Effects of Fluid Load on Human Urine Characteristics Related to Workplace Drug Testing*

Arvind K. Chaturvedi1,3, Jim L. Sershon1, Kristi J. Craft1, Patrick S. Cardona1, John W. Soper1, Dennis V. Canfield1, Kurt M. Dubowski1, James E. Whinnery1, Misti J. Leyva2, Christopher E. Aston2, Steve M. Blevins2, Jonelle E. Wright2,5, Albert D. Fraser3 and David J. Kuntz1,6

1Bioaeronautical Sciences Research Laboratory (AAM-610), Civil Aerospace Medical Institute, Federal Aviation Administration, U.S. Department of Transportation, P. O. Box 25082, Oklahoma City, Oklahoma 73125, 2The University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73117, 3Department of Laboratory Medicine, Saint John Regional Hospital, Saint John, New Brunswick E2E 5J6, Canada, 4Northwest Toxicology, Salt Lake City, Utah 84124, 5Leonard M. Miller School of Medicine, University of Miami, Miami, Florida 33136, and 6Clinical Reference Laboratory, 8433 Quivira Road, Lenexa, Kansas 66215

During workplace drug testing, urine is tested for dilution, substitution and adulteration. Donors argue that these findings are due to medical, health or working conditions or diet and genetic differences. There is a paucity of data correlating changes in urine characteristics after a fluid load to various body parameters. Therefore, five urine specimens (one in the morning, one prior to drinking 800 mL of a beverage, and three time intervals thereafter) from 12 males and 12 females were tested for four different beverages on separate occasions. Of the 480 samples, 376 were in sufficient amounts. Of these 376, 36 (10%) had creatinine <20 mg/dL but ≥2 mg/dL; 27 (75%) of 36 had specific gravity <1.0030 but >1.0010. Thus, these 27 samples can be considered to be dilute; 20 (74%) of 27 were from females. For males with at least one dilute sample, body fat was 11% less and resting metabolic rate (RMR) was 29% more than males with no dilute samples (p > 0.05); for females with at least one dilute sample, height was 8% less and weight 20% less than females with no dilute samples (p > 0.05). Individuals with a higher RMR appear to have a greater potential for producing dilute urine specimens than those with a lower RMR. Thus, a dilute sample does not necessarily indicate that it was intentionally diluted. Such samples must be carefully evaluated in consideration with recent consumption of liquid by donors to avoid false accusations.

Introduction

Under the United States Public Law 102-143 (1) and the Executive Order 12564 (2), individuals working in safety-related positions are subject to random drug testing. According to the procedures for transportation workplace drug testing (3, 4), randomly collected urine samples are tested for commonly abused drugs. However, drug users often find ways to defeat the intent of the testing by altering (diluting, substituting or adulterating) urine specimens. To address this issue, urine validity testing was designed to identify samples that were altered prior to drug testing (3, 4).

A urine sample is defined as dilute when its creatinine concentration is <20 mg/dL but ≥2 mg/dL and specific gravity is ≤1.0010 or ≥1.0200. A urine specimen is considered adulterated when certain substances are present in the sample at elevated levels. Commonly used adulterants are commercial products containing nitrite, chromium VI, iodide, peroxide, bleach, glutaraldehyde, soap and acids (5). The majority of adulterants are oxidants, and they act by destroying target analytes. Other adulterants, such as glutaraldehyde, soap and acids, may interfere with screening assays by destroying enzymes and other proteins, affecting antibody-antigen complexes, and/or altering pH. Drug analytical results of a urine sample with pH values in the range >3.0 and <4.5 or >9.0 and <11.0 will be reported as invalid (3, 6). Urine within either of these pH ranges would be considered non-physiological, because a normal physiological urinary pH range is 4.5 to 8.0 (7).

Some adulterants, such as nitrite, chromium and iodide, are endogenously present in urine at low concentrations. Consequently, sample donors accused of adulterating their samples argue that those substances are present in elevated concentrations due to a variety of factors, including the donor’s health, working conditions, dietary habits, or differences in body weight, gender, race and ethnicity. However, whether or not these factors have an effect in causing urine specimens to contain higher concentrations of endogenous substances has yet to be established.

The most controversial aspect of validity testing lies in the possibility that a person may have urine with low creatinine and/or specific gravity because of reasons—dietary habits, ethnic origin, health-related treatments, working conditions and/or genetic differences—other than an intentional dilution, substitution or adulteration of the urine sample. To address this issue, the 107th U.S. Congress asked the Department of Transportation (DOT) to conduct a study on drug and alcohol validity testing (8). Subsequently, the DOT Federal Aviation Administration (FAA) held a colloquium to address this issue during February 4–6, 2003 (9). Based upon the colloquium, several recommendations, including the lowering of creatinine value to <2 mg/dL for a sample to be considered substituted, were proposed for revising rules and regulations related to the validity testing (3, 4).

Criteria for urine specimen validity have been summarized in a review article published in 2000 (10). Ingestion of excessive amounts of water (3,785 L; 4 × 946 mL aliquots; 1 × 946 mL aliquot consumed each hour for 4 h) by human subjects has been reported to produce dilute urine samples based upon the
criteria of <20 mg/dL creatinine and <1.003 specific gravity (11). A fluid ingestion study (643–6059 mL consumption of water, beverages and soup over a 12 h period) with 14 volunteers revealed erratic changes in the urinary creatinine concentration with specific gravity of ≥1.003 (12). In a human study involving ingestion of at least 2.37 L of fluid over a 6 h period, none of the urine samples satisfied the substitution criteria that existed at the time (13). The substitution criteria were creatinine ≤5.0 mg/dL and specific gravity ≤1.001 or ≥1.020. However, 7.5% of the participants produced at least one specimen that could be considered to be dilute based upon creatinine <20.0 mg/dL and specific gravity <1.003 (13). Although some studies have been published addressing specimen validity related concerns, well-designed, focused scientific studies on this issue in conjunction with body composition parameters are lacking. In the present study, effects of the consumption of water and three types of commonly used commercial soft drinks by human participants were evaluated with regard to their urine characteristics associated with workplace drug testing. Attempts were also made to define a relationship between changes in such characteristics with human body composition parameters.

Materials and Methods

Human subjects

A total of 24 healthy subjects (12 males and 12 females) between the ages of 25 and 35 years participated in this study. Human subject participation was approved by the Institutional Review Boards of the University of Oklahoma Health Sciences Center (OUHSC) (Oklahoma City, OK), and of the FAA's Civil Aerospace Medical Institute (CAMI) (Oklahoma City, OK). These subjects were not on medications, with the exception of oral contraceptives and allergy remedies. However, the participants were asked not to take anti-allergy drugs on the days of the experiment.

Participant visits occurred between August 2003 and March 2006. Each subject visited the clinic four times for the experiments. The time spans over which these four visits occurred ranged from four to 43 days, with three extraordinary time spans of 98, 185 and 253 days.

Upon arrival at the OUHSC General Clinical Research Center (GCRC) for the initial visit, consent was obtained from the participants. Their body composition parameters were measured, which took an average of 1.5 h. Subsequently, subjects proceeded with liquid consumption and urine sampling.

Body composition parameters

Height, percent body fat and resting metabolic rate (RMR) of the subjects were determined on the day of their first visit, but body weight was measured on each of the four visits (Table I). This selective measurement approach was taken in the view that height, percent body fat and RMR would not significantly change within the scope of the study. Additionally, subjects would also have to go through less invasiveness during the experiments. The body weight determination on every visit would suggest any drastic change in the weight, and thereby maybe in other applicable parameter(s). The difference in weight of the subject who had almost a full year (253 days) between the visits was 4.9 Kg.

Percent body fat was measured by the dual-energy X-ray absorptiometry technique (HOLOGIC QDR 4500A, Delphi, Bedford, MA). RMR was measured by indirect calorimetry (MedGraphics, Cardiorespiratory Diagnostic Systems, Minneapolis, MN) after having the participant lie still on a bed for 30 min. Subjects were asked in advance to refrain from exercise in the 12 h period before the visit and to avoid strenuous activity when coming to GCRC.

Urine samples

Participants arrived at GCRC at approximately 8:00 A.M. nil per os (NPO; nothing by mouth, by self-report) for the previous 12 h. All participants brought their first morning void urine samples (average collection time: 7:00 A.M.) in urine sample collectors (Specimen Collector Commode, Medegen Inc., Ontario, CA) supplied by GCRC. For the morning sample (Sample 1), participants were instructed to place the collector’s cap under the lid of the toilet and collect all of the first morning urine specimen, record the volume, and then pour approximately 100 mL of the urine through the collector’s spout to a smaller container (Kendall Precision Mid-Stream Urine Collector Kit, Tyco 2001 Healthcare Group LP, Mansfield, MA), which was also provided by GCRC. Subsequent samples were similarly collected at GCRC. No preservatives were used in collecting the samples.

After the first morning void sample (Sample 1), urine samples were collected immediately before drinking the liquid (Sample 2), immediately after drinking the liquid (Sample 3), at predicted stomach clearance (Sample 4) (14) and at first urge to void (Sample 5), totaling five time-point samples during each visit. Following the collection of samples, volumes were measured. Specific gravity was determined by using a DiaScreen...

### Table I

<table>
<thead>
<tr>
<th>Participants (25–35 years)</th>
<th>Height (m)</th>
<th>Weight (kg)$^1$</th>
<th>Body mass index (kg/m$^2$)</th>
<th>Body fat$^3$ (%)</th>
<th>RMR$^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male ($n = 12$)</td>
<td>1.84 ± 0.05 (1.77–1.93)</td>
<td>86.5 ± 11.2 (74.0–106.8)</td>
<td>25.6 ± 2.3 (23.1–30.2)</td>
<td>19.7 ± 5.1 (8.4–27.6)</td>
<td>1404 ± 390 (904–2425)</td>
</tr>
<tr>
<td>Female ($n = 12$)</td>
<td>1.63 ± 0.07 (1.52–1.78)</td>
<td>63.9 ± 14.2 (44.1–84.0)</td>
<td>23.8 ± 4.7 (17.6–33.9)</td>
<td>30.5 ± 7.1 (21.5–41.9)</td>
<td>1085 ± 257 (672–1604)</td>
</tr>
<tr>
<td>Male and female** ($n = 24$)</td>
<td>1.73 ± 0.12 (1.52–1.93)</td>
<td>75.2 ± 16.9 (44.1–106.8)</td>
<td>24.7 ± 3.8 (17.6–33.9)</td>
<td>25.1 ± 8.2 (8.4–41.9)</td>
<td>1244 ± 359 (672–2425)</td>
</tr>
</tbody>
</table>

$^1$Note: Values are mean ± SD. Numbers in parentheses adjacent to the mean ± SD values represent the range.

$^2$Body weight of each subject was measured on each of the four visits. The average of these four determinations was calculated, giving a body weight value for each subject. The calculated values ($n = 12$ for male or female; $n = 24$ for male and female combined) were then used in the determination of the mean values.

$^3$Determined by the dual-energy X-ray absorptiometry technique.

$^4$Kilocalories expended in a 24h time period.

**Combined body composition parameter values of both male and female subjects.
Effects of Fluid Load on Human Urine Characteristics Related to Workplace Drug Testing

50 Urine Chemistry Analyzer (Hypoguard, Minneapolis, MN). This instrument provided the readings of the specific gravity values only up to the third decimal place with increments of 0.005 in the range of 1.010 to 1.025; any value <1.010 was reported as ≤ 1.005, and >1.025 as ≥ 1.030.

Initially, all urine samples collected in the small containers were stored at GCRC at 2–4°C. These urine samples were subsequently hand-delivered to CAMI with frozen gel bags in an insulated plastic box. All samples received at CAMI were stored at −20°C until they were sent for further analyses. The desired volume of each sample for additional analyses was 50 mL; however, that was not the case with all acquired samples, particularly Sample 3. Where available, 10–15 mL of the urine samples in screw-capped 20 mL glass culture tubes (Fisher Scientific, Pittsburgh, PA) were shipped to Northwest Toxicology (Salt Lake City, UT)—a U.S. Substance Abuse Mental Health Service Administration (SAMHSA) accredited laboratory—for analyses. These culture tubes were also free from preservatives. Samples were shipped with frozen gel bags in an insulated box by an air courier service for next-day delivery.

Included analytical parameters were associated with dilution, substitution and adulteration of urine. The analytical panel included measurements of creatinine, specific gravity (if creatinine level <20 mg/dL), pH, oxidants, glutaraldehyde and soap (4, 6). The methods used for these analyses were proprietary analytical methods of the accredited laboratory, but were consistent with the SAMHSA accreditation guidelines. Specific gravity in the samples with <20 mg/dL creatinine was measured up to the fourth decimal place, which is consistent with the current workplace drug testing guidelines (4).

Liquids consumed

On each visit, the participants were instructed to drink 800 mL of their assigned liquid within 5 min; the consumption was timed with a stopwatch for accuracy. The liquid types were water, obtained from a regular city drinking water supply tap and filtered (Brita Products Company, Oakland, CA) (Beverage A); a fully (3.2 volumes) carbonated, carbohydrate-rich beverage (100 g/L carbohydrate; 706 mOsm/L), Lemon-Lime Shasta (National Beverage Corporation, Ft. Lauderdale, FL) (Beverage B); a non-carbonated, electrolyte-rich beverage (60 g/L carbohydrate; 320 mOsm/L), Gatorade (Quaker Oats Company, PepsiCo, Chicago, IL) (Beverage C); and a lightly (1.15 volumes) carbonated beverage (83 g/L carbohydrate, 542 mOsm/L), Fanta Orange (Coca Cola Company, Atlanta, GA) (Beverage D). For a carbonated drink, one volume is equivalent to 1 L of carbon dioxide dissolved in 1 L of the drink at standard temperature and pressure (15). The stomach clearance times for these liquids, at which urine Sample 4 was collected, were calculated as 21, 107, 31 and 47 min, respectively (14). The participants were given tolerably cold beverages to comfortably complete drinking in 5 min. No ice pieces were added to the beverages.

Participants were randomized to begin with Beverages A, B, C or D at the first visit, and then followed the order of Beverages A, B, C and D at subsequent visits. In other words, the subjects who started with Beverage C had the sequence of the drinks C, D, A and B on consecutive visits. Six subjects began drinking experiments with A, four with B, eight with C, and six with D.

Statistics

Descriptive statistics are presented as mean ± standard deviation (SD). Statistical significance was determined using Student’s t-test for continuous measures (age, height, weight, body mass index, percent body fat and RMR) or chi-squared test for discrete measures (gender). Statistical analyses were performed with results on combined males and females and on separate males and females. Calculations were performed by using Microsoft Office Excel 2003 (Redmond, WA) or a Texas Instruments TI-60 Advanced Scientific Calculator (Texas Instruments Professional TI-60 Guide Book 1986, Lubbock, TX).

Results and Discussion

Twenty-four subjects of diverse ethnic backgrounds participated in the study: two African Americans, two Asians, 14 Caucasians, two Hispanics, and four Native Americans. Average values of various body composition parameters are given in Table I. The values of the parameters were consistent with those found in the general population: Percent body fat was greater in females than in males, and metabolic rate was higher in males than in females.

The number of collected urine samples with sufficient amounts for measuring volume and initial specific gravity was 469 of the expected 480 (24 subjects × four drink types × five time points) samples. There were 10 instances in which the sample volume was zero because urine could not be voided by the participant; at one instance, no successful analysis could be performed due to insufficient sample volume. These 11 samples were associated with the sample collection just after drinking the 800 mL liquid—that is, with Sample 3. As mentioned previously, average time gap was 1.5 h between the arrival of volunteers with Sample 1 and the collection of Sample 2. The time allowed for liquid consumption was 5 min and it took an additional 5 min to make arrangements for the collection of Sample 3. Therefore, the time gap between the collection before drinking the liquid (Sample 2) and after drinking the liquid (Sample 3) was short (10 min). This short time restriction also resulted in the collection of relatively smaller volumes for Sample 3. Typically, they were approximately 7% of the morning void samples (Table II). Thereafter, the volumes of Samples 4 and 5 increased to 36 and 46%, respectively, relative to the volumes of the morning void samples. The volumes for Sample 4 were dependent upon the clearance time of the respective liquids, that is, at 21, 107, 31 and 47 min (14). For example, the sample volume was less (approximately 9% of the morning void sample volume) with water (Beverage A), which had a reported stomach clearance of 21 min, and the sample volume was more (approximately 87%) with a fully carbonated, carbohydrate-rich beverage (Lemon-Lime Shasta, Beverage B), which had a reported stomach clearance of 107 min (14). At the time of the collection of Samples 4 and 5, subjects neither volunteered nor were asked if they had an urge to void before the stomach clearance. Therefore, it was not established whether any of the samples (Samples 4 and 5) might have been switched. The average time difference between Samples 4 and
5 for all subjects was 46 ± 1.3 min (n = 96; range: 15–115)—the difference for males was 48 ± 1.4 min (n = 48; range: 15–101) and for females 44 ± 1.2 min (n = 48; 17–115) with p = 0.213.

Not all urine samples were collected in sufficient amounts for the analyses of parameters associated with dilution, substitution and adulteration. Due to the limited volumes of collected samples, only 376 of the 469 samples were further analyzed for those parameters. Of the 376 samples, 199 were from male subjects and 177 from female subjects. None of these 376 samples was positive for oxidants (nitrite, chromium VI, iodine, peroxide and bleach) or for glutaraldehyde and soap (6). The absence of any adulterant was consistent with the absence of any substances that could adulterate their urine samples.

In accordance with the Department of Health and Human Services (DHHS) guidelines, findings from the drug analysis of a urine specimen with pH values of ≥3.0 and <4.5 or ≥9.0 and <11.0 could be reported as invalid results (3, 4, 6). In the present study, the pH values of the 376 samples were determined to be acceptable, ranging from 4.5 to 8.7, with the exception of a morning void sample that had a pH value of 9.1 and creatinine concentration of 49 mg/dL. This sample, in accordance with the DHHS guidelines, could have been called a specimen with invalid results (6). However, the declaration of the drug analysis results from a urine sample with pH ≥ 9.0 to be invalid must be based upon the final decision of the Medical Review Officer, after evaluating all possible medical scenarios of the urine sample donor. With regard to this high pH value, the sample was from the morning void urine of a female volunteer, was refrigerated or frozen prior to the pH determination, and there was no apparent motivation for the subject to alter pH. The observed high pH of the sample might be an isolated incident, particularly when other samples collected from the same volunteer later in the day and with other fluid load experiments were within the normal pH range. It is notable that none of the urine samples collected after the fluid load was outside of the valid sample pH range.

Based upon the DHHS guidelines (3, 4, 6), none of the 376 analyzed samples fell under the category of substituted, because their creatinine concentrations were not <2 mg/dL. However, 36 (10%) of 376 samples had creatinine concentrations <20 but ≥2 mg/dL, and 27 of these low creatinine samples had specific gravity <1.0030 but ≥1.0010 (Table III). The remaining nine samples had specific gravity ≥1.0030. Thus, 27 of the 376 urine samples could be considered dilute per DHHS guidelines. All 27 dilute samples were collected after the fluid load at the respective stomach clearance of the liquid (Sample 4) and/or at first urge to void thereafter (Sample 5). Twenty-one (78%) of the 27 dilute samples were collected at the time of urge (Sample 5). Interestingly, 20 (74%) of the 27 dilute samples were collected from female participants. Furthermore, six (50%) of 12 male subjects and eight (67%) of 12 female subjects produced at least one dilute specimen (Table III). These specific observations suggest that females, in comparison to males, have a higher potential to produce urine that could fall in the category of dilute. Females also have a lower muscle mass and lower creatinine values in general; thus, this is likely an additional reason why more females produced dilute urine.

The time interval between the collection of the urine samples and their analyses for creatinine determination was 12–494 days. Of the 96 experiments (24 subjects and four liquids), there were 33 experiments wherein the collection and analysis time interval was >100 days. There were 26 experiments in which the urine creatinine concentration in at least one sample was <20 but ≥2 mg/dL and specific gravity <1.0030 but >1.0010. In nine of these 26 experiments, the sample collection and creatinine analysis interval was >100 days. All collected urine samples were stored initially at 2–4°C and subsequently at −20°C prior to the analysis and were shipped for analyses with frozen gel bags in an insulated box by an air courier service for next-day delivery. Therefore, the effect of the time interval between the collection and analysis on the concentrations of creatinine was minimal, because this bio-marker has been reported to be essentially stable for at least four weeks if urine is refrigerated (16) and for at least 2 years if it is frozen (17, 18).

DOT urine specimen collection guidelines do not specifically indicate the use of any chemical preservatives in sample collection containers (19). This approach is possibly taken (i) to
prevent any chemical reaction between the preservatives and urine components, including possible drugs present in the collected urine; (ii) to prevent preservative-caused interference with the analytical methods themselves; and (iii) to diminish any possible argument that those preservatives caused false drug positive results.

The urine samples collected under the present study were stored at 2–4 and at –20°C to minimize any possible degradation of those samples with time. When the initial three-decimal specific gravity values were compared with those of the respective four-decimal specific gravity values of the samples with < 20 mg/dL creatinine, it was concluded that those values were consistent with each other in a broader sense. For example, 30 samples with specific gravity ≤ 1.005 were found to be less than 1.0050 when specific gravity was measured later after the creatinine measurement; two specimens with the initial specific gravity value of 1.010 were determined to be 1.0100; however, the final specific gravity values of the remaining four samples with the initial specific gravity value of 1.010 were 1.0016, 1.0025, 1.0025 and 1.0026 with the respective creatinine analysis intervals of 37, 352, 108 and 110 days. The initial specific gravity measurement was not as sensitive as the later measurement, because the former was determined in the increments of 0.005 up to only the third decimal place in the range of 1.010 to 1.025; thus, any value < 1.010 was reported as ≤ 1.005, and > 1.025 as ≥ 1.030. All of these observations suggested that the urine samples were fairly well preserved and that their specific gravity was not significantly affected by storing the samples for the periods within the scope of the study. Because none of the 376 specimens had creatinine concentration < 2 mg/dL and none of the 36 samples (< 20 mg/dL creatinine) had specific gravity ≤ 1.0010 or > 1.0200, it can be deduced that these examined urine specimens could not be considered to be specimens with invalid drug analysis results, as outlined in the DHHS guidelines (3, 4, 6).

In general, the type of liquid consumed did not appear to have an effect on risk of a dilute sample over all participants (Table III). However, among males, the types were restricted to water (Beverage A: four of the nine samples with < 20 mg/dL creatinine) and a lightly carbonated beverage (Beverage D: three of the nine samples with < 20 mg/dL creatinine). Among females, dilute samples were obtained following all liquid types (Beverage A: two, Beverage B: six, Beverage C: seven, and Beverage D: five of the 27 samples with < 20 mg/dL creatinine).

For the body composition parameters, no statistically significant differences were found between participants whose samples met the dilution criteria (creatinine < 20 mg/dL and specific gravity < 1.0030 but > 1.0010) and those
that did not. For male participants, the largest difference was for RMR, in which males with at least one dilute sample had a 29% higher RMR and had 11% lower percent body fat ($p > 0.05$). For female participants, the largest difference was for weight, in which females with at least one dilute sample weighed 20% less and were 8% shorter ($p > 0.05$). To show any difference as statistically significant may require a study with a larger subject sample size.

**Conclusions**

Findings from this study clearly suggest that obtaining a dilute urine sample from a donor does not necessarily indicate that the sample was intentionally diluted; it could be the result of natural physiological responses of the donor. In addition, individuals with higher resting metabolism (usually younger, trimmer and more muscular) have a greater potential for producing dilute urine specimens than individuals with lower resting metabolism. Certainly, in addition to the amount and type of fluid intake prior to the urine specimen collection, the role of factors—such as dietary habits, ethnic origin, health-related treatments, working conditions and/or genetic differences—in the production of dilute urine cannot be ruled out. For this very reason, the 107th U.S. Congress asked the DOT to study the effects of these factors on drug and alcohol validity testing (8, 9). The findings from the present study emphasize that dilute samples must be carefully evaluated by the regulatory authorities in consideration with the entire physiological and personal spectrum of dietary and personal habits and genetic differences affecting the characteristics of urine collected from a particular donor. Such evaluation must also include the necessity to inquire when (and how much of) any liquid was consumed by the donor before donating the urine sample. These steps will help to avoid false accusations of providing a dilute urine sample.

**References**


10 Chaturvedi et al.