Short Communication

The Optimal Timing of Blood Collection during the Menstrual Cycle for the Assessment of Endogenous Sex Hormones: Can Interindividua Differences in Levels over the Whole Cycle Be Assessed on a Single Day?\(^1\)

Naseer Ahmad, Tessa M. Pollard, and Nigel Unwin\(^2\)

Departments of Diabetes [N. U.] and Epidemiology and Public Health [N. A., N. U.], University of Newcastle, Newcastle NE2 4HH, United Kingdom, and Department of Anthropology, University of Durham, Durham, United Kingdom [T. M. P.]

Abstract

The objective of this study was to identify the optimal timing of sampling during the menstrual cycle for assessment of interindividual variation in exposure to endogenous sex hormones, including estradiol, progesterone, and the free androgen index. Twenty-four healthy premenopausal women with regular periods were recruited, and alternate day venous blood samples were taken in the morning throughout one menstrual cycle. Spearman rank correlation coefficients were calculated for the estimates of average hormone levels (based on area under the curve) over one menstrual cycle against values on single days within that cycle. Days within the menstrual cycle were identified that provided the best assessment of interindividual differences. The most consistent correlation for estradiol was seen between days 9 and 11 (e.g., \(r = 0.53\) and \(P = 0.01\), day 10), the most consistent correlation for progesterone was seen between days 17 and 21 (e.g., \(r = 0.80\) and \(P < 0.001\), day 20), and the most consistent correlation for free androgen index was seen between days 12 and 15 (e.g., \(r = 0.90\) and \(P < 0.001\), day 15). Post hoc analysis of estradiol and progesterone levels on days counted back from the start of the next menstrual cycle identified marginally stronger associations. On repeat hormone measurements (not done for progesterone) on days 10 and 15, two to five menstrual cycles later, correlation coefficients with the original hormone levels remained reasonable (\(r \\leq 0.55\) for most). In conclusion, a reasonable characterization of interindividual differences in premenopausal estradiol, androgen, and progesterone levels may be achieved with single blood samples taken on specific days. This provides a useful approach for future epidemiological studies.

Introduction

Endogenous sex hormones have been related to the risk of several diseases in women, including breast and ovarian cancer (1, 2), osteoporosis (3), and cardiovascular disease (4, 5). Most evidence for the effects of these hormones comes from comparing risk in women before and after the menopause [e.g., Rannevik et al. (6)], from our increasing understanding of the effects of postmenopausal hormone replacement therapy [e.g., Anonymous (7)], or from comparing risk according to endogenous postmenopausal hormone levels [e.g., Hankinson et al. (8)]. However, there is good evidence that variation between populations in premenopausal endogenous sex hormone levels correlates with variation in reproductive cancers in women (9) and also that within modern industrial populations, overall exposure to estrogen over the reproductive years is an important risk factor for the same cancers (8) and affects the degree of protection against cardiovascular disease (10).

It is important, therefore, that we understand the variation in premenopausal sex hormone levels between and within populations and its causes. Unfortunately, such comparisons are plagued by methodological difficulties associated with changes in sex hormones levels during the menstrual cycle (2). Thus, some researchers have chosen the labor-intensive course of collection of multiple samples of either serum or saliva through at least one menstrual cycle (11–13). Gann et al. (14) suggest that small and intensive studies may find repeated salivary samples, which are expensive, most useful, but that timed one-off serum samples are potentially more useful for larger studies. However, some studies have failed to address the timing of samples (15), whereas others have sampled at one stage of the menstrual cycle, with [e.g., Bernstein et al. (16)] or without [Kamath et al. (17)] attempts to justify their sampling strategy.

Michaud et al. (18) recognized the importance of evaluating whether a single plasma hormone measurement can be used to estimate stable differences between individuals, given that most epidemiological studies can only collect one sample per participant. They compared the reproducibility of plasma estrogen and progesterone levels over 1 year, with estrogens assessed either during the follicular phase (on the 3rd, 4th, or 5th day of the menstrual cycle) or in the luteal phase (7–9 days before the anticipated start of the next cycle), and progesterone assessed only in the luteal phase. They found that reproducibility was good, except for estradiol in the luteal phase, although after the exclusion of women who were anovulatory (identified by low progesterone levels) and those whose samples were collected <4 days or >10 days before their next menses, reproducibility in luteal phase samples was also moderately good. However, to our knowledge, no study has investigated in greater depth than this the extent to which interindividual differences can be reliably characterized using measurements on any one day during the menstrual cycle.
We suggest that it is important to demonstrate that levels on any day chosen for assessment can provide an indication of overall exposure during the menstrual cycle. Our aim, therefore, was to identify the optimal timing of sampling during the menstrual cycle for assessment of overall interindividual variation in exposure to endogenous estradiol, progesterone, free androgen, androstenedione, and DHEAS.3

Materials and Methods

Subjects
Twenty-four volunteers ages 18–45 years with regular menstrual cycles (i.e., the women claimed they could predict the onset of their menses to within 7 days) took part. They were recruited through a poster campaign within Newcastle University and Newcastle Royal Victoria Infirmary and by word of mouth. No woman had taken the oral contraceptive pill for at least 3 months.

Data Collection

Questionnaire and Anthropometry. Subjects completed a short questionnaire, which included basic demographic details and history of oral contraceptive use. Height and weight were measured.

Blood Samples. Our aim was to take blood samples at least every other morning (8:30 a.m. to 12 p.m.) from the beginning of one menstrual cycle until the start of the next. The only limitation on this was that for practical reasons blood was not taken on Sundays, and because of this, all women had blood taken Saturday and Monday mornings and then on Wednesday and Friday mornings. Thus, sampling was slightly more frequent than every other day. Venous blood was collected in two plain 10-ml tubes and sent the same morning to the biochemistry department, where the samples were spun and the serum-frozen at −70°C. The assays

3 The abbreviations used are: DHEAS, dehydroepiandrosterone sulfate; CV, coefficient of variation; SHBG, sex hormone-binding globulin; FAI, free androgen index.
were undertaken in several batches as part of the normal work of the regional steroid laboratory.

Assays
The following hormones were measured (a commercial kit was used), and the coefficients of variation were provided by the laboratory: (a) estradiol (Diagnostic Products Corp., Los Angeles, CA), intra-assay CV of 4.3% at 176 pg/liter and interassay CV of 6.8% at 168 pg/liter; (b) testosterone (Diagnostic Products Corp.), intra-assay CV of 7% at 100 ng/dl and interassay CV of 6.7% at 401 ng/dl; (c) androstenedione (Johnson & Johnson Clinical Diagnostics Ltd., Amersham, United Kingdom), intra-assay CV of 2.7% at 2.23 ng/ml and interassay CV of 4.8% at 2.07 ng/ml; (d) DHEAS (Johnson & Johnson Clinical Diagnostics Ltd.), intra-assay CV of 4.2% at 2245 ng/ml and interassay CV of 10% at 2190 ng/ml; (e) SHBG (Orion Corp., Orion Diagnostica), intra-assay CV of 0.56% at 24.8 nmol/liter and interassay CV between 1.6% and 8.3% at concentrations between 24.6 and 232 nmol/liter; and (f) progesterone (Johnson & Johnson Clinical Diagnostics Ltd.) interassay CV of 9.4% at 14.6 nmol/liter (no intra-assay CV was available).

Statistical Analysis
Data were entered, checked, cleaned, and then analyzed using SPSS software (19). We chose to define interindividual differences in hormone production using an estimate of total production over the whole cycle. For each woman, an average hormone level was computed by calculating the area under the curve for the whole menstrual cycle and then dividing it by the length of the cycle in days (20). These mean values were then correlated (using Spearman rank) with the values on the individual days. Values for days on which hormones were not measured were calculated as the mean of the measured values from the previous day and the following day. The FAI was computed as testosterone divided by SHBG multiplied by 100.

Repeat Measurements
Two days of the menstrual cycle were chosen as providing the closest and most stable agreement with the ranking based on the averages for the cycle, one for estradiol and one for the FAI. These days were chosen by visually inspecting the agreement as shown in Fig. 1, A–F. As far as possible, we looked for periods where the correlation coefficients were consistently high over several days and chose the day in the middle of such a period. Women were asked to return to have a repeat blood sample on these 2 days in a subsequent cycle (days 10 and 15). The ranking based on these repeat measurements was then correlated with the ranking based on the average levels of hormone production from the original cycle.

Results
The mean age of the 24 female subjects was 28.5 years (age range, 18–45 years), the mean cycle length was 27 days (range, 24–34 days), and the mean body mass index was 24.4 kg/m² (range, 20.4–30.3 kg/m²). Table 1 shows the means of the average (based on area under the curve for the whole menstrual cycle) hormone and SHBG levels for the 24 subjects. Fig. 1, A–F, shows the correlation coefficients for the five hormones, SHBG, and the FAI. Table 2 shows the actual values of the Spearman rank correlation coefficients for the best and worst days of agreement for each of the hormones and SHBG. Although the highest correlation coefficient for estradiol was on day 6, the most consistent correlations for estradiol, as shown in the figure, were between days 9 and 11 ($r = 0.53$ and $P = 0.01$, day 10). For the FAI, DHEAS, androstenedione, and SHBG, correlation was moderate to good over most of the cycle. Thus, between days 3 and 22, all $r$ values for FAI and androstenedione were greater than 0.6, and all $r$ values for DHEAS and SHBG were greater than 0.7. On repeat measurements (performed on 17 of the original 24 subjects) on days 10 and 15, two to five menstrual cycles later, correlations with the original ranking, apart for testosterone, remained reasonable to good ($r = 0.55$).

The relationships between average estradiol and progesterone levels and daily values counting backward from the start of the next menstrual cycle were assessed in post hoc analysis. These are shown in Fig. 2. There was reasonable to good agreement for both estradiol and progesterone at 5–7 days (inclusive) prior to the onset of menses.

Discussion
Measuring sex hormones on alternate days throughout a whole menstrual cycle is demanding on the subject and expensive, and it is for such reasons that our study was limited to 24 subjects. Implicit within the method chosen here is that it is the total hormone production over the course of the menstrual cycle that is of interest in characterizing interindividual differences, not, for example, peaks or troughs on individual days or parts of days. Within these confines, these results provide new information about optimal timing of sampling of sex hormones in premenopausal women. They suggest that timing for DHEAS assessment is not important and that samples obtained at any stage of the menstrual cycle can provide a good indication of overall levels, as well as being reliable across different menstrual cycles. The correlations for androstenedione, total testosterone, SHBG, and FAI all tended to be strongest toward the middle of the cycle, although correlations for all but testosterone were significant on all days. Reliability for measurements of all these hormones on day 15 of the cycle was good. Muti et al. (21) found that reliability over a 1-year interval for plasma samples taken on the 20th–24th cycle day was good for DHEAS and slightly lower for total testosterone. Our results suggest that such results are unlikely to be limited to these days of the cycle.

Estradiol and progesterone are the two hormones likely to be of the most widespread interest in epidemiological studies. Our results suggest that single assessments can be used to provide an estimate of overall interindividual differences, but that sample timing is critical. Estradiol and progesterone fluctuate markedly through the menstrual cycle, and it is not surprising, therefore, that correlations with overall levels varied considerably according to day of the cycle.

It seems that the optimal timing for one-off estradiol assessment, as identified by forward counting from the start of the previous menses, is probably between days 6 and 11 of

<table>
<thead>
<tr>
<th>Hormone</th>
<th>Mean (SD)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estradiol (pmol)</td>
<td>379.5 (76.88)</td>
<td>253–517</td>
</tr>
<tr>
<td>Androstenedione (nmol)</td>
<td>6.49 (1.98)</td>
<td>3.15–11.42</td>
</tr>
<tr>
<td>DHEAS (umol)</td>
<td>4.72 (1.92)</td>
<td>1.0–8.1</td>
</tr>
<tr>
<td>Progesterone (nmol)</td>
<td>12.0 (5.8)</td>
<td>3.0–22.0</td>
</tr>
<tr>
<td>Testosterone (nmol)</td>
<td>0.95 (0.4)</td>
<td>&lt;1–1.9</td>
</tr>
<tr>
<td>SHBG (nmol)</td>
<td>50.1 (21.8)</td>
<td>21.0–123.0</td>
</tr>
</tbody>
</table>

*Total production throughout the cycle, based on area under the curve, divided by the length of the menstrual cycle in days.
Short Communication: Interindividual Differences in Endogenous Sex Hormones

high; our results provide of the follicular phase, presumably following the commonsense by differences in stage of the cycle reached. This method allows cycles from different individuals are better aligned by this method, the luteal phase than in that of the follicular phase (18), so that estradiol. This probably reflects the lower variability in length of average hormone levels over the cycle for both progesterone and potential for higher correlations between daily hormone values and start of the next menses is an alternative strategy, giving the previous menses is likely to be optimal.

suggest that sampling between days 17 and 21 after the start of follicular phase, unsurprisingly, did not. In this case, our results correlated well with overall levels, but that levels during the luteal phase, as identified by forward counting, chosen by these authors. Our findings showed that progesterone differ in terms of cycle length, so that the window of opportunity is shorter (our data suggest that it may be as short as 3 days). In their study of 403 women in the United States, Waller et al. (23) found a within-person SD in menstrual cycle length of 2.4 days, suggesting that it will be difficult to use this method for estradiol assessment unless, like Michaud et al. (18), researchers are willing to exclude samples from women who did estimate the day of their next menses accurately, as well as, perhaps, samples from women identified as anovulatory (according to progesterone levels). An alternative strategy would be to take samples on more than one day to increase the possibility of obtaining a sample timed correctly in relation to the actual start of the menses.

These results should encourage the careful use of one-off assessments of sex hormone levels in premenopausal women as a useful method for the measurement of interindividual differences likely to have important implications for health. Recent work suggests that assays of these hormones from dried blood spots are sensitive and reliable and thus that sampling itself may become standardized without a significant increase in exposure to endogenous hormones has great potential for future epidemiological studies.

**Acknowledgments**

We thank Dr. Malcolm Watson for advice on and organization of the analyses of the samples and Dr. Salma Ayis for statistical advice.

**References**


23. Waller.


Correction


On page 147, the postal code for Dr. Tessa M. Pollard is DH1 3NH. We also wish to correct an incomplete reference on page 151. The reference should read:

The Optimal Timing of Blood Collection during the Menstrual Cycle for the Assessment of Endogenous Sex Hormones: Can Interindividual Differences in Levels over the Whole Cycle Be Assessed on a Single Day?

Naseer Ahmad, Tessa M. Pollard and Nigel Unwin