Title

Results of using multiplanar reconstructed CT images for assessing elbow joint osteoarthritis in dogs are consistent with results of radiographic assessment.

Journal

American Journal of Veterinary Research, 83(10)

ISSN

0002-9645

Authors

Shubert, Madison P
Filliquist, Barbro
Chou, Po-Yen
et al.

Publication Date

2022-08-01

DOI

10.2460/ajvr.22.04.0066

Peer reviewed

AJVR

AVMA

Results of using multiplanar reconstructed CT images for assessing elbow joint osteoarthritis in dogs are consistent with results of radiographic assessment

Madison P. Shubert, DVM[1]; Barbro Filliquist, DVM, MAS, DACVS-SA[2]; Po-Yen Chou, BVM, MVM, MS, DACVS-SA[2]; Amy S. Kapatkin, DVM, MAS, DACVS[2]; Mathieu Spriet, DVM, MS, DACVR[2]; Sun Young Kim, DVM, MS, DACVS[3]; Tanya C. Garcia, MS[2]; Denis J. Marcellin-Little, DEDV, DACVS, DACVSMR[2]*

[1]Veterinary Medical Teaching Hospital, School of Veterinary Medicine, University of California-Davis, Davis CA
[2]Department of Veterinary Surgical and Radiological Sciences and JD Wheat Veterinary Orthopedic Laboratory, School of Veterinary Medicine, University of California-Davis, Davis CA
[3]Department of Veterinary Clinical Sciences, College of Veterinary Medicine, Purdue University, West Lafayette, IN

*Corresponding author: Dr. Marcellin-Little (djmarcel@ucdavis.edu)

https://doi.org/10.2460/ajvr.22.04.0066

OBJECTIVE
To compare osteoarthritis scores assigned through radiographic evaluation of 18 anatomic regions in the elbow joint with scores assigned through evaluation of 3-D maximum intensity projection (MIP), 3-D surface rendering (TSR), and multiplanar reconstructed (MPR) CT images, and to evaluate intraobserver and interobserver agreement of radiographic and CT scoring.

SAMPLE
Radiographic and CT images of 39 elbow joints in 20 dogs.

PROCEDURES
Images were anonymized and graded independently by 5 observers. One observer graded 12 elbow joints 3 times. Intraobserver consistency and repeatability, interobserver agreement, consistency among methods, and bias between methods were calculated.

RESULTS
The most severe changes were observed at the proximal aspect of the anconeal process, and the medial and cranial aspects of the medial coronoid process. Intraobserver consistency was moderate or better for 11/16 regions with MIP images, 11/16 regions with TSR images, 17/18 regions with MPR images, and 14/18 regions with radiographic images. Interobserver agreement was moderate or better for 5/16 regions with MIP images, 9/16 regions with TSR images, 12/18 regions with MPR images, and 6/18 regions with radiographic images. Mean scores from CT-based methods were higher than mean radiographic scores.

CLINICAL RELEVANCE
Assessments of osteoarthritis severity in the elbow joints of dogs obtained by examining radiographic images were generally consistent with assessments obtained by examining CT scans. MPR scores were more consistent and more comparable to radiographic scores than were MIP or TSR scores.

Osteoarthritis (OA) of the elbow joint is common in dogs and most often results from elbow dysplasia. Detecting OA in the elbow joint is important, because the presence and severity of OA are the basis for breeding recommendations and are prognostic factors for cartilage damage, future limb function, and future quality of life.[1–5] Also, assessing the severity of elbow OA over time facilitates evaluation of treatments aimed at slowing the progression of OA.[6,7] Elbow OA in dogs has traditionally been assessed by means of radiography, with several available grading schemes, including the International Elbow Working Group (IEWG) grading scheme, based on the size of the largest osteophyte.[1,8–12] CT is now commonly used for assessing elbow OA,[10,13,14] but comparisons of elbow OA severity determined by means of radiography versus CT are limited.[12,15,16]

No scheme to grade elbow OA on the basis of CT images has been reported. Because several types of reconstructed CT images can be used to evaluate OA severity, including 3-D maximum intensity projection (MIP), 3-D surface rendering (TSR), and multiplanar reconstructed (MPR) images,[15–19] there are several options for schemes to grade elbow OA. However, comparisons of OA severity from MIP, TSR, and MPR images are lacking. A comparison of elbow

OA grading from radiographs and from various reconstructed CT images is warranted. The evaluation should compare radiographs and CT images acquired contemporaneously, so that OA is identical on radiographs and CT images.

The aims of our study were to score OA severity at landmarks around the elbow joint on radiographic and CT images of dogs, to evaluate intraobserver and interobserver agreement in radiographic and CT scoring of elbow OA, and to compare scores from radiographs and 3 types of reconstructed CT images (MIP, TSR, and MPR) with each other and with grades assigned with the IEWG grading scheme. We hypothesized that scoring of elbow OA with radiographic and CT images would be repeatable within and among observers. We also hypothesized that IEWG scores and scores from MIP, TSR, and MPR images would be higher than scores from radiographic images.

Materials and Methods

Patient selection

The study involved a retrospective review of radiographic and CT images of dog elbow joints that had been acquired contemporaneously. Data collection was covered by institutional animal care and use protocols of the University of California-Davis Veterinary Medical Teaching Hospital and the Michigan State University Veterinary Medical Center. Owners of dogs included in the study had signed an informed consent form authorizing the analysis of anonymized patient data. A convenience sample was selected randomly from the hospital populations that included elbow joints with OA ranging from mild to severe. The study objective was to include a sample with ≥ 10 elbow joints with mild OA, ≥ 10 elbow joints with moderate OA, and ≥ 10 elbow joints with severe OA, as determined on the basis of IEWG radiographic scores. Dogs were selected through electronic searches of medical records of patients presented to the University of California-Davis and Michigan State University between June 1, 2010 and December 31, 2018 that underwent contemporaneous radiography and CT of the elbow joints (defined as undergoing radiography and CT within 15 days of each other without surgery between imaging studies). Dogs were eligible for inclusion if a CT scan of the elbow joint with slice thickness ≤ 0.7 mm and mediolateral and craniocaudal radiographic projections in DICOM image format with a magnification marker were available and no fracture or neoplasia was present. Patients were identified in hospital information systems by conducting independent searches for radiographs and CT scans with the search terms elbow osteoarthritis, fragmented medial coronoid, ununited anconeal process, subchondral sclerosis, and elbow dysplasia, and then cross-referencing results to identify patients with contemporaneous imaging studies. Similar numbers of joints with mild, moderate, or severe OA were included.

Eligible radiographs were screened by 2 of the authors (MPS and BF) with an open-source DICOM image viewer (Horos, version 3.3.6; The Horos Project) who assigned a score to each joint with the IEWG grading system. Potential scores ranged from 0 to 3 points as follows: 0 point, normal (no evidence of osteophytes or sclerosis in the joint); 1 point, mild (osteophytes ≤ 2 mm anywhere in the joint or sclerosis of the ulnar trochlear notch); 2 points, moderate (osteophytes > 2 but ≤ 5 mm anywhere in the joint); and 3 points severe (osteophytes ≥ 5 mm anywhere in the joint).[15,20] One of the screeners was a board-certified veterinary surgeon with 11 years of postresidency experience; the other was a veterinary student. The IEWG scores from the 2 screening observers were used solely for the purpose of identifying imaging studies to include in the study and were not used for other analyses.

Image preparation, anonymization, randomization, and evaluation

All imaging studies selected for inclusion were anonymized. For each elbow joint, 3-D MIP studies were created with the DICOM viewer with a window width of 2,600 HU and window level of 600 HU. Forty views of an MIP rendering of the elbow joint with 360° rotation about the long axis of the radius were exported as a video file with a resolution of 1,024 X 1,024 pixels. TSR studies were also created with the DICOM viewer with a low threshold value of 500 HU (bone) without transparency. Forty views of TSR of the elbow joint with 360° rotation about the long axis of the radius were exported as a video file with a resolution of 1,024 X 1,024 pixels. All video files were viewed with video viewing software (QuickTime Player, version 10.5; Apple Corp). MPR studies were created with the radiographic software from unmodified DICOM images, and a reconstruction created with the MPR tool of the visualization software was provided. Observers were instructed to start the evaluation with the sagittal and frontal planes intersecting along the long axis of the antebrachium. Each observer was free to adjust the orientation of the planes. Radiographic views included a craniocaudal and flexed mediolateral view of the elbow joint. Order was randomized within each image type by means of randomizing software (randomizer.org).

Five board-certified clinicians (4 American College of Veterinary Surgeons [ACVS] diplomates and 1 American College of Veterinary Radiology diplomate) who had not participated in the patient selection process reviewed the images independently in the following order: MIP, TSR, MPR, and radiographic images. These individuals had 6, 9, 14, 28, and 31 years of postresidency experience. A day or more was allowed to elapse between scoring for each imaging modality. Each observer was blinded to results from the other observers. One observer (an ACVS diplomate with 28 years of experience) performed repeated evaluations of 12 elbow joints selected randomly from the sample 2 additional times. The repeated evaluations followed the same evaluation order, with ≥ 2 weeks between repeated evaluations.

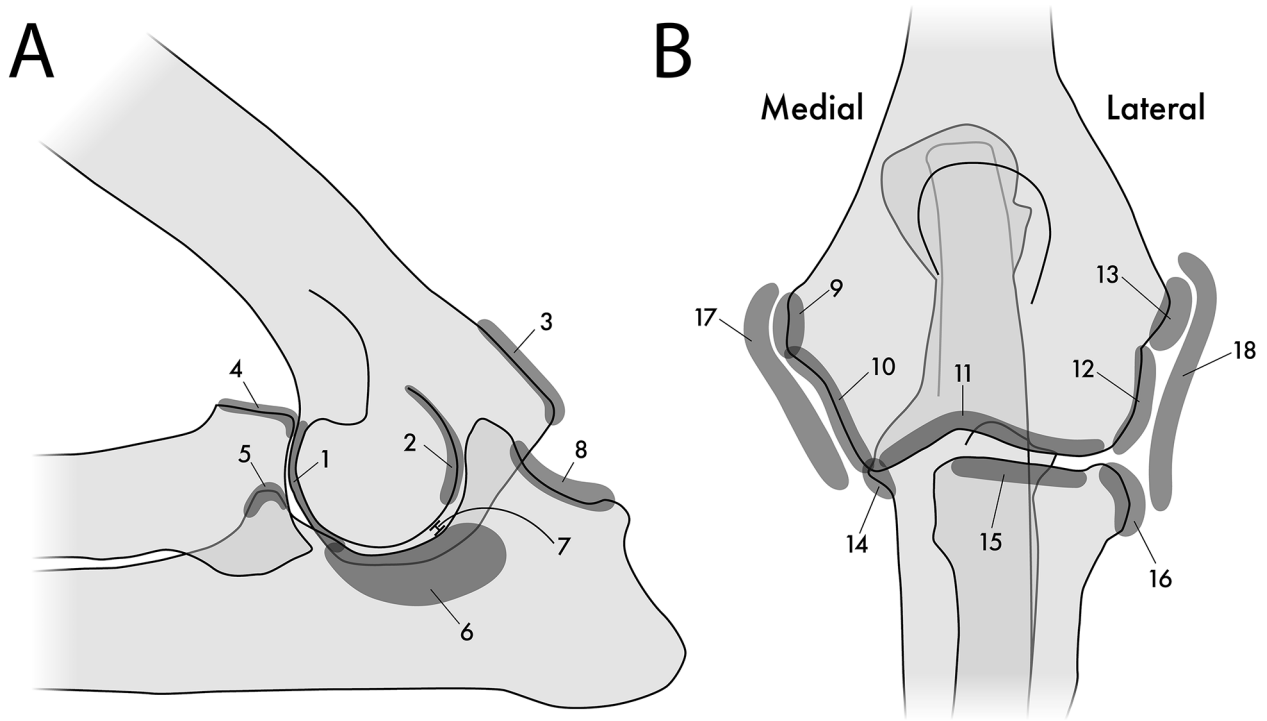

Figure 1—Illustrations of a medial or lateral view (A) and a cranial or caudal view (B) of the elbow joint in dogs showing 18 regions of interest when assessing severity of osteoarthritis. The regions were as follows: 1, the cranial aspect of the humeral condyle; 2, the lateral aspect of the humeral condyle; 3, the medial aspect of the humeral condyle; 4, the cranial aspect of the radial head; 5, the cranial aspect of the medial coronoid process; 6, the ulnar notch; 7, the humeroulnar articulation; 8, the proximal aspect of the anconeal process on the lateral or medial view; 9, the medial humeral epicondyle; 10, the medial aspect of the humeral condyle; 11, the distal aspect of the humeral condyle; 12, the lateral aspect of the humeral condyle; 13, the lateral epicondyle; 14, the medial aspect of the medial coronoid process; 15, the proximal aspect of the radial head; 16, the lateral aspect of the radial head; 17, the medial aspect of the joint capsule on the cranial or caudal view; and 18, the lateral aspect of the joint capsule on the cranial or caudal view.

Observers calibrated radiographs with a 100-mm-long magnification marker before reading. Observers were free to manipulate images as deemed necessary for optimal evaluation. Manipulations included stopping video animations and rotating and magnifying images. Each of 18 anatomic regions **(Figure 1)** was assigned a score of 0 point, normal; 1 point, mildly abnormal (changes < 2 mm); 2 points, moderately abnormal (changes ≥ 2 mm but ≤ 5 mm); or 3 points, severely abnormal (changes ≥ 5 mm).[13] The largest osteophyte observed in the joint was then used to assign an IEWG score, as described elsewhere.[15,20]

Statistical analysis

Studies of the left and right elbow joints of a dog were considered independent observations. Normality of distributions was evaluated with the Shapiro-Wilk test; standard statistical software (SAS, version 9.4; SAS Institute) was used. Intraobserver consistency was determined by calculating the intraclass correlation coefficient (ICCs) for triplicate measurements of each region and each scoring method and comparing them by means of ANOVA (Proc GLM, SAS Institute).[21] Intraobserver repeatability within a limb for each scoring method was calculated as the SD (ie, square root of the mean square error).[22,23] Interobserver agreement was determined by calculating the ICC of measurements of each anatomic region and each imaging method.[21] Consistency among scoring methods (radiographic, MIP, TSR, and MPR images) was determined by calculating the ICC for each anatomic region and observer.[20] MIP, TSR, MPR, and IEWG scores were compared with radiographic scores by calculating the mean score for all regions for each joint and then comparing mean scores to determine agreement between observers and consistency among scoring methods by calculating ICCs.[21] Mean radiographic, MIP, TSR, and MPR scores were compared pairwise with an ANOVA for ranked data. Spearman correlation coefficients were computed and compared statistically with the Fisher *z* transformation hypothesis test.[24] Bias and 95% limits of agreement between radiographic, MIP, TSR, and MPR scores were determined with Bland–Altman plots.[25] Differences between Bland–Altman plots and differences between the slopes of the regression lines were compared by means of ANCOVA. ICC values < 0.5 were considered to represent poor repeatability or consistency, values ≥ 0.5 and < 0.75 were considered to represent moderate repeatability or consistency, values ≥ 0.75 and < 0.9 were considered to represent good repeatability or consistency, and values ≥ 0.90 were considered to represent excellent

repeatability or consistency.[26] Severity maps of elbow joints with mild and severe OA were prepared. For all analyses, values of $P < .05$ were considered significant.

Results

Population

Twenty patients were included in the study: 12 neutered males, 4 sexually intact males, and 4 spayed females. Mean age (± SD) was 23 ± 17 months, and mean weight was 31 ± 13 kg. Images were available for both elbow joints for 19 dogs and for 1 elbow joint for 1 dog, for a total of 39 sets of images. The study included collection of 15,987 measurements. Interobserver calculations included 14,235 measurements (39 dogs, 18 regions, 4 scoring methods, and 5 observers in addition to 195 IEWG measurements), and intraobserver calculations included 1,752 additional measurements (12 dogs, 18 regions, 4 scoring methods, and 2 sets of measurements in addition to 24 IEWG measurements).

OA scores

Mean scores (± SD) for each anatomic region were calculated **(Table 1)**, and *P* values for comparisons between methods were summarized **(Supplementary Table S1)**. The anatomic regions with the highest radiographic OA scores were the proximal aspect of the anconeal process and the medial and cranial aspects of the medial coronoid process **(Figure 2)**. Means scores for all regions scored with all methods by all 5 observers ranged from 0.55 to 1.16.

Consistency and repeatability of OA score

Intraobserver consistency was moderate or better for 14 of the 18 regions when reading radiographs (median consistency, 0.75; median repeatability, 0.34), 11/16 regions when reading MIP images (median consistency, 0.62; median repeatability, 0.44), 11/16 regions when reading TSR images (median consistency, 0.71; median repeatability, 0.42), and 17/18 regions when reading MPR images (median consistency, 0.71; median repeatability, 0.44; **Supplementary Table S2**). Interobserver agreement was moderate or better for 6 of the 18 regions when reading radiographs, 5/16 regions when reading MIP images, 9/16 regions when reading TSR images, and 12/18 regions when reading MPR images **(Supplementary Table S3)**. Consistency was moderate or better for 2 of the 18 regions for observer 1, 8/18 regions for observer 2, 11/18 regions for observers 3 and 4, and 13/18 regions for observer 5 **(Supplementary Table S4)**. Observers 1 and 2 had 6 and 9 years of postresidency experience, respectively. Observers 3, 4, and 5 had 28, 31, and 14 years of experience, respectively.

Interobserver agreement (ICC [2,1]) for mean scores was 0.567 for MIP images, 0.679 for TSR

Table 1—Mean ± SD osteoarthritis (OA) severity scores assigned by 5 observers to 18 anatomic regions of the elbow joint for 39 dogs with elbow OA.

Anatomic region	Radiographs	MIP images	TSR images	MPR images
Flexed mediolateral view				
Humeral condyle, cranial aspect	0.54 ± 0.71 a	1.02 ± 0.94 b	0.96 ± 1.04 b	0.73 ± 0.88 c
Humeral condyle, lateral aspect	0.57 ± 0.81 a	0.88 ± 0.97 b	0.92 ± 1.00 b	0.72 ± 1.01 a
Humeral condyle, medial aspect	0.61 ± 0.87 a	0.77 ± 0.96 bc	0.64 ± 0.87 ac	0.82 ± 0.88 b
Radial head, cranial aspect	0.83 ± 0.97 a	0.79 ± 0.97 a	0.96 ± 1.00 b	0.80 ± 0.99 a
Medial coronoid process, cranial aspect	0.89 ± 0.84 a	0.95 ± 0.95 a	1.30 ± 1.17 b	1.26 ± 1.19 b
Trochlear notch of the ulna	0.85 ± 0.81 a	0.87 ± 0.88 a	1.36 ± 1.01 b	0.76 ± 0.89 a
Humeroulnar articulation	0.54 ± 0.79 ab	0.76 ± 0.94 c	0.42 ± 0.74 a	0.61 ± 0.84 b
Ulnar anconeal process, proximal aspect	0.91 ± 1.03 ab	0.98 ± 1.01 a	0.93 ± 1.05 ab	0.87 ± 1.10 b
Craniocaudal view				
Medial humeral epicondyle	0.76 ± 0.95	0.84 ± 0.92	0.80 ± 1.00	0.70 ± 0.88
Humeral condyle, medial aspect	0.71 ± 0.90 a	0.85 ± 0.90 b	0.73 ± 0.95 a	0.78 ± 0.94 ab
Humeral condyle, distal aspect	0.29 ± 0.55 a	0.47 ± 0.79 b	0.57 ± 0.80 b	0.48 ± 0.71 b
Humeral condyle, lateral aspect	0.43 ± 0.71 a	0.74 ± 1.05 b	0.45 ± 0.78 a	0.53 ± 0.75 a
Lateral humeral epicondyle	0.21 ± 0.49 a	0.51 ± 1.04 b	0.36 ± 0.72 bc	0.27 ± 0.57 ac
Medial coronoid process, medial aspect	0.85 ± 0.87 a	1.10 ± 1.02 b	1.38 ± 1.10 c	1.10 ± 1.12 b
Radial head, proximal aspect	0.22 ± 0.48 a	0.50 ± 0.83 b	0.75 ± 0.95 c	0.41 ± 0.72 b
Radial head, lateral aspect	0.19 ± 0.43 a	0.56 ± 0.83 b	0.83 ± 0.97 c	0.65 ± 0.96 b
Joint capsule, medial aspect	0.41 ± 0.75 a	0.03 ± 0.16 b	0.00 ± 0.00 b	0.40 ± 0.79 a
Joint capsule, lateral aspect	0.11 ± 0.36 a	0.00 ± 0.00 a	0.00 ± 0.00 a	0.23 ± 0.56 b
All regions	0.55 ± 0.51 a	0.79 ± 0.75 b	0.81 ± 0.71 b	0.68 ± 0.70 a

MIP = Maximum intensity projection; MPR = Multiplanar reconstructed; TSR = 3-D surface rendering.

Observers evaluated radiographic images and 3-D MIP, TSR, and MPR CT images and assigned a score for each region from 0 to 3, as follows: 0 point, normal (no evidence of osteophytes or sclerosis in the joint); 1 point, mild (osteophytes ≤ 2 mm anywhere in the joint or sclerosis of the ulnar trochlear notch); 2 points, moderate (osteophytes > 2 but ≤ 5 mm anywhere in the joint); and 3 points, severe (osteophytes ≥ 5 mm anywhere in the joint).

Within a row, mean scores with different letters are significantly ($P < .05$) different.

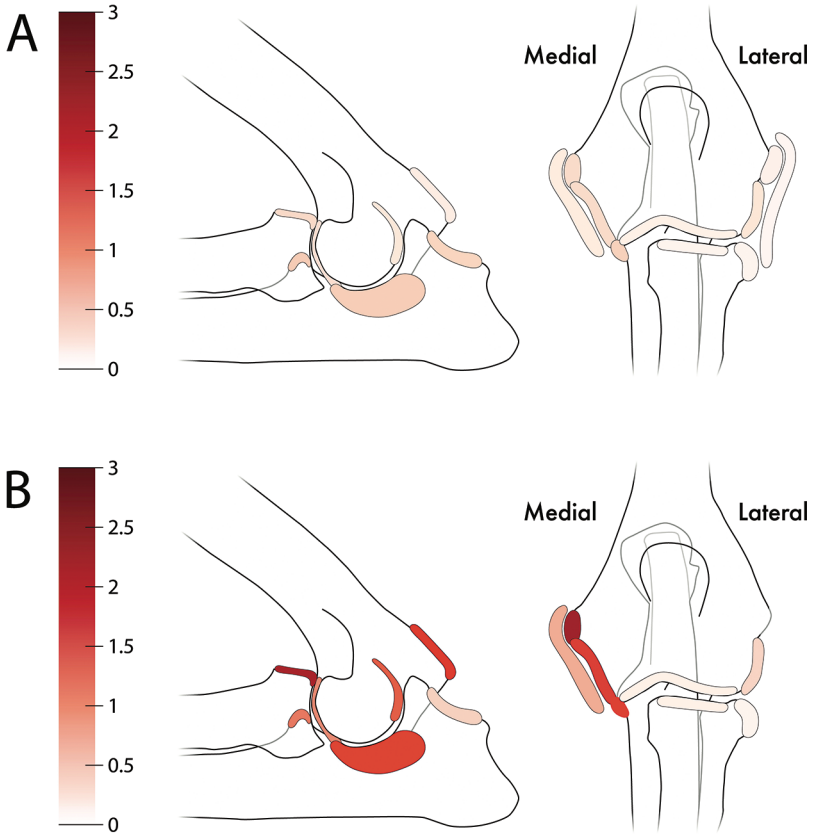

Figure 2—Severity maps showing mean osteoarthritis (OA) severity for the 18 regions illustrated in Figure 1 in elbow joints that received a mean International Elbow Working Group score of 1 point (mild OA; A) or 3 points (severe OA; B). In elbow joints with mild OA, changes were most visible at the medial and cranial aspects of the medial coronoid process, the ulnar notch, the proximal aspect of the anconeal process, and medial aspect of the humeral condyle. In elbows with severe OA, changes were most visible at the cranial aspect of the radial head, medial aspect of the condyle, and ulnar notch.

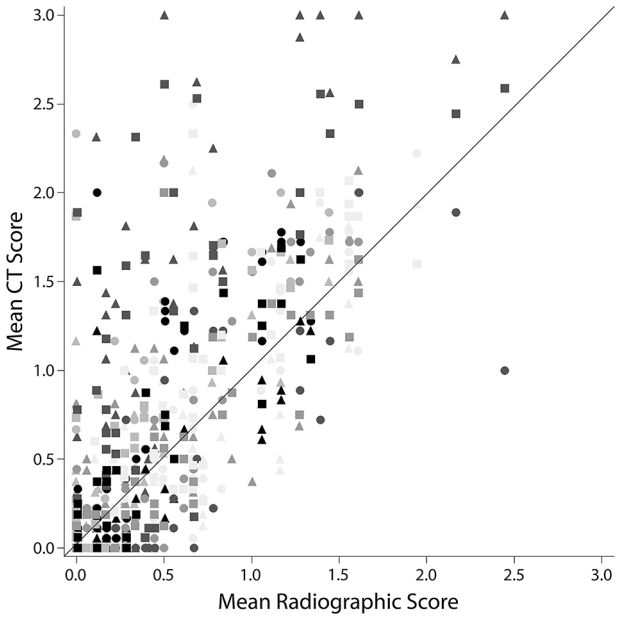

Figure 3—Scatterplot of mean CT scores versus mean radiographic scores of osteoarthritis severity in 39 canine elbow joints. Scores were assigned by 5 observers and represent mean scores for the 18 anatomic regions shown in Figure 1. Each observer assigned scores ranging from 0 to 3 points by reading radiographic images and 3-D maximum intensity projection (triangles), 3-D surface rendering (squares), and multiplanar reconstructed (circles) CT images. Scores assigned by observers 1 to 5 are shown in 5 increasingly darker shades of gray.

images, 0.814 for MPR images, 0.739 for radiographic images, and 0.554 for IEWG scores. Consistency among methods (ICC [3,1]) was 0.816 for observer 1, 0.706 for observer 2, 0.749 for observers 3 and 4, and 0.715 for observer 5 **(Figure 3)**.

Comparisons of scoring methods

MIP scores were significantly different from radiographic scores for 12/18 regions, TSR scores were significantly different from radiographic scores for 11/18 regions, and MPR scores were significantly different from radiographic scores for 8/18 regions. Median differences between radiographic and CT scores were 0.20 for MIP scores, 0.22 for TSR scores, and 0.11 for MPR scores.

Spearman correlation coefficients relative to radiographic scores for all observers combined were 0.653 for MIP images, 0.716 for TSR images, 0.742 for MPR images, and 0.881 for IEWG scores ($P < .001$ for all comparisons). Mean scores for all methods differed significantly; bias relative to radiographic scores, as determined from Bland–Altman plots, was +0.23 for MIP images ($P < .001$), +0.26 for TSR images ($P < .001$), +0.12 for MPR images ($P = .042$), and +0.95 for IEWG scores ($P < .001$; **Figure 4**). Bias was significantly greater for IEWG scores than for MIP, TSR, and MPR scores ($P < .001$ for all comparisons), and was significantly greater for TSR than MPR scores ($P = .042$). Overestimation of MIP, TSR, MPR, and IEWG scores relative to radiographic

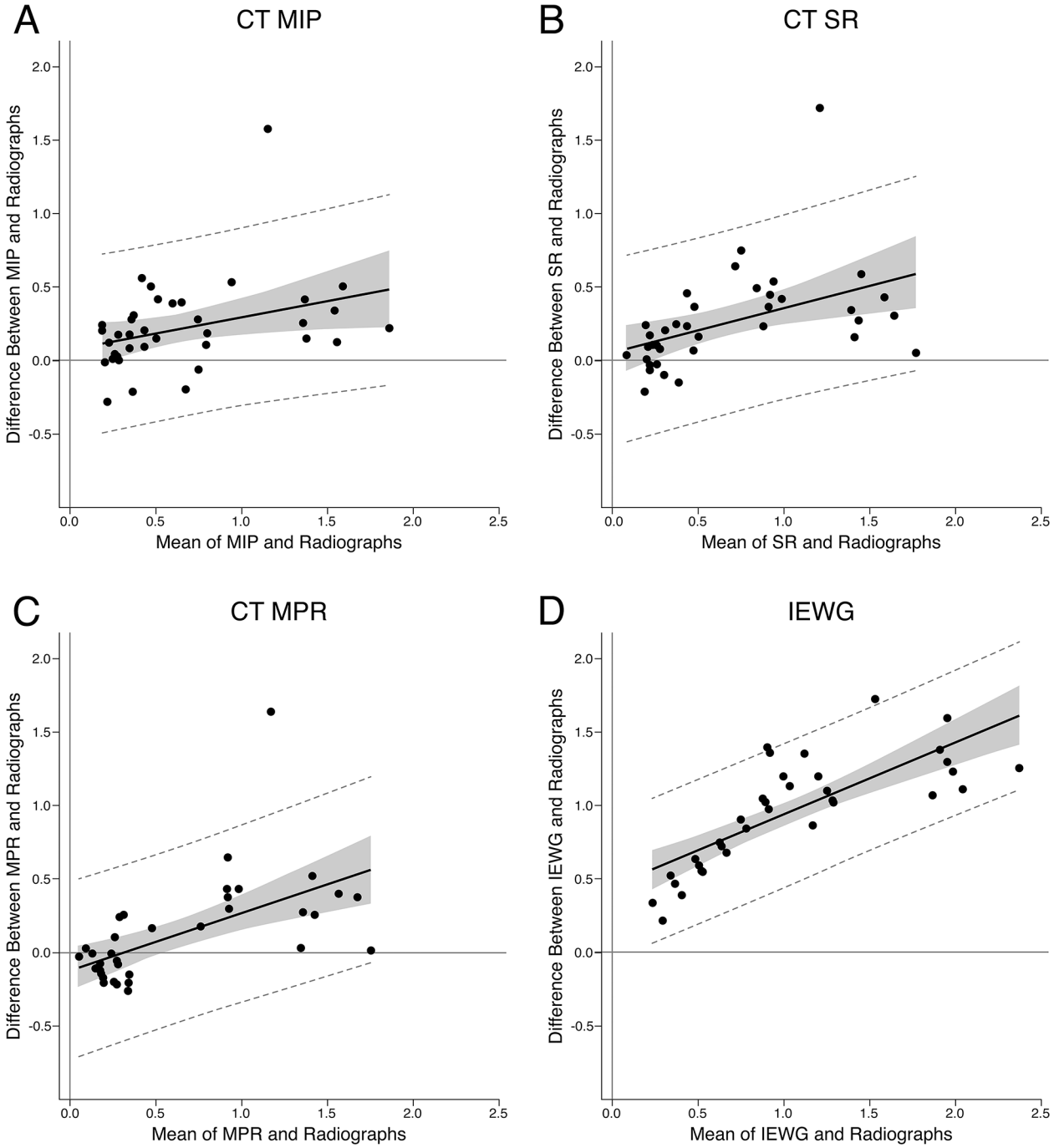

Figure 4—Bland–Altman plots (average of mean CT and mean radiographic scores for the 18 anatomic regions shown in Figure 1 vs the difference between mean CT and mean radiographic scores) of osteoarthritis scores for 39 canine elbow joints assigned by examining maximum intensity projection (MIP; A), 3-D surface rendering (TSR; B), and multiplanar reconstructed (MPR; C) CT images and a Bland–Alton plot of radiographic scores versus International Elbow Working Group (IEWG scores; D). Solid lines represent the regression line, shaded areas represent the 95% CI of the regression line, and dashed lines represent the 95% prediction interval. A positive bias of 0.23 was present for MIP images, 0.26 for TSR images, 0.12 for MPR images, and 0.95 for IEWG scores. For IEWG scores, overestimation of osteoarthritis (OA) scores, compared with radiographic scores, increased significantly (P = .025) as OA scores increased. SR = Surface rendering.

scores increased as scores increased. This increase was significant for IEWG scores (P = .025). Elbow joints with IEWG scores of 1 point had mean radiographic scores ranging from 0.06 to 0.83, elbow joints with IEWG scores of 2 points had mean radiographic scores ranging from 0.17 to 1.61, and elbows with IEWG scores of 3 points had mean radiographic scores ranging from 0.28 to 2.44 **(Figure 5)**.

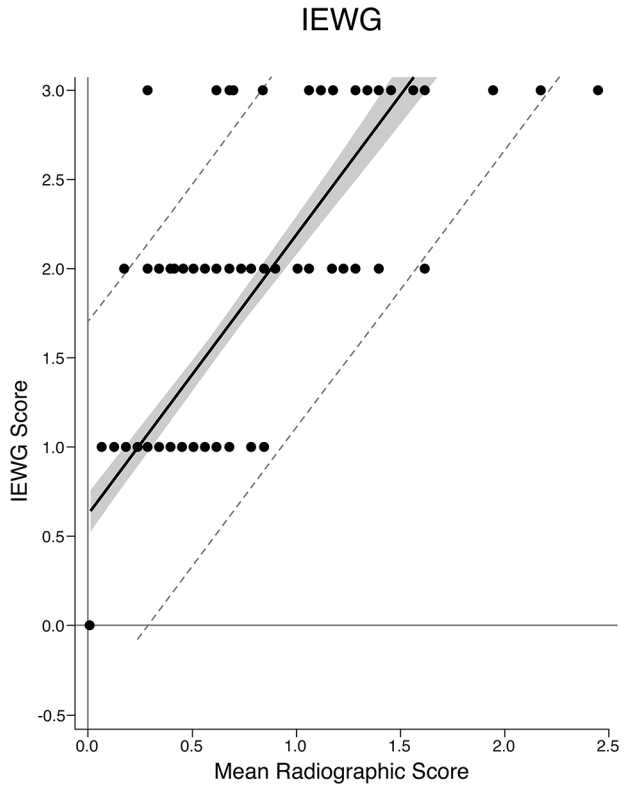

Figure 5—Scatterplot of International Elbow Working Group (IEWG) scores versus mean radiographic osteoarthritis scores for the 18 anatomic regions shown in Figure 1 for 39 canine elbow joints. Scores were assigned by 5 observers. Mean radiographic scores varied widely for dogs with IEWG scores of 1, 2, or 3 points.

Discussion

Findings of our study indicate that evaluation of elbow OA in dogs with MPR images is repeatable within and among observers, and consistent with radiographic evaluation when findings for 18 regions of the elbow are compared. Evaluation of TSR images was also repeatable within and among observers, and consistent with radiographic evaluation, albeit less so than evaluation of MPR images. Evaluation of MIP images appears less repeatable. Consistency between evaluation methods appears higher for observers with more than 10 years of experience than for observers with less than 10 years of experience.

The use of MIP images to evaluate elbow OA in our study led to a slight overestimation of the mean OA score (mean score difference, +0.23) compared with the radiographic score. Thus, we accepted our hypothesis that MIP scores would be higher than radiographic scores. Intraobserver consistency when reading MIP images was moderate or better for 11/16 regions evaluated, and interobserver consistency was moderate or better for only 5 of these 16 regions. We therefore rejected the hypothesis that CT scoring of elbow OA on the basis of MIP renderings was repeatable within or among observers. MIP images are not commonly used for CT assessment, compared with TSR and MPR images. Compared with TSR images, MIP images have the potential to demonstrate differences in attenuation, which could be pertinent for detecting subchondral sclerosis that would not be visible on TSR images. In our study, scores for the evaluation of the trochlear ulna notch, where subchondral sclerosis would be anticipated, were higher for TSR images than for images obtained with the other 3 methods. This was likely because bone changes at the surface of the notch were more visible on TSR images than on images obtained with other methods. Observers did not specifically grade subchondral sclerosis relative to the presence of osteophytes or enthesophytes. MIP renderings have been used to optimize positioning when acquiring radiographs of the human knee joint.[27] MIP contrast-enhanced magnetic resonance images have been shown to have high specificity (≥ 0.98) and to be superior to physical examination for detection of rheumatoid arthritis and prediction of joint destruction.[28–30]

The use of TSR images to evaluate elbow OA also led to a slight overestimation of mean OA score (mean score difference, +0.26) compared with radiographic scores. Thus, we accepted the hypothesis that TSR scores would be higher than radiographic scores. Intraobserver consistency was moderate or better for 11/16 regions evaluated, and interobserver consistency was moderate or better for 9 of these 16 regions. We therefore accepted the hypothesis that CT scoring of elbow OA on the basis of TSR images was repeatable within and among observers. Surface renderings have been used sparsely to evaluate OA. Compared with volumetric renderings, TSR images can detect surface changes and could facilitate the evaluation of anatomic relationships between bones, but they cannot detect changes occurring below the bone surface, such as sclerosis or cysts.[31,32] TSR images have been used to evaluate the integrity of the subchondral bone in joints with OA.[33]

OA scores obtained when reading MPR images were only slightly higher than radiographic scores, with MPR scores overestimating radiographic scores by only 0.12, on average. We rejected the hypothesis that MPR scores would be higher than radiographic scores. MPR images, however, revealed more changes at the cranial and medial aspects of the medial coronoid process (mean score difference, +0.37 and +0.25, respectively), most likely because of enhanced visualization of regions where the bones were superimposed when viewed on radiographs. The small bias between MPR and radiographic scores suggests that the progression of OA of the elbow joint in dogs could be evaluated with MPR renderings as well as radiographs. In a study[34] comparing the sensitivity and specificity of CT and radiography to detect elbow dysplasia in 180 dogs, both CT and radiography were highly sensitive (CT sensitivity, 100%; radiography sensitivity, 98%), but CT had a higher specificity (93%) than radiography (64% to 69%). In our study, intraobserver consistency was moderate or better for 17/18 regions, and interobserver consistency was moderate or better

for 12/18 regions when reading MPR images. Thus, we accepted the hypothesis that the interpretation of MPR images was repeatable within and among observers. MPR images are the standard CT rendering method to evaluate elbow OA in dogs.[35] Osteophytes, enthesophytes, and subchondral sclerosis are visible on all 3 MPR views.[13] Coronoid fragmentation and radioulnar fit are evaluated on transverse-plane images, radioulnar and humeroradial fit are evaluated on sagittal-plane images, and radioulnar fit is evaluated on frontal-plane images.[35] In a study[36] in which 8 observers read 84 MPR views of the coronoid process twice, intraobserver and interobserver repeatability of evaluations of abnormalities of the medial coronoid process, osteophytes, and subchondral sclerosis ranged widely from poor to excellent, depending on the parameter evaluated. Agreement was poor for grading subchondral sclerosis, fair to moderate for medial coronoid abnormalities, and moderate for detection of osteophytes.[36] Those findings were similar to the findings of our study. In human elbows, the evaluation of elbow OA from MPRs has been shown to have good to excellent intraobserver and interobserver repeatability, and to be more repeatable than radiographic assessment.[37] When evaluating human knee joints, reformatted coronal (frontal)-plane CT images were more sensitive and more accurate than radiographs for the detection of osteophytes and subchondral cysts.[38] When evaluating OA in human hip joints, a CT-based 6- to 8-mm-thick slab reconstructed in the coronal (frontal) plane had higher intra- and interobserver reliability than radiographs.[39]

When scoring radiographs with the IEWG grading system, the mean OA score was higher than the mean radiographic score by nearly a full grade (mean score difference, +0.95). Thus, we accepted the hypothesis that IEWG scores would be higher than radiographic scores. This overestimation is logical, because IEWG scoring is based on the most severely affected region, rather than the mean severity for all regions, and elbow OA is not distributed uniformly in the elbow joint. For example, OA is often more pronounced on the medial aspect than on the lateral aspect of the joint.[10] IEWG grading had the lowest interobserver agreement numerically (0.554) of all methods evaluated in our study.

Consistency of scoring across modalities appeared higher for the 3 observers with more than 10 years of experience since completing residency training, compared with consistency for the 2 observers with less than 10 years of experience. Expertise in reading medical images is complex and highly dependent on the number of cases read,[40] and is, therefore, dependent on the number of years of clinical experience. Accuracy differences among observers may also have resulted from differences in computer hardware or from differences in visualization techniques.

In our study, the scoring of OA severity in 18 regions allowed us to produce a severity map of OA for the dogs included in the study. OA was most severe on the medial and cranial aspects of the medial coronoid process, the proximal aspect of the anconeal process, and the medial aspect of the joint. Severity maps, also named heat maps, have been reported for arthritic human hips.[41] Severity maps may be a potential tool to predict clinical disease and assist in therapy development.[42,43] Severity maps are useful in the development of deep neural networks that can be used to evaluate radiographic signs of OA, as reported for human knees.[44]

We conclude from our results that scoring OA severity at 18 regions of the elbow joint in dogs can be done with acceptable intraobserver consistency and interobserver agreement on the basis of TSR and MPR images, and that MIP, TSR, and MPR scores are only slightly higher than scores obtained from reading radiographs. We also concluded that MPR scoring appeared more consistent and closer to radiographic assessments than MIP and TSR scoring.

Acknowledgments

No third-party funding or support was received in connection with this study or the writing or publication of the manuscript. The authors declare there were no conflicts of interest.

The authors thank Dr. Anke Langenbach, Veterinary Surgical Centers, Vienna, VA, for providing case data and Chrisoula Toupadakis Skouritakis, PhD, University of California-Davis, Davis, CA, for the illustrations.

References

1. Keller GG, Dziuk E, Bell JS. How the Orthopedic Foundation for Animals (OFA) is tackling inherited disorders in the USA: using hip and elbow dysplasia as examples. *Vet J.* 2011;189:197–202. doi:10.1016/j.tvjl.2011.06.019
2. Lewis TW, Blott SC, Woolliams JA. Comparative analyses of genetic trends and prospects for selection against hip and elbow dysplasia in 15 UK dog breeds. *BMC Genet.* 2013;14:16. doi:10.1186/1471-2156-14-16
3. Farrell M, Heller J, Solano M, Fitzpatrick N, Sparrow T, Kowaleski M. Does radiographic arthrosis correlate with cartilage pathology in Labrador Retrievers affected by medial coronoid process disease? *Vet Surg.* 2014;43:155–165. doi:10.1111/j.1532-950X.2014.12092.x
4. Theyse LF, Voorhout G, Hazewinkel HA. Prognostic factors in treating antebrachial growth deformities with a lengthening procedure using a circular external skeletal fixation system in dogs. *Vet Surg.* 2005;34:424–435. doi:10.1111/j.1532-950X.2005.00064.x
5. Bergstrom A, Johard S, Lee MH, Comin A. Long-term prognosis of quality of life in dogs diagnosed with mild to moderate elbow dysplasia in Sweden. *Front Vet Sci.* 2020;7:572691. doi:10.3389/fvets.2020.572691
6. Serrani D, Sassaroli S, Gallorini F, et al. Clinical and radiographic evaluation of short- and long-term outcomes of different treatments adopted for elbow medial compartment disease in dogs. *Vet Sci.* 2022;9:70. doi:10.3390/vetsci9020070
7. Huck JL, Biery DN, Lawler DF, et al. A longitudinal study of the influence of lifetime food restriction on development of osteoarthritis in the canine elbow. *Vet Surg.* 2009;38:192–198. doi:10.1111/j.1532-950X.2008.00487.x
8. Hou Y, Wang Y, Lu X, et al. Monitoring hip and elbow dysplasia achieved modest genetic improvement of 74 dog breeds over 40 years in USA. *PLoS One* 2013;8:e76390. doi:10.1371/journal.pone.0076390
9. Andersson A, Bergstrom A. Adaptation of the Canine Orthopaedic Index to evaluate chronic elbow

osteoarthritis in Swedish dogs. *Acta Vet Scand.* 2019;61:29. doi:10.1186/s13028-019-0465-1
10. Lappalainen AK, Molsa S, Liman A, Snellman M, Laitinen-Vapaavuori O. Evaluation of accuracy of the Finnish elbow dysplasia screening protocol in Labrador retrievers. *J Small Anim Pract.* 2013;54:195–200. doi:10.1111/jsap.12051
11. James HK, McDonnell F, Lewis TW. Effectiveness of canine hip dysplasia and elbow dysplasia improvement programs in six UK pedigree breeds. *Front Vet Sci.* 2019;6:490.
12. Lappalainen AK, Molsa S, Liman A, Laitinen-Vapaavuori O, Snellman M. Radiographic and computed tomography findings in Belgian shepherd dogs with mild elbow dysplasia. *Vet Radiol Ultrasound.* 2009;50:364–369. doi:10.1111/j.1740-8261.2009.01551.x
13. Moores AP, Benigni L, Lamb CR. Computed tomography versus arthroscopy for detection of canine elbow dysplasia lesions. *Vet Surg.* 2008;37:390–398. doi:10.1111/j.1532-950X.2008.00393.x
14. Groth AM, Benigni L, Moores AP, et al. Spectrum of computed tomographic findings in 58 canine elbows with fragmentation of the medial coronoid process. *J Small Anim Pract.* 2009;50:15–22. doi:10.1111/j.1748-5827.2008.00656.x
15. Kunst CM, Pease AP, Nelson NC, Habing B, Ballegeer EA. Computed tomographic identification of dysplasia and progression of osteoarthritis in dog elbows previously assigned OFA grades 0 and 1. *Vet Radiol Ultrasound.* 2014;55:511–520. doi:10.1111/vru.12171
16. Wennemuth J, Tellhelm B, Eley N, von Pückler K. Computed tomography enhances diagnostic accuracy in challenging medial coronoid disease cases: an imaging study in dog breeding appeal cases. *Vet Comp Orthop Traumatol.* 2020;33:356–362. doi:10.1055/s-0040-1714299
17. Tromblee TC, Jones JC, Bahr AM, Shires PK, Aref S. Effect of computed tomography display window and image plane on diagnostic certainty for characteristics of dysplastic elbow joints in dogs. *Am J Vet Res.* 2007;68:858–871. doi:10.2460/ajvr.68.8.858
18. Alves-Pimenta S, Ginja MM, Fernandes AM, Ferreira AJ, Melo-Pinto P, Colaço B. Computed tomography and radiographic assessment of congruity between the ulnar trochlear notch and humeral trochlea in large breed dogs. *Vet Comp Orthop Traumatol.* 2017;30:8–14. doi:10.3415/VCOT-16-03-0045
19. De Rycke LM, Gielen IM, van Bree H, Simoens PJ. Computed tomography of the elbow joint in clinically normal dogs. *Am J Vet Res.* 2002;63:1400–1407. doi:10.2460/ajvr.2002.63.1400
20. Ondreka N, Tellhelm B. Explanation of grading according to IEWG and discussion of cases. In: 31st Annual Meeting of the International Elbow Working Group (IEWG). 2017:33–34.
21. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psych Bull.* 1979;86:420–428. doi:10.1037/0033-2909.86.2.420
22. McAlinden C, Khadka J, Pesudovs K. Precision (repeatability and reproducibility) studies and sample-size calculation. *J Cataract Refract Surg.* 2015;41:2598–2604. doi:10.1016/j.jcrs.2015.06.029
23. Barnhart HX, Barboriak DP. Applications of the repeatability of quantitative imaging biomarkers: a review of statistical analysis of repeat data sets. *Transl Oncol.* 2009;2:231–235. doi:10.1593/tlo.09268
24. Weaver B, Wuensch KL. SPSS and SAS programs for comparing Pearson correlations and OLS regression coefficients. *Behav Res Methods.* 2013;45:880–895. doi:10.3758/s13428-012-0289-7
25. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat.* 2007;17:571–582. doi:10.1080/10543400701329422
26. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* 3rd ed. Prentice Hall; 2008.
27. Wang S, Xiao Z, Lu Y, Zhang Z, Lv F. Radiographic optimization of the lateral position of the knee joint aided by CT images and the maximum intensity projection technique. *J Orthop Surg Res.* 2021;16:581. doi:10.1186/s13018-021-02740-8
28. Mori G, Tokunaga D, Takahashi KA, et al. Maximum intensity projection as a tool to diagnose early rheumatoid arthritis. *Mod Rheumatol.* 2008;18:247–251. doi:10.3109/s10165-008-0043-2
29. Akai T, Taniguchi D, Oda R, et al. Prediction of radiographic progression in synovitis-positive joints on maximum intensity projection of magnetic resonance imaging in rheumatoid arthritis. *Clin Rheumatol.* 2016;35:873–878. doi:10.1007/s10067-016-3208-y
30. Taniguchi D, Tokunaga D, Oda R, et al. Maximum intensity projection with magnetic resonance imaging for evaluating synovitis of the hand in rheumatoid arthritis: comparison with clinical and ultrasound findings. *Clin Rheumatol.* 2014;33:911–917. doi:10.1007/s10067-014-2526-1
31. Udupa JK, Hung HM, Chuang KS. Surface and volume rendering in three-dimensional imaging: a comparison. *J Digit Imaging.* 1991;4:159–168. doi:10.1007/BF03168161
32. Cook CR, Cook JL. Diagnostic imaging of canine elbow dysplasia: a review. *Vet Surg.* 2009;38:144–153. doi:10.1111/j.1532-950X.2008.00481.x
33. Appleton CT, McErlain DD, Pitelka V, et al. Forced mobilization accelerates pathogenesis: characterization of a preclinical surgical model of osteoarthritis. *Arthritis Res Ther.* 2007;9:R13. doi:10.1186/ar2120
34. Villamonte-Chevalier A, van Bree H, Broeckx B, et al. Assessment of medial coronoid disease in 180 canine lame elbow joints: a sensitivity and specificity comparison of radiographic, computed tomographic and arthroscopic findings. *BMC Vet Res.* 2015;11:243. doi:10.1186/s12917-015-0556-9
35. Samoy Y, Gielen I, Van Caelenberg A, van Bree H, Duchateau L, Van Ryssen B. Computed tomography findings in 32 joints affected with severe elbow incongruity and fragmented medial coronoid process. *Vet Surg.* 2012;41:486–494. doi:10.1111/j.1532-950X.2011.00950.x
36. Shimizu N, Warren-Smith CM, Langley-Hobbs SJ, et al. Inter- and intraobserver agreement in interpretation of CT features of medial coronoid process disease. *J Small Anim Pract.* 2015;56:707–713. doi:10.1111/jsap.12411
37. Kwak JM, Kholinne E, Sun Y, Alhazmi AM, Koh KH, Jeon IH. Intraobserver and interobserver reliability of the computed tomography-based radiographic classification of primary elbow osteoarthritis: comparison with plain radiograph-based classification and clinical assessment. *Osteoarthritis Cartilage.* 2019;27:1057–1063. doi:10.1016/j.joca.2019.03.004
38. Segal NA, Nevitt MC, Lynch JA, Niu J, Torner JC, Guermazi A. Diagnostic performance of 3D standing CT imaging for detection of knee osteoarthritis features. *Phys Sportsmed.* 2015;43:213–220. doi:10.1080/00913847.2015.1074854
39. Turmezei TD, Fotiadou A, Lomas DJ, Hopper MA, Poole KE. A new CT grading system for hip osteoarthritis. *Osteoarthritis Cartilage.* 2014;22:1360–1366. doi:10.1016/j.joca.2014.03.008
40. Nodine CF, Mello-Thoms C. The nature of expertise in radiology In: Beutel J, Kundel HL, Van Metter RL, eds. *Handbook of Medical Imaging, Volume 1: Physics and Psychophysics.* International Society for Optics and Photonics; 2000:859–894.
41. Turmezei TD, Lomas DJ, Hopper MA, Poole KES. Severity mapping of the proximal femur: a new method for assessing hip osteoarthritis with computed tomography. *Osteoarthritis Cartilage.* 2014;22:1488–1498. doi:10.1016/j.joca.2014.03.007
42. Turmezei TD, Treece GM, Gee AH, Fotiadou AF, Poole KE. Quantitative 3D analysis of bone in hip osteoarthritis using clinical computed tomography. *Eur*

Radiol. 2016;26:2047–2054. doi:10.1007/s00330-015-4048-x
43. Turmezei TD, Treece GM, Gee AH, et al. Quantitative 3D imaging parameters improve prediction of hip osteoarthritis outcome. *Sci Rep.* 2020;10:4127. doi:10.1038/s41598-020-59977-2
44. Thomas KA, Kidzinski L, Halilaj E, et al. Automated classification of radiographic knee osteoarthritis severity using deep neural networks. *Radiol Artif Intell.* 2020;2:e190065. doi:10.1148/ryai.2020190065

Supplementary Materials

Supplementary materials are posted online at the journal website: avmajournals.avma.org