

Validation of photographic fecal scoring in puppies

Chloé Chauvel^{a,1}, Ilyas Bendahmane^{a,b,1}, Emma Vellard^a, Faouzi Lyazrhi^c,
Annabelle Meynadier^b, Hanna Mila^{a,*}, Aurélien Grellet^a

^a Neocare, Université de Toulouse, ENVT, Toulouse, France

^b GenPhySE, Université de Toulouse, INRAE, ENVT, Castanet Tolosan F-31326, France

^c Biostatistiques, Université de Toulouse, ENVT, Toulouse, France

ARTICLE INFO

Keywords:

Fecal score
Puppy
Feces
Diarrhea
Photograph
Method validation

ABSTRACT

Background: Diarrhea is a common issue in puppies and compromises the well-being of animals and their owners. To date, objective assessment of fecal consistency has relied on a fecal scoring scale applied to fresh feces; however, given the challenges associated with accessing fresh stool samples, digital scoring presents an interesting potential alternative. This study aimed to test the validity of using photographs to approximate the *in situ* fresh fecal scoring scale.

Methods: In a first experiment, 40 fresh feces were collected from puppies and scored *in-situ* with a 13-point fecal scale. Feces were then photographed and later assessed on picture. Agreement between the two scoring methods was evaluated using a Bland-Altman plot and kappa coefficients. In a second experiment, four observers scored independently 98 stool photographs. Intraclass correlation coefficient was calculated to assess inter-observer agreement. Intra-observer reproducibility was also explored with a Bland-Altman plot and kappa coefficients.

Results: Our findings showed good agreement between the two methods (weighted kappa coefficient of 0.93 (95 % CI [0.87, 0.98]) and no significant bias in the Bland-Altman plot). Our study also showed good inter-observer (ICC of 0.83 (95 % CI [0.78, 0.87])) and intra-observer (weighted kappa coefficient of 0.93 (95 % CI [0.91, 0.96]) and no significant bias in the Bland-Altman plot) agreements.

Conclusions: The puppy fecal scoring scale employed in the present study appeared to be reliably applicable to photographs for objectively assessing fecal consistency in puppies. It is a promising solution for veterinary practices and future research studies.

1. Background

Gastro-intestinal disorders are pervasive in dogs. One study showed that 29.5 % of dog consultations at English primary-care veterinary practices involved digestive disorders, making the digestive tract the most commonly affected organ system after the integument (O'Neill et al., 2014). Digestive disorders are particularly problematic in puppies, leading to a common presentation of diarrhea in this population (Sævik et al., 2012; Grellet et al., 2014; Pugh et al., 2017). A previous survey showed that diarrhea is the most frequent health issue observed in puppies within the first weeks after sale, reported by 32 % of French dog breeders (NeoCare, 2020). Diarrhea in puppies at the weaning period (i. e., between 5 and 14 weeks of age) can have many origins, such as parasites, viruses, dietary disorders, or environmental factors (Grellet et al., 2014). It causes significant discomfort, compromises owner's

well-being, and is associated with impaired growth and increased risk of mortality (Grellet et al., 2012). Objective evaluation of fecal consistency is useful both in veterinary practices and for research projects for standardizing diagnosis, monitoring clinical course of disease and assessing the efficacy of treatments. It can be achieved using a fecal scoring system (Rolfe et al., 2002; Westermarck, 2016; Cavett et al., 2021). A puppy scale, designed to adapt to specific stool consistency of puppies around weaning, has previously been published (Grellet et al., 2012). As young puppies generally have less formed feces than adults, the use of a specific scale prevents over-estimating the incidence of diarrhea. Indeed, this scale enables a more refined assessment of stool consistency, through a greater number of categories than adult scales, allowing more accurate monitoring of subtle variations in bowel function around weaning (Grellet et al., 2012).

Stool consistency scoring tools are intended for direct *in situ*

* Correspondence to: NeoCare Batiment 10.1, 23, chemin des capelles, CEDEX, Toulouse 31076, France.

E-mail address: hanna.mila@envt.fr (H. Mila).

¹ Co-first authors

observation of fresh stool (Grellet et al., 2012; Vandenplas et al., 2017). However, accessing fresh stool samples can be challenging because of reluctance to manipulate feces (Vandeputte et al., 2017; McSweeney et al., 2020). Additionally, transportation raises the question of sample preservation, and the effect of temperature variations and handling on stool consistency assessment, which has yet to be studied. Furthermore, it is challenging to collect liquid feces, in which case an *in-situ* assessment is necessary. Logistical constraints may restrict repeated access to stool samples, posing challenges for monitoring chronic cases or population-level observation.

Photography offers a convenient and readily available solution to capture the appearance of stool immediately after bowel movement. This would prevent transportation- and collection-related changes in the feces, save the owner time and travel, avoid stress in the animals if traveling to a clinic, and likely decrease zoonotic exposure. Indeed, veterinary telemedicine is becoming an increasingly widespread practice (Gyles, 2019; Kastelic and Ogilvie, 2021; Smiley, 2022). Dog owners could submit multiple pictures over time for remote assessment, giving veterinarians a useful tool for triage and a broader view of a puppy's health status between face-to-face consultations. However, image-based evaluation may introduce bias compared to direct observation of stool samples, necessitating validation of this method. Indeed, two-dimensional visualization could lead to a loss of visual information, and image pixilation could affect the estimation of stool texture and shape.

The present study aimed:

- To evaluate the agreement between direct and photographic assessments of stool consistency in puppies during the weaning period, using a validated fecal scoring scale (FSS).
- To assess the reproducibility of the photographic evaluation of stool consistency in puppies, including inter-observer agreement, the effect of a training session on inter-observer repeatability, and intra-observer agreement.

2. Methods

2.1. Evaluation of agreement between direct and photographic assessment of stool

This study was evaluated by a local ethical committee (CEEA SSA N°115).

In this first experiment, feces from 25 two-month-old (ranging from 7 to 11 weeks of life) Golden Retriever puppies from the Toulouse Gold Cohort were assessed (Toulouse GOLD Cohort: Accueil, 2025). These puppies were born and raised in a dedicated nursery as part of a French research program and benefited from regular veterinary care. From 4 weeks of age (nutritional weaning onset), they were fed the same solid food and received the same antiparasitic treatments and vaccination protocol. They were not yet separated from their mother at the time of the study. An operator evaluated 40 randomly selected stool samples immediately after spontaneous defecation between April 3rd and May 1st, 2024. The feces were scored according to their consistency and shape using the 13-point puppy FSS, ranging from 1 (completely liquid) to 13 (dry and hard feces) (Grellet et al., 2012). Immediately after scoring, pictures of the assessed feces were taken by the operator with a smartphone camera (iPhone 11 Pro, 12 Mpx camera, Apple®, Cupertino, California, USA), thereby creating a library of 40 still photographs (JPG files, 96 dpi, see examples in Additional file 2). The photographs were always taken in the same way, perpendicular to the stool surface, placing the phone at a distance of approximately 20–30 cm. The photographs were taken during the day to ensure sufficient natural light. The background was the floor on which the puppy had defecated: grey tiles or woodchips in the indoor kennel, which provided a color contrast with the feces. As all photographs were taken in field conditions and in the presence of the animals, no scale was included in the frame to assist

with measurements, color accuracy, or size fidelity. One stool photograph (n°29) was removed from the study due to deviation from the established methodology before photographing (handling, contamination with wood chips).

The same operator subsequently scored the photo library on his smartphone screen using the same FSS. To avoid recall bias (Campbell et al., 2015), this step was performed more than one year after fresh fecal scoring (July 9th, 2025).

All statistical analyses were performed using R software (version 4.5.1) (Team, 2014; R: The R Project for Statistical Computing, 2025). To assess agreement between the scoring method on fresh stool (FS) and photographed stool (PS), a Bland and Altman diagram was created using ‘BlandAltmanLeh’ package. In Bland-Altman plot, mean difference between the two measurement methods ($PS - FS$) is called bias (b). Assuming a normal distribution of differences, approximately 95 % of the differences are expected to be between $b \pm 2sd$, where sd is the standard deviation of the differences. The limits of this range are called limits of agreement (LoAs). 95 % confidence intervals of the b and LoAs are also shown in the Bland and Altman diagram (Martin Bland and Altman, 1986). Precise normality of differences is not required for this test, and thus only visual assessment was performed (Bland and Altman, 1999). For the present study, maximum acceptable b and LoAs were set at ± 1 and ± 3 points of the FSS, respectively. Weighted and unweighted kappa coefficients were then calculated using ‘psych’ package. The difference between weighted and unweighted kappa coefficients is that, in a weighted kappa, the higher the degree of scoring error, the more it will negatively impact the kappa value; in an unweighted kappa, each classification error is treated equally. The default quadratic weights were used. Kappa coefficient interpretation grid is shown in Table 1 (Landis and Koch, 1977).

2.2. Reproducibility of photographic stool assessment

2.2.1. Inter-observer reproducibility

A previously available library of 98 photographs of stool coming from puppies of various breeds and ages was used for this study (more details on the population are available in Grellet et al., 2012). Pictures were carefully chosen to ensure a selection of different consistencies representative of the entire FSS in proportions comparable with those observed in Grellet et al., 2012. Four observers (Obs1 to Obs4), including the observer involved in the previous experiment, three veterinarians and one non-veterinarian researcher, took part in this second experiment. Obs4 was an experienced user of the FSS with *in-situ* stool samples, while the three other observers had never used it before. Since the inter-observer repeatability study was conducted chronologically before the study comparing the two scoring methods, the observer in the previous study (Obs1) was still inexperienced at that stage. Each observer scored independently this photo library on a computer screen in Toulouse, France, around the same period, using the FSS described above. The fecal exam scoring sheet (pictures and descriptions) was provided to each observer in digital format; no further training was provided at this time.

For each picture, the difference between Obs1, Obs2 and Obs3 score and the Obs4 (reference observer) score was calculated ($ObsX - Obs4$). Inter-operator reproducibility for this scoring session (S1) was assessed

Table 1
Interpretation of Kappa coefficient according to Landis and Koch.

Kappa coefficient	Interpretation
< 0	Poor agreement
0–0.20	Slight agreement
0.21–0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.8	Substantial agreement
0.81–1.00	Almost perfect agreement

by calculating intraclass correlation coefficient (ICC) using 'performance' package. A linear mixed model was used ('lme4' package), with the observer as a fixed variable and the stool photograph as a random variable using the following formula: $\text{score} \sim \text{observer} + (1 | \text{photo-graph})$. ICC interpretation grid is shown in Table 2 (Koo and Li, 2016).

2.2.2. Influence of a training session on inter-observer reproducibility

In order to assess the influence of training on the agreement of assessments, Obs4 provided the three other observers with training on the practical application of the FSS two months after S1. During this training, Obs4 explained in more detail the differences in consistency described in the FSS, warned about the bias associated with stool color and the presence of mucus, and answered the observers' questions, using practical case studies illustrated with pictures. Immediately after the training, each observer scored for the second time the same 98 photographs of stool during a second scoring session (S2).

Post-training ICC (ICC2) was calculated in the same way as for the pre-training ICC (ICC1). The influence of training on inter-observer agreement was assessed based on the 95 % confidence interval overlap of ICCs.

2.2.3. Intra-observer reproducibility

Intra-observer reproducibility of the PS method was assessed by comparing the fecal scores assigned by Obs4 during S1 and S2. For each picture, the difference between S1 and S2 scores was calculated (S1 – S2). To assess intra-observer agreement, a Bland and Altman diagram was created in the same way as in the between-method agreement experiment. The same maximum acceptable b and LoAs were used as above. The normality test and calculation of kappa coefficients were done as previously.

3. Results

3.1. Evaluation of agreement between direct and photographic assessment of stool

Fecal scores of the 39 stool samples ranged from 3 to 10, with a median value of 7 for both methods (Fig. 1 A). All results of fecal scores are shown in an additional file (see Additional file 1).

Differences between PS and FS scores followed a normal distribution (Fig. 1B). Fecal scores evaluated on PS were equal to those on FS in 74.4 % of cases (29/39), and 1 point lower in 15.4 % of cases (6/39). According to the Bland-Altman diagram for between-method agreement, the b of the method was -0.10 (95 % CI [-0.31, 0.11]) (Fig. 1C). Lower and upper LoAs were -1.36 (95 % CI [-1.72, -1.00]) and 1.15 (95 % CI [0.79, 1.51]), respectively. Differences between PS and FS scores seemed visually more significant for higher fecal scores, suggesting better agreement for lower scores. Unweighted kappa coefficient was 0.69 (95 % CI [0.52, 0.85]) and weighted kappa coefficient was 0.93 (95 % CI [0.87, 0.98]).

3.2. Reproducibility of photographic stool assessment

The 98 feces studied during S1 had fecal scores ranging from 1 to 12 according to the reference observer (Obs4; Fig. 2A). Median fecal scores assigned by Obs1, Obs2, Obs3 and Obs4 during S1 were 8, 9, 8 and 7.5,

respectively (see Additional file 1 for all data). Fecal scores assigned by Obs1, Obs2 and Obs3 during S1 were equal to those assigned by Obs4 in 26.5 % (26/98), 28.6 % (28/98) and 40.8 % (40/98) of cases, respectively. Mean differences with Obs4 scores for Obs1, Obs2 and Obs3 scores ($\pm$ standard deviation) were 0.7 ± 1.7 , 1.1 ± 1.7 and 0.28 ± 1.2 , respectively (Fig. 2B). ICC1 was 0.83 (95 % CI [0.78, 0.87]).

Median fecal scores assigned by Obs1, Obs2, Obs3 and Obs4 during S2 were 8, 7, 8 and 8, respectively (Fig. 3A). ICC2 was 0.86 (95 % CI [0.80, 0.89]). Overlap of ICC1 and ICC2 95 % confidence intervals was marked (Fig. 3B), and the two ICCs were therefore considered not significantly different.

Median Obs4 fecal scores assigned during S1 and S2 were 7.5 and 8, respectively (Fig. 4A). Differences between S1 and S2 scores assigned by Obs4 followed a normal distribution (Fig. 4B). Fecal scores assigned by Obs4 during S2 were equal to those assigned during S1 in 59.2 % of cases (58/98). They were one point lower in 17.3 % of cases (17/98) and one point higher in 15.3 % of cases (15/98). According to the Bland-Altman diagram for intra-observer agreement (Obs 4 for S1 and S2), the bias was -0.08 (95 % CI [-0.26, 0.09]). Lower and upper LoAs were -1.79 (95 % CI [-2.09, -1.48]) and 1.62 (95 % CI [1.32, 1.92]), respectively (Fig. 4C). Differences between S1 and S2 scores seemed visually more significant for higher fecal scores, suggesting better agreement for lower scores. Unweighted kappa coefficient was 0.53 (95 % CI [0.42, 0.64]) and weighted kappa coefficient was 0.93 (95 % CI [0.91, 0.96]).

4. Discussion

4.1. Choice of the method

In the present study, fecal consistency was evaluated using fecal scoring system (FSS) developed by Grellet et al. (2012), one of the authors specifically for puppies at the weaning period. Fecal scoring systems are widely used in both human and veterinary medicine to provide a standardized and reproducible assessment of stool consistency, which in turn supports the diagnosis of diarrhea. For example, in infants, diarrhea is often defined in clinical guidelines by stool frequency, but this approach has been criticized as being too subjective; consequently, quantitative stool scales are now preferred (Levine et al., 2017; Vandenplas et al., 2017). In the same way, the FSS allows a more objective evaluation of fecal quality in puppies. Furthermore, comparison of fecal scores with objective measures such as stool water content, as performed during the development of the Bristol Stool Scale in humans (Blake et al., 2016), could further validate the method in veterinary research.

4.2. Photographic versus direct assessment of stool

This study highlighted the potential for using photograph-based scoring in daily medical practice and research projects to remotely evaluate fecal scores in puppies. Indeed, in the corresponding Bland and Altman plot, b fell within the maximum acceptable bias of ± 1 and the 95 % confidence interval of observed b included zero, indicating no significant difference in scoring between FS and PS methods. The lower and upper LoAs were both within the acceptable LoAs range of ± 3 . Although relatively wide, these LoAs were also considered clinically acceptable, as only differences of more than 3 points would alter fecal consistency classification. This interpretation is consistent with Grellet et al. (2012), who defined similar score groupings (1–3 liquid, 4–6 pasty, 7–9 formed, 10–12 drier formed, 13 dry and hard).

Additionally, in our study the weighted kappa coefficient indicated "almost perfect" agreement according to Landis and Koch (Landis and Koch, 1977). Quadratic weights were applied, as they assign progressively greater penalties for larger disagreements, thereby better reflecting the clinical importance of substantial discrepancies in stool consistency assessments. The use of quadratic weights therefore,

Table 2
Interpretation of intraclass correlation coefficient according to Koo and Li.

ICC	Interpretation
< 0.5	Poor reliability
0.5–0.75	Moderate reliability
0.75–0.9	Good reliability
> 0.9	Excellent reliability

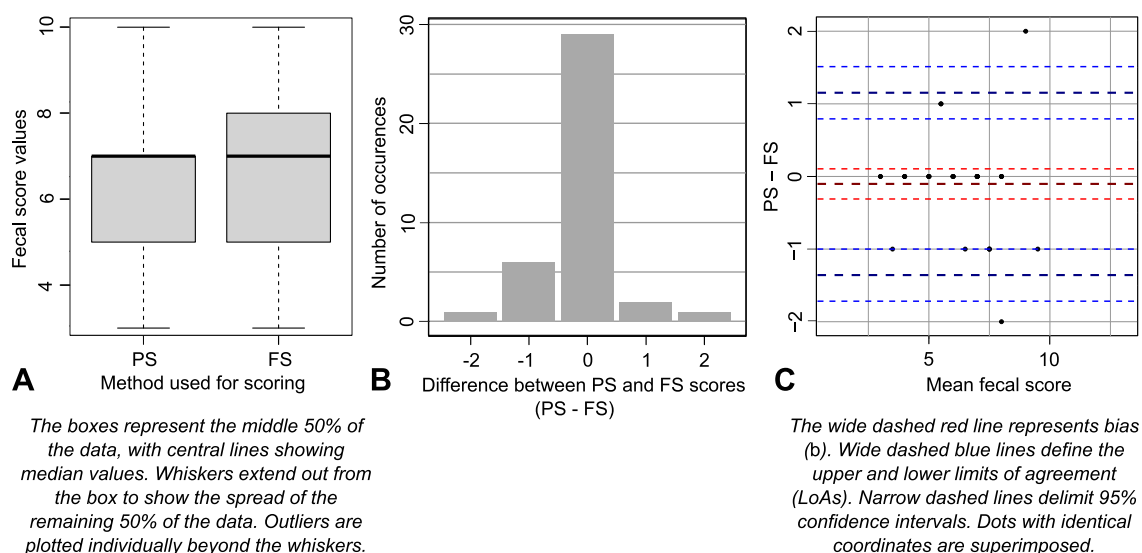


Fig. 1. Between-method agreement

Box plot represents fecal scores obtained on 40 stools according to the scoring method: on photographed stools (PS) and on fresh stools (FS) (1A). Histogram represents differences between PS and FS scores (PS - FS, 1B). And dot-plot represents the results of the Bland-Altman test on those differences in relation with the average score obtained using both methods (1C).

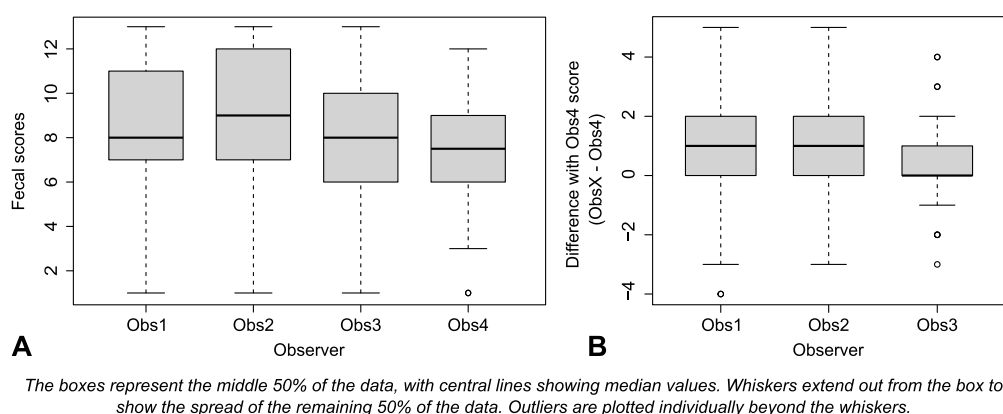


Fig. 2. Inter-observer agreement before a training session

Box and whisker plots represent fecal scores obtained on 98 stool photographs according to the observer (Obs1 to Obs4) (2A) and differences between fecal scores assigned by non-trained observers (Obs1 to Obs3) and those assigned by the reference observer (Obs4) (ObsX - Obs4, 2B).

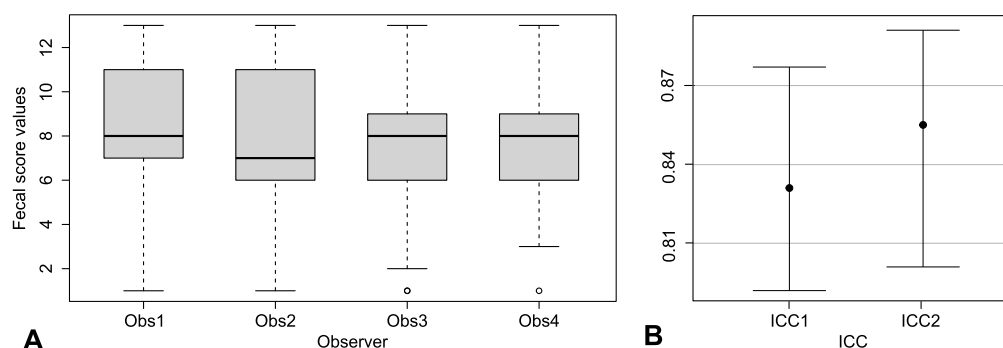


Fig. 3. Influence of a training session on inter-observer agreement

Box and whisker plot represents fecal scores obtained on 98 stool photographs during post-training evaluation according to the observer (Obs1 to Obs4) (3A). Dot-plot represents intraclass correlation coefficients (ICCs) before (ICC1) and after (ICC2) the training session according to their 95 % confidence intervals (3B).

appropriately emphasizes clinically meaningful disagreements and provides a more realistic measure of agreement. However, this weighting scheme may also yield slightly higher kappa values than linear

weighting, as it assigns greater tolerance to small, clinically negligible differences.

Based on visual observation of the diagram, greater variations in

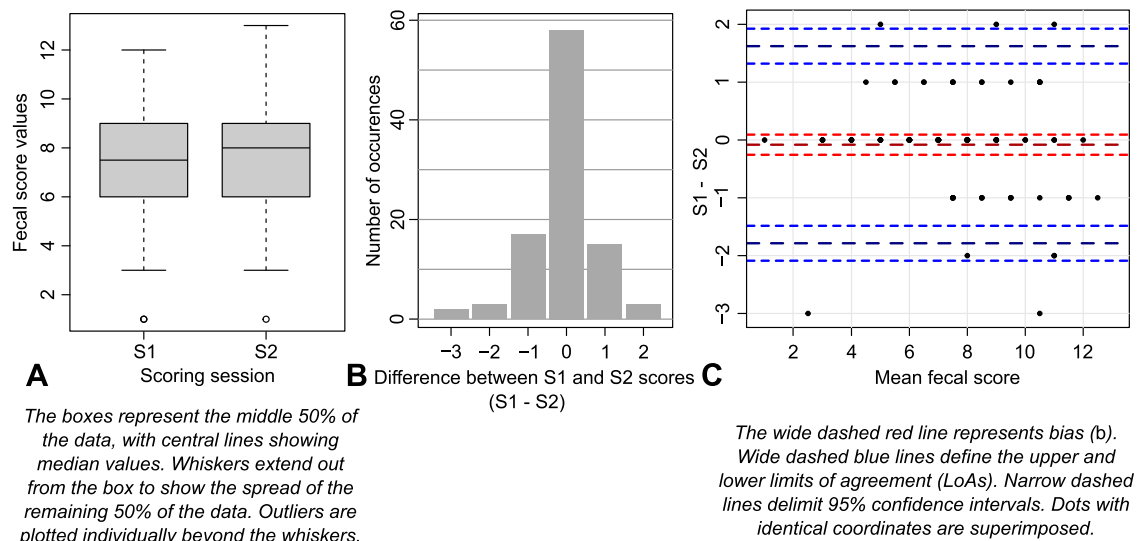


Fig. 4. Intra-observer agreement

Box and whisker plot of fecal scores attributed to 98 stool photographs by reference observer (Obs4) according to the scoring session (S1 and S2) (4A). Histogram represents differences between S1 and S2 scores attributed by Obs4 (S1 – S2, 4B). And dot-plot represents the results of the Bland-Altman test on those differences in relation with the average score obtained during the two sessions (4C).

scoring were noted between the two methods for higher fecal scores (>6) compared to lower scores (≤ 5). Although this observation is based solely on visual interpretation, it would appear that photographic assessment of liquid and thus potentially abnormal feces may be even more accurate than for firm feces.

To minimize bias associated with visual memory, the evaluations on fresh stool and photographed stool were performed more than a year apart. Recall bias is a well-documented phenomenon. For instance, a study demonstrated that pathologists could recall slides viewed four weeks earlier, with recall rates sufficient to influence intra-observer study design (Campbell et al., 2015). No published article allows us to verify that there is no memory bias after one year. Therefore, although unlikely, it remains possible that the agreement between methods is overestimated in this study. A significant limitation of this study is the small number of available samples (39) and the fact that only one evaluator looked at them *in situ*, which could result in wide confidence intervals. On the other hand, two different analyses were performed to assess agreement between the two studied methods (kappa and Bland-Altman plot), showing the consistency of our findings. In addition, as each feces emitted was included in this study, the entire scoring scale was not represented; no fecal score below 3 and above 10 was included, which means that the fidelity of the technique cannot be extrapolated to extreme values. Finally, feces of Golden Retriever puppies only were scored in the first part of our study. Future studies, on large sample sizes and on several different breeds, should address these missing scores to confirm their agreement with the reference method.

4.3. Intra and inter-observer reproducibility

We demonstrated that the photograph-based scoring method exhibited good intra-observer and inter-observer reproducibility. The latter was assessed using a Bland-Altman analysis. The 95 % confidence interval of observed b included zero, indicating no significant difference in scoring for Obs4 between S1 and S2. The lower and upper LoAs were both well within the maximum LoA range of ± 3 . The weighted kappa coefficient indicated “almost perfect” agreement according to Landis and Koch (Landis and Koch, 1977). These results confirmed excellent intra-observer reproducibility for this observer. More tests and more operators are needed to understand the actual variability of this test. Based on visual observation of the diagram, it would appear that higher

fecal scores (>7) were associated with greater variations in scoring between the two sessions compared to lower scores. This supports the hypothesis raised above that feces with a low fecal score (including scores ≤ 5 , considered pathological) are rated more accurately on photographs than those with a high fecal score, highlighting clinical potential of the remote method.

Our study also showed good inter-observer reproducibility. Using only the scoring grid without prior training, the ICC corresponded to the “good reliability” category defined by Koo and Li (Koo and Li, 2016). A training session seemed to help further improve the ICC, though the increase was not statistically significant, as 95 % confidence intervals for ICC1 and ICC2 overlapped substantially. Thus, the FSS proved easy to use without requiring additional training.

4.4. Perspectives

Taken together, our results suggest that fecal consistency assessment may be performed on photographs with seemingly satisfactory fidelity. This aligns with findings in infants and young children, for whom a specific stool scale was developed and shown to have good agreement with fresh stool scoring (Huysentruyt et al., 2019; Aman et al., 2021; Hofman et al., 2022). However, the kappa and ICC coefficients reported in these studies were derived from a much larger number of samples and observations, assessed by a greater number of observers, although using a scale with far fewer categories. Our findings on a limited number of subjects are thus preliminary in veterinary medicine, and further studies are needed to confirm these results.

Additionally, the use of digital photographs opens up the possibility of integrating artificial intelligence into fecal score interpretation in the future. This could significantly reduce the workload of research teams and accelerate advancements in the fields of veterinary gastroenterology and pediatrics.

5. Conclusion

In conclusion, the results of this study support the use of the puppy fecal scoring scale developed by Grellet et al. on stool photographs. This preliminary study, conducted with a single observer, showed fecal scores consistent with those obtained on fresh stool samples. Inter-observer reproducibility was satisfactory based on the small number

of observers involved in this study. This easy-to-use technique can be performed without prior training, making it accessible to all practitioners and researchers. Thus, this method offers a promising solution for monitoring canine fecal consistency to detect diarrhea in veterinary practices or research projects.

CRedit authorship contribution statement

Aurélien Grellet: Writing – review & editing, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Hanna Mila:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Annabelle Meynadier:** Writing – review & editing, Supervision, Conceptualization. **Faouzi Lyazrhi:** Writing – review & editing, Validation, Methodology, Formal analysis. **Emma Vellard:** Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ilyas Bendahmane:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization. **Chloé Chauvel:** Writing – original draft, Visualization, Resources, Methodology, Investigation, Formal analysis.

Author contributions

The authors contributed to the project as follows: Conceptualisation – IB, EV, AG, HM; Scoring – IB, EV, AH, HM; Data registration – IB, EV; Data Analysis – CC, EV, FL; Original writing – CC; Review and editing – AG, AM, HM, FL, IB. All authors read and approved the final manuscript.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used DeepL Translator in order to improve English wording. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Ethics, consent to participate, and consent to publish declarations

This study was approved by local ethics committee *Animal Health and Science Ethics Committee* (“Comité d’éthique science et santé animales”, CEEA SSA N°115). Some stool pictures were taken from client-owned animals as part of a previous study with informed consent from their owner. Other animals involved in this study were owned by the veterinary school.

Funding

This work was supported by the Occitania Region and VISIO Foundation.

Declaration of Competing Interest

Authors report no conflict of interest in this study.

Acknowledgements

The authors would like to express their gratitude the Visio Foundation (*Fondation VISIO*, no date) and Occitanie Region for their financial support on this project. Special thanks are extended to all NeoCare team members and volunteers who ensure the smooth running of the Toulouse Gold Cohort, as well as to all adopters and foster families involved in the associated projects.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.prevetmed.2025.106729](https://doi.org/10.1016/j.prevetmed.2025.106729).

Data Availability

The datasets supporting the conclusions of this article are included within the article (and its additional file).

References

- Aman, B.A., et al., 2021. Real time versus photographic assessment of stool consistency using the brussels infant and toddler stool scale: are they telling us the same? *Pediatr. Gastroenterol. Hepatol. Nutr.* 24 (1), 38–44. <https://doi.org/10.5223/pghn.2021.24.1.38>.
- Blake, M.R., Raker, J.M., Whelan, K., 2016. Validity and reliability of the Bristol Stool Form Scale in healthy adults and patients with diarrhoea-predominant irritable bowel syndrome. *Aliment. Pharmacol. Ther.* 44 (7), 693–703. <https://doi.org/10.1111/apt.13746>.
- Bland, J.M., Altman, D.G., 1999. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* 8 (2), 135–160. <https://doi.org/10.1177/096228029900800204>.
- Campbell, W.S., et al., 2015. Visual memory effects on intraoperator study design: determining a minimum time gap between case reviews to reduce recall bias. *Am. J. Clin. Pathol.* 143 (3), 412–418. <https://doi.org/10.1309/AJCPUC3TYMS3QOBM>.
- Cavett, C.L., et al., 2021. Consistency of faecal scoring using two canine faecal scoring systems. *J. Small Anim. Pract.* 62 (3), 167–173. <https://doi.org/10.1111/jsap.13283>.
- Grellet, A., et al., 2012. Validation of a fecal scoring scale in puppies during the weaning period. *Prev. Vet. Med.* 106 (3), 315–323. <https://doi.org/10.1016/j.prevetmed.2012.03.012>.
- Grellet, A., et al., 2014. Risk factors of weaning diarrhea in puppies housed in breeding kennels. *Prev. Vet. Med.* 117 (1), 260–265. <https://doi.org/10.1016/j.prevetmed.2014.07.016>.
- Gyles, C., 2019. Veterinary telemedicine. *Can. Vet. J.* 60 (2), 119–122.
- Hofman, Y.M.C., et al., 2022. Intra-rater variability of the brussels infants and toddlers stool scale (BITSS) Using Photographed Stools. *J. Pediatr. Gastroenterol. Nutr.* 75 (5), 584–588. <https://doi.org/10.1097/MPG.00000000000003586>.
- Huysentruyt, K., et al., 2019. The brussels infant and toddler stool scale: A study on interobserver reliability. *J. Pediatr. Gastroenterol. Nutr.* 68 (2), 207–213. <https://doi.org/10.1097/MPG.00000000000002153>.
- Kastelic, J., Ogilvie, T., 2021. Veterinary telemedicine is not only here to stay, it's poised to grow and likely exponentially. *Can. Vet. J.* 62 (12), 1277–1279.
- Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15 (2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>.
- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33 (1), 159–174.
- Levine, G.A., et al., 2017. Defining pediatric diarrhea in low-resource settings. *J. Pediatr. Infect. Dis. Soc.* 6 (3), 289–293. <https://doi.org/10.1093/jpids/pix024>.
- Martin Bland, J., Altman, DouglasG., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327 (8476), 307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8).
- McSweeney, B., et al., 2020. In search of stool donors: a multicenter study of prior knowledge, perceptions, motivators, and deterrents among potential donors for fecal microbiota transplantation. *Gut Microbes* 11 (1), 51–62. <https://doi.org/10.1080/19490976.2019.1611153>.
- NeoCare (2020) Relationship between breeder and buyer: NeoCare survey results. Disponible sur: (<https://neocare.pro/wp-content/uploads/2023/05/Enquete-NeoCare-17-Relationnel.pdf>) (Consulté le: 7 novembre 2024).
- O'Neill, D.G., et al., 2014. Prevalence of disorders recorded in dogs attending primary-care veterinary practices in England. *PLOS One* 9 (3), e90501. <https://doi.org/10.1371/journal.pone.0090501>.
- Pugh, C.A., et al., 2017. Incidence rates and risk factor analyses for owner reported vomiting and diarrhoea in labrador retrievers – findings from the dogslife cohort. *Prev. Vet. Med.* 140, 19–29. <https://doi.org/10.1016/j.prevetmed.2017.02.014>.
- R: The R Project for Statistical Computing (sans date). Disponible sur: (<https://www.r-project.org/>) (Consulté le: 13 janvier 2025).
- Rolfe, V.E., et al., 2002. Relationship between faecal character and intestinal transit time in normal dogs and diet-sensitive dogs. *J. Small Anim. Pract.* 43 (7), 290–294. <https://doi.org/10.1111/j.1748-5827.2002.tb00075.x>.
- Sævik, B.K., Skancke, E.M., Tranterud, C., 2012. A longitudinal study on diarrhoea and vomiting in young dogs of four large breeds. *Acta Vet. Scand.* 54 (1), 8. <https://doi.org/10.1186/1751-0147-54-8>.
- Smiley, A., 2022. Asynchronous veterinary telemedicine. *Vet. Clin. North Am. Small Anim. Pract.* 52 (5), 1135–1140. <https://doi.org/10.1016/j.cvsm.2022.05.003>.
- Team, R. (2014) R: A language and environment for statistical computing., *MSOR connections* [Preprint]. Disponible sur: (<https://www.semanticscholar.org/paper/R%3A-A-language-and-environment-for-statistical-Team/659408b243cec55de8d0a3bc51b81173007aa89b>) (Consulté le: 31 octobre 2024).

- Toulouse GOLD Cohort: Accueil (2025) Toulouse GOLD Cohort. Disponible sur: (<https://toulousegoldcohort.envt.fr/>) (Consulté le: 28 octobre 2024).
- Vandenplas, Y., et al., 2017. Development of the Brussels Infant and Toddler Stool Scale ('BITSS'): protocol of the study. *BMJ Open* 7 (3), e014620. <https://doi.org/10.1136/bmjopen-2016-014620>.
- Vandeputte, D., et al., 2017. Practical considerations for large-scale gut microbiome studies. *FEMS Microbiol. Rev.* 41 (1), S154–S167. <https://doi.org/10.1093/femsre/flux027>.
- Westermarck, E., 2016. Chronic Diarrhea in Dogs: what do we actually know about it? *Top. Companion Anim. Med.* 31 (2), 78–84. <https://doi.org/10.1053/j.tcam.2016.03.001>.