Characterizing Daily Urinary Hormone Profiles for Women at Midlife Using Functional Data Analysis

Peter M. Meyer1, Scott L. Zeger2, Siobán D. Harlow3, MaryFran Sowers3, Sybil Crawford4, Judith L. Luborsky5, Imke Janssen1, Daniel S. McConnell3, John F. Randolph6, and Gerson Weiss7

1 Department of Preventive Medicine, Rush University Medical Center, Chicago, IL.
2 Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD.
3 Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI.
4 Division of Preventive and Behavioral Medicine, University of Massachusetts Medical School, Worcester, MA.
5 Department of Pharmacology, Rush University Medical Center, Chicago, IL.
6 Department of Obstetrics and Gynecology, University of Michigan Medical School, Ann Arbor, MI.
7 Department of Obstetrics and Gynecology, New Jersey Medical School, Newark, NJ.

Received for publication January 14, 2005; accepted for publication October 9, 2006.

The availability of daily hormone values for entire menstrual cycles offers an opportunity to apply new analytic techniques that confirm current knowledge and provide new insights into patterns of changing hormone profiles in women as they transition to the menopause. The Study of Women’s Health Across the Nation (SWAN) collected urine samples during 1997–1999 from one menstrual cycle or up to 50 days from 848 women who live in seven cities across the United States. These samples were assayed for the urinary forms of estrogen, progesterone, follicle-stimulating hormone, and luteinizing hormone. The authors used functional data analysis to study variability in the hormone patterns of 572 of the 848 pre- and early-perimenopausal women with evidence of a luteal transition. Functional data analysis enabled the authors to identify asymmetries in women’s hormone patterns related to cycle length that are not captured with single hormone value comparisons. Longer cycles were characterized by increasing dyssynchrony between follicle-stimulating hormone and luteinizing hormone in the luteal phase.

hormones; menopause; menstrual cycle; principal component analysis

Abbreviations: FSH, follicle-stimulating hormone; PdG, pregnanediol-glucuronide; SWAN, Study of Women’s Health Across the Nation.

Female reproductive hormone concentrations vary cyclically throughout each menstrual cycle (1). As women near the end of their reproductive life, variability in these cyclic hormonal patterns increases (2). Statistical methods for analyzing hormonal patterns are needed that enable characterization of the variability in these nonlinear cyclic patterns and comparison across population subgroups.

The traditional approach to evaluation of reproductive capacity has been to measure serum follicle-stimulating hormone (FSH) during the early follicular phase (days 2–5) of a regular menstrual cycle (3, 4). Measurements from samples collected in the early follicular phase have been shown to provide a reliable estimate of ovarian reserve and fertility potential as women age (4, 5) and to assess reproductive capacity in population studies of menopause (6–8). Increasing levels of FSH in the early follicular period are associated clinically with progression toward ovarian failure and reduced ovarian feedback suppression of pituitary FSH (5, 9).
Single measurements are a limited proxy for a cyclic process. Early follicular phase samples do not reflect the amount and duration of a woman’s exposure to steroid hormones throughout the menstrual cycle (10). These limitations have led several investigators to adopt use of daily first-morning urine samples and to analyze patterns of daily hormone concentrations. Previous studies have used daily hormonal data to develop and validate algorithms to estimate the likelihood of ovulation, its date of occurrence, and the length of follicular and luteal phases (11–13).

Some studies have graphically displayed the daily hormone values, centering at the peak day of the luteinizing hormone surge and calculating the mean or geometric mean daily value for the study sample (5, 11, 14–17). This approach provides a useful visual assessment but does not evaluate within-woman patterns or readily permit formal testing of differences in hormonal profiles across groups. To test for group differences in hormone concentrations, most studies have calculated a summary score by estimating the mean hormone value or area under the curve for an individual woman for the total cycle, for cycle phases, or for critical time windows within the cycle (5, 11, 15, 17–19), while others calculated rates of change for specific intervals (16).

Regression approaches have been used to characterize cyclic hormone profiles more fully and to permit more formal population comparisons. Models based on splines have been used to compare hormone patterns in women with and without low peak bone mass (20). Bent stick regression models have been used to model the estrogen rise in the follicular phase (21). Zhang et al. (22, 23) have proposed a semiparametric mixed model that incorporates a periodic variance function and a smoothing function. These models include covariates as part of a linear predictor and, as such, model associations on a global scale.

This paper uses a functional data analysis approach to model hormonal patterns across the menstrual cycle, a methodology whose basic theory and methods have been summarized elsewhere (24, 25). We explore modes of variation by which cycles in our sample vary from the typical pre- and early perimenopausal patterns. These modes of variation are displayed as curves, each of which captures some aspect of typical departure from average cyclic behavior.

Evidence is mounting that cycle length changes are associated with characteristic changes in hormones as a woman enters and progresses through the menopausal transition. Gracia et al. (26) found differences among premenopausal women related to changes in cycle length in a comparison of the Study of Women’s Health Across the Nation (SWAN) and Stages of Reproductive Aging Workshop (STRAW) definitions of stages of the transition. Miro et al. (27, 28) also explored the relation between cycle length and hormone levels. Weinstein et al. (29) used approximate entropy to quantify variability in serially observed cycle lengths in data from The Tremin Trust (a databank containing information on menstrual cycles for more than 6,000 women older than age 65 years) and found it to be a strong predictor of the start of the menopausal transition.

The SWAN Daily Hormone substudy offers a unique opportunity to study patterns of four urinary hormones measured in samples collected daily during a menstrual cycle from women from five ethnic groups as they move through the menopausal transition. We use functional analysis to explore how differences in cycle length are associated with differences in the shapes of hormone patterns among pre- and early perimenopausal women. Functional analysis enables us to express individual women’s data as curves, retaining the ability to describe differences in specific time windows within the cycle while providing a framework to describe variation in shape over the entire cycle.

**MATERIALS AND METHODS**

SWAN is a multiethnic cohort study of 3,302 middle-aged women enrolled at seven sites throughout the United States; the study design has been reported previously (30). The Daily Hormone Study is a SWAN substudy that enrolled 848 women to collect first-morning-void urine samples daily for one complete menstrual cycle or 50 days (whichever comes first) annually, as described elsewhere (13). From the first annual collection (1997–1999), we analyzed levels of urinary FSH, urinary luteinizing hormone, estrone conjugates, and pregnanediol-glucuronide (PdG).

This study was approved by all of the sites’ institutional review boards, and written informed consent was obtained from each participant. The Daily Hormone Study sample was drawn from the parent SWAN baseline cohort of 1,550 Caucasian, 935 African-American, 250 Chinese, 281 Japanese, and 286 Hispanic women aged 42–52 years. Inclusion criteria for recruitment into the Daily Hormone Study were 1) an intact uterus and at least one ovary, 2) at least one menstrual period in the previous 3 months, 3) no use of exogenous sex steroid hormones in the 3 months prior to initial urine collection, and 4) not pregnant.

The reproductive status of all women in SWAN is expressed categorically as “menopausal status.” This definition is based on currently accepted nomenclature (3, 31). Premenopausal status was defined as menses in the past 3 months with no change over the past year in regularity of menstrual periods. Early perimenopausal status was defined as menses in the past 3 months with less predictable menstrual periods (3, 31). Menopausal status was assessed at the annual visit preceding entry into the Daily Hormone Study.

**Hormone assays**

Urinary luteinizing hormone, urinary FSH, estrone conjugates, and PdG were measured by using assays adapted to chemiluminescence technology on the ACS-180 Automated Chemiluminescence Analyzer (Bayer Diagnostics Corporation, Tarrytown, New York) (13). The assayed values were normalized for urinary creatinine excretion (32).

We applied validated algorithms (33, 34) to evaluate the menstrual cycles for features consistent with ovulation. The five nadir days of PdG in the follicular phase were identified. A threefold increase in PdG concentrations above this nadir for at least 3 consecutive days was considered evidence of luteal activity or presumptive ovulation. For cycles with evidence of luteal activity, another algorithm was used to determine the probable day of luteal transition, using a
modification of the method by Waller et al. (12) and Santoro et al. (13) that identified when the ratio of estrone conjugates to PdG decreased by 60 percent. To determine day of luteal transition, hormone concentrations for isolated missing days were interpolated (12). The goal of this analysis was to characterize ovulatory cycles; thus, cycles with no evidence of luteal activity were excluded ($n = 198$).

Statistical methods

We used a functional data analysis approach to model hormonal patterns across the menstrual cycle. Each cycle of daily hormone values is approximated by a weighted sum of a set of basis curves. Each basis curve is positive and non-zero over only a small portion of the cycle. As a result, the weights can be thought of as reflecting short time windows of the hormone pattern, and the overall pattern is approximated by the weighted sum of all basis functions. This approach is similar to the one of Zhang (22, 23) but without the shrinkage implicit in a mixed-model formulation. Functional principal components applies the traditional principal components methodology (35) to the weights associated with the basis curves to determine the most common ways in which the patterns vary from the average. In this paper, we explore modes of variation by which cycles in our sample vary from the typical pre- and early perimenopausal patterns and display them graphically.

Missing values, registration, and transformations. Luteinizing hormone and FSH values below the assay’s level of detection were imputed. Analyses were performed on the log$_{10}$ scale so that the random variability in the data would be more normally distributed. We begin with the daily observations on cycles of varying length (figure 1, step 1). The first day was omitted because many women whose bleeding began later in the day missed the first day urine collection ($185/572 = 32$ percent). Of the characteristics given in table 1, only cycle length (26.4 days with and 27.3 days without, $p = 0.004$) and timing of luteal transition (13.9 with and 14.8 without, $p = 0.003$) differed between women with and without a first day sample, as would be expected. Menstrual cycle days were rescaled so that the second day corresponded to time 0 and the last day to time 1 (step 2). Registration is the process of stretching or shrinking the time scale in the luteal and follicular phases so that each cycle’s day of ovulation corresponds to time 0.5 (step 3). This process was accomplished by using a family of two parameter exponential curves.

Study sample and training subsample. To ensure stability when estimating weights for the curves, we selected cycles with nonmissing time = 0 and time = 1 observations and with only isolated values missing during the cycle ($n = 572$). We further selected a subset of women to use as a training subsample, whose cycles were of typical length (22–34 days), who were not diabetic or did not have thyroid problems, who had no more than two missing values total, and who had 11 or more observations both before and after the day of ovulation ($n = 235$). The training subsample was used 1) to establish the appropriate amount of smoothing to use when determining the weights for the entire sample and 2) to calculate the principal components curves. This process provided a robust estimate of the amount of smoothing and the principal component curves, much the same as...
The differences in the patterns between premenopausal and early perimenopausal cycles were small. To demonstrate the utility of the approach in comparing hormone patterns across groups, we show differences associated with cycle length. Adjustments for covariates can be made in a natural way, extending analysis of variance to include functional data. We chose not to do so here because differences by race, menopausal status, and age were small and we wanted to avoid the additional layer of complexity that analysis of variance would introduce. Those interested can find examples of functional analysis of variance in the studies by Ramsay and Silverman (24, 25).

Analyses were conducted with S-Plus 6.2 (Insightful Corporation, Seattle, Washington) and R2.1.1 (R Foundation for Statistical Computing, Vienna, Austria) software. Some useful S-Plus functional data analysis functions from Jim Ramsay’s website were used in the analyses (available in zip-file format at http://ego.psych.mcgill.ca/pub/ramsay/FDAfuns/SPLUS/fdaS.zip).

RESULTS

Table 1 summarizes the characteristics of the study sample and the training subsample. The study sample and training subsample were similar except for cycle length variability.

Figure 2 shows mean curves by cycle length groups. Women with the shortest (18–23 days) and the longest (33–38 days) menstrual cycles differed visually from the other groups in a variety of ways across the four hormones.
The group with a cycle length of 33–38 days showed early rises in urinary FSH and urinary luteinizing hormone unaccompanied by an expected parallel change in estrone conjugates, suggesting the presence of an asynchrony between hormone values in the early follicular phase of long cycles.

Figure 3 displays the correlation of basis function weights for urinary luteinizing hormone and urinary FSH by cycle length groups. When urinary luteinizing hormone and urinary FSH are in synchrony, correlations are near 1 and the corresponding block is black. For typical-length cycles (24–32 days), we see a clear black line from upper left to lower right, indicating synchrony between urinary luteinizing hormone and urinary FSH throughout the cycle. Longer cycles display an increasing loss of synchrony across the entire cycle. To determine whether these differences might be explained by menopausal status, we examined the difference in mean cycle length between pre- and early-menopausal women. For the 551 cycles of length 18–38 days, the difference in mean cycle length between pre- and early-menopausal women is only 0.28 days, which approaches only marginal significance despite the large sample size ($p = 0.07$).

Functional principal components, displayed in figure 4, enable us to further examine modes of variation in the hormone patterns. We found three principal components adequate to capture the variability in each of the four hormones studied. The percentage of variability associated with a principal component is included in the subtitle for its plot. Each plot shows the mean curve for the training subsample (black line) and the amount of variation associated with a positive (“+” line) or negative (“−” line) one standard deviation change in the principal component scores of the respective principal components. As is common, the first principal component for each of the four hormones in our sample reflected differences in mean hormone levels across women, with a positive score associated with an increase and a negative score associated with a decrease in mean hormone level. For each hormone, differences in mean levels across the cycle accounted for more than half of the variability.

The second principal components for all but urinary luteinizing hormone reflected an asymmetry between the follicular and luteal phases of the cycle. For estrone conjugates, for example, a positive score is associated with relatively higher than average estrone conjugates values in the follicular phase and relatively lower values in the luteal phase, resulting in a pattern with a single rise. Negative scores were associated with an opposite shift in pattern, resulting in a clear second rise during the luteal phase. For urinary luteinizing hormone, the second largest component of variation was associated with a shift in the day of the midcycle surge.

For estrone conjugates and urinary luteinizing hormone, the third component was associated with either a blunted or an exaggerated midcycle surge. For urinary FSH, the third principal component—similar to the second principal component for urinary luteinizing hormone—represented a shift in timing of the surge. The total variability captured by the first three principal components (refer to the figures) was highest for PdG at 86.9 percent and roughly the same for the other three: estrone conjugates, 68.6 percent; urinary FSH, 67.8 percent; and urinary luteinizing hormone, 66.7 percent.

*FIGURE 2. Mean curves by cycle length groups (number of days) for estrone conjugates (E1C; ng/mg creatine), pregnanediol-glucuronide (PdG; μg/mg creatine), urinary follicle-stimulating hormone (uFSH; mIU/mg creatine), and urinary luteinizing hormone (uLH; mIU/mg creatine). Curves were fit by using a penalized cubic B-spline basis with 20 basis functions. Data are from the Study of Women’s Health Across the Nation (SWAN) Daily Hormone Study baseline collection (1997–1999). DLT, day of luteal transition.*
Principal components can be used to compare and test differences in modes of variation across groups. As with the familiar principal components (35), functional principal components associate a score for each principal component with each cycle. In functional principal components, this score represents the extent to which the pattern reflected by the principal component is present in an individual cycle. By comparing scores across groups, we can determine whether group differences are related to differences in specific modes of variation. Figure 5 displays box plots of the first three principal component scores for the four hormones, broken down by cycle length groups. In each of the 12 subplots, the principal component score differences among the cycle length groups are compared by using a Kruskal-Wallis test. Even with a strict Bonferroni adjustment for multiple comparisons (0.05/12 = 0.0042), scores for seven of the 12 principal components vary by cycle length group (estrone conjugates: 1, 2; PdG: 2, 3; FSH: 2, 3; luteinizing hormone: 3).

**DISCUSSION**

The current study displays the utility of modeling cycles of daily hormone concentrations as functional data curves, providing a natural characterization of both the overall cyclic hormonal pattern and patterns during specific time windows within the cycle. This paper also identifies principal modes of variation, providing comparisons of these modes across cycle-length subgroups. Most previous work has focused on assessing differences in hormone levels at specified times in the menstrual cycle or on summary hormonal characteristics. Using such techniques, prior studies, including a previous report from SWAN (36), have shown that early follicular sex steroid and urinary FSH levels vary by

---

**FIGURE 3.** Correlation of basis function weights for log urinary luteinizing hormone (uLH) and log follicle-stimulating hormone (uFSH) by cycle length groups. Axes represent the 20 basis functions for each hormone, and the shading indicates the correlation between coefficients for any pair of basis functions. Data are from the Study of Women's Health Across the Nation (SWAN) Daily Hormone Study baseline collection (1997–1999).
ethnicity and age. This paper enables us to further our understanding by identifying how hormone synchrony varies over the course of a cycle (urinary luteinizing hormone and urinary FSH) and by identifying typical modes of variation,
that is, how the shapes of hormone patterns vary across the population of cycles.

As an example, we demonstrated variations in the shape of hormone curves by cycle length. Shorter cycles are

FIGURE 5. Box plots (without outliers) of principal component (PC) scores by cycle length groups. Boxes show the interquartile range (IQR), and lines extend beyond the boxes to the most extreme values that are still within 1.5 IQR of the ends of the boxes. The $p$ value is from a Kruskal-Wallis test. The urinary hormones are estrone conjugates (E1C), pregnanediol-glucuronide (PdG), urinary follicle-stimulating hormone (uFSH), and urinary luteinizing hormone (uLH). Sample sizes of the cycle length groups are 18–20 days ($n = 18$), 21–23 days ($n = 63$), 24–26 days ($n = 197$), 27–29 days ($n = 182$), 30–32 days ($n = 54$), 33–35 days ($n = 30$), and 36–38 days ($n = 11$). Data are from the Study of Women’s Health Across the Nation (SWAN) Daily Hormone Study baseline collection (1997–1999).

Am J Epidemiol 2007;165:936–945
associated with higher estrone conjugates levels (principal component 1), suggesting that these cycles are early enough in the transition that estrone conjugates levels have not yet begun to drop. Longer cycles tend to show a second rise in estrone conjugates, while shorter cycles do not (principal component 2). Short cycles tend to have later PdG rises (principal component 3) that are smaller in magnitude (principal component 2). Short cycles show a higher peak FSH level compared with the start of the cycle (principal component 2), which occurs earlier than for cycles of typical and longer length (principal component 3). Shorter cycles also have relatively higher urinary luteinizing hormone peaks at the midcycle surge (principal component 3).

Functional data analysis provides insight into hormone behavior at specific times within the cycle through the weights associated with the basis functions. Because we used the same number of basis functions and the same time registration for all of the hormones, the correlations between the weights provide a measure of the synchrony or dysynchrony between the hormones during specific time windows.

Principal components analysis of functional data provides a tool for describing modes of variation in cyclic patterns. This tool validates earlier techniques based on means while yielding new insights. The first principal components reflected mean differences in hormone levels across cycles. Reliable estimates of mean hormone levels capture this major source of variation. However, our technique goes beyond mean structure and enables us to identify additional modes of variation. Two cycles may have the same overall mean but differ considerably in hormone levels during the middle of the follicular and luteal phases of the cycle. The second principal component for estrone conjugates, PdG, and urinary FSH and the third for urinary luteinizing hormone captured an asymmetry that is very different from the mean pattern. Particularly for estrone conjugates and PdG, where this mode of variation reflects over 10 percent of the variability, comparisons of midfollicular or midluteal measurements could lead to incorrect conclusions, which confirms the practice of sampling early in the cycle for pre- and early perimenopausal women, when asymmetric modes have less impact.

There are limitations to the current study. Cycles vary within woman, so studying one cycle per woman provides no measure of within-person variability. In addition, it is not length of cycle per se but change in cycle length that predicts stage of the transition (26, 37). We used a simple approach to variance, treating it as constant across the cycle. Since all women were either premenopausal or early perimenopausal, differences in modes of variation in this sample were small. The standardization of cycle length and the registration of the cycles at day of luteal transition have the greatest impact on the shortest and longest cycles. Because the mean curve for the shortest cycles (15–17 days) is very different from longer cycles (39–49 days) and both differ from cycles of length from 18 to 38 days, comparisons of modes for these most extremes of cycle length were not included. Other choices of normalization and registration are possible. These were chosen because they preserve the meaning of the day of luteal transition across cycles of various lengths.

Functional data analysis is a useful extension of past approaches to understanding urinary hormone patterns that offers new insights into cyclic hormone patterns, providing summary curves for individual cycles, giving useful descriptions of typical modes of variation, and enabling examination of both overall patterns and comparisons within specific time windows. The Daily Hormone Study continues to gather samples of urinary hormone cycles annually from the same women. Subsequent analyses will use these techniques to explore the associations between hormone patterns, menopausal symptoms, and risk factors as women progress through the menopausal transition.

ACKNOWLEDGMENTS

Dr. Meyer’s effort is funded by the National Heart, Lung, and Blood Institute (K25 HL68139-01A1).

The authors express their appreciation to J. Ramsay and B. Silverman for making their code for working with functional data objects available through the Internet.

SWAN was funded by the National Institutes of Health, Department of Health and Human Services, through the National Institute on Aging, the National Institute of Nursing Research, and the National Institutes of Health Office of Research on Women’s Health (grants NR004061; AG012505, AG012535, AG012531, AG012539, AG012546, AG012553, AG012554, AG012495).


Conflict of interest: none declared.

REFERENCES