

# Evaluation of Intra- and Inter-observer Measurement Variability of a Radiographic Stifle Osteoarthritis Scoring System in Dogs

Marlis Wessely<sup>1</sup> Andreas Brühshwein<sup>2</sup> Eva Schnabl-Feichter<sup>1</sup>

<sup>1</sup>Department for Companion Animals and Horses, University Clinic for Small Animals, Small Animal Surgery, University of Veterinary Medicine, Vienna, Austria

<sup>2</sup>Clinic of Small Animal Surgery and Reproduction, Centre of Clinical Veterinary Medicine, Ludwig-Maximilians-University Munich, Munich, Germany

Address for correspondence Marlis Wessely, DVM, Department for Companion Animals and Horses, University Clinic for Small Animals, Small Animal Surgery, University of Veterinary Medicine, Veterinärplatz 1, 1210 Vienna, Austria (e-mail: Marlis.Wessely@vetmeduni.ac.at).

Vet Comp Orthop Traumatol 2017;30:377–384.

## Abstract

**Objectives** To evaluate the intra- and inter-observer measurement variability of an existing osteoarthritis (OA) stifle scoring system.

**Methods** Paired caudocranial and mediolateral canine stifle radiographs were selected randomly. A total of 15 assessment points were evaluated independently and graded twice (integer numeric scale: 1–4) at an interval of 2 weeks by three observers with different levels of experience. The grades for each of the 15 factors were summed to obtain the OA score for each patient.

**Results** The 15 independent assessment points measured by the three observers showed high reproducibility and low intra-observer variability. Inter-observer variability was also low (mean:  $1.09 \pm 4.99$ , 95% CI [confidence interval]:  $-0.35$  to  $2.55$ ). The most discordant ratings among the three observers involved sesamoid bones of gastrocnemius muscle (assessment point 11 of 15) and popliteal surface of femur (assessment point 10 of 15).

**Clinical Significance** A validated and feasible OA scoring method is prerequisite for reliable radiographic assessment of OA progression. The low overall inter- and intra-observer variabilities among the 15 independent measures of the OA scoring system presented herein support its feasibility for application in clinical practice as an objective tool for radiographic scoring of stifle OA.

## Keywords

- ▶ stifle joint
- ▶ radiography
- ▶ osteoarthritis
- ▶ dogs
- ▶ cruciate ligament disease

## Introduction

Since a radiographic osteoarthritis (OA) grading scale was developed about three decades ago, several other clinical and epidemiological research veterinarians have sought to develop improved scoring systems to more accurately and conveniently assess morphological changes of stifle joints.<sup>1</sup> A review of the literature identified 22 different OA scoring systems for canine and feline stifle joints (▶ **Table 1**). More-

over, several of those studies showed substantial variability among the results obtained with the various radiographic OA scoring systems.<sup>2–6</sup> Thus, a scoring method is necessary to obtain comparable results from studies of stifle OA performed by different groups.

An inevitable prerequisite for the application of these grading scales is appropriate reproducibility of the features recorded by the radiographic technique. Among the studies

received  
September 16, 2016  
accepted after revision  
July 16, 2017

Copyright © 2017 Schattauer

DOI <https://doi.org/10.3415/VCOT-16-09-0134>.  
ISSN 0932-0814.

**Table 1** Overview of the published osteoarthritis (OA) scoring systems

Number	Year	Author(s)	AP	Grading	Observer	Evaluations	Intra-observer	Inter-observer
1	1984	Shires and colleagues <sup>1</sup>	3	0–3	n/a	n/a	No	No
2	1986	Schnell <sup>12</sup>	7	0–3	1	1	No	No
3	1988	Bennett and colleagues <sup>21</sup>	9	0–5	n/a	n/a	No	No
4	1989	Mullen and colleagues <sup>14</sup>	n/a	Mild/ moderate/ marked	n/a	n/a	No	No
5	1992	Vasseur and Berry <sup>3</sup>	19	0–3	1	1	No	No
6	1992	Brunnberg and colleagues <sup>15</sup>	7	0–3	1	1	No	No
7	1994	Mayer <sup>40</sup>	9	1–6	1	1	No	No
8	1995	Coetzee and colleagues <sup>41</sup>	n/a	Mild/ moderate/ advanced	n/a	n/a	No	No
9	1996	Garrels <sup>4</sup>	7	0–5	n/a	n/a	No	No
10	1996	Chauvet and colleagues <sup>13</sup>	n/a	1–4	1	1	No	No
11	1999	De Rooster and colleagues <sup>42</sup>	n/a	0–3	1	1	No	No
12	2000	Mager <sup>4</sup>	15	1–4	1	1	No	No
13	2003	Gordon and colleagues <sup>7</sup>	18	0–3	1	1	No	No
14	2004	Innes and colleagues <sup>6</sup>	n/a	varying	4	2	Yes	No
15	2005	Lineberger and colleagues <sup>43</sup>	n/a	1–4	3	1	No	No
16	2005	Lazar and colleagues <sup>8</sup>	32	0–3	2	1	No	No
17	2007	Hurley and colleagues <sup>9</sup>	30	0–3	2	1	No	No
18	2007	De Bruin and colleagues <sup>35</sup>	8	0–4	3	1	No	No
19	2010	Au and colleagues <sup>2</sup>	21	0–3	1	1	No	No
20	2010	Morgan and colleagues <sup>44</sup>	11	0–3	1	1	No	No
21	2011	Imholt	n/a	0–3	2	1	No	No
22	2012	Böddeker and colleagues	n/a	0–4	1	1	No	No

Abbreviations: AP, assessment points; n/a, not available.

mentioned above, however, there are also wide variations in study design, further complicating the comparability of the reported results. Information on number and qualification of involved observers, time intervals between repeat readings, and detailed descriptions of the assessed points are often different or even missing.<sup>5</sup>

Results from the OA scoring system devised by Vasseur and Berry<sup>3</sup> in 1992 are the most frequently reported among those studies. This stifle OA scoring system involves a total of 19 assessment points, with grade ranging from 0 (no changes) to 3 (severe changes). Currently, four modifications of the Vasseur and Berry scoring system<sup>3</sup> are in routine use, including the 18-point system by Gordon and colleagues,<sup>7</sup> the 21-point system by Au and colleagues,<sup>2</sup> the 32-point system by Lazar and colleagues,<sup>8</sup> and the 30-point system by Hurley and colleagues,<sup>9</sup> all of which grade severity from 0 to 3. Moreover, to the best of our knowledge, only one study<sup>6</sup> has evaluated the intra- and inter-observer variability of an OA scoring system to date, and the intention of those authors was to evaluate the radiographs with different parameters in the entire joint and not to use specific assessment points.

This study was designed to evaluate the intra- and inter-observer measurement variability of an OA scoring system with well-defined assessment points. The scoring system used in our study was developed by Mager<sup>4</sup> and Matis and colleagues<sup>10</sup> and consists of 15 assessment points. The same scoring system was used previously by Brahm<sup>11</sup> in a radiographic evaluation of the progression of OA. The scoring was conducted by three observers working independently to assess two radiographic examinations for single canine patients made at intervals of 2 weeks. The severity of arthritic change was graded according to the system by Schnell<sup>12</sup> (1: no radiological signs of arthritic changes; 2: mild radiological changes; 3: marked arthritic changes with mild exostosis and projection; and 4: marked exostoses and projections, and robust structural density of bone), which is currently the most commonly used method reported in the literature.<sup>3,13–15</sup>

By using an existing stifle OA scoring method, we aimed to provide a comparable monitoring method for stifle OA for future use in clinical studies as well as in daily practice. Our hypothesis was that scoring systems of Mager<sup>4</sup> and Matis and

colleagues<sup>10</sup> would have low intra- and inter-observer measurement variability and that the least consistency would be obtained by the least experienced observer.

## Materials and Methods

### Patients

Patient records of adult dogs that were presented to the Department of Small Animal Surgery at the University of Veterinary Medicine Vienna, Austria between 1 January and 15 June, 2015, were searched to identify all dogs with paired caudocranial and mediolateral stifle radiographs. Fifteen paired radiographs were chosen by a radiologist who did not participate in the study.

The inclusion criteria were set to include only adult or skeletally mature dogs with a minimum body weight of 15 kg, radiographs without evidence of fractures, bone tumours or patellar luxation, and with no previous surgical intervention. Skeletally immature dogs with open physes that could interfere with the radiographic assessment points used to evaluate the OA scoring system were excluded, as were radiographs with poor positioning (as evidenced by a double femoral condyle sign, tibial and femoral rotation on both views, or no intersection of the sesamoid bones on craniocaudal view).

### Radiographic Examination

All radiographs were obtained with the patient under heavy sedation or general anaesthesia, and by means of an AXICOM Iconos R200 X-ray unit (exposure of 73 kV and 3.6 mA). (Siemens Medical, Erlangen, Germany). To retrieve the standard mediolateral view, each dog was positioned in lateral recumbency with the stifle joint at 90 degrees flexion and the femoral condyles superimposed. To obtain the standard caudocranial view, each dog was positioned in sternal recumbency with the affected limb extended along the long axis of the femur and parallel to the long axis of the tibia. For all views, the pelvis was rotated slightly towards the affected limb.

### Radiographic Scoring

To determine the number of radiographs necessary for sufficient statistical analysis, a preliminary study was conducted with six radiographs under assessment by two observers working independently. A total of 15 radiographs was found to be sufficient for proper statistical analysis (two-tailed test, 80% power), with a range of variation of  $\pm 4$  points.

For the study, the radiographs were evaluated twice at a 2-week interval by three independent observers with different levels of training and experience, including a board-certified radiologist Diplomate European College of Veterinary Diagnostic Imaging (Dip ECVDI), a board-certified veterinary surgeon Diplomate European College of Veterinary Surgery (Dip ECVS) and a second-year surgical resident (second-year ECVS resident). The observers were blinded to age, weight, or breed of the patients for the radiographs they were assessing.

For each of the two assessments, a study controller presented the radiographs in a random order to each ob-

server, to help ensure that the observer was unaware on the second occasion of their initial score. Prior to initiation of the study, the observers discussed and agreed upon criteria for the radiographic evaluation of the 15 paired stifle radiographs and received a detailed drawing of all assessment points along with an accurate description of the 4-grade scale.

Following the system developed by Mager<sup>4</sup> and Matis and colleagues,<sup>10</sup> 15 variables were evaluated independently for each stifle (11 in the mediolateral view [–Fig. 1], 4 in the caudocranial view [–Fig. 2] and graded on an integer numeric scale from 1 to 4 [–Table 2]). The overall OA score was calculated as the sum of the scores for each of the 15 factors (range: 15–60).

### Statistical Analysis

The intra-observer variability estimate based on 95% confidence interval was defined as the minimal difference between two consecutive measurements made by the same observer. Intra-observer variability was determined by the method of residuals described by Bland and Altman<sup>16</sup> and Caylor and colleagues,<sup>17</sup> with the differences between measurements plotted against the mean. In addition, the mean differences in measurements between the first and second observations for each observer were determined. The inter-observer variability estimate based on 95% confidence interval was defined as the minimal difference between two consecutive measurements made by two different



**Fig. 1** Assessment points in the mediolateral view. 1, patellar apex; 2, patellar base; 3, proximal trochlear ridge; 4, distal trochlear ridge; 5, femoral condyle; 6, tibial tuberosity; 7, cranial aspect tibial plateau; 8, caudal aspect tibial plateau; 9, central aspect tibial plateau; 10, popliteal surface femur; 11, sesamoid bones.



**Fig. 2** Assessment points in the caudocranial view. 12, lateral tibial and femoral condyle; 13, medial tibial and femoral condyle; 14, intercondylar notch; 15, patella.

**Table 2** Grading system and corresponding radiographic changes

Grade	Severity	Changes
1	No	Radiographically normal/no evidence of sclerosis or osteophytes
2	Mild	Mild osteophytes and/or mild sclerosis (mild arthrosis)
3	Moderate	Moderate osteophytes and moderate sclerosis (moderate arthrosis)
4	Severe	Marked osteophytes and severe sclerosis (severe arthrosis)

observers.<sup>17</sup> The acceptable difference between two measurements was set within the range of  $\pm 3$  points according to the results of our pilot study. Inter-observer variability was determined for every single assessment point using

Pearson's correlation coefficient, with the mean of the calculated difference between both measurements compared using *t*-test. The requirement for standard distribution was determined by the Kolmogorov-Smirnov test. *p*-values  $\leq 0.05$  were considered to be significant. All statistical calculations were performed by the SPSS statistical software (v19; IBM, New York, USA).

## Results

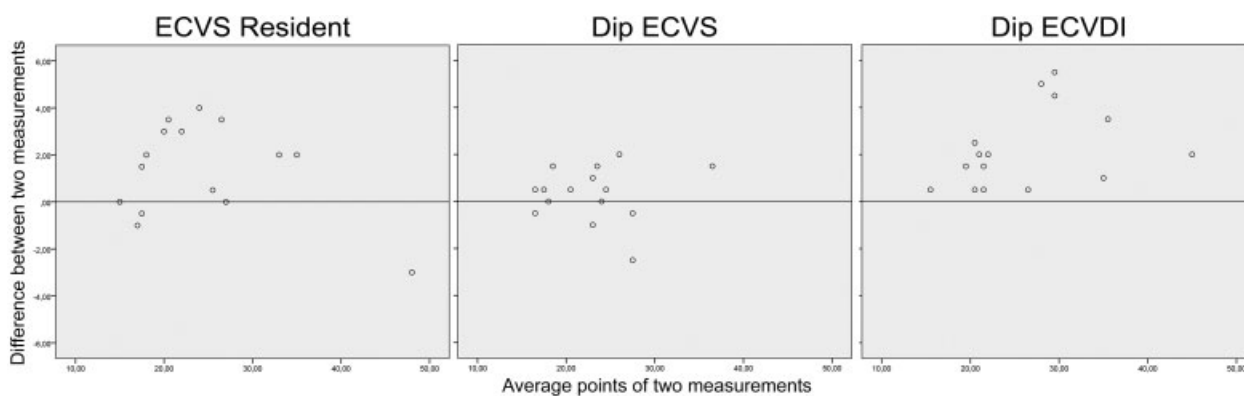
The features of the dogs for which radiographs were included were as follows: mean age of  $5.6 \pm 2.8$  years, mean weight of  $32.3 \pm 14.1$  kg, and breeds of mixed breed ( $n = 7$ ), Labrador Retriever ( $n = 3$ ), Golden Retriever ( $n = 2$ ), American Staffordshire Terrier ( $n = 1$ ), Rottweiler ( $n = 1$ ) and German Shepherd ( $n = 1$ ).

### Intra-observer Measurement

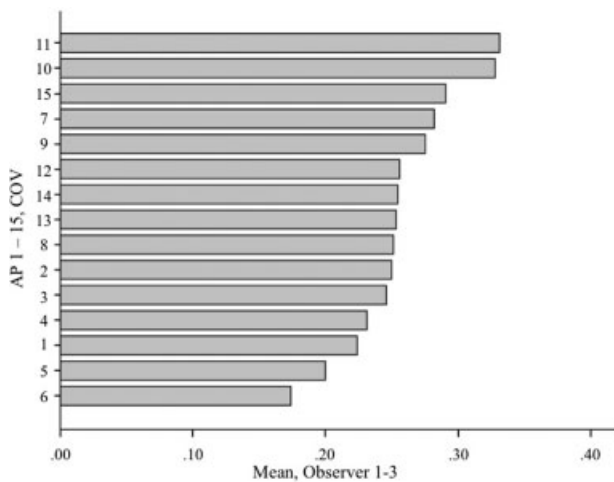
The variabilities between the two measurements, indicated by the mean differences between the first and second measurements for each observer, were  $-1.4 \pm 1.9$  (mean  $\pm$  standard deviation) points for the second-year ECVS resident,  $-0.3 \pm 1.2$  points for the Dip ECVS and  $-2.2 \pm 1.7$  points for the Dip ECVDI (**Fig. 3**). The second-year ECVS resident had the least concordance and therefore the lowest consistency in evaluation of the same radiograph (indicated by standard deviation of the difference). In general, the Dip ECVDI assigned the highest OA scores ( $26.1 \pm 7.8$  points), with the scores from the first evaluation being significantly higher (28.3 points) than those from the second evaluation (23.9 points), but with the difference reaching statistical significance. The second-year ECVS resident assigned the second highest overall OA scores ( $24.4 \pm 8.8$  points) and the Dip ECVS assigned the lowest ( $22.8 \pm 5.4$  points). In addition, there were significant differences between the scores of the first and second measurements for both the second-year ECVS resident ( $p = 0.014$ ) and the Dip ECVDI ( $p = 0.002$ ).

### Inter-observer Measurement

The correlation was high for all observers, with  $r = 0.845$  for the Dip ECVDI:second-year ECVS resident comparison;  $r = 0.868$  for Dip ECVS:second-year ECVS resident; and



**Fig. 3** Bland-Altman plots for intra-observer measurement variability of the second-year ECVS resident: MW (Marlis Wessely), Dip ECVS: ESF (Eva Schnabl-Feichter) and Dip ECVDI: AB (Andreas Brühshwein).



**Fig. 4** Coefficient of variation (CV) for all 15 assessment points (AP) at the 2-week intervals made by all three observers

$r = 0.837$  for Dip ECVDI:Dip ECVS; however, all correlations met the threshold for statistical significance ( $p < 0.001$ ). The mean difference between observers was  $1.09 \pm 4.99$  points with a 95% confidence interval for inter-observer variability, ranging from  $-0.37$  to  $2.55$  points.

The most discordant ratings among all observers were assigned for assessment point 11 (involving the sesamoid bones of the gastrocnemius muscle), closely followed by point 10 (involving the popliteal surface of the femur; ►Fig. 4).

## Discussion

The present study demonstrates that the radiographic stifle OA scoring system based on the system reported by Mager<sup>4</sup> and Matis and colleagues<sup>10</sup> is a feasible method to obtain and compare OA scores in the stifle with low intra- and inter-observer measurement variabilities. In our study, the lowest consistencies (indicated by standard deviation of the difference) in scoring were those assigned by the second-year ECVS resident, thereby confirming our hypothesis.

Several methods are currently available, and in routine use, for monitoring of OA; these include clinical assessment, joint fluid analysis, radiography, scintigraphy, arthroscopy, computed tomography and magnetic resonance imaging (MRI).<sup>5,6,18</sup> In daily clinical practice, conventional radiography is still the most commonly used imaging technique.<sup>19</sup> Radiographic changes that are observed in stifle OA include alterations in the periarticular soft tissues, joint effusion, osteophytosis, enthesopathy, intra-articular mineralization, subchondral sclerosis, subchondral cyst formation and narrowing of joint spaces.<sup>20,21</sup>

In human radiographic studies of OA, joint space narrowing has the greatest intra-observer reproducibility.<sup>22</sup> Obtaining accurate radiographic measurements of joint space by conventional radiography and MRI remains a challenge because the techniques involve only one line of contact, not allowing for assessment of differences in articular carti-

lage thickness over the entire contact region during active motion.<sup>23</sup> Weight-bearing radiography reportedly has improved accuracy of joint space measurements,<sup>6,22,24</sup> but standing radiographs are not routinely performed in small animal veterinary medicine. Edamura and colleagues<sup>25</sup> evaluated the effectiveness of weight-bearing standing radiography in dogs with complete cranial cruciate ligament rupture, and found that this technique could be easily performed in dogs that are able to bear their own weight on the affected limb. In another study, Anderst and colleagues<sup>23</sup> used a unique high-speed stereo radiographic system, combined with subject-specific computed tomography reconstructions, to obtain dynamic serial measurements of the joint space in vivo over a 2-year period; they concluded that this technique is a precise tool for measuring joint space serially in vivo under dynamic loading conditions. Controversy remains, however, as to whether standing radiographs are easy enough to carry out in daily clinical practice and sufficiently reproducible without sedation, especially for animals that are difficult to handle (temperamentally or physically).

Osteoarthritis scores are an important and common tool used for the assessment of radiographic OA progression, especially in cases following surgical treatment of cranial cruciate ligament injuries.<sup>2,4,6,7,10,11,14,20</sup> Since the first report of a radiographic grading system used in an epidemiological OA study,<sup>1</sup> a total of 22 different scores have been published to assess individual radiographic features and overall severity of stifle joint OA (►Table 1). These various OA scoring systems, however, have produced widely variable scores, complicating the development of a standardized evaluation method of OA in veterinary practice and research.

The OA scoring system developed by Vasseur and Berry<sup>3</sup> is currently the most cited system in the literature, but several groups<sup>2,7-9</sup> have applied this system with modifications that effectively preclude direct comparisons between their results. We did not use Vasseur and Berry's scoring system in this study for the following reasons. In their original study, Vasseur and Berry<sup>3</sup> included assessments of changes to the surrounding soft tissues; however, soft tissues have low contrast and are not always visible with conventional X-rays.<sup>26,27</sup> In addition, it is not yet empirically established as to whether (or to what extent) an examiner can manipulate digital radiographs to make the soft-tissue structures more or less visible and how that would affect the overall assessment of the other structures. The novel X-ray imaging technology of diffraction-enhanced imaging (DEI) can overcome this challenge, since it is capable of rendering images with absorption, refraction and scatter rejection qualities, allowing for detection of specific soft tissues based on small differences in tissue densities.<sup>26</sup> Unfortunately, this technique is currently not widely used in veterinary medicine, likely due to its relatively recent introduction. In our consideration of the scoring system by Vasseur and Berry,<sup>3</sup> we were also concerned by the fact that their original study assessed canine stifle joints that were divided into four compartments (the femoropatellar and the medial, central [intercondyloid] and lateral femorotibialis), with each

compartment assigned an individual secondary osteoarthritis score. Differences among the four compartments were not reported;<sup>3</sup> therefore, we decided not to divide the stifle joint into compartments.

An alternative system for evaluating radiographs was proposed by Innes and colleagues.<sup>6</sup> The parameters used in their study were a “global score for overall disease severity” (0–3), “joint effusion” (0–2), “osteophytosis” (0–3), “intra-articular mineralization” (0–2) and “tibial subchondral sclerosis” (0–1); moreover, the scales were interpreted subjectively and independently by four veterinary radiologists on a discontinuous ordinal scale. Innes and colleagues<sup>6</sup> reported excellent intra-observer agreements for global score, effusion and osteophytosis, and good inter-observer agreement for global score, effusion, osteophytosis and intra-articular mineralization. We considered using this system in our study, but decided against it (and in favour of Mager’s<sup>4</sup> and Matis and colleagues’<sup>10</sup> scoring system) because it is more complicated and requires a higher level of radiology knowledge, which may preclude its ease of adoption in daily clinical practice.

Mager<sup>4</sup> and Matis and colleagues<sup>10</sup> focused their score mainly on the presence and severity of osteophytes. Osteophyte formation accompanies joint space narrowing, subchondral sclerosis and subchondral cyst formation, the latter of which is one of the main radiographic features of OA and an important criterion for this disease.<sup>20,21</sup> It is not possible to manipulate the presence and severity of osteophytes with digital radiographic systems, and this fact suggests the potential advantage for use of a radiographic scoring system in evaluating mainly osteophytes. While the pathophysiology of osteophyte formation is not completely understood,<sup>28</sup> it is known that they occur early in the OA disease process, being consistent features of a destabilized stifle and representing a feature of disease progression that can be scored.<sup>29</sup>

In assessment of OA progression, osteophytes can be observed early and prior to development of joint space narrowing.<sup>30</sup> In human medical practice, osteophytes in stifle joints are important clinical features of OA used for measuring disease progression and predicting disease manifestation.<sup>31,32</sup> Studies of human OA cases have confirmed that osteophytes correlate well with the occurrence of pain.<sup>33</sup> Furthermore, studies in both humans and animals have demonstrated that radiographic assessment of osteophytes produces good intra-observer agreement.<sup>4,6,18,20,22,33–35</sup>

In the current study, the ability to reproduce measurements was high for all three observers, indicating an ability to consistently recognize subjectively determined assessment points on the radiographs, even by observers with limited previous experience. The high degree of reproducibility for the intra-observer measurements also suggests that all observers were able to consistently reproduce measurements for any single given radiograph. We specifically decided to not include only radiologists as observers, as one of our purposes was to assess consistency among surgeons and radiologists who represent the actual viewers of such radiographs in day-to-day clinical and research settings.

Although the intra-observer measurement variability was low in our study, the lowest consistency (indicated by the standard deviation of the differences) between the first and second measurement was seen for the second-year ECVS resident. These data could indicate that lower levels of training and experience can result in an inconsistent application of the measurement method. Similar observations have been made in different studies evaluating the intra- and inter-observer measurement variability of the tibia plateau angle (TPA) assessment in dogs and cats.<sup>36–38</sup> In those studies, significant differences were found among the more experienced and less experienced observers making the TPA measurements.

In our study, the variability of the overall inter-observer score was low ( $1.09 \pm 4.99$  points). The correlation was high for all observers, and all of them met the threshold for statistical significance.

The high coefficient of variation found particularly for assessment points 10 and 11 suggests that these points are more difficult to score consistently between and within observers. Variations in evaluating assessment points 10 and 11 may originate from a variety of sources, such as animal patient positioning for the limb radiograph, as well as observer error with this measurement. Sources of non-observer variability may include subtleties in individual anatomical variation. Moreover, subtle flexion, rotation, or sagittal variations in positioning of the stifle may alter the radiographic appearance of the assessment points required for evaluation of the OA score. Therefore, accurate medio-lateral radiographic projections of the stifle joint with superimposition of the femoral condyles are necessary.

No significant inter-observer differences were determined in the remaining 13 points of our study. Future considerations could include a more precise definition of the 4-grade scale; for instance, a scale with millimetre subdivisions has been implemented already by the International Elbow Working Group (IEWG) for the radiographic evaluation of elbow OA. A more precise grading according to the severity of osteophytes (e.g., grades being 1: no changes; 2: small osteophytes; 3: medium osteophytes; 4: large osteophytes) is definitely a future consideration for modification of this scoring system. An additional consideration is modification of Mager’s<sup>4</sup> and Matis and colleagues’<sup>10</sup> OA score to exclude points 10 and 11. A dropout of assessment scores for points 10 and 11 did not change the concordance of the results in the current study. Instead, merely the total score changed, but the variability of the results could not be reduced sufficiently. Multiple sources can underlie observed changes in point 11, such as chronic insertion tendinopathy of the gastrocnemius muscle, sesamoiditis, or secondary changes related to arthritis. Therefore, point 11 is not an exclusive marker for OA. For a modification of this scoring system, one could discuss whether point 11 should be excluded completely or if an evaluation in both views (mediolateral and craniocaudal) would facilitate the assessment and result, at least, in more concordant results.

This study that we performed was conducted in German, using the original grading system. One major limitation of the current study, however, is that the scoring system was not

translated, adapted and validated for cross-cultural research as described by Sousa and Rojjanasiriat.<sup>39</sup> To help mitigate this limitation, the reporting of our findings in English involved language translation by a professional medical translator with radiology as a main area of interest. Nevertheless, interpretation of our results needs to be done carefully.

In conclusion, we recommend the OA scoring system of Mager<sup>4</sup> and Matis and colleagues<sup>10</sup> for two reasons. First, it has been demonstrated in this study as an easily applicable and feasible system to be used by observers with differing levels of experience; this finding is supported by the findings of low inter- and intra-observer measurement variabilities. Second, the scoring system is not dependent on soft-tissue opacity and therefore cannot be manipulated easily with digital processing systems. Thus, the OA scoring system developed by Mager<sup>4</sup> and Matis and colleagues<sup>10</sup> appears to be a repeatable tool applicable for radiographic scoring of the stifle joint.

#### Note

Supplementary material for this article is available online at <https://doi.org/10.3415/VCOT-16-07-0134>.

#### Author Contributions

All authors were involved in the study conception and design, acquisition of data, data analysis and interpretation, and in the drafting or revising of the manuscript. All authors also approved of the submitted version.

#### Funding

The study was not supported by any dedicated funds from any agency in the public, commercial or not-for-profit sectors.

#### Conflict of Interest

The authors do not have any potential conflicts of interest to declare.

#### Acknowledgments

The authors thank Dr. Alexander Tichy for his assistance in the statistical analysis and Prof. Matis for her support.

#### References

- Shires PK, Hulse DA, Liu W. The under-and-over fascial replacement technique for anterior cruciate ligament rupture in dogs: a retrospective study. *JAAHA* 1984;20:69–77
- Au KK, Gordon-Evans WJ, Dunning D, et al. Comparison of short- and long-term function and radiographic osteoarthritis in dogs after postoperative physical rehabilitation and tibial plateau leveling osteotomy or lateral fabellar suture stabilization. *Vet Surg* 2010;39(02):173–180
- Vasseur PB, Berry CR. Progression of stifle osteoarthritis following reconstruction of the cranial cruciate ligament in 21 dogs. *J Am Anim Hosp Assoc* 1992;28:129–136
- Mager FW. Zur Kniegelenksarthrose des Hundes nach vorderer Kreuzbandruptur – ein retrospektiver Vergleich dreier Operationsmethoden [dissertation]. Munich: Ludwig-Maximilians University; 2000
- Günther KP, Sun Y. Reliability of radiographic assessment in hip and knee osteoarthritis. *Osteoarthritis Cartilage* 1999;7(02):239–246
- Innes JF, Costello M, Barr FJ, et al. Radiographic progression of osteoarthritis of the canine stifle joint: a prospective study. *Vet Radiol Ultrasound* 2004;45(02):143–148
- Gordon WJ, Conzemius MG, Riedesel E, et al. The relationship between limb function and radiographic osteoarthritis in dogs with stifle osteoarthritis. *Vet Surg* 2003;32(05):451–454
- Lazar TP, Berry CR, deHaan JJ, et al. Long-term radiographic comparison of tibial plateau leveling osteotomy versus extracapsular stabilization for cranial cruciate ligament rupture in the dog. *Vet Surg* 2005;34(02):133–141
- Hurley CR, Hammer DL, Shott S. Progression of radiographic evidence of osteoarthritis following tibial plateau leveling osteotomy in dogs with cranial cruciate ligament rupture: 295 cases (2001–2005). *J Am Vet Med Assoc* 2007;230(11):1674–1679
- Matis U, Brahm-Jorda T, Jorda C, et al. Radiographic evaluation of the progression of osteoarthritis after tibial plateau leveling osteotomy in 93 dogs. *Vet Comp Orthop Traumatol* 2005;18:A32
- Brahm T. Umstellungsosteotomie des Tibiaplateaus nach B. Slocum zur Behandlung der vorderen Kreuzbandruptur beim Hund – retrospektive Erhebungen bei 115 Fällen [dissertation]. Munich: Ludwig-Maximilians University; 2004
- Schnell EM. Drei Jahre Erfahrung mit einer modifizierten Kreuzbandplastik beim Hund [dissertation]. Munich: Ludwig-Maximilians University; 1986
- Chauvet AE, Johnson AL, Pijanowski GJ, et al. Evaluation of fibular head transposition, lateral fabellar suture, and conservative treatment of cranial cruciate ligament rupture in large dogs: a retrospective study. *J Am Anim Hosp Assoc* 1996;32(03):247–255
- Mullen HS, Matthiesen DT. Complications of transposition of the fibular head for stabilization of the cranial cruciate-deficient stifle in dogs: 80 cases (1982–1986). *J Am Vet Med Assoc* 1989;195(09):1267–1271
- Brunnberg L, Rieger I, Hesse EM. Sieben Jahre Erfahrung mit einer modifizierten „Over-the-Top“-Kreuzbandplastik beim Hund. *Kleintierpraxis* 1992;37:735–746
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307–310
- Caylor KB, Zumpano CA, Evans LM, et al. Intra- and interobserver measurement variability of tibial plateau slope from lateral radiographs in dogs. *J Am Anim Hosp Assoc* 2001;37(03):263–268
- Rayward RM, Thomson DG, Davies JV, et al. Progression of osteoarthritis following TPLO surgery: a prospective radiographic study of 40 dogs. *J Small Anim Pract* 2004;45(02):92–97
- Roemer FW, Eckstein F, Hayashi D, et al. The role of imaging in osteoarthritis. *Best Pract Res Clin Rheumatol* 2014;28(01):31–60
- Morgan JP. Radiological pathology and diagnosis of degenerative joint disease in the stifle joint of the dog. *J Small Anim Pract* 1969;10(09):541–544
- Bennett D, Tennant B, Lewis DG, et al. A reappraisal of anterior cruciate ligament disease in the dog. *J Small Anim Pract* 1988;29:275–297
- Cooper C, Cushnaghan J, Kirwan JR, et al. Radiographic assessment of the knee joint in osteoarthritis. *Ann Rheum Dis* 1992;51(01):80–82
- Anderst WJ, Les C, Tashman S. In vivo serial joint space measurements during dynamic loading in a canine model of osteoarthritis. *Osteoarthritis Cartilage* 2005;13(09):808–816
- Ravaud P, Ayral X, Dougados M. Radiologic progression of hip and knee osteoarthritis. *Osteoarthritis Cartilage* 1999;7(02):222–229
- Edamura K, Ueno F, Tanegashima K, et al. Practical use of weight-bearing standing radiography in dogs with rupture of the cranial cruciate ligament. Poster session presented at: 25th Annual conference of the European college of veterinary surgeons; 2016 Jul 7–9; Lisbon, Portugal
- Li J, Zhong Z, Lidtke R, et al. Radiography of soft tissue of the foot and ankle with diffraction enhanced imaging. *J Anat* 2003;202(05):463–470

- 27 Muehleman C, Zhong Z, Williams JM, et al. Diffraction enhanced X-ray imaging of the articular cartilage of experimental animals. Proceedings of the 48th Annual Meeting of the Orthopaedic Research Society; February 10–13, 2002; Dallas, TX, p.365
- 28 Junker S, Krumbholz G, Frommer KW, et al. Differentiation of osteophyte types in osteoarthritis: proposal of a histological classification. *Joint Bone Spine* 2016;83(01):63–67
- 29 Innes JF, Barr ARS. Clinical natural history of the postsurgical cruciate deficient canine stifle joint: year 1. *J Small Anim Pract* 1998;39(07):325–332
- 30 van der Kraan PM, van den Berg WB. Osteophytes: relevance and biology. *Osteoarthritis Cartilage* 2007;15(03):237–244
- 31 Croft P, Cooper C, Wickham C, et al. Defining osteoarthritis of the hip for epidemiologic studies. *Am J Epidemiol* 1990;132(03):514–522
- 32 Scott JC, Nevitt MC, Lane E, et al. Association of individual radiographic features of hip osteoarthritis with pain. *Arthritis Rheum* 1992;35:81
- 33 Felson DT, McAlindon TE, Anderson JJ, et al. Defining radiographic osteoarthritis for the whole knee. *Osteoarthritis Cartilage* 1997;5(04):241–250
- 34 Kellgren JH, Lawrence JS. Atlas of Standard Radiographs: The Epidemiology of Chronic Rheumatism. Vol. 2. Oxford, UK: Blackwell Scientific Publications; 1963
- 35 de Bruin T, de Rooster H, Bosmans T, et al. Radiographic assessment of the progression of osteoarthrosis in the contralateral stifle joint of dogs with a ruptured cranial cruciate ligament. *Vet Rec* 2007;161(22):745–750
- 36 Schnabl E, Reese S, Lorinson K, et al. Measurement of the tibial plateau angle in cats with and without cranial cruciate ligament rupture. *Vet Comp Orthop Traumatol* 2009;22(02):83–86
- 37 Unis MD, Johnson AL, Griffon DJ, et al. Evaluation of intra- and interobserver variability and repeatability of tibial plateau angle measurements with digital radiography using a novel digital radiographic program. *Vet Surg* 2010;39(02):187–194
- 38 Fettig AA, Rand WM, Sato AF, et al. Observer variability of tibial plateau slope measurement in 40 dogs with cranial cruciate ligament-deficient stifle joints. *Vet Surg* 2003;32(05):471–478
- 39 Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract* 2011;17(02):268–274
- 40 Mayer W. Traumatische Sehnen- und Bandverletzungen am Kniegelenk des Hundes und der Katze: Behandlung und Ergebnisse in den Jahren 1983–1991 (dissertation). Munich: Ludwig-Maximilians University; 1993
- 41 Coetzee GL, Lubbe AM. A prospective study comparing two fascial reconstruction techniques to stabilise the cranial cruciate deficient stifle in the dog. *Vet Comp Orthop Traumatol* 1995; 8:82–90
- 42 De Rooster H, Van Bree H. Use of compression stress radiography for the detection of partial tears of the canine cranial cruciate ligament. *J Small Anim Pract* 1999;40:573–576
- 43 Lineberger JA, Allen DA, Wilson ER, et al. Comparison of radiographic arthritic changes associated with two variations of tibial plateau leveling osteotomy. *Vet Comp Orthop Traumatol* 2005; 18:13–17
- 44 Morgan JP, Voss K, Damur DM, et al. Correlation of radiographic changes after tibial tuberosity advancement in dogs with cranial cruciate-deficient stifles with functional outcome. *Vet Surg* 2010; 39:425–432
- 45 Böddeker J, Drüen S, Meyer-Lindenberg A, et al. Computer-assisted gait analysis of the dog: Comparison of two surgical techniques for the ruptured cranial cruciate ligament. *Vet Comp Orthop Traumatol* 2012;25:11–21