Reproducibility of Plasma Hormone Levels in Postmenopausal Women over a 2–3-year Period[1]

Susan E. Hankinson,[2] JoAnn E. Manson, Donna Spiegelman, Walter C. Willett, Christopher Longcope, and Frank E. Speizer

Channing Laboratory [S. E. H., J. E. M., W. C. W., F. E. S.] and Division of Preventive Medicine [J. E. M.], Department of Medicine, Brigham and Women's Hospital and Harvard Medical School; Departments of Nutrition [W. C. W.], Epidemiology [S. E. H., D. S., W. C. W.], and Biostatistics [D. S.], Harvard School of Public Health, Boston, Massachusetts 02215; and the Departments of Obstetrics and Gynecology and Medicine [C. L.], University of Massachusetts Medical School, Worcester, Massachusetts

Abstract

We evaluated the reproducibility of plasma hormone levels over time in 79 healthy postmenopausal women, ages 51–69 years at baseline, who were not using postmenopausal hormones. Three blood samples were collected between 1989 and 1992 from each of these women. We assessed plasma levels of estradiol, free estradiol, percentage of free estradiol, bioavailable estradiol, percentage of bioavailable estradiol, estrone, estrone sulfate, sex hormone-binding globulin (SHBG), androstenedione, testosterone, dehydroepiandrosterone, dehydroepiandrosterone sulfate, and prolactin at each of three sample collections. The means and SD for each of the plasma estrogens, SHBG, and prolactin were similar at each collection. For the androgens, plasma levels tended to decrease over time consistent with an aging effect; decreases with increasing age were statistically significant for androstenedione, dehydroepiandrosterone, and dehydroepiandrosterone sulfate. Intraclass correlation coefficients (ICCs) ranged from 0.92 (95% confidence interval = 0.89–0.95) for SHBG to 0.53 (95% confidence interval = 0.43–0.69) for prolactin. Most correlations were at least 0.70. The ICCs did not vary by age or time since menopause. Women who changed weight over the course of the study tended to have lower ICCs for a number of the hormones, although these differences were not statistically significant. These data indicate that, for most of these plasma hormones, a single measurement can reliably categorize average levels over at least a 3-year period in postmenopausal women.

Introduction

Endogenous hormone levels are thought to influence risk of several major diseases in women, including breast (1) and endometrial cancer (2) and osteoporosis (3). Although average long-term hormone levels are often of primary interest in epidemiological studies of these relationships, for both economic and logistical reasons, frequently only one blood sample is collected per study subject. The degree to which a single sample reflects long-term hormone levels for an individual depends on the within-person variability of these levels over time. With greater variability, a single measurement will include a larger degree of measurement error and, as a consequence, observed associations (*e.g.*, relative risks) will be increasingly attenuated. Therefore, knowledge of how well one hormone measurement reflects longer term levels is important in conducting and interpreting epidemiological studies of these associations.

The within-person variation of plasma hormone levels in postmenopausal women has been addressed in few previous studies (4–7). Correlations for plasma estrogens over approximately a two-year period have ranged from 0.36 for estradiol (4) to 0.94 for percentage SHBG[3]-bound estradiol [100% − (percentage bioavailable estradiol)] (5). The only long-term study of plasma androgens found correlations of 0.77 for androstenedione and 0.74 for testosterone (7). Prolactin was reported to be reproducible with a correlation of 0.76 over a 2–4-year period (7). To our knowledge, the reproducibility of estrone sulfate, SHBG, DHEA, and DHEAS has not been published previously.

In the present study, we assess the reproducibility of plasma hormone levels in a subset of postmenopausal participants in the Nurses' Health Study, a large cohort study of women. We measured estradiol, free estradiol, percentage of free estradiol, bioavailable estradiol, percentage of bioavailable estradiol, estrone, estrone sulfate, SHBG, androstenedione, testosterone, DHEA, DHEAS, and prolactin over a 2–3-year period.

Materials and Methods

Blood Sample Collection. The NHS began in 1976 among 121,700 female registered nurses (8); the women have been followed biennially by mailed questionnaire since 1976. In 1989 and 1990, we collected blood samples from over 32,000 NHS participants. The women were 43–69 years of age at the time of baseline blood collection. Participants were sent a blood collection kit containing all supplies needed to have a blood sample drawn and returned to us (*e.g.*, needle, tourniquet, and blood collection tubes with sodium heparin). Each woman arranged to have her blood sample drawn and mailed to us, as whole blood, via overnight courier; samples were kept cool

Received 11/29/94; revised 5/16/95; accepted 5/18/95.

[1] This research was supported by National Cancer Institute Grants CA 40356 and CA 49449. Dr. Manson is a recipient of a Merck/Society for Epidemiological Research Grant Award.

[2] To whom requests for reprints should be addressed, at Channing Laboratory, 180 Longwood Avenue, Boston, MA 02215.

[3] The abbreviations used are: SHBG, sex hormone-binding globulin; DHEA, dehydroepiandrosterone; DHEAS, dehydroepiandrosterone sulfate; NHS, Nurses' Health Study; CV, coefficient of variation; BMI, body mass index; CI, confidence interval.

with an ice pack during transport. Ninety-seven % of the samples were received in our laboratory within 26 h of being drawn. The stability of estrogens, androgens, SHBG, and prolactin in whole blood for 24–48 h has been documented previously (9). On arrival in our laboratory the blood was centrifuged and aliquotted into plasma, RBC, and WBC components. The samples have been archived in continuously monitored liquid nitrogen freezers since collection.

Study Population. Three hundred-ninety NHS participants who gave a first blood sample in 1989–1990 were asked to collect two additional samples over the following 2 years. The women were postmenopausal, had not used postmenopausal hormones for at least 3 months, and had no previous diagnosis of cancer (except nonmelanoma skin cancer); these criteria were applied at each sample collection. Women were considered menopausal if they had no menses for at least 12 months and had a bilateral oophorectomy or, for women who had a hysterectomy without a bilateral oophorectomy, were at least 54 years of age if they were current smokers and at least 56 years if they were nonsmokers (these are the ages when natural menopause had occurred in 90% of the cohort). Of the 390 women, 186 (48%) sent two additional samples. A random sample of 80 of these women who had all three samples drawn between 6 a.m. and 12 p.m. was sent for hormone analysis. Ninety-four % of the samples were collected after a fast of at least 8 h.

We collected data on age and height on the 1976 study questionnaire. Menopausal status and smoking status have been asked on every biennial study questionnaire, and for menopausal status, it was also asked on the questionnaire completed at the time of blood collection. Data on current weight were obtained at the time of blood collection (for one woman who did not complete the questionnaire at the first blood collection, the weight reported on her 1990 study questionnaire was used). BMI (weight in kg/height in m^2) was used as the measure of obesity in these analyses. At the first blood collection, we had asked the women if they currently used steroids, antidepressants, or medications for a thyroid disorder. At the second and third sample collections, we asked women to list on their questionnaire all medications used in the previous week.

Laboratory Methods. With the exception of prolactin, SHBG, and estrone sulfate, which were assayed in the laboratory of Dr. Longcope, all plasma hormones were assayed by the Nichols Institute (San Juan Capistrano, CA). Estradiol (10), estrone (11), androstenedione (12, 13), testosterone (14), and DHEA (15) were assayed by RIA preceded by organic extraction and celite chromatography. DHEAS was assayed by RIA without a prior separation step (16). Percentage of free estradiol (*i.e.*, percentage nonprotein bound) was assayed by using equilibrium dialysis (17); the percentage of dialyzable estradiol was calculated as per the method of Vermuelen (17). Absolute levels of free estradiol were calculated by using the total estradiol and percentage of free estradiol. The percentage of bioavailable estradiol (*i.e.*, percentage of free plus percentage of albumin-bound estradiol) was assayed by using an ammonium sulfate precipitation (17, 18). Absolute levels of bioavailable estradiol were calculated by using the total estradiol and percentage of bioavailable estradiol. Estrone sulfate was assayed, after extraction of estrone, by RIA (of estrone) after enzyme hydrolysis, organic extraction, and separation by column chromatography (19). SHBG was measured by using an immunoradiometric kit from FARMOS Diagnostica (Orion Corp., Turku, Finland). Prolactin also was assayed by using a kit (Ciba-Corning, East Walpole, MA).

The assay detection limit was 2 pg/ml for estradiol, 0.5% for both percentage of free estradiol and percentage of bioavailable estradiol, 10 pg/ml for estrone, 5 pg/ml for estrone sulfate, 3 ng/dl for androstenedione, 2 ng/dl for testosterone, 10 ng/dl for DHEA, 5 μg/dl for DHEAS, 1 ng/ml for prolactin, and 6.25 nmol/liter for SHBG. In our data analyses, when plasma hormone values were reported as less than the detection limit, we set the value to one unit less than this limit [this occurred only for testosterone (all 3 values for 1 woman) and DHEAS (7 values among 4 women)].

All three samples from each woman were assayed in the same batch although the samples were ordered randomly in the batch and were labeled such that the laboratory could not identify which samples were from the same woman. In each batch of samples, we also interspersed replicate samples from a plasma pool that were labeled to preclude their identification by the hormone laboratory; these samples were used to assess laboratory precision. All inter- and intraassay laboratory CVs were $\leq$15% with the exception of estrone sulfate, which had an intraassay CV of 15% but an interassay CV of 28%. The higher interassay CV was due to a single run with high replicate values. Our results did not change when the 12 study samples in this run were excluded; thus, all samples are included in the results presented here.

Statistical Analysis. The natural logarithm of the plasma hormone values were used in the analyses as the transformed values were more normally distributed. To evaluate changes in mean hormone levels with age, we used the linear mixed model as implemented by SAS PROC MIXED,

$$\log(Y_{ij}) = \beta_0 + \beta_1 * \text{Age}_{ij} + b_i + \epsilon_{ij}$$

where Y_{ij} is the measured hormone value for participant i at sample collection j,

$$b_i \sim N(0, \sigma_B^2),\ \epsilon_{ij} \sim N(0, \sigma_W^2),\ \text{cov}(b_i, \epsilon_{ij}) = 0,$$

σ_B^2 is the between-person variance, σ_W^2 is the within-person variance, and *cov* = covariance.

By testing the hypothesis that $\beta_1 = 0$ in the usual manner, we assessed the evidence in the data for a significant change in hormone levels with increasing age.

We calculated ICC, where the ICC =

$$\frac{\sigma_B^2}{(\sigma_B^2 + \sigma_W^2)},$$

to assess reproducibility of hormone levels over time. The maximum likelihood estimates of σ_B^2 and σ_W^2 were obtained from linear mixed models as above to estimate the ICC. CIs (95%) for the intraclass correlations were calculated as shown in the Appendix. To assess change in the ICCs after accounting separately for age, years since menopause, and BMI, we calculated ICCs using the linear mixed models described above with age, time since menopause, and BMI included as continuous variables.

To assess the influence of weight change on the ICC, we conducted analyses among women who had a BMI change of $\leq$1 unit (ICC_R) and, separately, among those women who had a change of >1 unit (ICC_S). Stratum-specific estimates of σ_B^2 and σ_W^2 and their variance-covariance were obtained, and then to test H_0: $ICC_R = ICC_S$, the test statistic

$$Z^2 = \frac{(\hat{ICC}_R - \hat{ICC}_S)^2}{\text{var}(\hat{ICC}_R) + \text{var}(\hat{ICC}_S)}$$

Table 1 Mean (SD) plasma hormone levels from 79 postmenopausal women at each of the three sample collections

Hormone	Total no. of measurements[a]	Sample 1		Sample 2		Sample 3	
		Mean	(SD)	Mean	(SD)	Mean	(SD)
Estradiol (pg/ml)	221	7.4	(4.0)	7.8	(5.2)	7.7	(4.3)
Free estradiol (pg/ml)	221	0.12	(0.07)	0.13	(0.11)	0.13	(0.08)
% of free estradiol	237	1.6	(0.2)	1.6	(0.2)	1.6	(0.3)
Bio[b] estradiol (pg/ml)	203	2.6	(2.1)	3.0	(3.6)	2.8	(2.0)
% of bio estradiol	216	32.1	(13.6)	32.6	(13.1)	33.0	(13.8)
Estrone (pg/ml)	228	30.7	(12.9)	30.6	(12.3)	30.7	(12.6)
Estrone sulfate (pg/ml)	236	208	(122)	208	(116)	216	(113)
Androstenedione (ng/dl)	237	79.0	(34.9)	75.0	(39.9)	69.1	(33.6)
Testosterone (ng/dl)	237	24.0	(13.0)	22.1	(11.9)	21.8	(13.2)
DHEA (ng/dl)	237	249	(173)	230	(179)	205	(152)
DHEAS (μg/dl)	237	83.6	(63.3)	81.9	(72.2)	76.2	(68.7)
Prolactin (ng/ml)	236	11.3	(5.2)	11.1	(6.3)	11.3	(6.4)
SHBG (nmol/liter)	237	47.7	(27.5)	46.5	(27.9)	46.3	(28.5)

[a] Some values are less than 237 because of insufficient plasma for all analyses.
[b] Bio, bioavailable.

was referred to a χ^2 distribution with 1 degree of freedom where var = variance.

To quantify the influence of the observed measurement error in hormone levels on relative risk estimates from a cohort study, we calculated the relative risk that would be observed given relative risks of 1.5, 2.0, and 2.5 for the true exposure distribution. To obtain the observed relative risk, we multiplied the natural logarithm of the specified true relative risk by the ICC and exponentiated the result (20). This method assumes that the measurement error is random and that its variance is constant across hormone levels.

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.

Table 2 ICC and 95% CI for each plasma hormone based on three blood samples from each of 79 postmenopausal women

Hormone	ICC	95% CI
Estradiol (pg/ml)	0.68	(0.59–0.80)
Free estradiol (pg/ml)	0.73	(0.65–0.83)
% of free estradiol	0.80	(0.73–0.87)
Bio estradiol[a] (pg/ml)	0.78	(0.70–0.86)
% of bio estradiol[a]	0.86	(0.82–0.92)
Estrone (pg/ml)	0.74	(0.66–0.83)
Estrone sulfate (pg/ml)	0.75	(0.67–0.84)
SHBG (nmol/liter)	0.92	(0.89–0.95)
Androstenedione (ng/dl)	0.66	(0.57–0.78)
Testosterone (ng/dl)	0.88	(0.84–0.93)
DHEA (ng/dl)	0.75	(0.67–0.84)
DHEAS (μg/dl)	0.88	(0.84–0.93)
Prolactin (ng/ml)	0.53	(0.43–0.69)

[a] bio, bioavailable.

Results

Three blood samples from each of 80 women were sent for hormone analysis. However, we excluded from the data analyses one woman who had plasma testosterone levels approximately three times higher than the normal postmenopausal range. In addition, several women did not have a sufficient volume of plasma for all analyses. For this reason, only 78 women were included in the estrone, estradiol, and free estradiol analyses, 76 women were in the percentage of bioavailable estradiol analyses, and 74 women were in the bioavailable estradiol analyses. The following number of women were missing one of the three hormone values because of insufficient plasma: estrone sulfate (n = 1), prolactin (n = 1), estrone (n = 6), estradiol (n = 13), free estradiol (n = 13), percentage of bioavailable estradiol (n = 12), and bioavailable estradiol (n = 19). The total number of measurements per hormone is provided in Table 1.

At the first blood sample collection, women ranged in age from 51 to 69 years, with a mean age of 61 years. They had been postmenopausal for at least 2 years and for as long as 35 years. The average BMI was 25.9 kg/m^2 with a SD of 5.3 and a range from 17.8 to 43.7 kg/m^2. Only six (8%) of the women were current cigarette smokers. We also compared characteristics of these 79 women to the 390 randomly selected women initially invited to participate. The two groups were very similar in terms of age, BMI, and time since menopause. Women in the current analysis were somewhat less likely to be current smokers (8 *versus* 15%).

As shown in Table 1, the means and SDs for each of the plasma estrogens, SHBG and prolactin, were quite similar at each sample collection. For the androgens, plasma levels tended to decrease over time consistent with an aging effect. When we conducted regression analyses, using a continuous term for age, hormone levels decreased significantly with increasing age for androstenedione (8% decrease/5 year increase in age), DHEA (19% decrease/5 years), and DHEAS (25% decrease/5 years) only.

The ICC ranged from 0.92 (95% CI = 0.89–0.95) for SHBG to 0.53 (95% CI = 0.43–0.69) for prolactin (Table 2). Most correlations were at least 0.70. To assess to what extent reproducibility changed after accounting for the effects of age and time since menopause on between-person and within-person variability, we also calculated ICCs for each hormone controlling for these variables; all ICCs were essentially identical. When we controlled for BMI, the ICCs for several of the estrogens decreased, as might be expected given the contribution of obesity to the between-person component of variance in estrogen levels. The most substantial decreases were for the percentage of free estradiol (from 0.80 to 0.69) and bioavailable

Table 3 Intraclass correlation coefficients for each plasma hormone for postmenopausal women with (n = 32) and without (n = 47) a change in weight between the first and third blood sample collection

Plasma hormone	Women with ≤1 unit change in BMI	Women with >1 unit change in BMI	P value[a]
Estradiol (pg/ml)	0.80	0.51	0.08
Free estradiol (pg/ml)	0.84	0.57	0.07
% free estradiol	0.86	0.68	0.11
Bio estradiol[b] (pg/ml)	0.86	0.64	0.07
% bio estradiol[b]	0.90	0.80	0.07
Estrone (pg/ml)	0.81	0.64	0.14
Estrone sulfate (pg/ml)	0.77	0.70	0.45
SHBG (nmol/liter)	0.94	0.90	0.27
Androstenedione (ng/dl)	0.67	0.64	0.78
Testosterone (ng/dl)	0.90	0.85	0.30
DHEA (ng/dl)	0.70	0.79	0.29
DHEAS (μg/dl)	0.90	0.87	0.28
Prolactin (ng/ml)	0.60	0.41	0.21

[a] From χ^2 test of H_0: $ICC_1 = ICC_2$.
[b] Bio, bioavailable.

estradiol (from 0.78 to 0.66). ICCs for the androgens and prolactin were unchanged after controlling for BMI.

For prolactin, the plasma hormone with the lowest ICC, we recalculated the ICC after excluding the small number of women who fasted for <8 h before blood collection, as well as those women whose three samples were drawn at different times of day (defined as >2 h difference between any two samples), both potential contributors to the within-person component of variance. The ICC was essentially unchanged in these analyses. In addition, certain drugs are known to influence prolactin levels. When we recalculated the ICC (1) after excluding the 16 women who reported use of antidepressants, thyroid medications, or steroids on their first questionnaire or use of any medication suspected of influencing prolactin levels on their second or third questionnaires, or (2) after excluding these 16 women and using only data from the second and third sample collections for the remaining women, the ICC either was unchanged or somewhat decreased.

To assess the influence of a change in weight between the first and third blood collections on hormone reproducibility, we conducted analyses among women who had a BMI change of ≤1 unit and, separately, among those women who had a change of >1 unit. For a 5-foot 5-inch tall woman who weighed 159 pounds (the average height and weight in this sample), a unit increase in BMI (from 26 to 27 kg/m^2) would represent a 6-pound weight gain. The mean absolute change in BMI in the group with stable weights was 0.43 units, and in the group with a change in weight, 2.1 units. As shown in Table 3, women whose weight changed over the course of the study had lower ICCs for a number of the plasma hormones; this was particularly evident for the plasma estrogens. For example, correlation coefficients for estradiol and estrone were 0.80 and 0.81, respectively, among women with stable weights and 0.51 and 0.64, respectively, among women whose weight changed. Most of the difference in the ICCs was attributable to an increase in the within-person variance. None of these differences was statistically significant, however.

We calculated the estimated relative risk given true relative risks in the range 1.5–2.5, using the ICC of four of the hormones (chosen to represent the range of ICCs): SHBG, percentage of free estradiol, estradiol, and prolactin (Table 4). We observed little attenuation for SHBG and a moderate degree of attenuation for percentage of free estradiol and total estradiol. For example, for estradiol, the estimated relative risk was 1.3, 1.6, and 1.9, respectively, for true relative risks of 1.5, 2.0, and 2.5, a 15–25% decrease. The attenuation observed for prolactin was substantial and resulted in estimated relative risks of 1.2, 1.4, and 1.6, respectively.

Table 4 ICCs for the repeated measurement of plasma hormone levels and the resulting observed relative risk for specified true relative risks of 1.5–2.5

Plasma hormone	ICC	Observed relative risk given specified true relative risk		
		1.5[a]	2.0	2.5
SHBG (nmol/liter)	0.92	1.5[b]	1.9	2.3
Percentage of free estradiol	0.80	1.4	1.7	2.1
Estradiol (pg/ml)	0.68	1.3	1.6	1.9
Prolactin (ng/ml)	0.53	1.2	1.4	1.6

[a] True relative risk.
[b] Observed relative risk.

Table 5 Cross classification of postmenopausal plasma hormone levels: first blood sample by mean of samples two and three[a]

Estradiol (n = 64)[b]

		Quartiles for mean of samples 2 and 3				
		1	2	3	4	
Quartiles for sample 1	1	**12**[c]	4	1	1	18
	2	4	**5**	3	1	13
	3	0	7	**9**	1	17
	4	0	0	4	**12**	16
		16	16	17	15	

Testosterone (n = 79)

		Quartiles for mean of samples 2 and 3				
		1	2	3	4	
Quartiles for sample 1	1	**14**	4	0	0	18
	2	3	**11**	4	2	20
	3	2	5	**11**	3	21
	4	0	1	4	**15**	20
		19	21	19	20	

[a] Quartiles are approximate because of a number of women with identical hormone values.
[b] Fifteen women were missing one or two values for estradiol and are excluded from this analysis.
[c] Boldface, concordance between sample 1 and samples 2 and 3.

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 5 are matrices for two of the hormones of interest, estradiol (ICC = 0.68) and testosterone (ICC = 0.88). The quartiles are approximate because of a number of women with identical plasma hormone values. For estradiol, 38 of 64 (56%) were perfectly classified, 61 of 64 (95%) were off by one category or less and just 1 value (1.6%) was misclassified into an extreme category. Similar percentages for testosterone were 65, 94, and 0% (there was no extreme misclassification), respectively. 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.

Discussion

We observed correlations ranging from 0.92 to 0.53 for plasma hormone levels which were measured repeatedly over a 2–3-year period in postmenopausal women. These data indicate that, for most of the hormones assessed, a single plasma measure-

ment is sufficient to categorize long-term hormone levels in epidemiological studies. These correlations were essentially unchanged when age or time since menopause was accounted for in the analysis.

Although we observed no change in estrogen levels over time, we did observe significant decreases in several of the plasma androgens, *i.e.*, DHEA, DHEAS, and androstenedione. Decreases with increasing age have been reported for plasma DHEA and DHEAS (21–23,), whereas androstenedione levels in postmenopausal women are generally reported to decrease slightly or remain stable with increasing age (21, 23, 24). The changes in hormone levels are unlikely to be related to degradation of the stored samples because steroid and polypeptide hormone levels have been reported to be stable for 6 or more years in plasma stored in airtight vials and at −70°C or colder (5, 6, 22, 25). In addition, the observed androgen values were highest in the samples stored the longest.

These findings are likely to be representative of hormone levels, and their reproducibility, in our cohort of approximately 14,000 postmenopausal women (not on postmenopausal hormones) who provided blood samples. The distribution of age, obesity, and time since menopause in the 79 women was very similar to the 390 randomly selected women who were invited to participate. Several factors that might have decreased our ICCs by increasing the within-person component of variance were minimized in this study. We restricted all blood samples to morning collections (before noon), thereby reducing the circadian variation reported for prolactin (26) and testosterone (27). In addition, although not an inclusion criteria, 94% of our samples were drawn fasting; thus, we minimized the change in hormone levels associated with eating, such as is observed for prolactin (26). However, a majority of the samples we collected from our cohort also were drawn before noontime and after fasting.

Controlling for time since menopause did not alter the observed ICCs. This is not surprising, even for the estrogens, because all women in this study had been menopausal for at least 2 years, and 92% had been menopausal for at least 5 years. The substantial decrease in estrogens observed at menopause appears to stabilize within several years with little decrease thereafter (23, 24, 28). Controlling for age also did not influence our ICCs. Change in weight during the course of the study tended to decrease the level of reproducibility observed, particularly for the estrogens, although none of these differences was statistically significant. Most women who changed weight had gained rather than lost weight. Excluding the eight women who lost weight during the study had little effect on the ICC estimates in this weight change group.

For a number of the hormones, our results are similar to findings in previous studies. Our correlations for percentage of free estradiol and percentage of bioavailable estradiol are similar to those reported by Toniolo *et al.* (5). Micheli *et al.* (7) reported correlations of 0.77 for androstenedione and 0.74 for testosterone over a 1–3-year period; our correlations are qualitatively quite similar. Prolactin levels were observed previously to have a correlation of 0.76 over a 2–4-year period (6). However, when one woman with particularly high levels was removed from that analysis, the ICC decreased to 0.67, a somewhat higher value than what we observed. Our correlation did not increase when we restricted our analyses to women who had all samples drawn at the same time of day (±2 h) or after at least an 8-h fast, or when we excluded women who reported using medications that may influence prolactin levels.

The reproducibility of estradiol over a 1–2-year period has been reported to be 0.36 (4) and 0.51 (5), considerably lower than our ICC of 0.68. Similarly, a correlation of 0.56 was reported for estrone (4), which is lower than our estimate of 0.74. There are a number of factors that may have resulted in these differences. A change in weight over the course of our study tended to lower the ICCs, thus perhaps a greater proportion of women in the other studies changed weight resulting in a lower correlation. Alternatively, differences in assay sensitivity and specificity may have resulted in the measurement of somewhat different entities (that in turn may vary in their reproducibility). Assay differences are suggested by the substantial variation in mean postmenopausal estradiol levels between the studies: 4.1 pg/ml (4), 38.9 pg/ml (5), and 7.4 pg/ml (present study). Related to this, the between-person variance component also varied between the studies and could influence the ICC. For example, in the Cauley (4) study, the range in estradiol was 2.5–16.5 pg/ml and over 50% of the estradiol levels were less than the assay detection limit of 2.5 pg/ml. In contrast, only 8 (4%) of our estradiol values were <2.5 pg/ml. Although low precision in the laboratory assays will result in a larger within-person variance and, thus, lower ICCs, high precision in the assays was indicated in both previous reports and, thus, is unlikely to explain the different ICCs here. We also used precise assays: all intraassay CVs (and all but one interassay CV) were <15%. The laboratories we worked with had the lowest CVs among several academic and commercial laboratories evaluated (29). Of note, for several of the hormones, a commercial laboratory had the lowest CV.

These between-study differences in the ICC for estradiol and estrone suggest that investigators assessing hormone levels and disease should consider conducting a reproducibility study in their own study population with assays performed by the laboratory that will be analyzing their other study samples. These data also point to the difficulty in comparing studies of hormone levels and disease risk. In a recent study (30) where this issue was addressed, the authors observed that the hormone values provided by different laboratories not only varied but did not necessarily provide the same ranking of values, thus, indicating that the problem is not simply one of rescaling and raises questions of validity of different assays. The comparison of different laboratory methods against a gold standard would be helpful in resolving this issue, however, it is unclear which analytic method would be most appropriate as the gold standard.

It is useful to compare the level of reproducibility observed with hormone levels to that observed for other exposure measures of epidemiological interest. Correlations have been reported for serum cholesterol of 0.65 over 1 year (31) and 0.76 over 2 years (4), and for systolic blood pressure, of 0.60 over 2 years (4) and 0.64 over 4 years (32). The correlations observed for most of the plasma hormones compare very favorably to these. Correlations of approximately 0.65 or larger result in relatively modest decreases in the estimated relative risk, although the degree of attenuation that can be tolerated will depend on the size of the study. Correction for measurement error (or alternatively, obtaining more than one plasma sample for each study subject) is particularly important in the case of exposure measures, such as plasma prolactin (ICC = 0.53), where the level of error will result in appreciable attenuation of relative risk estimates.

These data indicate that in postmenopausal women, a

single plasma measurement can reliably categorize longer term hormone levels. Reproducibility data such as these also can be used to explicitly correct for measurement error in studies of plasma hormones and disease risk.

Appendix

$$\text{var}(\widehat{\text{ICC}}^{-1}) = \text{var}\left[\left(\frac{\hat{\sigma}_B^2 + \hat{\sigma}_W^2}{\hat{\sigma}_B^2}\right)\right] = \text{var}\left(1 + \frac{\hat{\sigma}_W^2}{\hat{\sigma}_B^2}\right).$$

By the delta method (33),

$$\widehat{\text{var}}\left[\left(1 + \frac{\hat{\sigma}_W^2}{\hat{\sigma}_B^2}\right)\right] \cong \frac{1}{(\hat{\sigma}_B^2)^2}\widehat{\text{var}}(\hat{\sigma}_W^2) + \frac{(\hat{\sigma}_W^2)^2}{(\hat{\sigma}_B^2)^4}\widehat{\text{var}}(\hat{\sigma}_B^2) - 2\frac{\hat{\sigma}_W^2}{(\hat{\sigma}_B^2)^3}\widehat{\text{cov}}(\hat{\sigma}_W^2, \hat{\sigma}_B^2)$$

where

$$\widehat{\text{var}}(\hat{\sigma}_W^2), \quad \widehat{\text{var}}(\hat{\sigma}_B), \quad \text{and} \quad \widehat{\text{cov}}(\hat{\sigma}_W^2, \hat{\sigma}_B^2)$$

are the usual variances and covariance of the maximum likelihood estimates of the corresponding parameters obtained from the linear mixed model as implemented by SAS PROC MIXED (34).

Then, a 95% CI for the ICC is:

$$\left(\widehat{\text{ICC}}^{-1} \pm 1.96\sqrt{\widehat{\text{var}}(\widehat{\text{ICC}}^{-1})}\right)^{-1}.$$

References

1. Bernstein, L., and Ross, R. K. Endogenous hormones and breast cancer risk. Epidemiol. Rev., *15:* 48–65, 1993.

2. Parazzini, F., LaVecchia, C., Bocciolone, L., and Franceschi, S. The epidemiology of endometrial cancer. Gynecol. Oncol., *41:* 1–16, 1991.

3. Cauley, J. A., Gutai, J. P., Sandler, R. B., LaPorte, R. E., Kuller, L. H., and Sashin, D. The relationship of endogenous estrogen to bone density and bone area in normal postmenopausal women. Am. J. Epidemiol., *124:* 752–761, 1986.

4. Cauley, J. A., Gutai, J. P., Kuller, L. H., and Powell, J. G. Reliability and interrelations among serum sex hormones in postmenopausal women. Am. J. Epidemiol., *133:* 50–57, 1991.

5. Toniolo, P., Koenig, K. L., Pasternack, B. S., Banerjee, S., Rosenberg, C., Shore, R. E., Strax, P., and Levitz, M. Reliability of measurements of total, protein-bound, and unbound estradiol in serum. Cancer Epidemiol., Biomarkers & Prev., *3:* 47–50, 1994.

6. Koenig, K. L., Toniolo, P., Bruning, P. F., Bonfrer, J. M. G., Shore, R. E., and Pasternack, B. S. Reliability of serum prolactin measurements in women. Cancer Epidemiol., Biomarkers & Prev., *2:* 411–414, 1993.

7. Micheli, A., Muti, P., Pisani, P., Secreto, G., Recchione, C., Totis, A., Fissi, R., Cavalleri, A., Panico, S., and Berrino, F. Repeated serum and urinary androgen measurements in premenopausal and postmenopausal women. J. Clin. Epidemiol., *44:* 1055–1061, 1991

8. Colditz, G. A., Willett, W. C., Stampfer, M. J., Rosner, B., Speizer, F. E., and Hennekens, C. H. Menopause and risk of coronary heart disease in women. N. Engl. J. Med., *316:* 1105–1110, 1987

9. Hankinson, S. E., London, S. J., Chute, C. G., Barbieri, R. L., Jones, L., Kaplan, L. A., Sacks, F. M., and Stampfer, M. J. Effect of transport conditions on the stability of biochemical markers in blood. Clin. Chem., *35:* 2313–2317, 1989.

10. Mikhail, G., and Chung, H. W. RIA of plasma estrogens. Use of polymerized antibodies. *In:* Immunological Methods in Steroid Determination, p. 113. 1970.

11. Judd, H. L., Lucas, W. E., and Yen, S. S. C. Serum 17 β-estradiol and estrone levels in postmenopausal women with and without endometrial cancer. J. Clin. Endocrinol. Metab., *43:* 272, 1976.

12. Gives, J. R. Hirsutism and hyperandrogenism. Adv. Intern. Med., *21:* 221, 1976.

13. Korth-Schutz, S., Virdis, R., Saenger, P., Chow, D. M., Levine, L. S., and New, M. I. Serum androgens as a continuing index of adequacy of treatment of congenital adrenal hyperplasia. J. Clin. Endocrinol. Metab., *46:* 452, 1978.

14. Kinouchi, T., Pages, L., and Horton, R. A. A specific RIA for plasma testosterone. J. Lab. Clin. Med., *82:* 309, 1973

15. DePeretti, E., and Forest, M. G. Unconjugated dehydroepiandrosterone plasma levels in normal subjects from birth to adolescence in human: the use of a sensitive RIA. J. Chem., *43:* 982, 1976.

16. Buster, J. E., and Abraham, G. E. RIA of plasma dehydroepiandrosterone sulfate. Anal. Lett., *5:* 543, 1972.

17. Moll, G. W., Jr., Rosenfield, R. L., and Helke, T. Estradiol-testosterone binding interactions and free plasma estradiol under physiological conditions. J. Clin. Endocrinol. Metab., *52:* 868–874, 1981.

18. Sodergard, R., Backstrom, T., Shanbhag, V., and Carstensen, H. Calculation of free and bound fractions of testosterone and estradiol-17β to human plasma proteins at body temperature. J. Steroid Biochem., *16:* 801–810, 1982.

19. Franz, C., Watson, D., and Longcope, C. Estrone sulfate and dehydroepiandrosterone sulfate concentrations in normal subjects and men with cirrhosis. Steroids, *34:* 563–573, 1979.

20. Rosner, B., Spiegelman, D., and Willett, W. C. Correction of logistic regression relative risk estimates and confidence intervals for random within-person measurement error. Am. J. Epidemiol., *136:* 1400–1413, 1992.

21. Roger, M., Nahoul, K., Scholler, R., and Bagrel, D. Evolution with aging of four plasma androgens in postmenopausal women. Maturitas, *2:* 171–177, 1980.

22. Orentreich, N., Brind, J. L., Rizer, R. L., and Vogelman, J. H. Age changes and sex differences in serum dehydroepiandrosterone sulfate concentrations throughout adulthood. J. Clin. Endocrinol. Metab., *59:* 551–555, 1984.

23. Meldrum, D. R., Davidson, B. J., Tataryn, I. V., and Judd, H. L. Changes in circulating steroids with aging in postmenopausal women. Obstet. Gynecol., *57:* 624–628, 1981.

24. Cauley, J. A., Gutai, J. P., Kuller, L. H., LeDonne, D., and Powell, J. G. The epidemiology of serum sex hormones in postmenopausal women. Am. J. Epidemiol., *129:* 1120–1131, 1989.

25. Cauley, J. A., Gutai, J. P., Kuller, L. H., and Dai, W. S. for the Multiple Risk Factor Intervention Trial Research Group. Usefulness of sex steroid hormone levels in predicting coronary artery disease in men. Am. J. Cardiol., *60:* 771–777, 1987.

26. Yen, S. S. C. Prolactin in human reproduction. *In:* S. S. C. Yen, R. B. Jaffe (eds.), Reproductive Endocrinology: Physiology, Pathophysiology and Clinical Management, pp. 357–388. Philadelphia: W. B. Saunders Co., 1991.

27. Panico, S., Pisani, P., Muti, P., Recchione, C., Cavalleri, A., Totis, A., and Berrino, F. Diurnal variation of testosterone and estradiol: a source of bias in comparative studies on breast cancer. J. Endocrinol. Invest., *13:* 423–426, 1990.

28. Longcope, C. Hormone dynamics at the menopause. Ann. NY Acad. Sci., *592:* 21–30, 1990.

29. Hankinson, S. E., Manson, J. E., London, S. T., Willett, W. C., and Speizer, F. E. Laboratory reproducibility of endogenous hormone levels in postmenopausal women. Cancer Epidemiol., Biomarkers & Prev., *3:* 51–56, 1993.

30. Potischman, N., Falk, R. T., Laiming, V. A., Siiteri, P. K., and Hoover, R. N. Reproducibility of laboratory assays for steroid hormones and sex-hormone binding globulin. Cancer Res., *54:* 5363–5367, 1994.

31. Shekelle, R. B., Shryrock, A. M., Paul, O. *et al.* Diet, serum cholesterol, and death from coronary heart disease: the Western electric study. N. Engl. J. Med., *304:* 65–70, 1981.

32. Rosner, B., Hennekens, C. H., Kass, E. H., and Miall, W. E. Age-specific correlation analysis of longitudinal blood pressure data. Am. J. Epidemiol., *106:* 306–313, 1977.

33. Rao, CR. Linear Statistical Inference and its Applications. New York: John Wiley & Sons, Inc., 1973.

34. SAS Users Guide: Version 6.09. Cary, NC: SAS Institute, 1994.