Comparison of Specific Gravity and Creatinine for Normalizing Urinary Reproductive Hormone Concentrations

Rebecca C. Miller,1 Eleanor Brindle,1 Darryl J. Holman,1 Jane Shofer,1 Nancy A. Klein,2 Michael R. Soules,2 and Kathleen A. O’Connor1*

Background: Specific gravity (SG) may perform as well as creatinine (CR) correction for adjusting urinary hormone concentrations, as well as offer some advantages. We compared the two methods and applied them to US and Bangladeshi specimens to evaluate their use in different populations.

Methods: Pearson correlations between serum concentrations and SG, CR, and uncorrected urinary concentrations were compared using paired daily urine and serum specimens from one menstrual cycle from 30 US women. Corrected urinary estrone conjugate and pregnanediol glucuronide concentrations were compared with serum estradiol and progesterone. Urine specimens across one menstrual cycle from 13 Bangladeshi women were used to evaluate the applicability of both methods to a nonindustrialized population. Linear mixed-effects models were used to compare CR and SG values in the Bangladeshi vs US specimens.

Results: There was no significant difference between SG-corrected vs serum and CR-corrected vs serum correlations for either assay. Usable CR results were obtained for all US specimens, but 37% of the Bangladeshi specimens were below the CR assay limit of detection. The Bangladeshi sample had significantly lower CR and higher inter- and intrasubject CR variability than the US sample.

Conclusions: SG is a potentially useful alternative to CR correction for normalizing urinary steroid hormone concentrations, particularly in settings where CR values are highly variable or unusually low.

© 2004 American Association for Clinical Chemistry

Urine specimens have several advantages over serum, saliva, and blood spots for monitoring reproductive hormone patterns. Urine collection is noninvasive, poses minimal infectious disease risk to participants and researchers, and provides sufficient volume for multiple assays and future research. Urine specimens are ideally suited for large studies because they can be collected and stored by participants, and compliance is high (1–3). Urine also provides integrated hormone measures without the confounding effects of pulsatile secretion (4). Finally, urinary concentrations of reproductive steroid hormone metabolites are generally higher than the circulating serum concentrations, providing better quantification of the lower end of the physiologic scale in humans, which can be close to or below the limit of detection in serum RIAs (5). Urinary hormone results must, however, be adjusted to account for the concentration of the specimen, which depends on an individual’s hydration status and time since last urination.

We examined the performance of two methods of adjustment: specific gravity (SG)3 correction and creatinine (CR) correction. Urine SG is the ratio of the density of a urine specimen to the density of water (6). SG increases with solute concentration and is most accurately measured by refractometry (6). CR, a byproduct of muscle activity, is cleared from the bloodstream by the kidneys and excreted in urine (7). Urinary CR concentrations are determined by colorimetric assay (8), and analyte concen-
trations are usually reported as a ratio of the analyte concentration to CR concentration.

On the basis of an early finding that daily individual CR excretion was fairly consistent (9), urinary CR became a common method of assessing kidney function in clinical settings and correcting for analyte concentrations, including reproductive hormones, in urine (4). Use of CR to normalize urinary analyte concentrations can be problematic, however, because there is evidence that CR excretion is not consistent: numerous studies have found considerable inter- and intrasubject variability in CR values and dependence on sex, age, activity, and diet (10–18). Population variation in CR excretion may also exist, but to the best of our knowledge, this topic has not been investigated.

SG is an alternative method with several advantages over CR, although it is not widely used, perhaps in part because of a lack of data evaluating its performance relative to CR. In this study we compared SG and CR correction methods on urinary hormone metabolite concentrations from healthy US women, using serum hormone measurements as the standard. We then applied both methods to spot specimens from Bangladeshi women to evaluate their applicability in a nonindustrialized population.

Materials and Methods

Samples and specimens
A total of 799 daily urine and serum specimens were collected over one menstrual cycle from 30 US women in 1997–1998. Thirteen women 20–25 years of age and 17 women 40–45 years of age were recruited for a study on reproductive aging. Monetary compensation was provided, participants provided written informed consent, and all procedures were approved by the Institutional Review Board of the University of Washington. All participants were normally cycling, in good health, had a mean body mass (SD) index of 22.6 (2.36) kg/m² (range 18.9–27.7 kg/m²), and were not using medications or hormones. Blood specimens 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. Serum specimens were immediately assayed, and all cycles were confirmed ovulatory by transvaginal ultrasound. Urine specimens were taken daily in the clinic, usually before 1200, at the same time as serum collection and immediately stored at −20 °C. Urine specimens remained frozen until thawing 2 years later for aliquoting, assay, and measurement of SG (19). The specimens underwent two to three more freeze-thaw cycles before CR assay in 2003.

For the Bangladeshi sample, 13 cycling women were selected randomly from a sample of women participating in a 9-month study of early pregnancy loss. All participants were married, noncontracepting residents in the nonintervention demographic surveillance region of the rural Matlab district in Bangladesh (20). No monetary compensation was provided, all participants provided written informed consent, and all procedures were approved by the Institutional Review Boards of The Pennsylvania State University and the International Centre for Diarrhoeal Disease and Research, Bangladesh (20). Spot urine specimens were collected by community healthcare workers every 3 or 4 days over the course of one menstrual cycle in 1993. The specimens were stored at the healthcare workers' homes in a cooler with ice packs for 1–3 days until they were transported to a field hospital and stored at 4 °C. One to three days later, the specimens were brought to room temperature, and SGs were taken. The specimens were preserved with 0.17 g/mL boric acid solution, stored at −20 °C, and transported to the US by frozen air freight. Specimens remained frozen until 1995, when they were thawed and assayed for CR and steroid hormone metabolites.

Assays
All serum specimens were assayed by RIA for estradiol (E2) and by immunofluorometric assay for luteinizing hormone (LH), but serum progesterone (P4) was measured only in the luteal phase. The RIA for E2 (ICN Biomedicals) cross-reacts 20% with estrone, 1.5% with estriol, and <1% with all other steroids. The inter- and intraassay CVs were 16% and 7%, respectively. The inter- and intraassay CVs for the LH immunofluorometric assay (Delphia) were 2.8% and 4.7%, respectively. The RIA for P4 (Diagnostic Systems Laboratories) cross-reacts <5% with all other steroids, and the inter- and intraassay CVs were 13% and 11%, respectively.

Urinary specimens were analyzed by competitive enzyme immunoassays (EIAs) for urinary steroid hormone metabolites. All US and Bangladeshi urine specimens were assayed for pregnanediol glucuronide (PDG) and estrone conjugate (E1Cs). The PDG and E1C EIAs have been described elsewhere (19). Briefly, the PDG assay uses the monoclonal antibody Q330 (Quidel Corporation) and reference calibrator 5β-pregnane-3α,20α-diol glucuronide (Sigma; cat. no. P3635). The inter- and intraassay CVs for the PDG assay were 10.3% and 9.2%, respectively (19). E1Cs were measured with the 155-B3 monoclonal capture antibody (Dr. Fortune Kohen, Weizmann Institute, Rehovet, Israel) and estrone-β-p-glucuronide reference calibrator (Sigma; cat. no. E1752). The inter- and intraassay CVs were 10.9% and 7.3% (19). The absorbance for the EIAs was measured with a Dynatech MR7000 Plate Reader (test wavelength, 405 nm; reference wavelength, 570 nm). All specimens were run in duplicate on microtiter plates, and hormone concentrations were estimated from absorbance by use of a four-parameter logistic model in Biolinx 1.0 Software (Dynex Laboratories, Inc).

Urinary CR was measured by reaction with sodium hydroxide and picric acid in the method described by Jaffe (8) with calibrators purchased from Sigma (cat. no. 925-11). The absorbance of specimens was read with a
Dynatech MR7000 Plate Reader (test wavelength, 490 nm; reference wavelength, 630 nm), and CR concentrations were estimated with a four-parameter logistic model in Biolinx 1.0 Software. The inter- and intraassay CVs for the CR assay were 1.6% and 14%, respectively.

The CR correction formula applied to each sample was as follows:

\[
\text{CR-corrected concentration (nmol/nmol)} = \frac{\text{raw hormone concentration (nmol/volume)}}{\text{CR (nmol/volume)}}
\]

SG measurements were taken with a hand-held urine SG refractometer (Atago Uicon-PN; NSA Precision Cells, Inc). The SG of Bangladeshi urine specimens was measured before freezing, whereas the SG for the US specimens was measured after the first freeze-thaw cycle. The correction formula applied to each hormone result was as follows:

\[
\text{SG-corrected concentration} = \frac{(\text{SG}_{\text{target}} - 1.0)}{(\text{SG}_{\text{sample}} - 1.0)} \times \text{raw hormone concentration (nmol/volume)}
\]

where \(\text{SG}_{\text{target}}\) is a population mean SG (20). The target SG used was 1.020 (21) for US specimens and 1.015 (20) for Bangladeshi specimens.

**Statistical Analyses**

Pearson correlations between serum hormone values and raw, SG-corrected, and CR-corrected urinary hormone results were calculated with an optimum lag day. Individual cycles were aligned by day of midcycle serum LH peak (day 0), and a mean hormone value for each cycle day was calculated (\(n = 34\) paired urine/serum cycle days from 30 cycles). The effect of time between appearance of hormones in serum and clearance into the urine was evaluated by correlating serum results with urine results lagged by \(-1, 0, 1, 2,\) and 3 days behind serum. The optimum lag day was selected on the basis of the highest correlation with serum results.

The correlations of serum E2 with urinary E1Cs and serum P4 with urinary PDG were weighted by the number of specimens available for each cycle day. The 95% confidence intervals for the difference between SG-corrected urine vs serum, CR-corrected urine vs serum, and uncorrected urine vs serum mean correlations were obtained by use of bootstraps. The hypothesis that two correlations were the same was rejected when the confidence interval did not include zero.

We used a linear mixed-effects model to test for differences in mean CR or SG by population (US vs Bangladesh). Mixed-effects models allow the significance of a fixed effect (in this case, the mean CR or SG difference between the US and Bangladesh populations) to be tested while accounting for within-subject repeated measures by modeling a random effect for each individual as follows:

\[
y = a + bx + c_i + e_{ij}
\]

where \(y\) is either a CR or SG value; \(x\) is 0 for US and 1 for Bangladesh; \(a\) is the estimated CR or SG for the US sample; \(b\) is the estimated difference between CR or SG between Bangladesh and the US; \(c_i\) is a random effect estimated for the \(i\)th individual; and \(e_{ij}\) is the residual error associated with the \(j\)th specimen from the \(i\)th individual. We also used linear mixed-effects models to compare between- and within-subject (residual) variability for each population in separate models by estimating the SD for the \(c_i\) term, which represents between-subject variability, and the \(e_{ij}\) term, the within-subject variability.

**Results**

For both assays, the highest correlation between urine and serum values occurred on the same lag day for uncorrected, SG-corrected, and CR-corrected values (Table 1). PDG and E1C results lagged 1 day behind their respective serum values. When optimum lag day was used, weighted correlations between serum and SG-corrected urinary hormone values were high for both assays, ranging from 0.94 to 0.97 (Table 2). Correlations were also high for CR-corrected (0.93–0.98) and uncorrected urine results (0.92–0.93; Table 2). There was no significant difference in the correlations between SG-corrected urine vs serum values and CR-corrected urine vs serum for either hor-

| Table 1. Pearson correlations between serum and lagged urinary hormone values in US women.\(^a\) |
|-----------------------------------------------|----------------|----------------|----------------|
| **Urine concentrations**                       | **Uncorrected** | **CR-corrected** | **SG-corrected** |
| Lag                                            | E1C vs E2      | PDG vs P4       | E1C vs E2      | PDG vs P4       |
| Urine 1 day before serum                       | 0.47          | 0.79            | 0.45          | 0.85            | 0.51          | 0.83            |
| None                                           | 0.72          | 0.91            | 0.45          | 0.85            | 0.76          | 0.95            |
| Urine 1 day after serum                        | 0.91          | 0.95            | 0.91          | 0.98            | 0.93          | 0.98            |
| Urine 2 days after serum                       | 0.84          | 0.91            | 0.80          | 0.92            | 0.84          | 0.93            |
| Urine 3 days after serum                       | 0.61          | 0.57            | 0.61          | 0.53            | 0.61          | 0.55            |

\(^a\) For E1C vs E2, \(n = 34\) mean paired urine and serum cycle days from 30 US women; for PDG vs P4, \(n = 17\) mean paired urine and serum cycle days from 30 US women.

| Table 2. Weighted correlations between serum and urinary hormone values in US women.\(^a\) |
|-----------------------------------------------|----------------|----------------|----------------|
| **Urine concentration**                       | **Uncorrected** | **SG-corrected** | **CR-corrected** |
| E1C vs E2                                     | 0.92          | 0.93            | 0.94            |
| PDG vs P4                                     | 0.92          | 0.98            | 0.97            |

\(^a\) Correlations incorporate a 1-day lag of urine behind serum. For E1C vs E2, \(n = 34\) mean paired urine and serum cycle days from 30 US women; for PDG vs P4, \(n = 17\) mean paired urine and serum cycle days from 30 US women.
Table 3. Mean differences (95% confidence intervals) between correlations of serum with urinary hormone values, corrected by CR or SG and without correction, from US specimens.a

<table>
<thead>
<tr>
<th></th>
<th>CR-corrected vs uncorrected</th>
<th>SG-corrected vs uncorrected</th>
<th>SG-corrected vs CR-corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1C vs E2</td>
<td>0.013 (−0.032 to 0.079)</td>
<td>0.019 (−0.007 to 0.058)</td>
<td>0.006 (−0.024 to 0.043)</td>
</tr>
<tr>
<td>PDG vs P4</td>
<td>0.063 (0.011–0.197)</td>
<td>0.056 (0.017–0.179)</td>
<td>−0.007 (−0.036 to 0.015)</td>
</tr>
</tbody>
</table>

a Correlations incorporate a 1-day lag of urine behind serum. For E1C vs E2, n = 34 mean paired urine and serum cycle days from 30 US women; for PDG vs P4, n = 17 mean paired urine and serum cycle days from 30 US women.

SG and CR normalization of urinary analyte concentrations (22, 23). We demonstrated that SG adjustment is applicable in specimens with CR concentrations that are not sufficiently high to correct metabolite concentrations. We have also provided evidence of a population concentration difference in urinary CR values and concurrent population concentration similarities in SG measurements. Currently, CR correction is widely used to adjust for urine analyte concentration in spot and first-morning specimens, although several studies have concluded that it offers no advantage over unadjusted results (24–27). Our results did indicate a small but significant improvement in PDG correlation with serum P4 values when adjusted by CR and SG. Although CR and SG perform similarly, SG offers some advantages, particularly for large-scale research on reproductive hormones.

Discussion

Our results show that SG performs as well as CR in correcting steroid hormone metabolite results for urine concentration, in agreement with other studies comparing SG and CR normalization of urinary analyte concentrations (22, 23). We demonstrated that SG adjustment is applicable in specimens with CR concentrations that are not sufficiently high to correct metabolite concentrations. We have also provided evidence of a population concentration difference in urinary CR values and concurrent population concentration similarities in SG measurements. Currently, CR correction is widely used to adjust for urine analyte concentration in spot and first-morning specimens, although several studies have concluded that it offers no advantage over unadjusted results (24–27). Our results did indicate a small but significant improvement in PDG correlation with serum P4 values when adjusted by CR and SG. Although CR and SG perform similarly, SG offers some advantages, particularly for large-scale research on reproductive hormones.

SG is inexpensive; a handheld refractometer and transfer pipettes are the only equipment needed, whereas the CR assay is a microtiter plate-based assay. The CR assay uses picric acid, a harsh physical irritant that is explosive when dry, so special use and storage conditions are necessary. Finally, although multiple samples can be assayed simultaneously for CR, it does not offer a time advantage over SG. An equal number of samples can be measured on the refractometer in the time needed to complete the CR assay.

SG refractometers measure urine density via the ratio of light refraction between air and a urine specimen; refraction increases with solute concentration of the specimen (28). Urine density, however, varies with the total mass of solutes, which depends not only on the number of particles present, but also on their molecular weight. Therefore, SG is affected more by the presence of heavy molecules such as glucose, albumin, phosphates, sulfates, radiocontrast media, and heavy metals than it is by low-molecular-weight substances such as sodium, chloride, and urea (6, 29, 30). Thus, SG correction may not be an appropriate method for individuals with diabetes mellitus and nephrotic syndrome, which cause high concentrations of glucose and protein in urine, increasing SG and underestimating urine analyte concentrations (30). Diabetes, starvation, and dehydration produce ketones from fat metabolism. The ketones are excreted in urine and erroneously lower SG readings because they are less dense than water (28). Several authors have also indicated that urine solute concentrations given by SG correction may be inaccurate if the urine specimen is very dilute or low in solute content.
very concentrated (31, 32). However, a minimum acceptable SG of 1.010 (16, 33) seems too high; 31% (247 of 799) of our specimens from healthy US women had SG values lower than this, and our averaged SG-corrected urine vs serum correlations were very high. SG readings are also affected by temperature fluctuations, which make specimens expand or contract, altering their density (16); therefore, readings should be taken at a consistent temperature.

CR correction of urinary hormone values is problematic primarily because CR excretion exhibits inter- and intrasubject variations and is influenced by time of day, age, sex, diet, body mass, and activity level. This variation could complicate interpretations of analytes reported in

Fig. 1. Distribution of CR values from one menstrual cycle for 30 US (△) and 13 Bangladeshi (●; ○) participants. Each line shows all CR measurements for one participant from one menstrual cycle with CR concentration plotted on the x axis.
ratio to CR. Because CR is a byproduct of muscle use, its production is expected to vary with body composition and activity. Edwards and Whyte (34) reported a correlation between lean body mass and urinary CR of 0.65 from a group of 31 men and women. Bleiler and Schedl (35) found the same measure to be 0.47–0.48 in 11 women and 0.53–0.55 in 51 men and reported correlations between urinary CR and weight and body surface area in 24-h specimens. Muscularity also contributes to observed sex differences in CR concentrations. Men produce more CR than women and have a higher clearance rate (17). Kesteloot and Joossens (18) reported a mean CR clearance of \( \frac{1101}{H_1} \) mL/min in 2075 men and 86.9 mL/min in 1933 women. This and other studies (36, 37) also showed a decrease in CR clearance with age, leading to increased serum and decreased urinary concentrations at older ages, making CR correction particularly questionable for adjusting urinary hormone values in research concerning aging, such as the transition to menopause.

Diets with substantial amounts of particular kinds of meat, such as beef, can also affect urinary CR concentrations. Meat contains creatine, the precursor of CR, which can accumulate in the body and lead to a gradual increase in CR excretion (7). Cooking meat converts creatine to CR, which is quickly excreted and can cause considerable short-term CR increases in the hours after ingestion (10, 15).

Population differences in activity, nutritional status, and body composition may account for the differences observed between the Bangladeshi and US samples, and they challenge the use of CR as a correction method for this population. Most women in Bangladesh suffer from chronic undernutrition and infectious disease and have limited access to healthcare (38). They also have very low body mass indexes: the mean (SD) body mass index for a large random sample of nonpregnant women 15–45 years of age in 1992 in Matlab, Bangladesh was 18.8 (1.9) kg/m\(^2\) (38). Thus, the low and variable CR concentrations in our Bangladeshi samples may reflect these factors. It should be noted that the effects of differences in specimen treatment conditions between the Bangladeshi and US samples on detected CR was not tested and cannot be ruled out. However, several studies have shown that CR is very stable in refrigerated and frozen samples (39, 40), and we did not find any effect of specimen treatment conditions on urinary concentrations of E1Cs and PDG in a test of a range of treatment conditions (19).

Additionally, although the original use of CR was to check the completeness of 24-h urine specimens, many studies have shown considerable intraindividual variation in daily CR excretion. Alessio et al. (16), using four consecutive 24-h urine collections, observed variation ranging between 9.2% and 79.4% in the extreme values of
16 individuals. Similarly, Greenblatt et al. (14) found a range of 63–244% in 24-h collections from eight individuals. In a summary of studies on 24-h CR excretion, Curtis and Fogel (12) showed that some individuals had relatively consistent daily CR excretion with individual CV around 5%, whereas others had highly variable excretion with individual CV exceeding 20%.

CR correction is particularly questionable when applied to spot samples because CR excretion over short intervals also shows considerable variation. Verstergaard and Leverett (11) showed that subsequent 2-h interval samples varied by >100%, and several studies have reported that spot-sample CR variation is several times higher than variation for 24-h values (12, 13). The inap-

Fig. 4. Distribution of SG values from one menstrual cycle for 30 US (△) and 13 Bangladeshi (B; ○) participants. Each line shows all SG measurements for one participant from one menstrual cycle with SG plotted on the x axis.
plicability of CR correction to spot specimens presents a major problem for large prospective studies on reproductive hormones. Frequent collections are needed to accurately characterize hormonal patterns, and daily 24-h urine collections are impractical, difficult to obtain, and would present storage and shipping problems.

In conclusion, this study supports the use of SG as an alternative to CR for urinary hormone concentration correction. Both methods show high correlation with serum hormone values. However, SG correction offers several practical advantages over CR. Given these advantages, we recommend SG as an alternative to CR for adjusting urinary steroid hormone metabolite concentrations, particularly in populations with very low or highly variable urinary CR concentrations.

This research was supported by the NSF (DBS-9218734 and DBS-9600690), NIA R01 AG15141, NIA R01 AG14579, the Mellon Foundation, the Hill Foundation, and the American Institute for Bangladesh Studies.

References

33. Trevisan A. Concentration adjustment of spot samples in analysis
34. Edwards OM, Whyte HM. Creatinine excretion and body composi-
35. Bleiler RE, Schedl HP. Creatinine excretion: variability and rela-
36. Carrieri M, Trevisan A, Battista Bartolucci G. Adjustment to
concentration-dilution of spot urine samples: correlation between
specific gravity and creatinine. Int Arch Occup Environ Health
37. Hall Moran V, Leathard HL, Coley J. Urinary hormone levels during
the natural menstrual cycle: the effect of age. J Endocrinol
2001;170:157–64.
38. Holman DJ, O’Connor KA. Bangladeshis. In: Ember M, Ember C,
eds. Encyclopedia of medical anthropology: health and illness in
albumin, n-acetylglucosaminidase, and creatinine in urine sam-
40. Riboli E, Haley NJ, De Waard F, Saracci R. Validity of urinary
biomarkers of exposure to tobacco smoke following prolonged