Observer variability and sensitivity of radiographic diagnosis of canine medial coronoid disease*

F. C. Rau; A. Wigger; B. Tellhelm; M. Zwick; S. Klump; A. Neumann; B. Oltersdorf; K. Amort; K. Failing; M. Kramer

1Department of Veterinary Clinical Sciences, Small Animal Clinic – Surgery, Justus-Liebig-University Giessen, Germany; 2Department of Veterinary Medicine, Unit of Biomathematics and Data Processing, Justus-Liebig-University, Giessen, Germany

Key words
- Fragmented medial coronoid process, intraobserver agreement, interobserver agreement, radiographic assessment and sensitivity, arthroscopy, computed tomography, dog

Summary
Objective: Medial coronoid disease (MCD) is a very common form of elbow joint disease and it’s radiographic diagnosis can be challenging since it is frequently based on the detection of rather subtle primary or secondary changes than on a large primary lesion. We hypothesized that accuracy of radiographic diagnosis of MCD is highly dependent on training and experience level. Methods: Radiographs of 102 canine elbows were evaluated for MCD by four observers with different levels of training and experience. All elbows underwent CT scans and arthroscopy. Sensitivity and specificity of radiographic and CT interpretation was determined using arthroscopy as a gold standard. Inter- and intraobserver agreement (reliability and repeatability) were assessed by using Cohen’s Kappa (κ) statistic. Results: The sensitivity (92.4–96.7%) of the two experienced observers was almost comparable to that of CT (100%) and significantly higher than that of the two less experienced observers (77.2–80.4%). Reliability of the radiographic diagnosis of MCD was better between observers with higher experience level (κ = 0.74) than between observers of lower or different experience levels (κ = 0.07–0.42). Repeatability was better in experienced (κ = 0.73–0.88) than in less experienced observers (κ = 0.31–0.42). Conclusion: Our results confirm that training and experience play important roles in reaching high sensitivity, reliability and repeatability for the radiographic diagnosis of MCD.

Clinical relevance: Although radiography is inferior to CT in imaging of the medial coronoid process itself, sensitivity of radiographic diagnosis MCD can be significantly improved with observer experience almost reaching that of CT. Therefore, it is advised that radiographic screening for MCD should be performed by specialists experienced in the radiographic evaluation of elbow joint disease.

Correspondence to
Dr. med. vet. Friederike C. Rau, Resident ECVDI
Department for Veterinary Clinical Sciences
Small Animal Clinic – Surgery
Justus-Liebig-University Giessen
Frankfurter Strasse 108
35392 Giessen, Germany
Email: friederike.c.rau@gmail.com

* Dedicated to Prof. Dr. Ulrike Matis.

Schlüsselwörter
- Koronoiderkrankung, radiologische Beurteilung, Intra- und Interobservererübereinstimmung, Arthroskopie, Computertomographie, Hund

Zusammenfassung
Ziel: Die Koronoiderkrankung (MCD) ist eine sehr häufig auftretende Form der Elbogengelenkerkrankung. Ihre Röntgendiagnose gestaltet sich häufig schwierig, da sie in vielen Fällen mehr auf der subjektiven Interpretation von geringen oder sekundären Veränderungen als auf die Detektion der primären Läsion selbst beruht. Ziel der Studie war, den Einfluss des Ausbildungs- und Erfahrungsgrades des Auswerters auf die Exaktheit der Röntgendiagnose MCD zu untersuchen. Methoden: Vier Auswerter mit verschiedenem Ausbildungs- und Erfahrungsgrad werteten Röntgenaufnahmen von 102 Hundeelbows im Hinblick auf MCD aus. Für alle Ellbogen lagen die Ergebnisse der computertomographischen (CT) und arthroskopischen Untersuchung vor. Als Goldstandard zur Bestimmung der Sensitivität und Spezifität der Röntgen- und CT-Diagnose diente die Arthroskopiediagnose. Die Inter- und Intraauswerterkonkordanz der Röntgendiagnose MCD wurden mit Hilfe des Cohens’s Kappa (κ) Statistiktest ermittelt. Ergebnisse: Die Sensitivität (92,4–96,7%) der beiden erfahrensten Auswerter war fast vergleichbar hoch mit der der CT (100%) und signifikant höher als die der weniger erfahrenen Auswerter (77,2–80,4%). Die Inter- und Intraauswerterkonkordanz der Röntgendiagnose MCD wurden mithilfe des Cohens’s Kappa (κ) Statistiktest ermittelt. Schlussfolgerung: Die Inter- und Intraauswerterkonkordanz der Röntgendiagnose MCD war besser als zwischen Auswertern mit weniger oder unterschiedlich großer Erfahrung (κ = 0.07–0.42). Ebenso ergab sich zwischen erfahrenen Auswertern, dass sie eine bessere Interobserverkonkordanz (Wiederholbarkeit) (κ = 0.73–0.88) als zwischen weniger erfahrenen Auswertern (κ = 0.31–0.42) aufwiesen. Die Sensitivität der CT-Diagnose erhöhte sich von 77,2–80,4% auf 92,4–96,7%. Klinische Relevanz: Obwohl das Röntgen im Gegensatz zur CT eine überlagerungsfreie Darstellung des medialen Koronoids ermöglicht, kann durch eine entsprechende Auswertererfahrung die Sensitivität der Röntgendiagnose MCD signifikant verbessert werden und fast die Sensitivität der CT-Diagnose erreichen. Folglich sollte bei Hunden das “Röntgenscreening” für MCD von Spezialisten durchgeführt werden.

Tierärztliche Praxis Kleintiere 5/2011
Introduction

Elbow dysplasia (ED) is one of the most common heritable diseases in dogs (17) and the most frequent cause of canine elbow osteoarthritis (13). The fragmented medial coronoid process (FCP) is the most common form of ED (11) causing forelimb lameness in juvenile dogs of large breeds, particularly Labrador retrievers and Rottweilers (10, 13, 24, 25). Thus, it is important that FCP is diagnosed as early as possible in order to allow appropriate therapy. The term medial coronoid process (MCD) has been introduced by some other authors (10, 25) to cover the entire spectrum of the medial coronoid process (MCP) lesions: fissuring and fragmentation of the medial coronoid process, and subchondral bone and/or cartilage pathology (10). Subchondral bone pathology often is clinically demonstrated by avascularity and softening of bone. Cartilage pathology of the MCP has been described as chondromalacia-like lesions (35, 36) or as “medial compartment erosions” (36) if the entire medial part of the elbow joint is severely affected, in absence of fragmentation or fissuring. Different diagnostic imaging techniques can be used to assess the medial coronoid process (MCP) of the ulna and establish a diagnosis of MCD such as radiography (5, 16, 25), computed tomography (CT) (28), magnetic resonance imaging (MRI) (31) and scintigraphy (34). However, an invasive method like arthroscopy is frequently required to confirm diagnosis (5, 7).

In spite of the fact that radiographic diagnosis of MCD is difficult and radiography is inferior to advanced imaging modalities such as CT and MRI in imaging the MCP (4, 31) radiography still remains the primary imaging modality in general practice because it is widely available, requires no or short sedation of the patient and is inexpensive. Screening for elbow dysplasia (ED) according to the International Elbow Working Group (IEWG) guidelines (32) is routinely performed by evaluating radiographs for ED lesions and osteoarthritis.

A distinct disadvantage in the radiographic diagnosis of MCD remains the inability to define the cranioproximal and medial border of the MCP without superimposition of other bony structures on any radiographic projection (15), which leads to a low detection rate of free fragments usually seen in subsequent CT or arthroscopic exams (4, 10, 14, 25, 39). Thus, presumptive radiographic diagnosis MCD is frequently based on the detection of secondary arthritic changes like periarticular osteophytosis, rather than the detection of primary lesions. Osteophyte assessment proves to be only sensitive as an indicator of late signs of osteoarthritis (1, 19). However, more recent studies show that the inclusion of other radiographic changes such as loss of definition of the MCP outline on the mediolateral radiographs of elbow joints (16) as well as an increased troclear notch sclerosis (TNS) (3, 9, 16, 30) significantly increases the sensitivity (3, 9, 20) above previously reported sensitivities (4, 19, 28) of radiographic detection of MCD.

In our opinion, detection and interpretation of radiographic signs of MCD, and thus sensitivity of the presumptive radiographic diagnosis, is highly dependent on the training and experience level of the observer. The aim of our study was to investigate the sensitivity of observers with different experience levels in the radiographic detection of MCD compared to CT, using arthroscopy as a gold standard. Further, we were interested in evaluating inter- and intraobserver agreement.

Our hypotheses were:
1) Sensitivity of the radiographic diagnosis of MCD would increase with the experience level of the observer.
2) Agreement between radiographic and CT diagnosis of MCD would increase with the experience level of the observer.
3) Interobserver agreement of observers of higher experience levels would be better than that of observers with less experience.
4) Intraobserver agreement would be better in experienced than less experienced observers.
5) Intraobserver agreement would be better than interobserver agreement in experienced observers.

Material and methods

Study population

In this retrospective study, the clinical database of the Clinic for Small Animals of the University Giessen, Germany, was searched for dogs that had undergone radiographic evaluation, CT and arthroscopic assessment of MCD. Inclusion criteria were thoracic limb lameness localized to the elbow joint, diagnostic quality radiographs, CT scans and arthroscopy. Cases treated for any prior elbow pathology were excluded.

Radiographic, computed tomographic and arthroscopic assessment of MCD

The radiographic projections evaluated were extended cranio 15° lateral-caudomedial oblique (Cr15L-CdMO) and two mediolateral radiographs, one with elbow in 135° flexion (extended ML view) and one with 90° flexion (flexed ML view). Radiographs were obtained with a vertical beam and same focus-film and object-film distance.

Sensitivity, specificity and inter- and intraobserver agreement of the radiographic detection of MCD were determined by four observers of different experience levels (one last-year student, one first-year doctoral student in small animal surgery, one 4th year ECVDI resident and one ECVDI diplomate with extensive experience in scrutinizing ED). The radiographs of 102 elbows were independently assessed for MCD and osteophyte formation to determine interobserver agreement. For intraobserver agreement a subset of the radiographs (60 of 102 elbows) were assessed by all four observers on two separate occasions at least 2 weeks apart. All observers assessed suspicion or absence of MCD or definitive presence of MCD which was determined by fragmentation or fissuring (Table 1). According to the current IEWG guidelines (32) suspi-
Assessment of radiographic, computed tomographic and arthroscopic findings in medial coronoid disease (MCD) and osteoarthrosis. (MCP, medial coronoid process).

**Table 1**

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Grade</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiographic MCD grade</td>
<td>0</td>
<td>normal MCP</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>suspicion of MCD*</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>definitive MCD (fragment/fissure visible)</td>
</tr>
<tr>
<td>Radiographic osteophyte grade</td>
<td>0</td>
<td>no osteophytes</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>mild contour changes; &quot;borderline&quot;</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>maximal osteophyte height &lt; 2 mm</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>maximal osteophyte height 2–5 mm</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>maximal osteophyte height &gt; 5 mm</td>
</tr>
<tr>
<td>Computed tomographic MCD grade</td>
<td>0</td>
<td>normal MCP</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>suspicion of MCD**</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>definitive MCD (fragment/fissure visible)</td>
</tr>
<tr>
<td>Arthroscopic MCD grade</td>
<td>0</td>
<td>normal MCP</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>MCD (fragment, fissure, other MCP pathology)</td>
</tr>
</tbody>
</table>

* Changes in shape, outline, radiopacity of the MCP, increased trochlear notch sclerosis, elbow joint incongruity
** Changes in shape, density of MCP, increased trochlear notch sclerosis, irregular radial incisure, elbow joint congruity

Statistical analysis

Data were analyzed by two-way and multiway frequency tables. Sensitivity was defined as the proportion of elbows that were positive for MCD based on arthroscopy and correctly diagnosed as MCD suspicious or MCD positive with radiography and CT, respectively.

Specificity was defined as the proportion of elbows, that were negative for MCD in arthroscopy and correctly diagnosed as MCD negative with radiography or CT. Specificity and sensitivity were calculated for each observer. Two by two tables were used for sensitivity and specificity calculations. Additionally, 95% confidence intervals were determined for both values.

Inter- and intraobserver agreement was assessed by using Cohen’s Kappa (κ) statistic. Interpretation (18) of the Cohen’s Kappa (κ) value was performed as shown in Table 2. Statistical analyses were performed with help of a statistical software program (BMDP/Dynamic Statistical Soft ware Release 8.2 BMDP4F) developed for two-way or multiway frequency tables (8).

Results

Study population

A total of 102 (51 left and 51 right) elbows of 65 dogs met all inclusion criteria. The median age of the dogs was 14 months (range: 5–122 months) and the median weight was 32 kg (range 19–56 kg).
Forty-five dogs were male (69.2%) and 20 were female (30.8%). Labrador Retrievers (29.2%) were the most commonly represented among 17 breeds followed by Golden Retrievers (13.8%) and Rottweilers (12.3%). Fifty-one of 65 (78.5%) dogs were bilaterally and 14 of 65 (21.5%) were unilaterally affected with MCD.

Radiographic, computed tomographic and arthroscopic findings for MCD

Out of the 102 elbows 92 (90.2%) were positive for MCD and 10 (9.8%) were free of any abnormalities affecting the MCP in arthroscopy. The majority of the elbows (78.4%) in our study showed displaced (20.6%) and non-displaced (18.6%) fragments and fissures (39.2%). Only 13 (12.7%) elbows showed subchondral bone and cartilage pathology of the MCP tip exclusively, without any overt fragmentation or fissuring (Table 3).

Depending on the degree of radiological experience of the observer the number of MCD positive elbows on radiographs varied between 79 to 95. Based on CT, 100 out of 102 elbows were rated as MCD positive (Table 3).

The detection rate for fragmentation and fissuring ranged from 12–23 of 102 elbows in radiography, whereas the detection was higher in CT (61 of 102) and arthroscopy (79 of 102). However, the number of elbows that were considered as suspicious for MCD (56–83 of 102) was higher with radiography than with CT (39 of 102) and with arthroscopy (13 of 102) (Table 3).

Radiographically, no osteophyte formation was detected in 14–31 (14.4–30.4%) of the elbows and 11–30 (10.7–19.6%) of the elbows were scored due to minimal contour changes predominantly occurring on the dorsal aspect of the anconeal process as “borderline”, depending on the individual observer. CT detected periartricular osteophytes exceeding minimal contour changes on any aspect of the elbow joint, classified as “borderline” changes, in 82.5% of the cases.

Sensitivity and specificity

Sensitivity of the radiographic assessment for MCD using arthroscopy as gold standard increased with higher level of experience up to 96.7% for the most experienced observer (Table 4). The second most experienced observer (ECVDI resident) had a sensitivity of 92.4% and the second least experienced observer (doctoral student) showed a sensitivity of 80.4%. The least experienced observer (student) showed the lowest sensitivity of 77.2%. Assessment of MCD by reviewing CT scans had a sensitivity of 100%.

Correspondingly, we found the highest specificity of 40% for the two most experienced observers, followed by the least experienced observer with a specificity of 20%. The second least experienced observer reached a specificity of 10% by rating 9 of 10 elbows false positive. Specificity of CT assessment was 30% with 7 out of 10 false positives.

The sensitivity and specificity of CT for detection of fragments was 70% and 60.9%. Radiographic detection of fragments showed significantly lower sensitivity (28.2–38.5%) than CT, independent of the level of experience (Table 5).

### Table 2
Interpretation of Cohen’s Kappa statistic (18) for strength of agreement.

<table>
<thead>
<tr>
<th>Cohen’s Kappa statistic (κ)</th>
<th>Strength of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.00</td>
<td>Poor</td>
</tr>
<tr>
<td>0.00–0.20</td>
<td>Slight</td>
</tr>
<tr>
<td>0.20–0.40</td>
<td>Fair</td>
</tr>
<tr>
<td>0.41–0.60</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.61–0.80</td>
<td>Substantial</td>
</tr>
<tr>
<td>0.81–1.00</td>
<td>Almost perfect</td>
</tr>
</tbody>
</table>

### Table 3
Results of the radiographic assessment of the different observers, computed tomographic (CT) and arthroscopic assessment (Arthros) of the medial coronoid process (MCP) for medial coronoid disease (MCD) in absolute numbers (n).

<table>
<thead>
<tr>
<th>MCP</th>
<th>Student</th>
<th>Doctoral</th>
<th>Resident</th>
<th>Diplomate</th>
<th>CT</th>
<th>Arthros</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total MCD-positive</td>
<td>79</td>
<td>83</td>
<td>91</td>
<td>95</td>
<td>100</td>
<td>92</td>
</tr>
<tr>
<td>Fragmented or fissured MCP</td>
<td>23</td>
<td>17</td>
<td>17</td>
<td>12</td>
<td>61</td>
<td>79</td>
</tr>
<tr>
<td>Suspicion of MCD</td>
<td>56</td>
<td>66</td>
<td>74</td>
<td>83</td>
<td>39</td>
<td>13</td>
</tr>
<tr>
<td>MCD-negative</td>
<td>23</td>
<td>19</td>
<td>11</td>
<td>7</td>
<td>2</td>
<td>10</td>
</tr>
</tbody>
</table>

### Table 4
Sensitivity and specificity of radiographic assessment of each observer and computed tomographic assessment for the diagnosis of medial coronoid disease (MCD) using arthroscopic findings as gold standard; confidence interval to 95th percentile (CI95) in brackets.

<table>
<thead>
<tr>
<th>Assessment MCD</th>
<th>Student</th>
<th>Doctoral</th>
<th>Resident</th>
<th>Diplomate</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%)</td>
<td>77.2</td>
<td>80.4</td>
<td>92.4</td>
<td>96.7</td>
<td>100</td>
</tr>
<tr>
<td>CI95</td>
<td>(0.67–0.85)</td>
<td>(0.71–0.88)</td>
<td>(0.85–0.97)</td>
<td>(0.91–0.99)</td>
<td>(0.97–1.0)</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>20.0</td>
<td>10.0</td>
<td>40.0</td>
<td>40.0</td>
<td>30.0</td>
</tr>
<tr>
<td>CI95</td>
<td>(0.03–0.55)</td>
<td>(0.00–0.45)</td>
<td>(0.12–0.74)</td>
<td>(0.12–0.74)</td>
<td>(0.07–0.65)</td>
</tr>
</tbody>
</table>
However, if fragments were detected by the observers, they were detected with a high specificity (82.5–98.4%).

**Inter- and intraobserver agreement of radiographic assessment**

Interobserver agreement for the radiographic assessment using two categories (presence or absence of MCD) showed the best results with a “substantial” agreement of $\kappa = 0.74$ between the two most experienced observers, while the interobserver agreement among all other observers was only “slight to moderate” ($\kappa = 0.07–0.42$) (Fig. 1). Interestingly, the “moderate” ($\kappa = 0.42$) agreement was seen between the two least experienced observers. When the interobserver agreement was compared using three categories (presence, absence or suspicion of MCD), the kappa values ($\kappa$) decreased mildly but not significantly for all observers.

The intraobserver agreement was determined to assess the repeatability and consistency of the individual observer in the radiographic interpretation. The two most experienced observers showed “substantial to almost perfect” intraobserver agreement ($\kappa = 0.73–0.88$), while that of the two less experienced observers was “fair to moderate” ($\kappa = 0.31–0.42$) (Fig. 2). When the intraobserver agreement was compared using the three categories (presence, absence or suspicion of MCD), the kappa values ($\kappa$) decreased slightly but not significantly for all observers.

### Discussion and conclusion

The purpose of this study was to evaluate the impact of observer experience on the sensitivity, reliability (interobserver agreement) and repeatability (intraobserver agreement) of radiographic diagnosis of MCD. Sensitivity significantly increased with experience and came close to the sensitivity of CT. Reliability was better between observers with high experience than between observers of lower or different experience level. Repeatability positively correlated with the experience level as well. Intraobserver agreement

### Table 5

<table>
<thead>
<tr>
<th>Assessment categories</th>
<th>Student</th>
<th>Doctoral student</th>
<th>Resident</th>
<th>Diplomate</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%) CL95</td>
<td>30.8 (0.17–0.48)</td>
<td>30.8 (0.17–0.47)</td>
<td>38.5 (0.23–0.55)</td>
<td>28.2 (0.15–0.45)</td>
<td>70.0 (0.53–0.83)</td>
</tr>
<tr>
<td>Specificity (%) CL95</td>
<td>82.5 (0.71–0.91)</td>
<td>92.1 (0.82–0.97)</td>
<td>96.8 (0.89–0.99)</td>
<td>98.4 (0.91–0.99)</td>
<td>60.9 (0.54–0.65)</td>
</tr>
</tbody>
</table>

### Fig. 1 Interobserver agreement for the radiographic diagnosis of medial coronoid disease (MCD). Interobserver agreement was best between the two most experienced observers with substantial agreement, while it was only slight to moderate between the other observers. No significant differences were observed regardless of the use of two categories (MCD present or absent) or three categories (MCD present, suspicious, absent). Observers: 1: student; 2: doctoral student; 3: resident and 4: diplomate. Error bars represent asymptotic standard error.

### Fig. 2 Intraobserver agreement for the radiographic assessment of the diagnosis medial coronoid disease (MCD). A high kappa value in the intraobserver agreement was positively associated with increasing experience. The two most experienced observers showed “substantial to almost perfect” intraobserver agreement, while that of the two less experienced observers was “fair to moderate”. When the intraobserver agreement was compared by using three categories (MCD present, suspicious, absent) the most severe decrease in agreement was seen in the least experienced observer by one category from “slight” to “poor” agreement. Observers: 1: student; 2: doctoral student; 3: resident and 4: diplomate. Error bars represent asymptotic standard error.
was better than interobserver agreement in the case of the most experienced observer.

The median age and weight of the dogs in our study group was similar to other studies (9, 22, 25, 39). Also consistent with others studies (10, 12, 22, 25, 39) male dogs were more commonly affected by MCD than females (ratio 2.3 to 1) and the majority of dogs (78.5%) were bilaterally affected (9, 10, 26, 39). The overrepresentation of Labrador Retrievers, Golden Retrievers and Rottweilers is consistent with many other studies (9, 12, 22, 25).

Arthroscopic findings including fissuring and fragmentation (77.5%) of the MCP were in mildly or moderately higher frequency present than in other studies (10, 25, 35). Subchondral and cartilage pathology without overt fragmentation or fissuring was identified in 12.7% of the cases compared with “chondromalacia-like lesions” and grossly intact MCP with cartilage damage previously reported in 18.2% and 17.6%, respectively (35, 10). The MCP was intact in 9.8% of the cases and was lower compared with the reported 15% in a recent study (25).

In 38.2% of the cases CT raised a suspicion for MCD without detecting signs of fragmentation or fissuring of MCP, which was comparable with the reported incidence of 30% of elbows showing only MCP sclerosis or lucency without fragmentation or fissuring (25). Consistent with the low reported incidence of 1% of elbows showing no pathological changes on CT in that study (25), only 2% of the evaluated elbows in our study showed no abnormalities on CT. Osteophyte formation was detected by CT in 82.5% of the cases. This was comparable with 84% found in the above mentioned study (25).

It has been confirmed that preoperative radiographs are clinically helpful for ruling elbow disease in or out (5), however, similar to other studies (10, 15, 26), fragmentation of the MCP was often not visible on radiographs. In 54.0–81.4% of the elbows the observers could only raise suspicion of MCD due to the lack of visibility of the primary lesion on x-rays, whereas the primary lesion could be visualized in nearly 60% of CT scans. While the sensitivity of radiographic detection of fragments in all observers was low (< 40%), the specificity (82.5–98.4%) was higher than that of CT (60.9%). The high specificity of the most experienced observer (98.4%) indicates that if fragments are detected on radiographs, they can be confirmed in arthroscopy with a high probability in contrast to CT, suggesting overinterpretation of fissure lines or misinterpretation of incompletely mineralized osteophytes associated with the joint capsule as MCP fragments by CT (4, 25).

The sensitivity of radiographic detection of fragments in our study was at 28.2–38.5%, which is more than double than 12.6% reported in a recent study (10). This is most likely due to the fact that we used the Cr15L-CdMO projection instead of the straight craniodorsal projection. The Cr15L-CdMO projection improves the visibility of the MCP by isolating it from other bony structures (15, 39) and most reliably allows for identification of secondary degenerative changes, similar to craniodorsal and mediolateral extended projection (39). Consequently, we used mediolateral extended and flexed as well as Cr15L-CdMO radiographic projections as recommended by the IEWG guidelines for screening ED.

More importantly, the overall sensitivity of radiography to detect pathologic changes of the MCP including fragmentation, fissuring and other subchondral bone and cartilage pathologies reached sensitivities between 77.2–96.7% and increased with the level of training and experience. This is in contrast to previous
studies reporting significantly lower sensitivities ranging between 10–62% (4, 15, 39). However, more recent studies reported high sensitivities for the radiographic detection of arthritis secondary to MCD ranging between 51–89% and 77–96% (3, 9). In those studies, like in our study, the radiographic assessment included besides fragmentation of the MCP and periarticular osteophytosis more subtle changes like TNS. Thus, the sensitivity of detecting MCD was increased, particularly when combined with the other radiographic signs (9).

The absence of osteophytosis in up to 30% of MCD affected dogs in our study and reported by others (10) suggests that screening tests based only on periarticular osteophytosis may fail to identify some dogs with MCD. Moreover, a lack of periarticular osteophytes at 2 years of age does not rule out the presence of MCD (24, 27) as previously suggested (37). In a very recent study (36), 19.6% of dogs older than 6 years diagnosed with MCD lacked osteoarthrosis. In fact, we found that elbows without arthrotic changes or minimal contour changes graded as “borderline” according to the scheme of the IEWG were significantly overrepresented in the group of false negatives.

Like most radiographic changes, TNS often relies on subjective interpretation by the observers rather than on objective measurements. Interobserver agreement for the visual assessment of TNS was found to be only “fair” and was best between diagnostic imaging diplomats only. Intraobserver agreement was “moderate to substantial” in that study. In another previous study (9) interobserver agreement for TNS descriptive assessment were “moderate to substantial” between the two more experienced observers and only “slight” between the least experienced observer and the other observers.

In our study we found a better inter- and intraobserver agreement for the diagnosis of MCD between experienced observers.
We speculate that experience helps to detect subtle pathological radiographic changes, to distinguish between “normal” and “abnormal” and to avoid under- or overinterpretation of the changes. Furthermore, it leads to a greater certainty and repeatability of the diagnosis as it is confirmed by the “substantial to almost perfect” intraobserver agreement of the experienced observers. “Slight to poor” intraobserver agreement of the least experienced observers supports this assumption. Reliability of the diagnosis among the observers tested by interobserver agreement was “substantial” if observers had experience in ED scoring. This may be favoured by the training of the resident by the diplomate over several years.

It is still a subject of debate if objectifying subtle radiological changes like TNS by measurement would improve accuracy, reliability and repeatability of the radiological diagnosis of MCD, especially in less experienced observers. A system to determine the radiopacity of TNS using digital analysis has been published (2). Increase in radiopacity throughout the ulnar trochlear notch region measured by quantitative digital analysis was a radiological finding associated with MCD in Labrador Retrievers. Since we used the same print-outs for each observer to avoid bias, digital analysis of the degree of sclerosis was not possible.

A TNS ratio that is easy to apply in general practice in contrast to x-ray absorptiometry was described in another study (9). However, the authors (9) still recommend using a combination of descriptive and ratio assessment, because the extent of sclerosis (ratio) may not increase to a degree that can be measured, whereas there may be an increase of subjective detectable radiopacity.

To improve the interpretation of more subtle radiographic signs of MCD, in our experience it is very valuable to compare radiographic, CT and arthroscopic findings in cases that are evaluated as MCD free in the radiographic assessment, but turn out to be MCD affected on CT or arthroscopy.

With all of these attempts to increase sensitivity of the radiographic assessment of MCD, it is important to keep in mind that elbows radiographically considered as “normal”, whether subjectively or objectively, can still clinically be affected by MCD (27). In our clinic the contralateral elbow is imaged regardless of whether the dog is clinically affected only unilaterally in order to prevent diseased but clinically “silent” elbows from being missed. However, even if CT shows no abnormalities, a MCP may histologically be considered as “abnormal” (33).

This was a retrospective study and therefore has its limitations. The population was not randomly selected and these cases do not represent all cases presented with elbow disease at our clinic during the study period, because inclusion criteria asked for elbows of dogs that had undergone CT and arthroscopy. We used arthroscopy as gold standard in our study because it allows the direct observation of MCD and there is no definitive “gold standard test” for the diagnosis of MCD. Certainly histopathology would have been optimal as a gold standard test since it has been proven (7) that fissuring in subchondral bone occurs before fibrillation of articular cartilage develops. Collection of tissue at the time of arthroscopy for histopathology would have been ideal, but was unacceptable due to the associated morbidity. It is still subject of debate whether arthroscopy or CT can serve by itself as a gold standard test. A recent comparison (25) of CT with arthroscopy for the assessment of MCD showed that these procedures were complementary for MCD assessment.

The definition of MCD in this study and thus the classification of elbows with cartilage pathology in absence of fragmentation and fissuring as MCD positives may be considered controversial, but has been reported by other authors before (7, 10, 25, 36). It is debatable, if those elbows are more accurately classified as medial coronoid disease (MCD). a) Normal elbow joint: Smooth, regular and normally colored cartilage surface. b) MCD: Displaced fragment (star) and outerbridge grade 5 erosions are visible as racing stripes on the medial humeral condyle (arrow) and MCP (arrow). c) MCD: Non-displaced fragment (star). d) MCD: Avital soft bone and chondromalazia-like lesions: Irregular, soft and discolored cartilage (star) without a fissure or fragment of the MCP. Softening of the subchondral bone of the tip of the MCP was detected by probing. MCP = medial coronoid process of the ulna; R = radius; U = ulna; H = humerus.
Clinical relevance

Radiography is still the primary imaging modality for MCD screening because of its low costs, wide availability and rare need for anaesthesia. However, radiographic diagnosis of MCD is reported to be inferior to advanced diagnostic imaging techniques such as CT and MRI offering superimposition free inspection of the MCP. In contrast to CT, radiographic diagnosis of MCD is based on mainly subjective interpretation of rather subtle primary and secondary changes than that of a large primary lesion. Thus, radiographic diagnosis of MCD is challenging and depends highly on training and experience. When read by specialists, radiographic diagnosis of MCD can reach very high sensitivity, almost comparable to CT, with high reliability and high repeatability. While most of the breeding associations require radiographic ED evaluation by specialists, there is no recommendation existing for the growing individual canine patient. Therefore, radiographic screening for MCD in the individual patient is advised to be performed at least proof-read by trained specialists to increase chance of early detection and therapy.

compartment disease. We are convinced that such cases should be included within the diagnosis of MCD since it is reported that excised coronoid specimens of elbows with comparable changes were permeated by subchondral microcracks and diffuse microdamage in the histomorphometrical examination (10).

Our study population was biased due to the lack of availability of arthroscopies of dogs without MCD. Nevertheless, we determined the specificity of radiographic and CT diagnosis of MCD although we were aware that the low numbers of “MCD-free” elbows in arthroscopy were limiting to the interpretation of specificity. Consequently, our study has a lack of ability to obtain estimates for accuracy and predictive values for radiographic diagnosis of MCD of the single observers.

In the authors’ and others’ opinion (9), sensitivity is more important than specificity in detecting MCD. While false positive results may occur in practice due to inexperience, it is preferable to over diagnose and pursue a further diagnostic step in order to start therapy in a timely manner than to wait for the occurrence of radiographic signs of osteoarthrosis.

Investigator factors may represent a further study limitation. Arthroscopic and CT assessments were performed by a single observer and not by a consensus diagnosis of several observers. This could be a source of inaccuracy and bias since interpretation of arthroscopic and CT findings were at least partly subjective. Further, the different levels of radiographic interpretation experience were represented by only a single observer instead of groups of several observers.

In summary, the results of this study confirm our hypotheses that training and experience are essential to reach high sensitivity, reliability, and repeatability because radiographic diagnosis of MCD is based on mainly subjective interpretation of different assessment criteria. Therefore, radiographic screening for MCD is advised to be performed by specialists experienced in ED screening to ensure optimal outcome.

Conflict of interest

The authors confirm that they do not have any conflict of interest.

References