INTER- AND INTRAOBSERVER AGREEMENT FOR DIAGNOSING PRESUMPTIVE ISCHEMIC MYELOPATHY AND ACUTE NONCOMPRESSIVE NUCLEUS PULPOSUS EXTRUSION IN DOGS USING MAGNETIC RESONANCE IMAGING

Joe Fenn*, Randi Drees, Holger A. Volk, Steven De Decker

Department of Veterinary Clinical Sciences, Royal Veterinary College, Hawkshead Lane, North Mymms, Hertfordshire, AL9 7TA

*Corresponding author: E-mail – jfenn@rvc.ac.uk, Tel: (+44)1707 663666

Keywords: Ischemic myelopathy, acute non-compressive nucleus pulposus extrusion, fibrocartilaginous embolism, MRI, high velocity-low volume

Running Head: Magnetic resonance imaging evaluation of IM and ANNPE

(Results presented as a poster at the 27th Annual Symposium of the European College of Veterinary Neurology, Madrid 2014)

Abstract

Ischemic myelopathy (IM) and acute noncompressive nucleus pulposus extrusion (ANNPE) are common spinal emergencies in dogs with similar clinical presentations. Magnetic resonance imaging (MRI) criteria for a presumptive antemortem diagnosis have been reported, however inter- and intraobserver agreement for use of these criteria has not been established. The aim of this retrospective, descriptive, cross-sectional study was to describe inter- and intraobserver agreement for using previously published MRI criteria to diagnose presumptive IM and ANNPE in a sample dogs. Dogs with a presumptive diagnosis of IM or ANNPE and available MRI scan data were retrieved from medical record archives during the period of 2009 and 2013. A total of 127 dogs were identified. From this sample, MRI scans for 60 dogs were randomly selected and duplicated for intraobserver analysis, giving a total of 187 anonymized studies that were presented to two blinded assessors (one board-certified veterinary neurologist, one board-certified veterinary radiologist). Assessors were asked to diagnose lesions as IM or ANNPE based on previously published MRI characteristics. Interobserver agreement in diagnosing IM or ANNPE was moderate (Kappa = 0.56) and intraobserver agreement was moderate to good (Assessor 1 Kappa = 0.79, Assessor 2 Kappa = 0.47). Agreement was strongest for detecting presence of lesions overlying a vertebral body (94% of lesions that were diagnosed as IM) or overlying an intervertebral disk (85% of lesions that were diagnosed as ANNPE). Findings indicated that use of previously published MRI criteria yields moderate inter- and moderate to good intraobserver agreement for a presumptive diagnosis of IM or ANNPE in dogs.

Introduction

Ischemic myelopathy (IM) and acute non-compressive nucleus pulposus extrusion (ANNPE) are increasingly recognised as a common cause of acute myelopathy in dogs.[1-12] They share a characteristic clinical presentation of a hyperacute onset of non-deteriorating, often markedly lateralising, non-painful paresis or plegia. Typically, clinical signs will occur suddenly following strenuous exercise or traumatic injury, and after an initial short period of deterioration, a static or improving clinical course will follow.[2-4,8] The most commonly identified cause of IM in dogs is a fibrocartilaginous embolism (FCE) within the spinal cord vasculature, of material histologically indistinguishable from the nucleus pulposus.[4,6,12-14] ANNPE on the other hand, represents an acute extrusion of normal, nondegenerate, nucleus pulposus material, causing minimal to no spinal cord compression. It is hypothesized that the spinal cord injury as a consequence of ANNPE differs from IM, in that the impact of an explosive extrusion of nucleus pulposus causes a mainly contusive as opposed to ischemic injury, followed by variable secondary injury and oedema. The ability to achieve an antemortem differentiation between IM and ANNPE is important to allow comparison between clinical characteristics and outcome of these two conditions.

As spinal cord histopathology is unlikely to obtained in a clinical scenario, antemortem diagnosis of IM (or presumed FCE) and ANNPE is based on the presence of the characteristic clinical presentation, in combination with established magnetic resonance (MR) imaging criteria.[2,3,8-10] Previous studies have identified several MR imaging features associated with a diagnosis of IM or ANNPE.[2,3,5,8-10,15] There have been well-defined criteria established to make an MR imaging based diagnosis of ANNPE consisting of; 1) a focal area of intramedullary spinal cord hyperintensity on T2-weighted (T2W) images that overlies an intervertebral disk space, 2) a reduction in volume of the T2W hyperintense nucleus pulposus signal, 3) mild

narrowing of the associated disk space, and 4) extradural material or signal intensity change with minimal or no spinal cord compression at this level.[8] MR imaging diagnosis of IM, or presumed FCE, is based on the presence of a focal, relatively well-demarcated intramedullary T2W hyperintense lesion, mainly affecting grey matter, with an absence of the above criteria used to diagnose ANNPE.[2,3,5]

Although these MR imaging criteria have been shown to help achieve a presumptive diagnosis of IM or ANNPE in dogs,[2,3,8] inter- and intraobserver agreement has not yet been established. As well as supporting the clinical reliability of MR imaging in the differentiation between IM and ANNPE, this is also an important step to support further comparative studies into differences in clinical presentation and outcome. Therefore, aims of this retrospective, descriptive, cross-sectional study were to describe inter- and intraobserver agreement for diagnosing IM and ANNPE in dogs using the previously established MRI characteristics, and to describe agreement for detecting presence or absence of each MRI characteristic. Authors hypothesized that there would be moderate to good interobserver and good intraobserver agreement in differentiating between IM and ANNPE using the established MRI criteria.

Materials and Method

Medical records of the Royal Veterinary College (RVC), University of London were retrospectively reviewed for dogs that underwent MR imaging leading to a presumptive diagnosis of either ANNPE or IM, between November 2009 and December 2013. Electronic clinical records were searched for the diagnoses “ischemic myelopathy”, “fibrocartilaginous embolism”, “acute noncompressive nucleus pulposus extrusion”, “traumatic intervertebral disk extrusion” and “high velocity low volume disk extrusion”. Information retrieved from medical

records included breed, age, gender, clinical history, general physical examination and neurological examination findings.

Inclusion criteria for the study were as follows: dogs with an acute onset myelopathy that was nondeteriorating after 24 hours, and MRI performed at the RVC leading to a diagnosis of either IM or ANNPE. Exclusion criteria were as follows: incomplete or inadequate quality MRI sequences, incomplete clinical history, and concurrent spinal disease (fractures, Hansen Type I disk disease).

Magnetic resonance imaging studies for all dogs were anonymized using image analysis freeware (Osirix Dicom viewer, Osirix Foundation v5.5.2, Geneva, Switzerland) and randomized using a random number generator (Microsoft Excel for Mac 2011 v14, Microsoft Corporation, Redmond, WA). Sixty studies were duplicated and added to the original studies in a randomized order to facilitate analysis of intraobserver agreement. The anonymized MRI studies were then given to one board-certified veterinary neurologist (Assessor 1) and one board-certified veterinary radiologist (Assessor 2) for independent assessment. The assessors were provided with written instructions and specific criteria to use in making a presumptive diagnosis of IM, ANNPE, or "other" for each study (full questionnaire available on request from the corresponding author).

The assessors were first asked to identify and record the vertebral level of the lesion. They were asked to assess the presence of an intramedullary T2W hyperintense lesion (present or not, with intensity defined as compared to normal spinal cord parenchyma), any lateralisation of this lesion (left, right or symmetrical), whether the lesion affected predominantly grey matter, white matter or both, and if the lesion was overlying an intervertebral disk space, vertebral body or both. The length of the lesion as a ratio of lesion length to C6 or L2 vertebral body length (C6 for cervical, L2 for thoracolumbar lesions) was calculated,[2,3,8] as well as the presence or

absence of contrast enhancement on post-contrast T1W sequences, where available. The associated intervertebral disk space was evaluated for the presence or absence of narrowing and a reduction in T2W hyperintense nucleus pulposus volume. The associated intervertebral disk was also evaluated for evidence of disk degeneration, using a previously reported scoring system (0 = homogenous T2W hyperintense nucleus pulposus signal, 1 = heterogenous loss of T2W hyperintense signal, 2 = complete loss of T2W hyperintense signal). The degree of spinal cord compression was also evaluated according to a previously reported scoring system[16] (0 = no compression, 1 = partial ventral subarachnoid space compression, 2 = complete ventral subarachnoid space compression without spinal cord compression, 3 = spinal cord compression with deviation or distortion of the spinal cord), with a score recorded separately for compression caused by intervertebral disk protrusion or extraneous extradural material. The presence or absence of extradural T1W or T2W signal intensity change, extraneous material, and spinal cord swelling were also assessed. For the purpose of this study, spinal cord swelling was assessed subjectively on T2W sagittal images. Any evidence of external trauma such as epaxial or hypaxial muscle signal intensity changes was recorded. Assessors were finally asked to make a diagnosis based on the criteria outlined above of IM, ANNPE or "other". All nominal (such as presence of lateralisation) and ordinal data (such as degree of spinal cord compression) were assigned numerical values for statistical analysis.

All statistical analyses were performed by one of the authors (J.F.), using a commercially available statistical software program, with significance established as $P < 0.05$ (two-sided) where relevant (SPSS Statistics v19, IBM SPSS Inc., Chicago, IL). Interobserver agreement was analyzed using the data obtained from each assessor for all retrieved randomized studies, with intraobserver analysis performed using the results for the 60 duplicated studies. Kappa (κ) analysis was performed for nominal data and weighted κ values obtained for ordinal data with less than possible scores (intervertebral disk degeneration and degree of spinal cord

compression). The strength of agreement was evaluated based on the resulting κ values, with values between 0.81 and 1.00 indicating very good agreement, values of 0.61 to 0.80 indicating good agreement, values of 0.41 to 0.60 indicating moderate agreement, values 0.21 to 0.40 indicating fair agreement and values ≤ 0.20 suggesting poor agreement.[17] A minimum threshold for agreement was established in accordance with previous studies, as a combination of $\kappa > 0.4$ and 75% agreement.[18,19] Prevalence indices were also calculated to characterise population homogeneity, using a previously described method, as this has been reported to influence the interpretation of κ statistics.[18] Agreement for continuous variables (lesion length to C6 or L2 vertebral body length ratio) was evaluated using Bland-Altman analysis,[20,21], with an independent samples t-test used to compare lesion length ratio between IM and ANNPE. Chi square tests were performed to evaluate the relationship between agreement for specific MRI variables and diagnosis.

Results

A total of 127 dogs fulfilled the inclusion criteria, including: 22 Staffordshire Bull Terriers, 19 cross-breed dogs, 16 Labrador Retrievers, 10 Border Collies, 9 Whippets and 30 other breeds (full list of breeds available on request). Age at diagnosis ranged from 0.6 years to 12.4 years (mean ± SD, 6.5 ± 2.6 years), with 81 male dogs (63.8%) and 46 female dogs (36.2%) included.

Magnetic resonance imaging had been performed in all dogs using a 1.5 T unit (Intera 1.5 T, Philips Healthcare, Eindhoven, Netherlands), under general anesthesia. Anesthetic protocols varied on an individual patient basis, as assessed by the attending veterinary anesthetist. T2-weighted and T1-weighted (T1W) fast spin echo (FSE) sequences were obtained in sagittal and transverse planes in all dogs, with T1W FSE postgadolinium (Gadovist 1.0 mmol/ml, Bayer,

Newbury, UK), gradient echo, half Fourier acquisition single shot turbo spin echo and additional plane sequences performed at the request of the attending clinicians. Dogs were positioned in dorsal recumbency. Images for the transverse plane were aligned parallel to the respective intervertebral disks. Slice thickness was 2 mm in the sagittal plane and 2.5–3 mm in the transverse plane. Magnetic resonance imaging included the C1 to T2 vertebrae in dogs with a neuroanatomical localization of C1-C5 or C6-T2 spinal cord segments, and T3 to S3 vertebrae with a neuroanatomical localization between T3-L3 or L4-S3 spinal cord segments.

Kappa statistics revealed moderate interobserver agreement in presumptively differentiating between IM and ANNPE in the 187 studies assessed ($\kappa = 0.56$) with agreement in 77.8% of cases and a prevalence index of 0.56. The MR imaging variables with at least moderate agreement based on Kappa statistics were: whether the lesion was overlying a vertebral body ($\kappa = 0.55$), the presence of lateralisation ($\kappa = 0.53$), the presence of a T2W hyperintense lesion ($\kappa = 0.48$) and the presence of extradural material or signal changes ($\kappa = 0.45$). Two variables satisfied the established threshold for agreement: whether the lesion was overlying a vertebral body ($\kappa = 0.55$, 79.7% agreement) and the presence of a T2W hyperintense lesion ($\kappa = 0.48$, 96.8% agreement). The poorest interobserver agreement was seen in identifying intervertebral disk space narrowing ($\kappa = 0.08$), the presence of contrast enhancement ($\kappa = 0.09$) and the degree of intervertebral disk degeneration ($\kappa = 0.07$). The full list of interobserver agreement statistics is shown in Table 1.

There was moderate to good intraobserver agreement in diagnosis of IM or ANNPE based on the 60 duplicated studies (Overall $\kappa = 0.63$), with agreement in 81.5% of cases and a prevalence index of 0.63. Individual intraobserver agreement in diagnosis was good for Assessor 1 ($\kappa = 0.79$, with 90% agreement and 0.80 prevalence index) and moderate for Assessor 2 ($\kappa = 0.47$, with 73% agreement and 0.46 prevalence index). All MR imaging variables showed at least

moderate overall intraobserver agreement, apart from assessing the presence of spinal cord swelling (κ = 0.29) and grey-white matter lesion distribution (κ = 0.26). Two variables (intervertebral disk space narrowing and epaxial muscle changes) did not yield a Kappa value but had high percentage agreements (83.3% and 82.5%) and prevalence indices of >0.5. Intraobserver agreement statistics are shown in Table 2.

Bland-Altman analysis revealed that the 95% limits of agreement between observers in measuring lesion length to C6 or L2 vertebral length ratio ranged from -1.55 to 0.97, with a mean bias of -0.29. The 95% limits of agreement for Assessor 1 ranged from -1.00 to 0.98, with a mean bias of -0.01, with 95% limits of agreement for Assessor 2 ranging from -1.29 to 1.16, with a mean bias of -0.07 (Figure 2).

There were significant associations between agreement on diagnosis and agreement on the following MR imaging variables: whether the lesion was overlying a vertebral body or an intervertebral disk, lesion lateralisation, reduced nucleus pulposus volume and the presence of extradural material or signal changes ($P < 0.05$ – Table 3). There was also an association between the lesion location (C1-C5, C6-T2, T3-L3 or L4-S3 spinal cord segments) and agreed diagnosis (Table 3). Out of 21 C1-C5 lesions with agreed diagnoses, just 2 (9.5%) were presumptively diagnosed as IM, with the other 19 (94.5%) being presumed ANNPE lesions. Conversely, of 13 C6-T2 and 8 L4-S3 lesions, the majority were diagnosed as presumed IM (76.9% and 75%, respectively). Of the 101 T3-L3 lesions, 53 (52.5%) were presumed IM lesions and 48 (47.5%) were presumed ANNPE lesions.

An independent samples t-test also showed that cases with an agreed diagnosis of IM had significantly greater lesion length ratio (mean = 1.91) than cases with an agreed diagnosis of ANNPE (mean = 0.72, $P < 0.001$). There were no significant associations between agreed diagnosis and any of the other MRI variables evaluated.

Discussion

The results of this study suggest that there is moderate to good intraobserver agreement and moderate interobserver agreement in making a presumptive diagnosis of IM or ANNPE in dogs using established MR imaging criteria. To the authors' knowledge this is the first study to evaluate the reliability of MR imaging in making an ante-mortem differentiation between these two conditions. Definitive histopathological diagnosis was not available in the cases included in this study, similar to most previous studies into MR imaging findings of ANNPE and IM.[2,3,5,8,9] The lack of histopathological diagnosis is representative of the most common clinical scenario, where the diagnosis is most often presumptive based on clinical findings and MRI characteristics. The previously reported imaging criteria for use in the diagnosis of ANNPE (lesion overlying an intervertebral disk, reduced volume of nucleus pulposus, extradural material or signal change and intervertebral disk space narrowing[8]) were among the MR imaging variables with the strongest inter- and intraobserver agreement (Tables 1 and 2). Of these variables, the strongest interobserver agreement was found in identifying extradural material or extradural intensity changes (κ = 0.45, 72.2% agreement) and reduced nucleus pulposus volume (κ = 0.39, 80.8% agreement). The intraobserver agreement for these variables was greater, with the strongest agreement found for identifying reduced nucleus pulposus volume (κ = 0.69, 90.8% agreement).

Assessment of intervertebral disk degeneration and intervertebral disk space narrowing were both associated with poor interobserver agreement (κ = 0.07 and 0.08 respectively, Table 1). This may reflect the difficulty in interpreting these two MR imaging variables, as they are both inherently subjective in nature, despite the defined criteria outlined for the assessors in this study. On MR imaging, intervertebral disk degeneration is associated with reduced T2W signal

intensity, which in some cases may be challenging to differentiate from a reduced nucleus pulposus volume of normal signal intensity.[22] Poor interobserver agreement in assessing intervertebral disk space narrowing is in accordance with a previous study in which both myelography and computed tomography-myelography were found to have superior interobserver agreement compared to MR imaging in detecting disk space narrowing.[23] It was hypothesised in that study that increased variation could arise due to the difficulty in delineating the border between hypointense annulus fibrosus and similarly hypointense vertebral endplates.

Intraobserver agreement in making a presumptive diagnosis of IM or ANNPE was found to be good for Assessor 1 ($\kappa = 0.79$) and moderate for Assessor 2 ($\kappa = 0.47$). This discrepancy between the intraobserver agreement for the two assessors could reflect differences in previous clinical experiences, training background and technique. It is also possible that Assessor 1 reviewed the images over a shorter period of time contributing to more consistency, which is difficult to standardise due to individual clinical duties and time constraints. However, whilst there was a difference between assessors, both achieved at least moderate intraobserver agreement in differentiating between the two conditions. All of the assessed MR imaging variables also demonstrated stronger intraobserver agreement, compared to interobserver agreement (Table 2). This might be expected, as regardless of the variable or imaging modality being assessed, intraobserver agreement is typically greater than interobserver agreement.[23-25] It is also likely that each observer will aim to be particularly consistent in identifying the specific criteria that are used to directly inform their diagnosis. Nonetheless, the level of inter- and intraobserver agreement seen for these four previously established MR imaging criteria supports their use in differentiating between presumptive IM and ANNPE, in dogs with consistent clinical signs.

This study illustrates the advantages, but also the limitations of using Kappa statistics to evaluate agreement. Although it is considered an accurate and useful statistical method to determine inter- and intraobserver agreement, several of the assessed variables in this study demonstrated a high percentage agreement, with a comparatively low Kappa value. For example, in the case of identifying intervertebral disk space narrowing, intraobserver agreement revealed 83.3% agreement despite not yielding a valid Kappa value, with similar discrepancies seen in assessing the presence of muscle changes (interobserver $\kappa = 0.01$, 80.2% agreement) and the presence of spinal cord swelling (intraobserver $\kappa = 0.29$, 81.7% agreement). This is because in a more homogenous population, where for example, observers are unlikely to observe paraspinal muscle signal changes, the probability of agreement by chance is very high. This is demonstrated by the high prevalence indices for these three variables (0.67, 0.60 and 0.63, respectively). Previous reports have shown that as prevalence index (an index of population homogeneity) increases above 0.4, Kappa values will decrease accordingly.[18,19] As a result, low Kappa values should be interpreted in the context of the prevalence index and percentage agreement.

Lesions with an agreed diagnosis of IM in this study were significantly more likely to be agreed to be overlying a vertebral body (94%, $P<0.05$) and to demonstrate no lateralisation (40.3%, $P < 0.05$). In contrast, lesions with an agreed diagnosis of ANNPE were more likely to be agreed to overly an intervertebral disk, with that disk showing a reduced volume of nucleus pulposus, to demonstrate lateralisation of the intramedullary lesion and to be associated with extradural material or extradural signal changes (Table 3). This is to be expected as these include the previously reported imaging criteria used by the assessors to make a diagnosis of IM or ANNPE. However, it is interesting that an agreed diagnosis of ANNPE was significantly more likely to be associated with a lateralised intramedullary lesion on MR imaging (51.3%,

compared to 26.4% of agreed IM lesions; P = 0.002). There have been no previous studies comparing the incidence of MR imaging asymmetry in IM and ANNPE lesions, but in a previous MR imaging study, asymmetry of presumed IM lesions was reported in 31 out of 39 lesions.[3] A study involving only dogs with ANNPE found asymmetry of clinical signs in 26 out of 42 cases, but the prevalence of symmetry of MR imaging lesions was not reported for these cases.[8] A smaller case series of 11 dogs with traumatic intervertebral disk extrusions, presumed to be analogous to ANNPE, found asymmetry in all 11 intramedullary lesions on MR imaging.[9] It has been established that the asymmetrical arterial blood supply to the spinal cord explains the presence of lateralisation in IM, or presumed FCE lesions,[1,4,26] and there may be an anatomical explanation for the degree of lateralisation seen also in ANNPE. It is possible that the dorsal longitudinal ligament provides protection against the explosive extrusion of small volumes of nucleus pulposus seen with ANNPE, making a lateralised extrusion more likely. This finding of an increased tendency for presumed ANNPE lesions to be lateralised in comparison to IM lesions warrants further corroboration in future studies.

Lesion length, as a ratio of lesion length to C6 or L2 vertebral body length, was also shown to be significantly greater in cases with an agreed diagnosis of IM, compared to ANNPE. Lesion length is not one of the previously reported imaging criteria used to differentiate between a diagnosis of IM and ANNPE. The results of ROC curve analysis suggest a fair to good ability for lesion length ratio to differentiate between cases with an agreed presumptive diagnosis of IM or ANNPE (Figure 2). Although a lesion length ratio of 1 was associated with the highest combined sensitivity and specificity to differentiate between presumptive IM and ANNPE, a value of <0.5 was 93% specific for a presumptive ANNPE, but with only 29% sensitivity, and a lesion length ratio <1.9 was 95% sensitive in diagnosing ANNPE, with a relatively low specificity of 40%. A decreased lesion length in ANNPE lesions compared to IM could reflect a difference in spinal cord pathology, with ANNPE potentially associated with a more focal

contusive injury, whereas IM represents an ischemic lesion, with the potential for a more diffuse lesion depending on the nature of the initial embolism or emboli. This could represent a clinically important difference between the two conditions, as lesion length has previously been associated with prognosis in IM,[3] and is worthy of further investigation. These results should however be interpreted in the light of the variation in measurements between observers shown by Bland-Altman analysis (Figure 2). Assessing the relevance of inter- and intraobserver variation in measurements found by Bland-Altman analysis is a question of clinical judgement, as opposed to objective statistical significance. The limits of agreement for lesion length ratio measurements were large enough to raise concerns regarding clinical usefulness, particularly evident at larger lesion length ratio values (Figure 2).

The association found between lesion location and diagnosis in this study has also not been reported before in IM and ANNPE (Table 3). Lesions affecting the C1-C5 region were significantly more likely to be interpreted as ANNPE (90.5%) than IM, whereas lesions affecting the C6-T2 or L4-S3 regions were more likely to be interpreted as IM (75% of C6-T2 and 76.9% of L4-S3 lesions). There was no significant predisposition for either diagnosis in the T3-L3 region, the category into which the majority of lesions fell (70.6% of cases). It has previously been hypothesised that the tendency for ANNPE to occur more commonly near the thoracolumbar junction may be explained by the variation in biomechanical forces that the canine vertebral column is subjected to in this region.[8,30] Whilst earlier studies have reported a higher prevalence of IM and presumed FCE lesions affecting the lumbosacral intumescence, more recent studies demonstrate, in agreement with the results presented here, a higher prevalence of IM lesions affecting the T3-L3 spinal cord segments.[2,4,6] Comparison of the relative lesion distribution of IM and ANNPE cases should also be interpreted with care in the light of the low overall number of lesions affecting the C6-T2 or L4-S3 regions in the current study (n=21, 14.7%).

In conclusion, although there was variability in agreement amongst the individual MR imaging variables assessed, the results of this study support the use of the previously identified and described MR imaging criteria[8] in making a presumptive diagnosis of IM and ANNPE in dogs with consistent clinical signs. However, it is important to recognise that there is a cohort of cases in which differentiating between the two conditions using MR imaging may be challenging (Figure 1). The results of this study also suggest that, alongside the previously reported imaging criteria for differentiating IM and ANNPE, both lesion length ratio and lateralisation could be potentially useful additional indicators. Further studies may be beneficial to identify clinical factors that could aid in differentiating between IM and ANNPE, although these will likely face familiar limitations in terms of the challenge of achieving a final diagnosis without a gold standard ante-mortem diagnostic test or histopathology.

References

1. De Risio L, Platt SR. Fibrocartilaginous embolic myelopathy in small animals. Vet Clin North Am Small Anim Pract 2010;40:859-869.

2. De Risio L, Adams V, Dennis R, et al. Magnetic resonance imaging findings and clinical associations in 52 dogs with suspected ischemic myelopathy. J Vet Intern Med 2007;21:1290-1298.

3. De Risio L, Adams V, Dennis R, et al. Association of clinical and magnetic resonance imaging findings with outcome in dogs suspected to have ischemic myelopathy: 50 cases (2000-2006). J Am Vet Med Assoc 2008;233:129-135.

4. Gandini G, Cizinauskas S, Lang J, et al. Fibrocartilaginous embolism in 75 dogs: clinical findings and factors influencing the recovery rate. J Small Anim Pract 2003;44:76-80.

5. Abramson CJ, Garosi L, Platt SR, et al. Magnetic resonance imaging appearance of suspected ischemic myelopathy in dogs. Vet Radiol Ultrasound 2005;46:225-229.

6. Cauzinille L, Kornegay JN. Fibrocartilaginous embolism of the spinal cord in dogs: review of 36 histologically confirmed cases and retrospective study of 26 suspected cases. J Vet Intern Med 1996;10:241-245.

7. Cauzinille L. Fibrocartilaginous embolism in dogs. Vet Clin North Am Small Anim Pract 2000;30:155-167, vii.

8. De Risio L, Adams V, Dennis R, et al. Association of clinical and magnetic resonance imaging findings with outcome in dogs with presumptive acute noncompressive nucleus pulposus extrusion: 42 cases (2000-2007). J Am Vet Med Assoc 2009;234:495-504.

9. Chang Y, Dennis R, Platt SR, et al. Magnetic resonance imaging of traumatic intervertebral disc extrusion in dogs. Vet Rec 2007;160:795-799.

10. Henke D, Gorgas D, Flegel T, et al. Magnetic resonance imaging findings in dogs with traumatic intervertebral disk extrusion with or without spinal cord compression: 31 cases (2006-2010). J Am Vet Med Assoc 2013;242:217-222.

11. Griffiths IR. A syndrome produced by dorso-lateral "explosions" of the cervical intervertebral discs. Vet Rec 1970;87:737-741.

12. Griffiths IR. Spinal cord infarction due to emboli arising from the intervertebral discs in the dog. J Comp Pathol 1973;83:225-232.

13. Gill CW. Case report: fibrocartilaginous embolic myelopathy in a dog. Can Vet J 1979;20:273-278.

14. Gilmore D, De Lahunta A. Necrotizing myelopathy secondary presumed or confirmed fibrocartilaginous embolism in 24 dogs. J Am Anim Hosp Assoc 1986;23:373-376.

15. Lu D, Lamb CR, Wesselingh K, et al. Acute intervertebral disc extrusion in a cat: clinical and MRI findings. J Feline Med Surg 2002;4:65-68.

16. Grunenfelder FI, Weishaupt D, Green R, et al. Magnetic resonance imaging findings in spinal cord infarction in three small breed dogs. Vet Radiol Ultrasound 2005;46:91-96.

17. da Costa RC, Parent JM, Partlow G, et al. Morphologic and morphometric magnetic resonance imaging features of Doberman Pinschers with and without clinical signs of cervical spondylomyelopathy. Am J Vet Res 2006;67:1601-1612.

18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-174.

19. Burn CC, Weir AA. Using prevalence indices to aid interpretation and comparison of agreement ratings between two or more observers. Vet J 2011;188:166-170.

20. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85:257-268.

21. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. Ultrasound Obstet Gynecol 2003;22:85-93.

22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-310.

23. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.

24. Obuchowski NA. Receiver operating characteristic curves and their use in radiology. Radiology 2003;229:3-8.

25. Levitski RE, Lipsitz D, Chauvet AE. Magnetic resonance imaging of the cervical spine in 27 dogs. Vet Radiol Ultrasound 1999;40:332-341.

26. De Decker S, Gielen IM, Duchateau L, et al. Intraobserver, interobserver, and intermethod agreement for results of myelography, computed tomography-myelography, and low-field magnetic resonance imaging in dogs with disk-associated wobbler syndrome. J Am Vet Med Assoc 2011;238:1601-1608.

27. De Decker S, Gielen IM, Duchateau L, et al. Intraobserver and interobserver agreement for results of low-field magnetic resonance imaging in dogs with and without clinical signs of disk-associated wobbler syndrome. J Am Vet Med Assoc 2011;238:74-80.

413 28. De Decker S, Gielen IM, Duchateau L, et al. Agreement and repeatability of linear
414 vertebral body and canal measurements using computed tomography (CT) and low field
415 magnetic resonance imaging (MRI). Vet Surg 2010;39:28-34.

416 29. Caulkins SE, Purinton PT, Oliver JE, Jr. Arterial supply to the spinal cord of dogs and
417 cats. Am J Vet Res 1989;50:425-430.

418 30. Shores A. Fractures and Luxations of the Vertebral Column. Veterinary Clinics of
419 North America-Small Animal Practice 1992;22:171-180.

420

421

422

423

424

425

426

427

428

429

430

431

Tables

Table 1 – Results of interobserver Kappa statistics for evaluation of 187 MR imaging studies by two assessors.

MR Imaging Variable	% Agreement	Prevalence Index	Kappa
IM or ANNPE	**77.83†**	**0.56**	**0.56†**
Overlying vertebral body	79.68†	0.59	0.55†
Lateralisation	68.98	0.38	0.53*
T2W hyperintensity	96.79†	0.94	0.48†
Extradural material / signal changes	72.19	0.44	0.45*
T2W reduced volume of NP	80.75	0.61	0.39
Overlying IVD	74.87	0.50	0.33
Grey Vs white matter	71.66	0.43	0.19
Spinal cord swelling	68.98	0.38	0.18
Contrast enhancement	58.82	0.18	0.09
Narrowed disc space	44.92	0.10	0.08
Muscle changes	80.21	0.60	0.01
IVD degeneration‡	33.69	0.33	0.07
SC compression (IVD)‡	21.93	0.56	-
SC compression (material)‡	55.61	0.11	0.24

IVD = Intervertebral disc, NP = nucleus pulposus, SC = Spinal cord, * = Moderate agreement, † = Exceed minimum threshold for agreement, ‡ = Weighted Kappa analysis

Table 2 – Results of intraobserver Kappa statistics for evaluation of 60 MR imaging studies by two assessors.

MR Imaging Variable	OVERALL			ASSESSOR 1			ASSESSOR 2		
	%	P.I	Kappa	%	P.I	Kappa	%	P.I	Kappa
IM or ANNPE	**81.51†**	**0.63**	**0.63†**	**90.00†**	**0.80**	**0.79†**	**72.88**	**0.46**	**0.47***
T2W reduced volume of NP	90.83†	0.82	0.69†	93.33†	0.87	0.82†	88.33†	0.77	0.56†
Overlying vertebral body	86.67†	0.73	0.65†	93.33†	0.87	0.86†	80.00†	0.60	0.43†
Contrast enhancement	95.00†	0.90	0.60†	100.00†	1.00	1.00†	90.00	0.80	0.20
Overlying IVD	85.00†	0.70	0.59†	88.33†	0.77	0.60†	81.66†	0.63	0.58†
Lateralised	73.33	0.47	0.55*	75.00†	0.50	0.61†	71.67	0.43	0.49*
T2W hyperintensity	98.33†	0.97	0.49†	100†	1.00	1.00†	96.67	0.93	-0.02
Extradural material / signal changes	72.50	0.45	0.46*	76.66†	0.53	0.54†	68.33	0.37	0.37
Spinal cord swelling	81.67	0.63	0.29	73.33	0.47	0.37	90.00	0.80	0.20
Grey Vs white matter	76.67	0.53	0.26	68.33	0.37	0.18	85.00	0.70	0.34
Narrowed disc space	83.33	0.67	-	71.67	0.43	0.42*	95.00	0.90	-
Muscle changes	82.50	0.65	-	66.67	0.33	0.16	98.33	0.97	-
IVD degeneration‡	85.00†	0.70	0.60†	88.33†	0.77	0.80†	81.66†	0.63	0.40†
SC compression (IVD)‡	66.67	0.33	0.41*	80.00†	0.60	0.69†	53.33	0.07	0.13
SC compression (material)‡	74.17	0.48	0.57*	70.00†	0.40	0.62†	78.33†	0.57	0.51†

IVD = Intervertebral disc, NP = nucleus pulposus, SC = Spinal cord, P.I = Prevalence index, % = % of agreement, * = Moderate-good agreement, † = Exceed minimum threshold for agreement, ‡ = Weighted Kappa analysis

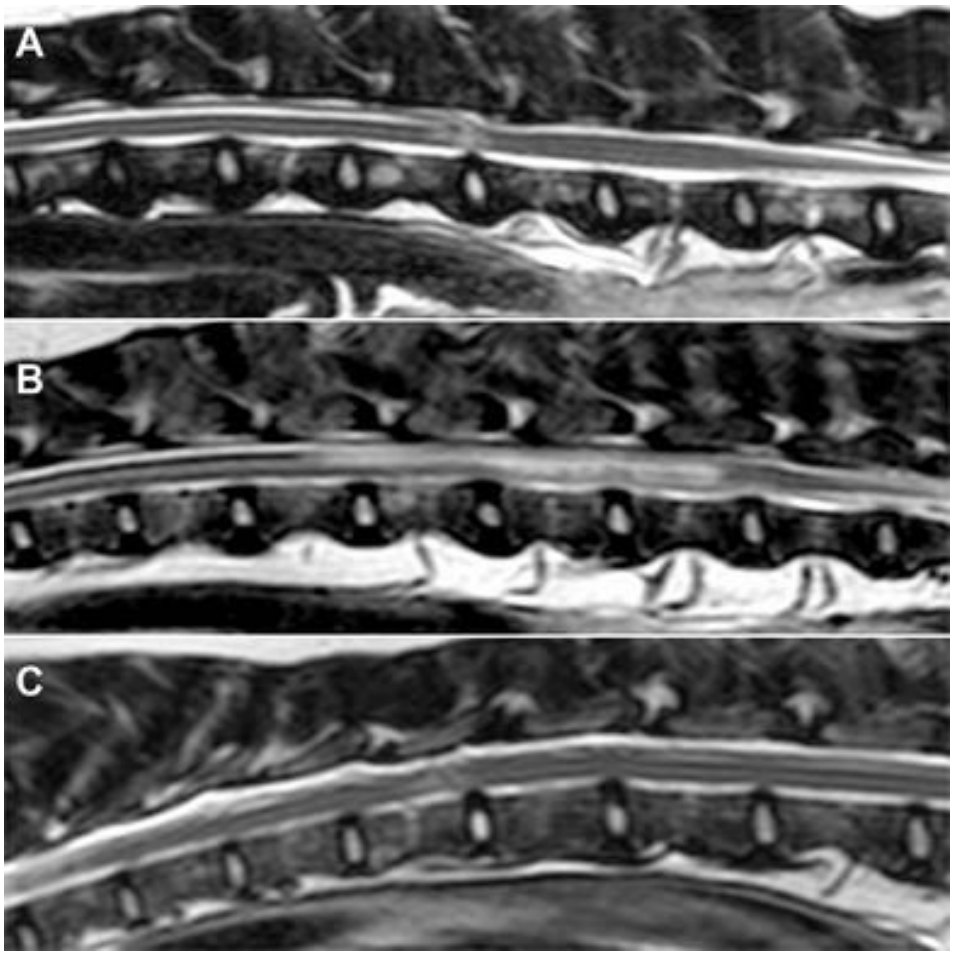
Table 3 – Chi-square analysis showing significant associations between agreement in lesion characteristics and diagnosis in 144 MR imaging studies with an agreed diagnosis of IM or ANNPE.

MR Imaging Variable		Agreed IM cases (n=72)	Agreed ANNPE cases (n=72)	χ^2	P
Lesion overlying vertebral body	**Yes (%)**	68 (94.4)	15 (20.8)	71.08	<0.001
	No (%)	0 (0)	38 (52.8)		
Lesion overlying IVD	**Yes (%)**	38 (52.8)	61 (84.7)	13.41	<0.001
	No (%)	14 (19.4)	2 (2.8)		
Lateralised lesion	**Left (%)**	13 (18.1)	23 (31.9)	9.32	0.002
	Right (%)	6 (8.3)	14 (19.4)		
	Neither (%)	29 (40.3)	16 (22.2)		
Reduced volume of NP	**Yes (%)**	17 (23.6)	65 (90.3)	65.53	<0.001
	No (%)	36 (50.0)	0 (0)		
Extradural material / signal change	**Yes (%)**	7 (9.7)	40 (55.6)	43.37	<0.001
	No (%)	50 (69.4)	14 (19.4)		
Vertebral level of lesion	**C1-C5 (%)**	2 (2.8)	19 (26.4)	19.77	<0.001
	C6-T2 (%)	10 (14.1)	3 (4.2)		
	T3-L3 (%)	53 (74.6)	48 (47.5)		
	L4-S3 (%)	6 (8.5)	2 (25.0)		

IVD = intervertebral disk, NP = nucleus pulposus

Figure legends

Figure 1 – Sagittal T2-weighted MR images of the thoracolumbar spine representing examples of an agreed diagnosis of ANNPE (A), an agreed diagnosis of IM (B), and a case with interobserver disagreement in diagnosis (C).



1. De Risio L, Platt SR. Fibrocartilaginous embolic myelopathy in small animals. *Vet Clin North Am Small Anim Pract* 2010;40:859-869.
2. De Risio L, Adams V, Dennis R, et al. Magnetic resonance imaging findings and clinical associations in 52 dogs with suspected ischemic myelopathy. *J Vet Intern Med* 2007;21:1290-1298.
3. De Risio L, Adams V, Dennis R, et al. Association of clinical and magnetic resonance imaging findings with outcome in dogs suspected to have ischemic myelopathy: 50 cases (2000-2006). *J Am Vet Med Assoc* 2008;233:129-135.
4. Gandini G, Cizinauskas S, Lang J, et al. Fibrocartilaginous embolism in 75 dogs: clinical findings and factors influencing the recovery rate. *J Small Anim Pract* 2003;44:76-80.
5. Abramson CJ, Garosi L, Platt SR, et al. Magnetic resonance imaging appearance of suspected ischemic myelopathy in dogs. *Vet Radiol Ultrasound* 2005;46:225-229.
6. Cauzinille L, Kornegay JN. Fibrocartilaginous embolism of the spinal cord in dogs: review of 36 histologically confirmed cases and retrospective study of 26 suspected cases. *J Vet Intern Med* 1996;10:241-245.
7. Cauzinille L. Fibrocartilaginous embolism in dogs. *Vet Clin North Am Small Anim Pract* 2000;30:155-167, vii.
8. De Risio L, Adams V, Dennis R, et al. Association of clinical and magnetic resonance imaging findings with outcome in dogs with presumptive acute noncompressive nucleus pulposus extrusion: 42 cases (2000-2007). *J Am Vet Med Assoc* 2009;234:495-504.
9. Chang Y, Dennis R, Platt SR, et al. Magnetic resonance imaging of traumatic intervertebral disc extrusion in dogs. *Vet Rec* 2007;160:795-799.
10. Henke D, Gorgas D, Flegel T, et al. Magnetic resonance imaging findings in dogs with traumatic intervertebral disk extrusion with or without spinal cord compression: 31 cases (2006-2010). *J Am Vet Med Assoc* 2013;242:217-222.
11. Griffiths IR. A syndrome produced by dorso-lateral "explosions" of the cervical intervertebral discs. *Vet Rec* 1970;87:737-741.
12. Griffiths IR. Spinal cord infarction due to emboli arising from the intervertebral discs in the dog. *J Comp Pathol* 1973;83:225-232.
13. Gill CW. Case report: fibrocartilaginous embolic myelopathy in a dog. *Can Vet J* 1979;20:273-278.
14. Gilmore D, De Lahunta A. Necrotizing myelopathy secondary presumed or confirmed fibrocartilaginous embolism in 24 dogs. *J Am Anim Hosp Assoc* 1986;23:373-376.
15. Grunenfelder FI, Weishaupt D, Green R, et al. Magnetic resonance imaging findings in spinal cord infarction in three small breed dogs. *Vet Radiol Ultrasound* 2005;46:91-96.
16. da Costa RC, Parent JM, Partlow G, et al. Morphologic and morphometric magnetic resonance imaging features of Doberman Pinschers with and without clinical signs of cervical spondylomyelopathy. *Am J Vet Res* 2006;67:1601-1612.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174.
18. Burn CC, Weir AA. Using prevalence indices to aid interpretation and comparison of agreement ratings between two or more observers. *Vet J* 2011;188:166-170.
19. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005;85:257-268.
20. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. *Ultrasound Obstet Gynecol* 2003;22:85-93.

21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-310.
22. Levitski RE, Lipsitz D, Chauvet AE. Magnetic resonance imaging of the cervical spine in 27 dogs. *Vet Radiol Ultrasound* 1999;40:332-341.
23. De Decker S, Gielen IM, Duchateau L, et al. Intraobserver, interobserver, and intermethod agreement for results of myelography, computed tomography-myelography, and low-field magnetic resonance imaging in dogs with disk-associated wobbler syndrome. *J Am Vet Med Assoc* 2011;238:1601-1608.
24. De Decker S, Gielen IM, Duchateau L, et al. Intraobserver and interobserver agreement for results of low-field magnetic resonance imaging in dogs with and without clinical signs of disk-associated wobbler syndrome. *J Am Vet Med Assoc* 2011;238:74-80.
25. De Decker S, Gielen IM, Duchateau L, et al. Agreement and repeatability of linear vertebral body and canal measurements using computed tomography (CT) and low field magnetic resonance imaging (MRI). *Vet Surg* 2010;39:28-34.
26. Caulkins SE, Purinton PT, Oliver JE, Jr. Arterial supply to the spinal cord of dogs and cats. *Am J Vet Res* 1989;50:425-430.