

## Comparison of canine urine specific gravity measurements between various refractometers in a clinical setting

Kelly du Preez BSc BVSc (Hons)

Section of Clinical Pathology, Department of Companion Animal Clinical Studies, Faculty of Veterinary Science, University of Pretoria, Onderstepoort 0110, Pretoria, South Africa.

Keagan Boustead BSc BVSc

Department of Companion Animal Clinical Studies, Faculty of Veterinary Science, University of Pretoria, Onderstepoort 0110, Pretoria, South Africa.

Yolandi Rautenbach BVSc (Hons) MMedVet

Section of Clinical Pathology, Department of Companion Animal Clinical Studies, Faculty of Veterinary Science, University of Pretoria, Onderstepoort 0110, Pretoria, South Africa.

Amelia Goddard BVSc (Hons) MMedVet PhD

Section of Clinical Pathology, Head of Department of Companion Animal Clinical Studies, Faculty of Veterinary Science, University of Pretoria, Onderstepoort 0110, Pretoria, South Africa.

Emma H Hooijberg BVSc PhD DipECVCP

Section of Clinical Pathology, Department of Companion Animal Clinical Studies, Faculty of Veterinary Science, University of Pretoria, Onderstepoort 0110, Pretoria, South Africa.

### Correspondence

Kelly du Preez BSc BVSc (Hons)

Institution address: Department of Companion Animal Clinical Studies, Faculty of Veterinary Science, Private Bag X04, Onderstepoort, Pretoria, South Africa, 0110.

Email address: [kelly.dupreez@up.ac.za](mailto:kelly.dupreez@up.ac.za)

## Abstract

**Background:** Veterinary facilities might use multiple refractometers and individuals to measure urine specific gravity (USG). Previous comparison studies show conflicting results. Furthermore, the clinical significance of measurement differences and interobserver variabilities have not been assessed.

**Objectives:** We aimed to determine statistically and clinically significant differences between four refractometers in measuring canine USG and subsequent categorization of urine concentrations and azotemia and determine the variability between different observers performing USG measurements.

**Methods:** Fifty-nine specimens were included for the USG measurements with four refractometers by different observers. Each refractometer pair was compared using Spearman's rank correlation, Bland-Altman difference plots, and Deming regression analyses. Calculated bias was compared to set performance goals. Interobserver agreement was evaluated, and intraclass correlation coefficients were used to determine differences in the categorization of urine concentrations and azotemia (pre-renal or renal).

**Results:** There was excellent correlation ( $r = 0.99 - 1.00$ ) between refractometers. All comparisons involving R4 showed significant constant and proportional biases. Mean bias met the clinical performance goal for all refractometers, except for comparisons with R4, where up to 17 results were outside the allowable bias. There was almost perfect agreement ( $r = 0.999$ ) between observers and excellent agreement ( $r = 0.96 - 0.99$ ) for the classification of urine concentration. In azotemic patients (22%), there was perfect agreement ( $r = 1.00$ ) for the categorization of azotemia.

**Conclusions:** In most cases, three of the refractometers evaluated in this study can be used interchangeably at all USG values, without affecting clinical decision-making. Multiple observers did not significantly affect decision-making.

**Keywords:** azotemia, dog, interobserver variability, method comparison, urine concentration

## Introduction

Refractometers are commonly used and invaluable tools for assessing urine in clinical practice. Previous studies in dogs have shown the suitability of urine specific gravity (USG) as a proxy for measuring urine concentration, based on its strong agreement with urine osmolality (gold standard).<sup>1-3</sup> The measurement of USG can also assist in differentiating between renal and non-renal azotemia, which helps guide clinicians in their diagnostic workup.<sup>4</sup>

In clinical practice, different refractometers might be used at different time points for analyzing urine from the same patient. Refractometers intended for USG measurements could differ in terms of the manufacturer, means of measurement (optical vs digital), conversion factors, specificity for veterinary species (refractometers produced for use in people vs those developed for use in a particular species, such as in dogs), and measuring range. Furthermore, in multi-veterinarian practices and practices employing veterinary nurses or technical staff, different individuals might read USG measurements in the same patient.

Previous studies have shown conflicting results regarding the appropriateness of using different refractometers interchangeably in dogs.<sup>3,5,6</sup> One study comparing a veterinary digital refractometer (designed for use in cats) with a medical optical refractometer, found an unacceptable negative proportional bias in the digital refractometer when measuring the USG of canine urine.<sup>5</sup> These findings are similar to those of another study comparing a combination of five veterinary (dog- and cat-specific) and medical digital and optical refractometers to a reference method.<sup>6</sup> A negative proportional bias was found in all refractometers compared to the estimated specific gravity from dried total solids and was greatest in the cat-specific refractometers (both optical and digital) when measuring the USG of canine urine.<sup>6</sup> In contrast, a study in dogs only, that compared a medical digital and optical refractometer showed that, in 91.5% of dogs, USG measurements differed by < 0.002 units on the USG scale and concluded that these refractometers could be used interchangeably.<sup>3</sup> Furthermore, these studies either involved only one observer to reduce variability<sup>3</sup>

or cited possible interobserver variability as a study limitation when more than one observer was included.<sup>5,6</sup>

The aims of this study were first, to compare USG readings of four different refractometers using standard method comparison statistics; second, to determine whether any differences were clinically significant and relevant by comparing the refractometers in their classification of urine concentration categories and types of azotemia; and third, to determine the magnitude of variability between observers. We hypothesized that the refractometers would show significantly different readings to each other over a range of USGs and, subsequently, that azotemic patients would be classified differently by the various refractometers. Furthermore, we hypothesized that significant variability would exist between observers.

## Materials and Methods

### Patients

This prospective study was a method comparison study, based on previously published guidelines for method comparison in the clinical laboratory.<sup>7</sup> Urine specimens were collected from canine patients presented to the Onderstepoort Veterinary Academic Hospital (OVAH), University of Pretoria, from May 2017 to September 2018. The dogs were presented with a variety of historical and clinical findings warranting, in all cases, a complete urinalysis and determination of serum creatinine concentration as part of routine diagnostic procedures. No urine specimens or serum samples were collected exclusively for use in this study. There was no selection for breed, age, sex, presenting complaint, or final diagnosis. However, patients receiving fluid therapy or with macroscopic hematuria or pigmenturia at the time of urine and/or blood collection were not included in the study. The study was approved by the Animal Ethics Committee of the University of Pretoria in May 2017 (certificate number V036-17).

## *Comparison of refractometric USG measurement*

### Study Specimens

Urine was collected via free-flow, catheterization, or cystocentesis as part of a routine diagnostic workup. Complete urinalyses were performed in the clinic within 30 minutes of collection and included organoleptic evaluations (color, turbidity, and odor), USG determinations, urine dipstick examinations, and microscopic evaluations of urinary sediments. Surplus urine ( $\geq 1$  mL) from this procedure was submitted to the clinical pathology laboratory for use in the study. Urine specimens were considered for inclusion if a complete urinalysis was performed by the attending clinician, and if serum creatinine concentration was determined within two hours of urine collection. All the results of a complete urinalysis were recorded and submitted with the specimen. Specifically included in this record was the turbidity of the specimen and the dipstick results for protein, glucose, and heme protein. Each submitted specimen was stored at 2 – 8°C for up to 12 hours<sup>8</sup> until the USG could be measured, which was usually at the end of the working day. Blood for serum creatinine concentrations was collected either at the same time or within two hours of urine collections.

### Specimen Analysis

Four refractometers were used to measure the USG (Table 1). All the selected instruments were hand-held optical refractometers used by final-year veterinary and veterinary nursing students and clinicians across various hospital clinics (Figure 1). Additionally, all four refractometers were designed for USG measurements in human urine, and none were veterinary specific. The refractometers were calibrated to 1.000 using distilled water at room temperature before each measurement, as recommended.<sup>9</sup> A group of six observers took part in the study, and for each specimen, two observers from this group were randomly selected to perform the analyses. The observers included certified laboratory technologists, residents, and specialist clinical pathologists. The urine specimen was left to warm to room temperature (22 – 25°C) before analyses. Urine specimens were resuspended, and USGs were determined by placing one to two drops of room temperature, non-centrifuged urine on each of the four refractometers. The USG was measured on

*Comparison of refractometric USG measurement*

each instrument by two independent observers, in the same room, and one after the other, using the same drop(s) of urine. The observers were blinded to the USG measured in the complete urinalyses and to the USG measurements determined by the other observer. Concentrated urine specimens with a USG beyond the analytical range of the refractometer were assigned the value of the upper limit of the measurement. Dilution of concentrated urine specimens reportedly yields spuriously elevated USG values, and for this reason, dilutions were not performed.<sup>6</sup> Serum creatinine concentrations were measured with a modified Jaffe method using an automated chemistry analyzer (Cobas Integra 400 Plus, Roche Products [Pty] Ltd., Basel, Switzerland). Serum creatinine concentrations measured beyond the analytical range of the assay (> 1300 µmol/L or > 14.7 mg/dL) was achieved using automatic dilutions.

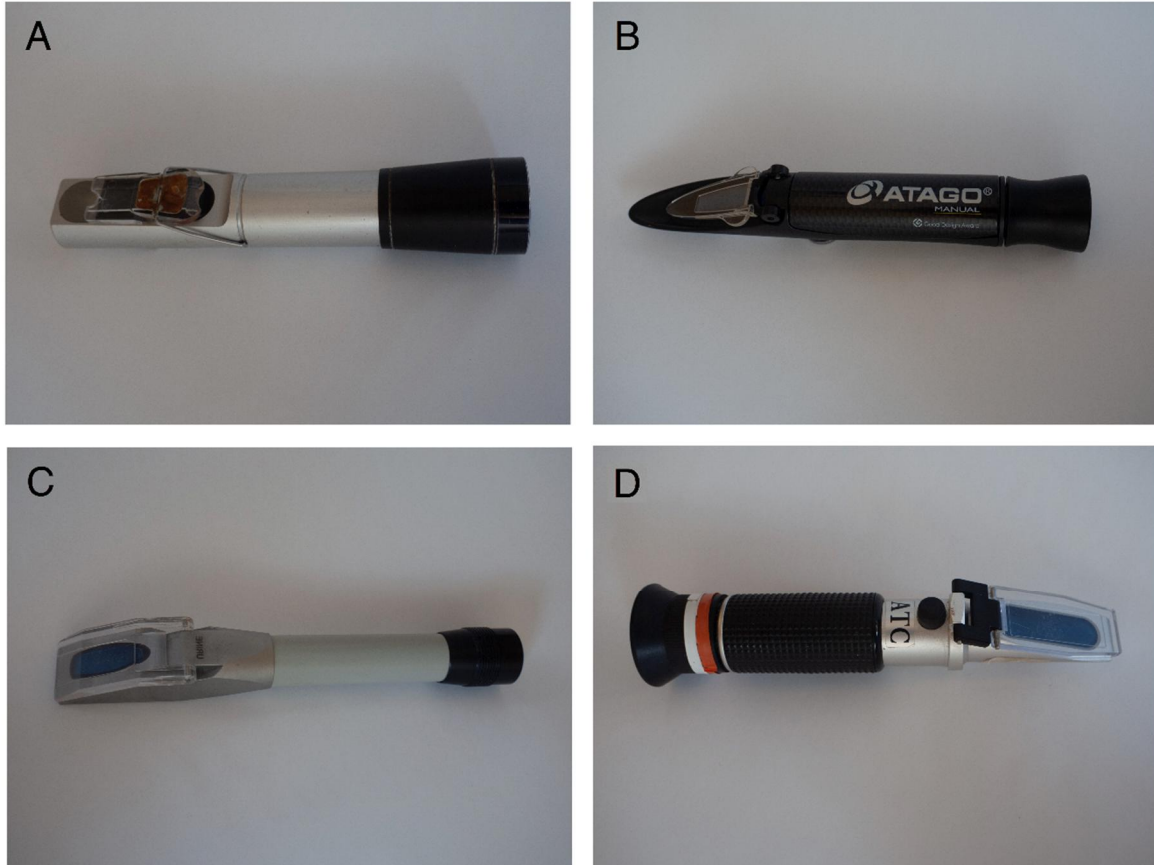
Table 1. Details of the four refractometers used to measure urine specific gravity (USG) in canine urine specimens. The details include the refractometer's designated study number, manufacturer, and model of each refractometer, whether the instrument is temperature-compensated, and the instrument's USG measurement range.

Refractometer study number	Refractometer manufacturer Refractometer model	Temperature compensation	USG measurement range
R1	Atago MASTER-SUR/NM	No	1.000 – 1.060
R2	American Optical TS Meter (TS-B/10400)	Yes	1.000 – 1.035
R3	Bellingham and Stanley The Field Refractometer (21-24)	Yes	1.000 – 1.040
R4	JorVet Protein/Specific Gravity refractometer (J-0351)	Yes	1.000 – 1.050

R, refractometer

## Comparison of refractometric USG measurement

Figure 1. Photographs showing the four medical optical hand-held refractometers used to measure urine specific gravity in canine urine specimens. (A) Atago MASTER-SUR/NM (R1); (B) American Optical TS Meter (R2); (C) Bellingham and Stanley The Field Refractometer (R3); (D) JorVet Protein/Specific Gravity refractometer (R4).



### Data analysis

#### *Interobserver agreement and method comparison*

All data were tested for normality using the Kolmogorov-Smirnov normality test. Interobserver agreement for each refractometer was evaluated by calculating the Spearman's rank correlation coefficient for duplicate readings. The coefficient of variation (CV) of duplicate measurements was used to determine interobserver CVs.

### *Comparison of refractometric USG measurement*

For method comparison statistics, the mean of the duplicate readings (mean of the two readings, one from each observer) for each refractometer was used. In addition, specimens with USG values above the upper limit of the refractometer with the narrowest measuring range, as measured by this refractometer, were excluded from analysis so that only exact USG values were used. For example, if R2 was compared to R1, all specimens with USG values > 1.035 (as measured by R2) were excluded. All refractometers were compared with one another using the Spearman's rank correlation, Bland-Altman difference plots, and Deming regression analyses.<sup>7</sup> The regression analyses yield a linear equation for the regression line in the form of

$$y = ax + b$$

where  $y$  is the result for the second refractometer in the pair,  $x$  is the USG value of the first refractometer,  $a$  is the slope (proportional bias), and  $b$  is the y-intercept (constant bias), of the regression line.

A significant constant bias was identified if the confidence interval of the regression line y-intercept did not include zero, and a significant proportional bias was identified if the confidence interval of the regression line slope did not include 1.0.<sup>7</sup>

The bias (absolute and percentage) between each pair of refractometers was calculated using the regression equation, where  $y$  is the USG measured by the second refractometer in the pair, at a USG of 1.010, 1.020 and 1.030, as measured by the first refractometer ( $x$ ) and then applying the following formulas:

$$\textit{Absolute Bias} = x - y$$

$$\% \textit{ Bias} = \frac{(x - y)}{y} \times 100$$



### *Comparison of refractometric USG measurement*

This bias was compared with the performance goals for analytical bias drawn from two sources: (1) derived from biological variation in humans – 0.27%<sup>10</sup>; and (2) consensus reached by a panel of small animal internal medicine specialists at the OVAH – 0.002 USG units.

#### *Determination of clinically relevant differences in the classification of urine concentrations and azotemia*

For each refractometer, patients were placed into four defined concentration categories according to the mean of the duplicate USG readings. These categories were as follows: highly concentrated (USG > 1.030), moderately concentrated (USG of 1.013 – 1.029), isosthenuric (USG of 1.008 – 1.012) and hyposthenuric (USG < 1.008) urine.<sup>4,11</sup>

Patients were also categorized as azotemic and non-azotemic. Azotemia was defined as a serum creatinine concentration greater than the laboratory upper reference limit of 109  $\mu\text{mol/L}$  (1.2 mg/dL). The subset of azotemic patients was further classified according to the USG findings for each refractometer as follows: pre-renal azotemia (USG  $\geq$  1.030) or renal azotemia (USG < 1.030).<sup>11</sup>

The level of reliability (agreement and correlation) between refractometers for each concentration and azotemia category was determined using the intraclass correlation coefficient based on a mean rating ( $k = 2$ ), absolute agreement, and two-way model.<sup>12</sup> The correlation coefficients were interpreted as follows: < 0.50 is poor; 0.50 – 0.75 is moderate; 0.75 – 0.9 is good, and > 0.90 is excellent reliability.<sup>12</sup>

Statistics were performed using MedCalc for Windows (version 16.4.3) (MedCalc Software, Ostend, Belgium). Statistical significance was set at a  $P$ -value < 0.05.

## Results

### Patients and Specimen Analyses

Surplus urine was obtained from 60 canine patients. Information regarding signalment (breed, age, and sex) of the dogs included in this study is provided in Table S1. One urine specimen was excluded

### Comparison of refractometric USG measurement

due to the presence of macroscopic pigmenturia, and 59 were included. Most of the specimens were clear in appearance (42/59; 71%) and the rest showed varying degrees of turbidity (17/59; 29%). The median USGs, using the mean of the duplicate readings, as measured by each refractometer, is shown in Table 2. The median pH of the specimens was 6 (range 5 – 8). Only 4 (7%) of the specimens were glucosuric, and 46 (78%) of the specimens were proteinuric, although the majority of the proteinuric specimens (29/46; 63%) showed only a 1+ on the protein patch of the urine dipstick. Thirty-one (53%) of the specimens showed hematuria, hemoglobinuria, or myoglobinuria, as measured by the blood patch on the urine dipstick. However, none of these specimens showed macroscopic pigmenturia or hematuria, based on the exclusion criteria for this study.

Thirteen (22%) of the enrolled patients were azotemic and had serum creatinine concentrations ranging from 113 – 1703  $\mu\text{mol/L}$  (1.3 – 19.3 mg/dL).

Table 2. Descriptive statistics (median and interquartile range [IQR]) of the urine specific gravity (USG) values as measured by each of the four refractometers (R1, R2, R3, and R4)

Refractometer label	Median USG	IQR
R1	1.024	1.015 – 1.042
R2	1.023	1.015 – 1.035
R3	1.021	1.014 – 1.040
R4	1.024	1.016 – 1.050

R, refractometer

### Data analysis

#### *Interobserver agreements and method comparisons*

All data showed non-Gaussian distributions ( $P < 0.05$ ). Perfect to almost perfect agreement between observers was determined for all refractometers ( $r = 0.998 - 1.000$ ,  $P < 0.0001$ ). The calculated interobserver CVs were 0.03% for R2, 0.04% for R1 and R3, and 0.05% for R4.

*Comparison of refractometric USG measurement*

Table 3. Spearman's rank correlation rho and Deming regression plot y-intercept (with 95% confidence intervals [CI]) and slope (with 95% CIs) values for each pair of refractometers.

First refractometer	Second refractometer	n	Correlation (rho)	P-value	y-intercept	95% CI (y-intercept)	Slope	95% CI (slope)
R1	R2	39	0.990	< 0.0001	-0.026	-0.064 to 0.012	1.026	0.988 to 1.063
	R3	44	0.994	< 0.0001	0.009	-0.016 to 0.033	0.991	0.967 to 1.014
	R4	46	0.996	< 0.0001	-0.189	-0.232 to -0.146	1.187	1.145 to 1.229
R2	R3	39	0.992	< 0.0001	0.001	-0.034 to 0.037	0.997	0.962 to 1.032
	R4	39	0.990	< 0.0001	-0.132	-0.176 to -0.088	1.131	1.088 to 1.174
R3	R4	44	0.994	< 0.0001	-0.188	-0.247 to -0.129	1.188	1.130 to 1.246

R, refractometer

### *Comparison of refractometric USG measurement*

Using the mean of the duplicate readings, there was an excellent correlation between each pair of refractometers (Table 3). Bland-Altman difference plots and Deming regression plots comparing each pair of refractometers are shown in Figures 2 and 3, respectively. The final number of specimens included in each comparison analysis, after the removal of highly concentrated specimens, is indicated in Table 3. Statistically significant negative constant and positive proportional biases were identified for all comparisons with R4 (Table 3). No statistically significant bias was identified between the other three refractometer pairs (R1 vs R2, R1 vs R3, and R2 vs R3) (Table 3). Mean bias estimates from the Bland-Altman difference plots were  $< 0.002$  for all comparisons except R1 vs R4 and R3 vs R4 (Figure 2).

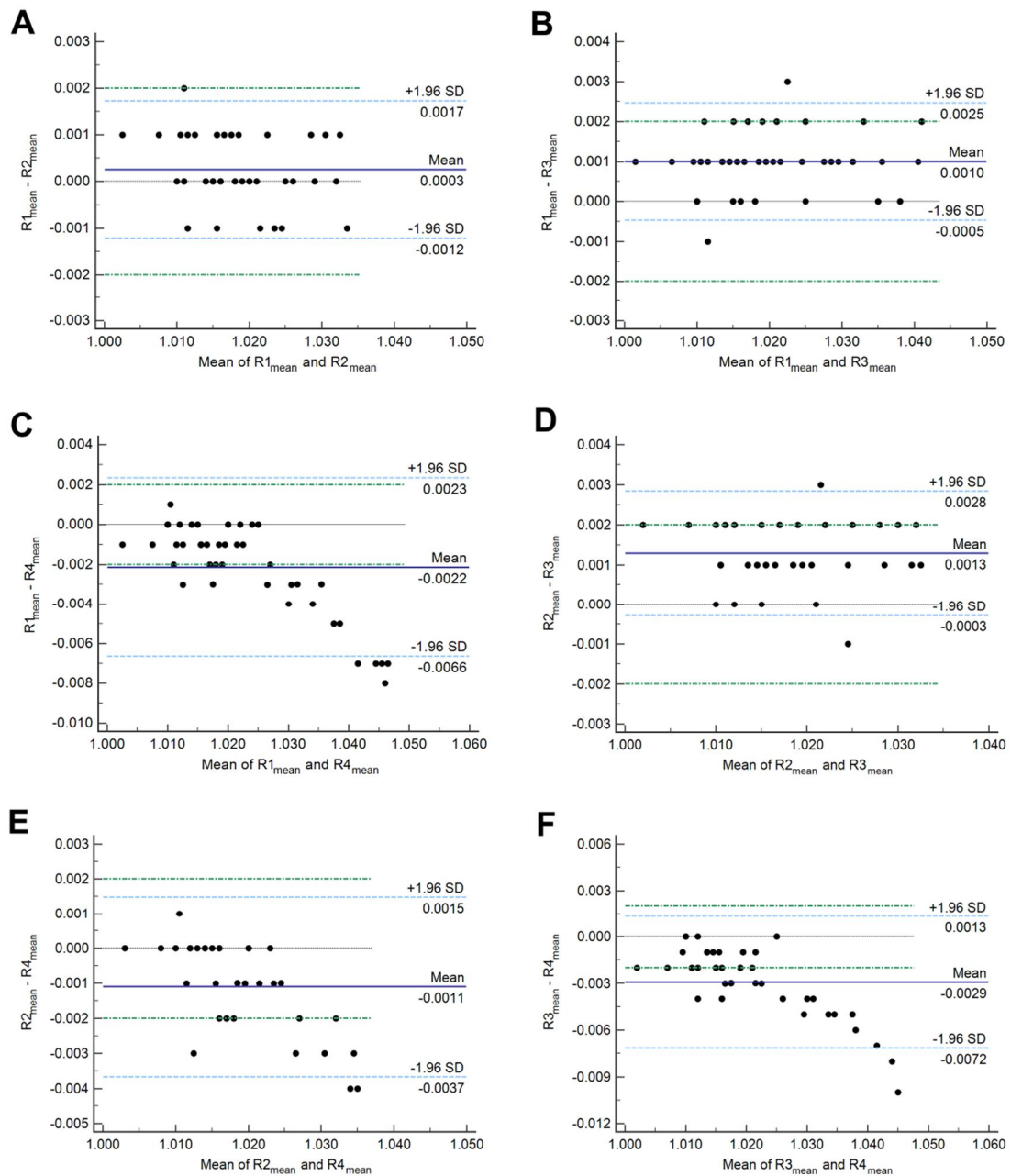
When comparing the calculated bias to both of the set performance goals, a clinically unacceptable positive bias was present between R1 vs R4 and R2 vs R4 at a USG value of 1.030, and between R3 vs R4 at USG values of 1.020 and 1.030 (Table 4). All other comparisons met the set performance goals at all USG values.

### *Determination of clinically relevant differences in the classification of urine concentrations and azotemia*

Urine specimens over a range of USG values were included in the study so that each concentration category was represented (Table S2). The overall agreement and correlation between all the refractometers for categorization of urine concentration was excellent ( $r = 0.988$ ; 95% confidence interval [CI] = 0.982 – 0.992). The agreement and correlation between each pair of refractometers for the categorization of urine concentrations were also excellent (Table 5). Across all six refractometer pairs, there were a total of 18 discordant concentration classifications: nine highly concentrated vs moderately concentrated, six moderately concentrated vs isosthenuric, and three isosthenuric vs hyposthenuric. The agreement and correlation between all refractometers for the categorization of azotemia was perfect ( $r = 1.00$ ).

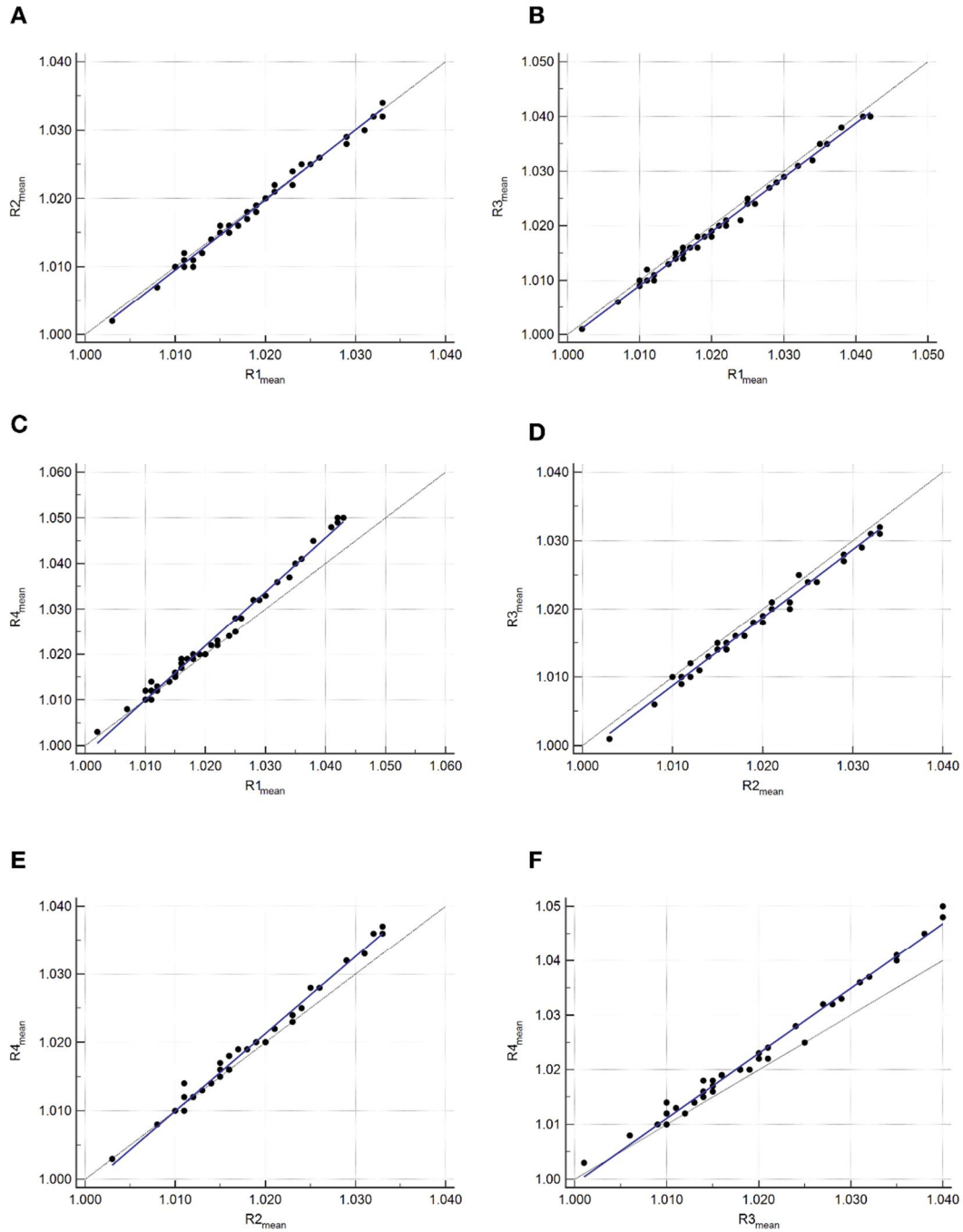
## Comparison of refractometric USG measurement

Figure 2. Bland-Altman difference plots comparing each of the six pairs of refractometers in their measurement of urine specific gravity (USG) in 39 to 46 canine urine specimens. (A) R1 vs R2; (B) R1 vs R3; (C) R1 vs R4; (D) R2 vs R3; (E) R2 vs R4; (F) R3 vs R4. The plots include the lines of zero difference (black dotted), mean absolute difference (solid blue), limits of agreement ( $\pm 1.96$  SD from the mean difference) (light blue dashed), and the maximum allowable clinically relevant difference of 0.002 (green dash-dotted) lines. The black solid dots represent each specimen.



### Comparison of refractometric USG measurement

Figure 3. Deming regression plots comparing each of the six pairs of refractometers in their measurement of urine specific gravity (USG) in 39 to 46 canine urine specimens. (A) R1 vs R2; (B) R1 vs R3; (C) R1 vs R4; (D) R2 vs R3; (E) R2 vs R4; (F) R3 vs R4. The plots include the lines of identity (black dotted) and regression lines (solid blue). The black solid dots represent each specimen.



*Comparison of refractometric USG measurement*

Table 4. Absolute and percentage biases between each pair of refractometers when the first refractometer measures the urine specific gravity (USG) at 1.010, 1.020 and 1.030.

First refractometer	Second refractometer	Bias	First refractometer USG measurement		
			1.010	1.020	1.030
R1	R2	Measured USG	1.010	1.021	1.031
		Absolute bias	0.000	0.001	0.001
		Bias (%)	0.03	0.05	0.08
	R3	Measured USG	1.010	1.020	1.030
		Absolute bias	0.000	0.000	0.000
		Bias (%)	-0.01	-0.02	-0.03
	R4	Measured USG	1.010	1.022	1.034
		Absolute bias	0.000	0.002	0.004
		Bias (%)	-0.01	0.17	0.35
R2	R3	Measured USG	1.008	1.018	1.028
		Absolute bias	-0.002	-0.002	-0.002
		Bias (%)	-0.20	-0.20	-0.20
	R4	Measured USG	1.010	1.022	1.033
		Absolute bias	0.000	0.002	0.003
		Bias (%)	0.03	0.16	0.28
R3	R4	Measured USG	1.012	1.024	1.036
		Absolute bias	0.002	0.004	0.006
		Bias (%)	0.19	0.37	0.55

R, refractometer

### *Comparison of refractometric USG measurement*

Table 5. Correlation and agreement (and 95% confidence intervals [CI]) between each pair of refractometers (R1, R2, R3, and R4) for the categorization of urine concentrations as measured using intraclass correlations based on a mean rating ( $k = 2$ ), absolute agreement, and two-way model.

First refractometer	Second refractometer	Correlation ( $r$ )	95% CI ( $r$ )
R1	R2	0.986	0.976 – 0.992
	R3	0.993	0.989 – 0.996
	R4	0.964	0.934 – 0.979
R2	R3	0.978	0.963 – 0.987
	R4	0.977	0.961 – 0.986
R3	R4	0.956	0.921 – 0.975

R, refractometer

### Discussion

This study found no clinically significant differences between three of the four medical optical refractometers in their measurement of USGs. The fourth refractometer, R4, showed clinically significant differences at higher USG values when compared with the three other refractometers. In the majority of cases, and in all azotemic cases, there was no clinically significant difference between observers and between the refractometers in their classification of patients into urine concentration categories. This study differs from others conducted in dogs in that all the refractometers compared were medical optical hand-held refractometers, were different models to those previously compared and, were not veterinary specific (digital or optical).<sup>3,5,6</sup> In two previous studies, one optical and digital refractometer (one veterinary specific and one medical) were compared with one another.<sup>3,5</sup> Both found that the digital refractometer measured lower USG values than the optical refractometer, but in the study using a medical digital refractometer, these differences were within acceptable limits.<sup>3,5</sup> Tvedten et al compared five different refractometers - three medical optical



### *Comparison of refractometric USG measurement*

refractometers and two veterinary-specific digital refractometers - in dogs and cats. The authors found that the refractometers reported different results from one another.<sup>6</sup> Our study aimed to compare the refractometers currently used in the clinics of our hospital to determine whether the same was true, specifically to determine if clinically significant differences that could affect decision-making were present. As no digital or species-specific refractometers were used in the hospital clinics, these types of refractometers were not included in this study, and to our knowledge, the chosen refractometers have not been included in any previous method comparison studies.

The urine specimens included in this study varied in their physical and chemical properties, and a wide range of USG measurements was analyzed. No specimen dilutions were performed, based on Tvedten et al's findings that dilution resulted in spurious increases in USG.<sup>6</sup> Although this study aimed only to compare refractometers and not to determine accuracy or precision, there are no studies, to the authors' knowledge, that assess whether this spurious increase in USG is predictable and equal in magnitude across various refractometers. Additionally, this study aimed to assess clinically relevant implications of differences between refractometers and, in dogs, the degree that urine is concentrated above 1.030 does not provide clinically relevant information.

Two of the above-mentioned previous studies excluded specimens that were turbid or had active urine sediments.<sup>3,6</sup> Specimen turbidity is the result of suspended particles such as cells and is not expected to affect the refractive index of urine. One of the studies also excluded glucosuric or significantly proteinuric urine ( $\geq 2+$  on a urine dipstick).<sup>3</sup> Glucosuria and significant proteinuria are both known to spuriously increase USGs, although these conditions do not appear to significantly affect the relationship between USG and urine osmolality in dogs.<sup>1,3</sup> Because the aim of this study was not to assess refractometer accuracy but to determine how the refractometers compared to one another, conditions that could spuriously affect the USGs were not excluded. Macroscopic pigmenturia or hematuria, on the other hand, can cause difficulty in reading an exact USG

### *Comparison of refractometric USG measurement*

measurement. For the purposes of this study, exact measurements were needed, resulting in one specimen with macroscopic pigmenturia being excluded.

Three of the instruments were temperature-compensated, and one (R1) was not. As the temperature of a solution increases above a certain threshold, the refractive index decreases due to the expansion of solutes.<sup>9</sup> Temperature compensation is a means of measurement standardization so that only solute types and concentrations will affect the refractive index and, thus, the measured USGs. According to the manufacturer of R1, the temperature threshold above which USG could decrease is 20°C.<sup>13</sup> In this study, varying biases in USGs were found between R1 and the other three refractometers (R2, R3, and R4). If the temperature of the specimens analyzed by R1 affected the USG, one would expect a significant constant positive bias in all three temperature-compensated refractometers compared with R1. On the contrary, the biases found were insignificant for R2 and R3; and positive (constant and proportional) only for R4. Thus, the absence of temperature compensation is unlikely to have affected comparisons including R1 but cannot be ruled out as a factor contributing to the differences seen.

There was no difference between observers in their measurement of USG. Two previous studies comparing the performance of different refractometers ensured only one observer analyzed the specimens to reduce inter-reader variability.<sup>3,14</sup> In addition, other previous studies mentioned inter-reader variability as a study design limitation but this variability was not measured.<sup>5,6</sup> In clinical practice, various individuals such as veterinary nurses, veterinary technologists, and veterinarians may be reading USG values, and keeping observers consistent is often not possible. To replicate this scenario, we involved six different observers with various levels of experience and knowledge, and perfect to almost perfect agreement was found between them. Apart from the subjectivity when reading USG values that fall between two lines of measurement and those where there is no clear line of demarcation, analysis of USG is largely objective, which makes this result unsurprising.

### *Comparison of refractometric USG measurement*

The agreement between each pair of refractometers was excellent to perfect. Using regression analysis, statistically significant constant and proportional biases were identified between all the refractometer pairs, which included R4. These positive biases were considered clinically relevant at higher USG values, with the greatest bias when R4 was compared with R3. This indicates that the bias, particularly the proportional bias, will affect clinical decisions. The mean systematic bias between all other refractometer pairs (whether negative or positive) was never greater than the set performance goals, and in some pairs, was undetectable (Table 4). Furthermore, the mean difference between most refractometer pairs was  $< 0.002$  (Figure 2), except for R1 vs R4 and R3 vs R4, where it was  $> 0.002$ . This is likely due to the proportionally larger differences seen at USG values above 1.030. This trend is not apparent when evaluating R2 vs R4, because of the narrower measuring range of R2 (up to 1.035), which resulted in more highly concentrated specimens being removed for statistical analysis.

No cases for R1 vs R2 and only one case each for R1 vs R3 and R2 vs R3 were outside the maximum allowable clinically relevant difference as determined by the Bland-Altman difference plots (Figure 2). These findings are similar to those previously reported, where 91.5% of specimens showed mean differences in USG values of  $\leq 0.002$  between a medical optical and digital refractometer.<sup>3</sup> It can be concluded that R1, R2, and R3 compare favorably with one another and can be used interchangeably with no clinically significant differences. Conversely, comparisons with R4 (R1 vs R4, R2 vs R4, and R3 vs R4) exhibited substantially more cases outside the maximum allowable clinically relevant difference of 0.002 (15, 6, and 17 cases, respectively) (Figure 2). In all these cases, the USG values measured by R4 were higher than the other refractometer. Of these cases, two dogs for R2 vs R4, and three dogs for R1 vs R4 and R3 vs R4, would have had a discordant result around the clinical decision cut-off of 1.030. In the remaining cases, the classification and clinical interpretation were not affected. Based on the discordant results and clinically significant biases in comparisons involving R4, it can be concluded that this instrument should not be used interchangeably with the other refractometers in the present study.

### *Comparison of refractometric USG measurement*

These conclusions regarding interchangeability differ from those made by Tvedten et al, who concluded that, with increasing USG values, the results of all five refractometers included in their study did not increase consistently compared with the reference method and with one another.<sup>6</sup> Thus, these instruments could not be used interchangeably.<sup>6</sup> However, when only evaluating the three refractometers with the same specifications as the refractometers used in the present study, it was determined that all three had a mean bias of  $\leq 0.002$  (positive or negative) compared with the reference method.<sup>6</sup> Therefore, it is possible that these medical optical refractometers compared similarly to each other, but no method comparison statistics were performed between the refractometers, only between the refractometers and the reference method.<sup>6</sup>

Lastly, the agreement between the refractometers in their categorization of patient urine concentrations and azotemia was excellent to perfect. To the authors' knowledge, this is the first study that has assessed the differences in refractometers based on the clinical significance of results. This finding does not, however, diminish the value of repeated USG measurements. One USG value should not be used in isolation and further diagnostics and follow up measurements are necessary to confirm the categorizations of urine concentrations and azotemia. The number of discordant results between R4 and the other three refractometers strengthens this recommendation. These discordant results could lead to misclassification and, if not followed up, could result in misdiagnoses and incorrect patient management.

A limitation of this study is that the complete urinalyses were performed by several different people. The number of proteinuric, glucosuric, and turbid specimens might have been different had one individual performed all complete urinalyses, as there is a degree of subjectivity involved when assessing these variables. Another limitation is that analyses with warmer urine or at ambient temperatures greater than 25°C were not performed, and so conclusions regarding how the refractometers compared in warmer conditions could not be drawn. Also, only a small number of azotemic dogs were included. Further studies are needed to confirm the similarity of refractometers

### *Comparison of refractometric USG measurement*

in their categorization of azotemic patients. Lastly, the minimum recommended sample size of 40 for method comparison studies<sup>7</sup> was not achieved in the comparisons with R2 (39 specimens included). Because specimens were collected prospectively and were not excluded based on the degree of urine concentration, many urine specimens concentrated beyond the measuring range of R2 were initially included. Nevertheless, the comparisons that included R2 show similar trends to the other comparisons and the range ratio, an additional indicator of the appropriate number of samples to include<sup>7</sup>, calculated from a range of USG values of 1.003 – 1.033 is 11. This ratio is high, indicating that small sample sizes are sufficient to obtain acceptable results.<sup>7</sup>

In conclusion, refractometer comparisons that included R1, R2, and R3 did not show clinically relevant differences in most cases, and these refractometers can be used interchangeably at all USG values. In contrast, R4 should not be used interchangeably with the other three refractometers, despite its excellent to perfect correlation in the categorization of urine concentrations and azotemia. The reason for this conclusion is based on the number of discordant results, the number of cases outside the maximum allowable clinical difference, and the failure to meet the performance goals at USG values  $\geq 1.020$ . Additionally, given that interobserver variability is negligible, regardless of the observer's level of experience, measurements by different observers will also not show clinically relevant differences in almost all cases. These findings have positive implications in clinical practice, where several different refractometers could be used by different observers during a patient's hospitalization and at follow up visits.

### Acknowledgments

The authors are grateful to the students, veterinary nurses and clinicians working at the Onderstepoort Veterinary Academic Hospital for supplying samples for this study and to Mrs Carien Muller (University of Pretoria) and Dr Zandri Whitehead (University of Pretoria), who assisted the authors in USG measurement.

## Conflict of Interest

The authors state that there are no conflicts of interest.

## References

1. Ayoub JA, Beaufriere H, Acierno MJ. Association between urine osmolality and specific gravity in dogs and the effect of commonly measured urine solutes on that association. *Am J Vet Res.* 2013;74:1542-1545.
2. Dossin O, Germain C, Braun JP. Comparison of the techniques of evaluation of urine dilution/concentration in the dog. *J Vet Med A.* 2003;50:322-325.
3. 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;170:463.
4. International Renal Interest Society: Using urine specific gravity. Watson ADJ, Lefebvre HP, Elliott J, ed. 2015. Available at: [http://www.iris-kidney.com/education/urine\\_specific\\_gravity.html](http://www.iris-kidney.com/education/urine_specific_gravity.html). Accessed October 23, 2019.
5. Tvedten HW, Noren A. Comparison of a Schmidt and Haensch refractometer and an Atago PAL-USG Cat refractometer for determination of urine specific gravity in dogs and cats. *Vet Clin Pathol.* 2014;43:63-66.
6. Tvedten HW, Ouchterlony H, Lilliehook IE. Comparison of specific gravity analysis of feline and canine urine, using five refractometers, to pycnometric analysis and total solids by drying. *N Z Vet J.* 2015;63:254-259.
7. Jensen AL, Kjelgaard-Hansen M. Method comparison in the clinical laboratory. *Vet Clin Pathol.* 2006;35:276-286.
8. Albasan H, Lulich JP, Osborne CA, Lekcharoensuk C, Ulrich LK, Carpenter KA. Effects of storage time and temperature on pH, specific gravity, and crystal formation in urine samples from dogs and cats. *J Am Vet Med Assoc.* 2003;222:176-179.

*Comparison of refractometric USG measurement*

9. George JW. The usefulness and limitations of hand-held refractometers in veterinary laboratory medicine: An historical and technical review. *Vet Clin Pathol.* 2001;30:201-210.
10. Westgard QC: Desirable Specifications for Total Error, Imprecision, and Bias, derived from intra- and inter-individual biologic variation. Ricos C, Alvarez V, Cava F, et al., ed. 2014. Available at: <https://www.westgard.com/biodatabase1.htm>. Accessed October 17, 2018.
11. Stockham SL, Scott MA. *Fundamentals of Veterinary Clinical Pathology*. 2nd ed. Ames, IA: Wiley-Blackwell; 2008: 415-494.
12. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med.* 2016;15:155-163.
13. Atago Co. Ltd.: Refractometer Data Book - Temperature Correction Chart. ed. 2019. Available at: [https://www.atago.net/en/databook-refractometer\\_correction.php](https://www.atago.net/en/databook-refractometer_correction.php). Accessed February 11, 2019.
14. 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;13:152-154.

*Comparison of refractometric USG measurement*

Supplementary material

Table S1. Signalment (breed, age, and sex) of the dogs from which 59 urine specimens were collected.

	Number of individuals
Breeds	
Jack Russell Terrier	7
Dachshund	5
Mixed Breed	4
Yorkshire Terrier	4
Miniature Schnauzer	3
Miniature Doberman Pinscher	3
Labrador Retriever	3
Pitbull Terrier	3
Beagle	2
Maltese Poodle	2
Cocker Spaniel	2
Pekingese	2
Boxer	1
English Bulldog	1
Siberian Husky	1
Malinois	1
Great Dane	1
Standard Poodle	1
Mexican Hairless Dog	1
Fox Terrier	1



*Comparison of refractometric USG measurement*

	Bichon Frise	1
	Scottish Terrier	1
	Rottweiler	1
	Shar Pei	1
	Irish Wolfhound	1
	Staffordshire Bull Terrier	1
	Shih Tzu	1
	Corgi	1
	Bull Terrier	1
	Weimaraner	1
	Italian Greyhound	1
Age	Mean	7.4 years
	Range	6 months to 14.5 years
Sex	Intact Males	13
	Neutered Males	14
	Intact Females	7
	Neutered Females	25

*Comparison of refractometric USG measurement*

Table S2. The number of urine specimens classified into each concentration category by each of the four refractometers (R1, R2, R3, and R4)

Refractometer label	Highly concentrated (USG > 1.030)	Moderately concentrated (USG 1.013 – 1.029)	Isosthenuric (USG 1.008 – 1.012)	Hyposthenuric (USG < 1.008)
R1	24	25	8	2
R2	24	26	8	1
R3	23	26	8	2
R4	26	25	7	1

R, refractometer; USG, urine specific gravity