Menopause: The Journal of The North American Menopause Society
Vol. 12, No. 5, pp. 567-577
DOI: 10.1097/01.gme.0000172265.40196.86
© 2005 The North American Menopause Society
♾ Text printed on acid-free paper.

Monitoring reproductive aging in a 5-year prospective study: aggregate and individual changes in steroid hormones and menstrual cycle lengths with age

Rebecca J. Ferrell, PhD,[1] Kathleen A. O'Connor, PhD,[2] Germán Rodríguez, PhD,[3] Tristan Gorrindo, BS,[1] Darryl J. Holman, PhD,[2] Eleanor Brindle, MA,[2] Rebecca C. Miller, MA,[2] Deborah E. Schechter, MA,[2] Lauren Korshalla, BS,[1] James A. Simon, MD, PhD,[4] Phyllis K. Mansfield, PhD,[5] James W. Wood, PhD,[6] and Maxine Weinstein, PhD[1]

Abstract

Objective: We describe a 5-year prospective study of reproductive aging, and present analyses of steroid hormone and menstrual cycle changes with age.

Design: Participants were college-educated white women, primarily of northern European ancestry, recruited from the Tremin Research Program on Women's Health (n = 156, 25-58 years). In each of 5 consecutive years, they collected daily urine specimens for 6 months and recorded menstrual bleeds for all months. Urine specimens were assayed for estrone-3-glucuronide (E1G) and pregnanediol-3-glucuronide (PDG), urinary metabolites of estradiol and progesterone. Using multilevel models, we estimated hormone and cycle-length trajectories for individual women and within- and between-woman variance by age.

Results: At the aggregate level, PDG declined beginning in the 30s, E1G increased into the 40s before declining, and cycle length became more variable with age. Individual-level models revealed substantial hormonal variation across women, in both absolute levels and rates of change. Most women showed declining E1G by the late 40s, declining PDG in the 30s, and increasing mean cycle length in the 40s. Hormonal variation decreased with age; cycle length variation decreased and then increased. Within individual women, cycle lengths were highly variable while hormone levels were more stable. Women differed more from each other in hormone levels than for cycle lengths.

Conclusions: Aggregate-level analyses show general changes in steroid hormones and cycle length but cannot show variation within and across women. Individuals' cycle lengths were too variable to predict hormone levels. Clinicians should obtain more data on individual women's hormonal patterns when determining fertility or menopause treatments.

Key Words: Reproductive aging – Menopause – Menstrual cycle – Urinary steroid hormones – Estrone-3-glucuronide – Pregnanediol-3-glucuronide.

Received October 22, 2004; revised and accepted January 17, 2005.

[1]Center for Population and Health, Georgetown University, Washington, DC; [2]Department of Anthropology and Center for Studies in Demography and Ecology, University of Washington, Seattle, WA; [3]Office of Population Research, Princeton University, Princeton, NJ; [4]George Washington University School of Medicine, Washington, DC; [5]Department of Women's Studies and TREMIN Research Program on Women's Health, Pennsylvania State University, University Park, PA; and [6]Department of Anthropology and Population Research Institute, Pennsylvania State University, University Park, PA.

This study was supported by National Institutes of Health (Grants: NICHD 1R01HD034159, NIA 1R01AG015141, NIA 5T32AG000208, NICHD 2P30HD028263, and NICHD 5F32HD007994-02), the Population Research Institute, Pennsylvania State University, and the Center for Studies in Demography and Ecology, University of Washington.

Address correspondence to: Rebecca J. Ferrell, Center for Population and Health, 313 Healy Hall, Box 571197, Georgetown University, Washington, DC 20057-1197. E-mail: rft2@georgetown.edu.

Changes in hormonal and menstrual cycle patterns that occur as women age have not been fully characterized.[1,2] Data collected at infrequent intervals from large numbers of women,[3-9] or at frequent intervals from small numbers of women,[10-15] do not capture the full range of variation that exists within and across women. Describing this variation at both the individual and aggregate levels and for a wide range of reproductive and postreproductive ages remains an important goal for understanding when and how women make the transition from reproductive to postreproductive life.

The Biodemographic Models of Reproductive Aging (BIMORA) project was a 5-year prospective study of the hormonal and menstrual cycle characteristics of reproductive aging, particularly the menopausal transition. BIMORA participants were recruited from the TREMIN Research Program on Women's Health (TREMIN), an ongoing longitudinal study of menstruation and health for which participants keep a prospective record of menstruation and answer annual or biannual health questionnaires.[16-18] The TREMIN women were originally recruited from the University of Minnesota (cohort I from 1934-1939, n = 2,350; cohort II from 1961-1963, n = 1,367; continued recruitment after cohort II) and were college-educated, midwestern, white women, many of northern European ancestry. Women selected for the BIMORA project were either part of cohort II or in the pool of younger women recruited into TREMIN after cohort II.

BIMORA participants provided daily urine specimens and menstrual cycle data for 6-month intervals in each of the 5 years from 1998 through 2002; they continued to provide TREMIN with prospective menstrual and health data. We used these complementary sources of data to: (1) investigate how the aggregate distributions of steroid hormone levels vary with age and how menstrual cycle frequency varies with age; and (2) model age-related changes in steroid hormone levels and menstrual cycle lengths for individual BIMORA participants, and estimate age-specific within- and between-woman variance for observed hormone concentrations and menstrual cycle lengths.

METHODS

Eligibility and enrollment

Women who participated in TREMIN during 1997 or who had participated in the past up until menopause, and who were at least 18 years old in 1997, were initially identified to be contacted about participating in BIMORA. Of those women, 272 declined participation and 40 could not be contacted. The remaining 436 potentially eligible women were all younger than 60 years old and were sent a questionnaire about participating in BIMORA; 225 were eligible, 204 were using hormones, 5 were pregnant, and 2 were breastfeeding. Those who were using hormones or were pregnant or breastfeeding were considered for later recruitment if they became eligible. A total of 156 women ultimately participated in the BIMORA protocol.

Participants had at least one intact ovary, were not using exogenous prescription reproductive hormones (eg, oral contraceptives, hormone therapy, fertility drugs), and were not pregnant, breastfeeding, or receiving cancer treatment. Women who reported ever having fibroids and women who later had a hysterectomy and bilateral oophorectomy were included because excluding them did not improve model fit in our analyses. We used continuous enrollment to maintain participant numbers across the 5-year study: some women enrolled or re-enrolled after cessation of pregnancy, breastfeeding, or exogenous hormone use. Participants received compensation of $150 per year. All participants provided written informed consent, and the institutional review boards of Georgetown University, the University of Utah, the Pennsylvania State University, and the University of Washington approved all procedures.

Data collection

TREMIN data

BIMORA participants continued to record data for TREMIN in the form of calendar cards and health surveys. The women recorded information about menstrual bleeding and events such as surgeries, pregnancies, births, and menstrual or reproductive problems on the calendar cards on a daily basis. On the health surveys, they reported information on menstruation, reproduction, contraception and hormone use, surgeries and clinical procedures, illness, and medications.

BIMORA data

Each day, for a total of 6 months (January 15 to July 14), in each of the 5 study years from 1998 to 2002, women collected urine specimens and information on menstrual bleeding, prescription and over-the-counter supplement and drug use, and health conditions and medical procedures. During each month of participation, women received a collection kit containing a styrofoam cooler and coldpack, daily urine collection sponge vials, and a label record sheet with peel-off vial labels. Participants collected a first-morning urine specimen each day and placed specimens in a home

freezer. (Less than 1% of specimens were recorded as non-first-morning samples. These were "spot" collections, collected at whatever time of day was convenient for the participants.) Women were asked to list on the label record sheet, for each day, medication and supplement use, bleeding, major medical events, and how a specimen was treated if it was not frozen within an hour. Each month, subjects shipped the frozen specimens via overnight courier to the BIMORA laboratory.

Specimens were stored at −20°C for up to 3 months before processing. Aliquots were taken from thawed specimens, preserved with 17 mg/mL of boric acid, and refrigerated at 4°C until they were assayed. All urine specimens were assayed in duplicate for estrone-3-glucuronide (E1G), a urinary metabolite of estradiol, and pregnanediol-3-glucuronide (PDG), a urinary metabolite of progesterone. These metabolites closely parallel the serum levels of estradiol and progesterone.[19] Enzyme immunoassays (EIAs) using monoclonal capture antibodies were used to quantify the urinary levels of E1G and PDG.[20,21] Inter- and intra-assay coefficients of variation were 9.2% and 10.3%, respectively, for the PDG EIA, and 4% and 3.6%, respectively, for the E1G EIA. The PDG EIA cross reacts 100% with PDG, 187% with 20α hydroxy-4-pregnen-3-one, 13.4% with pregnanediol, 4.3% with 20β hydroxy-4-pregnen-3-one, and less than 1% with other progestins.[20] The E1G EIA cross-reacts 100% with E1G, 83% with estradiol-3-glucuronide, and less than 5% with other estrogens.[21]

Hormone concentrations were estimated from optical density (Dynatech MR7000 MicroPlate Reader, test wavelength 405 nm, reference 570 nm) using a four-parameter logistic model[22] in Biolinx 1.0 software (Dynex Laboratories, Inc., Chantilly, VA). All concentrations were corrected for hydration status using specific gravity,[23] measured with a handheld urine-specific gravity refractometer (Uricon-PN, NSG Precision Cells, Inc., Farmingdale, NY). E1G concentrations were corrected statistically for slight assay nonparallelism, using a 1:5 dilution as the standard to which all values were corrected.[21]

We excluded data collected during, and in the 3 months after, any event reported to TREMIN or BIMORA that was known to affect menstruation or hormone levels including (1) medications or supplements, (2) pregnancy or breastfeeding, and (3) major medical diagnoses, procedures, or treatments. Project staff coded all medications, supplements, and medical events according to their documented or suspected effects on the ovarian cycle, menstrual bleeding, or endogenous reproductive hormone levels.

Variable definitions

Bleed Episode

Following the World Health Organization's definition[24] as modified by Harlow et al,[25] a bleed episode consisted of at least 2 days of bleeding in a 3-day interval (ie, 3 consecutive days of bleeding or a nonbleed day between 2 bleed days), and was preceded by at least 2 bleed-free days.

Cycle Length

Cycle day "1" was the first day of a bleed episode. Cycle length was the number of days from day 1 of one bleed episode through the day before day 1 of the next bleed episode. Data for incomplete menstrual cycles—those not fully observed during the project—were excluded from our individual-level analysis of cycle length.

Week

We defined 7-day intervals to calculate hormone averages. The first week began with the woman's first date of participation, regardless of day, and the last week included a woman's last date of participation. (Mean hormone values calculated from as few as 2 days in the week were highly correlated [more than 0.9] with the same mean calculated from 7 days of data; using fewer days did not significantly increase variance. Less than 0.01% of weeks had to be dropped because they lacked 2 or more days of hormone values.)

Data analysis

We analyzed hormones and menstrual cycle lengths at both the aggregate and individual levels. In the aggregate-level analyses, we combined women's data to look at the ranges of daily hormone values or menstrual cycle lengths by age across all participants. In the individual-level analyses, we used multilevel modeling to investigate each woman's changes in hormone levels and cycle lengths with age and to describe age-specific within- and between-woman variation.

For aggregate hormone analyses, we used daily hormone data from all BIMORA women, regardless of age and menstrual cycle characteristics (n = 145; 11 women were excluded entirely because of medications or medical procedures). We divided data into 1-year age intervals, with each woman's daily age rounded to the nearest integer; calculated the 5th, 25th, 50th, 75th, and 95th percentile values of log E1G and of log PDG for each age interval; and plotted the percentiles using a LOWESS smoothing function.[26] For the aggregate

menstrual cycle length analysis, we used menstrual bleed data reported to TREMIN or BIMORA. For BIMORA participants who had at least one uncensored cycle (n = 89), we plotted uncensored menstrual cycle lengths versus mean age per cycle. We used log values for consistency with the individual-level analyses as discussed below.

For individual-level hormone analyses, we used data from all BIMORA women (n = 145); for the individual-level menstrual cycle length analyses, we used BIMORA women with at least one uncensored cycle (n = 89). We modeled average log E1G value per week, average log PDG value per week, and log menstrual cycle length as quadratic functions of a woman's age, using a longitudinal multilevel model with random intercept and slope (see Appendix A for model specifications); the models were fitted using the statistical package MLwiN.[27] A log transform was used in each case to obtain a less skewed distribution of hormone values or menstrual cycle lengths. We used MLwiN to plot fitted trajectories of individuals' hormonal and menstrual cycle length patterns and to compute the within-woman and between-women components of variance as a function of age.

RESULTS

One hundred fifty-six women ranging from 26 to 58 years of age at the start of the study participated in BIMORA. Table 1 shows a comparison of the TREMIN sample (n = 748; 731 with age information available), those who did not participate or could not be contacted (n = 295), the TREMIN eligible subsample (n = 436), and the BIMORA sample (n = 156). The first three groups did not differ significantly from one another in average age. The final BIMORA sample (n = 156) was younger and had a higher percentage of women who were still experiencing menstrual bleeds than the other samples.

As shown in Figure 1, 53 women participated for the full 30 months of the study (five 6-month collection intervals); the average length of participation was 21 months. Most women who dropped out of the study did so at the end of a study year, accounting for the peaks in duration of participation at 6, 12, 18, and 24 months. Table 2 shows the number and average age of participants in each study year, along with the number of women who withdrew or were recruited each year. Total sample size decreased an average of 11% per year. Average age increased fewer than 5 years during the 5-year study because of ongoing recruitment and withdrawals. Hormone therapy for menopause symptoms was the most common reason for leaving the study, but represented just under half of withdrawals (Table 3). There was no significant difference in average age (at baseline) between those who withdrew and those who participated for the full 30 months (withdrew: 48.4 years, SD 8.4 years; 30-month participation: 49.0 years, SD 7.7). The average ages and standard deviations of each subgroup that withdrew from the study were not significantly different from the average age of those who participated for the entire study, except for the pregnancy group, which was younger (Table 3).

Table 4 shows the number of months of urine samples and menstrual data returned by BIMORA participants and the resulting steroid hormone readings by year. We report participation in woman-months rather than as numbers of menstrual cycles, because not all women in our sample were menstruating. The numbers of daily

TABLE 1. *Comparison of TREMIN and BIMORA samples*

	n	Mean age (SD)	Still menstruating[a] (%)
All TREMIN women >18 y contacted about participating in BIMORA	731[b]	52.0 (7.9) y	43.9
TREMIN women who declined participation or could not be contacted	295[b]	53.8 (9.6) y	25.8
TREMIN women who were eligible for BIMORA	436	51.0 (6.7) y	56.2
BIMORA	156	47.6 (8.1) y	65.4

[a]"Still menstruating" means at least one menstrual bleed in 1997 at the time of recruitment (for TREMIN categories) or during the 5-year BIMORA study (for BIMORA participants).
[b]Of the 748 women we attempted to contact, we did not have age information for 17 of the women who declined participation or could not be contacted (17 of 312).

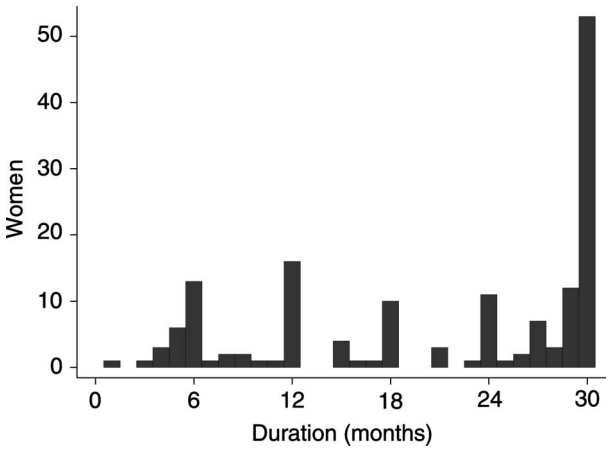

FIG. 1. Duration of individual participation in BIMORA (n = 156).

TABLE 2. *BIMORA sample characteristics, recruitment, and withdrawals by study year*

Study year	n	Mean age (SD)	Number who withdrew during study year[a]	Number who were recruited or re-enrolled during study year[a]
1	141	49.0 (7.5) y	18	17
2	134	49.7 (8.1) y	26	11
3	110	50.1 (8.2) y	16	2
4	96	51.7 (7.8) y	12	2
5	89	52.5 (8.3) y	1	5
Total			73	37

[a]During the 6-month BIMORA collection period for the study year, or before the beginning of the following 6-month collection period.

TABLE 3. *Reasons women gave for discontinuing BIMORA participation*

Reason	No. of women discontinuing	Mean age (SD)
Exogenous hormone use	33	47.8 (6.9) y
Medical procedure or illness	9	48.9 (2.3) y
Breast cancer, (n = 3)		
Hysterectomy/oophorectomy, (n = 1)		
Too busy/complicated/tired	12	48.7 (9.0) y
Pregnancy	4	34.5 (7.0) y
No response/unknown	10	47.6 (8.3) y
Temporary drops (pregnancy, hormone use)	5	
Total	73	

TABLE 4. *Steroid hormone and menstrual bleed data collected for the BIMORA project*

Study year	BIMORA woman-months[a]	No. of samples with valid E1G and PDG results	TREMIN woman-months[b]
1	770	21,918	459
2	771	21,852	434
3	653	18,506	434
4	572	16,195	339
5	530	14,921	18[c]
Total	3,296	93,392	1,684

[a]Total months of bleed data and urine samples from the BIMORA collection interval.
[b]Total months of additional bleed data from TREMIN during months when BIMORA data were not being collected.
[c]TREMIN bleed data were not available for all women in study year 5.

hormone readings per year for E1G and PDG are lower than expected based on the number of woman-months of participation, because women occasionally missed a day of collection and because some of the urine samples did not contain sufficient volume for assay. Three monthly boxes of urine samples (approximately 90 specimens total) were lost during shipment.

Two hundred fourteen boxes of samples were more than 1 day in transit, and 41 boxes were sent without a coldpack or contained samples that were unfrozen sometime during the month they were collected. Thirteen boxes were room temperature or warmer when unpacked. O'Connor et al[20] found that PDG and E1G showed little decline in immunoreactivity at room temperature for 8 days and for up to 10 freeze-thaw cycles. Therefore, samples that were warm after transit were still included in this analysis.

Aggregate-level analyses

Aggregate E1G levels (Fig. 2A) were at a maximum and had the widest range between ages 35 and 40 years, after which they declined throughout the older ages. Above age 55 years, values of E1G were low and had a narrow range. Many of the very highest individual E1G values (>95th percentile), however, occurred over age 50 years. Aggregate PDG levels (Fig. 2B) were highest and had the widest range at the early ages and declined thereafter, although some very high individual values did occur above age 40 years.

Figure 2C shows that cycle length variation declined from the earliest ages to age 40 years, after which shorter and longer cycles became more common. Cycle length ranged from 18 to 82 days at ages 25 to 30 years (mean = 33.6 days, SD = 4.6, n = 137 cycles), from 6 to 53 days at ages 30 to 40 years (mean = 29.5, SD = 3.0, n = 420 cycles), from 9 to 438 days at ages 40 to 50 years (mean = 33.4 days, SD = 15.2, n = 1,213 cycles), and from 7 to 386 days for women older than 50 years (mean = 40.1, SD = 22.4, n = 496 cycles). (For by-woman mean, we calculated each woman's average cycle length for the interval and took the mean of those average values.)

Individual-level analyses

Appendix A presents parameter estimates from our multilevel model, for average log E1G per week, average log PDG per week, and log menstrual cycle length by age.

Figures 3A through C show estimated trajectories of steroid hormones and cycle lengths for individual women, based on the model parameters. Each line represents an individual woman; lines vary in length because women varied in length of participation and because rapid, large changes in hormonal levels or cycle length resulted in comparatively long, steep lines.

Figure 3A shows that for most individual women, E1G levels rose until about age 45 years and then declined. This finding contrasts with the aggregate E1G

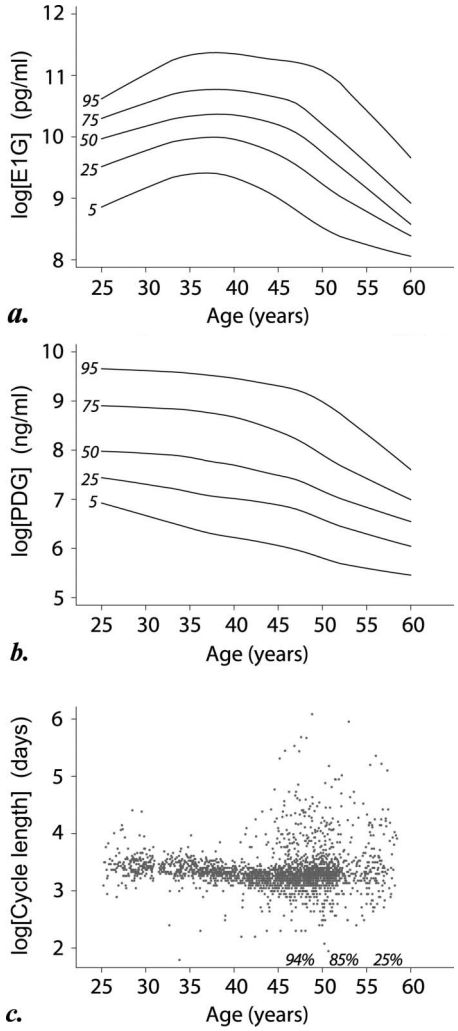

FIG. 2. Aggregate-level changes in steroid hormones and menstrual cycle length with age. Values are presented in log scale for consistency with individual-level models in Figure 3; approximate non-log equivalents are given below. (**A**) distribution (in percentiles) of daily E1G values, across all women and study years, by age (n = 145 women); approximate non-log values: log 8 = 3000 pg/mL (6.40 nmol/L), 9 = 8,100 pg/mL (17.3 nmol/L), 10 = 22,000 pg/mL (50 nmol/L), 11 = 60,000 pg/mL (128 nmol/L), 12 = 163,000 pg/mL (348 nmol/L). (**B**) distribution (in percentiles) of daily PDG values, across all women and study years, by age (n = 145); approximate non-log values: 5 = 150 ng/mL (302 nmol/L), 6 = 400 ng/ml (806 nmol/L), 7 = 1,100 ng/mL (2,215 nmol/L), 8 = 3,000 ng/mL (6,041 nmol/L), 9 = 8,100 ng/mL (16,311 nmol/L), 10 = 22,000 ng/mL (44,301 nmol/L). (**C**) scatter plot of menstrual cycle length by mean age per cycle (n = 89); percentages indicate BIMORA participants that are still menstruating and contributing to the plot in each 5-year interval; approximate non-log values: 2 = 7.4 days, 3 = 20 days, 4 = 55 days, 5 = 150 days, 6 = 400 days.

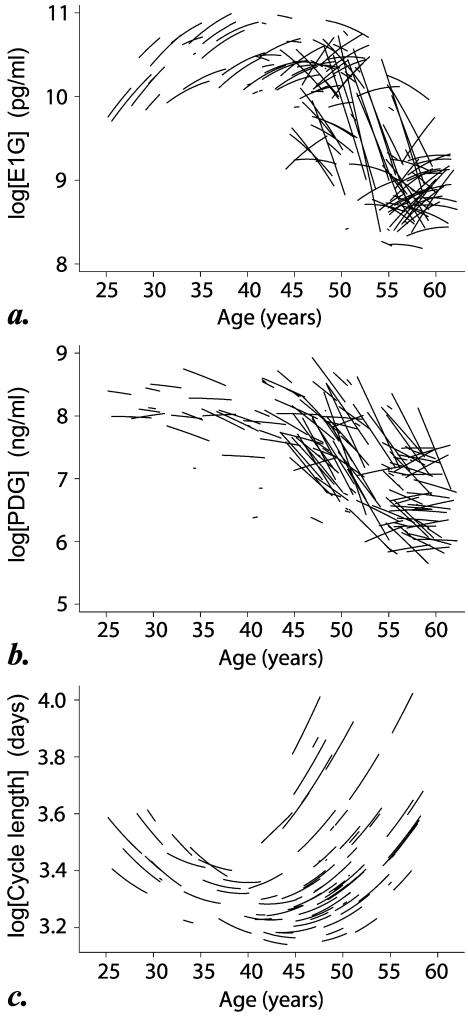

FIG. 3. Estimated individual trajectories of steroid hormones and menstrual cycle lengths. Log-transform values used in the individual-level quadratic model are presented here; approximate non-log values are given below. (**A**) mean E1G per woman-week (n = 145 women), (**B**) mean PDG per woman-week (n = 145 women), and (**C**) menstrual cycle length (n = 89 women) with age. E1G, 8 = 3,000 pg/mL (6.4 nmol/L), 9 = 8,100 pg/mL (17.3 nmol/L), 10 = 22,000 pg/mL (50 nmol/L), 11 = 60,000 pg/mL (128 nmol/L); PDG, 5 = 150 ng/mL (302 nmol/L), 6 = 400 ng/mL (806 nmol/L), 7 = 1,100 ng/mL (2,215 nmol/L), 8 = 3,000 ng/mL (6,041 nmol/L), 9 = 8,100 ng/mL (16,311 nmol/L); cycle length, 3.2 = 24.5 days, 3.4 = 30.0 days, 3.6 = 36.6 days, 3.8 = 44.7 days, 4.0 = 54.6 days.

results in which the E1G 95th percentile values peaked between ages 35 and 40 years. Most women's trajectories, although differing in absolute hormone levels, were fairly uniform in the early and late ages but could differ in the late 40s, when E1G levels were still rising for some women but had started to decline for others. At the oldest ages, there was a clear group of women with E1G values much lower than the rest of the women; for some of these women, E1G values appeared to be rising slightly.

PDG levels (Fig. 3B) appeared constant until the mid-30s and then started to decline for most women, with a clear acceleration at later ages. By the mid-50s,

levels were declining for all, with individual differences only in how steep the decline was. Similar to the E1G results, there was a group of women at the oldest ages who had low values of PDG; the PDG levels for these women showed flatter trajectories and were more variable compared with the low E1G trajectories.

For most women (Fig. 3C), cycle length declined with age up to the mid-40s and then increased, with a smooth transition and some individual variation in the steepness of the decline or subsequent increase. Curves reached a minimum—where the slope changes from negative to positive—for individuals aged 37 to 45 years.

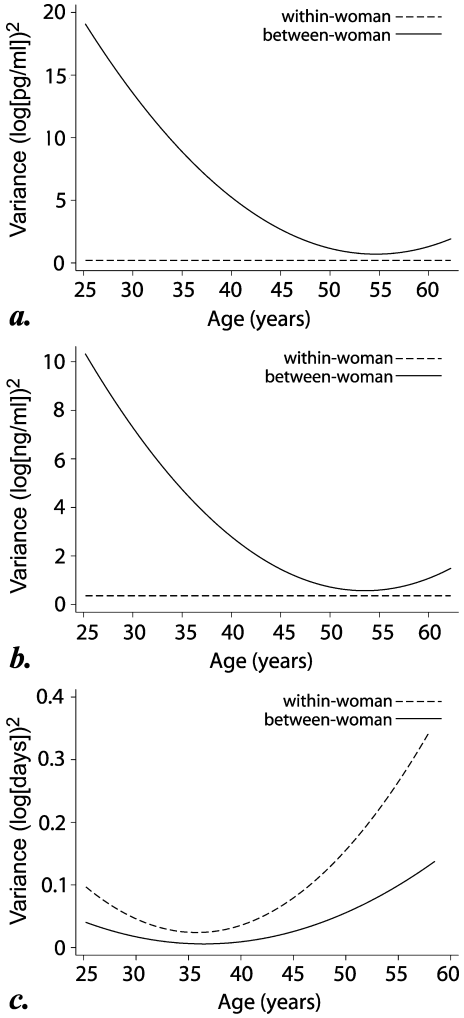

FIG. 4. Components of variance of (**A**) mean E1G per woman-week (n = 145 women), (**B**) mean PDG per woman-week (n = 145 women), and (**C**) menstrual cycle length (n = 89 women) by age.

Figures 4A through C show the two components of variance (within-woman and between-woman) for the steroid hormones and cycle length as a function of age.

Variation in E1G and PDG levels declined with age (Figs. 4A and 4B). The variation across women was largest at age 25 years and declined over time, only increasing again at ages older than 55 years. The variation within women was estimated to be zero across all ages (see Appendix A).

Figure 4C shows a different pattern for cycle length variance: a U-shaped relationship with age and within-woman variance larger than between-woman variance. Both within- and between-woman variance reached a minimum between ages 33 and 40 years At the oldest ages, within-woman variance was double the between-woman component, and total variance was nearly four times its value at the youngest ages.

DISCUSSION

Steroid hormones and menstrual cycle lengths vary a great deal within and across women with age.[10] To date, comparing results across cross-sectional and longitudinal studies to understand the nature of this variation has been challenging. The BIMORA dataset included daily observations collected over a 5-year period, allowing us to simultaneously analyze steroid hormone and menstrual cycle length changes at both the aggregate and individual levels and to estimate within- and between-woman variance with multilevel models. Although there are limitations to the study (below), our findings support previous research that PDG declines beginning in the 30s,[10,11,28-30] E1G stays at reproductive levels until closer to the final menstrual period,[10,11,28,31] and that the steroid hormones and menstrual cycle length become more variable after age 40.[11,25,32-34] The BIMORA data also provide important information on the differences between aggregate and individual results and on how cycle length and steroid hormones vary within and across women.

Steroid hormone and menstrual cycle changes at the aggregate level do not fully reflect the changes occurring in individual women. The aggregate menstrual cycle data (Fig. 2C) show that cycle length increases in variability over age 40 years, but only the individual trajectories (Fig. 3C) show that some women have a much higher average cycle length than others at those later ages. Individual E1G trajectories (Fig. 3A) show that E1G peaks for most women in their mid- to late 40s, not between 35 and 45 years as suggested by aggregate results; some women older than 40 years continue to have PDG values comparable to those of

younger women (Fig. 3B). These differences occur because aggregate distributions of hormone values (Figs. 2A and 2B) begin to decline at the age when a few women first show declining hormone levels, even though other women at those ages continue to maintain higher hormone levels. Also, our aggregate level analyses only show the distribution of values up to the 95th percentile. We noted that some of the very highest E1G and PDG values in the aggregate analyses occurred above age 40, but we could not tell if those values came from many women or just a few. Based on the individual trajectories, there were a few women over 40 who had E1G or PDG distributions that spanned much higher absolute values; these are likely to be the individuals contributing high values to the aggregate E1G distribution. Discrepancies between aggregate and individual-level analyses may explain why previous study results have differed about whether older reproductive-aged women have higher[10] or lower[4] E1G (or estradiol) levels than younger women, and why Klein et al[35] reported higher levels of follicular-phase PDG among women aged 40 to 45 years compared with women aged 20 to 25 years.

We found that the variance structures for cycle length and steroid hormones are different. Each woman has a wide range of cycle lengths that cannot be easily distinguished from other women's ranges of cycle lengths. This pattern is indicated by the close spacing between individual cycle length trajectory lines in Fig. 3C and by the larger within-woman component of variance in Fig. 4C. This finding is consistent with a previous study of cycle lengths for women in the first cohort of the TREMIN project.[25] In contrast, women are more different from one another in steroid hormone levels; this pattern is represented by the wide spacing between individual trajectories in Figs. 3A, B and by the larger between-woman variance component in Figs. 4A, B. These variance plots support an earlier report that steroid hormones have greater between-woman variance than within-woman variance.[36] The difference in the variance of menstrual cycle lengths and steroid hormones has clinical significance: although menstrual cycle length and hormone changes have both been used as indicators of proximity to menopause,[37] our results suggest that individual cycle length may be too variable to be used as a direct predictor of hormone levels.

The primary advantages of the BIMORA study included participants who had a wide range of ages and a history of cooperation. They provided daily data over a long period of time, had previous health and menstrual data for up to 30 years, and, aside from excluding hormone users, were not chosen for any reproductive characteristics. The large size and age range of the dataset allowed us to use statistical methods that would not otherwise be possible. The inclusion of women regardless of reproductive status allowed us to include women across a wide interval of reproductive and postreproductive life. The existence of other health data for these women will allow us to compare hormone patterns with concurrent and previous health data in future analyses.

Of course, our study also had limitations. Women from TREMIN were white, middle class, and had higher-than-average education. Most of the BIMORA women came from the second TREMIN cohort, and in 1997, some of these women had already reached menopause; this is reflected in Table 1 by the differences between the BIMORA and TREMIN samples in average age and percentage still menstruating. Although we included many of these women in the project as postmenopausal participants, we cannot know whether they had different hormone patterns during perimenopause that were associated with an early menopause. We also lacked data on age at final menstrual period (FMP) for some women, because not all BIMORA participants had reached their FMP by the end of the 5-year study. We therefore used chronological age for our analyses. Chronological age is not a good marker of menopausal status[38] because the length and timing of the menopausal transition varies substantially across women.[39] However, the wide range of ages in our sample ensured that we captured the transition to menopause at the aggregate level. We reanalyzed the hormone data for just BIMORA women with known FMP (n = 42, postmenopausal) and found that the trajectories and variance for E1G and PDG were not significantly changed in our model when "age" was replaced by "time since FMP." Presenting variation in hormones and cycle length by age may be more relevant to clinical practice, given that women know their age but not their future FMP date.

Many women who did not participate at all in BIMORA or dropped out during the study did so to use hormone therapy. Hormone therapy users could have different menstrual cycle patterns, but recent research suggests that women do not often use oral contraceptives or menopausal hormone therapy for cycle irregularities.[40,41] Women do commonly take hormones during perimenopause for vasomotor symptoms,[41] and in one study, women who lacked peaks in estrogen during menstrual cycles (or a 50-day interval) had more vasomotor symptoms than others.[42] Such women may be close to reaching their FMP and experiencing elongated cycles with long periods of low estrogen[43]; if such

individuals are underrepresented in our sample, then we would expect that our E1G and cycle length results would overestimate E1G levels and underestimate cycle lengths in perimenopausal women. Long cycle length was also underestimated because we were not able to include women who were still experiencing intermittent bleeds, but who had only censored cycles in our dataset. Women who declined to participate or dropped out because they were too busy, too tired, or had medical problems or illness may also have had different hormone or cycle patterns, but it was not possible to determine such differences in the current analysis; those groups did not differ in age from women who participated for the entire study (Table 3), and our individual-level analyses controlled for age.

There are other factors such as body mass index and ethnicity[44,45] that may explain some of the variation we found in hormone levels and cycle lengths. These factors have not been addressed here for several reasons. First, our primary focus was to describe variation that exists relative to age across our whole sample of women, with the understanding that part of the variation we observe may be related to a large number of additional factors. Second, we had no variation in ethnicity in our sample. Finally, our measure of body mass index was taken at a single time—for most women during the third year of the BIMORA project, but for others up to 5 years before the study began. We found that adding body mass index as a fixed covariate in our time-varying individual-level analyses did not improve model fit or change the overall appearance of the predicted trajectories or variance plots.

The BIMORA results have two important implications for clinical practice. The difference between individual trajectories and aggregate trends in our data suggests that clinicians may need to test women's hormone levels more often and over longer periods when they are seeking fertility treatment, help with menopausal symptoms, or assessment of health risks related to menstruation and hormone levels. And, because cycle lengths vary a great deal in each woman whereas hormone levels vary much less, it may not be enough to assess a woman's hormonal status based on menstrual cycle characteristics.

CONCLUSIONS

Using daily E1G, PDG, and cycle length data for a large group of women spanning a wide age range, we have presented aggregate distributions of steroid hormone and cycle length data with age, individual trajectories of steroid hormone levels and cycle length with age, and an age-specific estimation of the within- and between-woman variance for these data. Our results support generally reported trends for E1G, PDG, and cycle length, but provide specifics on the extent of variation within and across women and on how individual-level patterns relate to aggregate-level summaries of hormone and cycle length data. The differences between aggregate and individual results, and the fact that individuals' cycle lengths were too variable to predict hormone levels, suggest that clinicians should obtain more detailed information about individual women's hormonal patterns when determining fertility or menopause treatments.

Acknowledgments: The authors thank S. Barsom, G. Chorover, A. Gregor, M. Carey, and the Georgetown University Graduate School of Arts and Sciences. Profound gratitude is extended to the Tremin Trust participants for their dedicated and consistent contributions to the Biodemographic Models of Reproductive Aging Study.

REFERENCES

1. Hee J, MacNaughton J, Bangah M, Burger H. Perimenopausal patterns of gonadotrophins, immunoreactive inhibin, oestradiol and progesterone. *Maturitas* 1993;18:9-20.
2. Metcalf M, Donald RA, Livesey J. Pituitary-ovarian function in normal women during the menopausal transition. *Clin Endocrinol (Oxf)* 1981;14:245-255.
3. Dennerstein L, Dudley E, Burger H. Well-being and the menopausal transition. *J Psychosom Obstet Gynaecol* 1997;18:95-101.
4. Burger H, Dudley E, Hopper J, et al. The endocrinology of the menopausal transition: a cross-sectional study of a population-based sample. *J Clin Endocrinol Metab* 1995;80:3537-3545.
5. Burger H, Dudley E, Hopper J, et al. Prospectively measured levels of serum follicle-stimulating hormone, estradiol, and the dimeric inhibins during the menopausal transition in a population-based cohort of women. *J Clin Endocrinol Metab* 1999;84:4025-4030.
6. Overlie I, Moen M, Morkrid L, Skjaeraasen J, Holte A. The endocrine transition around menopause-a five years prospective study with profiles of gonadotropines, estrogens, androgens and SHBG among healthy women. *Acta Obstet Gynecol Scand* 1999;78:642-647.
7. Rannevik G, Carlstrom K, Jeppsson S, Bjerre B, Svanberg L. A prospective long-term study in women from pre-menopause to post-menopause: changing profiles of gonadotrophins, oestrogens and androgens. *Maturitas* 1986;8:297-307.
8. Longcope C, Franz C, Morello C, Baker R, Johnston C Jr. Steroid and gonadotropin levels in women during the peri-menopausal years. *Maturitas* 1986;8:189-196.
9. Trevoux R, De Brux J, Castanier M, Nahoul K, Soule J, Scholler R. Endometrium and plasma hormone profile in the peri-menopause and post-menopause. *Maturitas* 1986;8:309-326.
10. Santoro N, Brown J, Adel T, Skurnick J. Characterization of reproductive hormonal dynamics in the perimenopause. *J Clin Endocrinol Metab* 1996;81:1495-1501.
11. Shideler S, Devane G, Kalra P, Benirschke K, Lasley B. Ovarian-pituitary hormone interactions during the perimenopause. *Maturitas* 1989;11:331-339.
12. Metcalf M, Donald R, Livesey J. Pituitary-ovarian function before, during and after the menopause: a longitudinal study. *Clin Endocrinol (Oxf)* 1982;17:489-494.
13. Klein N, Soules M. Endocrine changes of the perimenopause. *Clin Obstet Gynecol* 1998;41:912-920.

14. Sherman B, West J, Korenman S. The menopausal transition: analysis of LH, FSH, estradiol, and progesterone concentrations during menstrual cycles of older women. *J Clin Endocrinol Metab* 1976;42:629-636.
15. Reyes F, Winter J, Faiman C. Pituitary-ovarian relationships preceding the menopause. I. A cross-sectional study of serum follicle-stimulating hormone, luteinizing hormone, prolactin, estradiol, and progesterone levels. *Am J Obstet Gynecol* 1977; 129:557-564.
16. Mansfield P, Bracken S. TREMIN: A History of the World's Oldest Ongoing Study of Menstruation and Women's Health. Lemont, PA: East Rim Publishers; 2003.
17. Treloar A, Boynton R, Behn B, Brown B. Variation of the human menstrual cycle through reproductive life. *Int J Fertil* 1967;12:77-126.
18. Voda A, Morgan J, Root J, Smith K. The Tremin Trust Research Program. In: Taylor NF, Taylor D, eds, trans. *Menstrual Health and Illness.* New York: Hemisphere Press; 1991.
19. Munro C, Stabenfeldt G, Cragun J, Addiego L, Overstreet J, Lasley B. Relationship of serum estradiol and progesterone concentrations to the excretion profiles of their major urinary metabolites as measured by enzyme immunoassay and radioimmunoassay. *Clin Chem* 1991;37:838-844.
20. O'Connor K, Brindle E, Holman D, et al. Urinary estrone conjugate and pregnanediol 3-glucuronide enzyme immunoassays for population research. *Clin Chem* 2003;49:1139-1148.
21. O'Connor K, Brindle E, Shofer J, et al. Statistical correction for non-parallelism in a urinary enzyme immunoassay. *J Immunoassay Immunochem* 2004;25:259-278.
22. Rodbard D. Statistical quality control and routine data processing for radioimmunoassays and immunoradiometric assays. *Clin Chem* 1974;20:1255-1270.
23. Miller R, Brindle E, Holman D, et al. Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations. *Clin Chem* 2004;50:924-932.
24. Belsey E, Farley T. The analysis of menstrual bleeding patterns: a review. *Contraception* 1988;38:129-156.
25. Harlow S, Lin X, Ho M. Analysis of menstrual diary data across the reproductive life span applicability of the bipartite model approach and the importance of within-woman variance. *J Clin Epidemiol* 2000;53:722-733.
26. Cleveland W. LOWESS: a program for smoothing scatterplots by robust locally weighted regression. *Am Stat* 1981;35:54.
27. MLwiN Version 1.10.0007 [computer program]. Multilevel Models Project, Institute of Education; 2001.
28. Muttukrishna S, Child T, Lockwood G, Groome N, Barlow D, Ledger W. Serum concentrations of dimeric inhibins, activin A, gonadotrophins and ovarian steroids during the menstrual cycle in older women. *Hum Reprod* 2000;15:549-556.
29. Metcalf M, Livesey J. Pregnanediol excretion in fertile women: age-related changes. *J Endocrinol* 1988;119:153-157.
30. Lipson S, Ellison P. Normative study of age variation in salivary progesterone profiles. *J Biosoc Sci* 1992;24:233-244.
31. Lee S, Lenton E, Sexton L, Cooke I. The effect of age on the cyclical patterns of plasma LH, FSH, estradiol and progesterone in women with regular menstrual cycles. *Hum Reprod* 1988;3:851-855.
32. Zapantis G, Santoro N. Ovarian ageing and the menopausal transition. *Best Pract Res Clin Obstet Gynaecol* 2002;16:263-276.
33. Burger H, Dudley E, Robertson D, Dennerstein L. Hormonal changes in the menopause transition. *Recent Prog Horm Res* 2002;57:257-275.
34. Weinstein M, Gorrindo T, Riley A, et al. Timing of menopause and patterns of menstrual bleeding. *Am J Epidemiol* 2003;158:782-791.
35. Klein N, Battaglia D, Miller P, Branigan E, Giudice L, Soules M. Ovarian follicular development and the follicular fluid hormones and growth factors in normal women of advanced reproductive age. *J Clin Endocrinol Metab* 1996;81:1946-1951.
36. Gann P, Giovanazzi S, Van Horn L, Branning A, Chatterton R Jr. Saliva as a medium for investigating intra- and interindividual differences in sex hormone levels in premenopausal women. *Cancer Epidemiol Biomarkers Prev* 2001;10:59-64.
37. Soules M, Sherman S, Parrott E, et al. Executive summary: stages of reproductive aging workshop (STRAW) Park City, Utah, July 2001. *Menopause* 2001;8:402-407.
38. Burger H, Dudley E, Mamers P, Groome N, Robertson D. Early follicular phase serum FSH as a function of age: the roles of inhibin B, inhibin A and estradiol. *Climacteric* 2000;3:17-24.
39. McKinlay S. The normal menopause transition: an overview. *Maturitas* 1996;23:137-145.
40. Wegienka G, Baird D. Potential bias due to excluding oral contraceptive users when estimating menstrual cycle characteristics. *Am J Epidemiol* 2003;158:947-950.
41. Den Tonkelaar I, Oddens B. Determinants of long-term hormone replacement therapy and reasons for early discontinuation. *Obstet Gynecol* 2000;95:507-512.
42. Weiss G, Skurnick J, Goldsmith L, Santoro N, Park S. Menopause and hypothalamic-pituitary sensitivity to estrogen. *JAMA* 2004; 292:2991-2996.
43. Miro F, Parker S, Aspinall L, Coley J, Perry P, Ellis J. Origins and consequences of the elongation of the human menstrual cycle during the menopausal transition: the FREEDOM Study. *J Clin Endocrinol Metab* 2004;89:4910-4915.
44. Lasley B, Santoro N, Randolf J, et al. The relationship of circulating dehydroepiandrosterone, testosterone, and estradiol to stages of the menopausal transition and ethnicity. *J Clin Endocrinol Metab* 2002;87:3760-3767.
45. Santoro N, Lasley B, McConnell D, et al. Body size and ethnicity are associated with menstrual cycle alterations in women in the early menopausal transition: The Study of Women's Health across the Nation (SWAN) Daily Hormone Study. *J Clin Endocrinol Metab* 2004;89:2622-2631.

APPENDIX A

Parameter estimates and standard deviations for the model of age effects on hormone levels and cycle lengths.

The hormone or cycle length outcome Y_{ij} for woman *i* in week *j* (for hormones) or cycle *j* (for cycle length) is modeled as

$$Y_{ij} = b_{0ij} + b_{1ij}\ age_{ij} + \beta_2\ age_{ij}^2 \qquad (1)$$

where age_{ij} is the mean age of woman *i* in week *j* (cycle *j*). Age_{ij} is centered on the overall mean age of the sample, and age^2_{ij} is its square. The parameters b_{0ij}, b_{1ij} and β_2 are the intercept and the linear and quadratic coefficients for age. We let the intercept and slope vary across women and weeks (cycles) but kept the curvature constant. For the intercept, $b_{0ij} = \beta_0 + u_{0i} + e_{0ij,}$ where β_0 is the expected outcome at the mean age averaged across all women, u_{0i} is a woman-specific residual reflecting the extent to which her outcome can be higher or lower than the population average, and e_{0ij} is a week-specific (cycle-specific) residual reflecting the fact that a woman's outcome in a week (cycle) can be higher or lower than her own average, after controlling for age. For the slope, $b_{1ij} = \beta_1 + u_{1i} + e_{1ij}$, where β_1 is the

TABLE 1. *Parameter estimates and SD for the model of age effects on hormone levels and cycle lengths*

Dependent variable (Y_{ij})	Unit of measure (N)	Constant (SE)($\hat{\beta}_0$)	Age (SE)($\hat{\beta}_1$)	Age^2 (SE)($\hat{\beta}_2$)
log E1G	12,683 (weeks)	9.594 (0.093)	−0.032 (0.013)	−0.006 (0.001)
log PDG	12,683 (weeks)	7.263 (0.076)	−0.090 (0.011)	−0.001 (0.001)
log cycle length	2,266 (cycles)	3.315 (0.022)	0.014 (0.003)	0.002 (0.000)

	Woman-level variance			Week-level or cycle-level variance		
	Constant (SE) ($\hat{\sigma}^2_{u0}$)	Age (SE) ($\hat{\sigma}^2_{u1}$)	Covariance (SE) ($\hat{\sigma}^2_{u01}$)	Constant (SE) ($\hat{\sigma}^2_{e0}$)	Age (SE) ($\hat{\sigma}^2_{e1}$)	Covariance (SE) ($\hat{\sigma}^2_{e01}$)
log E1G	1.093 (0.143)	0.021 (0.003)	−0.091 (0.016)	0.197 (0.002)	0.000 (0.000)	0.000 (0.000)
log PDG	0.695 (0.096)	0.012 (0.002)	−0.039 (0.010)	0.364 (0.005)	0.000 (0.000)	0.000 (0.000)
log cycle length	0.024 (0.005)	0.000 (0.000)	0.002 (0.001)	0.074 (0.003)	0.001 (0.000)	0.006 (0.000)

effect of age at the mean age for the sample, averaged across all women, u_{1i} is a woman-specific residual reflecting the extent to which age can affect her outcomes more or less than the average, and e_{1ij} is a week-specific (cycle-specific) residual that allows age to have a stronger or weaker effect in a given week (cycle) than is typical for the woman. The curvature, β_2, represents the extent to which the effect of age itself varies with age, and is assumed to be the same for all women.

The model in equation (1) can then be written in more detailed form as

$$Y_{ij} = (\beta_0 + u_{0i} + e_{0ij}) + (\beta_1 + u_{1i} + e_{1ij})age_{ij} + \beta_2\, age^2_{ij} \quad (2)$$

The woman-level residuals u_{0i} and u_{1i} are assumed to be drawn from a bivariate normal distribution with means zero, variances $\sigma^2{}_{u0}$ and $\sigma^2{}_{u1}$ and covariance σ_{u01}. The week-level (cycle-level) residuals e_{0ij} and e_{1ij} are assumed to be independent of the woman-level residuals and are drawn from a bivariate normal distribution with means zero, variances $\sigma^2{}_{e0}$ and $\sigma^2{}_{e1}$ and covariance σ_{e01}. The variance of the outcome in this model can be written as

$$Var(Y_{ij}) = (\sigma^2_{u0} + 2\sigma_{u01}\, age_{ij} + \sigma^2_{u1}\, age^2_{ij}) + (\sigma^2_{e0} + 2\sigma_{e01}\, age_{ij} + \sigma^2_{e1}\, age^2_{ij}), \quad (3)$$

and can be seen to depend on age and its square.

The fitted trajectories for individual women are obtained from equation 2 by setting the population and woman-level parameters to their estimated values:

$$\hat{Y}_{ij} = (\hat{\beta}_0 + \hat{u}_{0i}) + (\hat{\beta}_1 + \hat{u}_{1i})age_{ij} + \hat{\beta}_2\, age^2_{ij}. \quad (4)$$

Our estimates of week-level (within-woman) variance were zero at all ages for both E1G and PDG. We had expected that using week-long mean hormone values would introduce some within-woman variance into the model since, at least in reproductive-aged women, different weeks are capturing different portions of a menstrual cycle. Most likely, the use of log hormone values made the differences between week-long means for a given woman very small; the concurrent estimation of within-woman variance and the comparatively much higher between-woman variance probably resulted in these within-woman variance estimates of zero. In an earlier version of these individual-level steroid hormone analyses, we had used mean hormone values per menstrual cycle instead of mean values per week. We found that within-woman variance was very low relative to between-woman variance, but that it was still greater than zero; for both E1G and PDG, within-woman variance decreased to a minimum around age 40 years and then increased slightly. This previous analysis had a smaller number of hormone mean values than the current analysis and included only women who were still menstruating.