Veterinary Clinical Pathology
An International Journal of Laboratory Medicine

# Evaluation of the analytical variability of dipstick protein pads on canine urine

| | |
|---|---|
| Journal: | *Veterinary Clinical Pathology* |
| Manuscript ID | Draft |
| Manuscript Type: | Original Article |
| Date Submitted by the Author: | n/a |
| Complete List of Authors: | Giraldi, Marco; University of Milan, Department of Veterinary Medicine; University of Milan, Veterinary Teaching Hospital<br>Paltrinieri, Saverio; University of Milan, Department of Veterinary Medicine; University of Milan, Veterinary Teaching Hospital<br>Zatelli, Andrea; Medical Consultancy Services, |
| Key Words: | Dogs, Proteinuria, Reagent strips, Urinalysis |
| | |

# Evaluation of the analytical variability of dipstick protein pads on canine urine

Dipstick analysis of canine proteinuria

Marco Giraldi,[a,b] DVM

Saverio Paltrinieri,[a,b]* DVM PhD Dipl ECVCP

Andrea Zatelli[c], DVM PhD

[a]*Department of Veterinary Medicine – University of Milan, Via Celoria, 10 - 20133 Milan, Italy*

[b]*Veterinary teaching hospital – University of Milan, Via dell'Università 6 – 26900 Lodi, Italy*

[c]*Medical Consultancy Services, Ta'Xbiex, Malta*

* Corresponding author: Dr. Paltrinieri

*Department of Veterinary Medicine – University of Milan,*

*Via Celoria, 10 - 20133 Milan, Italy*

Tel.: +39 0250318103.

*E-mail address:* saverio.paltrinieri@unimi.it

**Abstract**

Background: The dipstick is a first line and inexpensive test to exclude the presence of proteinuria in dogs. No information is available on the analytical variability of dipstick analysis

Objectives: The aim of this study was to assess the analytical variability of two dipsticks and the inter-operator variability in dipstick interpretation

Methods: Canine urine supernatants (n=174) were analyzed with two commercially available dipstick. Two observers evaluated each result blinded to the other observer and to the results of the other dipstick. Intra and inter-assay variability were assessed on 5 samples (corresponding to the 5 different semi-quantitative results) tested for 10 consecutive times and 5 consecutive days, respectively. The variability between observer and between dipsticks was evaluated with Cohen's k test.

Results: Intra-assay repeatability was good (≤3/10 errors), whereas inter-assay variability was higher (from 1/5 to 4/5 discordant results). The concordance between operators (k=0.68 and 0.79 for the two dipsticks) and between dipsticks (k=0.66 and 0.74 for the two operators) was good. However, one observer and one dipstick overestimated the results compared with the other observer or dipstick. In any case, discordant results accounted for a single unit of the semi-quantitative scale.

Conclusions: As for any other method, analytical variability may affect the semi-quantification of urinary proteins with dipstick. Subjective interpretation of the pad and, to a lesser extent, intrinsic staining properties of the pads could affect the results. Further studies are warranted in order to evaluate the effect of this variability on clinical decision.

***Keywords***: Dogs, Proteinuria, Reagent strips, Urinalysis

## Introduction

In clinical practice proteinuria is defined as the increased amount of proteins in the urine. The presence of persistent proteinuria of renal origin has a diagnostic and prognostic value for chronic kidney disease (CKD) in dogs. The detection of proteinuria is therefore a milestone in the management of dogs with CKD.[1]

The recommended methods to evaluate proteinuria in dogs is the quantitative evaluation of the urinary proteins and urinary creatinine, followed by the calculation of the urinary protein to creatinine (UPC) ratio, that allows to correct the magnitude of proteinuria by the dilution of urine.[1,2] However, in clinical practice, dry reagent test strips (dipsticks) are rapid and inexpensive methods that allow, along with other urinary physico-chemical or cytological parameters, a first evaluation of the presence or absence of proteins in urine in a point-of-care setting.[3] Proteins (mainly albumin), when present, react with the pad yielding a variable color change whose intensity is proportional to the protein concentration. Results are then expressed semi-quantitatively, usually as negative, trace or 1+ to 4+ (corresponding, for most of the commercially available dipsticks to 15 to 2000 mg/dL of proteins), comparing the pad against the chart on the side of the dipstick package or, alternatively, loading the strip in an automated spectrophotometric reader.

Although the evaluation of protein excretion using the dipstick must be considered a screening test, erroneous interpretations of the pad may affect clinical decisions. In human medicine information from quality assurance programs revealed that the rates of misclassification of one and two scores above or below the expected value were 9.7% and 2.3%, respectively,[4] and that the intra- and inter-observer agreement is moderate (k=0.53) to very good (k=1).[5] Moreover, the use of automated reader is recommended to minimize the observer-related errors[6] despite automated readings

only slightly improved the reproducibility of dipstick analysis.[5] In veterinary medicine, although similar results are likely, and quality control programs for urinalysis and dipstick tests are recommended,[2] little is known about the influence of inter-operator variability on the analytical variability of dipstick testing or about the intrinsic performances of the dipstick from different manufacturers, that may have their peculiar analytical sensitivity and range or different semi-quantitative interpretation charts.

Therefore, the aim of this study was to assess the analytical variability of dipstick analysis in the evaluation of the presence/absence of proteins in canine urine, through the comparison of results obtained using two commercially available dipsticks by two different independent operators.

## Materials and Methods

A total of 174 canine urine supernatants were included in this study.

Urine samples were collected over a period of 12 months from dogs of different age, sex and breed, presented for diagnostic investigation at the internal medicine unit of the Department of Veterinary Medicine (DIMEVET, University of Milan), by mean of cystocenthesis (n=113), free catch (n=47) or unspecified method (n=14). Samples were collected for diagnostic purposes under informed consent of the owner and therefore, according to the institutional Ethical Committee (deliberation number: 2/2016) a formal approval of the study from the Institutional Animal Care and Use Committee was not necessary.

Since this was a validation study not focused on the impact of the results on the clinical diagnosis, samples were included irrespective of health status of the dogs.

All urine samples underwent complete urinalysis (including USG, dipstick evaluation and sediment examination) within two hours from collection. In order to perform sediment analysis, 5 mL of sample were placed in sterile conical 10 mL tube and centrifuged 5 minutes at 1250 rpm (450 G). Supernatants were aliquoted in a 1.5 mL tube and stored at -20°C within 4 hours from collection for a maximum of 12 months. At the time of analysis, supernatants were gently thawed by transferring tubes at 4°C the day before analysis and then at room temperature one hour before analysis.

Each sample was tested with two commercially available dipsticks (Dipstick 1: U-11 Urine Strips, Mindray, Shenzhen, China; Dipstick 2: Multistix 10 SG Reagent Strips, Siemens, Siemens Healthcare Diagnostics Inc, Tarrytown, NY /Siemens Healthcare Diagnostics, Eschborn, Germany) after a preliminary assessment of intra- and interassay variability of each dipstick (see below). Moreover, each dipstick was evaluated by two operators with similar experience in urinalysis procedures.

In each analysis, the dipstick was kept out from the case no more than 2 minutes before the use. Fifty microliters of urines supernatant were applied with a dispensable pipette only on the protein pad and, in order to prevent cross-contamination by dyes from close pads the contact of urine with adjacent pads was avoided; then, excess urine was gently discarded hitting the dipstick on a clean paper towel.

In order to avoid bias in interpretation of the second dipstick by each operator and to avoid excess time between urine application and reading, samples were analyzed in batches (8 samples per batch), thus allowing the evaluation within 60-120 seconds, as recommended by the manufacturer's instruction of both the dipstick kits, and urine samples were applied to the second dipstick in a different random order compared with that used for the first dipstick. Each operator interpreted the dipsticks results

blinded to the results of the other operator. Also, due to the different random order described above, the interpretation of the second dipstick was blinded to results of the first dipstick.

Interpretation of each dipstick was performed by comparing the color of the protein pads with the corresponding color chart provided by each manufacturer.

Because, as expected, some samples yielded a color reaction with a chromatic intensity not identical to those proposed on the chart, the following reading method was chosen: each pad was compared with one reference color block at a time; when an almost perfect match between the pad and the block on the chart was identified, the corresponding result was assigned (i.e negative or trace - N/T – or positive: 1+, 2+, 3+ or 4+); when the color intensity was intermediate between two blocks on the chart, the results corresponding to the nearest reference color blocks (lighter or darker) was assigned. However, in these cases, the samples were also recorded as "difficult".

**Intra-assay and inter-assay precision**

Five urine samples yielding results corresponding to the 5 different semi-quantitative scores of proteinuria (namely N/T = <30 mg/dL, 1+ = 30 mg/dL, 2+ = 100 mg/dL, 3+ = 300 mg/dL, 4+ = ≥2 g/dL) were selected and used for analytical precision tests.

For intra-assay evaluation, the 5 samples were tested with both dipsticks 10 consecutive times.

Then, inter-assay variability was assessed testing the same samples 5 times in 5 consecutive days by both dipsticks, storing samples at 4°C overnight between the evaluations.

For both intra- and inter-assay evaluations, interpretation of pads was performed with the same method described for the whole set of samples.

For each semi-quantitative score, numbers of different results were counted and the percentage of the results lower (underestimation) or higher (overestimation) than the first reading were calculated.

Imprecision was expressed as the percentage of different results out of the sum of the two operators (i.e 20 reading/results for intra-assay evaluation and 10 results for inter-assay evaluation)

**Statistical analysis**

Concordance between operators and between dipsticks was tested with Cohen's Kappa test[7] and was calculated either for the whole set of results (N/T, 1+, 2+, 3+, 4+) or grouping results as ≤1 and ≥2.

Moreover, for the evaluation of the concordance between dipsticks, the results of the two operators were combined in order to reach a consensus and, in case of discordant results, further intermediate categories were created (N/T−1+, 1+−2+, 2+−3+, 3+−4+).

The k coefficients were used to define the concordance as follows: 0.00−0.20, 0.21−0.40, 0.41−0.60, 0.61−0.80 and 0.81−1.00 represented poor, fair, moderate, good and very good concordance[8], respectively.

In order to quantify the rate of discordant results at different level of positivity, the number of discordant sample yielded between two scores (e.g. between N/T and 1+ or between 1+ and 2+ etc.) were counted and the percentage was calculated out of the total number of samples found among the two scores evaluated.

Descriptive statistics were performed with Excel software and the Analyze-it statistical software (Analyse-it Software Ltd, Leeds, West Yorkshire, England) was used to assess the level of concordance (Cohen's k)

## Results

### Intra-assay variability

Dipstick 1 always provided the same results recorded at first reading, except in two cases: operator 1 overestimated one 3+ sample (10%) and Operator 2 underestimated one 2+ sample (10%) Difficult interpretations were more frequent for Operator 1 (5/10 at 1+; 2/10 at 2+ and 3+) than for Operator 2 (4/10 at 2+).

Using Dipstick 2, Operator 1 overestimated one 1+, one 2+ and one 3+ sample (10% each), while Operator 2 underestimated three 2+ and three 3+ samples (30% each). Only Operator 1 recorded difficult interpretations (3/10 at 1+ and 2+ and 1/10 at 3+).

### Inter-assay variability

With Dipstick 1, Operator 1 overestimated one 1+ sample (20%) and Operator 2 overestimated one 1+ sample (20%) and underestimated three 2+ samples (60%) and one 3+ sample (20%). Difficult interpretations were rare either for Operator 1 (1/10 at N/T and at 2+) or for Operator 2 (1/10 at 3+).

Imprecision was more frequent for Dipstick 2: Operator 1 overestimated four 1+ (40%) and four 3+ samples (80%), while Operator 2 overestimated one 1+ (10%), one 2+ (10%) and three 3+ samples (30%). Only Operator 1 recorded difficult interpretations (1/10 at 2+ and 3+).

### Analysis of samples

Samples covered all possible results of the dipsticks but results with scores lower than 2+ were more frequent for both the dipsticks, as follows.

Using Dipstick 1, samples recorded as N/T, 1+, 2+, 3+ and 4+, were respectively 96, 50, 17, 10 and 1 for Operator 1 and 112, 35, 16, 10 and 1 for Operator 2 (Table 1). Operators 1 and 2 recorded 23 (13,2%) and 22 (12.6%) difficult interpretations, respectively. These uncertain interpretations were mostly between negative and trace or between trace and 1+ either for Operator 1 (n=10, and n=10, respectively) or Operator 2 (n=12 and n=6, respectively). The total number of samples that were difficult for both operators was 39 (22.4%).

Using Dipstick 2, samples recorded as N/T, 1+, 2+, 3+ and 4+, were respectively 84, 53, 19, 12 and 6 for Operator 1 and 105, 37, 17, 11 and 4 for Operator 2 (Table 1). Operators 1 and 2 recorded 23 (13,2%) and 32 (18.4%) difficult interpretations, respectively. These uncertain interpretations were mostly between negative or trace, between trace and 1+ and between 1+ and 2+ either for Operator 1 (n=7, n=8, and n= 6, respectively) or 2 (n=11, n=14 and n=3, respectively). The total number of samples that were difficult for both operators was 46 (26.4%).

**Concordance between operators**

The concordance between operators was “good” for both dipsticks, and slightly higher for Dipstick 1 (k=0.79) than for Dipstick 2 (k=0.68). Discordant results were found mostly at lower protein concentrations for dipstick 1 (N/T vs 1+: 11% of misclassifications, 1+ vs 2+: 6% of misclassifications; 2+ vs 3+: 8% of misclassifications) but at all the protein concentrations for dipstick 2 (N/T vs 1+: 17% of misclassifications, 1+ vs 2+: 14% of misclassifications; 2+ vs 3+: 12% of

misclassifications; 13% of misclassifications). Moreover, for both the dipsticks Operator 1 tended to provide higher scores compared with Operator 2.

When results were grouped as ≤1+ and ≥2, concordance improved consistently to the "very good" category (Dipstick 1 k=0.94; Dipstick 2 k=0.87) and only 3 and 7 discordant results were misclassified with Dipstick 1 and Dipstick 2, respectively.

**Concordance between dipsticks**

The concordance between the two dipsticks was "good" for both operators (Operator 1 k=0.66; Operator 2 k=0.74). Similarly to the inter-operator variability described above, discordant results were more frequent for Observer 1 (N/T vs 1+: 12% of misclassifications, 1+ vs 2+: 16% of misclassifications; 2+ vs 3+: 30% of misclassifications; 3+ vs 4+: 45% of misclassifications) than for Observer 2 (N/T vs 1+: 11% of misclassifications, 1+ vs 2+: 11% of misclassifications; 2+ vs 3+: 17% of misclassifications; 3+ vs 4+: 27% of misclassifications)

When results were grouped as ≤1+ and ≥2+, concordance improved to "very good" category (k=0.83 and =0.90 for Operator 1 and Operator 2, respectively), showing 9 and 5 samples misclassified by dipsticks (≥2 using Dipstick 2 and ≤1 using Dipstick 1) with Operator 1 and Operator 2, respectively.

Overall Dipstick 2 tended to provide higher scores compared with Dipstick 1.

Concordance between dipsticks using the consensual agreement between operators (Table 2) was defined as "moderate" (k=0.59). Misclassifications were 22.5% in the interval N/T vs 1+, 42.8% in 1+ vs 2+, 66.6% in 2+ vs 3+ and 66.6% in 3+ vs 4+.

Again, grouping results as <1, 1-2, >2, concordance improved to "good" (k=0.76).

**Discussion**

This study, using two commercially available dipsticks and two operators, demonstrated that a variable imprecision in the evaluation of the concentration of urinary proteins exist. The study design adopted to assess the analytical variability (e.g. application of urine by the same operator, disposal of urine only on the protein pad, analysis of 8 samples per batch, randomization of readings etc) prevented the effect of other possible sources of error and variability such as oxidation of dipsticks, insufficient amount of urine on the pad, excess urine with contamination from other pads, colour changes due to delay of readings, interpretative biases due to the sequential analysis of samples by the same operator, etc.[4,9,10]

Intra-assay imprecision, however, was similar in magnitude to that reported in human medicine where the reproducibility of visual reading was 68-85%,[11] and was variable between the operators. As for any other test,[12] inter-assay variability was even higher. Nevertheless, this linter-assay variability has a low clinical importance because rarely repeated readings are performed during different days.

The imprecision may depend on intrinsic characteristics of the pads, on the visual perception of the operators, on environmental factors (e.g. different light during days) or, as regards inter-assay variability, on changes of pH or of protein concentration induced by refrigeration, as shown in studies on the protein to creatinine ratio.[13] However, the dipstick is analytically less sensitive than wet chemistry and no effects of storage on pH were reported in dogs[14] and therefore storage artifacts are unlikely.

Moreover, imprecision tests evidenced a high frequency of samples that were difficult to interpret, especially at low scores (i.e. N/T and 1+) and with Dipstick 2. For both the dipsticks, the two operators found difficulty with different samples. This points out that the difficult interpretation was operator-dependent instead of sample- or dipstick-dependent. However, to our impression, the two dipstick had slight differences in the

colour of the pad before the application of urines and in the hue after reaction with the same samples (especially at low protein concentration). This difference may complicate the interpretation of results. Whether sample-dependent factors (e.g. physical or chemical properties of urine such as urine dilution or presence of pigments) could affect the agreement between dipsticks needs further evaluation.

This study showed also that the inter-operator concordance was not perfect, due to discordant results at all the levels of positivity. The degrees of concordance recorded in this study were similar to that found in people, where a k coefficient of 0.82 was found[5] but lower than that reported in dogs (k=0.92).[15]

The majority of the discordant results were due to an overestimation of results by one operator compared with the other. Also in human medicine a tendency to overestimate or underestimate protein pads by single operators was demonstrated,[11] likely due to the different visual perception mentioned above. Moreover, although the reading method was standardized between the operators, each operator could consistently perceive as closer to the lower or the higher score the color reactions that were intermediate to those shown on the chart, as already demonstrated.[10]

Similarly to the results of inter-operator variability, sub-optimal concordances between dipsticks were found with both operators. Interestingly, Dipstick 2 tented to provide higher scores compared with Dipstick 1, and about a quarter of samples ≥2+ by Dipstick 2 were classified as 1+ by Dipstick 1. This result points out that, although the two manufactures declared the same analytical sensitivity (15 mg/dL) and reported equal protein concentrations for the 4 blocks on the chart, slight differences in biochemical reaction may exist between different commercially available dipsticks. According to a previous study in dogs,[16] samples with negative dipstick are likely non proteinuric, samples with 2+ or more are likely proteinuric and samples with 1+ may

or may not be proteinuric depending on the USG. Therefore, misclassification of samples as N or 1+ and 1+ or 2+ could be of clinical significance, while misclassification of samples with 2+, 3+, or 4+ may be less relevant on a clinical point of view, since the calculation of the UPC ratio, that is more accurate that dipstick, is recommended for any sample with results ≥2+.

The increase of agreement grouping results as ≤1+ or ≥2+ showed that the inter-operator variability decreases. In other words, samples with more than 2+, likely corresponding to proteinuric dogs according to a previous study,[16] can be correctly identified independently on the operator, while this results confirms that discrepancies mainly occurred between N/T results and 1+, that according to the study cited above should be considered respectively as definitely on proteinuric and dubious, i.e. proteinuric or not depending on the USG. In practical terms this may indicate that the analytical and inter-operator variability does not affect the sensitivity of the dipstick but it may affect the specificity of the method. Therefore, further research on samples with known USG and UPC ratio, which were not available in all the cases included in the present study, is needed in order to evaluate the diagnostic performances of the two dipsticks as well as to understand which of the two dipsticks employed in this study is more accurate and which over- or underestimate proteinuria compared with the UPC ratio.

Interestingly, in any case errors were higher than one score. Studies in human medicine reported that errors higher than one score are possible in clinical practice and may account up to 2.4%[4,9,11] However, those studies evaluated not only the analytical variability (as in our study) but also the effect of preanalytical and postanalitycal errors on variability. Evaluation of such a variability was beyond the aims of this study but it could be speculated that also in veterinary practice,

preanalytical errors may occur and, along with the analytical variability reported in the present study, may induce misclassifications higher than one score of positivity.

In conclusion, although dipstick is considered simple and intuitive test, analytical variability may affect the interpretation of results, as well as for any other diagnostic tests. Both the imprecision and the difficulty of interpretation may depend either on intrinsic factors of the pads or on different capability of the operators. The effect of these variables could be considerable in misclassification of samples between two contiguous scores at any level of positivity but misclassification of results between N/T, 1+ and 2+ could be of clinical significance and therefore should be interpreted with caution and confirmed with more sensitive methods such as the UPC ratio. Further studies are warranted in order to evaluate whether automated readers may reduce this variability, or to determine the accuracy of these dipsticks in comparison with a gold standard method such as the UPC ratio and the effect of the variability quantified in this study on clinical decisions.

**Conflict of interest statement**

None of the authors of this paper has a financial or personal relationship with other people or organizations that could inappropriately influence or bias the content of the paper.

**Acknowledgements:**

The Authors are grateful to Prof. Paola Scarpa that provided most of the samples included in this study and Dr. Tiziana Vitiello that performed routine urinalysis.

**References**

1. Grauer GF. IRIS: Proteinuria. Available at: http://www.iris-kidney.com/education/proteinuria.html. Accessed January 22th, 2016.
2. Gunn-Christie RG, Flatland B, Friedrichs KR, et al. ASVCP quality assurance guidelines: control of preanalytical, analytical, and postanalytical factors for urinalysis, cytology, and clinical chemistry in veterinary laboratories. *Vet Clin Pathol*. 2012;41:18-26
3. Stockham SL, Scott MA. *Fundamentals of veterinary clinical pathology*. 2nd ed. Ames, IA: Wiley-Blackwell; 2008:415-494
4. Tighe P. Laboratory-based quality assurance program for near-patient urine dipstick testing, 1990-1997: development, management and results. *Br J Biomed Sci*. 1999;56:6-15.
5. Winkens RA, Leffers P, Degenaar CP, Houben AW. The reproducibility of urinalysis using multiple reagent test strips. *Eur J Clin Chem Clin Biochem*. 1991;29:813-818.
6. European Confederation of Laboratory Medicine. European urinalysis guidelines. *Scand J Clin Lab Invest Suppl*. 2000;231:1-86.
7. Kouri T, Laippala P, Kutter D, Gant V. Quality specifications for ordinal scale measurements with multiproperty (multiple) urine test strips. *Scand J Clin Lab Invest*. 1999;59:523-526
8. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 1977;33:159–174.
9. Tighe P. Urine dry reagent strip “error” rates using different reading methods. *Accred Qual Assur*. 2000;5:488-490.
10. Berg B, Hellsing K, Jagenburg R, Kallner A. Guidelines for evaluation of reagent strips. Exemplified by analysis of urine albumin and glucose concentration using

visually read reagent strips. International Federation of Clinical Chemistry, Scientific Division. *Scand J Clin Lab Invest*. 1989;49:689-699.

11. Rumley A. Urine dipstick testing: comparison of results obtained by visual reading and with the Bayer CLINITEK 50. *Ann Clin Biochem*. 2000;37:220-221.
12. Westgard JO. *Basic method validation*. 2nd ed. Madison, Wis: Westgard QC Inc, 2003.
13. Rossi G, Giori L, Campagnola S, Zatelli A, Zini E, Paltrinieri S. Evaluation of factors that affect analytic variability of urine protein-to-creatinine ratio determination in dogs. *Am J Vet Res*. 2012;73:779-788
14. Froom P, Bieganiec B, Ehrenrich Z, Barak M. Stability of common analytes in urine refrigerated for 24 h before automated analysis by test strips. *Clin Chem*. 2000;46:1384-6.
15. Bauer N, Rettig S, Moritz A. Evaluation the Clinitek status™ automated dipstick analysis device for semiquantitative testing of canine urine. *Res Vet Sci*. 2008;85:467-472
16. Zatelli A, Paltrinieri S, Nizi F, Roura X, Zini E. Evaluation of a urine dipstick test for confirmation or exclusion of proteinuria in dogs. *Am J Vet Res*. 2010;71:235-240.

## Tables and Figures

Table 1 Contingency table of the 174 canine urine samples assayed with dipstick 1 (above) and dipstick 2 (below) and interpreted by the two operators. Concordant results are in bold, discordant results are in italics

| | | | Operator 2 | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | N/T | 1 | 2 | 3 | 4 | Total |
| Operator 1 | Dipstick 1 | N/T | **96** | 0 | 0 | 0 | 0 | 96 |
| | | 1+ | *16* | **33** | *1* | 0 | 0 | 50 |
| | | 2+ | 0 | *2* | **14** | *1* | 0 | 17 |
| | | 3+ | 0 | 0 | *1* | **9** | 0 | 10 |
| | | 4+ | 0 | 0 | 0 | 0 | **1** | 1 |
| | | Total | 112 | 35 | 16 | 10 | 1 | **174** |
| | Dipstick 2 | N/T | **83** | *1* | 0 | 0 | 0 | 84 |
| | | 1+ | *22* | **30** | 1 | 0 | 0 | 53 |
| | | 2+ | 0 | *6* | **13** | 0 | 0 | 19 |
| | | 3+ | 0 | 0 | *3* | **9** | 0 | 12 |
| | | 4+ | 0 | 0 | 0 | *2* | **4** | 6 |
| | | Total | 105 | 37 | 17 | 11 | 4 | **174** |

393 Table 2 Raw data of agreement between the two dipsticks in scoring protein pads using the results of the consensus between
394 operators. Concordant results are in bold, discordant results are in italics
395

| | | Dipstick2 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N/T | N/T or 1+ | 1+ | 1+ or 2+ | 2+ | 2+ or 3+ | 3+ | 3+ or 4+ | 4+ | Total |
| Dipstick1 | N/T | **82** | *11* | *3* | 0 | 0 | 0 | 0 | 0 | 0 | 96 |
| | N/T or 1+ | 0 | **9** | *7* | 0 | 0 | 0 | 0 | 0 | 0 | 16 |
| | 1+ | *1* | *3* | **20** | *7* | *2* | 0 | 0 | 0 | 0 | 33 |
| | 1+ or 2+ | 0 | 0 | 0 | **0** | *3* | 0 | 0 | 0 | 0 | 3 |
| | 2+ | 0 | 0 | 0 | 0 | **8** | *3* | *3* | 0 | 0 | 14 |
| | 2+ or 3+ | 0 | 0 | 0 | 0 | 0 | **0** | *2* | 0 | 0 | 2 |
| | 3+ | 0 | 0 | 0 | 0 | 0 | 0 | **4** | *1* | *3* | 8 |
| | 3+ or 4+ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **1** | 0 | 1 |
| | 4+ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **1** | 1 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | TOTAL | 83 | 23 | 30 | 7 | 13 | 3 | 9 | 2 | 4 | **174** |

396