Saliva as a Medium for Investigating Intra- and Interindividual Differences in Sex Hormone Levels in Premenopausal Women[1]

Peter H. Gann,[2] Susan Giovanazzi, Linda Van Horn, Amy Branning, and Robert T. Chatterton, Jr.

Department of Preventive Medicine [P. H. G., S. G., L. V. H., A. B.], Robert H. Lurie Comprehensive Cancer Center [P. H. G., L. V. H., R. T. C.], and Department of Obstetrics and Gynecology [R. T. C.], Northwestern University Medical School, Chicago, Illinois 60611

Abstract

Repeated measurement of ovarian steroids in saliva could provide an advantage in studies estimating long-term sex steroid exposure in premenopausal women, by reducing the measurement error associated with collection of serum or urine samples. We previously reported on characteristics of ultrasensitive RIAs adapted for extraction-free measurement of estradiol (E_2) and progesterone (PG) in saliva. The purpose of the present study was to evaluate the consistency of E_2 and PG levels in saliva in the same women across menstrual cycles, and to compare this with the variation observed between women. We also evaluated the effect of altering the number of consecutive daily samples considered and the method for locating a particular cycle day in relation to ovulation (day 0). Study participants included 12 healthy women who provided daily saliva samples for two consecutive, ovulatory menstrual cycles. A single midluteal serum sample was collected 7–8 days after detection of a luteinizing hormone (LH) peak in urine. We plotted individual cycle profiles and computed intraclass correlation coefficients (ICC) for various definitions of peak and cumulative daily hormone level. For peak PG, determined as the maximal running 3-day mean, ICC was 0.68. For cumulative PG, based on 8 consecutive cycle days (+2 to +9), ICCs were 0.72–0.76 when reverse dating LH peak or rise in salivary PG determined day 0. For E_2, ICCs ranged from 0.74 to 0.79 by various dating methods for the 5 preovulatory days (-4–0), and from 0.85 to 0.92 for the 15 days about the center of the cycle (-6 to +8). With exclusion of just the first 5 days of the cycle, the ICC for E_2 was 0.91. For both E_2 and PG, selection of 5 or 7 days for the estimation of the midluteal mean level provided separation of within and between subject variance that was comparable with a LH-timed serum sample. These results indicate that daily saliva samples can be combined to clarify the interindividual differences in E_2 and PG levels in premenopausal women, and that these interindividual differences may be greater than previously imagined.

Introduction

According to the dominant paradigm, a woman's risk of developing breast (and perhaps endometrial) cancer is affected by long-term exposure to ovarian sex steroid hormones, primarily E_2[3] and PG (1). Interindividual differences in cumulative exposure to these steroids are, therefore, presumably important. It has proved difficult, however, to test aspects of this model in either etiological or intervention studies involving premenopausal women because of the fluctuation in these hormone levels across the menstrual cycle (2). Previous studies have reported a low correlation between two luteal-phase serum E_2 samples obtained from the same woman over a 1-year interval, although sampling occurred approximately the same number of days after the onset of menses (3, 4). Standardized timing of blood samples can be improved by linking sampling to detection of ovulation or LH peaks, but such testing can be elaborate and/or time-consuming. We have adapted ultrasensitive RIAs for the direct measurement of E_2 and PG in saliva, and have previously reported on the sensitivity, reliability, and serum-saliva correlations for these assays (5). Saliva has several advantages over blood as a sampling medium: it can be easily collected by subjects themselves at repeated intervals; it requires no special collection or storage equipment; and the steroid concentrations measured exclude the fraction tightly bound to serum proteins and thus unavailable for biological action. Most importantly, consecutive daily samples can be grouped for analysis after the length of the menstrual cycle is known.

The aim of the present study was to determine how salivary levels of E_2 and PG track within women from cycle to cycle and to evaluate how cumulative levels of salivary steroid, based on daily sampling, can be used to discriminate interindividual from intraindividual differences. The ultimate goal is to develop a protocol for salivary measurement that reflects sustained exposure of target tissues to ovarian steroids, and thus provides a tool either for etiological studies relating this expo-

Received 6/14/00; revised 10/24/00; accepted .
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.
[1] This work was supported by National Cancer Institute Grants RO1 CA66695 and KO7 CA66185, United States Army Medical Research and Materiel Command Grant DAMD 17-94-J-4203, National Center for Health Resources (NIH) Grant RR-00048, and the Carol Gollob Breast Cancer Research Foundation.
[2] To whom requests for reprints should be addressed, at Department of Preventive Medicine, Northwestern University Medical School, 680 North Lake Shore Drive, Suite 1102, Chicago, IL 60611. Phone: (312) 908-8432; Fax: (312) 908-9588; E-mail: pgann@northwestern.edu.

[3] The abbreviations used are: E_2, estradiol; PG, progesterone; LH, luteinizing hormone; GCRC, General Clinical Research Center (at Northwestern Memorial Hospital); ICC, intraclass correlation coefficient; CV, coefficient(s) of variation.

sure to disease risk or for experimental studies of dietary or pharmacological interventions designed to reduce exposure.

Materials and Methods

Study Population. We recruited female participants from the Chicago area with the following eligibility criteria: ages 20–40, no use of exogenous hormones within 6 months, regular menstrual cycles, and at least 12 months since previous childbirth or lactation. Potentially eligible participants were invited to the GCRC at Northwestern University Medical School, where specially trained staff reviewed eligibility, obtained informed consent, and collected baseline data on body size and past medical history. At the initial visit, participants also received supplies and instructions regarding collection of the biological samples during the ensuing two consecutive menstrual cycles. A total of 20 women were recruited, and 19 completed the two-cycle protocol.

Sample Collection. During the two menstrual cycles, each participant provided samples of saliva, urine, breast fluid, and blood. Beginning on the first day after the onset of menstrual bleeding, participants deposited their own saliva, while at home, into sequentially numbered empty plastic vials that were stored in cardboard boxes in the home freezer. The women were instructed to collect the saliva daily between 8 and 11 p.m. Toothbrushing prior to sample collection was forbidden, to reduce the risk of contaminating the saliva sample with blood. After thorough rinsing of the mouth with water, participants chewed sugarless gum to increase salivation while depositing 7–10 ml of saliva into the appropriate vial. The gum was demonstrated previously to have no effect on salivary steroid measurements (6). Saliva collection was continued through the onset of vaginal bleeding marking the end of the second menstrual cycle. Completed boxes of saliva samples were returned to the GCRC, where they were immediately catalogued and stored at −20°C.

For each participant, study staff calculated the expected day of ovulation based on the usual menstrual cycle length. Five days prior to this predicted day, participants began testing their first-void morning urine specimens using the Ovu-Quick cassette for urinary LH (Quidel, San Diego, CA). Previous studies have validated this method for the detection of LH in urine (7). On the day the LH peak was detected, participants noted the day on a calendar provided and called the GCRC to schedule a midluteal appointment for 7 days after the LH peak. Midluteal appointments falling on Sunday were scheduled for the following day. In the event of a failure to detect a clear LH peak, participants were instructed to come to the GCRC 7 or 8 days after the predicted ovulatory day based on the usual cycle length. Venous blood samples were drawn in the morning after an overnight fast. Serum was separated and stored at −70°C.

Laboratory Methods. All of the samples from a given individual were assayed in duplicate in the same run. Unidentifiable quality control samples were inserted in each batch to allow measurement of blinded intra- and interassay CV. Before analysis, frozen saliva was thawed and centrifuged at 1500 × *g* for 1 h, with the supernatant saved for assay. Salivary PG was measured with a modification of the competitive immunoassay described previously (8). Briefly, this assay uses tritium-labeled PG and an antiserum prepared by one of the authors (R. T. C.) that has known cross-reactivities of 0.5% with 17-hydroxy PG, 2.9% with pregnenelone, 0.95% with corticosterone, 11.4% with 5α-pregnanedione, 1.9% with 5β-pregnanolone, and <0.1% with seven other steroids tested. Standards were prepared in a special gelatinized buffer [0.1 M PBS (pH 7.0), containing 0.015 M NaN_3 and 0.1% gelatin]. The volume of sample used was 0.2 ml. Intra-assay and interassay CV were 13.5 and 18%, respectively.

Salivary E_2 was measured with a double antibody RIA, also described previously (5). Antiserum and ^{125}I-labeled E_2 tracer were obtained from Diagnostic Services Laboratories (Webster, TX). The antiserum has cross-reactivities of 2.4% with estrone, 0.01% with estrone sulfate, 0.21% with 16-keto-E_2, 2.6% with E_2-3-glucuronide, 0.64% with estriol, and <0.1% with nonphenolic steroids tested. In the assay, the antiserum was diluted to give ~40% binding. Standards were prepared by diluting a methanolic stock solution of E_2 with the same gelatin buffer used for PG. A precipitating antibody solution was prepared by titrating the amount of sheep antirabbit gamma globulin required for precipitation of 0.1 ml of rabbit serum, and adding this to propylene glycol (4.8 g/dl). The total volume of sample required was 0.4 ml. Intra- and interassay CV were 9.9 and 11.6%, respectively.

Data Analysis. We assayed salivary PG in all available samples, except those collected in the first 10 days of the cycle, from all of the 19 women who completed both cycles. For each cycle, midcycle day 0 was initially defined in two ways: as the day the urine LH peak was detected or by reverse dating, as the day located 14 days prior to onset of the subsequent menses. After daily salivary PG levels were measured, we identified the day 0 by onset of the rise in PG, defined as the middle day of the first 3-day running mean with an increase of at least 20% over the previous 3-day set, and an average daily level of at least 10 pg/ml. For PG, we defined two consecutive daily segments of interest *a priori*. The first was the 8-day segment from day +2 to day +9, and the second was the 1-, 3-, 5-, or 7-day segment surrounding the midluteal center. Midluteal center was defined as the midpoint between day 0 and the end of the cycle. These definitions were arbitrary but were chosen to include as many samples as possible while accommodating women with relatively short menstrual cycles. We computed 3-day running means for PG, and we defined peak PG as the middle day of the highest running mean.

If a single daily saliva sample was missing, we imputed a value by averaging levels from the preceding and following days. If two or more consecutive daily samples were missing, or if a given segment contained more than two missing samples, we excluded the segment from the analysis. Of the 19 women whose salivary PG levels were assayed, two had anovulatory cycles (one each) identified by the absence of LH peak and a serum PG of <3 ng/ml. Five other women had at least one cycle with segments disqualified by missing samples. Two women had one cycle apiece in which LH was not detected in the urine, although serum and salivary PG suggested that ovulation had occurred.

For E_2, we defined four segments of interest: the 5-day segment from day −4 to day 0 just prior to ovulation, the 15-day segment from day −6 to day +8, the mean daily E_2 from all samples excluding the first 5 days, and the midluteal segment encompassing 1, 3, 5, or 7 days. We assayed all of the available saliva samples from 10 women, excluding those with anovulatory cycles or excessive numbers of missing samples. Peak E_2 was not determined because a single peak was not expected.

We used repeated measures ANOVA (SAS-PC, SAS Institute, Inc., Cary, NC) to compute between-person and within-person variances. ICCs were computed as the between-person variance divided by the sum of between- and within-person

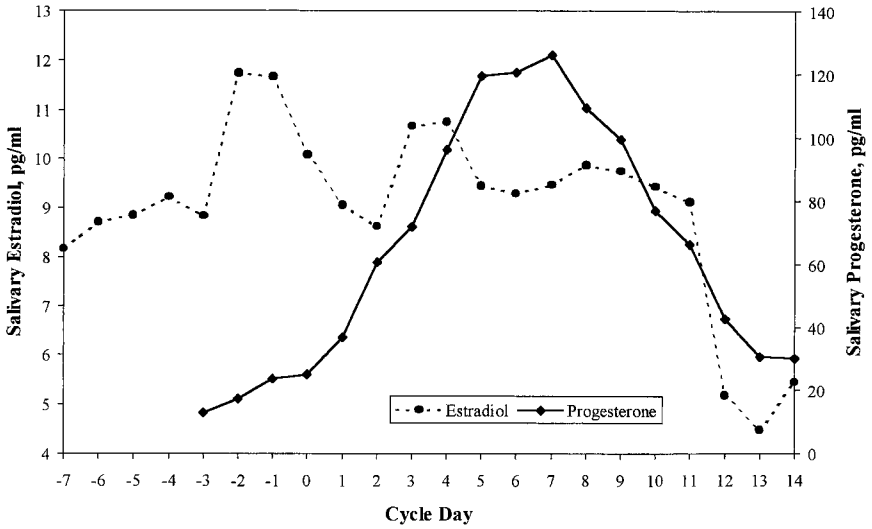

Fig. 1. Mean daily salivary E_2 and PG concentrations aligned with midcycle day 0 determined by urine LH testing (10 women, 20 cycles).

variances (9). We calculated confidence intervals (95%) for the ICCs based on the expected distribution of the *F* statistic (9).

Results

Among the 12 participants whose cycles were analyzed, mean age was 35 years (range, 26–39 years), 6 were nulliparous, mean body mass index was 23.7, and mean age at menarche was 12.6 years. Mean daily saliva concentrations of E_2 and PG across the menstrual cycle, aligned with the urine LH peak as day 0, are shown in Fig. 1. The E_2 and PG curves correspond to the daily profiles normally observed in blood but at far lower concentrations and with peaks of smaller amplitude. Day 0 by LH testing corresponded to the day immediately before the onset of the rise in mean PG, and to the day immediately after the steep midcycle drop in mean E_2. Fig. 2 shows daily profiles of E_2 and PG for consecutive cycles from selected individual participants, aligned by LH peak. These representative curves show greater unexplained day-to-day variation in E_2, leading to profiles that contain a discernible preovulatory peak, but few other features typical of a classical serum profile. For PG, however, salivary profiles always contained a distinct rise in PG at the onset of the luteal phase, followed by a relatively smooth, broad luteal peak. Individual and group cycle profiles were similar when reverse dating or rise in PG was used for alignment. Interindividual differences in cumulative levels of both E_2 and PG from cycle to cycle are readily apparent.

Cycle-to-cycle variability in peak and cumulative salivary PG is summarized in Table 1. For 24 eligible cycles, the ICC for the peak level of PG was 0.68, meaning that only 32% of the total variance in peak PG was attributable to within-woman variation. For cumulative PG measured over 8 consecutive luteal days (+2 to +9), between-woman differences accounted for three times as much total variance as within-woman differences. These ICCs are all significantly different from the null hypothesis of equal within- and between-woman variance ($P < 0.01$).

Table 2 shows the comparison of within- and between-cycle variance for various cumulative measures of salivary E_2. For a segment of 5 consecutive preovulatory days (−4 to 0), ICCs ranged from 0.74 to 0.79, depending on the method of locating the midcycle day. For a 15-day segment from the center of the cycle, the ICCs were higher, ranging from 0.85 to 0.92 by method of alignment. For the mean daily E_2, computed across the entire cycle excluding only the first 5 days, the ICC was 0.91. Again, all of the ICCs in this table have a very low likelihood under the null hypothesis of equal within- and between-woman variance.

The effects on the ICC of varying the number of salivary samples used to estimate the mean midluteal steroid levels are shown in Table 3. The center of the midluteal period was defined as the midpoint between the midcycle day 0 and the end of the cycle. As expected, ICCs generally rise for both PG and E_2 as one progresses from considering 1 to 7 days in estimating the mean. The ICCs for a single midluteal serum sample, obtained 7–8 days after appearance of the urine LH peak, were 0.77 and 0.81 for PG and E_2, respectively.

Discussion

We have found that measurement of consecutive daily salivary E_2 and PG provides estimates of cumulative and peak concentrations that are consistent within women from one cycle to the next, and are relatively distinct between women. The ICC for peak PG was 0.68 and ranged from 0.72–0.76 for an 8-day luteal segment, depending on the method used for locating ovulation. A single midluteal saliva sample generally gave low ICCs for PG, but increasing the number of consecutive daily samples considered to 3, 5, or 7 produced ICCs that were comparable with that observed with a single serum carefully timed in relation to the urine LH peak. For E_2, we observed substantial unexplained day-to-day within-person variation in salivary concentration, but despite this, the combination of consecutive daily saliva samples provided ICCs comparable with, or greater than, that provided by an LH-timed serum sample. When all but the first 5 cycle days were considered, only 9% of the total variance in mean daily salivary E_2 was attributable to within-woman variance. This represents much higher within-person correlation than observed for long-established measures used in population studies, such as serum total cholesterol measured 1 year apart (*r*, 0.65; Ref. 10).

It appears that the variation in daily salivary E_2 coexists with an even greater variation in time-integrated levels between women. Our data indicate that repeated sampling of saliva can be used to reveal potentially important differences in area-under-the-curve or time-integrated exposure of individual pre-

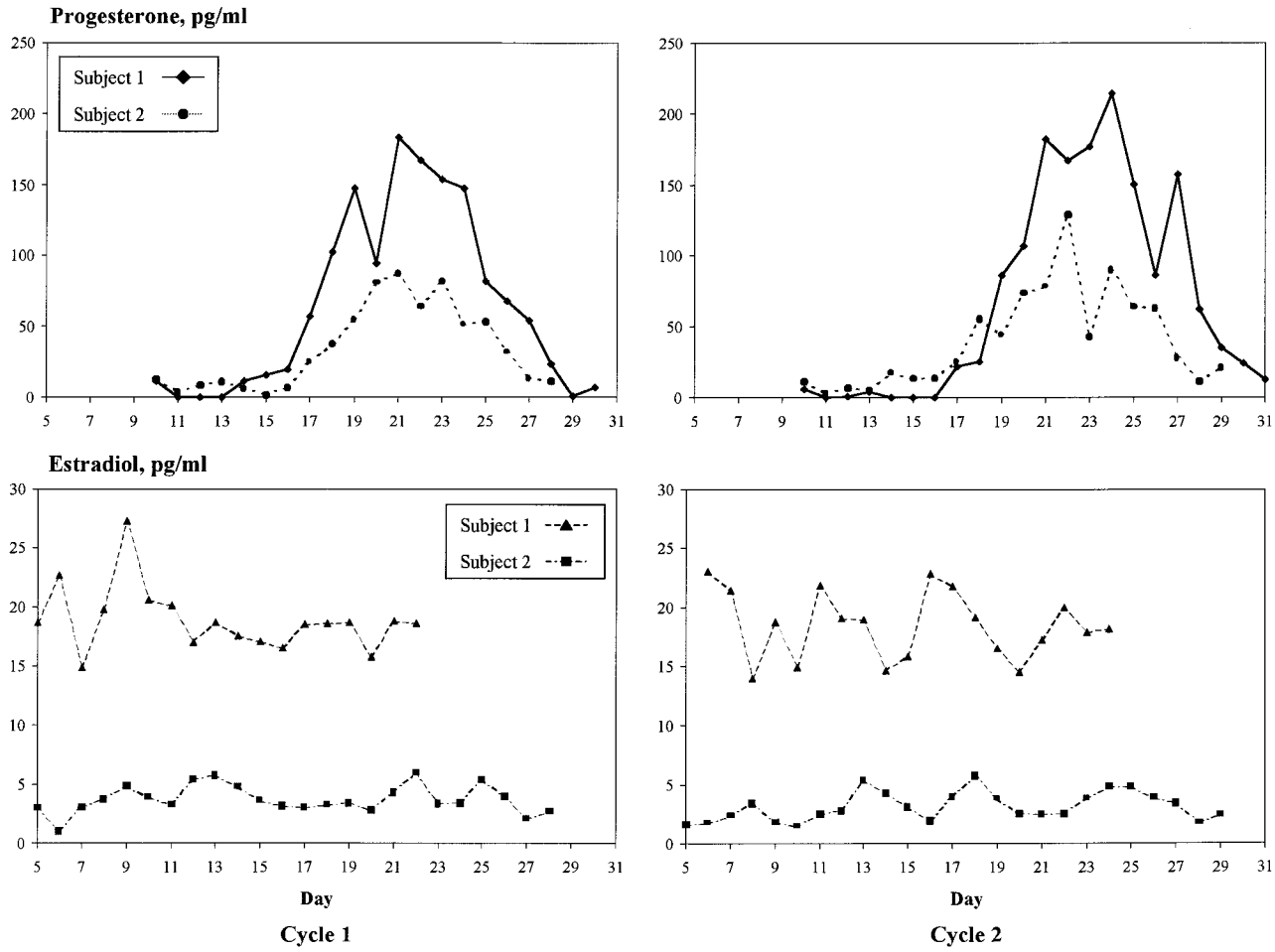

Fig. 2. Individual daily salivary PG and E_2 profiles for pairs of cycles for selected subjects.

Table 1 Variation within *versus* between women in peak and cumulative salivary PG, measured during two menstrual cycles

	No. of cycles (subjects)	First cycle mean, pg/ml (s.e.)	ICC	95% CI[a]
Peak PG (3-day running mean)	24 (12)	120.2 (5.8)	0.68	0.22–0.89
Cumulative PG (days +2 to +9); midcycle (day 0) set by:				
Reverse dating[b]	24 (12)	722.6 (55.5)	0.72	0.30–0.91
Urine LH testing	20 (10)	776.0 (67.8)	0.76	0.32–0.93
Rise in salivary PG	22 (11)	754.3 (61.8)	0.76	0.37–0.92

[a] CI, confidence interval.
[b] Reverse dating identifies day 0 by subtracting 14 days from the onset of the subsequent menses.

menopausal women to endogenous sex steroids. The importance of even small differences in usual daily exposure, relative to cancer risk, are amplified because such differences are repeated month after month during a woman's reproductive years (11). Cross-sectional studies have reported substantially higher ovarian steroid hormone profiles in affluent Western women compared with women from nonindustrialized cultures, which provides evidence that the former have a high cumulative exposure to ovarian steroids attributable to higher exposure during a typical cycle as well as to a greater total number of menstrual cycles (12). Within some agrarian cultures such as the Lese of Africa, seasonal variation in ovarian hormone levels suggests that large changes in caloric intake could be an important factor (13). In the U.S., one study involving measurement of estrogen and PG in daily urine samples from a single cycle in 175 women found lower PG levels in women with early menarche and higher body weight and lower estrogen levels in cigarette smokers but no other notable associations with reproductive variables associated with breast cancer risk (14). Regular exercise, whether vigorous or moderate, has also been associated with lower salivary PG profiles (15). Polymorphisms in genes encoding key enzymes involved in sex steroid synthesis or metabolism are also under investigation as possible determinants of interindividual differences in long-term exposure (16). Thus far, however, it appears that most of the inter-woman variation in ovarian steroid levels remains unexplained.

Table 2 Variation within *versus* between women in cumulative and mean daily E_2, measured during two menstrual cycles

	No. of cycles (subjects)	First cycle mean, pg/ml (SE)	ICC	95% CI[a]
Cumulative E_2				
Preovulatory (days −4 to 0)				
Reverse dating	18 (9)	51.5 (7.5)	0.79	0.35–0.95
Urine LH testing	16 (8)	55.1 (7.9)	0.77	0.27–0.95
Rise in salivary PG	20 (10)	51.9 (6.8)	0.74	0.27–0.93
Center cycle (days −6 to +8)				
Reverse dating	20 (10)	148.8 (19.4)	0.85	0.54–0.96
Urine LH testing	16 (8)	152.5 (23.7)	0.92	0.68–0.98
Rise in salivary PG	20 (10)	148.1 (19.3)	0.86	0.56–0.96
Mean daily E_2 (entire cycle minus first 5 days)	20 (10)	9.5 (1.4)	0.91	0.68–0.98

[a] CI, confidence interval.

Table 3 ICCs for midluteal[a] PG and E_2, measured by a single, LH-timed serum sample *versus* varying numbers of consecutive saliva samples

	No. of consecutive daily samples			
	1	3	5	7 (95% CI)[b]
Saliva PG				
Reverse dating	0.21	0.63	0.65	0.66 (0.20–0.89)
Urine LH	0.73	0.83	0.68	0.73 (0.22–0.93)
Rise in salivary PG	0.16	0.58	0.75	0.81 (0.42–0.95)
Serum progesterone[c]				
Urine LH (95% CI)	0.77 (0.41–0.92)			
Saliva E_2				
Reverse dating	0.27	0.35	0.53	0.67 (0.15–0.90)
Urine LH	0.26	0.41	0.48	0.62 (0.03–0.90)
Rise in salivary PG	0.23	0.32	0.60	0.69 (0.19–0.91)
Serum E_2[c]				
Urine LH (95% CI)	0.81 (0.46–0.95)			

[a] Midluteal day determined as midway between day 0 and end of menstrual cycle.
[b] CI = confidence interval.
[c] Based on single fasting blood sample obtained 7–8 days after urine LH peak.

We are aware of no previous studies comparing within- and between-woman variability in salivary E_2. A single study of salivary PG in six cycles from each of eight women reported that interindividual variance was approximately three times the intraindividual variance, a result similar to ours (17). Several studies, however, have examined reproducibility of sex steroid hormone levels in serum or urine samples. Studies involving repeated measures in blood are few and tend to be small because of the obvious practical constraints on performing many serial venipunctures. The largest such study measured daily plasma hormone levels during two cycles from 17 women (18). These investigators calculated the within-person correlation between cycles, roughly equivalent to the ICC, for plasma E_2 and PG during the interval +2 to +8 days after the LH peak. The correlation for PG was 0.80, very close to the estimate of 0.76 that we obtained in saliva for nearly the identical interval. The between-cycle correlation for plasma E_2 was 0.49; although we did not assess this interval in saliva, the plasma result is very comparable with the ICCs that we obtained for a 5-day midluteal segment (0.48–0.60) and is considerably lower than we obtained for longer cycle segments.

In a study involving a single-luteal-phase serum sample, repeated 1 year later, Muti *et al.* reported an ICC for E_2 of only 0.06 (3). More recently, Michaud *et al.*, using a similar design, obtained an ICC of only 0.19 for a single luteal serum sample (4). However, these investigators timed each blood sample collection to coincide with the anticipated midluteal day based on the individual's usual cycle length and also had subjects report back on the onset of the subsequent menstrual cycle. They found that with the exclusion of women with probable anovulatory cycles or samples obtained outside a 4–10-day window before the subsequent period, the E_2 ICC increased to 0.62. We note that our results, which also excluded anovulatory cycles, indicate a somewhat higher ICC (0.81) for a single serum sample timed more precisely, 7–8 days after detection of the urine LH peak. The method for timing a single blood sample used by Michaud *et al.* would have obtained a usable blood sample in 63% of all cycles. Use of saliva might exclude fewer cycles and provide higher ICCs over long segments, but it is not clear that these advantages outweigh the potentially higher costs of collecting daily saliva samples in all conceivable study applications.

Measurement of E_2 and PG metabolites in daily first-void urine samples provides some of the same benefits as salivary measurements in field study settings. Baird *et al.* (19) demonstrated that monitoring the urinary ratio of estrone-3-glucuronide and pregnanediol-3-glucuronide can be used to estimate the day of ovulation. Other investigators have reported menstrual cycle profiles for urinary E_2 and pregnanediol glucuronides, aligned by basal body temperature readings, which resemble those obtained in serum (14). Finally, the feasibility of daily urine collection was demonstrated in a study of 403 California women, which used creatinine-adjusted estrogen and PG metabolite levels to examine variation in follicular and luteal phase lengths (20). Urine samples require measuring conjugated metabolites one or more steps removed from the active hormone and, thus, could introduce possible additional error attributable to extraneous variations in metabolic activity.

Previous research has documented several advantages of saliva over blood or urine as a medium for frequent measurement of hormone levels (21). Salivary samples can be collected over a matter of minutes by subjects themselves and can be stored conveniently in home freezers, although studies have shown that most steroids remain stable for days in saliva even at room temperature (6). Because it is easy to collect samples over an entire cycle, it is possible to select samples for assay afterward, once the length of the cycle is known. E2 and PG enter saliva by diffusion and represent the fraction of steroid not bound to carrier proteins. In principle, this means that salivary levels could provide a better reflection of the diffusable fraction available to target tissues such as the breast; however, this remains to be fully investigated. The salivary PG assay has been established for some time; our method is precise at levels as low as 15 pg/ml and provides a good correlation between synchronous saliva and serum samples (r, 0.80). Development of the E_2 assay, however, proved frustrating for many investigators. The evolution of immunoassay techniques permitted us to develop an assay for E_2 with good sensitivity and precision and no required extraction step (5). We found that serum-saliva correlations for E_2 (measured in serum as total or non-SHBG-bound fraction) were not very high (r, 0.21) when serum-saliva pairs from many women were compared. However, median correlation within women was substantially higher (r, 0.71), suggesting that the ability of E_2 to move from serum to saliva varies between women.

Apart from the use of sensitive and precise direct assays for salivary hormones, the strengths of this study include estimation of ovulation by urine LH peak and rise in salivary PG as well as by reverse dating. Locating ovulation by detecting the rise in daily salivary PG is an established technique (22), as

is detection of the midcycle decrease in salivary E_2 (23). We found that the rise in salivary PG, the drop in salivary E_2, and reverse dating all gave similar results to LH testing in identifying midcycle day 0. In only one cycle did the selected day 0 differ by more than 3 days between any two methods. Measurement of individual daily saliva samples is relatively expensive. Therefore, in some field settings, it might be more practical to use self-detection of the urine LH peak; and, in fact, our data show that a single serum sample collected at a fixed interval after LH detection is highly reproducible across cycles. However, some women with ovulatory cycles are not able to detect an LH peak with this method, and the daily urine testing around midcycle places a definite burden on study subjects and staff. The ICCs that we observed for location of saliva samples by reverse dating were comparable with those obtained with LH or rise in PG dating. Therefore, we believe that it is most practical to collect samples from an entire cycle and then select samples from a broad consecutive segment located by reverse dating. The required sample volumes for a single assay for E_2 and PG are 400 μl and 200 μl, respectively. It is possible to pool equal aliquots from consecutive days and, thus, perform one assay to measure the cumulative or mean daily level over the segment.

We recognize that our data also have several limitations. The study size was not large, in part because it was necessary to assay many samples per subject in this methodological effort. In addition, we excluded women with anovulatory cycles or excessive missing samples. The criteria for missing samples were quite strict; it is possible that they could be relaxed to include more women and still provide adequate reproducibility, particularly for long cycle segments. Our definitions of cycle segments were somewhat arbitrary. In general, we would expect ICCs to increase as the number of samples that are considered increases. On the other hand, segments that are too long could exclude some women with unusually short cycles. Our study sampled saliva from two consecutive menstrual cycles. It is certainly possible that cumulative levels of sex steroids are more highly correlated in consecutive cycles than those farther apart in time. Individual women do have variations in E_2 from cycle to cycle, and this can even play a role in fecundity during specific cycles (23). However, Lenton *et al.* (18) reported that the separation between women in plasma PG during the interval +2 to +9 days after the LH peak was easily apparent even when they examined five cycles per woman spanning 3–6 years. Studies ongoing within our group will analyze reproducibility of consecutive daily segments from four cycles occurring over a one-two year period.

In conclusion, these results support the use of salivary measurements for exploring changes in cumulative exposure to endogenous E_2 and PG in premenopausal women. As a recent study has shown, a single midluteal serum sample could be adequate in a large prospective cohort setting in which reverse dating can be used to exclude out-of-phase samples and for which the cost per subject is not great. On the other hand, in trials or small observational studies with a high cost investment per subject, salivary measures could be more efficient by excluding fewer subjects and providing higher ICCs. More methodological research regarding the application of salivary methods is needed. In the context of breast cancer, the relevance of salivary steroid levels to the exposure of breast target tissue itself is a subject of great interest.

Acknowledgments

We thank the staff of the General Clinical Research Center at Northwestern Memorial Hospital for their help in data collection, Kelly Hodges for her excellent work conducting assays, and Laura Studee and Linda Schiffer for their assistance in preparation of the manuscript.

References

1. Pike, M. C., Spicer, D. V., Dahmoush, L., and Press, M. F. Estrogens, progestogens, normal breast cell proliferation and breast cancer risk. Epidemiol. Rev., *15:* 17–35, 1993.

2. Kuller, L. H. The etiology of breast cancer—from epidemiology to prevention. Public Health Rev., *23:* 157–213, 1995.

3. Muti, P., Trevisan, M., Micheli, A., Krogh, V., Bolelli, G., Sciajno, R., and Berrino, F. Reliability of serum hormones in premenopausal and postmenopausal women over a one-year period. Cancer Epidemiol. Biomark. Prev., *5:* 917–922, 1996.

4. Michaud, D. S., Manson, J. E., Spiegelman, D., Barbieri, R. L., Sepkovic, D. W., Bradlow, H. L., and Hankinson, S. E. Reproducibility of plasma and urinary sex hormone levels in premenopausal women over a one-year period. Cancer Epidemiol. Biomark. Prev., *8:* 1059–1064, 1999.

5. Lu, Y., Bentley, G. R., Gann, P. H., Hodges, K. R., and Chatterton, R. T. Salivary estradiol and progesterone levels in conception and nonconception cycles in women: evaluation of a new assay for salivary estradiol. Fertil. Steril., *71:* 863–868, 1999.

6. Lipson, S. F., and Ellison, P. T. Development of protocols for the application of salivary steroid analyses to field conditions. Am. J. Hum. Biol., *1:* 249–255, 1989.

7. Chiu, T. T., Tam, P. P., and Mao, K. R. Evaluation of a semiquantitative urinary LH assay for ovulation detection. Int. J. Fertil., *35:* 120–124, 1990.

8. Lu, Y. C., Chatterton, R. T. J., Vogelsong, K. M., and May, L. K. Direct radioimmunoassay of progesterone in saliva. J. Immunoassay, *18:* 149–163, 1997.

9. Fleiss, J. L. The Design and Analysis of Clinical Experiments, pp. 1–32. New York: Wiley Inc., 1986.

10. Shekelle, R. B., Shryock, A. M., Paul, O., Lepper, M., Stamler, J., Liu, S., and Raynor, W. J. J. Diet, serum cholesterol, and death from coronary heart disease. The Western Electric study. N. Engl. J. Med., *304:* 65–70, 1981.

11. Key, T. J. A., and Pike, M. C. The role of oestrogens and progestagens in the epidemiology and prevention of breast cancer. Eur. J. Cancer Clin. Oncol., *24*: 29–43, 1988.

12. Ellison, P. T., Lipson, S. F., O'Rourke, M. T., Bentley, G. R., Harrigan, A. M., Panter-Brick, C., and Vizthum, V. J. Population variation in ovarian function. Lancet, *342:* 433–434, 1993.

13. Bentley, G. R., Harrigan, A. M., and Ellison, P. T. Dietary composition and ovarian function among Lese horticulturalist women of the Ituri Forest, Democratic Republic of Congo. Eur. J. Clin. Nutr., *52:* 261–270, 1998.

14. Westhoff, C., Gentile, G., Lee, J., Zacur, H., and Helbig, D. Predictors of ovarian steroid secretion in reproductive-age women. Am. J. Epidemiol., *144:* 381–388, 1996.

15. Ellison, P. T., and Lager, C. Moderate recreational running is associated with lowered salivary progesterone profiles in women. Am. J. Obstet. Gynecol., *154:* 1000–1003, 1986.

16. Feigelson, H. S., Ross, R. K., Yu, M. C., Coetzee, G. A., Reichardt, J. K. V., and Henderson, B. E. Genetic susceptibility to cancer from exogenous and endogenous exposures. J. Cell. Biochem., *25S:* 15–22, 1996.

17. Sukalich, S., Lipson, S. F., and Ellison, P. T. Intra and interwomen variation in progesterone profiles. Am. J. Phys. Anthropol., *18* (Suppl.)*:* 191–191, 1994.

18. Lenton, E. A., Lawrence, G. F., Coleman, R. A., and Cooke, I. D. Individual variation in gonadotrophin and steroid concentrations and in the lengths of the follicular and luteal phases in women with regular menstrual cycles. Clin. Reprod. Fertil., *2:* 143–150, 1983.

19. Baird, D. D., McConnaughey, R., Weinberg, C. R., Musey, P. I., Collins, D. C., Kesner, J. S., Knecht, E. A., and Wilcox, A. J. Application of a method for estimating day of ovulation using urinary estrogen and progesterone metabolites. Epidemiology, *6:* 547–550, 1995.

20. Waller, K., Swan, S. H., Windham, G. C., Fenster, L., Elkin, E. P., and Lasley, B. L. Use of urine biomarkers to evaluate menstrual function in healthy premenopausal women. Am. J. Epidemiol., *147:* 1071–1080, 1998.

21. Read, G. F. Status report on measurement of salivary estrogens and androgens. Ann. NY Acad. Sci., *694:* 146–160, 1993.

22. Walker, R. F., Wilson, D. W., Truran, P. L., Read, G. F., Richard, G., Walker, S. M., and Riad-Fahmy, D. Characterization of profiles of salivary progesterone concentrations during the luteal phase of fertile and subfertile women. J. Endocrinol., *104:* 441–446, 1985.

23. Lipson, S. F., and Ellison, P. T. Comparison of salivary steroid profiles in naturally occurring conception and non-conception cycles. Hum. Reprod., *11:* 2090–2096, 1996.

Saliva as a Medium for Investigating Intra- and Interindividual Differences in Sex Hormone Levels in Premenopausal Women

Peter H. Gann, Susan Giovanazzi, Linda Van Horn, et al.

Cancer Epidemiol Biomarkers Prev 2001;10:59-64.