2015

Reductions in Urinary Collection Frequency for Assessment of Reproductive Hormones Provide Physiologically Representative Exposure and Mean Concentrations when Compared with Daily Collection

Heather C.M. Allaway

The Pennsylvania State University, hca111@psu.edu

Nancy J. Williams

The Pennsylvania State University, niw1@psu.edu

Rebecca J. Mallinson

The Pennsylvania State University, mallinsonrj@gmail.com

Karsten Koehler

University of Nebraska - Lincoln, kkoehler3@unl.edu

Mary Jane De Souza

~~Follow this and additional works at <http://digitalcommons.unl.edu/nutritionfacpub>~~

Part of the [Human and Clinical Nutrition Commons](#), [Molecular, Genetic, and Biochemical Nutrition Commons](#), and the [Other Nutrition Commons](#)

Allaway, Heather C.M.; Williams, Nancy J.; Mallinson, Rebecca J.; Koehler, Karsten; and De Souza, Mary Jane, "Reductions in Urinary Collection Frequency for Assessment of Reproductive Hormones Provide Physiologically Representative Exposure and Mean Concentrations when Compared with Daily Collection" (2015). *Nutrition and Health Sciences -- Faculty Publications*. 104.
<http://digitalcommons.unl.edu/nutritionfacpub/104>

This Article is brought to you for free and open access by the Nutrition and Health Sciences, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Nutrition and Health Sciences -- Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Reductions in Urinary Collection Frequency for Assessment of Reproductive Hormones Provide Physiologically Representative Exposure and Mean Concentrations when Compared with Daily Collection

Heather C.M. Allaway, Nancy I. Williams, Rebecca J. Mallinson, Karsten Koehler, and Mary Jane De Souza

Department of Kinesiology, Women's Health and Exercise Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802

Corresponding author – Mary Jane De Souza, Department of Kinesiology, Women's Health and Exercise Laboratory, The Pennsylvania State University, 104 Noll Laboratory, University Park, Pennsylvania 16802, United States of America, email mjd34@psu.edu

Abstract

Objective – To determine if reducing the frequency of urinary sample collection from daily to 5, 3, or 2 days per week during a menstrual cycle or 28-day amenorrheic monitoring period provide accurate representations of the reproductive hormone metabolites estrone-1-glucuronide (E1G) and pregnanediol glucuronide (PdG) exposure and mean concentrations.

Methods – Exercising women presenting with eumenorrhea or exercise-associated menstrual disturbances collected daily urine samples for the assessment of E1G and PdG concentrations. After enzyme immunoassay analysis of the daily samples, E1G and PdG data were systematically removed from each menstrual cycle or amenorrheic monitoring period to mimic three reduced collection frequencies, representing 5, 3, and 2 days per week. Exposure and mean concentration were calculated for both hormones and all four urinary collection frequencies.

Results – E1G and PdG exposure and mean cycle concentrations derived from reduced collection frequencies were not different from daily collection ($P > 0.05$), independent of whether menstrual cycles and monitoring periods were analyzed together or separately. Bland-Altman analysis indicated acceptable agreement between each reduced collection frequency and daily collection.

Conclusions – Compared with daily urinary collection, a reduced collection frequency of 5, 3, or 2 days each week provides accurate E1G and PdG profiles of collection periods of various lengths and types of menstrual function. Reduction of urinary sample collection frequency may enable researchers to reduce participant burden and costs, increase compliance, and study a wider range of study populations.

Hormonal evaluation of the reproductive potential of women in population based studies most often uses samples of blood (Ahrens et al., 2014; Jeppesen et al., 2013; Mumford et al., 2012), urine (De Souza et al., 2010; Direito et al., 2013; McConnell et al., 2002; O'Connor et al., 2003), or saliva (Direito et al., 2013; Ellison et al., 2007; Kesner et al., 1992; Ziomkiewicz et al., 2008). These assessments are generally limited by participant training, compliance, and the cost of the assays. Indeed, the optimal method should yield good participant tolerance and compliance with sample collection while still providing reliable, sensitive, and specific information about the participants' reproductive status (Kesner et al., 1992).

In many cases, the use of the gold standard for repeated assessment of reproductive function, i.e., daily serum sampling which involves invasive blood draws, is not feasible; therefore, the collection of daily urine samples is an advantageous alternative due to its noninvasive and self-collectable nature (Kesner et al., 1995). Daily collection of urine samples during a menstrual cycle or specified monitoring period provides information about reproductive hormone exposure and clinical endpoints of reproductive status, such as ovulation (Baird et al., 1995; Direito et al., 2013; Kassam et al., 1996; McConnell et al., 2002; O'Connor et al., 2006), pregnancy

(Cate et al., 2013; Greene et al., 2013; Holman et al., 1998; O'Connor et al., 1998), and menstrual cycle status (De Souza et al., 2010; Direito et al., 2013; Mallinson et al., 2013b; Mumford et al., 2012; O'Connor et al., 2011; Williams et al., 2010) in humans and non-human primates. Exposure to reproductive hormones has more recently been shown to be important predictors of general health and disease risk. Measures of reproductive hormone exposure from daily urinary samples have been associated with cardiovascular function in amenorrheic exercising women, specifically endothelial dysfunction, bradycardia, low systolic blood pressure, reduced regional blood flow, increased local vascular resistance, and an unfavorable lipid profile (O'Donnell et al., 2007; Rickenlund et al., 2005; Zeni Hoch et al., 2003). Further, reduced exposure to estrogen in amenorrheic exercising women, as assessed by daily urine sample collection, has been associated with increased concentrations of osteoprotegerin, an important regulator of bone resorption (West et al., 2009), and clinical measures of bone health (Mallinson et al., 2013a; Scheid et al., 2011). Risk of ovarian cancer (Parazzini et al., 1989) and breast cancer (Kossmann et al., 2011; Parazzini et al., 1993) have also been associated with exposure to reproductive hormones.

Unsupervised participants can easily collect urine samples, thus facilitating monitoring of ovarian function over extended time periods (Hall Moran et al., 2001). However, it has been noted that daily urine sample collection presents a substantial participant burden (Mumford et al., 2011). The substantial cost of time and participant burden can contribute to increased non-compliance and higher dropout rates (Mumford et al., 2011). Compliance with daily urinary sample collection is typically high in short-term studies (1–3 months). The potential for reduced compliance increases over time (Mumford et al., 2011); however, specific data on participant compliance to urinary collection is scant in publications. For studies lasting between 1 and 3 months, compliance to daily urinary sample collection is in the range of 92 to 97% (Kesner et al., 1992; Windham et al., 2002; Wright et al., 1992). For example, Kesner et al. (1992) reported that during a time period of two complete menstrual cycles, 97% of all scheduled samples were collected. In the Women's Reproductive Health Study, 93% of all daily urine samples were collected over the course of two consecutive menstrual cycles (Waller et al., 1998).

In studies lasting between 5 and 12 months, compliance to urinary sample collection is more variable and somewhat lower, ranging between 50 and 100% (Kravitz et al., 2005; Liu et al., 2004; Santoro et al., 1996). For example, retrospective analysis of the Study of Women Across the Nation Daily Hormone substudy, only 680 of 848 eligible participants had collected 80% of the required samples (Kravitz et al., 2005). In the Semi-Conductor Health Study, where participants were asked to collect urinary samples daily for five cycles, only 57% of all cycles had fewer than 3 days of missing data in any 5-day rolling window (Liu et al., 2004). In our laboratory, the participants who completed 4 or more months of a 12-month study collected an average of 90% of the requested samples; however, individual compliance ranged from 61 to 100% (unpublished data).

The design of any experiment needs to balance data quantity and quality while reducing participant burden and project cost and increasing compliance. To our knowledge, the only attempt to validate a reduced sampling frequency for use with urine specimens was conducted by O'Connor et al. (2006), who evaluated the specificity and sensitivity of reduced collection frequencies to determine the presence of ovulation with progesterone glucuronide based algorithms. The every other day reduced collection frequency accurately and precisely detected day of ovulation (O'Connor et al., 2006). Thus, a reduced collection frequency could be useful in conducting research in populations who may be hesitant to participate in research projects that involve daily urine sampling, such as children or adolescents, and may aid in collection of urinary samples in locations with limited cooling and storage capacity. In large-scale and long-term research studies, reduced collection frequencies would not only reduce project cost and participant burden but would also enable researchers to recruit from a larger demographic area due to the reduced need for storage.

To reduce the burden of collecting daily urinary hormone specimens and reduce project costs, our goal was to evaluate if a reduction in the number of collection days from 7 days per week (i.e., daily sample collection) to 5 (i.e., weekday sample collection), 3 (i.e., Monday/Wednesday/ Friday), or 2 (i.e., Monday/Thursday) days per week would provide an

accurate representation of estrone-1-glucuronide (E1G) and pregnanediol glucuronide (PdG) exposure and mean concentration during an entire menstrual cycle/monitoring period. We chose to evaluate the impact of reduced collection frequencies on E1G and PdG exposure and mean concentrations because both measures are important predictors of bone health (Mallinson et al., 2013a; Scheid et al., 2011), cardiovascular health (O'Donnell et al., 2007; Rickenlund et al., 2005; Zeni Hoch et al., 2003), and ovarian (Parazzini et al., 1989) and breast cancer risk (Iversen et al., 2011; Kossman et al., 2011; Parazzini et al., 1993; Terry et al., 2005; Whelan et al., 1994). We also sought to evaluate if the validity of the reduced sample collection frequencies would be affected by cycle type (eumenorrheic or amenorrheic) or by variability of cycle lengths (20–45 days range or 26–36 days range). As such, the purpose of this analysis was to explore the average and individual agreement of daily urine sample collection versus sample collection for 5 days, 3 days, or 2 days per week for the following variables: E1G exposure (area under the curve; AUC), E1G mean concentration, PdG exposure, and PdG mean concentration. We hypothesized that E1G and PdG cycle AUC and mean concentration would be similar when samples were collected daily versus 5 days, 3 days, or 2 days per week for a 28-day monitoring period or a menstrual cycle with an intermenstrual interval ranging from 20 to 45 days.

Materials and Methods

Experimental design

This study utilizes menstrual cycle data from subjects participating in a study conducted at two sites, University of Toronto (UT) and the Pennsylvania State University (PSU) over 8 years. Subjects included women with severe exercise-associated menstrual disturbances (EAMD), including oligomenorrhea (long and inconsistent menstrual cycle lengths of 36–90 days) and functional hypothalamic amenorrhea (the absence of menses for >90 days). The study also included a eumenorrheic exercising group (EU) that served as a control group. Concentrations of reproductive hormone metabolites were assessed in daily urinary sample collections. The study was approved by the Research Ethics Board at the UT and the Institutional Review Board at the PSU. All participants signed an approved informed consent document.

Participants

Women reporting regular menstrual cycles of 26 to 35 days for the previous 6 months before the study were recruited for the EU group, while women reporting no menses in the previous 3 months, or less than six cycles in the previous 12 months were recruited for the EAMD group. Eligibility criteria for the study included, (1) age 18 to 35 years; (2) weight stability (± 2 kg) for at least 3 months; (3) body mass index (BMI) 16 to 25 kg/m²; (4) good health as determined by medical exam and no history of any serious medical conditions; (5) no chronic illness, including hyperprolactinemia and thyroid disease; (6) not currently dieting; (7) no current clinical diagnosis of an eating disorder or psychiatric disorder; (8) non-smoking, (9) not taking any form of hormonal therapy for at least 6 months; (10) currently participating

in ≥ 2 h/week of purposeful exercise; (11) no history of a clinical diagnosis of polycystic ovarian syndrome (PCOS); (12) not pregnant, lactating, or planning a pregnancy; (13) no medication use that would alter metabolic or reproductive hormone concentrations; and (14) no other contraindications that would preclude participation in the study.

Participant grouping categories

Classification of participant menstrual status (eumenorrheic, oligomenorrheic, or amenorrheic) was based on self-reported menstrual histories and menstrual calendars. Menstrual status was confirmed by urinary concentrations of the reproductive hormone metabolites, E1G, PdG, and luteinizing hormone (LH).

Demographic assessment

Height (to the nearest 1.0 cm) and weight (to the nearest 0.1 kg) were measured and participants completed questionnaires to assess medical history, exercise and menstrual history, eating behaviors and psychological health. A physical exam and blood sample were performed to determine overall health.

Urinary collection procedures

Participants in the EAMD group collected daily urinary samples for a 28-day monitoring period and EU participants collected daily specimens for an entire menstrual cycle. The EAMD group initiated urinary collection on an arbitrary day in the study, while the EU group initiated urinary collection on day 1 or 2 of the menstrual cycle subsequent to demographic assessment. All participants utilized calendars to record menses and time of urine collection. All urine specimens were labeled with calendar date, cycle/monitoring period number, and cycle/monitoring period day. Participants stored urine specimens in their household freezers between drop offs at the laboratory. Frozen ice packs and insulated lunch packs were used to keep samples cold during transport to the laboratory. In the laboratory urine samples were stored in a -20°C freezer until analyzed.

Urinary measurement of E1G and PdG

Microtiter plate competitive enzyme immunoassays (EIA) were used to measure E1G and PdG, as previously described (De Souza et al., 2010). The secretion of these estrogen and progesterone metabolites in the urine parallels serum concentrations of the parent hormones (Munro et al., 1991; O'Connor et al., 2003). Urinary concentrations of E1G and PdG were corrected for specific gravity, determined using a hand refractometer (NSG Precision Cells, Inc., Farmingdale, NY), to account for hydration status (Boeniger et al., 1993; Haddow et al., 1994; Miller et al., 2004).

Selection of eligible cycles

A flow chart is presented (Fig. 1) to describe the design of the study and the contribution of participants and cycles from each study site. A total of 116 participants and 572 cycles/monitoring periods were evaluated for eligibility for this analysis. There were 63 amenorrheic participants with 267 28-day monitoring periods and 79 eumenorrheic participants with 305 menstrual cycles within the range of 20 to 45 days. Menstrual cycle length was defined as the number of days from day 1 of menses to the day before the

Figure 1. Number of participants and menstrual cycles/monitoring periods which contributed to and were excluded from the current analysis. Data from participants recruited at two study sites, University of Toronto (UT) and the Pennsylvania State University (PSU), were included. Initially there were 572 menstrual cycles/28-day monitoring periods from 116 participants evaluated for complete sample collection. Three hundred eighty-two menstrual cycles/monitoring periods from 61 participants were excluded from the analysis due to missing collection days. Included in the analysis were 190 menstrual cycles/monitoring periods from 55 participants, which included 90 menstrual cycles/monitoring periods from 27 participants enrolled at the UT site and 100 menstrual cycles/monitoring periods from 28 participants enrolled at the PSU site.

(contained no missing samples) that were collected during the 12-month study were used in this analysis. In addition, 41 menstrual cycles with no more than 3 missing days in the first 6 days of the cycle were also used. In these cases, concentrations of E1G and PdG for the missing days were estimated by averaging the concentrations from the day before and after the missing day. If the missing day was the first day of the cycle, days 2 to 4 of the cycle were averaged to estimate the concentration for the missing day. Menstrual cycles included in the complete sample analysis were a combination of ovulatory ($n = 23$), luteal phase defect (LPD; $n = 43$), and anovulatory ($n = 54$) cycle classifications. Classifications of menstrual cycles were completed from the original daily specimens and conducted to ensure inclusion of all types of menstrual cycles in the analysis. Specific hormonal criteria for classification of ovulatory, LPD, and anovulatory cycles has been described previously (De Souza et al., 2010). Sixty-one participants and their 382 menstrual cycles/monitoring periods were excluded from the analysis (Fig. 1). Monitoring periods (used for amenorrheic women) were excluded if there were any missing samples. Menstrual cycles (used for eumenorrheic and oligomenorrheic women) were excluded if there were more than 3 missing days in the first 6 collection days or any missing days beyond the first 6 days. Menstrual cycles were also excluded if the cycle length was outside the 20- to 45-day range included in this analysis. The UT site contributed 27 participants and 90 menstrual cycles/monitoring periods to this analysis while the PSU site contributed 28 participants and 100 menstrual cycles/monitoring periods to this analysis (Fig. 1).

Simulation of reduced number of sample collection days

The daily samples per menstrual cycle/monitoring period ($n = 190$) collected by the participants were referenced to day of the menstrual cycle/monitoring period and collection calendar date. To determine if fewer days of urine collection would provide accurate and precise data for E1G and PdG exposure and mean concentration during the menstrual cycle/monitoring period, E1G and PdG data were systematically removed from each menstrual cycle or 28-day monitoring period to mimic a reduced frequency of sample collection for participants with 100% compliance to daily collection. The reduced sample collection frequencies were selected to reduce participant burden and represented three different collection frequencies as follows: 5 days of urine collection each week, 3 days of urine collection each week, and 2 days of urine collection each week. Specifically, for the simulation of collecting five urinary samples per week, E1G and PdG data for Saturday and Sunday each week of the menstrual cycle/monitoring period were systematically, leaving only the E1G and PdG concentrations from the weekdays for analysis. For the simulation of collecting three urinary samples per week, E1G and PdG data were systematically removed for Tuesday, Thursday, Saturday, and Sunday each week of the menstrual cycle/monitoring period leaving E1G and PdG concentrations for Monday, Wednesday, and Friday for the analysis. For simulation of collecting two urinary samples per week, E1G and PdG data were systematically removed for Tuesday, Wednesday, Friday, Saturday, and Sunday each week of the menstrual cycle/monitoring period leaving only E1G and PdG concentrations for Monday and Thursday for analysis.

Urinary hormone assessment calculations

E1G and PdG exposures across the menstrual cycle or monitoring period were determined by calculating the AUC for daily, 5 days, 3 days, and 2 days per week collection frequencies using Kaleidagraph Software (Synergy Software, Reading, PA). Mean E1G and PdG concentrations across the cycle or monitoring period for all collection frequencies were also calculated.

Statistical analyses

The data presented were obtained at two different locations, the UT and PSU, over 8 years. E1G and PdG data were analyzed as a merged group of eumenorrheic cycles of 20 to 45 days in length and 28-day amenorrheic monitoring periods (complete sample analysis; $n = 190$). Subanalyses of eumenorrheic cycles of 26 to 36 days in length (which is the most common range of intermenstrual intervals among regularly menstruating women; $n = 94$) alone and 28-day amenorrheic monitoring periods ($n = 70$) alone. Data screening was conducted before statistical analysis in order to identify whether the data met the assumptions required by the specific statistical techniques in this analysis. Data screening involved examination of variable distributions within each of the three analysis groupings for all four collection frequencies (daily, 5 days, 3 days, and 2 days) and all four hormone variables (E1G AUC, PdG AUC, E1G mean, and PdG mean). All hormonal variables were found to be not normally distributed. However, logarithmic transformation did not improve normality of these variables. In addition, logarithmic

transformation was not considered as a practical approach for Bland-Altman analysis, as the limits of agreement (LOA) are expressed as multiples of the measured concentration following logarithmic transformation (Euser et al., 2008). All data are presented as means $\pm$ SD, unless otherwise indicated. Linear mixed model ANOVA was used to compare all ovarian steroid (E1G AUC, E1G mean, PdG AUC, or PdG mean) data between daily urinary collection and each reduced urinary collection frequency for the complete samples analysis and the eumenorrheic and amenorrheic subanalyses. Since the same individual provided multiple cycles and/or monitoring periods these data were considered to be of nested nature; therefore, the participant identifier was included as a random effect in the linear model. A significance level of $\alpha = 0.05$ was used to detect differences, and for multiple comparisons, α was adjusted using Bonferroni correction. Bland Altman analysis was performed to determine the 95% LOA and to identify potential mean and proportional bias for both AUC and mean concentration (Bland and Altman, 1995). Errors were calculated as the difference between daily urinary collection data and each reduced urinary collection data since daily urinary sample collection was regarded as the criterion method. For convenience, mean error and lower and upper LOA are also reported as percent of the average of daily and each reduced urinary collection values. All analyses were conducted using R statistical software (Revolution Analytics, Palo Alto, CA).

Results

Complete sample analysis

In total, there were 55 participants and 190 menstrual cycles (20–45 days in length) and 28-day monitoring periods with complete data. There were 120 menstrual cycles and 70 28-day amenorrheic monitoring periods. The participants were aged 22.6 ± 64.3 years, weighed 57.0 ± 6.6 kg, were 164.3 ± 6.6 cm tall, and had a BMI of 21.1 ± 2.0 kg/m². The average age at menarche was 13.2 ± 1.6 years and the mean gynecologic age was 9.5 ± 4.6 years.

Mixed model analysis

Composite graphs of the average E1G and PdG concentrations, respectively, across the entire cycle/monitoring period for daily, 5-day, 3-day, and 2-day collection frequencies are shown in Figure 2A, B. The average AUC and cycle mean concentration for each urinary collection frequency are displayed in inset bar graphs within the composite graphs and in Table 1. There were no significant differences detected between daily collection and each reduced collection frequency ($P > 0.99$) with regard to E1G mean concentration; however, E1G AUC for 2 days per week collection was significantly lower when compared with E1G AUC for daily collection ($P > 0.046$). There were no significant differences detected between daily collection and each reduced collection frequency (5 days, 3 days, or 2 days per week) with regard to PdG AUC ($P > 0.050$) or mean concentration ($P > 0.99$).

Bland Altman analysis

On average, the reduced urine collection frequencies for the complete sample analysis underestimate the daily E1G AUC by 1.4% for the 5-day collection frequency, 3.2% for the 3-day collection frequency, and by 8.2% for the 2-day

Figure 2. Composite graphs of urinary reproductive hormone excretion of E1G (A) and PdG (B) for eumenorrheic cycles of 20 to 45 days in length and 28-day amenorrheic monitoring periods for daily, 5-day/week, 3-day/week, and 2-day/week urinary collection frequencies. The inset bar graphs depict the cycle AUC and mean concentration for E1G and PdG for daily, 5-day/week, 3-day/week, and 2-day/week collection frequencies. Cycle day 1 represents the first day of menses. The data points for the central cycle/monitoring period days have an $n = 5$ to 190. Cycle/monitoring period days at the beginning and end of the composite graphs have a variable “ n ” due to the varying lengths of the cycles included in the sample. E1G, estrone-1-glucuronide; PdG, pregnanediol glucuronide; AUC, area under the curve. *Statistical difference ($P < 0.05$) between daily collection and 2 days per week collection frequency.

collection frequency. The 5-day collection frequency demonstrated the lowest degree of underestimation and 2-day collection frequency demonstrated the greatest degree of underestimation when compared with daily sample collection. The E1G AUC for all reduced collection frequencies demonstrate good agreement with daily urine collection as indicated by the inclusion of zero in the 95% LOA for the Bland Altman analyses. The best agreement was observed with the 5-day collection frequency, indicated by the smallest range for the 95% LOA (see Fig. 3A–C). A proportional bias was observed in all reduced collection frequencies ($P < 0.010$) indicating larger AUC values are, on average, underestimated more than smaller AUC values in all reduced

collection frequencies compared with daily urine collection (see Table 2).

On average, reduced sample collection underestimated the daily PdG AUC by 0.6% for the 5-day collection frequency, 2.9% for the 3-day collection frequency, and by 10.8% for the 2-day collection frequency (see Fig. 3D–F and Table 3). The 5-day collection frequency demonstrated the lowest degree of underestimation of daily sampling, while the 2-day collection frequency demonstrated the greatest degree of underestimation. The PdG AUC for all reduced collection frequencies demonstrate good agreement with daily urine collection as indicated by the 95% LOA including zero. The best agreement was observed with 5-day

collection frequency, indicated by the smallest range for the 95% LOA. A proportional bias was not observed in the 5-day and 3-day collection frequencies ($P = 0.10$), while a proportional bias was observed in the 2-day collection frequency

($P < 0.000$) indicating larger AUC values are, on average, underestimated more than smaller AUC values in the 2-day collection frequency compared with daily urine collection (see Table 3).

Table 1. Reproductive steroid hormone metabolite parameters for the all cycles, eumenorrheic cycles only, and amenorrheic monitoring periods only analyses for all collection frequencies

	Reduced collection frequency			
	Daily	5-d/wk	3-d/wk	2-d/wk
	Mean±SE	Mean±SE	Mean±SE	Mean±SE
All cycles and monitoring periods (n = 190)				
E1G cycle mean AUC (ng × day/mL)	856.0±36.3	844.0±35.3	829.0±35.1	789.0±34.2 ^a
E1G cycle mean concentration (ng/mL)	29.5±1.2	29.6±1.2	29.7±1.2	29.6±1.2
PdG cycle mean AUC (mg × day/mL)	50.0±2.7	49.6±2.8	48.5±2.7	44.8±2.5 ^b
PdG cycle mean concentration (mg/mL)	1.7±0.1	1.8±0.1	1.8±0.1	1.7±0.1
Eumenorrheic cycles (n = 94)				
E1G cycle mean AUC (ng × day/mL)	1,037.5±50.7	1,027.5±50.5	1,014.7±50.7	972.5±49.2
E1G cycle mean concentration (ng/mL)	35.7±1.7	35.9±1.7	36.2±1.7	36.2±1.8
PdG cycle mean AUC (mg × day/mL)	68.4±4.2	68.4±4.2	67.0±4.1	62.2±4.0 ^b
PdG cycle mean concentration (mg/mL)	2.4±0.1	2.4±0.2	2.4±0.2	2.4±0.2
Amenorrheic periods (n = 70)				
E1G cycle mean AUC (ng × day/mL)	539.2±27.7	532.6±27.3	515.5±26.5	488.6±27.5
E1G cycle mean concentration (ng/mL)	20.0±1.0	20.0±1.0	20.0±1.0	20.0±1.1
PdG cycle mean AUC (mg × day/mL)	25.6±1.6	25.2±1.6	24.2±1.5	22.7±1.5
PdG cycle mean concentration (mg/mL)	1.0±0.1	1.0±0.1	0.9±0.1	0.9±0.1

Reproductive steroid metabolite parameters include estrone-1-glucunonide (E1G) and pregnanediol glucuronide (PdG) area under the curve (AUC) and mean concentration. d/wk, days per week.

- a. Significant differences ($P<0.05$) between the reduced collection frequency and daily collection frequency.
- b. A trend for a difference ($P<0.1$) between the reduced collection frequency and daily collection frequency.

Figure 3. Bland Altman plots for all cycles 20 to 45 days in length and 28-day monitoring periods. The difference between daily and reduced collection frequencies are plotted against the mean of the daily and reduced collection frequency in the 190 paired measurements from the all cycles/monitoring periods analysis. The comparison of daily and 5 days/week collection frequency is in column 1, daily and 3 days/week collection frequency is in column 2, and daily and 2 days/week collection frequency is in column 3. Differences between daily and the reduced collection frequencies for E1G AUC (A-C) and PdG AUC (D-F) are demonstrated.

On average, daily E1G mean concentration was overestimated by 0.5% for the 5-day collection frequency, 0.8% for the 3-day collection frequency, and by 0.4% for the 2-day collection frequency (see Fig. 4A–C and Table 4). The 2-day collection frequency demonstrated the lowest degree of overestimation of daily sampling, while the 3-day collection frequency demonstrated the greatest degree of overestimation. The E1G cycle mean concentration for all reduced collection frequencies demonstrate good agreement with daily urine collection as indicated by the 95% LOA including zero. The best agreement was observed with the 5-day collection frequency, indicated by the smallest range for the 95% LOA. A proportional bias was not observed in the 5-day and 3-day collection frequencies ($P > 0.60$), while a proportional bias was observed in the 2-day collection frequency ($P = 0.040$) indicating greater cycle mean concentrations are, on average, overestimated more than smaller mean cycle concentrations in 2-day collection frequency compared with daily collection (see Table 4).

Daily PdG mean concentration was overestimated by 2.2% for the 5-day collection frequency and 2.1% for the 3-day collection frequency, while the 2-day collection frequency underestimated the daily PdG mean concentration by

0.2% (see Fig. 4D–F and Table 5). The 5-day collection frequency demonstrated the greatest degree of overestimation of daily sample collection, while the 2-day collection frequency demonstrated underestimation of daily sample collection. The PdG cycle mean concentration for all reduced collection frequencies demonstrate good agreement with daily urine collection as indicated by the 95% LOA including zero. The best agreement was observed with the 5-day collection frequency, indicated by the smallest range for the 95% LOA. A proportional bias was observed in the 5-day and 3-day collection frequencies ($P < 0.010$) indicating greater cycle mean concentrations are, on average, overestimated compared with daily urine collection, while a trend toward a proportional bias was observed in the 2-day collection frequency ($P = 0.060$; see Table 5).

Subanalysis of eumenorrheic cycles

A subanalysis of eumenorrheic menstrual cycles of 26 to 36 days in length included 31 participants and 94 menstrual cycles with complete data. This subanalysis included anovulatory ($n = 14$), LPD ($n = 31$), and ovulatory ($n = 49$) menstrual cycles, which were classified using daily collection frequency hormonal measurements. The participants were

Table 2. Bland-Altman analysis for all menstrual cycle/monitoring period analyses of E1G AUC

Method	Bland-Altman analysis						Proportional bias
	Mean difference		Lower limit of agreement		Upper limit of agreement		
	(ng x day/mL)	%	(ng x day/mL)	%	(ng x day/mL)	%	
All cycles and monitoring periods (n = 190)							
Daily vs. 5-d/wk	12.0	1.4	-133.8	-15.7	157.9	18.6	0.0132
Daily vs. 3-d/wk	27.2	3.2	-141.8	-16.8	196.2	23.3	0.0140
Daily vs. 2-d/wk	67.0	8.2	-136.8	-16.6	270.8	32.9	0.0001
Eumenorrheic cycles (n = 94)							
Daily vs. 5-d/wk	10.0	1.0	-106.1	-10.3	126.1	12.2	0.6988
Daily vs. 3-d/wk	22.7	2.2	-122.1	-11.9	167.5	16.3	0.9908
Daily vs. 2-d/wk	65.0	6.5	-127.4	-12.7	257.3	25.6	0.1528
Amenorrheic periods (n = 70)							
Daily vs. 5-d/wk	6.6	1.2	-65.9	-12.3	79.1	14.8	0.4515
Daily vs. 3-d/wk	23.7	4.5	-60.4	-11.5	107.8	20.4	0.0405
Daily vs. 2-d/wk	50.6	9.8	-84.1	-16.4	185.2	36.0	0.8165

E1G, estrone-1-glucuronide; AUC, area under the curve; d/wk, days per week.

Table 3. Bland-Altman analysis for all menstrual cycle/monitoring period analyses of PdG AUC

Method	Bland-Altman analysis						Proportional bias
	Mean difference		Lower limit of agreement		Upper limit of agreement		
	(mg x day/mL)	%	(mg x day/mL)	%	(mg x day/mL)	%	
All cycles and monitoring periods (n = 190)							
Daily vs. 5-d/wk	0.3	0.6	-8.8	-17.7	9.5	19.0	0.1819
Daily vs. 3-d/wk	1.4	2.9	-12.7	-25.8	15.6	31.7	0.4675
Daily vs. 2-d/wk	5.1	10.8	-12.4	-26.3	22.7	47.9	0.0000
Eumenorrheic cycles (n = 94)							
Daily vs. 5-d/wk	0.1	0.1	-12.1	-17.6	12.2	17.8	0.5247
Daily vs. 3-d/wk	1.4	2.0	-17.8	-26.3	20.5	30.3	0.3789
Daily vs. 2-d/wk	6.2	9.6	-15.8	-24.3	28.3	43.4	0.0541
Amenorrheic periods (n = 70)							
Daily vs. 5-d/wk	0.4	1.7	-3.5	-13.8	4.3	17.1	0.1814
Daily vs. 3-d/wk	1.4	5.6	-3.7	-15.0	6.5	26.3	0.0016
Daily vs. 2-d/wk	2.8	11.8	-4.4	-18.2	10.1	41.8	0.0314

PdG, pregnanediol glucuronide; AUC, area under the curve; d/wk, days per week.

aged 24±4.5 years, weighed 57.1±6.0 kg, were 164.5±6.5 cm tall, and had a BMI of 21.1±1.7 kg/m². The average age at menarche was 12.5±.3 years and the mean gynecologic age was 11.5±4.3 years.

Mixed model analysis

Composite graphs of the average E1G and PdG concentrations, respectively, across the entire cycle for daily, 5-day, 3-day, and 2-day collection frequencies are shown in Figure 5A,B. The classic characteristics of ovulatory cycles, i.e. the mid-cycle E1G peak and luteal phase PdG peak, are

evident. The average AUC and cycle mean concentration for each urinary collection frequency are displayed in inset bar graphs within the composite graphs and in Table 1. There were no significant differences detected between daily collection and each reduced collection frequency (5 days, 3 days, or 2 days per week) with regard to E1G AUC ($P > 0.46$) or mean concentration ($P > 0.99$). There were no significant differences detected between daily collection and each reduced collection frequency (5 days, 3 days, or 2 days per week) with regard to PdG AUC ($P > 0.27$) or mean concentration ($P > 0.99$).

Figure 4. Bland Altman plots for all cycles 20 to 45 days in length and 28-day monitoring periods. The difference between daily and reduced collection frequencies are plotted against the mean of the daily and reduced collection frequency in the 190 paired measurements from the all cycles/monitoring periods analysis. The comparison of daily and 5 days/week collection frequency is in column 1, daily and 3 days/week collection frequency is in column 2, and daily and 2 days/week collection frequency is in column 3. Differences between daily and the reduced collection frequencies for E1G mean (A–C) and PdG mean (D–F) are demonstrated.

Table 4. Bland-Altman analysis for all menstrual cycle/monitoring period analyses of E1G mean

Method	Bland-Altman analysis						Proportional bias
	Mean difference		Lower limit of agreement		Upper limit of agreement		
	(ng/mL)	%	(ng/mL)	%	(ng/mL)	%	
All cycles and monitoring periods (n = 190)							
Daily vs. 5-d/wk	-0.1	-0.5	-3.8	-13.0	3.6	12.1	0.9569
Daily vs. 3-d/wk	-0.2	-0.8	-5.2	-17.7	4.8	16.1	0.5725
Daily vs. 2-d/wk	-0.1	-0.4	-6.7	-22.6	6.4	21.8	0.0352
Eumenorrheic cycles (n = 94)							
Daily vs. 5-d/wk	-0.3	-0.8	-3.8	-10.6	3.3	9.1	0.4931
Daily vs. 3-d/wk	-0.5	-1.4	-5.3	-14.7	4.2	11.8	0.1312
Daily vs. 2-d/wk	-0.5	-1.5	-7.1	-19.7	6.0	16.7	0.0067
Amenorrheic periods (n = 70)							
Daily vs. 5-d/wk	0.0	0.0	-1.8	-9.1	1.8	9.1	0.5498
Daily vs. 3-d/wk	0.0	-0.2	-2.9	-14.3	2.8	14.0	0.8204
Daily vs. 2-d/wk	0.0	-0.2	-5.7	-28.7	5.7	28.4	0.0066

E1G, estrone-1-glucuronide; d/wk, days per week.

Bland Altman analysis

The reduced collection frequencies for eumenorrheic cycles of 26 to 36 days in length, on average, underestimate the daily E1G AUC as shown in Table 2. The E1G AUC for all reduced collection frequencies demonstrate good agreement with daily urine collection as indicated by the 95% LOA including zero. The best agreement is observed with 5-day collection, indicated by the tighter range for the 95% LOA. A proportional bias was not observed in any of the reduced collection frequencies ($P > 0.20$), as shown in Table 2. As shown in Table 3, the reduced collection frequencies for eumenorrheic cycles alone, on average, underestimated the daily PdG AUC. The smallest range for the 95% LOA was observed in the 5-day collection frequency indicating the best agreement with daily sample collection. A proportional bias was not observed in any of the reduced collection frequencies ($P > 0.050$), as shown in Table 3.

Reduction of the collection frequency for eumenorrheic cycles of 26 to 36 days in length, on average, overestimated the E1G mean concentration as shown in Table 4. The 5-day collection frequency demonstrated the lowest degree of overestimation of E1G cycle mean concentration, while the 2-day collection frequency demonstrated the greatest degree of overestimation. A proportional bias was not observed in the 5-day and 3-day collection frequencies ($P > 0.10$); however, a proportional bias was observed in the 2-day collection frequency ($P = 0.007$), as shown in Table 4. On average, the daily PdG cycle mean concentration was overestimated by the reduced collection frequencies for eumenorrheic cycles of 26 to 36 days in length, as shown in Table 5. The best agreement was observed with 5-day collection frequency, indicated by the smaller range for the 95% LOA. A proportional bias was not observed in the 3-day collection frequency ($P = 0.20$); however, a proportional bias was observed in the 5-day and 2-day collection frequencies ($P < 0.030$), as shown in Table 5.

Subanalysis of amenorrheic monitoring periods

In a subanalysis of amenorrheic monitoring periods of 28 days there were 19 participants and 70 monitoring periods with complete data. The participants were aged 21.2 ± 3.5 years, weighed 57.3 ± 7.8 kg, were 165.1 ± 0 cm tall, and had

a BMI of 21.1 ± 2.4 kg/m². The average age at menarche was 14.1 ± 1.4 years and the mean gynecologic age was 7 ± 3.9 years.

Mixed model analysis

Composite graphs of the average E1G and PdG concentrations, respectively, across the entire monitoring period for daily, 5-day, 3-day, and 2-day collection frequencies are shown in Figure 6A,B. The chronic suppression of E1G and PdG that is characteristic of reproductive hormone concentrations among amenorrheic women is clearly evident. Within the inset bar graphs and Table 1, the average AUC and cycle mean concentrations are displayed for each frequency of urinary collection. There were no significant differences detected between daily collection and each reduced collection frequency (5 days, 3 days, or 2 days per week) with regard to E1G AUC ($P > 0.12$) or mean concentration ($P > 0.99$). There were no significant differences detected between daily collection and each reduced collection frequency (5 days, 3 days, or 2 days per week) with regard to PdG AUC ($P > 0.080$) or mean concentration ($P > 0.99$).

Bland Altman analysis

In amenorrheic 28-day monitoring periods daily sample collection E1G AUC was, on average, underestimated by the reduced collection frequencies (see Table 2). The E1G AUC for all reduced collection frequencies demonstrate good agreement with daily urine collection as indicated by the 95% LOA including zero. The best agreement was observed with the 5-day sample collection frequency, indicated by the smaller range for the 95% LOA. A proportional bias was not observed in the 5-day and 2-day collection frequencies ($P > 0.20$), but a proportional bias was observed in the 3-day collection frequency ($P = 0.04$; see Table 2). Daily PdG AUC was, on average, underestimated by the reduced collection frequencies for 28-day monitoring periods (see Table 3). The smallest range for the 95% LOA was observed in the 5-day collection frequency, indicating the best agreement with daily sample collection. A proportional bias was not observed in the 5-day collection frequency ($P > 0.20$); however, a proportional bias was observed in the 3-day and 2-day collection frequencies ($P < 0.030$; see Table 3).

Table 5. Bland-Altman analysis for all menstrual cycle/monitoring period analyses of PdG mean

Method	Bland-Altman analysis						Proportional bias
	Mean difference		Lower limit of agreement		Upper limit of agreement		
	(ng/mL)	%	(ng/mL)	%	(ng/mL)	%	
All cycles and monitoring periods (n 5 190)							
Daily vs. 5-d/wk	0.0	-2.2	-0.3	-16.5	0.2	12.2	0.0000
Daily vs. 3-d/wk	0.0	-2.1	-0.5	-29.0	0.4	24.7	0.0115
Daily vs. 2-d/wk	0.0	0.2	-0.6	-37.3	0.7	37.7	0.0635
Eumenorrheic cycles (n = 94)							
Daily vs. 5-d/wk	-0.1	-2.9	-.4	-16.3	0.3	10.6	0.0004
Daily vs. 3-d/wk	-0.1	-3.1	-.7	-29.2	0.6	23.1	0.2167
Daily vs. 2-d/wk	0.0	-1.3	-.9	-36.2	0.8	33.5	0.0296
Amenorrheic periods (n = 70)							
Daily vs. 5-d/wk	0.0	-0.6	-0.1	-11.8	0.1	10.5	0.2336
Daily vs. 3-d/wk	0.0	0.7	-.2	-16.9	0.2	18.3	0.3317
Daily vs. 2-d/wk	0.0	1.6	-.2	-26.3	0.3	29.4	0.4256

PdG, pregnanediol glucuronide; d/wk, days per week.

Figure 5. Composite graphs of urinary reproductive hormone excretion of E1G (A) and PdG (B) for eumenorrheic cycles of 26 to 36 days in length for daily, 5-day/week, 3-day/week, and 2-day/week urinary collection frequencies. The inset bar graphs depict the cycle AUC and mean concentration for daily, 5-day/week, 3-day/week, and 2-day/week collection frequencies. Cycle day 1 is the first day of menses. The data points for the central cycle/monitoring period days have an n 5 94. Cycle/monitoring period days at the beginning and end of the composite graph have a variable “n” due to the varying lengths of the cycles included in the subsample. E1G, estrone-1-glucuronide; PdG, pregnanediol glucuronide; AUC, area under the curve.

The 5-day collection frequency for 28-day monitoring periods, on average, estimated the daily E1G cycle mean concentration, while 3-day and 2-day collection frequencies overestimated E1G cycle mean concentration (see Table 4). The calculated E1G cycle mean concentration for all 28-day monitoring periods and reduced collection frequencies demonstrate good agreement with daily urine collection as indicated by the 95% LOA including zero. A proportional bias was not observed in the 5-day and 3-day collection frequencies ($P > 0.50$); whereas, a proportional bias was observed in the 2-day collection frequency ($P < 0.007$; see Table 4). The 3-day and 2-day collection frequencies for 28-day monitoring periods underestimate the daily PdG cycle mean concentration, on average, while the 5-day collection frequency overestimated the PdG mean cycle concentration (see Table

5). The best agreement is observed with 5-day collection frequency, indicated by the smaller range for the 95% LOA. A proportional bias was not observed in any of the reduced collection frequencies ($P > 0.20$; see Table 5).

Discussion

The present study was designed to assess the level of agreement between reduced urinary collection frequencies (5 days, 3 days, or 2 days per week) and the urinary gold standard of daily specimen collection in a sample of exercising eumenorrheic and amenorrheic women and in specific subpopulations of eumenorrheic and amenorrheic exercising women. This report is the first to provide detailed information on the accuracy and precision of quantifying

Figure 6. Composite graphs of urinary reproductive hormone excretion of E1G (A) and PdG (B) for 28 day amenorrheic monitoring periods for daily, 5-day/week, 3-day/week, and 2-day/week urinary collection frequencies. The inset bar graphs depict the monitoring period AUC and mean concentration for daily, 5-day/week, 3-day/week, and 2-day/week collection frequencies. Monitoring period day 1 is the first day of sample collection. The data points for the monitoring period days have an $n = 70$. E1G, estrone-1-glucuronide; PdG, pregnane-20 α -diol glucuronide; AUC, area under the curve.

reproductive hormone exposure using a reduced sampling schedule. This report supports and builds upon the work of O'Connor et al. (2006), which tested the presence and day of ovulation during reduced urinary collection frequencies. All comparisons between daily collection and each of the reduced collection frequencies (5 days, 3 days, or 2 days per week) for the average AUC and mean concentration for both E1G and PdG were not different when eumenorrheic cycles and amenorrheic monitoring periods were analyzed together and when analyzed separately; however, a difference was observed between daily collection and 2 days per week collection frequency with regard to the average E1G AUC in the complete cycle analysis. In general the 5-day collection

frequency demonstrated the best agreement with daily sample collection in all three analyses. The 3-day and 2-day collection frequencies also showed good agreement with daily sample collection in all three analyses of exercising women. The LOA for individual AUC and mean concentrations, however, appeared to increase with reduced sample collection frequencies. Depending on the required level of accuracy and precision, researchers may choose to use varying sampling frequencies. The present analysis is the first to quantify mean error as well as individual agreement for urinary reproductive hormones.

The Bland Altman analysis was used to compare the daily and reduced collection frequencies (Bland and Altman,

1986). This technique was used to assess the agreement between reduced collection frequencies (5 days, 3 days, or 2 days each week) and daily sample collection. If the differences between the reduced collection frequencies and daily collection are not large enough to change the interpretation of the research findings, daily collection could be replaced by the reduced collection frequencies or the collection frequencies could be used interchangeably. Though there were small over- or underestimations of the AUC and mean concentrations observed in the Bland Altman analysis, good agreement among the reduced collection frequencies compared with daily collection frequency was observed according to the inclusion of zero in the Bland Altman LOA. The Bland Altman analysis indicated that the AUC for E1G and PdG from reduced urinary collection frequencies were within 9.8% and 11.8% of the daily measures, respectively, in the complete samples analysis and separate eumenorrheic and amenorrheic analyses. The AUC percent differences would have enabled O'Donnell et al. (2007) and West et al. (2009) to detect the observed 33 to 98% differences in E1G AUC between sedentary and exercising eumenorrheic women and exercising amenorrheic women. The mean concentrations for E1G and PdG from reduced urinary collection frequencies were within 1.5% and 3.1% of the daily measures, respectively, for all three analyses. Differences in mean concentrations between ethnicities, which have been reported to lie between 1.0 and 22.3% (E1G) and 3.0 and 31.7% (PdG) (Santoro et al., 2004), would have been observed using 2 days per week as collection frequency. However, observing subtle differences of 1 to 5% in E1G and PdG AUC, as reported between sedentary and exercising eumenorrheic women (O'Donnell et al., 2007; West et al., 2009) would require the urinary collection frequency to be no fewer than 3 days per week (3.2% for E1G and 2.9% for PdG mean percent difference). Use of reduced urinary collection frequencies as few as 2 days per week to assess group differences in cycle AUC and mean concentration provides comparable values to daily urinary collection. Thus, the use of reduced collection frequencies would still allow detection of differences between experimental groups.

In a study by Mumford et al. (2012) 8 serum samples were collected during biologically relevant windows timed to a standardized 28-day menstrual cycle. The serum samples were utilized to assess the exposure to reproductive hormones in women with short (<26 days), normal (26–35 days), and long (>35 days) menstrual cycle lengths. Unfortunately, the authors did not conduct a comparison of the eight samples to daily sampling, but the AUC values presented by Mumford et al. (2012) demonstrated a high standard error indicating eight samples per menstrual cycle to be too large a reduction in sample collection. The authors commented that collecting only eight serum samples throughout the menstrual cycle reduced the generalizability of their findings in women with longer and more irregular cycles (Mumford et al., 2012). The basic shape of the E1G and PdG curves across the menstrual cycles in our analysis indicate that reducing the urinary collection frequency as low as eight samples per cycle (2 days per week in a 28-day menstrual cycle or monitoring period) increases the likelihood of missing the peak hormonal concentrations, thus greatly altering the AUC and mean concentration calculations. In our analysis a reduced collection frequency of less than 5 days per week was

associated with an increased range for the LOA due to the potential of missing peak hormonal concentrations, which may also compromise the use of reduced collection frequencies in studies evaluating individual subjects and clinical outcomes, such as day of ovulation.

The shape and timing of the peaks of E1G and PdG greatly affect the determination of exposure and cycle mean concentration. For example, the E1G peak is generally narrow and rises quickly to a peak, while the PdG peak is a broad curve. Hence, the narrow peak for E1G is more likely to be missed by a reduced sample collection frequency compared with the peak of a broad curve, like PdG. Differences in capture of the E1G and PdG peaks are shown in the composite graphs (Figs. 2, 5, and 6). Figure 5 shows that the broad PdG peak was not influenced by the reduced collection frequencies; however, the shape of the E1G peak is varied, though not visually different, between the reduced collection frequencies and daily collection hormone profile. Within a group analysis a proportion of the narrow peaks would be captured with the reduced collection frequencies, thus the increased LOA; however, when analyzing a cycle from an individual participant it is highly likely the E1G peak concentration would have occurred on a non-sample collection day. When evaluating individual cycles for a case report, as seen with Mallinson et al. (2013), a reduction in the collection frequency as low as 2 days per week would not influence the characterization of the cycle via hormonal exposure in a participant who has had amenorrhea for a long period of time. In a participant who had amenorrhea for a short period of time a reduction in sample collection to 2 days a week would underestimate the exposure of the participant to reproductive hormones due to higher variations in monitoring period peaks of E1G and PdG across a year (Mallinson et al., 2013b).

Urinary samples are self-collectable, noninvasive, can be easily stored and transported, and have been shown to tolerate a wide variety of non-perfect experimental conditions in the field (O'Connor et al., 2006; Kesner et al., 1995). Initially it was assumed that daily samples could be collected for prolonged periods of time with a high degree of compliance, however, many recent studies have shown compliance to be highly variable depending on study length (Kesner et al., 1992; Kravitz et al., 2005; Liu et al., 2004; Santoro et al., 1996; Waller et al., 1998; Wright et al., 1992). The level of compliance to sample collection may be due to personal or housemate comfort level with storage of specimens or ability to store samples when away from the primary residential address. Participants in early studies which evaluated perceptions of urinary sample collection reported that the benefits of increased knowledge about their body outweighed the uncomfortable nature of urine collection (Wilcox et al., 1987; Wright et al., 1992). According to Wright et al. (1992), urine collection was one of the least objectionable of the eight methods used to assess reproductive function (transvaginal ultrasound, basal body temperature, salivary electrical resistance, blood sampling, salivary samples, vaginal mucus electrical resistance, and manual cervical mucus consistency). The study by Wright et al. (1992) was one of the first to assess the attitudes of the general population, instead of nurses, to reproductive hormone collection methods. Anecdotal evidence from our laboratory has indicated that participant travel over weekends and for vacations lead to large gaps in sample collection and decreased

compliance, thus the reduced collection frequencies evaluated in this article provided the participants with weekends free from sample collection. In fact, for this analysis we excluded 67% of available cycles due to reduced compliance. In studies of long duration it is unavoidable to have participants run into vacation time such as Thanksgiving, family summer holidays, or winter break, when reduced compliance to sample collection is more likely to occur.

One limitation of the present analysis was the restriction of the data set to only include subjects who were 100% compliant to the reduced collection frequencies. The presence of missed collection days within the reduced sample collection frequencies would decrease the accuracy of the urinary reproductive hormone metabolite profile across the menstrual cycle or monitoring period. Reduced collection frequency strategies still require the use of a valid menstrual cycle calendar in order to create an accurate presentation of the hormonal profile. Not all urinary collections for eumenorrheic cycles will begin on the first day of the menstrual cycle nor will the final collection always be on the last day of the menstrual cycle. Another limitation of the study was that we did not evaluate clinical reproductive outcomes in this analysis; however, such outcomes are integral in the usefulness of reduced collection frequencies. Future studies should evaluate the validity of the reduced collection frequencies evaluated in this article in detecting luteal sufficiency and assessing this strategy in an independent sample.

The strengths of this analysis are that there were a large number of cycles included, individual participants provided multiple cycles, and the cycles used for reducing the collection frequency were based on menstrual cycles, and 28-day monitoring periods that had all samples collected. In the data set, there are near equal numbers of short, normal, and long menstrual cycles in the eumenorrheic participant cycles and 28-day monitoring periods. Within our eumenorrheic cycles we intentionally included all cycle classifications (LPD, anovulatory, and ovulatory cycles) to show that even with cycles of highly variable hormone levels the reduced collection frequencies continue to have good agreement with daily urinary sample collection.

Though daily urinary collection is the most accurate reflection of reproductive hormone production when using urinary analysis, our results demonstrate that accurate E1G and PdG profiles of menstrual cycles of various lengths (20–45 days) and types (eumenorrheic and amenorrheic) can be measured with reduced urinary collection frequencies. This work supports and builds on the work of O'Connor et al. (2006) demonstrating that a reduced sampling schedule can provide useful and accurate information in a manner that is comparable to that obtained from daily sampling of urine regarding ovarian hormone exposure and mean concentrations. The accuracy in quantifying exposure allows reduced collection frequency strategies to be utilized in under researched populations and in less developed regions around the world, where the capacity to store samples in a cold environment may be limited. The reduced collection frequencies produced composite E1G and PdG profiles for an entire cycle or monitoring period that were similar to the composite graphs of daily urinary collection. We have shown that daily ovarian steroid levels are not necessarily required to quantify AUC or mean values E1G or PdG across the entire menstrual cycle or monitoring period when conducting

group examinations for the assessment of disease risk, such as osteoporosis, endothelial dysfunction, ovarian cancer, and breast cancer, in large populations. Further research is needed to evaluate whether clinical outcomes, such as luteal sufficiency, are possible to determine through these specific reduced collection frequencies. We suggest reducing the urinary collection frequency in an effort to reduce the participant burden, increase compliance, and decrease project costs, depending on the accuracy and precision required to answer the reproductive questions of interest.

Funding — United States Department of Defense (DOD); Contract grant number: PR054531 (to M.J.D. and N.I.W.); Contract grant sponsor: National Institutes of Health; Contract grant number: M01RR10732.

Author contributions — H.C.M.A. was responsible for study execution, data collection, subanalysis design, data analysis and interpretation, and manuscript writing. N.I.W. and M.J.D. designed the study and obtained study approvals, supervised data collection and analysis, and participated in data interpretation. M.J.D. supervised the design of the subanalysis and writing of the manuscript. R.J.M. participated in study execution, data analysis and interpretation, and manuscript revision. K.K. participated in statistical analysis and interpretation, and manuscript revision.

Acknowledgments — The authors thank the participants whose dedication to sample collection was invaluable for the completion of this analysis.

Literature cited

- Ahrens KA, Vladutiu CJ, Mumford SL, Schliep KC, Perkins NJ, Wactawski-Wende J, Schisterman EF. 2014. The effect of physical activity across the menstrual cycle on reproductive function. *Ann Epidemiol* 24:127–134.
- Baird DD, McConaughy DR, Weinberg CR, Musey PI, Collins DC, Kesner JS, Knecht EA, Wilcox AJ. 1995. Application of a method for estimating day of ovulation using urinary estrogen and progesterone metabolites. *Epidemiology* 6:547–550.
- Bland JM, Altman DG. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310.
- Bland JM, Altman DG. 1995. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet* 346: 1085–1087.
- Boeniger MF, Lowry LK, Rosenberg J. 1993. Interpretation of urine results used to assess chemical exposure with emphasis on creatinine adjustments: a review. *Am Ind Hyg Assoc J* 54:615–627.
- Cate FL, Moffett C, Gronowski AM, Grenache DG, Hartmann KE, Woodworth A. 2013. Analytical and clinical validation of the Immulite 1000 hCG assay for quantitative analysis in urine. *Clin Chim Acta* 421: 104–108.
- De Souza MJ, Toombs RJ, Scheid JL, O'Donnell E, West SL, Williams NI. 2010. High prevalence of subtle and severe menstrual disturbances in exercising women: confirmation using daily hormone measures. *Hum Reprod* 25:491–503.
- Direito A, Bailly S, Mariani A, Ecochard R. 2013. Relationships between the luteinizing hormone surge and other characteristics of the menstrual cycle in normally ovulating women. *Fertil Steril* 99:279–285.

- Ellison PT, Lipson SF, Jasienska G, Ellison PL. 2007. Moderate anxiety, whether acute or chronic, is not associated with ovarian suppression in healthy, well-nourished, Western women. *Am J Phys Anthropol* 134: 513–519.
- Euser AM, Dekker FW, le Cessie S. 2008. A practical approach to Bland-Altman plots and variation coefficients for log transformed variables. *J Clin Epidemiol* 61:978–982.
- Greene DN, Schmidt RL, Kamer SM, Grenache DG, Hoke C, Lorey TS. 2013. Limitations in qualitative point of care hCG tests for detecting early pregnancy. *Clin Chim Acta* 415:317–321.
- Haddow JE, Knight GJ, Palomaki GE, Neveux LM, Chilmontczyk BA. 1994. Replacing creatinine measurements with specific gravity values to adjust urine cotinine concentrations. *Clin Chem* 40:562–564.
- Hall Moran V, Leathard HL, Coley J. 2001. Urinary hormone levels during the natural menstrual cycle: the effect of age. *J Endocrinol* 170:157–164.
- Holman DJ, Rasheed FN, Stroud CM, Brindle E, O'Connor KA, Campbell KL. 1998. A commercial pregnancy test modified for field studies of fetal loss. *Clin Chim Acta* 271:25–44.
- Iversen A, Thune I, McTiernan A, Emaus A, Finstad SE, Flote V, Wilsgaard T, Lipson SF, Ellison PT, Jasienska G, Furberg AS. 2011. Ovarian hormones and reproductive risk factors for breast cancer in premenopausal women: the Norwegian EBBA-I study. *Hum Reprod* 26: 1519–1529.
- Jeppesen JV, Anderson RA, Kelsey TW, Christiansen SL, Kristensen SG, Jayaprakasan K, Raine-Fenning N, Campbell BK, Yding Andersen C. 2013. Which follicles make the most anti-Müllerian hormone in humans? Evidence for an abrupt decline in AMH production at the time of follicle selection. *Mol Hum Reprod* 19:519–527.
- Kassam A, Overstreet JW, Snow-Harter C, De Souza MJ, Gold EB, Lasley BL. 1996. Identification of anovulation and transient luteal function using a urinary pregnanediol-3-glucuronide ratio algorithm. *Environ Health Perspect* 104:408–413.
- Kesner JS, Knecht EA, Krieg EF Jr. 1995. Stability of urinary female reproductive hormones stored under various conditions. *Reprod Toxicol* 9:239–244.
- Kesner JS, Wright DM, Schrader SM, Chin NW, Krieg EF Jr. 1992. Methods of monitoring menstrual function in field studies: efficacy of methods. *Reprod Toxicol* 6:385–400.
- Kossmann DA, Williams NI, Domchek SM, Kurzer MS, Stopfer JE, Schmitz KH. 2011. Exercise lowers estrogen and progesterone levels in premenopausal women at high risk of breast cancer. *J Appl Physiol* (1985) 111: 1687–1693.
- Kravitz HM, Janssen I, Santoro N, Bromberger JT, Schocken M, Everson-Rose SA, Karavolos K, Powell LH. 2005. Relationship of day-to-day reproductive hormone levels to sleep in midlife women. *Arch Intern Med* 165:2370–2376.
- Liu Y, Gold EB, Lasley BL, Johnson WO. 2004. Factors affecting menstrual cycle characteristics. *Am J Epidemiol* 160:131–140.
- Mallinson RJ, Williams NI, Hill BR, De Souza MJ. 2013a. Body composition and reproductive function exert unique influences on indices of bone health in exercising women. *Bone* 56:91–100.
- Mallinson RJ, Williams NI, Olmsted MP, Scheid JL, Riddle ES, De Souza MJ. 2013b. A case report of recovery of menstrual function following a nutritional intervention in two exercising women with amenorrhea of varying duration. *J Int Soc Sports Nutr* 10:34.
- McConnell HJ, O'Connor KA, Brindle E, Williams NI. 2002. Validity of methods for analyzing urinary steroid data to detect ovulation in athletes. *Med Sci Sports Exerc* 34:1836–1844.
- Miller RC, Brindle E, Holman DJ, Shofer J, Klein NA, Soules MR, O'Connor KA. 2004. Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations. *Clin Chem* 50:924–932.
- Mumford SL, Schisterman EF, Gaskins AJ, Pollack AZ, Perkins NJ, Whitcomb BW, Ye A, Wactawski-Wende J. 2011. Realignment and multiple imputation of longitudinal data: an application to menstrual cycle data. *Paediatr Perinat Epidemiol* 25:448–459.
- Mumford SL, Steiner AZ, Pollack AZ, Perkins NJ, Filiberto AC, Albert PS, Mattison DR, Wactawski-Wende J, Schisterman EF. 2012. The utility of menstrual cycle length as an indicator of cumulative hormonal exposure. *J Clin Endocrinol Metab* 97:E1871–E1879.
- Munro CJ, Stabenfeldt GH, Cragun JR, Addiego LA, Overstreet JW, Lasley BL. 1991. Relationship of serum estradiol and progesterone concentrations to the excretion profiles of their major urinary metabolites as measured by enzyme immunoassay and radioimmunoassay. *Clin Chem* 37:838–844.
- O'Connor KA, Brindle E, Holman DJ, Klein NA, Soules MR, Campbell KL, Kohen F, Munro CJ, Shofer JB, Lasley BL, Wood JW. 2003. Urinary estrone conjugate and pregnanediol 3-glucuronide enzyme immunoassays for population research. *Clin Chem* 49:1139–1148.
- O'Connor KA, Brindle E, Miller RC, Shofer JB, Ferrell RJ, Klein NA, Soules MR, Holman DJ, Mansfield PK, Wood JW. 2006. Ovulation detection methods for urinary hormones: precision, daily and intermittent sampling and a combined hierarchical method. *Hum Reprod* 21:1442–1452.
- O'Connor KA, Brindle E, Shofer J, Trumble BC, Aranda JD, Rice K, Tatar M. 2011. The effects of a long-term psychosocial stress on reproductive indicators in the baboon. *Am J Phys Anthropol* 145:629–638.
- O'Connor KA, Holman DJ, Wood JW. 1998. Declining fecundity and ovarian ageing in natural fertility populations. *Maturitas* 30:127–136.
- O'Donnell E, Harvey PJ, Goodman JM, De Souza MJ. 2007. Long-term estrogen deficiency lowers regional blood flow, resting systolic blood pressure, and heart rate in exercising premenopausal women. *Am J Physiol Endocrinol Metab* 292:E1401–1409.
- Parazzini F, La Vecchia C, Negri E, Franceschi S, Tozzi L. 1993. Lifelong menstrual pattern and risk of breast cancer. *Oncology* 50:222–225.
- Parazzini F, La Vecchia C, Negri E, Gentile A. 1989. Menstrual factors and the risk of epithelial ovarian cancer. *J Clin Epidemiol* 42:443–448.
- Rickenlund A, Eriksson MJ, Schenck-Gustafsson K, Hirschberg AL. 2005. Amenorrhea in female athletes is associated with endothelial dysfunction and unfavorable lipid profile. *J Clin Endocrinol Metab* 90:1354–1359.
- Santoro N, Brown JR, Adel T, Skurnick JH. 1996. Characterization of reproductive hormonal dynamics in the perimenopause. *J Clin Endocrinol Metab* 81:1495–1501.

- Santoro N, Lasley B, McConnell D, Allsworth J, Crawford S, Gold EB, Finkelstein JS, Greendale GA, Kelsey J, Korenman S, Luborsky JL, Matthews K, Midgley R, Powell L, Sabatine J, Schocken M, Sowers MF, Weiss G. 2004. Body size and ethnicity are associated with menstrual cycle alterations in women in the early menopausal transition: the Study of Women's Health across the Nation (SWAN) Daily Hormone Study. *J Clin Endocrinol Metab* 89:2622-2631.
- Scheid JL, Toombs RJ, Ducher G, Gibbs JC, Williams NI, De Souza MJ. 2011. Estrogen and peptide YY are associated with bone mineral density in premenopausal exercising women. *Bone* 49:194-201.
- Terry KL, Willett WC, Rich-Edwards JW, Hunter DJ, Michels KB. 2005. Menstrual cycle characteristics and incidence of premenopausal breast cancer. *Cancer Epidemiol Biomarkers Prev* 14:1509-1513.
- Waller K, Swan SH, Windham GC, Fenster L, Elkin EP, Lasley BL. 1998. Use of urine biomarkers to evaluate menstrual function in healthy premenopausal women. *Am J Epidemiol* 147:1071-1080.
- West SL, Scheid JL, De Souza MJ. 2009. The effect of exercise and estrogen on osteoprotegerin in premenopausal women. *Bone* 44:137-144.
- Whelan EA, Sandler DP, Root JL, Smith KR, Weinberg CR. 1994. Menstrual cycle patterns and risk of breast cancer. *Am J Epidemiol* 140: 1081-1090.
- Wilcox AJ, Baird DD, Weinberg CR, Armstrong EG, Musey PI, Wehmann RE, Canfield RE. 1987. The use of biochemical assays in epidemiologic studies of reproduction. *Environ Health Perspect* 75:29-35.
- Williams NI, Reed JL, Leidy HJ, Legro RS, De Souza MJ. 2010. Estrogen and progesterone exposure is reduced in response to energy deficiency in women aged 25-40 years. *Hum Reprod* 25:2328-2339.
- Windham GC, Elkin E, Fenster L, Waller K, Anderson M, Mitchell PR, Lasley B, Swan SH. 2002. Ovarian hormones in premenopausal women: variation by demographic, reproductive and menstrual cycle characteristics. *Epidemiology* 13:675-684.
- Wright DM, Kesner JS, Schrader SM, Chin NW, Wells VE, Krieg EF Jr. 1992. Methods of monitoring menstrual function in field studies: attitudes of working women. *Reprod Toxicol* 6:401-409.
- Zeni Hoch A, Dempsey RL, Carrera GF, Wilson CR, Chen EH, Barnabei VM, Sandford PR, Ryan TA, Gutterman DD. 2003. Is there an association between athletic amenorrhea and endothelial cell dysfunction? *Med Sci Sports Exerc* 35:377-383.
- Ziomkiewicz A, Ellison PT, Lipson SF, Thune I, Jasienska G. 2008. Body fat, energy balance and estradiol levels: a study based on hormonal