Agreement of Urine Specific Gravity Measurements Between Manual and Digital Refractometers

Dawn M. Minton, MS, ATC*; Eric Kyle O'Neal, PhD†; Toni Marie Torres-McGehee, PhD, ATC*

*Department of Physical Education and Athletic Training, University of South Carolina, Columbia; †Department of Health, Physical Education, and Recreation, University of North Alabama, Florence

Context: Urine specific gravity (U_{sg}), measured by a handheld manual refractometer (MAN), has been recognized as a valid and practical means of assessing hydration status. Newer, digital refractometers are faster and more user friendly but have not been validated against the traditional MAN.

Objective: To compare the reliability and validity of 2 digital refractometer models and a MAN.

Design: Descriptive laboratory study.

Setting: Research laboratory.

Patients or Other Participants: Sample of convenience was recruited from the local university and surrounding community ($n = 82$).

Intervention(s): Participants provided multiple urine samples ($n = 124$) over a 5-month period under various hydration conditions.

Main Outcome Measure(s): Urine specific gravity was compared among a MAN, a digital refractometer requiring the prism to be dipped (DIP) into a urine sample, and a digital refractometer that requires urine to be pipetted (PIP) onto its prism for analysis.

Results: The MAN measurements were strongly correlated with the DIP ($r = 0.99$, $P < .001$) and PIP ($r = 0.97$, $P < .001$) measurements. Bland-Altman analyses revealed slight mean underestimation (95% upper and lower levels of agreement) between MAN and DIP (-0.0012 [0.0028] and PIP -0.0011 [0.0035], respectively) and trends toward increased underestimation at higher U_{sg} . Measurement error $\geq .005$ was greater for PIP (4/124, 3.2%) than for DIP (2/124, 1.6%).

Conclusions: Negligible differences were exhibited between PIP and DIP, with both displaying acceptable reliability and validity compared with the MAN. However, the Bland-Altman analysis suggests underestimation bias for the DIP and PIP as U_{sg} increases, with the potential for rare but substantial underestimation when using PIP that should be recognized by clinicians, particularly when used as a screening measure in weight-class sports.

Key Words: hydration assessment, reliability, validity, weight-class sports

Key Points

- Compared with a traditional manual refractometer, digital refractometers offer a practical, easy-to-use measure of urine specific gravity.
- Digital refractometers may underestimate hydration status, especially as dehydration increases. This potential underestimation should be recognized when precise urine specific gravity measures are needed, as in weight-class sports.

Accurately assessing hydration status is an important factor in detecting hypohydration and preventing subsequent performance deficits, such as decreased time to fatigue,^{1,2} increased rate of perceived exertion,³ decreased resistance-exercise performance,⁴ increased thermal strain,^{5,6} and increased cardiovascular strain.^{7,8} Plasma osmolality is sometimes referred to as the criterion measure of hydration status; however, it requires collecting a blood sample, which must be centrifuged and analyzed using an expensive osmometer, and the results can be influenced by a variety of factors.⁹ Practical methods to monitor hydration status should be noninvasive and require little technical expertise. Additionally, to be useful in clinical settings, assessment instruments should be portable and inexpensive and provide results in a short amount of time. Therefore, the National Athletic Trainers' Association and American College of Sports Medicine advocate educating athletes to monitor hydration status through body mass changes and

urine color.^{10,11} These are also the most common methods used to measure hydration status in athletes.^{9,10} Another measurement that clinicians commonly use to assess acute hydration status is urine specific gravity (U_{sg}).¹⁰

Urine specific gravity refers to the density (mass per volume) of a sample in comparison with distilled water.^{12,13} Urinary concentration is determined by the number of particles (electrolytes, phosphate, urea, uric acid, proteins, glucose, and radiographic contrast media) per unit of urine volume with consideration for urine temperature.^{13,14} Fluid denser than water has a specific gravity > 1.000 .¹² Urine reagent strips and urinometers have been used to assess U_{sg} ; however, refractometry is the only technique that has consistently been shown to be highly correlated with urine osmolality,^{15–20} a more technical and expensive hydration-assessment technique.

Provided no errors are made by the measurer, the traditional manual clinical refractometer (MAN) offers a

Table 1. Comparisons of Agreement Between Measurements (n = 124) With the Refractometers

Refractometers Compared	Spearman ρ Correlation			Bland-Altman Mean Bias $\pm$ 95% Levels of Agreement	Samples With Significant Error ^b
	<i>r</i>	<i>R</i> ²	<i>P</i> Value ^a		
Manual and pen-like digital	0.99	0.99	<.001	−0.0012 $\pm$ 0.0028	1.6% (2/124)
Manual and pipetted digital	0.97	0.94	<.001	−0.0011 $\pm$ 0.0035	3.2% (4/124)
Pen-like digital and pipetted digital	0.97	0.94	<.001	NA	NA

Abbreviation: NA, not available.

^a Significant at $P < .05$.

^b Significant error was defined as a difference of 0.005 or greater from the reading of the manual refractometer.

fairly quick (<1 minute per sample) and precise reading of U_{sg} . Possible user errors include poor calibration, sample contamination, not cleaning the daylight plate or prism surface completely between samples, and misreading the metered scale displayed in the viewfinder. Digital refractometers have been introduced as a more convenient method to measure U_{sg} that decreases the time needed to assess the value. The automaticity of the digital refractometer reduces the risk of some of the potential user errors. Yet we are unaware of any published studies that have examined the validity of digital refractometers compared with the MAN. The purpose of this research was to compare the reliability and validity of 2 popular digital refractometers and a MAN.

METHODS

Participants

We recruited a sample of convenience from a local university in the southeastern United States ($n = 82$; males = 37, females = 45; age range = 18–40 years). Participants included students, faculty, and athletes who were approached by a member of the investigative team and asked if they could provide a urine sample. No instructions were provided to participants regarding hydration practices. We sought heterogeneous and spontaneous samples from individuals in their natural environment throughout the day. Participants, in various hydration conditions, provided either 1 or multiple urine samples over 5 months. There were no inclusion or exclusion criteria for participation in this study. Institutional review board approval was obtained from the primary author's institution (University of South Carolina, Columbia), and participants read and signed an approved informed consent before the study.

Instruments and Procedures

We used 2 popular digital refractometers and a MAN in this investigation. The first digital type features a “pen-like” body (DIP; model PEN-PRO; Atago Co, Ltd, Tokyo, Japan) with the prism located at the tip of the refractometer, requiring the tester to dip the prism into a sample of urine for analysis. The second digital refractometer technique requires a urine sample ≥ 0.3 mL to be pipetted (PIP; model PAL-10S; Atago Co, Ltd) on the prism. The measurement values for both digital refractometers are available in less than 3 seconds, allowing samples to be tested rapidly, and both devices automatically control for sample temperature.^{21,22} The MAN (model REF312; Atago Co, Ltd) requires a urine sample to be pipetted onto the prism, the cover plate to be closed, and the prism pointed toward a

light source. The researcher looks through a viewfinder at a metered scale to determine the specific gravity of the sample. Each urine sample was analyzed on each refractometer according to the manufacturer's instructions by the same researcher in duplicate, with the average of the 2 measures being used for analysis. Before use, each refractometer was calibrated with distilled water, and samples were analyzed immediately after collection.

Statistical Analyses

We calculated Spearman ρ correlations to determine relationships among the 3 refractometer models. Bland-Altman analysis was used to determine mean estimation bias and direction and 95% limits of agreement among refractometers. The SPSS statistical software (version 19.0; IBM Corp, Armonk, NY) was used for all statistical analysis. Additionally, the diagnostic validities of the DIP and PIP were compared with the MAN, which served as the criterion measurement; a U_{sg} value of 1.020 was the upper limit classification of euhydration. Sensitivity and specificity were depicted in a receiver operating characteristic (ROC) curve and a contingency table. An α level of .05 was used to determine significance for all analyses.

RESULTS

We assessed a total of 124 urine samples using all 3 refractometers. Correlations between each of the refractometer comparisons were strong and significant (Table 1). Mean error bias was minimal for both the DIP and PIP; however, examination of individual data revealed multiple samples with significant error (ie, difference ≥ 0.005) for both digital refractometers (Table 1).

Results for the Bland-Altman analysis can be seen in Figures 1 and 2. The downward slopes of the lines of best fit for the DIP (Figure 1; $r = -0.89$; $P < .001$) and PIP (Figure 2; $r = -0.84$; $P < .001$) show underestimation bias for the digital refractometers, with increasing U_{sg} values compared with the MAN. Underestimation bias was particularly prevalent when U_{sg} approached 1.025. However, the tight 95% upper and lower level-of-agreement lines suggest that the amount of error (± 0.003) would not make a significant practical difference when a clinician was simply trying to determine an athlete's general hydration state in a field environment. Conversely, the disagreement between refractometers is great enough to advocate that the MAN would be preferable in laboratory settings when greater precision is desired.

Accuracy is also important when determining hydration status for participation eligibility in weight-class competitions such as wrestling. The agreement between MAN

Figure 1. Bland-Altman analysis for measurements from the pen-like digital refractometer (DIP) compared with the manual refractometer (MAN). Number of urine samples analyzed = 124. x = Sample (n = 6) in which the DIP indicated euhydration (≤ 1.020) when the MAN indicated hypohydration.

Figure 2. Bland-Altman analysis for the pipetted digital refractometer (PIP) compared with the manual refractometer (MAN). Number of urine samples used for analysis = 124. x = Sample (n = 7) in which the PIP indicated euhydration (< 1.020) when the MAN indicated hypohydration. o = Sample (n = 1) in which the PIP indicated hypohydration (> 1.020) when the MAN indicated euhydration.

Table 2. Contingency Table for Agreement of Measurements from the Manual Refractometer and Digital Refractometers^a (n = 124)

Manual Refractometer	Pen-Like Digital Refractometer			Pipetted Digital Refractometer		
	Euhydration	Dehydration	Total (n)	Euhydration	Dehydration	Total (n)
Euhydration	True positive 59.7% (n = 74)	False positive 0% (n = 0)	74	True positive 58.9% (n = 73)	False positive 0.8% (n = 1)	74
Dehydration	False negative 4.8% (n = 6)	True negative 35.5% (n = 44)	50	False negative 5.6% (n = 7)	True negative 34.7% (n = 43)	50
Total (n)	80	44	124	80	44	124

^a Based on the classification that ≤ 1.020 = euhydration and > 1.020 = dehydration.

compared with the DIP and PIP using $U_{sg} \leq 1.020$ as the definition of *euhydration* is shown in Table 2.^{10,11} Less than 5% and 7% of the DIP and PIP samples, respectively, produced classification inaccuracies compared with MAN. All but 1 misclassification erred on the side of underestimating dehydration with the DIP and PIP in comparison with MAN, which are denoted with special data symbols in Figures 1 and 2. The specificity and associated sensitivity are also displayed in the ROCs in Figure 3. The ROCs represent the predicted relationship between true-positive (ie, individuals classified as dehydrated by PIP or DIP and confirmed by MAN) and false-positive rates (ie, individuals classified as euhydrated by PIP and DIP but dehydrated by MAN). On an ROC curve, the straight line from the lower left corner to the upper right corner represents a diagnostic tool that was no better a predictor than pure chance (ie, 50% chance of correct diagnosis). Any portion of the curve that passes underneath this line represents a test that was a worse predictor. The closer the ROC moves to the upper left corner of the chart, the greater the chances are that the test will represent the correct diagnosis. The ROCs for both DIP and PIP demonstrate the strong prediction performance of each technique in estimating hydration status consistent with the standard MAN technique. Further, the DIP and PIP methods, which provided ROC areas of 0.999 and 0.992,

respectively, did not differ in their diagnostic performance ($P = .20$).²³

DISCUSSION

Athletic trainers are encouraged to monitor hydration status through body mass changes and U_{sg} .¹⁰ Monitoring hydration status may reduce the risk of developing heat illnesses,¹⁰ decrease performance decrements associated with hypohydration,¹⁻⁸ and prevent disqualification for weight-categorized sports.^{14,24,25} Researchers^{10,13,15,16,19,24} have repeatedly demonstrated that refractometry is the gold standard for measuring U_{sg} compared with other instruments, such as urinometers or reagent strips. Traditionally, refractometry has been assessed using a MAN; however, this method adds subjectivity to the measure and lends itself to possible user errors. Digital refractometers mitigate or eliminate many of these error sources and allow for numerous U_{sg} measurements to be assessed quickly when turnover time is critical, such as during registration in weight-class sport competitions. Digital refractometers are more costly than the MAN, and limited data²⁶ are available on the reliability and validity between the MAN and digital refractometers. With their rapid result generation and user-friendly design, digital refractometers have become increasingly popular. Our purpose was to investigate the validity of 2 digital refractometers compared with a MAN.

Measurements from both the DIP and PIP were strongly correlated with those from the MAN and displayed acceptable levels of mean error bias (Table 1), supporting the manufacturer's accuracy claim of ± 0.001 .^{21,22} However, despite the strong correlations, as U_{sg} measures increased (ie, indicating greater levels of dehydration), our results showed a greater tendency for both digital refractometers, and particularly the PIP, to underestimate U_{sg} compared with the MAN (Figures 1 and 2). Although inconsistencies between the DIP and the MAN were evident in our investigation, they appeared to be less pronounced than noted by Niemann,²⁶ who found that the DIP measurement was correlated with urine osmolality ($r = 0.81$) but not as strongly as with the MAN measurement ($r = 0.94$). Not unexpectedly, Niemann's results²⁶ and ours demonstrate that discrepancies in measurement of U_{sg} will be found between MAN and digital refractometers.

Although few researchers have studied the use of digital refractometers in human populations, comparisons within veterinary medicine have been investigated. Bennett et al²⁷ found that a MAN refractometer measured significantly higher feline U_{sg} than a digital model, with a mean difference of 0.003 indicated by Bland-Altman analysis. Similar results were found in canines, where MAN

Figure 3. Receiver operating characteristic curve for the pen-like digital refractometer (DIP) and the pipetted digital refractometer (PIP) in comparison with the manual refractometer. The curves are presented together for comparison. There was no difference ($P = .20$) in area under the curve (AUC) for the DIP (AUC = 0.999) and PIP (AUC = 0.992). The specificity and sensitivity for the DIP = 1.0 and 0.93 and PIP = 0.98 and 0.91, respectively.

measurements were higher than digital measurements (mean difference = 0.001, $P < .001$).²⁸ Additionally, measurements from both MAN and the digital refractometers were strongly correlated with osmolality in canines ($r = 0.98$ and $r = 0.98$, respectively)²⁸ and felines ($r = 0.96$ and $r = 0.97$, respectively).²⁷ Although animal urine is not identical to human urine, these studies show statistical differences between the measurements obtained with the MAN versus the digital refractometers.

For athletic trainers who require precise measurements (ie, clinicians working with wrestlers and researchers using hydration indices as outcome measures), the underestimation and potential error from digital refractometers should be concerns. The ROC analysis reveals excellent diagnostic agreement between measurements obtained from the MAN and the digital refractometers (Figure 3), but approximately 5% of the samples for both the PIP and DIP were misclassified when we used a MAN U_{sg} marker of 1.020 to define euhydration. All but 1 misclassification erred on the side of underestimation by the digital refractometers, which was not unexpected given the bias in agreement exhibited in the Bland-Altman analyses (Figures 1 and 2). Clinically, these results suggest that a small percentage of athletes who would be classified as hypohydrated with the MAN would not be classified as hypohydrated with the digital refractometers.

We experienced some problems using each of the digital refractometers. First, despite following the manufacturer's directions, we sometimes had to make several attempts before obtaining a PIP reading. This appeared to be due to the amount of urine on the prism or the brightness of lighting in different areas of the laboratory (ie, despite being in the same laboratory, the refractometer showed a lighting error, which required the researcher to change the angle of the prism so that it was directly under the fluorescent lighting). Second, the DIP requires a greater amount of sample in the specimen cup for the prism to be adequately covered, compared with the 2 to 3 drops needed by the MAN or PIP. This is a drawback when measurements are needed from individuals who are hypohydrated and cannot provide a substantial amount of urine. As with the PIP, we experienced difficulty obtaining a reading from the DIP when measuring in areas of the laboratory with different types of lighting.

Limitations

The major limitation of our investigation and of previous research²⁶ is that we were unable to compare U_{sg} measures with plasma osmolality, which some consider a more accurate hydration criterion reference.²⁹ As such, we were able to compare only the results from the DIP and PIP with those from the MAN, which is a valid and reliable method to assess hydration status when plasma osmolality cannot be determined.^{13,29} Additionally, we did not control for the presence of glucose, protein, or blood in the urine, which may alter U_{sg} results¹³; however, because we were comparing 3 refractometer measures, we believe that this factor had little effect on our results.

Future Research

Urine specific gravity is suggested to lag behind plasma measures when significant alterations in hydration status

occur over a brief time period.²⁹ This lag could explain the differences between U_{sg} and plasma osmolality if fluid intake and hydration status are not controlled. Therefore, controlled laboratory studies comparing U_{sg} measures in euhydrated and hypohydrated participants should be conducted. Future researchers should also examine the reliability and validity of digital refractometers compared with plasma osmolality in a variety of clinical and field settings. For example, athletic trainers may measure U_{sg} in locker rooms, bathrooms, or gyms, where lighting is poor and other confounding factors (eg, hydration status, diet, supplements) may alter the accuracy of these measures.

CONCLUSIONS

Our study was unique in that we examined measurements from 2 digital refractometer models that have not, to our knowledge, previously been validated in the literature. Overall, both the DIP and PIP measurements were reliable and valid compared with the MAN measurements; the DIP displayed a trend toward less error and misclassification of hydration status. Despite a strong correlation, the Bland-Altman analysis suggests a bias for the DIP and PIP measurements as U_{sg} increases. Clinicians using the PIP or DIP should be aware of potential error during hydration assessment, especially when determining qualification for participation (ie, wrestling).

ACKNOWLEDGMENTS

We thank Matt Laurent and Adam Fullenkamp for their help in analyzing the data.

REFERENCES

1. Bigard AX, Sanchez H, Claveyrolas G, Martin S, Thimonier B, Arnaud MJ. Effects of dehydration and rehydration on EMG changes during fatiguing contractions. *Med Sci Sports Exerc.* 2001;33(10):1694–1700.
2. Montain SJ, Smith SA, Mattot RP, Zientara GP, Jolesz FA, Sawka MN. Hypohydration effects on skeletal muscle performance and metabolism: a ³¹P-MRS study. *J Appl Physiol.* 1998;84(6):1889–1894.
3. Sawka MN, Coyle EF. Influence of body water and blood volume on thermoregulation and exercise performance in the heat. *Exerc Sport Sci Rev.* 1999;27:167–218.
4. Judelson DA, Maresh CM, Farrell MJ, et al. Effect of hydration state on strength, power, and resistance exercise performance. *Med Sci Sports Exerc.* 2007;39(10):1817–1824.
5. Sawka MN, Young AJ, Francesconi RP, Muza SR, Pandolf KB. Thermoregulatory and blood responses during exercise at graded hypohydration levels. *J Appl Physiol.* 1985;59(5):1394–1401.
6. Montain SJ, Coyle EF. Influence of graded dehydration on hyperthermia and cardiovascular drift during exercise. *J Appl Physiol.* 1992;73(4):1340–1350.
7. Casa DJ. Exercise in the heat, I: fundamentals of thermal physiology, performance implications, and dehydration. *J Athl Train.* 1999;34(3):246–252.
8. Casa DJ, Stearns RL, Lopez RM, et al. Influence of hydration on physiological function and performance during trail running in the heat. *J Athl Train.* 2010;45(2):147–156.
9. Armstrong LE. Assessing hydration status: the elusive gold standard. *J Am Coll Nutr.* 2007;26(suppl 5):S575–S584.
10. Casa DJ, Armstrong LE, Hillman SK, et al. National Athletic Trainers' Association position statement: fluid replacement for athletes. *J Athl Train.* 2000;35(2):212–224.

11. Sawka MN, Burke LM, Eichner ER, Maughan RJ, Montain SJ, Stachenfeld NS. American College of Sports Medicine position stand: exercise and fluid replacement. *Med Sci Sports Exerc.* 2007;39(2):377–390.
12. Armstrong LE. Hydration assessment techniques. *Nutr Rev.* 2005;63(6, pt 2):S40–S54.
13. Chadha V, Garg U, Alon US. Measurement of urinary concentration: a critical appraisal of methodologies. *Pediatr Nephrol.* 2001;16(4):374–382.
14. Oppliger RA, Bartok C. Hydration testing of athletes. *Sports Med.* 2002;32(15):959–971.
15. Armstrong LE, Maresh CM, Castellani JW, et al. Urinary indices of hydration status. *Int J Sport Nutr.* 1994;4(3):265–279.
16. Armstrong LE, Soto JA, Hacker FT Jr, Casa DJ, Kavouras SA, Maresh CM. Urinary indices during dehydration, exercise, and rehydration. *Int J Sport Nutr.* 1998;8(4):345–355.
17. Armstrong LE, Pumerantz AC, Fiala KA, et al. Human hydration indices: acute and longitudinal reference values. *Int J Sport Nutr Exerc Metab.* 2010;20(2):145–153.
18. Costa CE, Bettendorff C, Bupo S, Ayuso S, Vallejo G. Comparative measurement of urine specific gravity: reagent strips, refractometry, and hydrometry. *Arch Argent Pediatr.* 2010;108(3):234–238.
19. Eberman LE, Minton DM, Cleary MA. Comparison of refractometry, urine color, and urine reagent strips to urine osmolality for measurement of urinary concentration. *Athl Train Sports Health Care.* 2009;1(6):267–271.
20. Oppliger RA, Magnes SA, Popowski LA, Gisolfi CV. Accuracy of urine specific gravity and osmolality as indicators of hydration status. *Int J Sport Nutr Exerc Metab.* 2005;15(3):236–251.
21. PAL-10S Digital Hand-Held “Pocket” Refractometer, version 4410-E05 [instruction manual]. Tokyo, Japan: Atago Co, Ltd; 2010.
22. PEN Wrestling Refractometer, version 3749-E01 [instruction manual]. Tokyo, Japan: Atago Co, Ltd; 2009.
23. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* 1983;148(3):839–843.
24. Stuempfle KJ, Drury DG. Comparison of 3 methods to assess urine specific gravity in collegiate wrestlers. *J Athl Train.* 2003;38(4):315–319.
25. Bubb RG, Smith T. *National Collegiate Athletic Association 2011–2012 and 2012–2013 Wrestling Rules and Interpretations.* Indianapolis, IN: National Collegiate Athletic Association; 2011.
26. Niemann AJ. *The Effect of Instrument Type on the Measure of Hydration Status* [master’s thesis]. Terre Haute: Indiana State University; 2012.
27. Bennett AD, McKnight GE, Dodkin SJ, Simpson KE, Schwartz AM, Gunn-Moore DA. Comparison of digital and optical hand-held refractometers for the measurement of feline urine specific gravity. *J Feline Med Surg.* 2011;13(2):152–154.
28. Paris JK, Bennett AD, Dodkin SJ, Gunn-Moore DA. Comparison of a digital and an optical analogue hand-held refractometer for the measurement of canine urine specific gravity. *Vet Rec.* 2012;170(18):463.
29. Popowski LA, Oppliger RA, Lambert GP, Johnson RF, Johnson AK, Gisolfi CV. Blood and urinary measures of hydration status during progressive acute dehydration. *Med Sci Sports Exerc.* 2001;33(5):747–753.

Address correspondence to Dawn M. Minton, MS, ATC, Department of Physical Education and Athletic Training, University of South Carolina, Blatt PE Center, Room 218, Columbia, SC 29208. Address e-mail to dawn.minton@my.athens.edu.