

# Comparison of 3 Methods to Assess Urine Specific Gravity in Collegiate Wrestlers

Kristin J. Stuempfle; Daniel G. Drury

Gettysburg College, Gettysburg, PA

*Kristin J. Stuempfle, PhD, ATC, contributed to conception and design; acquisition and analysis and interpretation of the data; and drafting, critical revision, and final approval of the article. Daniel G. Drury contributed to analysis and interpretation of the data and critical revision and final approval of the article.*

*Address correspondence to Kristin J. Stuempfle, PhD, ATC, Health and Exercise Sciences Department, Campus Box 432, Gettysburg College, Gettysburg, PA 17325. Address e-mail to kstuempf@gettysburg.edu.*

**Objective:** To investigate the reliability and validity of refractometry, hydrometry, and reagent strips in assessing urine specific gravity in collegiate wrestlers.

**Design and Setting:** We assessed the reliability of refractometry, hydrometry, and reagent strips between 2 trials and among 4 testers. The validity of hydrometry and reagent strips was assessed by comparison with refractometry, the criterion measure for urine specific gravity.

**Subjects:** Twenty-one National Collegiate Athletic Association Division III collegiate wrestlers provided fresh urine samples.

**Measurements:** Four testers measured the specific gravity of each urine sample 6 times: twice by refractometry, twice by hydrometry, and twice by reagent strips.

**Results:** Refractometer measurements were consistent between trials ( $R = .998$ ) and among testers; hydrometer mea-

surements were consistent between trials ( $R = .987$ ) but not among testers; and reagent-strip measurements were not consistent between trials or among testers. Hydrometer ( $1.018 \pm 0.006$ ) and reagent-strip ( $1.017 \pm 0.007$ ) measurements were significantly higher than refractometer ( $1.015 \pm 0.006$ ) measurements. Intraclass correlation coefficients were moderate between refractometry and hydrometry ( $R = .869$ ) and low between refractometry and reagent strips ( $R = .573$ ). The hydrometer produced 28% false positives and 2% false negatives, and reagent strips produced 15% false positives and 9% false negatives.

**Conclusions:** Only the refractometer should be used to determine urine specific gravity in collegiate wrestlers during the weight-certification process.

**Key Words:** refractometer, hydrometer, reagent strips, hydration status

In 1997, three collegiate wrestlers died while attempting to reduce weight by dehydration.<sup>1-3</sup> To prevent a recurrence of this tragedy, the National Collegiate Athletic Association (NCAA) introduced new rules in 1998 that discourage dangerous weight-cutting practices.<sup>1-3</sup> These new rules include a weight-certification process that requires the determination of hydration status. The NCAA selected urine specific gravity as the most practical, cost-efficient hydration measure to use during the weight-certification process.<sup>4</sup>

Urine specific gravity is a measure of the ratio of the density of urine to the density of water. Urine specific-gravity measurements normally range from 1.002 to 1.030.<sup>5</sup> The NCAA selected a urine specific-gravity measurement of  $\leq 1.020$  to indicate euhydration.<sup>4</sup> Wrestlers with a urine specific gravity  $\leq 1.020$  are considered euhydrated and may have their body composition assessed to determine their minimal weight for competition, whereas wrestlers with a urine specific gravity  $> 1.020$  are considered to be dehydrated and may not proceed to body-composition testing on that day.

Refractometry, hydrometry, and reagent strips are commonly used to assess urine specific gravity. In 1998, the NCAA allowed the use of all 3 methods.<sup>4</sup> However, in 1999, the use of reagent strips was eliminated.<sup>6</sup> Previous reports have indicated that refractometry is the criterion measure for urine specific gravity<sup>7</sup> and that urine specific gravity measured by re-

fractometry is a valid indication of hydration status.<sup>8-11</sup> In 2 papers assessing urinary indices of hydration status, Armstrong et al<sup>8,9</sup> reported that urine specific-gravity measurement by refractometry was a more sensitive indication of hydration status than blood measurements, including plasma osmolality, plasma sodium, or hematocrit. Popowski et al<sup>11</sup> also concluded that measurement of urine specific gravity by refractometry was a valid assessment of hydration status, although it may lag somewhat behind plasma osmolality during progressive acute dehydration. Finally, the recent National Athletic Trainers' Association position statement on fluid replacement for athletes stated that urine specific gravity measured by a refractometer should be used to determine the hydration status of athletes.<sup>10</sup>

Research to assess the validity of hydrometry and reagent strips compared with refractometry to determine urine specific gravity has provided mixed results. McCrossin and Roy<sup>12</sup> described the overall correlation between refractometry and hydrometry as "good." In studies comparing refractometry with reagent strips, several researchers have suggested that reagent strips are an acceptable alternative to refractometry,<sup>13-15</sup> whereas others have concluded that they are not.<sup>12,16-19</sup> Although these findings are contradictory, the interclass Pearson correlation coefficients reported in these previous validity studies might not be the most appropriate statistical approach

to use when comparing 2 methods of measurement.<sup>20,21</sup> Furthermore, none of these authors assessed the reliability of refractometry, hydrometry, or reagent strips.

Therefore, our purposes were to assess the reliability of refractometry, hydrometry, and reagent-strip measurements across multiple trials and testers and to investigate the validity of hydrometer and reagent-strip measurements compared with refractometry using a variety of statistical approaches.

## METHODS

Reliability of refractometry, hydrometry, and reagent strips was assessed between 2 trials and among 4 testers. The validity of hydrometry and reagent strips was assessed by comparison with refractometry, the criterion measure for urine specific gravity.<sup>8-11</sup>

### Subjects

Twenty-one healthy members of an NCAA Division III wrestling team (age =  $20.0 \pm 1.29$  years, height =  $174.7 \pm 8.27$  cm, mass =  $82.8 \pm 18.00$  kg) each provided one urine sample. The subjects provided written informed consent, and the study was approved by the institution's institutional review board.

### Testers

Three certified athletic trainers and one athletic training student who were experienced in refractometer, hydrometer, and reagent-strip measurements served as testers.

### Instruments

The following instruments were used to assess urine specific gravity. The Schuco Clinical Refractometer (model 5711-2021; Williston Park, NY) has a temperature-compensating dial and graduated intervals of 0.005 units with a scale ranging from 1.000 to 1.040. It was calibrated with distilled water before use.

Urine samples greater than 60 mL were measured in an Assistant Urinprober hydrometer (model 242; Sondheim/Rhon, Germany), which is graduated in intervals of 0.001 units with a scale ranging from 1.000 to 1.060. Urine samples less than 60 mL were measured in a smaller Assistant Urinprober hydrometer (model 248), which is graduated in intervals of 0.002 units with a scale ranging from 1.000 to 1.060. Distilled water provided the calibration standard for the hydrometers before use. Hydrometer results were adjusted for temperature by adding or subtracting 0.001 specific-gravity units for each 3°C above or below 20°C, respectively.<sup>22</sup>

The N-MULTISTIX 10 SG Reagent Strips (Miles Laboratories, Inc, Elkhart, IN) have a specific-gravity scale ranging from 1.000 to 1.030, with color blocks in intervals of 0.005 units. The reagent strips were placed in the urine and removed and specific gravity read between 45 and 60 seconds after removal.

A standard laboratory thermometer measured urine temperature at the time of assessment.

### Procedures

Testing was done before the start of the wrestling season, and subjects had not exercised within 24 hours of testing. Sub-

jects were not given any special instructions concerning fluid consumption before testing. Urine samples were collected between 1:00 PM and 3:00 PM Greenwich Mean Time and analyzed immediately. Urine samples were divided into 2 subsamples (trials A and B) for analysis. Each tester measured the specific gravity of each urine subsample in the following order: reagent strip, hydrometer, and refractometer. All specific-gravity measurements were read to the nearest 0.001 specific-gravity unit. Therefore, the refractometer, the smaller Assistant Urinprober hydrometer, and the reagent strips were read to a precision of less than one marked unit. A recorder wrote down the specific-gravity measurements on the data sheets so that testers could not compare the values of the measurements.

### Statistical Analysis

We calculated a  $3 \times 2 \times 4$  (method  $\times$  trial  $\times$  tester) repeated-measures analysis of variance for significant mean differences in urine specific-gravity measurements among methods, trials, and testers with alpha set at  $P < .05$  and followed up with Scheffé post hoc comparisons. Intraclass correlation coefficients were calculated between methods and trials. The strength of the intraclass correlation coefficients was assessed according to the rating scale of Vincent.<sup>23</sup> We prepared Bland-Altman plots<sup>21</sup> to evaluate agreement between measurement methods for urine specific gravity. Refractometer measurements were accepted as the true indication of hydration status.<sup>8-11</sup> When the hydrometer or reagent strips indicated results contradictory to the refractometer, the measurements were recorded as false positives or false negatives. A false positive was defined as a measurement with the hydrometer or reagent strip indicating dehydration (value exceeded 1.020), when the refractometer indicated euhydration (value did not exceed 1.020). A false negative was defined as a measurement by the hydrometer or reagent strip indicating euhydration (value did not exceed 1.020), when the refractometer indicated dehydration (value exceeded 1.020). False positives and false negatives are reported as a percentage of all samples. The sensitivity of a test indicates how well a test finds disease positives (ie, those with the condition, such as dehydration). The sensitivity of the hydrometer and reagent strips was calculated as the number that were both disease positive and test positive, divided by the number that were disease positive, times 100.<sup>24</sup> The specificity of a test indicates how well a test excludes disease negatives (ie, those without the condition, such as dehydration). The specificity of the hydrometer and reagent strips was calculated as the number that were both disease negative and test negative, divided by the number that were disease negative, times 100.<sup>24</sup>

## RESULTS

The method  $\times$  trial  $\times$  tester interaction was significant ( $F_{6,160} = 4.085$ ,  $P = .0008$ ) (Table 1). The refractometer was reliable between trials for each tester and among all testers. The hydrometer was reliable between trials for each tester but was not consistent among testers. The reagent strips were not reliable between trials for the testers or among testers.

A significant interaction occurred between method and trial ( $F_{2,160} = 4.079$ ,  $P = .0187$ ) (Table 2). The refractometer and hydrometer were reliable between trials, but the reagent strips were not. The tester  $\times$  trial interaction also was significant ( $F_{3,80} = 3.423$ ,  $P = .0211$ ) (Table 2). Tester 1 and tester 2

**Table 1. Method × Trial × Tester Interaction (Mean ± SD)**

	Trial A	Trial B
Refractometer		
Tester 1	1.015 ± 0.006	1.015 ± 0.006
Tester 2	1.015 ± 0.006	1.015 ± 0.006
Tester 3	1.015 ± 0.006	1.015 ± 0.006
Tester 4	1.015 ± 0.006	1.015 ± 0.006
Hydrometer		
Tester 1	1.018 ± 0.006	1.018 ± 0.007
Tester 2	1.017 ± 0.005	1.017 ± 0.005
Tester 3	1.017 ± 0.006	1.017 ± 0.007
Tester 4	1.018 ± 0.006	1.018 ± 0.006
Reagent strips		
Tester 1	1.018 ± 0.007	1.016 ± 0.007
Tester 2	1.015 ± 0.006	1.015 ± 0.008
Tester 3	1.015 ± 0.006	1.019 ± 0.007
Tester 4	1.018 ± 0.006	1.019 ± 0.008

**Table 2. Method × Trial and Tester × Trial Interactions (Mean ± SD)**

	Trial A	Trial B
Method × trial interaction		
Refractometer	1.015 ± 0.006	1.015 ± 0.006
Hydrometer	1.018 ± 0.006	1.018 ± 0.006
Reagent strips	1.016 ± 0.006	1.018 ± 0.007
Tester × trial interaction		
Tester 1	1.017 ± 0.006	1.017 ± 0.006
Tester 2	1.016 ± 0.006	1.016 ± 0.006
Tester 3	1.016 ± 0.006	1.017 ± 0.007
Tester 4	1.017 ± 0.006	1.018 ± 0.007

**Table 3. Method and Trial Main Effects (Mean ± SD)**

Method main effect	
Refractometer	1.015 ± 0.006
Hydrometer	1.018 ± 0.006*
Reagent strips	1.017 ± 0.007*
Trial main effect	
Trial A	1.016 ± 0.006
Trial B	1.017 ± 0.007†

\*Significantly different from refractometer ( $P < .05$ ).

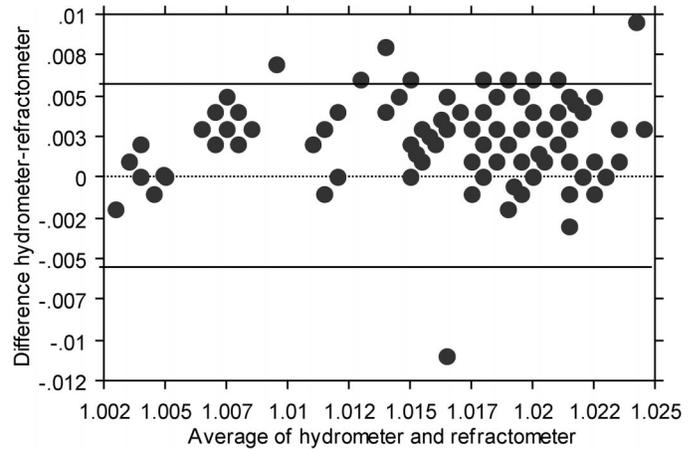
†Significantly different from trial A ( $P < .05$ ).

were consistent between trials, but tester 3 and tester 4 were not.

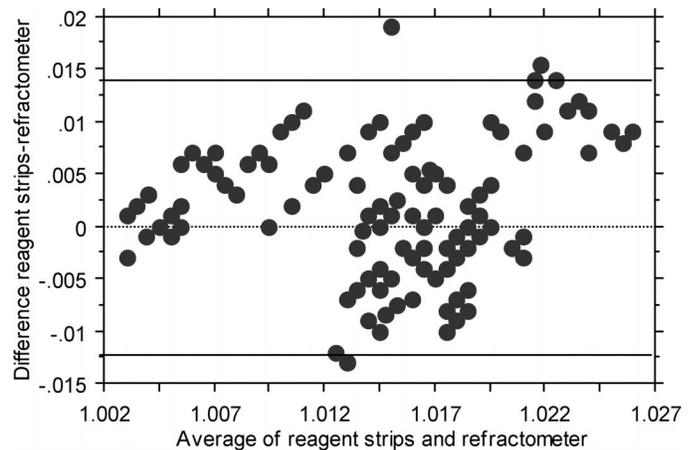
Main effects were significant for both method and trial (Table 3). Refractometer measurements were significantly lower than hydrometer or reagent-strip measurements ( $F_{2,160} = 8.993, P = .0002$ ). Trial A measurements were significantly lower than trial B measurements ( $F_{1,80} = 5.769, P = .0186$ ).

Intraclass reliability between trials was high for refractometry ( $R = .998$ ) and hydrometry ( $R = .987$ ) and moderate for reagent strips ( $R = .854$ ). Intraclass coefficients between refractometry and hydrometry were moderate ( $R = .869$ ) and low between refractometry and reagent strips ( $R = .573$ ).

Hydrometer measurements were consistently greater than refractometry measurements (mean difference =  $0.002 \pm 0.003$ ) (Figure 1). The calculated 95% limits of agreement indicate that, for 95% of observations, hydrometer values will



**Figure 1. Bland-Altman plot for refractometry and hydrometry. The dotted line represents zero difference between the 2 methods and the 2 solid lines ±2 standard deviations of the difference between methods.**



**Figure 2. Bland-Altman plot for refractometry and reagent strips. The dotted line represents zero difference between the 2 methods and the 2 solid lines ±2 standard deviations of the difference between methods.**

be 0.004 less than or 0.008 greater than refractometer values. Reagent-strip measurements also tended to be greater than refractometer measurements (mean difference =  $0.002 \pm 0.007$ ) with 95% limits of agreement indicating that reagent-strip values are expected to be 0.012 less than or 0.016 greater than refractometer measurements (Figure 2).

False positives (pass refractometer, fail other method) occurred with both hydrometry (47/168, 28%) and reagent strips (25/168, 15%). False negatives (fail refractometer, pass other method) also occurred with the hydrometer (3/168, 2%) and reagent strips (15/168, 9%). Sensitivity and specificity of the hydrometer were 88% and 67%, respectively. Sensitivity of the reagent strips was 38% and specificity was 83%.

## DISCUSSION

Two types of questions can be asked in method-comparison studies. First, what are the characteristics of each method? How repeatable are the measurements for each method? The reliability of a method is typically determined by the test-retest method.<sup>23</sup> Second, how do the methods compare? Do the



Figure 3. Trial-by-tester interaction for refractometer.



Figure 4. Trial-by-tester interaction for hydrometer.

methods measure the same thing? The validity of a method may be determined by comparing it with another method known to be valid.<sup>23</sup>

To our knowledge, this is the first study to examine both the reliability and validity of methods commonly used to measure urine specific gravity. Reliability of refractometry, hydrometry, and reagent strips was assessed between trials and among testers. Validity of the hydrometer and reagent strips was assessed by comparing them with refractometry, the criterion measure for urine specific gravity.<sup>7-11</sup>

Examination of the method  $\times$  trial  $\times$  tester interaction (see Table 1, Figure 3) reveals that the refractometer was reliable by trial and tester, which was further confirmed by a high intraclass correlation between trials. This is not surprising because determining the urine specific-gravity value using a refractometer is very objective. Hydrometer readings were consistent between trials for each tester but were not consistent among testers (see Table 1, Figure 4). This may reflect tester subjectivity and the difficulty in determining the density-indicating meniscus. Finally, analysis of the data in Table 1 and Figure 5 reveals that reagent-strip measurements were not consistent between trials for the testers or among testers. This was expected because a key disadvantage of reagent strips is that the visual interpretation of color change on the reagent strip often is difficult and very subjective.

Previous researchers have focused on the validity of hy-

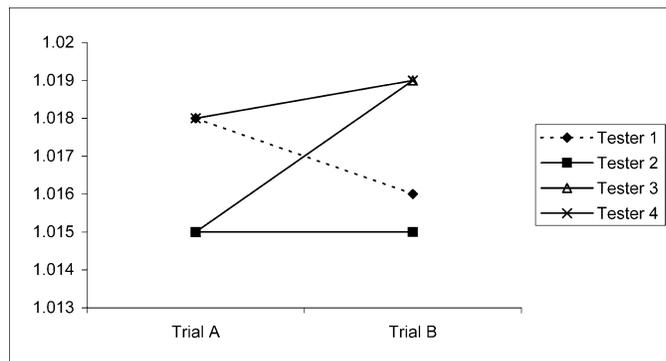


Figure 5. Trial-by-tester interaction for reagent strips.

drometry and reagent strips compared with refractometry, rather than on the reliability of the 3 methods. McCrossin and Roy<sup>12</sup> reported an overall correlation of .96 between hydrometry and refractometry in 69 urine samples from hospitalized children. However, the correlation coefficient decreased as the urine specific gravity increased. Several investigators<sup>13-15</sup> have suggested that reagent strips are an acceptable alternative to refractometry. Gounden and Newall<sup>14</sup> compared results using reagent strips with 2 refractometers in 12 normal subjects and reported correlations of .906 and .911. Guthrie et al<sup>15</sup> collected urine samples from 279 hospital outpatients and found  $r = .88$  between reagent strips and refractometry when no correction was made for the presence of glucose in the urine and  $r = .92$  when a correction was made for glucose. They suggested that the correction for glucose was appropriate but probably academic. Finally, scientists from Miles Laboratories,<sup>13</sup> which produces N-MULTISTIX reagent strips, compared their reagent strips with refractometry in 791 nonhospitalized subjects. They reported a correlation of .796 between the reagent strips and refractometry and reported that the reagent strips were not affected by glucose in the urine, but that urine protein may have had an effect. In contrast, other researchers<sup>12,16-19</sup> assessing the validity of reagent strips with refractometry have reported that the reagent strips are not an acceptable alternative to refractometry. McCrossin and Roy<sup>12</sup> and Adams<sup>16</sup> reported correlation coefficients of  $r = .82$  and  $r = .80$ , respectively, between reagent strips and refractometry and concluded that the reagent strips were not a valid measure of urine specific gravity. Similar conclusions were made by Zack ( $r = .791$ )<sup>19</sup> and Brandon ( $r = .7246$ )<sup>17</sup> when comparing reagent-strip and refractometer measurements. Finally, Dorizzi and Caputo<sup>18</sup> compared 2 different reagent strips with refractometry in 1725 urine samples from hospital inpatients and outpatients and reported correlation coefficients of only  $r = .663$  and  $r = .514$ . In addition to concluding that reagent strips cannot replace refractometry, these researchers also reported that urine glucose and protein had no effect on reagent-strip specific-gravity measurements.

Investigators in the previously described validity studies all compared hydrometry or reagent strips with refractometry using interclass Pearson correlation coefficients. However, interclass correlations may not be the appropriate statistical technique because the same variable is being correlated. It has been suggested that, when multiple tests are given for the same variable, intraclass correlation coefficients should be used.<sup>20</sup> Furthermore, Bland and Altman<sup>21</sup> have proposed that the best statistical approach when comparing measurement devices is the Bland-Altman plot. This approach plots the differences between 2 mea-

surement devices against the mean of the measurement devices. If the measurements using the 2 devices are comparable, the differences on the plot should be small and centered on zero.

Using a variety of statistical approaches (analysis of variance, intraclass correlation coefficients, Bland-Altman plots, and the calculation of false-positive and false-negative readings), we found that neither hydrometry nor reagent strips was a valid measure of urine specific gravity in collegiate wrestlers.

Hydrometer and reagent-strip measurements were significantly greater than refractometer measurements, a finding also shown with the Bland-Altman plots (see Figures 1 and 2). In both plots, data points occurred outside the 95% limits of agreement, which potentially influenced the analysis-of-variance results (decreased methods effect sizes) and the intraclass correlation coefficients (decreased correlation coefficients). Intraclass correlations between reagent strips or hydrometry and refractometry ranged from low to moderate. Because urine specific gravity determines a wrestler's hydration status during the weight-certification process, consideration of false-positive and false-negative hydrometer and reagent-strip results takes on added importance. False positives occurred with both the hydrometer (28%) and reagent strips (15%). These wrestlers were euhydrated, but the hydrometer or reagent-strip measurement indicated that they were dehydrated, negating establishment of a weight class. False negatives occurred to a lesser extent with the hydrometer (2%) and reagent strips (9%). These wrestlers were dehydrated, but the hydrometer or reagent-strip measurement indicated euhydration, allowing the establishment of a weight class. Because this would be determined with the wrestler dehydrated, the weight class for that wrestler would be set incorrectly low. This is a health risk for the wrestler, which is the very situation the NCAA was trying to prevent with the establishment of the new wrestling rules. Although it is difficult to directly compare our results with the findings of others because of different statistical approaches, our overall conclusion that hydrometry and reagent strips are not acceptable substitutes for refractometry is consistent with the findings of other validity studies discussed previously.<sup>12,16-19</sup>

A limitation of our study is that none of the subjects was severely dehydrated, so we did not assess the reliability and validity of the 3 methods in this situation. As described previously, McCrossin and Roy<sup>12</sup> found that the correlation between refractometry and hydrometry diminished in the high range of urine specific-gravity measurements. Furthermore, we did not determine if the urine samples contained glucose or protein. Although some researchers have suggested that glucose<sup>15</sup> or protein<sup>13</sup> may affect urine specific-gravity measurements with reagent strips, others have concluded that reagent-strip measurements are not affected by the presence of glucose<sup>13,18</sup> or protein,<sup>18</sup> as discussed previously.

In conclusion, our results suggest that the refractometer is a reliable measure of urine specific gravity. In contrast, the hydrometer and reagent strips were not reliable, nor were they valid measures of urine specific gravity when compared with refractometry. When the new wrestling rules were first introduced in 1998,<sup>4</sup> the NCAA allowed the use of all 3 methods to assess urine specific gravity. Subsequently, the use of reagent strips was eliminated.<sup>6</sup> Our data support this decision and further suggest that the use of the hydrometer also should be eliminated. The refractometer is reliable, fast, accurate, and technically easy to use and requires only a single drop of urine. Therefore, we suggest that refractometry should be the NCAA's only choice for measuring a collegiate wrestler's urine

specific gravity during the weight-certification process. Future research is warranted to assess whether a urine specific-gravity measurement of  $\leq 1.020$  as selected by the NCAA is an appropriate cut-off value to indicate euhydration.

## ACKNOWLEDGMENTS

We thank Sharon Birch for her statistical expertise and Michael Cantele, Joseph Donolli, and Kristin Petrovia for their assistance in data collection.

## REFERENCES

1. Benson M. Committee refines wrestling safety rules. *NCAA News*. April 20, 1998;35:1,14.
2. Hunt V. Sport pins down new regulations to help ensure safe competition. *NATA News*. October 1998;10-12.
3. Schnirring L. NCAA strengthens weight-cutting rules. *Physician Sports-med*. 1998;26(7):16.
4. National Collegiate Athletic Association Wrestling Rules Committee. 1998-1999 wrestling weight certification program. Indianapolis, IN: National Collegiate Athletic Association; 1998.
5. Dirckx JH. *Stedman's Concise Medical Dictionary for the Health Professions*. 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2001.
6. National Collegiate Athletic Association Wrestling Rules Committee. 1999-2000 Wrestling weight certification program. Indianapolis, IN: National Collegiate Athletic Association; 1999.
7. Ross DL, Neely AE. *Textbook of Urinalysis and Body Fluids*. Norwalk, CT: Appleton-Century-Crofts; 1983.
8. Armstrong LE, Maresh CM, Castellani JW, et al. Urinary indices of hydration status. *Int J Sport Nutr*. 1994;4:265-279.
9. Armstrong LE, Soto JAH, Hacker FT Jr, Casa DJ, Kavouras SA, Maresh CM. Urinary indices during dehydration, exercise, and rehydration. *Int J Sport Nutr*. 1998;8:345-355.
10. Casa DJ, Armstrong LE, Hillman SK, et al. National Athletic Trainers' Association position statement: fluid replacement for athletes. *J Athl Train*. 2000;35:212-224.
11. Popowski LA, Oppliger RA, Patrick Lambert G, Johnson RF, Kim Johnson A, Gisolfi CV. Blood and urinary measures of hydration status during progressive acute dehydration. *Med Sci Sports Exerc*. 2001;33:747-753.
12. McCrossin T, Roy LP. Comparison of hydrometry, refractometry, osmometry, and Ames N-Multistix SG in estimation of urinary concentration. *Aust Paediatr J*. 1985;21:185-188.
13. Burkhardt AE, Johnston KG, Waszak CE, Jackson CE, Shafer SR. A reagent strip for measuring the specific gravity of urine. *Clin Chem*. 1982; 28:2068-2072.
14. Gounden D, Newall RG. Urine specific gravity measurements: comparison of a new reagent strip method with existing methodologies, as applied to the water concentration/dilution tests. *Curr Med Res Opin*. 1983;8:375-381.
15. Guthrie RM, Lott JA, Kriesel S, Miller IL. Does the dipstick meet medical needs for urine specific gravity? *J Fam Pract*. 1987;25:512-514.
16. Adams LJ. Evaluation of Ames Multistix-SG for urine specific gravity versus refractometer specific gravity. *Am J Clin Pathol*. 1983;80:871-873.
17. Brandon CA. Urine specific gravity measurement: reagent strip versus refractometer. *Clin Lab Sci*. 1994;7:308-310.
18. Dorizzi RM, Caputo M. Measurement of urine relative density using refractometer and reagent strips. *Clin Chem Lab Med*. 1998;36:925-928.
19. Zack JF Jr. Evaluation of a specific gravity test strip. *Clin Chem*. 1983; 29:210.
20. Thomas JR, Nelson JK. *Research Methods in Physical Activity*. 3rd ed. Champaign, IL: Human Kinetics; 1996.
21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-310.
22. Donnersberger A, Lesak A. *A Laboratory Textbook of Anatomy and Physiology*. 7th ed. Sudbury, MA: Jones and Bartlett Publishers; 2000.
23. Vincent WJ. *Statistics in Kinesiology*. 2nd ed. Champaign, IL: Human Kinetics; 1999.
24. Bland M. *An Introduction to Medical Statistics*. 3rd ed. Oxford, UK: Oxford University Press; 2000.