Reproducibility of Plasma Steroid Hormones, Prolactin, and Insulin-like Growth Factor Levels among Premenopausal Women over a 2- to 3-Year Period

Stacey A. Missmer,[1,2,3] Donna Spiegelman,[3,4] Elizabeth R. Bertone-Johnson,[5] Robert L. Barbieri,[2] Michael N. Pollak,[6] and Susan E. Hankinson[1,3]

[1]Channing Laboratory, Department of Medicine, and [2]Department of Obstetrics, Gynecology, and Reproductive Biology, Brigham and Women's Hospital and Harvard Medical School; Departments of [3]Epidemiology and [4]Biostatistics, Harvard School of Public Health, Boston, Massachusetts; [5]Department of Public Health, University of Massachusetts, Amherst, Massachusetts; and [6]Department of Oncology, McGill University and Lady Davis Research Institute, Montreal, Quebec, Canada

Abstract

Few studies have evaluated whether a single blood hormone measurement, as is available in most epidemiologic studies, sufficiently characterizes a premenopausal woman's long-term hormone levels; there is particular concern whether sex steroid hormones, which fluctuate during the menstrual cycle, are reliable. We conducted a prospective study within the Nurses' Health Study II to examine the reproducibility of plasma estrogens, androgens, progesterone, prolactin, sex hormone binding globulin, insulin-like growth factor–I (IGF-I), and IGF binding protein–3 (IGFBP-3). One blood sample per year over 3 years was collected from 113 premenopausal women during both the follicular and luteal phases of the menstrual cycle. We calculated intraclass correlation coefficients (ICC) across the three samples for all women. Among estrogens, ICCs ranged from 0.38 (estradiol) to 0.60 (estrone sulfate) in the follicular phase and from 0.44 (estrone) to 0.69 (estrone sulfate) in the luteal phase. Among androgens, ICCs ranged from 0.58 (androstenedione) to 0.94 [dehydroepiandrostenedione sulfate (DHEAS)] in the follicular phase and from 0.56 (testosterone) to 0.81 (DHEAS) in the luteal phase. When values were averaged across the follicular and luteal phases, the ICC for prolactin was 0.64 whereas ICCs for IGF-I and IGFBP-3 were 0.86 and 0.82, respectively. The ICC for progesterone in the luteal phase was only 0.29. These data suggest that for androgens, estrone sulfate, prolactin, IGF-I, and IGFBP-3, a single measurement can reliably categorize average levels over at least a 3-year period in premenopausal women. For estrone and estradiol, where ICCs were relatively low, it is important to use reproducibility data such as those to correct for measurement error in epidemiologic studies. (Cancer Epidemiol Biomarkers Prev 2006;15(5):972–8)

Introduction

Endogenous hormone concentrations are strongly associated with breast cancer (1), endometrial cancer (2), and osteoporosis risk (3, 4) among postmenopausal women; however, investigation of the associations among premenopausal women has been limited (5-10). In part, this is due to the concern that given the fluctuations in plasma hormone levels during the menstrual cycle, a single blood measurement (such as is available in most epidemiologic studies) is not a sufficient indicator of long-term levels—the exposure generally being of greatest interest.

Previously, we assessed the reproducibility of plasma estradiol, estrone, estrone sulfate, and progesterone among 87 premenopausal women over a 1-year period (11). However, evaluation of the reproducibility of these and additional hormones over a longer period of time has not been undertaken. Therefore, we conducted a prospective study within the Nurses' Health Study II to examine the reproducibility of plasma estradiol, free estradiol, estrone, estrone sulfate, androstenedione, testosterone, dehydroepiandrostenedione (DHEA), dehydroepiandrostenedione sulfate (DHEAS), progesterone, sex hormone binding globulin (SHBG), prolactin, insulin-like growth factor–I (IGF-I), free IGF-I, and IGF binding protein–3 (IGFBP-3)—endogenous factors that may be related to chronic disease risk—in both the follicular and luteal phases of the menstrual cycle among 113 premenopausal women over a 2- to 3-year period.

Materials and Methods

Study Population. The Nurses' Health Study II was established in 1989 when 116,671 female registered nurses, 25 to 42 years of age, completed and returned a mailed questionnaire. The cohort continues to be followed every 2 years by questionnaire to update exposure status and to identify cases of newly diagnosed disease. The Nurses' Health Study II blood cohort includes 29,611 women who contributed blood samples from 1996 to 1999; for 19,092 women, the collected samples were timed within the menstrual cycle.

Each blood collection kit contained all of the supplies needed to have blood samples drawn by a local laboratory or a colleague (e.g., needle, tourniquet, and blood collection tubes with sodium heparin). Premenopausal participants were asked to have their first blood sample drawn on the 3rd, 4th, or 5th day of their menstrual cycle (''follicular'' blood draw) and to have the second blood sample drawn 7 to 9 days before the anticipated start of their next cycle (''luteal'' blood draw). Timing of the luteal sample from the estimated first day of the next menstrual cycle is generally more accurate than counting forward from day 1 of the current cycle because the length of the follicular phase is more variable than the length of the luteal phase (12, 13). Participants placed their follicular blood samples in a

Received 11/7/05; revised 2/8/06; accepted 2/21/06.

Grant support: NIH grants CA67262 and CA50385, and Training grant in cancer epidemiology, 5 T32 CA090001-281, from the National Cancer Institute (S.A. Missmer).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Stacey A. Missmer, Channing Laboratory, Brigham and Women's Hospital, 181 Longwood Avenue, Boston, MA 02115. Phone: 617-525-2021; Fax: 617-525-2008 alternate fax: 617-525-4597. E-mail: stacey.missmer@channing.harvard.edu

doi:10.1158/1055-9965.EPI-05-0848

refrigerator for 8 to 24 hours after it was drawn; they then aliquoted the plasma into cryotubes. The plasma was kept in the participant's home freezer until the second (luteal) blood collection. Then the woman arranged to have both samples shipped, via overnight courier and with an ice pack, to our laboratory; on arrival, the luteal whole blood sample was processed and aliquoted into labeled cryotubes. All samples have been stored in the vapor phase of continuously monitored liquid nitrogen freezers since collection.

We previously reported that plasma hormones remained stable when collected in the manner used for the luteal phase samples (14, 15). We additionally conducted a pilot study to determine if hormone levels remained stable with our follicular phase processing method. Two tubes of blood were collected in the follicular phase from each of 16 premenopausal women. One tube was processed and frozen immediately and the second was treated identically to the processing method described above. Estradiol, free and bioavailable estradiol, estrone, and testosterone were assayed on all samples. For all hormones, the mean and SD for each of the processing methods were almost identical and, with the exception of estrone, the ICC ranged from 0.93 to 0.98. For estrone, the ICC was 0.8, due primarily to a single replicate (values of 48 and 70 pg/mL); without this replicate, the ICC was 0.87.

Among blood study participants, invitation to participate in the ''hormone stability study'' was extended to a random sample of responders who were premenopausal, not using exogenous hormones (e.g., oral contraceptives), and who were not currently nor planning to be pregnant or lactating. Second and third blood collection kits were mailed to women who returned the first kit without reminding (i.e., excellent responders) and remained eligible to participate. Of the 412 invited women, 74% (n = 304) provided a second set of samples and, of these, 236 (57%) sent a third set. These 236 women did not significantly differ by age, race, parity, body mass index, or cigarette use from the 412 who received the initial invitation (data not shown). Overall, six blood samples—a follicular and a luteal in each of the three sets—were collected from each of 236 women over the 3-year period.

For each menstrual cycle sampled, a questionnaire was sent with the blood collection kit on which to record the first day of the menstrual cycle during which the blood samples were drawn and the dates of both blood draws. Details on the number of hours since last food intake before the two blood draws, time of day and month of blood collection, and the participant's current weight, menstrual cycle length, exercise frequency, and smoking status were also collected. Finally, a postcard on which to record the first day of the next menstrual cycle was provided; all but one postcard (99.7%) was returned.

For these analyses, based on a desired final sample size of 100 women and given the estimated proportion of women who would be found to have had anovulatory cycles and limiting to women whose luteal phase samples were each collected between 3 and 11 days before the start of her next menstrual period, blood samples collected from a total of 113 women were selected for hormone and SHBG level analysis. For financial reasons and to decrease plasma use, follicular samples from only 50% of the women were submitted for assays where the correlation across the menstrual cycle was expected to be relatively high (i.e., free estradiol, DHEA, DHEAS, IGF-I, free IGF-I, and IGFBP-3).

The study was approved by the Committee on the Use of Human Subjects in Research at the Brigham and Women's Hospital.

Laboratory Methods. Assays were conducted by three different laboratories. Estradiol, free estradiol, estrone, estrone sulfate, progesterone, androstenedione, testosterone, dehydroepiandrosterone, and dehydroepiandrosterone sulfate were assayed at Quest Diagnostics-Nichols Institute (San Juan Capistrano, CA). SHBG and prolactin were assayed by Dr. P. Sluss at the Reproductive Endocrinology Unit Laboratory of the Massachusetts General Hospital (Boston, MA). IGF-I, free IGF-I, and IGFBP-3 were assayed by Dr. M. Pollak at McGill University (Quebec, Canada).

Hormone and growth factor assay methods have previously been described in detail (15-17). In brief, samples were extracted with hexane-ethyl acetate, the steroids were eluted from celite columns, and the fractions were then assayed by radioimmunossay (18-22). DHEAS was assayed by radioimmunossay without a prior separation step (23). After extraction of estrone, estrone sulfate was assayed by radioimmunossay of estrone, after enzyme hydrolysis, organic extraction, and separation by column chromatography (24). Prolactin levels were assayed using the AxSYM Immunoassay system (Abbott Diagnostics, Chicago, IL). IGF-I and IGFBP-3 levels were assayed by ELISA with reagents from Diagnostic Systems Laboratory (Webster, TX).

All of the follicular and luteal samples from a single woman were assayed together; the samples were ordered randomly and labeled so that the laboratory could not identify samples from the same woman. To assess laboratory precision, quality control replicates of 10% of all samples assayed were randomly interspersed and labeled to preclude their identification. Overall, with the exception of progesterone, within-batch laboratory coefficients of variation for the assays ranged from 4% for IGF-I to 14% for estrone sulfate. Progesterone had a coefficient of variation of 40% due primarily to a single outlier among 53 samples tested; when this value was excluded, the coefficient of variation was 14%.

Statistical Analyses. We used the extreme studentized deviate many-outlier procedure (25, 26) to assess for outliers in each set of laboratory results. This resulted in the removal of 5 estradiol, 1 free estradiol, 5 estrone, 10 estrone sulfate, 1 androstenedione, 2 DHEA, 3 testosterone, 12 prolactin, 1 SHBG, and 1 IGFBP-3 values from among the 678 possible values per hormone (113 women $\times$ 2 blood samples $\times$ 3 time periods). The final number of samples available for each phase of the menstrual cycle within each collection is provided in Table 1.

The natural logarithm of the plasma hormone and growth factor values were used in analyses because the transformed values were more normally distributed. Between-person and within-person variances were estimated from the three sets of hormone measurements by random effect models, as implemented by SAS statistical software (SAS Institute, Cary, NC). To assess reproducibility across the three samples for all women combined, we calculated intraclass correlation coefficients (ICC) by dividing the between-person variance by the sum of the between-person and within-person variances (27). Ninety-five percent confidence intervals were calculated for the ICCs (28). Spearman correlations were calculated to evaluate the correlation both between hormones and between the follicular and luteal phases within individual hormones (27).

To assess the utility of a single hormone measurement to correctly classify longer term hormone levels into quartile categories, we compared quartiles of hormone levels (as measured by the first blood sample) to quartiles as defined by the mean of the second and third blood samples. Quartile cut points were defined separately and thus were not necessarily the same.

Results

A total of 113 women (n = 339 menstrual cycles) contributed data to these analyses. They ranged in age at first blood draw

Table 1. Blood draw–specific sample size and geometric means among all women

Endogenous factor	Geometric mean (10th-90th percentile)					
	Follicular blood draws			Luteal blood draws		
	1	2	3	1	2	3
Estradiol (pg/mL)	63 (30-109)*	60 (29-104)*	62 (26-117)	140 (87-203)	150 (87-210)	145 (88-223)
Free estradiol (pg/mL)	1.2 (0.6-2.3)†	1.2 (0.6-2.1)†	1.3 (0.6-2.0)†	2.4 (1.5-3.5)	2.6 (1.5-3.7)	2.5 (1.6-3.7)
Estrone (pg/mL)	48 (28-72)*	45 (28-65)*	47 (32-68)*	86 (52-128)	87 (50-134)	80 (46-116)
Estrone sulfate (pg/mL)	793 (307-1,379)*	669 (250-1,249)*	651 (258-985)*	1,471 (573-2,383)*	1,475 (521-2,552)*	1,332 (496-2,301)*
Progesterone (ng/dL)	—	—	—	1,319 (475-2,179)	1,202 (436-1,948)	1,216 (466-2,135)
Testosterone (ng/dL)	23 (12-33)	20 (11-32)	21 (11-31)*	27 (16-40)	26 (16-39)	27 (15-41)
Androstenedione (ng/dL)	127 (57-214)	107 (57-168)*	116 (54-196)*	150 (78-255)	135 (70-213)	136 (72-217)
DHEA (ng/dL)	367 (128-701)†	356 (114-626)†	352 (128-663)†	378 (150-724)	365 (137-669)	370 (139-739)
DHEAS (μg/dL)	177 (64-303)†	166 (66-272)†	163 (61-261)†	179 (82-284)	163 (69-266)	161 (64-268)
SHBG (nmol/L)	70 (37-108)	72 (37-110)	69 (34-113)	74 (38-119)	73 (40-116)	73 (38-115)
Prolactin (ng/mL)	15 (7-25)	16 (8-27)	16 (8-28)	18 (8-32)	18 (8-36)	18 (8-31)
IGF-I (ng/mL)	176 (105-276)†	176 (108-256)†	173 (115-253)†	226 (134-323)	224 (151-308)	224 (139-320)
Free IGF-I (ng/mL)	1.0 (0.5-1.5)†	1.0 (0.5-1.7)†	1.1 (0.4-1.9)†	1.1 (0.7-1.8)	1.3 (0.7-1.8)	1.3 (0.8-2.1)
IGFBP-3 (ng/mL)	3,158 (2,388-4,258)†	3,197 (2,210-3,859)†	3,070 (2,342-4,064)†	3,422 (2,555-4,427)	3,250 (2,469-4,106)	3,252 (2,452-4,240)

NOTE: N = 100 to 113 samples were assayed unless otherwise noted.
Abbreviations: DHEA, dehydroepiandrostenedione; DHEAS, dehydroepiandrostenedione sulfate; SHBG, sex hormone binding globulin; IGF-I, insulin-like growth factor I; IGFBP-3, IGF binding protein–3.
*N = 75-99 samples were assayed.
†N = 37-54 samples were assayed.

from 34 to 49 years, with a mean age of 41 years. Weight ranged from 90 to 255 pounds and the mean body mass index was 24.5 kg/m^2. Fifty-eight percent of follicular samples and 68% of luteal samples were drawn at least 8 hours since the woman's last meal. Only eight women were current smokers and three of these women quit smoking between their second and third blood collection. Nineteen women contributed at least one blood sample during an anovulatory menstrual cycle, identified as having a progesterone level <400 ng/dL. Among the three luteal phase samples, the time from blood draw to the first day of the woman's next menstrual cycle ranged from 3 to 11 days (mean = 7 days). The mean interval between the first and the third blood collection was 34 months (range, 24-46 months).

Whereas differences in absolute levels between the follicular and luteal phases at each blood sample were most evident for the estrogens, androgens and growth factors also differed significantly across the menstrual cycle (paired t test $P < 0.05$; Table 1). The correlations between the follicular and luteal phase levels based on the first blood sample, with the exception of free estradiol, were all statistically significant (Table 2), suggesting that women tended to rank similarly between the two menstrual cycle phases. The strongest correlations were for SHBG and DHEAS ($r > 0.8$) and secondarily for testosterone, DHEA, IGF-I, free IGF-I, and IGFBP-3 ($r > 0.6$). If the mean of all three samples was used, all correlations tended to increase slightly (data not shown). The notable exceptions to this were androstenedione and IGF-I where the increase in the correlation was substantial (androstenedione, r = 0.52 to 0.79; IGF-I, r = 0.65 to 0.86).

The correlations between hormones within the follicular and luteal phases are presented in Table 3. In the follicular phase, the androgens were positively correlated with each other whereas testosterone was also significantly correlated with estrone, SHBG, and prolactin. IGF-I was correlated with androstenedione, DHEA, and IGFBP-3. Estrone sulfate was correlated only with estrone and DHEAS. In the luteal phase, whereas prolactin was not correlated with any other hormone, the androgens were only correlated with each other. The IGFs were inversely correlated with SHBG. Estradiol was significantly correlated with estrone, progesterone, and SHBG but not with estrone sulfate.

For all women combined, we observed relatively high ICCs (>0.60) for the androgens, IGF-related factors, and, in both menstrual cycle phases, estrone sulfate (Table 4). Estradiol and estrone yielded ICCs of 0.38/0.45 and 0.42/0.44, respectively, in the follicular and luteal phases. The ICC for luteal progesterone was poor (0.29). Averaging the follicular and luteal phase measurements within each cycle improved the ICC for testosterone, androstenedione, IGF-I, IGFBP-3, and most dramatically for prolactin. As expected, the estrogen ICCs did not improve with averaging; however, the biological relevance of an average across the menstrual cycle is questionable given the marked differences in levels between the two cycle phases. After adjustment for age at blood draw (continuous in years), time of day of draw (continuous), fasting status of draw (>8 hours since last meal versus more recent intake), luteal day (e.g., time to start of next menstrual cycle) of draw (continuous), average exercise frequency with heavy perspiration over the past month (<1, 1, 2-3, 4-6, and >6 times per week), current body mass index (continuous kg/m^2), and average menstrual cycle length over the past 6 months (<21, 21-25, 26-31, 32-39, 40-50, >50 days, and irregular), the ICCs remained essentially unchanged (data not shown).

To evaluate possible sources of within-person variation in hormones where the ICCs were relatively low, we conducted several subanalyses excluding women/samples whose potential hormone level–related characteristics changed over the 3-year data collection period (Table 5). ICCs remained similar

Table 2. Spearman correlation coefficient (95% confidence interval) between the follicular and luteal phase hormones and SHBG levels in the first cycle of blood collection among all women

Endogenous factor	Correlation (95% confidence interval)
Estradiol	0.22 (0.02-0.42)
Free estradiol	−0.09 (−0.41-0.23)
Estrone	0.38 (0.18-0.58)
Estrone sulfate	0.50 (0.31-0.69)
Progesterone	— (—)
Testosterone	0.63 (0.48-0.78)
Androstenedione	0.52 (0.35-0.69)
DHEA	0.60 (0.37-0.83)
DHEAS	0.82 (0.66-0.98)
SHBG	0.86 (0.76-0.96)
Prolactin	0.55 (0.39-0.71)
IGF-I	0.65 (0.44-0.86)
Free IGF-I	0.67 (0.47-0.87)
IGFBP-3	0.66 (0.46-0.86)

NOTE: See Table 1 for abbreviations.

Table 3. Spearman correlation in the first cycle of blood collection between the luteal phase and follicular phase hormones and SHBG among all women

Endogenous factors	Estradiol	Estrone	Estrone sulfate	Progesterone	Testosterone	Androstenedione	DHEA	DHEAS	SHBG	Prolactin	IGF-I	IGFBP-3
Estradiol	*1.0*	0.50*	0.07	0.39*	0.17	0.11	0.01	−0.07	0.26*	0.07	−0.01	−0.05
Estrone	**0.35***	*1.0*	0.43*	0.05	0.36*	0.34*	0.15	0.06	−0.01	0.10	−0.04	−0.08
Estrone sulfate	**0.16**	**0.54***	*1.0*	−0.04	0.17	0.23*	0.18	0.35*	−0.17	−0.14	0.07	−0.04
Progesterone	—	—	—	*1.0*	0.00	0.07	−0.02	0.03	0.29*	−0.24	−0.00	0.06
Testosterone	**0.07**	**0.30***	**0.14**	—	*1.0*	0.77*	0.56*	0.38*	0.24*	0.09	0.11	−0.06
Androstenedione	**−0.01**	**0.08**	**0.03**	—	**0.81***	*1.0*	0.72*	0.41*	0.16	0.12	0.19	0.02
DHEA	**−0.03**	**−0.31**	**0.13**	—	**0.64***	**0.78***	*1.0*	0.68*	0.10	0.21	0.14	0.02
DHEAS	**−0.18**	**−0.04**	**0.40***	—	**0.40***	**0.54***	**0.73***	*1.0*	−0.08	−0.02	0.11	0.07
SHBG	**0.40***	**0.21***	**−0.08**	—	**0.20***	**0.08**	**−0.20**	**−0.30***	*1.0*	−0.05	−0.20*	−0.30*
Prolactin	**−0.07**	**0.06**	**−0.06**	—	**0.21***	**0.30***	**0.26**	**−0.03**	**−0.22***	*1.0*	0.04	0.02
IGF-I	**0.20**	**−0.10**	**0.02**	—	**0.18**	**0.29***	**0.33***	**0.05**	**−0.19**	**0.20**	*1.0*	0.61*
IGFBP-3	**−0.00**	**−0.19**	**−0.05**	—	**−0.03**	**0.15**	**0.22**	**0.03**	**−0.15**	**0.07**	**0.79***	*1.0*

NOTE: See Table 1 for abbreviations. Follicular phase correlations (bottom left portion of the table) are bolded whereas the luteal phase correlations (top right portion of the table) are not.
*$P < 0.05$.

after excluding anovulatory cycles, those who changed smoking status (n = only three women; data not shown), those who reported being postmenopausal on the 2001 or 2003 Nurses' Health Study II main questionnaire, those who gained or lost >10 pounds in the past year (n = 19; data not shown), those whose body mass index changed by >1 kg/m^2, those whose luteal collection varied by >±2 days across the three blood samples, and those who had not contributed a fasting blood sample. For example, the ICC across three blood samples for luteal-phase estradiol among all women was 0.45. After exclusion of anovulatory cycles, the ICC for luteal phase estradiol was 0.51. Similarly, after excluding women who became postmenopausal within 4 years of the final blood sample, the ICC was 0.48, or those whose body mass index changed by >1 kg/m^2 (ICC = 0.43), or those whose luteal collection day varied by >±2 days (ICC = 0.52). After exclusion of nonfasting blood samples [n = 248 samples (37%)], the follicular-phase ICC for prolactin was 0.52 whereas the luteal-phase ICC was 0.41. Similar analyses were conducted for the other hormones, and there too, results were essentially unchanged (data not shown).

Additionally, in Bland-Altman analysis comparing within-person variation as a function of the within-person means, we found that there was little evidence for an association (29). In the cases where the association was significant, less relative variation occurred with greater intraindividual levels. For example, in the luteal phase, these included progesterone (P < 0.001), estradiol (P = 0.004), prolactin (P = 0.007), DHEAS (P = 0.007), and DHEA (P = 0.008).

Finally, we addressed how well a single sample would classify women into the appropriate quartile of exposure using the mean of the second and third samples as the ''gold standard.'' Shown in Table 6 are matrices for four of the hormones of interest that represent a range of ICCs: follicular estradiol (ICC = 0.38), follicular testosterone (ICC = 0.68), follicular SHBG (ICC = 0.83), and follicular DHEAS (ICC = 0.94). The quartiles are approximate because of a number of women with identical plasma hormone values. For estradiol, 38 of 92 (41%) were perfectly classified, 79 of 92 (86%) were off by one category or less, and just 3 values (3.3%) were misclassified into an extreme category. Percentages for testosterone were 55%, 90%, and 1.3%, respectively. For SHBG, 57% were perfectly classified, 93% were off by one category or less, and none were misclassified into an extreme category. For DHEAS, 79% were perfectly classified whereas 100% were off by one category or less. Although concordance was high, these results will tend to underestimate agreement with true long-term levels as two, rather than a large number of replicates, were used as the gold standard.

Table 4. ICCs and 95% confidence intervals among all women across three blood collections

Endogenous factor	ICCs (95% confidence intervals)		
	Follicular samples	Luteal Samples	Cycle Average
Estradiol	0.38 (0.26-0.51)	0.45 (0.34-0.57)	0.45 (0.34-0.56)
Free estradiol	0.22 (0.08-0.49)	0.43 (0.31-0.55)	0.38 (0.27-0.50)
Estrone	0.42 (0.29-0.56)	0.44 (0.33-0.56)	0.33 (0.22-0.46)
Estrone sulfate	0.60 (0.49-0.70)	0.69 (0.60-0.77)	0.57 (0.47-0.67)
Progesterone	— (—)	0.29 (0.18-0.42)	— (—)
Testosterone	0.68 (0.59-0.76)	0.56 (0.46-0.66)	0.73 (0.65-0.79)
Androstenedione	0.58 (0.48-0.68)	0.56 (0.46-0.66)	0.66 (0.57-0.74)
DHEA	0.73 (0.60-0.82)	0.66 (0.56-0.74)	0.69 (0.61-0.76)
DHEAS	0.94 (0.90-0.96)	0.81 (0.75-0.85)	0.86 (0.81-0.89)
SHBG	0.83 (0.78-0.87)	0.83 (0.78-0.87)	0.89 (0.85-0.92)
Prolactin	0.55 (0.45-0.65)	0.41 (0.30-0.53)	0.64 (0.54-0.72)
IGF-I	0.69 (0.56-0.79)	0.83 (0.77-0.87)	0.86 (0.82-0.90)
Free IGF-I	0.47 (0.32-0.63)	0.58 (0.48-0.67)	0.59 (0.50-0.68)
IGFBP-3	0.69 (0.56-0.79)	0.76 (0.70-0.82)	0.82 (0.76-0.87)

NOTE: See Table 1 for abbreviations.

Discussion

Our study of hormone reproducibility among premenopausal women suggests that a single measure of androgens, IGFs, and estrone sulfate during either the follicular or luteal phase sufficiently quantifies circulating levels over at least a 3-year period. The average of a follicular and luteal measurement can improve this categorization, particularly for prolactin. However, the ICCs for estradiol and estrone were relatively low such that correction of calculated relative risk estimates for random within-person variation should be employed in studies of hormone/disease associations (30). All ICC calculations were robust to potential hormone-related factors or subgroup analyses.

We observed little change in the within-woman characteristics across the 3-year blood sampling period. Only three current smokers became past smokers and only 11 women (10%) became postmenopausal in the follow-up period (through June 2003) after their final blood draw. Therefore, analyses excluding these women did not alter the observed ICCs appreciably. Even analyses taking into account weight gain—common in this study population—did not alter the ICCs. This is likely because among premenopausal women, estrogens derive primarily from the ovary, whereas among postmenopausal women, where adipose is the primary source of estrogens, an effect of weight change would be expected (31).

Most previous studies have measured hormone reliability over time among postmenopausal women (28, 32-35). Within the Nurses' Health Study, we observed ICCs of 0.68 and 0.74 for estradiol and estrone, 0.75 for DHEA, and 0.88 for testosterone, suggesting substantial reproducibility over a 2- to 3-year period. The ICC observed for prolactin was 0.53 and for SHBG was 0.92 (28). Of course, among postmenopausal women, endogenous levels of steroid hormones are not fluctuating in response to the menstrual cycle. In the current study of premenopausal women, we observed similar ICCs for endogenous hormones that do not change with the postmenopausal transition, such as DHEA and DHEAS, whereas we observed quite different ICCs for hormones such as estrogens that change dramatically from premenopause to postmenopause.

Few studies have evaluated reproducibility over time among premenopausal women (11, 32, 35). Two have reported ICCs for plasma androgens ranging from 0.60 to 0.85 over a 1-year period (32, 35); our ICCs over a 2- to 3-year period were similar to these. Only our previous study ($n = 87$ women) evaluated the ICC for both the follicular and luteal phases, reporting ICCs over a 1-year period for estradiol, estrone, estrone sulfate, and progesterone (11). With the exception of estradiol in the luteal phase, observed ICCs in that study ranged from 0.52 to 0.71. The ICC for luteal estradiol was 0.19, which increased to 0.49 after restricting to ovulatory cycles and to 0.62 when further restricting luteal timing of the sample to 4 to 10 days. The latter analysis, however, included only 39 women. In the current study, the ICC for the intervals between years 1 and 2 and between years 2 and 3 were similar to the ICCs observed across the entire 3-year period, suggesting that this between-study difference may be due to chance. Of note, in our previous study, 30 of 174 (18%) cycles from the 87 women were anovulatory with a cut point of <300 ng/dL of progesterone, whereas here only 25 of 339 (7%) cycles from the 113 women were identified as anovulatory with a more specific cut point of <400 ng/dL, which may explain why the current results change very little when excluding anovulatory cycles.

The poor reproducibility of luteal progesterone contradicts our previous findings over a 1-year period (ICC = 0.54), despite sampling from women within the same cohort with similar personal characteristic distributions (e.g., age, body mass index). When we limited our current analyses to include only the first and second blood samples, the ICC for progesterone improved slightly to 0.38, whereas when we limited the analyses to include only the second and third blood samples, the ICC for

Table 5. ICCs among select groups of women across three blood collections

Endogenous factor	ICCs	
	Follicular Samples	Luteal samples
Excluding anovulatory cycles* ($n = 3$ women excluded; 25 cycles excluded)		
Estradiol	0.36	0.51
Free estradiol	0.22	0.48
Estrone	0.39	0.49
Estrone sulfate	0.59	0.72
Progesterone	—	0.27
Excluding women who were postmenopausal within 4 y of the final blood sample ($n = 11$ women excluded)		
Estradiol	0.40	0.48
Free estradiol	0.22	0.45
Estrone	0.44	0.48
Estrone sulfate	0.59	0.69
Progesterone	—	0.29
Excluding women with >1 kg/m^2 change in body mass index across the three blood samples ($n = 66$ women excluded)		
Estradiol	0.26	0.43
Free estradiol	0.08	0.34
Estrone	0.39	0.51
Estrone sulfate	0.43	0.69
Excluding women whose luteal collection varied by >±2 d across the three blood samples ($n = 69$ women excluded)		
Estradiol	—	0.52
Free estradiol	—	0.49
Estrone	—	0.48
Estrone sulfate	—	0.64
Progesterone	—	0.33

NOTE: See Table 1 for abbreviations.
*Anovulatory cycles were defined by a progesterone level of <400 ng/dL ($n = 22$ cycles from 19 women). Three women were anovulatory during two sampled menstrual cycles and were excluded from these analyses.

Table 6. Cross-classification of follicular phase hormone levels among premenopausal women: first blood sample by mean of samples two and three

Estradiol ($n = 92$)*

Quartiles for sample 1	Quartiles for mean of samples 2 and 3				
	1	2	3	4	
1	**14**†	9	1	3	27
2	7	**4**	8	1	20
3	3	5	**8**	7	23
4	0	5	5	**12**	22
	24	23	22	23	

Testosterone ($n = 80$)*

Quartiles for sample 1	Quartiles for mean of samples 2 and 3				
	1	2	3	4	
1	**12**	6	1	0	19
2	6	**8**	6	1	21
3	2	3	**11**	5	21
4	1	3	2	**13**	19
	21	20	20	19	

SHBG ($n = 112$)*

Quartiles for sample 1	Quartiles for mean of samples 2 and 3				
	1	2	3	4	
1	**21**	6	1	0	28
2	5	**13**	6	4	28
3	2	9	**12**	5	28
4	0	1	9	**18**	28
	28	29	28	27	

DHEAS ($n = 34$)*

Quartiles for sample 1	Quartiles for mean of samples 2 and 3				
	1	2	3	4	
1	**9**	0	0	0	9
2	0	**6**	2	0	8
3	0	2	**5**	2	9
4	0	0	1	**7**	8
	9	8	8	9	

NOTE: Quartiles are approximate because of a number of women with identical hormone values.
*Women missing one or two values are excluded from these analyses.
†Boldface, concordance between sample 1 and samples 2 and 3.

progesterone was 0.28—neither approached the coefficient previously observed. Progesterone secretion is somewhat pulsatile (36) and exhibits a rapid and steep increase and then a decrease in the luteal phase, factors that would contribute to the observed lower reproducibility. If, as seems most likely, the difference between the ICC observed in this and in our previous study is due to chance, then a combination of the estimates is most appropriate. Regardless, it remains likely that a single measurement of progesterone is at least valid for identification of anovulation within a single menstrual cycle as, in this instance, only very low values are being identified.

Whereas the absolute levels of IGF-I and IGFBP-3 were observed to be slightly greater in the luteal phase than in the follicular phase, these growth factors seem to be quite stable across the 3-year period with follicular ICCs of 0.7 and luteal ICCs of 0.8. Indeed, when IGF-I levels were averaged across the follicular and luteal phases, the ICC approached 0.9. These results are similar to a study where the ICCs for two samples collected 1 year apart in 59 women were 0.8 for IGF-I and 0.6 for IGFBP-3, although when analyses were restricted to premenopausal women, the ICC for IGF-I was 0.6 (37). Spearman rank correlations reported from reproducibility studies conducted across shorter periods (2-8 weeks) among men and women of ages 50 to 97 years ranged from 0.92 (38) to 0.97 (39).

For reference, cholesterol is generally accepted as measured reasonably well with a single blood sample, and the observed ICC over a 1- to 2-year period ranges from 0.65 to 0.76 (40-42). Correlations in this range result in relatively modest decreases in the estimated relative risk, although the degree of attenuation will depend on the magnitude of the relative risk and the sample size (28). For example, measurement error in a variable with an ICC of 0.68 will underestimate a true relative risk of 2.0 and 2.5 to 1.6 and 1.9, respectively (28). For a variable with an ICC of 0.45 (as we observed for the estrone and estradiol in the luteal phase), these true relative risks would be lowered further to 1.4 and 1.5. Besides providing important information on the reproducibility of a variable, ICCs can be used to correct relative risk estimates for random within-person measurement error in epidemiologic studies (30).

The follicular-phase and luteal-phase ICCs for estradiol and for estrone were similar, suggesting that measurement during neither phase is superior in this regard. Not surprisingly, averaging across the phases worsened the ICC. In addition, we calculated an ICC for each using all six sample levels, as this estimate of reliability mimics that underlying past case-control studies in premenopausal women that collected a random blood sample and did not match cases and controls on cycle day of collection. In fact, these ICCs were worse than those of the averaged values (estradiol, 0.02; estrone, 0.08). Interestingly, in both menstrual cycle phases, estrone sulfate was the most reliably measured estrogen, perhaps due to its longer half life compared with estrone and estradiol (31). Our results suggest that estrone sulfate may be the most accurate measure of estrogen levels among premenopausal women. Interestingly, the correlation of estrone sulfate with estradiol was only 0.16 and 0.07 in the follicular and luteal phases, respectively; however, this may be due to the small sample size.

These data among premenopausal women suggest that for androgens, estrone sulfate, prolactin, and IGFs, a single blood measurement can reliably categorize average levels over at least a 3-year period in premenopausal women and is valid for use in the investigation of the relation between endogenous hormone levels and disease risk. For estradiol and estrone, where ICCs were somewhat low, it will be particularly important to use these reproducibility data to correct relative risks (e.g., when assessing plasma hormones and breast cancer risk) for measurement error. We are currently using these data to correct relative risks in our ongoing study of the relation between endogenous premenopausal hormones and growth factors and breast cancer risk.

Acknowledgments

We thank Victor Pontes, Helena Judge Ellis, Jeanne Sparrow, Rachel Meyer, and Ellen Hertzmark for their expert assistance; Drs. Graham Colditz, David Hunter (Project Director of the NHSII cohort), and Walter Willett (Principal Investigator of NHSII) for their valuable input and insights; and the participants of the Nurses' Health Study II for their longstanding contribution.

References

1. Key T, Appleby P, Barnes I, et al. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst 2002;94:606–16.
2. Kaaks R, Lukanova A, Kurzer MS. Obesity, endogenous hormones, and endometrial cancer risk: a synthetic review. Cancer Epidemiol Biomarkers Prev 2002;11:1531–43.
3. Guthrie JR, Lehert P, Dennerstein L, Burger HG, Ebeling PR, Wark JD. The relative effect of endogenous estradiol and androgens on menopausal bone loss: a longitudinal study. Osteoporos Int 2004;15:881–6.
4. Rapuri PB, Gallagher JC, Haynatzki G. Endogenous levels of serum estradiol and sex hormone binding globulin determine bone mineral density, bone remodeling, the rate of bone loss, and response to treatment with estrogen in elderly women. J Clin Endocrinol Metab 2004;89:4954–62.
5. Potischman N, Hoover RN, Brinton LA, et al. Case-control study of endogenous steroid hormones and endometrial cancer. J Natl Cancer Inst 1996;88:1127–35.
6. Thijssen JH, Blankenstein MA. Endogenous oestrogens and androgens in normal and malignant endometrial and mammary tissues. Eur J Cancer Clin Oncol 1989;25:1953–9.
7. Helzlsouer KJ, Alberg AJ, Bush TL, Longcope C, Gordon GB, Comstock GW. A prospective study of endogenous hormones and breast cancer. Cancer Detect Prev 1994;18:79–85.
8. Secreto G, Toniolo P, Pisani P, et al. Androgens and breast cancer in premenopausal women. Cancer Res 1989;49:471–6.
9. Wysowski DK, Comstock GW, Helsing KJ, Lau HL. Sex hormone levels in serum in relation to the development of breast cancer. Am J Epidemiol 1987; 125:791–9.
10. Thomas HV, Key TJ, Allen DS, et al. A prospective study of endogenous serum hormone concentrations and breast cancer risk in premenopausal women on the island of Guernsey. Br J Cancer 1997;75:1075–9.
11. Michaud DS, Manson JE, Spiegelman D, et al. Reproducibility of plasma and urinary sex hormone levels in premenopausal women over a one-year period. Cancer Epidemiol Biomarkers Prev 1999;8:1059–64.
12. Lenton EA, Landgren BM, Sexton L. Normal variation in the length of the luteal phase of the menstrual cycle: identification of the short luteal phase. Br J Obstet Gynaecol 1984;91:685–9.
13. Lenton EA, Landgren BM, Sexton L, Harper R. Normal variation in the length of the follicular phase of the menstrual cycle: effect of chronological age. Br J Obstet Gynaecol 1984;91:681–4.
14. Hankinson SE, London SJ, Chute CG, et al. Effect of transport conditions on the stability of biochemical markers in blood. Clin Chem 1989;35: 2313–6.
15. Hankinson SE, Willett WC, Colditz GA, et al. Circulating concentrations of insulin-like growth factor-I and risk of breast cancer. Lancet 1998;351: 1393–6.
16. Hankinson SE, Willett WC, Manson JE, et al. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. J Natl Cancer Inst 1998;90:1292–9.
17. Hankinson SE, Willett WC, Michaud DS, et al. Plasma prolactin levels and subsequent risk of breast cancer in postmenopausal women. J Natl Cancer Inst 1999;91:629–34.
18. Abraham GE, Tulchinsky D, Korenman SG. Chromatographic purification of estradiol-17 for use in radio-ligand assay. Biochem Med 1970;3:365–8.
19. Mikhail G, Chung HW. Radioimmunoassay of plasma estrogens. Use of polymerized antibodies. In: Peron FG, Caldwells BV, editors. Immunological methods in steroid determination. New York: Appleton-Century-Croft; 1970. p. 113.
20. Mikhail G, Wu CH, Ferin M, Vande Wiele RL. Radioimmunoassay of plasma estrone and estradiol. Steroids 1970;15:333–52.
21. Kinouchi T, Pages L, Horton R. A specific radioimmunossay for testosterone in peripheral plasma. J Lab Clin Med 1973;82:309–16.
22. McNatty KP, Smith DM, Makris A, Osathanondh R, Ryan KJ. The microenvironment of the human antral follicle: interrelationships among the steroid levels in antral fluid, the population of granulosa cells, and the status of the oocyte *in vivo* and *in vitro*. J Clin Endocrinol Metab 1979;49: 851–60.
23. Buster JE, Abraham GE. Radioimmunoassay of serum DHEAS. Analytic Letters 1972;543.
24. Franz C, Watson D, Longcope C. Estrone sulfate and dehydroepiandrosterone sulfate concentrations in normal subjects and men with cirrhosis. Steroids 1979;34:563–73.
25. Carey VJ, Walters EE, Colditz GA, et al. Body fat distribution and risk of non-insulin-dependent diabetes mellitus in women. The Nurses' Health Study. Am J Epidemiol 1997;145:614–9.

26. Rosner B. Percentage points for generalized ESD many-outlier procedure. Technometrics 1983;25:165–72.
27. Rosner B. Fundamentals of biostatistics. Belmont (CA): Wadsworth Publishing Co; 1993.
28. Hankinson SE, Manson JE, Spiegelman D, Willett WC, Longcope C, Speizer FE. Reproducibility of plasma hormone levels in postmenopausal women over a 2-3-year period. Cancer Epidemiol Biomarkers Prev 1995;4:649–54.
29. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
30. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for random within-person measurement error. Am J Epidemiol 1992;136:1400–13.
31. Yen SSC, Jaffe RB. Reproductive endocrinology. Philadelphia: W.B. Saunders Company; 1991.
32. Muti P, Trevisan M, Micheli A, et al. Reliability of serum hormones in premenopausal and postmenopausal women over a one-year period. Cancer Epidemiol Biomarkers Prev 1996;5:917–22.
33. Toniolo P, Koenig KL, Pasternack BS, et al. Reliability of measurements of total, protein-bound, and unbound estradiol in serum. Cancer Epidemiol Biomarkers Prev 1994;3:47–50.
34. Koenig KL, Toniolo P, Bruning PF, Bonfrer JM, Shore RE, Pasternack BS. Reliability of serum prolactin measurements in women. Cancer Epidemiol Biomarkers Prev 1993;2:411–4.
35. Micheli A, Muti P, Pisani P, et al. Repeated serum and urinary androgen measurements in premenopausal and postmenopausal women. J Clin Epidemiol 1991;44:1055–61.
36. Filicori M, Butler JP, Crowley WF, Jr. Neuroendocrine regulation of the corpus luteum in the human. Evidence for pulsatile progesterone secretion. J Clin Invest 1984;73:1638–47.
37. Muti P, Quattrin T, Grant BJ, et al. Fasting glucose is a risk factor for breast cancer: a prospective study. Cancer Epidemiol Biomarkers Prev 2002;11:1361–8.
38. Milani D, Carmichael JD, Welkowitz J, et al. Variability and reliability of single serum IGF-I measurements: impact on determining predictability of risk ratios in disease development. J Clin Endocrinol Metab 2004;89:2271–4.
39. Goodman-Gruen D, Barrett-Connor E. Epidemiology of insulin-like growth factor-I in elderly men and women. The Rancho Bernardo Study. Am J Epidemiol 1997;145:970–6.
40. Shekelle RB, Shryock AM, Paul O, et al. Diet, serum cholesterol, and death from coronary heart disease. The Western Electric study. N Engl J Med 1981;304:65–70.
41. Willett W. Nutritional epidemiology. New York: Oxford University Press, Inc.; 1998.
42. Rosner B, Hennekens CH, Kass EH, Miall WE. Age-specific correlation analysis of longitudinal blood pressure data. Am J Epidemiol 1977;106:306–13.

Reproducibility of Plasma Steroid Hormones, Prolactin, and Insulin-like Growth Factor Levels among Premenopausal Women over a 2- to 3-Year Period

Stacey A. Missmer, Donna Spiegelman, Elizabeth R. Bertone-Johnson, et al.

Cancer Epidemiol Biomarkers Prev 2006;15:972-978.