### Background
Well-recognized markers for static (one time) or dynamic (monitoring over time) dehydration assessment have not been rigorously tested for their usefulness in clinical, military, and sports medicine communities. Objective: This study evaluated the components of biological variation and accuracy of potential markers in plasma, urine, saliva, and body mass, for static and dynamic dehydration assessment. Design: Eighteen healthy volunteers (13M, 5F) were studied while carefully controlling hydration and numerous pre-analytical factors. Biological variation was determined over three consecutive days using published methods. Atypical values based on statistical deviations from a homeostatic set-point were examined. Measured deviations in body fluid were produced using a separate, prospective dehydration experiment and evaluated by ROC analysis to quantify diagnostic accuracy. Results: All dehydration markers displayed substantial individuality and half displayed marked heterogeneity of intra-individual variation. Decision levels for all dehydration markers were within one standard deviation of the ROC criterion values and most were nearly identical to the regression group mean of dehydration volunteers but 1.5-7.0% of body mass. However, only...
Biological variation and diagnostic accuracy of dehydration assessment markers

Samuel N Cheuvront, Brett R Ely, Robert W Kenefick, and Michael N Sawka

ABSTRACT

Background: Well-recognized markers for static (one time) or dynamic (monitoring over time) dehydration assessment have not been rigorously tested for their usefulness in clinical, military, and sports medicine communities.

Objective: This study evaluated the components of biological variation and the accuracy of potential markers in plasma, urine, saliva, and body mass (Bm) for static and dynamic dehydration assessment.

Design: We studied 18 healthy volunteers (13 men and 5 women) while carefully controlling hydration and numerous preanalytic factors. Biological variation was determined over 3 consecutive days by using published methods. Atypical values based on statistical deviations from a homeostatic set point were examined. Measured deviations in body fluid were produced by using a separate, prospective dehydration experiment and evaluated by receiver operating characteristic (ROC) analysis to quantify diagnostic accuracy.

Results: All dehydration markers displayed substantial individuality and one-half of the dehydration markers displayed marked heterogeneity of intr-individual variation. Decision levels for all dehydration markers were within one SD of the ROC criterion values, and most levels were nearly identical to the prospective group means after volunteers were dehydrated by 1.8–7.0% of values, and most levels were nearly identical to the prospective dehydration experiment and evaluated by receiver operating characteristic (ROC) analysis to quantify diagnostic accuracy.

Conclusions: P_{osm} is the only useful marker for static dehydration assessment. P_{osm}, U_{sg}, and B_m are valid markers in the setting of dynamic dehydration assessment.

INTRODUCTION

Healthy humans maintain homeostatic control of the balance of body fluids and physiologic and behavioral adaptations (1, 2). However, when fluids are limited, illness strikes, or there is exposure to extreme environments, the accumulation of fluid deficits can threaten homeostasis, health, and performance (3, 4). The efficacy of any assessment marker (5) for detecting disturbances in homeostasis depends critically on the nature of body fluid losses (6–8). Clinical dehydration is a state of hypertonic hypovolemia that occurs in response to a net loss of hypotonic body fluids (9–11). The rise in extracellular tonicity is a hallmark clinical feature of dehydration that provides diagnostic distinction from isotonic or hypotonic hypovolemia (6–8, 10). Dehydration predictably modulates renal function and urine composition in accordance with the deficit of body water (12). Thus, the fundamental principles of body fluid regulation provide the framework for using plasma osmolality (P_{osm}) (2, 10, 13, 14), urine osmolality (U_{osm}) (15–17), urine specific gravity (U_{sg}) (14, 15), urine color (U_{col}) (15), and, possibly, saliva osmolality (S_{osm}) (18, 19) as body fluid markers of dehydration. Similarly, because humans maintain a relatively stable total-body water pool, despite diverse factors that affect water requirements (eg, climate, activity, and dietary solute load) (2, 20), acute changes in body mass (B_m) may be used to accurately measure dehydration across both clinical and sports medicine disciplines (8, 21–25).

When the balance between fluid intake and output is carefully controlled experimentally, a consensus exists as to what constitutes euhydrated values for plasma and urine (4, 23) and surrogate B_m stability (23, 24). In contrast, it is well appreciated (1, 20) that modest perturbations in body fluid balance may be well tolerated and difficult to detect biochemically until some functional threshold (≥2% B_m) (4) is breached. For example, dehydration thresholds (in mmol/kg) of 586 (26), 716 (17), 830 (14), and 1052 (15) have all been proposed for U_{osm}. Because dehydration was experimentally assured across studies, strong population heterogeneity may be responsible for the apparent discordance of values. Indeed, there can be substantial heterogeneity in healthy individuals with respect to homeostatic control of the body fluid balance (1, 27), and the reference interval of typical population body fluid concentrations is wide (28). In accordance with a statistical distribution theory (29), the degree of overlap between typical and atypical values will determine the nosologic sensitivity of any dehydration assessment marker (29–32).

1 From the Thermal and Mountain Medicine Division, US Army Research Institute of Environmental Medicine, Natick, MA.
2 The opinions and assertions contained in this article are the private views of the authors and should not be construed as official or reflecting the views of the US Army or the Department of Defense.
3 Supported by the US Army Medical Research and Materiel Command.
4 Address correspondence to SN Cheuvront, Thermal and Mountain Medicine Division, US Army Research Institute of Environmental Medicine, Kansas Street, Natick, MA 01760-5007. E-mail: samuel.n.cheuvront@us.army.mil.

Received March 8, 2010. Accepted for publication June 29, 2010.
First published online July 14, 2010; doi: 10.3945/ajcn.2010.29490.
The degree to which any marker can correctly classify an individual as dehydrated can be described by its diagnostic quality or accuracy (33). Although good diagnostic accuracy is a prerequisite for any marker to have a practical value, it is inadequate alone for describing the potential usefulness or application of a test (33). Additional explanatory power can be gained with knowledge of the biological variability that underlies the population reference intervals. Very widely applied methods (29, 31) for isolating the analytic (CV₁), intraindividual (CVᵢ), and interindividual (CVᵢ) components of variation in body fluids can be used to understand the diagnostic usefulness (33) of a static value from one individual compared with a population-based reference interval (29, 31, 32). In addition, data on biological variation allow interpretation of the temporal nature of homeostatic control (34, 35) so that the importance of any dynamic change in serial measurements in an individual can be understood in longitudinal terms (29–31, 35). A knowledge of variation can also afford cautious assignment of probabilistic decision levels (29) and may even provide the necessary framework (29, 31) for determining the elusive gold standard marker for dehydration (5). Although simple reports of CVᵢ or CVᵢ can be found for P_osm (1), U_osm (17), U_sg (36), and Bₘ (24), few studies (37–40) have used the necessary methodology (29, 31) for extracting appropriate decision-making information. A study of the biological components of variation for multiple dehydration markers measured simultaneously from the same set of subjects is absent from the literature (41).

Components of biological variation (29, 31) for potential markers in plasma, urine, saliva, and Bₘ were evaluated to understand their usefulness in relation to static and dynamic dehydration assessments. The diagnostic accuracy (33) of dehydration markers was quantified separately by using prospective and comprehensive dehydration experiments. The latter experiment complemented the former experiment by providing actual physiologic measurements of dehydration against which the validity of probabilistic decision levels could be evaluated. A limited diagnostic characterization of dehydration markers is available, (26) and several unique markers have, to our knowledge, never been characterized (15, 18). This information holds important potential for reducing the effect and complications of dehydration across broad medical disciplines (4, 6–8, 10, 42, 43).

SUBJECTS AND METHODS

Subjects

Eighteen active soldier volunteers took part in this 2-wk study [13 men and 5 women; age (mean ± SD): 24 ± 4 y; Bₘ: 78.2 ± 12.1 kg; height: 172.2 ± 9.2 cm; body mass index (in kg/m²): 26.4 ± 3.9]. All volunteers passed the Army Physical Fitness Test within the previous 6 mo and received a general medical clearance before participation: thus, all volunteers were considered physically fit and healthy. The use of alcohol, dietary supplements, and any medication other than an oral contraceptive was prohibited. No female volunteer was or became pregnant during the course of the study. Volunteers were provided informational briefings and gave voluntary, informed, written consent to participate. Investigators adhered to Army Regulation 70–25 on the use of volunteers in research. The US Army Research Institute of Environmental Medicine Human Use Review Committee approved this study.

Experimental design

Each volunteer completed 7 d of testing over 2 wk. The study was broken into 2 phases. In phase I, 3 consecutive days of testing were used to study biological variation. In phase II, 2 consecutive days of testing were completed twice separated by 1 wk.

Phase I

One day before the start of testing, an informational briefing was given to provide a detailed study schedule, to make anthropometric measurements, and to provide instructions regarding fluid intake. To achieve a euhydrated state before the first day of testing, volunteers were given 3.0 L fluid to consume over 24 h in addition to ad libitum beverage consumption and habitual occasional dietary practices. Volunteers were instructed to consume a premeasured 1.0 L water between waking and 1800 and another 2.0 L prepackaged sports drink for consumption between 1800 and 2200. It was also estimated that food intake would provide an additional 0.6 L fluid each day (2), which would bring daily fluid-intake totals to ≈3.6 L (2). Physical exercise was permitted but was restricted to a short list of allowable activities and work durations. No food or drink was permitted between 2200 and 0600 the next morning (8-h fast). All measurements began at precisely 0630 after a 30-min opportunity for personal hygiene. This procedure was repeated 3 times over as many consecutive days to establish the day-to-day biological variation of several markers while volunteers were euhydrated. In all, 54 measures of P_osm, U_osm, S_osm, U_sg, U_col, and Bₘ were planned in phase I.

Phase II

The daily procedure used in phase I was repeated again in phase II on day 1 and followed by a return to the laboratory in the early evening for dehydration and follow-up measurements on day 2. Dehydration was achieved by performing 3 to 5 h of work and rest cycles (50 min of work and 10 min of rest) that included treadmill (1.56 m/s; 4–7% grade) or cycle ergometer exercise (85–120 W) inside an environmental chamber set to 40°C and 20% relative humidity with a 1-m/s laminar wind flow. The purpose of the exercise–heat exposure was to increase body heat storage and induce sweating (hypotonic fluid loss) to achieve a state of hypertonic hypovolemia (9–11). The goal for dehydration spanned a functionally important range of 2.0–7.0% of Bₘ (2, 4, 20). After exiting the environmental chamber, volunteers were provided with a small, standardized meal (450 kcal; 57% carbohydrate, 30% fat, 13% protein, and 450 mg Na⁺) and 0.2 L water or apple juice. No additional food or water was permitted, and volunteers were kept in supervised housing until morning. The phase II procedure was performed twice. After the measure of dehydration markers on the following morning, volunteers were rehydrated by oral and/or intravenous fluids to return Bₘ to within 1% of the euhydrated baseline. Measurements of P_osm, U_osm, S_osm, U_sg, U_col, and Bₘ were made on each day for a planned total of 36 euhydrated and dehydrated samples in phase II.
The sequence of events for all measurements remained consistent throughout the study. Upon reporting to (phases I and II of day 1 or awakening at the laboratory (phase II of day 2), volunteers voided a urine sample at 0630 and were weighed nude. Then, volunteers sat quietly for 30 min and remained seated while having their blood drawn, after which they provided a saliva sample. To examine the potential effect of simple oral artifacts (ie, food, drink, tobacco, and gum) on SOsm measures, a subset of 4 dehydrated volunteers provided an additional saliva sample (n = 8 samples) 1 min after a 10-s mouthrinse with 50 mL water.

**Analytic measures**

**Plasma volume and body fluid osmometry**

A 3-mL sample of venous blood was collected without stasis in lithium-heparin tubes (Sarstedt Inc, Newton, NC). Blood samples were centrifuged, and plasma was separated for analysis without delay (44). The first void morning urine was collected in a sterile, inert polypropylene cup (Tyco Healthcare Group, Mansfield, MA), and a small aliquot was transferred into a 1.5-mL polypropylene cryvue vial (Wheaton, Milville, NJ) and briefly vortexed before an analysis of osmolality. The remaining sample volume was analyzed for USg and UCol. A 2-mL sample of unstirred whole saliva was collected into a single-use polypropylene Falcon tube (Voigt Global Distribution Inc, Lawrence, KS) and centrifuged before analysis. Plasma and saliva samples were centrifuged at 3500 x g at 4°C for 10 min.

POsm, USg, and SOsm were measured by freezing-point depression on 3 dedicated osmometers (Fiske Micro-Osmometer, model 210; Fiske, Norwood, MA) and always by the same 3 technicians assigned to the study. In addition to daily calibrations, each osmometer was calibrated in triplicate by using standards within the dedicated ranges for plasma (290 mmol/kg), saliva (50 mmol/kg), and urine (850 mmol/kg) before each new biological sample. The fluid samples were also run in triplicate, and the median value was taken as final. If any of the intrasample triplicate fluid measures differed by >1.0%, the median of 5 samples was used. This approach was recommended on the basis of the ordinate scale of the readings (44), and the desired imprecision was based on instrument resolution and the potential physiologic importance of small fluctuations (≥1.0%) in POsm to hormonal fluid regulation (12, 27). Plasma tonicity (effective osmolality) (13, 45) reflects the potential for osmotic water movement across cell membranes and should not include solutes that contribute to osmolality without affecting tonicity, such as plasma urea (13, 45). However, even after exclusion of major osmotically active moieties such as sodium, glucose, and urea, other ionic solutes may still account for deviations of ≥5–10 mmol/kg between measured osmolality and calculated (effective) tonicity (13, 45–47), regardless of what formula is chosen (45). This observation may explain, in part, why sodium alone is not an ideal dehydration assessment marker (2, 23). Therefore, the direct measure of POsm was considered synonymous with total plasma tonicity in this study of healthy volunteers, despite the small plasma urea contribution to osmolality but not tonicity.

**USg and UCOL**

After a small aliquot of urine was used for osmolality determinations, USg was measured in duplicate for each specimen with a refractometer (1110400A TS Meter; AO Reichert Scientific Instruments, Keene, NH) that was calibrated from 1.00 to 1.03 against liquid preparations of known relative density (mass/volume). UCol was determined in a naturally lit room against a standardized 8-point color chart produced by Armstrong et al (15). A single technician performed all USg and UCOL analyses to avoid between-observer variability. A blinded, triplicate analysis of 3 specimens (9 observations) on a single day was used to test single-rater reliability for UCOL.

**Bm and percentage dehydration**

In response to exercise-heat exposure, water (ie, sweat and urine) volume and Bm losses were considered equivalent (1 mL = 1 g) (22) after correction for carbon exchange and respiratory water loss, which was estimated at 2 g/min during heat exposure (21). The level of dehydration was calculated from the corrected change in nude Bm over 24 h (from 0630 on day 1 to 0630 on day 2) and expressed as a percentage of starting Bm. The starting 0630 nude Bm used in the denominator of the calculation was considered normal for an individual if it was within ±1.0% of his/her individual phase I mean (24), which was labeled as edehyrdated on the basis of previous 24-h fluid consumption (2), and subsequently confirmed as edehyrdated on the basis of meeting at least one additional plasma or urine concentration criterion (4, 23). The imprecision of the platform scale used (model WSI-600; Mettler Toledo, Toledo, Ohio) was checked daily at 25, 70, and 95 kg, and ±50 g was considered an acceptable performance.

**CVa**

The widely accepted methodology (29, 31) for generating and applying data on biological variation was used. In the case of assessing CVa, the third approach described by Fraser and Harris (31) was used based on the instability of samples, freeze-thaw concerns, and the typical approach to osmolality assessment (44). Thus, the analytic imprecision (CVa) reported for osmolality (POsm, USg, and SOsm) represented the between-batch analytic imprecision by using osmolality control standards (average CV of 18 triplicate measures). The CVa reported for all other markers was determined from the average CV of 3 duplicate (USg and Bm) or triplicate (UCol) calibration or specimen samples, respectively. All CVs were calculated as below:

\[
\text{CVa} = \left( \frac{\text{SD}}{\text{mean}} \right) \times 100
\]

Preanalytic variation (in vitro and in vivo) (29, 31) was minimized to the greatest extent possible. Therefore, it was assumed that biological variation around a homeostatic set point could be attributed to factors largely independent of preanalytic influences. The menstrual cycle phase could not be standardized and was accepted as one intrinsic source of biological variation.

**Biological variation**

One representative measurement, which was made daily for each volunteer, provided 54 potential POsm, USg, SOsm, USg, UCOL, and Bm values for analysis. No outliers were detected by using Reed’s criterion with an exclusion ratio ≥1.3. Gaussian data distributions were confirmed by the Shapiro-Wilk test; thus, all data are
expressed as means ± SDs. The total CV\textsubscript{I} and CV\textsubscript{G} were computed by analysis of variance according to published methods (31).

With the use of the 3 components of variation, the usefulness of a static measure against population-based reference intervals and the magnitude of dynamic changes required for statistical significance were determined. The ratio

$$\frac{(CV^2_A + CV^2_I)^{1/2}}{CV_G}$$  \hspace{1cm} (2)

[i.e., the index of individuality (II)] was calculated whereby a low value (<0.60: high individuality) indicated a low usefulness of population-based reference intervals, whereas a high value (>1.40: low individuality) indicated a high usefulness of population-based reference intervals (31, 32, 48, 50), particularly when an unusual result was repeated for verification (48). The magnitude of dynamic change required to make a difference statistically significant [i.e., the reference change value (RCV)] was calculated as

$$2^{1/2} \times 1.65 \times (CV^2_A + CV^2_I)^{1/2}$$  \hspace{1cm} (3)

for a one-sided test with a 95% probability level, which was consistent with a “very likely” clinical dehydration scale (29, 49). However, the RCV was not valid if the CV\textsubscript{I} was heterogeneous in volunteers (31, 34). The index of heterogeneity (IH) was calculated as the CV of

$$(SD^2_A + SD^2_I)^{1/2}$$  \hspace{1cm} (4)

divided by the theoretical CV if no heterogeneity existed:

$$\frac{2}{(n - 1)^{1/2}}$$  \hspace{1cm} (5)

where n is the number of specimens collected from each vol- untee (31, 34, 35). Because the number of replicates for each individual was precisely 3 in this study, the denominator was simplified to 1.0. Under the hypothesis of nonheterogeneity, the RCV was valid if

$$IH < \{1 + 2[1/(2n)^{1/2}]\}$$  \hspace{1cm} (6)

or 1.82 when n was set to 3 (31, 34, 35). We estimated that the potential for correlations between successive values to affect the RCV (29) was marginal because only RCV outcomes with a constant true variance (IH) were considered useful. A decision level for clinical judgments about dehydration was determined (29, 49, 50) by adding the RCV (units) to the euhydration grand mean. An SD for the decision level was calculated as below:

$$2^{1/2} \times SD_I (50)$$  \hspace{1cm} (7)

Judgments concerning the usefulness of decision levels were based on the consideration of II. The validity of the RCV for dehydration markers was based on the consideration of IH. When valid, the RCV for each marker was expressed across 4 probability levels (80%, 90%, 95%, and 99%) to form a simple clinical likelihood scale (49) for dehydration monitoring.

**Statistical analyses**

The diagnostic accuracy of \(P_{osm}, U_{osm}, S_{osm}, U_{sg}\), and \(U_{col}\) for detecting dehydration was determined by using receiver operating characteristic (ROC) analysis (GraphPad, 4.0; San Diego, CA). \(B_m\) was not included as a variable in the ROC analysis because the acute change in \(B_m\) itself serves to distinguish dehydration from euhydration (8, 22–25). The area under the ROC curve (AUC) represents the accuracy of the test for discriminating the presence of dehydration significantly better than chance (33). It also yields 2 valuable components of diagnostic accuracy, namely sensitivity and specificity or the fraction of true positive and true negative determinations, respectively. No precise tradeoff between sensitivity and specificity was assigned as long as both sensitivity and specificity were \(\geq 80\%\) (33), which represented odds of 4 to 1 in favor of a correct classification. The criterion value for optimal diagnostic performance was determined thereafter by the value with the highest combination of sensitivity and specificity.

When \(\geq 2\) ROC curves are generated by using the same set of subjects, a method analogous to the paired t test (51) can be used to compare performance diagnostics. Within each ROC analysis, the potential for repeated measures to reduce the SE and modestly inflate diagnostic accuracy was equitably ignored across all hydration measures (52). \(P_{osm}\) was used as the gold standard for comparison (2, 10, 13, 14). To control for \(x\) inflation of the 4 comparisons to \(P_{osm}\) (ie, \(U_{osm}, S_{osm}, U_{sg}\), and \(U_{col}\)), Holm’s method of sequential Bonferroni corrections for interdependent tests was used. \(S_{osm}\) rinse measures were analyzed by using repeated measures analysis of variance and Tukey’s post hoc procedure. Significance was accepted at the 95% probability level. All data are presented as means ± SDs.

**RESULTS**

All 18 volunteers completed phases I and II of the study. All 54 possible samples were obtained on each measured variable in phase I with the exception of \(P_{osm}\) for which there were 51 samples. In phase II, a complete set of 36 measurement pairs was obtained for ROC analysis on all measures except \(P_{osm}\) \((n = 33)\) and \(S_{osm}\ \((n = 35)\). Lost samples were due to phlebotomy problems with one volunteer, handling error, or an inability of one volunteer to produce enough saliva while dehydrated by \(\approx 6.0\%\) of \(B_m\).

**Phase I**

A descriptive summary (mean ± SD) for the 3 d of euhydrated measurements in presented in Table 1. The group means were below (for \(U_{osm}, U_{sg}, U_{col}\), and \(B_m\)) or slightly above (for \(P_{osm}\)) the euhydration cutoff values, but all of these variables were \(\leq 1\, SD\) of the cutoff, which did not suggest an overt body water deficit or dilution. The \(U_{osm}\)-to-\(P_{osm}\) ratio (2.1 ± 0.8) supported this conclusion as well (12, 16, 27). For variables where typical population intervals have been reported (ie, for \(P_{osm}, U_{osm}, U_{sg}\)) (28), results fell within these intervals.
VARIATION AND ACCURACY OF DEHYDRATION MARKERS

Euhydrated Sosm and Ucol measures and typical percentage fluctuations in daily Bm were consistent with limited reports on a smaller scale (15, 18, 24).

The basic calculated components of variation (ie, CV_A, CV_I, CV_G) in this study are expressed as percentage CVs in Table 2. Rarely was a triplicate calibration CV for Posm, Sosm, or Uosm >0.5%. Duplicate or triplicate measures for Usg, Bm, and Ucol were commonly identical. The CV_A for all measures was ultimately well below 0.5 × CV_I, as recommended (29, 31).

CV_I was only slightly smaller than previous reports (0.43–0.50) (37, 38) and a negligible (zero) value for Bm was also obtained by Widjaja et al (39). CV_I was smallest (0.4–1.3%) for Posm, Usg, and Bm and largest (9.5–30.9%) for Sosm, Uosm, and Ucol. The CV_I and CV_G values for Posm and Bm were also very comparable with Fraser et al (37) and Widjaja (39).

CV_I was similar to Shephard et al (40), but the CV_G was higher.

CV_I was smaller than CV_G among all markers (Table 2), which was consistent with the general trend for most biological constituents of medical interest (29, 31, 32). As a result, the II was <1.4 for all markers and <0.60 for all but Posm (0.90) and Ucol (0.65). Only Posm was considered to be of potential static dehydration assessment value (31, 32, 48, 50). For IH, only Posm, Usg, and Bm were considered to be of potential dynamic dehydration monitoring values (IH < 1.82) (31, 34, 35). All other variables displayed significant heterogeneity. The RCV, or magnitude of dynamic change in a serial result that would make a difference significant, is provided (Table 3) for 95% probability (dehydration very likely) as a percentage change in the marker, as a unit change, and decision level. The unit RCVs at 4 probability levels for valid markers of dynamic dehydration monitoring only are presented in Table 4.

Phase II
Volunteers were dehydrated by 3.7 ± 1.0% of Bm (range: 1.8–7.0%). Individual and mean values for Posm, Uosm, Sosm, Usg, and Ucol when subjects were euhydrated and dehydrated are depicted in Figure 1 (A–E). Mean euhydrated values were nearly identical to measured values in Table 1 and again <1 SD of the suggested dehydration cutoff (23). The mean dehydration values were nearly identical to the 95% probabilistic decision levels (Table 3) with the exception of Sosm. The dotted lines in each panel represent the criterion values in Table 5, all of which were <1 SD of the probabilistic decision levels (Table 3). The Usg criterion value was the same as that reported by Bartok et al (26), whereas Posm and Uosm were substantially higher. Although Sosm was the only analyte that was statistically inferior to Posm, the diagnostic accuracy of all analytes was significantly better than chance (Table 5), and the criterion value for each was balanced by a minimum 80% threshold for both sensitivity and specificity. In the Sosm-rinse study subgroup, euhydrated Sosm (88 ± 32 mmol/kg) increased significantly (P < 0.05) after dehydration (147 ± 88 mmol/kg) but dropped back to euhydrated values (81 ± 25 mmol/kg) 1-min after the mouthrinse even though volunteers remained dehydrated.

DISCUSSION
To our knowledge, this is the first study to carry out the determination and application of data on biological variation (29, 31) by using multiple dehydration assessment markers (23). Atypical values on the basis of statistical deviations from a homeostatic set point were examined alongside measured body fluid deficits induced by using a prospective dehydration study. We provide decision levels for multiple clinical markers of dehydration at the 95% probability level that can be used as a starting point to guide decision making for dehydration assessment across broad medical disciplines.

TABLE 1
Plasma, urine, and saliva quantity estimates when euhydrated

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Euhydration</th>
<th>Measured (n = 18)</th>
<th>Range</th>
<th>Reference interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posm (mmol/kg)</td>
<td>0.4</td>
<td>1.3</td>
<td>1.5</td>
<td>NA</td>
</tr>
<tr>
<td>Sosm (mmol/kg)</td>
<td>1.0</td>
<td>9.5</td>
<td>35.8</td>
<td>NA</td>
</tr>
<tr>
<td>Uosm (mmol/kg)</td>
<td>0.7</td>
<td>28.3</td>
<td>57.9</td>
<td>NA</td>
</tr>
<tr>
<td>Usg (units)</td>
<td>Negligible</td>
<td>0.4</td>
<td>1.0</td>
<td>NA</td>
</tr>
<tr>
<td>Ucol (units)</td>
<td>Negligible</td>
<td>30.9</td>
<td>47.4</td>
<td>NA</td>
</tr>
<tr>
<td>Bm (percentage change)</td>
<td>Negligible</td>
<td>1.1</td>
<td>26.6</td>
<td>NA</td>
</tr>
</tbody>
</table>

1 All values are percentages. CV_I, intraindividual variation; CV_G, interindividual variation; P_posm, plasma osmolality; S_saliva osmolality; U_osm, urine osmolality; U_sgr, urine specific gravity; U_col, urine color; B_m, body mass.

TABLE 2
Components of analytic variation (CV_A) and biological variation

<table>
<thead>
<tr>
<th>Quantity</th>
<th>CV_A</th>
<th>CV_I</th>
<th>CV_G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posm (mmol/kg)</td>
<td>0.4</td>
<td>1.3</td>
<td>1.5</td>
</tr>
<tr>
<td>Sosm (mmol/kg)</td>
<td>1.0</td>
<td>9.5</td>
<td>35.8</td>
</tr>
<tr>
<td>Uosm (mmol/kg)</td>
<td>0.7</td>
<td>28.3</td>
<td>57.9</td>
</tr>
<tr>
<td>Usg (units)</td>
<td>Negligible</td>
<td>0.4</td>
<td>1.0</td>
</tr>
<tr>
<td>Ucol (units)</td>
<td>Negligible</td>
<td>30.9</td>
<td>47.4</td>
</tr>
<tr>
<td>Bm (percentage change)</td>
<td>Negligible</td>
<td>1.1</td>
<td>26.6</td>
</tr>
</tbody>
</table>

1 All values are percentages. CV_I, intraindividual variation; CV_G, interindividual variation; P_posm, plasma osmolality; S_saliva osmolality; U_osm, urine osmolality; U_sgr, urine specific gravity; U_col, urine color; B_m, body mass.

indicated by a small ratio as can be seen in equation (2) and consistent with the general trend for most biological moieties of medical interest (29, 31, 32). Volunteers with a small II (<0.60) can experience a change in any dehydration assessment marker that is very atypical for them, but their values would still fall within the typical population-based reference interval (31, 32, 48, 50). When II is >1.40, atypical individual values will be seen with greater frequency. However, the probability of identifying an atypical individual result within a population-based reference interval increases rapidly as II exceeds 0.60 and approaches 1.00 (32), particularly if a confirmatory measurement is made (48). The II for P\textsubscript{osm} (0.90) was very near the values of 0.93 and 1.40 reported by Fraser et al (37) and Jahan et al (38), respectively.

Although there is disagreement on how to interpret a P\textsubscript{osm} between 295 and 300 mmol/kg (42), a P\textsubscript{osm} decision level of 301 ± 5 mmol/kg offers a 95% probability of dehydration and was close to the group mean for dehydration achieved prospectively (304 ± 6 mmol/kg) (Figure 1A). This P\textsubscript{osm} decision level also exceeds the upper 295 mmol/kg population-based reference limit (10, 28) and encompasses the sometimes proposed overt dehydration value of 300 mmol/kg (8, 42). The ROC criterion value of 297 mmol/kg (Figure 1A), which displayed high sensitivity and specificity for dehydration (Table 5), was ≤1 SD of the decision level and identical when an 80% probability level was applied to the RCV calculation (Table 4). When decision levels were calculated for S\textsubscript{osm}, U\textsubscript{osm}, U\textsubscript{col}, and B\textsubscript{m} (Table 3), they were also very near the group means for experimental dehydration and ≤1 SD of the ROC criterion values in Table 5. Although this seems to support a narrow range of values as good general starting points, II data negate their relevance for static dehydration assessment (31, 32, 48, 50) in high-risk treatment situations. However, a practical use in circumstances of lower-risk assessment may still be warranted (4, 5, 23).

Dynamic dehydration assessment was best accomplished by using P\textsubscript{osm}, U\textsubscript{sg}, and B\textsubscript{m} (Table 3) on the basis of the prerequisite homogeneity of CV\textsubscript{r} (31, 34, 35). Because of the study design and the number of specimens collected for each volunteer, the validity of the RCV was only accepted when IH was <1.82 (Table 3). Whether serial changes were calculated by using the 95% probability or some other level of confidence (Table 4), the corroborating data suggested that serial changes would be detectable whether they occurred over hours or days, and possibly even over several months (30). On the other hand, the objective assessment that CV\textsubscript{r} was not constant for S\textsubscript{osm}, U\textsubscript{osm}, and U\textsubscript{col} (Table 3) casts doubts on their use in this manner. The fact that U\textsubscript{sg} was useful for serial measures may seem counterintuitive given that renal concentrating mechanisms produce strong relations among U\textsubscript{sg}, U\textsubscript{osm}, and U\textsubscript{col} (15). However, strong relations contain substantial scatter, and U\textsubscript{osm} may vary by as little as 81 (U\textsubscript{sg} : 1.012) or as much as 533 mmol/kg (U\textsubscript{sg} : 1.021) at any given U\textsubscript{sg} measured across a broad range (1.005–1.028) (53). The categorical nature of the U\textsubscript{col} chart led to measurement imprecision on the scale required for clinical usefulness, although it may still be of practical use for low-risk population assessment (15, 23). Importantly, any urine concentration measure must be interpreted carefully because disease states, time of day of urine collection (first morning void preferred), and drinking behaviors can all confound urine interpretation for hydration (4, 14, 15, 17, 23, 28, 53). S\textsubscript{osm} results were similar to U\textsubscript{osm} and U\textsubscript{col} results when IH and IH were concerned, and S\textsubscript{osm} also performed to the minimum ROC standard (Table 5; Figure 1B). The potential effect of simple oral artifacts (ie, food, drink, tobacco, and gum) to mar the diagnostic usefulness of S\textsubscript{osm} (18, 19) is also worrisome considering the subgroup S\textsubscript{osm}-rinse outcomes.

The results of this study in young, healthy volunteers should have relevant application to military (43) and sports medicine (4) populations. Fraser et al (37) clearly showed that RCVs generated from young subjects can be applied to healthy older subjects (> 70 y) because biological variation does not change with age if health is maintained. Indeed, studies of biological variation require observations from a healthy population (29, 31). Thus, the RCVs of 9 mmol/kg (P\textsubscript{osm}), 0.010 (U\textsubscript{sg}), and 2.5% (B\textsubscript{m}) could be used broadly for dynamic dehydration assessment as long as baseline values are within the typical population-based reference intervals. It is also possible that a P\textsubscript{osm} of 301 ± 5 mmol/kg could be used as a decision level for static dehydration assessment within both a healthy and stable, but diseased, clinical population (29, 31). Because clinical assessment of an individual with an unknown health status involves the calculation of an osmolar gap (measured – calculated) (13), a P\textsubscript{osm} of 301 ± 5 mmol/kg in the presence of a gap <10 mmol /kg would also indicate dehydration because the pathologic glucose or urea values (3 times upper reference limit) required to raise a P\textsubscript{osm} above the decision level would be simultaneously ruled in or out by their direct measurement (13, 45). Because

### Table 3

Indexes derived from analytic and biological variation data

<table>
<thead>
<tr>
<th>Quantity</th>
<th>II</th>
<th>IH</th>
<th>RCV</th>
<th>RCV</th>
<th>Decision level</th>
</tr>
</thead>
<tbody>
<tr>
<td>P\textsubscript{osm} (mmol/kg)</td>
<td>0.90</td>
<td>1.35</td>
<td>3.1</td>
<td>9</td>
<td>301 ± 5</td>
</tr>
<tr>
<td>S\textsubscript{osm} (mmol/kg)</td>
<td>0.27</td>
<td>9.53</td>
<td>22.2</td>
<td>45</td>
<td>86 ± 9</td>
</tr>
<tr>
<td>U\textsubscript{osm} (mmol/kg)</td>
<td>0.49</td>
<td>28.32</td>
<td>65.9</td>
<td>404</td>
<td>1018 ± 245</td>
</tr>
<tr>
<td>U\textsubscript{sg} (units)</td>
<td>0.44</td>
<td>0.43</td>
<td>1.0</td>
<td>0.010</td>
<td>1.028 ± 0.006</td>
</tr>
<tr>
<td>U\textsubscript{col} (units)</td>
<td>0.65</td>
<td>30.90</td>
<td>71.9</td>
<td>2.7</td>
<td>6.5 ± 1.7</td>
</tr>
<tr>
<td>B\textsubscript{m} (percentage change)</td>
<td>0.04</td>
<td>1.08</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5 ± 1.5</td>
</tr>
</tbody>
</table>

1. Index of individuality; II, index of heterogeneity; RCV, reference change value; P\textsubscript{osm}, plasma osmolality; S\textsubscript{osm}, saliva osmolality; U\textsubscript{osm}, urine osmolality; U\textsubscript{sg}, urine specific gravity; U\textsubscript{col}, urine color; B\textsubscript{m}, body mass.

### Table 4

Reference change value (RCV) and dehydration likelihood scale

<table>
<thead>
<tr>
<th>Probability</th>
<th>Unidirectional z score</th>
<th>Dehydration likelihood</th>
<th>RCV units</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>P\textsubscript{osm}</td>
</tr>
<tr>
<td>80%</td>
<td>0.84</td>
<td>Likely</td>
<td>5</td>
</tr>
<tr>
<td>90%</td>
<td>1.28</td>
<td>More likely</td>
<td>7</td>
</tr>
<tr>
<td>95%</td>
<td>1.65</td>
<td>Very likely</td>
<td>9</td>
</tr>
<tr>
<td>99%</td>
<td>2.33</td>
<td>Virtually certain</td>
<td>13</td>
</tr>
</tbody>
</table>

1. Concept adapted from Fraser et al (49). P\textsubscript{osm}, plasma osmolality; U\textsubscript{sg}, urine specific gravity; B\textsubscript{m}, body mass.
any unusual Posm result is likely to be repeated, II >1.0 should decrease the number of false-positive confirmatory measures (48). A Posm of 301 ± 5 mmol/kg would be valid only for diagnosing hypertonic hypovolemia (6, 7, 10, 13). A gold-standard method for diagnosing hypovolemia of isotonic or hypotonic origins remains elusive (7).

All biological variation and ROC analysis data were predicated on the fact that volunteers were euhydrated daily (phase I) and before dehydration was imposed (phase II). Euhydration can be supported on the basis of 24-h fluid intakes ≥3.0 L (2) and on the grounds that all volunteers were below euhydration cutoff values for ≥2 hydration measures each day (4, 23). The mean values for all volunteers were also never >1 SD below the cutoff, and the Uosm-to-Posm ratio was not close to 1.0 (dilution) (12, 16, 27) or >3.0 (concentrated) as observed after a 14-h fast (53). However, the group mean for Posm was slightly higher (292 ± 3 mmol/kg) (Table 1; Figure 1A) than the corresponding 290 mmol/kg euhydration cutoff. Many of the studies used to support the <290 mmol/kg cutoff used a testing methodology whereby fluid was consumed by subjects in the morning before blood sampling. The grand mean Posm for these studies was 284 mmol/kg (2). A Posm in this range or lower would not alter the usefulness of the RCV, but it would reduce the sensitivity of the 301 ± 5-mmol/kg decision level. However, fewer than 2% of free-living people have a Posm ≤285 mmol/kg, and most free-living people have a Posm between 285 and 295 mmol/kg when they consume ≥3.0 L fluid each day (42). A Posm ≤285 mmol/kg is also commonly associated with a Uosm <400 mmol/kg (27) and, therefore, results in a Uosm-to-Posm ratio consistent with mild dilution. Indeed, fluid volumes vary (when reported), but data from Robertson et al (12) support a 1 mmol/kg drop in Posm for each 175 mL water consumed after an overnight fast. There is substantial individual variation in the Posm set point brought about by an even larger variation in the thirst threshold and arginine vasopressin sensitivity (27). Taken together, a Posm of 292 ± 3 mmol/kg is a defensibly euhydrated value that reflects an overnight fluid-restriction methodology coupled with ordinary biological variation.
In conclusion, the sports medicine literature provides a consensus about what threshold values constitute dehydration by using a variety of dehydration assessment markers (4, 23). We add to this literature by providing decision levels for multiple body fluids that can be used as starting points for the diagnosis and treatment of dehydration. Values that occur between euhydrated levels of hydration and treatment of dehydration. Values that occur between euhydrated levels of hydration and treatment of dehydration. 

In conclusion, the sports medicine literature provides a consensus about what threshold values constitute dehydration by using a variety of dehydration assessment markers (4, 23). We add to this literature by providing decision levels for multiple body fluids that can be used as starting points for the diagnosis and treatment of dehydration. Values that occur between euhydrated levels of hydration and treatment of dehydration. Values that occur between euhydrated levels of hydration and treatment of dehydration. 

### REFERENCES