Urinary enzyme-immunoassays for population research on reproduction: Estrone conjugates and pregnanediol-3-glucuronide

by

Kathleen O’Connor
University of Washington
Title:

Urinary enzyme-immunoassays for population research on reproduction: Estrone conjugates and pregnanediol-3-glucuronide.

Running Title:

Urinary EIA’s for population research: E1C and PDG

Authors and Institutions:

Kathleen A. O’Connor
Eleanor Brindle
Darryl J. Holman
Nancy A. Klein
Michael R. Soules
Kenneth L. Campbell
Fortüne Kohen
Coralie J. Munro
William L. Lasley
James W. Wood

1 Department of Anthropology and Center for Studies in Demography and Ecology, University of Washington, Seattle WA 98195
2 Department of Obstetrics and Gynecology, University of Washington, Seattle WA 98195
3 Department of Biology, University of Massachusetts, Boston MA 02125
4 Department of Biological Regulation, Weizmann Institute of Science, Rehovet 76100 Israel
5 Department of Population Health and Reproduction, University of California, Davis, CA 95616
6 Department of Obstetrics and Gynecology, University of California, Davis, CA 95616
7 Department of Anthropology and Population Research Institute, Pennsylvania State University, University Park, PA 16802

Total Number of Pages: 22
Number of Figures: 9
Number of Tables: 5

Keywords: E1G, E1C, EIA, PDG, urinary reproductive steroids, Quidel 330, validations, Bangladesh, 3F11 clone, 155B3 clone, assay validation, urine specimen stability, specific gravity

Corresponding Author:

Kathleen A. O’Connor
Department of Anthropology
Box 353100
University of Washington
Seattle, Washington 98195
phone: (206) 543-9605
fax: (206) 543-3285
email: oconnork@u.washington.edu

Date: 10/19/2002
ABSTRACT

Background: Detailed monitoring of steroid hormone patterns across the ovulatory cycle, pregnancy, the postpartum period and the transitions to menarche and menopause requires frequent hormone measurements. Our aim in this study is to establish and validate enzyme immunoassays for urinary metabolites of estradiol and progesterone for population level and cross-cultural research in reproductive biology.

Methods: Three microtiter plate based enzyme-immunoassays (EIAs) are developed and validated for measuring the main urinary metabolites of estradiol and progesterone. Specificity, sensitivity, parallelism, accuracy, and precision of the assays are determined, and the urinary hormone data are compared with serum hormone measures. Performance of the assays using samples from populations in different ecological settings is examined with daily or twice weekly urine specimens collected in the field from US and Bangladeshi women.

Results: The urinary EIA’s are specific, accurate, sensitive, precise, and provide hormone profiles parallel to the serum parent hormones. The assays and specimens are stable and reliable for prospective studies and population level research. The pregnanediol-3-glucuronide (PDG) EIA is useful for a wide range of populations, including those where PDG levels may be quite low. Different urinary estrone metabolite assays are necessary for Bangladesh and US samples, a consequence of population differences in metabolite levels.

Conclusions: These urinary EIA’s are well suited for cost-effective and efficient processing of the large numbers of specimens used in population level research on ovarian function. Population variation in hormones and their metabolites requires consideration when developing and applying urinary assay methods as indicators of ovarian function.

INTRODUCTION

Detailed population-level monitoring of steroid hormone patterns across the menstrual cycle, pregnancy, the postpartum period and the transitions to menarche and menopause requires frequent hormone measurements. Frequent monitoring also is necessary for research examining inter-population and inter-individual variation in reproductive function, and relationships with demographic, health, environmental, sociocultural and biological covariates. The objectives of this paper are to develop and optimize enzyme immunoassays for the urinary metabolites of estradiol and progesterone for population level research on reproductive function, and to examine the benefits and limitations of the assays in two different population settings. Enzyme immunoassays for the urinary metabolites of estradiol and progesterone that are currently widely used in epidemiological and clinical research ((1)) are hampered first by the use of polyclonal antibodies, which limits the lifespan of the use of the assays, and second by lack of evaluation of the assays in different population settings.

Urine samples are easy to collect and ideally suited for population-based field research. Urine has several advantages over saliva, serum, and blood spots for extracting repeated steroid hormone measures over long periods of time: collection of urine is noninvasive; urine poses minimal infectious disease risk to subjects or researchers; samples can be self-collected and stored by subjects without research personnel present; daily samples over long periods of time are easily obtained; subject compliance is high; and a sufficient volume of sample can be collected for multiple assays and future research (2)(3)(4)(5)(6). Additionally, urine has the advantage of providing integrated hormone measures without the confounding effects of pulsatile secretion (1)(6); this eliminates a potential source of variability over serum or blood spot samples when monitoring daily and monthly patterns of hormone secretion. Finally, urinary hormone assays are better able to quantify the lower end of the physiological scale in humans than many serum hormone assays (7).
Potential disadvantages of using urine for measuring reproductive steroids include: the need for relatively constant storage temperatures from time of collection to time of assay (8); inability to examine pulsatility or other microvariation in secretion patterns (6), and the necessity of correcting for hydration status of the subject (4).

The principal steroids regulating reproduction in the human female are estradiol and progesterone. These steroids are primarily produced in the ovarian follicles and corpora lutea, with some peripheral but minor production in the adrenal glands and adipose tissue (9). Circulating serum estradiol and progesterone are metabolized in the liver, where they are transformed and conjugated to glucuronic or sulfuric acid prior to excretion in urine (9). Stable metabolites of estradiol in urine are free estrone (E1), and the estrone conjugates (E1C): estrone sulfate (E1S) and estrone-3-glucuronide (E1G); the principal stable urinary metabolite of progesterone is pregnanediol-3-gulcuronide (PDG) (9)(10)(11). The urinary levels of these metabolites closely parallel serum levels of estradiol and progesterone, after correction for hydration status (1)(12).

Urinary E1C and PDG have long been used as indicators of ovarian status and function, in studies of reproductive toxicology (3)(13)(14), ovulation and conception(6)(10)(15)(16)(17), reproductive aging (7)(18)(19), luteal phase function (20)(21), energetics (22)(23); breast cancer risk (24), and infertility (25). A range of capture and labeling methods have been used to detect and quantify urinary E1C and PDG: gas chromatography (11) (26); radio-immunoassay (3)(11)(12)(27), immunofluorometric assay (28)(29), chemiluminescent assay (30)(31)(32), and enzyme immunoassay (1)(33)(34). Of these various approaches, we have found the enzyme immunoassay (EIA) to be well-suited and cost-effective for population level research. The equipment and reagents are affordable, no sample preparation (e.g. extraction) is needed, no hazardous or radioactive materials are used, and the assays are reliable and sensitive across the physiological range of hormone levels.

The EIA format of the PDG assay (1) we present in this paper has been widely used in research examining luteal phase function (14)(14)(35)(36)(37)(38), infertility (39), and characteristics of the peri-menopause and menopause (40). In the majority of these applications, the PDG assay employs a polyclonal antibody, of which there is a finite supply, limiting the lifespan of use of the assay. Our first objective in this paper was to evaluate the performance of the PDG EIA using a monoclonal antibody, and optimize this assay for population level and cross cultural research.

A second objective was to evaluate and optimize the performance of two different estrone conjugate assays for population based research. The extensively used R522 E1C EIA (41) (42) (37) (38)refs uses a polyclonal antibody, thus limiting long term use of the assay. The estrone conjugate assays developed in the present report have the same general format as the PDG EIA. The assay protocol we report here has not been widely used to date [e.g. (40)], although the monoclonal antibodies used in these assays have been well characterized (43) and have been used in immunofluorometric assays (28)(29). Unlike progesterone, there are several conjugated and unconjugated forms of estrogen in the urine (9), and estrogen EIAs can vary in their specificity to different estrogen metabolites.

A final objective was to illustrate the benefits and limitations of the use of these assays for specimens collected in different population settings. A growing body of research has documented population differences in reproductive hormone levels (44)(45)(46)(47)(48)(49). Although the cause of this variation is unknown, it has been associated with differences in body composition, diet, metabolism, disease and route of excretion (46)(50). One consequence of this population variation is that assays optimized in Western or industrialized populations may have performance or clinical limitations when applied to populations living under different cultural and ecological conditions.
MATERIALS AND METHODS

Subjects and Specimens

Urine specimens were collected from women of various adult reproductive statuses to assess the utility and limitations of the EIAs using samples collected in clinical, home, and field settings in two different populations. In both Bangladesh and the US, urine specimens were ‘spot samples,’ collected at whatever time of day was convenient for the participants and/or researchers. All subjects provided written informed consent, and all procedures were approved by the institutional review boards of the Pennsylvania State University, the International Centre for Diarrhoeal Disease Research, Bangladesh, and the University of Washington.

A random sample of resident, married non-contracepting women in the nonintervention demographic surveillance region of the rural Matlab district of Bangladesh were invited to participate in a research study on early pregnancy loss (51). Reproductive statuses of the subjects included adult women who were pregnant, breastfeeding, cycling, peri-menopausal, or postmenopausal. A total of 19,033 urine specimens were collected in 1993 on a twice per week schedule from 841 subjects aged 18-50 years (51). No monetary compensation was provided for participation in the study. Immediately after collection, the Bangladeshi urine specimens were placed in coolers with ice packs and then transported within two days to a research hospital (51). Samples were preserved with 17 mg/mL boric acid and kept at 4ºC for up to one week and then frozen at –20ºC. The Bangladeshi samples were transported via frozen air freight to the US. During transport and storage prior to assay in 1996, these specimens underwent up to five freeze-thaw cycles, and variable numbers of hours at refrigerated or ambient temperatures and variable numbers of days at refrigerated temperatures.

Urine and serum specimens were collected daily from thirty US women in 1997-1998. Thirteen women aged 20-25 years and seventeen women aged 40-45 years were recruited for a study on reproductive aging (ref). All participants had regular 25-35 day menstrual cycles, were in good health, had a BMI between 18-24 kg/m², and were using no medications or hormones. Daily blood samples were obtained by venipuncture, beginning with the first day of menstrual bleeding and continuing until the first day of menstrual bleeding of the subsequent cycle. Daily transvaginal ultrasound was performed on all subjects from the mid- to late follicular phase until evidence of ovulation was observed. Monetary compensation was provided for participation. Daily urine specimens were frozen immediately after collection, and remained frozen at –20ºC until thawing two years later for aliquoting, preservation and assay. Specimens were preserved with 17 mg/mL boric acid after the first thaw (52), and underwent one to three freeze-thaw cycles prior to assay.

Given the various field conditions under which specimens were collected in the US and Bangladesh, we examined the effects on hormone recovery of various exposure to ambient temperature and multiple freeze-thaws. A single 50 mL specimen was collected from each of two healthy normally cycling US women. No preservative was added to these specimens. The specimens were subdivided into the following treatments: 0, 2, 4, and 7 days at room temperature. Each of these treatments were refrigerated and then subjected to 0, 1, 3, 5, 7 and 10 freeze thaw cycles. Six replicates of each treatment were assayed in the 3F11 and PDG EIAs.

Assay Reagents and Protocols

A competitive microtiter plate solid phase enzyme-immunoassay (EIA) for pregnanediol 3-glucuronide was developed using the Quidel anti-pregnanediol 3-glucuronide monoclonal antibody, clone 330. The general protocol is similar to Munro et al. (1). The purified antibody is coated onto Nunc Immunosorp 96-well microtiter plates (50 µL/well; working concentration is typically close to a 1:4,000 dilution, but varies by batch) and incubated at 4ºC for 18 hours or up to 5 days. The plates are then washed (0.15 M NaCl; 0.05% Tween 20) and blocked with 0.1% BSA (50 µL/well) for one-half to three hours at room temperature. This is followed by the addition of standards, samples and controls.
(20µL/well). The competitor, pregnanediol 3-glucuronide conjugated to horseradish-peroxidase (HRP) (50 µL/well; working dilution varies around 1:40,000; see Munro et al. [1]), is added immediately after the test antigens. After an overnight incubation at 4ºC, the plates are washed and then developed in citrate buffer combined with ABTS (Sigma) and 0.0064% hydrogen peroxide (100µL/well). Optical density measures of the amount of competitor present per sample are made with a Dynatech MR7000 Plate Reader (test wavelength 405 nm, reference wavelength 570 nm) and quantified using log-linear sigmoid regression in Biolinx 1.0 Software (Dynex Laboratories, Inc.). Commercial standards (5β-pregnane-3α, 20α-diol glucuronide, Sigma Catalog No. P3635) and in-house urine controls are used in all assays. Standards, samples and controls are run in duplicate on every plate.

Two different monoclonal anti-estrone antibodies were evaluated in the EIA format, the 3F11 rat-derived clone, and the 155B3 mouse-derived clone, both from F. Kohen (28)(43). The 3F11 assay is highly specific to E1G, while the 155B3 also cross-reacts with E1S and E1 (Table 1). In these assays, microtiter plates are pre-coated at 10µg/mL with anti-immunoglobulin G (IgG) (50µL/well; rabbit-anti-rat IgG, 0.5µg/mL for the 3F11 assay; rabbit-anti-mouse IgG, 0.5µg/well for the 155B3 assay; both from Jackson Immuno-research). After an overnight 4ºC incubation, the unpurified ascites fluid containing the monoclonal antibodies is applied to plates (50 µL/well; working dilution approximately 1:100,000 for 155B3 and 1:15,000 for 3F11). From here on, the protocol is identical to PDG EIA except the competitor, controls and standards differ, 40µL/well (instead of the 20µL/well in the PDG assay) of standards, samples and controls are added, and there is a one-half hour delay between adding standards, samples, controls and the addition of competitor. Estrone 3-glucuronide conjugated to horseradish peroxidase is used as the competitor [see Munro et al (1)]; working dilution is around 1:20,000 dilution for the 3F11 assay and 1:10,000 for 155B3 assay), and estrone-ß-D-glucuronide (Sigma Catalog No. E1752) is used for the standard curve.

### TABLE 1. Specificity of Urinary Estrone Conjugate Assays

<table>
<thead>
<tr>
<th>Steroid</th>
<th>Monoclonal 3F11 (% cross-reaction)</th>
<th>Monoclonal 155B3 (% cross-reaction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>estrone 3-glucuronide</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>estrone</td>
<td>&lt;0.1</td>
<td>100</td>
</tr>
<tr>
<td>estrone 3-sulfate</td>
<td>&lt;0.1</td>
<td>100</td>
</tr>
<tr>
<td>estradiol 3-glucuronide</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>estriol 3-glucuronide</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>estradiol</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>estriol</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

1Data are from Barnard et al 1989, page 557.

Sample and in-house control dilutions are made in de-ionized, distilled water (ddH2O). Estimated urinary hormone concentrations are corrected for hydration status using specific gravity rather than creatinine, as the latter exhibits high intra and inter-individual variability (53)(54)(55)(56)(57)(58)(59). Specific gravity of Bangladesh urine specimens was taken after collection and before freezing of the samples. Specific gravity of the US specimens was taken after the first freeze-thaw cycle. Specific gravity was measured by placing a drop of urine on a hand-held urine specific gravity refractometer (Atago, Uricon-PN). The correction formula is applied to each hormone result:

\[
\text{corrected hormone concentration}_{\text{sample}} = \frac{\text{raw hormone concentration}_{\text{sample}} \times (\text{specific gravity}_{\text{target}} - 1.0)}{\text{specific gravity}_{\text{sample}} - 1.0}
\]
where specific gravity_{target} is a population mean specific gravity. We use a population mean of 1.015 for the Bangladesh samples (51) and 1.020 for the US samples (58).

Validations

Assay performance was evaluated for each assay by demonstrating parallelism between a standard curve (estrone-β-D-glucuronide or pregnanediol-3-glucuronide) and serially diluted urine samples. A total of five samples from Bangladesh and U.S. women were used in a series of dilutions to test for parallelism. Results are expressed as percent of antibody sites bound by competitor (E1G-HRP or PDG-HRP).

Accuracy for each assay was determined as percent of added mass (known standard dose) recovered from ddH2O and from a urine matrix. Pre-pubertal male urine (PPMU), charcoal-stripped to remove endogenous steroid hormones, was used for the urine matrix. PPMU and ddH2O were spiked with serial dilutions of either E1C or PDG standard. The standards were prepared in ddH2O and then added as 10% of the sample volume in either PPMU or ddH2O.

Specificity of the PDG assay was measured as the percent cross-reaction with hormones of similar molecular structure. The cross-reactivity of several commercially available steroids and steroid metabolites structurally related to PDG was assessed by EIA. The 50% inhibition point of respective dose-response curves was expressed as (picograms of PDG/picograms of steroid or steroid metabolite) × 100%. Specificities for the 3F11 and 155B3 monoclonal antibodies have previously been determined in time-resolved fluoroimmunoassay (28).

Sensitivity for each assay was estimated as the minimum detectable level of hormone, three standard errors below the response (optical density) at the zero standard dose.

Precision was estimated by examining intra-and inter-assay variation. In-house high and low hormone concentration urine control pools were run in duplicate and monitored over 20 randomly selected plates to estimate the inter-assay coefficient of variation (CV). Twenty replicates each of high and low controls in a single assay were used for the intra-assay CV. Long-term reproducibility of results over the course of one year was estimated using these same controls across three hundred microtiter plate assays.

The performance of the urinary 3F11, 155B3 and PDG EIAs was further evaluated by comparing results from paired urine and serum specimens. Daily urine and serum specimens were collected from 30 normally cycling US women for one complete menstrual cycle (n = 808 samples). Serum estradiol (E2) and progesterone (P4) were measured by RIA. The RIA for E2 (ICN Biomedicals, Costa Mesa, CA) cross-reacts 20% with estrone, 1.5% with estriol and < 1% with all other steroids. The inter-assay and intra-assay coefficients of variation are 16% and 7%, respectively. The RIA for P4 (Diagnostic Systems Laboratories, Webster, TX) cross-reacts < 5% with all other steroids. The inter-assay and intra-assay coefficients of variation are 13% and 11%, respectively. Urinary E1C and PDG were measured in the EIAs as outlined above. Cycles were aligned by day of the midcycle serum luteinizing hormone (LH) peak (day 0), and the day of ovulation as determined from ultrasound using specific criteria including follicle collapse. Serum LH was measured by a solid phase two-site fluoroimmunometric assay (DELFIA, Pharmacia, Gaithersburg, MD), with intra- and interassay coefficients of variation of 2.8% and 4.7%.

Assay Applications

Urine specimens from cycling, menopausal, and pregnant women were run in each steroid EIA to examine the use of the assays for examining the physiological range of hormones encountered in these reproductive states. To illustrate the benefits and limitations of the assays for samples collected from different populations, US and Bangladesh urine specimens were compared in the PDG and two estrone conjugate assays.
Statistical Analyses

Mean E1C and PDG concentrations were estimated from duplicate wells using log linear regression in Biolinx software (Dynex Technologies, Inc). Optical density results with standard deviations ≥ 0.085 or outside the limits of detection of the assay were re-run. A CV for each sample was estimated from the mean optical density of duplicate wells.

The paired urine and serum data were examined for each hormone by Pearson correlation coefficient, for the total sample and for the averaged cycle days of the paired urine/blood data (n=34 cycle days or paired urine/serum means). Correlations were calculated using uncorrected and specific gravity corrected urinary hormone values.

Parallelism was evaluated for each assay by comparing a regression of percent hormone bound in a standard curve to four serially diluted urine samples. For each of the three assays, log-linear regression was fit for each sample as well as a dilution series based on standards. Paired t-tests were used to compare the slope of the standard curve with those of the serially diluted samples.

Serum/urinary hormone profiles, accuracy profiles, and sample treatment effects were examined graphically using the mean ± 2SEM.

RESULTS
Assay Characteristics

Specificity of the urinary PDG EIA is presented in Table 2. The high cross-reactivity of the 20α hydroxy-4-pregnen-3-one in the PDG assay is not of concern as this metabolite occurs is uncommon in urine specimens. Specificities for each of the estrone conjugate monoclonal antibodies are shown in Table 1 [from (28)]. The 155B3 EIA cross-reacted with a broader range of metabolites than did the 3F11 EIA.

<table>
<thead>
<tr>
<th>Steroid</th>
<th>Monoclonal Quidel 330 (%) cross-reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>pregnanediol 3-glucuronide</td>
<td>100</td>
</tr>
<tr>
<td>20α hydroxy-4-pregnen-3-one</td>
<td>187</td>
</tr>
<tr>
<td>2β hydroxy-4-pregnen-3-one</td>
<td>4.3</td>
</tr>
<tr>
<td>17α hydroxy-4-pregnen-3-one</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>5α pregnan-3,20-dione</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>pregnanediol</td>
<td>13.4</td>
</tr>
<tr>
<td>prenenolone</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>estradiol 17β</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>progesterone</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>cortisol</td>
<td>&lt;0.1</td>
</tr>
</tbody>
</table>

Analytical recovery for each of the assays is shown in Figure 1. Average recovery was 115% in PPMU and 113% in ddH20 for the 3F11 assay, 80% in PPMU and 68% in ddH20 for the PDG assay, and 157% in PPMU and 105% in ddH20 for the 155B3 assay.

Dose-response curves of urine samples exhibited parallelism with standard curves for each of the EIAs (Figure 2). A total of five samples from Bangladesh and U.S. women were used in a series of dilutions to test for parallelism. For each of the three assays, log-linear regression was fit for each sample as well as a dilution series based on standards. None of the sample slopes differed significantly from the corresponding standard dilution slope (p < 0.05).
Figure 1. Accuracy of PDG EIA (top), 3F11 E1G EIA (middle), and 155B3 E1C EIA (bottom). Mean (±2 SEM) recovery determined from 6 replicates per dose.
Figure 2. Parallelism between a curve made from serial dilutions of standard (triangles), and four subject samples (all other symbols), for each assay; PDG EIA (top), 155B3 E1C EIA (middle) and 3F11 E1G EIA (bottom). Dilutions for the subject samples are: neat, 2, 5, 10, 20, 30, 40, 50, 60, 70, 85, and 100. Dilutions in which the resulting concentration was outside of the valid range of the assay were removed. Solid lines are estimated from a linear regression on log dilution. Within each assay, none of the sample slopes differed significantly from the standard slope.
The minimum detectable dose for the PDG EIA was 23.3 ng/mL (±0.34SE), based on 626 different plates. For the 3F11 assay it was 1177.8 pg/mL (±9.67 SE; n=1,954 plates), and for the 155B3 assay it was 187.2 pg/mL (±9.3 SE; n=624 plates).

Intra-assay and inter-assay CVs are shown in Table 3. Long term reproducibility for each of the assays is shown in Figure 3 (1,296 PDG plates, 1,292 3F11 plates, 390 155B3 plates).

<table>
<thead>
<tr>
<th>TABLE 3. Precision of the PDG and Estrone-Conjugate EIA’s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td><strong>Intraassay % CV</strong></td>
</tr>
<tr>
<td>High control</td>
</tr>
<tr>
<td>Low control</td>
</tr>
<tr>
<td><strong>Interassay % CV</strong></td>
</tr>
<tr>
<td>High control</td>
</tr>
<tr>
<td>Low control</td>
</tr>
</tbody>
</table>

\(^1 N = 20\) replicates for each control
\(^2 N = 20\) separate assays for each control

Scatterplots of the paired urinary and serum estrogen and progesterone measures are shown in Figure 4. Pearson correlations between paired serum and urine measures (corrected for specific gravity) were 0.54 for E2-3F11 \((N=792)\), 0.39 for E2-155B3 \((N=791)\), and 0.57 for P4-PDG \((N=406)\). The correlations were lower when uncorrected urinary measures were used: E2-3F11=0.42 \((N=792)\), E2-155B3=0.32 \((N=791)\), P4/PDG = 0.43 \((N=406)\). All of these correlations were significantly greater than zero \((p<0.01)\). Paired urinary and serum hormone profiles across the menstrual cycle showed nearly identical patterns (Figure 5). Urinary and serum cycles were aligned by day of serum luteinizing hormone (LH) peak (day 0). Although there are serum E2, urinary E1C, and urinary PDG data across the entirety of each of the 30 cycles, only the luteal phase of each was assayed for serum P4. Comparing the averaged serum and urinary data across the 30 cycles gave Pearson correlations of 0.86 for E2-3F11 \((N=34\) cycle days), 0.74 for E2-155B3 \((N=34\) cycle days), and 0.98 for P4-PDG \((N=17\) cycle days). These correlations were significant \((p<0.01)\). When considering the entire cycle, a lag between serum and urinary measures was not evident for PDG, but there was, on average, a one-day lag for E1C (Table 4).

The hormone metabolites E1G and PDG were stable in urine samples at room temperature for up to 2 days, and were robust for up to 5 freeze-thaws (Figure 6). Some degradation of these metabolites was clear after 2 days at room temperature, and between 5 and 10 freeze thaws.

<table>
<thead>
<tr>
<th>TABLE 4. Pearson Correlations Between Urine and Serum with Time Lags</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag</td>
</tr>
<tr>
<td>Lag urine one day before serum</td>
</tr>
<tr>
<td>No Lag</td>
</tr>
<tr>
<td>Lag urine one day after serum</td>
</tr>
<tr>
<td>Lag urine two days after serum</td>
</tr>
</tbody>
</table>

\(^1 N = 34\) for 3F11-E2 and 155B3-E2; \(N = 17\) for PDG-P4
Figure 3. Long term reproducibility for the PDG EIA (top), 155B3 E1C EIA (middle), and 3F11 E1G EIA (bottom). The hormone levels recovered from both high and low hormone concentration in-house controls are plotted over time. Scale is representative of the standard curve used in each assay.
Figure 4. Scatterplots of paired urinary and serum hormone measures. Urinary values are corrected by specific gravity; $N=792$ for E2-3F11, $N=791$ for E2-155B3 and $N=406$ for P4-PDG.
Figure 5. Mean (± 2 SEM) steroid profiles for 30 menstrual cycles from 30 women, based on daily paired urine and serum samples. Top panel: serum estradiol (E2) and urinary 3F11 E1G; Middle panel: serum E2 and urinary 155B3 E1C; Bottom panel: serum progesterone (P4) and urinary pregnanediol-3-glucuronide (PDG). Cycles are aligned by day zero, the day of the serum LH peak. Black triangles indicate the average day of follicle collapse, as determined by ultrasound. Because of variation in length of cycle, the number of observations varies by cycle day, with a minimum of 4 observations on cycle days -17 and 16 and a maximum of 28 to 30 observations for cycle days -10 through 11. Urinary values are corrected by specific gravity.
Figure 6. Effects of freeze-thaw cycles and room temperature (RT) storage effects on 3F11 E1G and PDG for two US samples (mean ±2 SEM). No preservatives were added to these samples.
Assay Applications

The results of the three assays when applied to regularly cycling US women are shown in Figure 5. Application of the 3F11E1G and PDG assays to samples from early pregnancy and postmenopause (US subjects) is shown in Figure 7. The high levels of E1G and PDG across later pregnancy can be quantified with these assays by diluting samples with ddH2O. Similarly, assaying samples neat allows quantification of the lower end of the physiological range of these hormones, as seen, for example, in postmenopausal women.

In samples from Bangladeshi women the 3F11 EIA had an important limitation: hormone values were frequently below the limit of detection. After optimizing the 3F11 assay for a lower limit of detection and using a twelve point standard curve ranging from 10,000 to zero pg/mL, the minimum detectable dose was 237.6 pg/mL (N=14 separate assays). Precision, accuracy and parallelism were essentially the same as reported above (data not shown). Even using these modifications, many of the Bangladesh samples were still below the limit of detection (Figure 8). The 155B3 EIA assay reacted with a broader range of estrone metabolites than did the 3F11 assay (Table 2) making it more useful for the Bangladesh samples. All of the Bangladesh samples were above the limit of detection for the 155B3 EIA (Figure 8).

The PDG assay was useful for both Bangladesh and US samples: all samples in both populations were above the limit of detection of the assay. However, the Bangladesh samples generally exhibited very low PDG levels when compared to US samples (Figure 9). Although the Bangladesh cycle shown in Figure 9 has surprisingly low PDG levels, it was a conception cycle that ended in a full term pregnancy.

DISCUSSION

This paper presents validation and application data on three urinary reproductive steroid EIA’s designed for population-based research. The microtiter plate format and in-house reagent assembly make these assays cost effective and efficient for processing large numbers of specimens, when compared to kit-based commercial assay systems.

The 3F11, 155B3 and PDG EIAs showed excellent accuracy, precision and sensitivity. Hormone profiles generated from these urinary EIAs were parallel to the E2 and P4 profiles generated from serum RIAs. The higher correlation of specific gravity corrected samples with the serum data indicates that specific gravity correction of the urinary hormone values is necessary. Urinary peak E1C and E1G lags behind the serum E2 peak by one day. This has implications for algorithms estimating the day of ovulation [e.g., (60)(61)].

The urinary assays validated in the present paper compare favorably to widely used serum E2 and P4 assays. We examined Pearson correlations of the mean values from paired urine/serum samples in other studies that analyzed daily samples across one complete menstrual cycle (Table 5). The correlations between urinary and serum values are high across most of the studies, even though the urinary and serum assays are measuring different forms of the hormones.

Despite the high correlation of serum and urinary measures, the urinary PDG EIA and 155B3 EIA accuracy results are suggestive of some non-specific interaction in pre-pubertal male urine that acts to create a systematic bias toward apparent higher concentrations. We assume that a similar non-specific interaction exists in adult female urine samples, although we did not specifically test for this. In any event, the bias is minor and appears to be consistent across dose levels. For population based research this will not compromise data analysis or interpretation, but it may be of concern if the assays are applied in clinical settings.
Figure 7. Examples of 3F11 E1G and PDG assay applications in early pregnancy (top panel; cycle day taken from estimated date of ovulation) and post-menopause (bottom panel; 12 years post menopause). Both examples are from US women. Hormone values are corrected by specific gravity.

Figure 8. Comparison of two estrone conjugate assays across one ovarian cycle from a Bangladeshi woman. Samples were collected twice per week across the cycle. Cycle Day 1 is the first day of menses. Hormone values are corrected by specific gravity.
Figure 9. Comparison of PDG for ovulatory cycles in a Bangladesh and a US subject. Samples were collected twice per week. Hormone values are corrected by specific gravity. The US cycle ends at the start of the next menses, the Bangladesh cycle ends in a full term pregnancy.

The long term reproducibility of the assays and stability of urine samples make these urinary EIAs useful for large scale prospective research on ovarian function. Urine controls are stable for at least one year (stored at –20C), and urine specimens retain their baseline PDG and E1G after being stored for 2 days at room temperature, and being subjected to 5 freeze thaw cycles. Urine specimens assayed up to 3 years after collection (Figure 5) clearly retain E1G and PDG patterns seen in the serum hormone profiles (the latter were assayed at the time of collection). This robust stability is necessary for samples collected in the remote or rural field settings often encountered in cross-cultural research (e.g. Bangladesh, or (62)).

The PDG and 3F11 EIAs work well for the physiological range observed in US women. Samples from US women are rarely if ever below the limit of detection for these assays, and dilutions can be made to measure high hormone concentrations, including those seen during pregnancy. The 3F11 E1G assay is, however, limited in its use in populations with very low levels of urinary estrone-3-glucuronide. The 155B3 E1C EIA picks up a broader range of estrogen metabolites than the 3F11 EIA and has accuracy and precision suitable for population level research. In fact, the 155B3 EIA performs similarly to, and

<table>
<thead>
<tr>
<th>Study (sample size)</th>
<th>E</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present Study (N=30)</td>
<td>0.858</td>
<td>0.977</td>
</tr>
<tr>
<td>Munro et al 1991 (N=10)</td>
<td>0.88</td>
<td>0.94</td>
</tr>
<tr>
<td>Stanczyk et al 1980 (N=7)</td>
<td>0.81</td>
<td>0.92</td>
</tr>
<tr>
<td>Kesner et al 1992 (N=10)</td>
<td>0.63</td>
<td>0.71</td>
</tr>
<tr>
<td>Kesner et al 1994 (N=10)</td>
<td>0.73</td>
<td>0.89</td>
</tr>
</tbody>
</table>

*1 N = number of subjects giving daily paired samples.*
cross-reacts with the same metabolites as the widely used R522 polyclonal antibody, which is in limited supply (1)(37)(38)(41)(42). The 155B3 assay thus provides a reliable alternative for monitoring reproductive function in a broad range of populations.

An emerging literature points to population variation in the different forms of estrogens (44)(45), (46)(47)(48)(49). Although the findings are not consistent across all studies, there is a trend in these studies and the present report, of higher levels of serum, salivary and urinary estrone and estradiol, and lower levels of estriol, in American women. The factors contributing to cross-population differences in hormone levels are unknown, but associations with diet, reproductive cancer risk, body composition, and metabolic factors (including route of excretion) have been observed or hypothesized (46)(50)(63). Most women in Bangladesh suffer from chronic under-nutrition and infectious disease, little formal education, and limited access to health care (64)(65). The average BMI for a large random sample of non-pregnant women between 15-45 years of age in 1992 in Matlab Bangladesh was 18.8 +/-1.9 (65). The total fertility rate in the Matlab nonintervention area in 1992 was 4.03 (66), indicating that reproductive function was robust despite the challenging living conditions. The differences in hormone levels observed between the Bangladesh and US urine samples in the present report are not due to field conditions or sample treatment problems; in the specimen stability experiments even the extreme conditions of 7 days at room temperature coupled with 10 freeze-thaw cycles (Figure 6) did not produce the wide gap in hormone levels observed between the Bangladesh and US samples. It is thus likely that ecological (diet, disease and work load) and/or genetic factors contributed to the low reproductive steroid hormone levels of the Bangladeshi women.

Our data suggest that population variation in steroid hormone levels should be taken into consideration when developing and applying immunoassay methods in cross-cultural research on ovarian function. In particular, specific estrogen metabolite assays may need to be optimized for each population under study. Additionally, care should be taken when applying clinical standards cross-culturally(67), for example, luteal phase function criteria based on PDG levels (35)(36)(38). The data presented here indicate that clinical thresholds for PDG levels in Western populations are not likely to be applicable in Bangladesh. Finally, in addition to examination of ovarian function across the ovarian cycle and transitions between reproductive states, the characteristics and clinical and ecological covariates of population variation can be investigated in more detail using these urinary EIA’s.

CONCLUSION

We have validated three urinary EIA’s that are well suited for anthropological, biodemographic, epidemiologic and clinical research in which large numbers of specimens are examined for reproductive steroid hormones. The assays and urine specimens are reliable and stable for large scale and prospective field research.

Population variation in hormones and their metabolites need to be taken into consideration when developing and applying urinary hormone assay methods for population level and cross-cultural research on ovarian function. The PDG EIA reported on here is likely to be useful across a wide range of applications and populations, including those where PDG levels may be quite low. On the other hand, different urinary estrone metabolite assays are necessary for Bangladesh and US samples, a consequence of population differences in metabolite levels. For populations with levels of urinary metabolites of estradiol similar to those found in US women, the assay of choice should be the 3F11 EIA, given that it has somewhat better performance characteristics than the 155B3 EIA. For applications in populations with lower levels of urinary metabolites of estradiol the 155B3 EIA provides a suitable alternative.
ACKNOWLEDGMENTS

This research was supported by NSF (DBS-9218734, DBS-9600690), NIH (NIA 5 T32 AG00208, NIA RO1 AG15141, NIA RO1 AG14579, NICHD 2 P30 HD28263 and NICHD F32 HD 07994-02), the Mellon Foundation, the Hill Foundation, the American Institute for Bangladesh Studies, the Population Council, and the Population Research Institute of Pennsylvania State University. Thanks to SH Barsom, G Davis, PK Mansfield, R Miller, S Palmer, DE Schechter and CM Stroud for their contributions to this work.

REFERENCES


