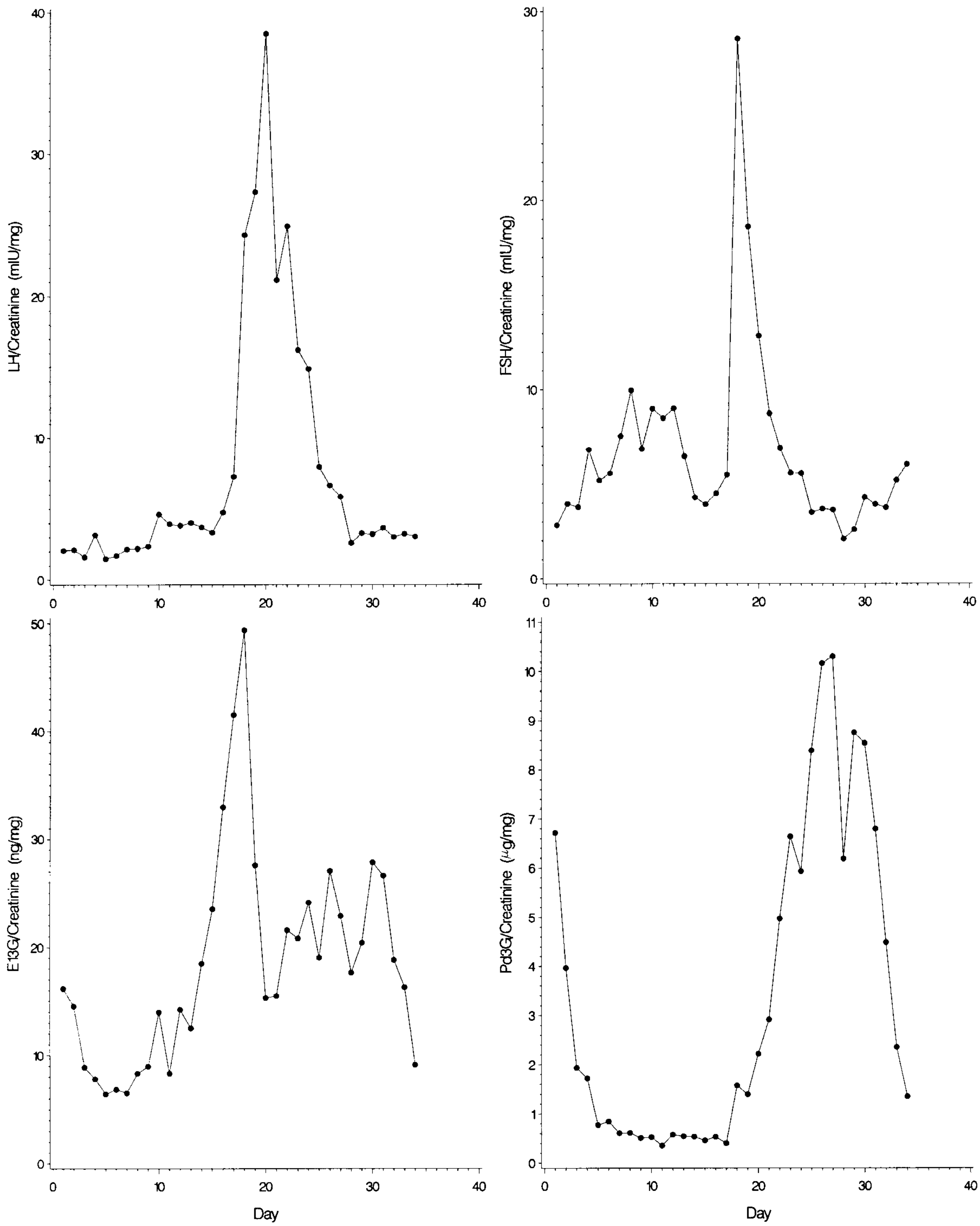

Fig. 1. An example of a normal cycle.

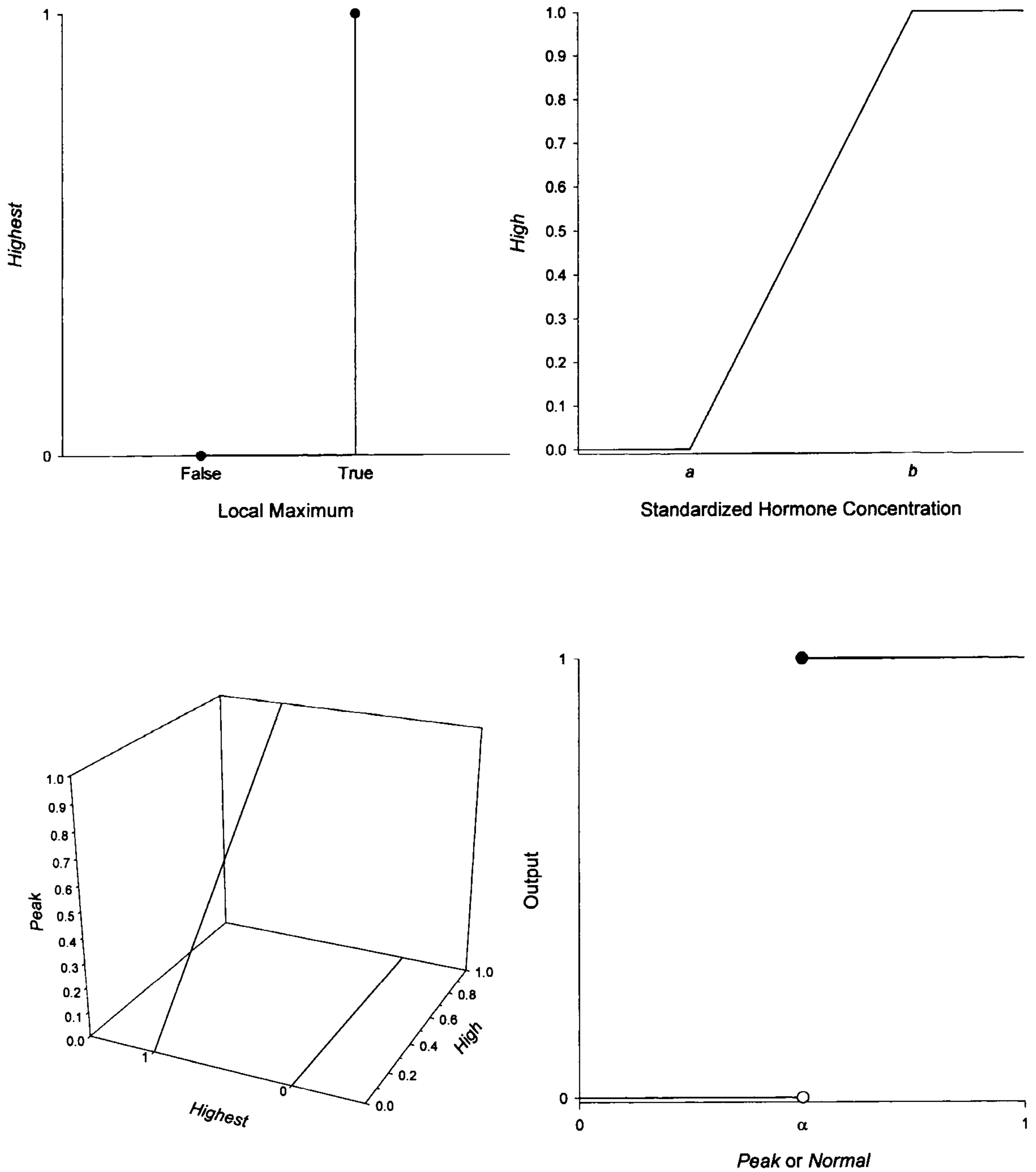

Fig. 2. The membership functions *Highest* (top left) and *High* (top right) as a function of their input variables, the membership function *Peak* as a function of *Highest* and *High* (bottom left), and the output variable as a function of *Peak* or the membership function *Normal* (bottom right).

A normal cycle was defined as a cycle with an LH peak, an FSH peak within −2 to +0 days of the LH peak, an $E_1$3G peak within −3 to +1 days of the LH peak, and a Pd3G peak within +2 to +14 days of the LH peak. To calculate a membership value for the fuzzy set *Normal*, the union of *Peak* was determined for each hormone over the appropriate range of days, then the intersection of the four unions was calculated,

$$Normal = \bigcap_{h \in H} \left[\bigcup_{t \in T_h} Peak(h,t) \right]. \quad (2)$$

Here, H represents the set of four hormones, and T_h is the set of days appropriate for hormone h.

An example of a normal cycle is shown in Fig. 1. The LH peak occurs on day 20, the FSH peak occurs on day 18, the $E_1$3G peak occurs on day 18, and the Pd3G peak occurs on day 27.

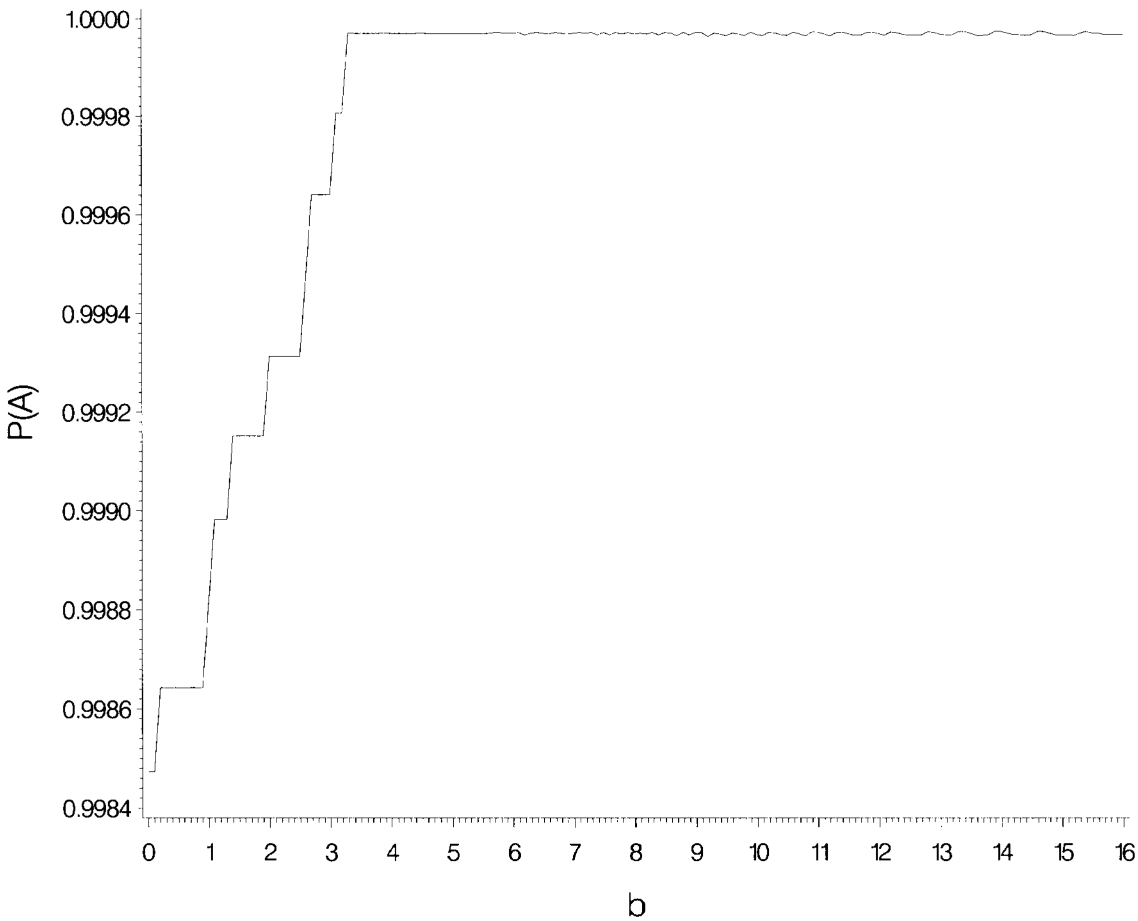

Fig. 3. An example of the area under the ROC curve, $P(A)$, plotted as a function of the parameter b from Eq. (1).

Fig. 2 shows graphical representations of *Highest* and *High* as a function of their input variables, and *Peak* as a function of *Highest* and *High*.

Missing data were handled in the following way. If the standardized hormone value on a given day was missing, then *Highest, High*, and *Peak* were set to missing for that day. *Peak* was also set to missing on a given day if there was a missing datum within ± 1 day of the standardized value for that day. *Normal* was set to missing if there was no value of *Peak* for LH, or if all the values of *Peak* for one of the other hormones in the applicable time window were missing.

2.3. Step 3—searching for peaks and normal cycles

In the final step, peaks and normal cycles are detected. The algorithm proceeds from the first to the last day of a series of measurements. If *Peak* or *Normal* is greater than or equal to a specified value, α, a response is indicated for that day and the algorithm skips k days ahead

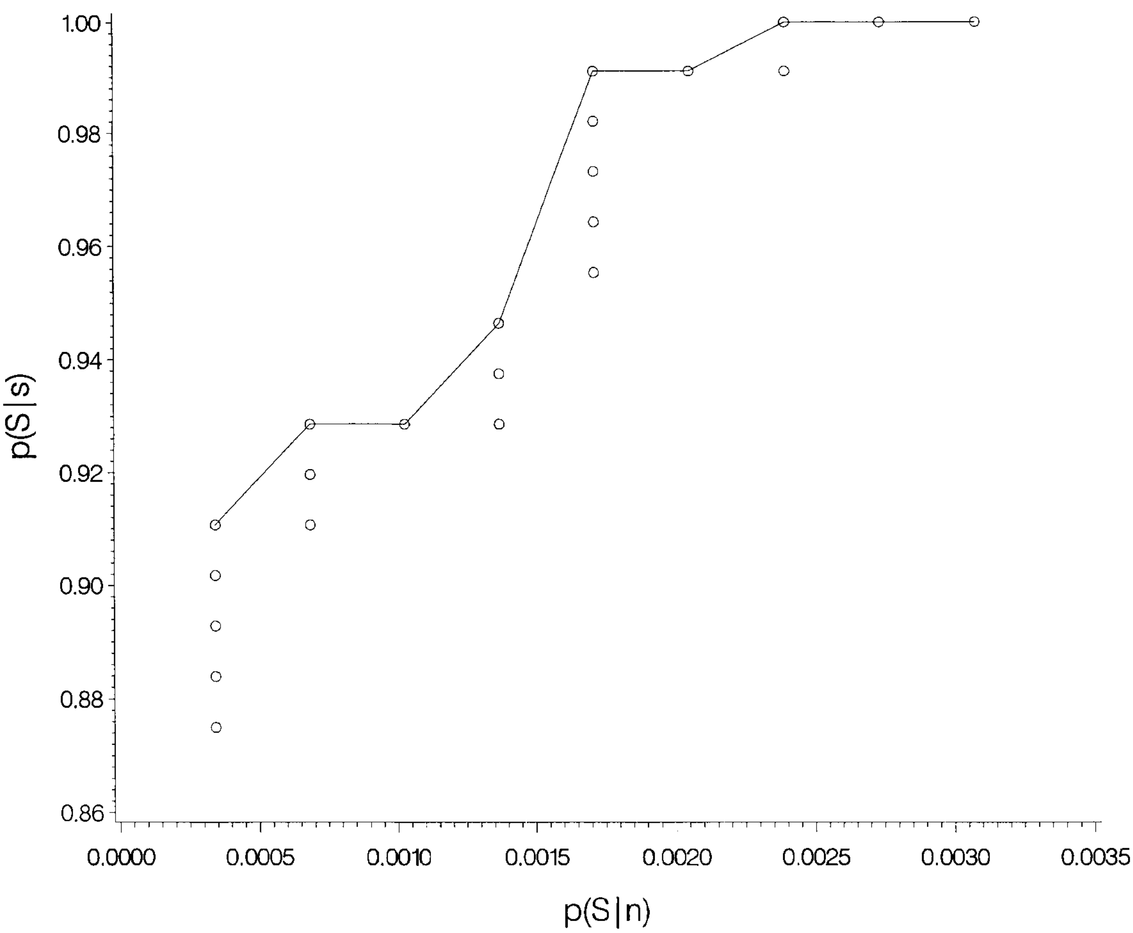

Fig. 4. An example of an ROC curve.

and begins searching again. An example of the output as a function of *Peak* or *Normal* is also shown in Fig. 2.

3. Optimizing the algorithm

The algorithm was optimized to detect LH peaks, FSH peaks which coincided with the LH peak, $E_1$3G and $E_1$3G/Pd3G peaks which coincided with the LH peak, and Pd3G peaks which followed the LH peak. The authors determined which values and cycles were signals.

Values for the parameters *d* and *k* were also determined by the authors. The initial values were chosen based on experience, and the values were adjusted based on where the algorithm made mistakes. The parameter *a* from the membership function *High* was always set to 0.

The optimal value for the parameter *b* was determined by varying it from 0.0 to 16.0 in

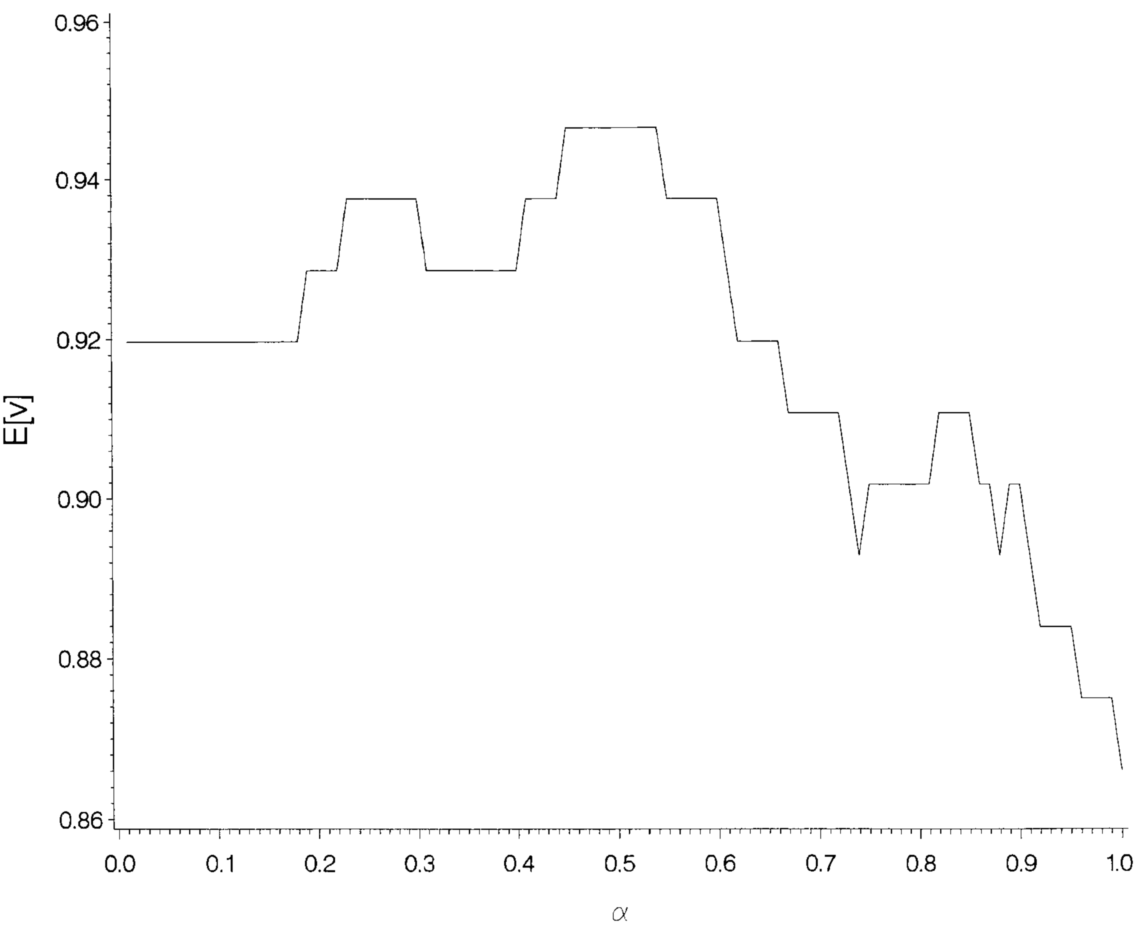

Fig. 5. An example of $E[V]$ from Eq. (4) plotted as a function of α.

increments of 0.1 and calculating the area under the ROC curve, $P(A)$, for each value of b. An example of $P(A)$ as a function of b is shown in Fig. 3.

An ROC curve for a value of b was created by varying α from 0.01 to 1.00 by 0.01 and calculating the probability of a hit, $P(S \mid s)$, and the probability of a false alarm, $P(S \mid n)$, for each value of α. An example of an ROC curve is shown in Fig. 4. There were often multiple values of $P(S \mid s)$ for a unique value of $P(S \mid n)$. The largest value was always used to calculate the area using the equation

$$P(A) = \sum_{i=1}^{N+1} \frac{1}{2}[P_i(S \mid n) - P_{i-1}(S \mid n)][P_i(S \mid s) + P_{i-1}(S \mid s)]. \tag{3}$$

Here, N is the number of unique values of $P(S \mid n)$, and $P_i(S \mid n) \geq P_{i-1}(S \mid n)$. When $i = 1$, $P_{i-1}(S/n) = P_{i-1}(S \mid s) = 0$, and when $i = N+1$, $P_i(S \mid n) = P_i(S \mid s) = 1$.

Once an optimal value of b was selected, an optimal value of α was chosen for that b by maximizing the percentage of correct responses by maximizing

$$E[V] = P(S \mid s) - \beta P(S \mid n), \tag{4}$$

where

$$\beta = \frac{P(n)}{P(s)}. \tag{5}$$

$P(s)$ is the a priori probability of a signal, and $P(n)$ is the a priori probability of no signal. The calculations were done by incrementing α from 0.01 to 1.00 in steps of 0.01 and calculating $E[V]$ for each value of α. An example of $E[V]$ as a function of α is shown in Fig. 5.

Table 1
Algorithm parameters

Hormone	*a*	*b*	*d*	*k*
Initial optimization and test				
LH	0	6.1	12	12
FSH	0	6.5	12	12
$E_1$3G	0	3.0	6	12
Pd3G	0	1.1	12	12
$E_1$3G/Pd3G	0	3.5	12	12
Normal	–	–	–	12
Re-optimization				
LH	0	6.1	12	12
FSH	0	5.0	13	13
$E_1$3G	0	3.3	5	12
Pd3G	0	2.0	12	12
$E_1$3G/Pd3G	0	3.4	12	12
Normal	–	–	–	12

In practice, there was always a range of values of b which produced an optimal $P(A)$, so a value of b was selected in the optimal range such that $E(V)$ was at a maximum in a region surrounding $\alpha = 0.50$.

The fuzzy set *Normal* was optimized after the value of b was determined for each of the hormones. The optimal α for *Normal* was determined by varying α from 0.01 to 1.00 by 0.01 and calculating $E[V]$. An α in or near the middle of the range of values that produced the maximum $E[V]$ was selected.

4. Results of optimization and testing

Initially, data from 18 women were used to optimize the algorithm. Only 10 of these women had data for FSH. The length of the measurement sequences for individual women ranged from 22 to 149 days. Table 1 shows the algorithm parameters after the first optimization. Table 2 shows the performance of the algorithm.

A second set of data was used to test the algorithm. This data consisted of the hormone profiles from 70 women. The length of the measurement sequences for individual women ranged from 8 to 69 days. The performance of the algorithm on the test data is summarized in Table 2.

Table 2
Algorithm performance

Hormone	$P(A)$	Optimal α	$P(S \mid s)$	$P(S \mid n)$	β	$E[V]$
Initial optimization						
LH	1.0000	[0.50,0.51]	1.0000	0.0000	24.37	1.0000
FSH	0.9894	[0.49,0.50]	0.7500	0.0120	62.75	0.0000
$E_1$3G	0.9782	[0.45,0.56]	1.0000	0.0041	29.44	0.8800
Pd3G	1.0000	[0.42,0.54]	1.0000	0.0014	30.58	0.9583
$E_1$3G/Pd3G	1.0000	[0.47,0.52]	1.0000	0.0000	29.04	1.0000
Normal	–	[0.01,0.32]	1.0000	0.0000	59.88	1.0000
Test						
LH	1.0000	–	0.9718	0.0004	31.22	0.9578
FSH	0.9985	–	0.7627	0.0036	37.78	0.6271
$E_1$3G	0.9840	–	0.9559	0.0072	32.65	0.7206
Pd3G	0.9989	–	1.0000	0.0027	24.85	0.9318
$E_1$3G/Pd3G	0.9992	–	1.0000	0.0036	32.46	0.8824
Normal	–	–	0.9800	0.0005	41.74	0.9600
Re-optimization						
LH	1.0000	[0.50,0.50]	0.9802	0.0003	29.19	0.9703
FHS	0.9976	[0.49,0.50]	0.9552	0.0066	40.76	0.6866
$E_1$3G	0.9880	[0.50,0.51]	0.9677	0.0051	31.78	0.8064
Pd3G	0.9997	[0.45,0.54]	0.9911	0.0017	26.08	0.9464
$E_1$3G/Pd3G	0.9994	[0.48,0.52]	1.0000	0.0026	31.46	0.9167
Normal	–	[0.01,0.40]	1.0000	0.0000	44.24	1.0000

After testing, the algorithm was re-optimized using all the data. The performance of the algorithm after re-optimization is shown in Table 2. The algorithm parameters are shown in Table 1.

Note that the best estimates of the algorithm's performance, and the main results of the study, are the results of the test that occurred between the optimization and the re-optimization. The performance of the algorithm, as indicated by $E[V]$, decreased from the initial optimization to the test (with the exception of FSH), and then increased slightly after the re-optimization.

5. Discussion

The algorithm described here can detect hormone peaks and normal hormone cycles defined in terms of the peaks. It can be applied to large data sets where it would be tedious for an individual to go through the data looking for peaks or normal cycles and determining on what day they occurred. The algorithm will work for a series of measurements of arbitrary length and arbitrary starting point in the menstrual cycle. It will work in the presence of missing data, even if there are large gaps in the sequences of measurements. Variations of the algorithm can be used for blood measurements, different assay methods, or to detect different characteristics of hormone profiles. The methods used here can be applied to other types of physiological measurements and combinations of different types of measurements.

The current version of the algorithm is written in SAS® code and is not interactive. Ideally, a user would interact with the algorithm and the algorithm would learn from user inputs. For a given hormone profile, the algorithm would identify the peaks and normal cycles, the user would correct the algorithm's output, and, after one or a series of hormone profiles, the algorithm would re-optimize.

In a fuzzy system, crisp inputs are transformed into fuzzy values, operations are preformed on the fuzzy values resulting in fuzzy outputs, and the fuzzy outputs are converted to crisp outputs. The first step is called fuzzification and the last step is called defuzzification. In many applications the desired output is a single, crisp value from a continuous scale. One way of achieving this is to take the weighted sum of the fuzzy outputs of the system. In our application a discrete, binary output, i.e. 'peak' or 'not a peak', or 'normal cycle' or 'not a normal cycle', was determined by using a value of α to transform a fuzzy set into two mutually exclusive crisp sets.

The algorithm also included an operation on a combination of fuzzy and non-fuzzy values, i.e. the intersection of *Highest* and *High*, and it used fuzzy set theory and probability theory together in a compatible way. These theories are concerned with different types of uncertainty. The former concerns itself with the vagueness and imprecision of meaning, while the latter, as used here, concerns itself with the frequency with which random events occur. The theories are often viewed as being antagonistic and incompatible.

The performance of the algorithm compares favorably to the performance of other methods that have been used to detect features of the human menstrual cycle. Royston and Abrams [7] used a cumulative sum test to identify an increase in basal body temperature that occurs near the middle of the menstrual cycle. They report a true positive rate of 100% based on 137 cycles from 21 women. Schiphorst, Collins and Royston [8] applied a cumulative sum test to estrone glucuronide measurements made in urine samples in order to identify the limits of the fertile period and the time of maximum conception probability. They reported true positive rates of 89% and 82%, respectively. An algorithm based on numerical rules has been used to identify the day of luteal transition, which can be defined as the day after the peak of the urinary $E_1$3G to Pd3G ratio. Two studies [4,9] have reported a true positive rate of 88%. Another algorithm based on numerical rules was used to identify anovulatory cycles using measurements of urinary Pd3G [10]. This algorithm had a sensitivity (true positive rate) of 75% and a specificity (true negative rate) of 89.5% or 92.2% depending on which variation of the algorithm was used.

Of the individual hormones, the LH peak is the most reliably detected, followed by Pd3G, and then $E_1$3G and FSH. FSH profiles often have peaks near the end of menstruation which can cause the algorithm to make errors. $E_1$3G profiles sometimes have large peaks during the second half (luteal phase) of the menstrual cycle. The $E_1$3G/Pd3G peaks were detected less reliably than the LH and Pd3G peaks and more reliably than the $E_1$3G and FSH peaks. Normal cycles were detected most reliably of all because information was used from all the hormones. These results are dependent on the type of measurements being made, i.e. in urine, and the particular assays used to make the measurements.

Once the peaks are identified and the days of the peaks are known, one can estimate the day of ovulation and calculate the quantities of hormones produced in certain time windows and the relative timing of peaks. These values can be used to characterize the female reproductive system and her fecundity. To determine which hormone or combination of hormones is best for determining the day of ovulation would require a standard indicator for ovulation.

6. Summary

This paper describes an algorithm that identifies qualitative endocrine features of the human menstrual cycle. The algorithm detects hormone peaks and normal cycles. It was developed and tested using urinary values for luteinizing hormone, follicle stimulating hormone, estrone 3-glucuronide ($E_1$3G), and pregnanediol 3-glucuronide (Pd3G). It also detects the peak of the $E_1$3G to Pd3G ratio, which is used to identify the day of transition from follicular production of estrogen to luteal production of progesterone. Normal cycles are detected by combining information from the individual hormones.

Fuzzy sets are used to identify features in the hormone profiles. Signal detection theory was used to optimize performance. The algorithm can be applied to large data sets where it would be tedious for an individual to go through the data looking for peaks or normal cycles and determining on what day they occurred. The methods described here can be applied to other types of physiological measurements and combinations of different types of measurements. The algorithm can be re-optimized after new data is processed.

Acknowledgements

The authors wish to thank Stanley Shulman (NIOSH, USA) and Anca Ralescu (University of Cincinnati, USA) for critical reviews of the manuscript. Allen Wilcox (National Institute of Environmental Health Sciences, USA), and Barbara Grajewski and Elizabeth Whelan (NIOSH, USA) contributed urine samples used to develop and evaluate the algorithm. The authors are also grateful for the generous gifts of immunoassay reagents used to generate the data analyzed in this report: $E_1$3G monoclonal antibody from Dr Fortüne Kohen (Weizmann Institute of Science, Israel), Pd3G monoclonal antibody from Dr Mohamed Gani (Unilever Research Colworth Laboratory, UK), and europium-labeled $E_1$3G and Pd3G from Dr Heikki Mikola (Wallac Oy, Finland).

References

[1] J.S. Kesner, E.A. Knecht, E.F. Krieg Jr, Time-resolved immunoflourometric assays for urinary luteinizing hormone and follicle stimulating hormone, Analytica Chimica Acta 285 (1994) 13–22.
[2] J.S. Kesner, E.A. Knecht, E.F. Krieg Jr, G. Barnard, J.H. Mikola, F. Kohen, M.M. Gani, J. Coley, Validation of time-resolved flouroimmunoassays for urinary estrone 3-glucuronide and pregnanediol 3-glucuronide, Steriods 59 (1994) 205–211.
[3] H.H. Taussky, A microcolorimetric determination of creatine in urine by the Jaffe reaction, Journal of Biological Chemistry 208 (1954) 853–861.
[4] D.D. Baird, C.R. Weinberg, A.J. Wilcox, D.R. McConnaughey, P.I. Musey, Using the ratio of urinary oestrogen and progesterone metabolites to estimate day of ovulation, Statistics in Medicine 10 (1991) 255–266.
[5] G.J. Klir, T.A. Folger, Fuzzy sets, uncertainty, and information, Prentice Hall, Englewood Cliffs, New Jersey, 1988.
[6] D.M. Green, J.A. Swets, Signal detection theory and psychophysics, John Wiley & Sons, New York, 1966.
[7] J.P. Royston, R.M. Abrams, An objective method for detecting the shift in basal body temperature in women, Biometrics 36 (1980) 217–224.
[8] L.E.M. Schiphorst, W.P. Collins, J.P. Royston, An estrogen test to determine the times of potential fertility in women, Fertility and Sterility 44 (1985) 328–334.
[9] D.D. Baird, D.R. McConnaughey, C.R. Weinberg, P.I. Musey, D.C. Collins, J.S. Kesner, E.A. Knecht, A.J. Wilcox, Application of a method for estimating day of ovulation using urinary estrogen and progesterone metabolites, Epidemiology 6 (1995) 547–550.
[10] A. Kassam, J.W. Overstreet, C. Snow-Harter, M.J. De Souza, E.B. Gold, B.L. Lasley, Identification of anovulation and transient luteal function using a urinary pregnanediol-3-glucuronide ratio algorithm, Environmental Health Perspectives 104 (1996) 408–413.

Edward F. Krieg, Jr. is married to a woman from Peru and has two children who are five years old. He was born in Albany, New York in 1956 and moved many times because his father worked for IBM. He received a bachelor's degree from the University of Bridgeport in Bridgeport, Connecticut, and master's and doctoral degrees from Tulane University in New Orleans, Louisiana, all in psychology. Since 1987, he has been a statistician at the Division of Biomedical and Behavioral Science, National Institute for Occupational Safety and Health in Cincinnati, Ohio. He does research in quantitative methods in biology and psychology, and likes to visit his relatives in South America.

James S. Kesner was born and raised on a family farm. He received his bachelor's degree in Agricultural Science from the University of Illinois. His doctorate dissertation at Michigan State University, under the guidance of Edward Convey, described the endocrine control of ovulation in cattle. Post-graduate research in the laboratories of Ernst Knobil at the University of Pittsburgh and University of Texas described the electrical activity in primate hypothalamic centers which controls reproduction. At USDA, he established procedures for studying reproductive neuroendocrinology in pigs. He currently directs research at NIOSH to assess the effects of occupational hazards on the reproductive health of women.

Edwin A. Knecht is a native of northern Kentucky, and earned both bachelor's and master's degrees from the University of Cincinnati. He joined the Commissioned Corps of the United States Public Health Service in 1974, and has worked in a number of laboratory assignments with the National Institute for Occupational Safety and Health. For the past nine years, he has managed the Institute's Reproductive Endocrinology Laboratory, and participated in collaborative field studies investigating the effects of potential occupational hazards on the reproductive health of America's working men and women. Ed has been married to his wife, Marilyn, for 23 years and they have three grown daughters.