Evaluation of Magnetic Resonance Imaging Guidelines for Differentiation Between Thoracolumbar Intervertebral Disk Extrusions And Intervertebral Disk Protrusions

Steven De Decker; Sergio A Gomes; Rowena MA Packer; Patrick J Kenny; Elsa Beltran; Birgit Parzefall; Joe Fenn; Devi Nair; George Nye; Holger A Volk

Department of Veterinary Clinical Science and Services, Royal Veterinary College, University of London, Hawkshead lane, AL9 7TA North Mymms, Hatfield, England

Dr. Gomes’ current address is: The Queen’s Veterinary School Hospital, University of Cambridge, CB3 OES, Cambridge, England

Running head: MRI guidelines for intervertebral disk disease

Keywords: intervertebral disk herniation, accuracy, reliability, intervertebral disk disease, agreement

Name, address, and e-mail address of the corresponding author:
Steven De Decker, sdedecker@rvc.ac.uk, Department of Veterinary Clinical Science and Services, Royal Veterinary College, University of London, North Mymms, Hertfordshire, AL97TA, England

Funding Sources: None

The results of this study have been presented in abstract form (Poster) for the 28th

Acknowledgments: None
Abstract

Four MRI variables have recently been suggested to be independently associated with a diagnosis of thoracolumbar intervertebral disk extrusion or protrusion. Midline intervertebral disk herniation, and partial intervertebral disk degeneration were associated with intervertebral disk protrusion, while presence of a single intervertebral disk herniation and disk material dispersed beyond the boundaries of the intervertebral disk space were associated with intervertebral disk extrusion. The aim of this retrospective, cross sectional study was to evaluate if using these MRI variables improves differentiation between thoracolumbar intervertebral disk extrusions and protrusions. Eighty large breed dogs with surgically confirmed thoracolumbar intervertebral disk extrusions or protrusions were included. Randomized MRI studies were presented on two occasions to six blinded observers, which were divided into three experience categories. During the first assessment, observers made a presumptive diagnosis of thoracolumbar intervertebral disk extrusion or protrusion without guidelines. During the second assessment they were asked to make a presumptive diagnosis with the aid of guidelines. Agreement was evaluated by Kappa-statistics. Diagnostic accuracy significantly improved from 70.8% to 79.6% and inter-observer agreement for making a diagnosis of intervertebral disk extrusion or intervertebral disk protrusion improved from fair (κ = 0.27) to moderate (κ = 0.41) after using the proposed guidelines. Diagnostic accuracy was significantly influenced by degree of observer experience. Intra-observer agreement for the assessed variables ranged from fair to excellent and inter-observer agreement ranged from fair to moderate. The results of this study suggest that the proposed imaging guidelines can aid in differentiating thoracolumbar intervertebral disk extrusions from protrusions.
Two types of degenerative thoracolumbar intervertebral disk disease have been recognized; intervertebral disk extrusion or Hansen Type-I, and intervertebral disk protrusion or Hansen Type-II intervertebral disk disease.1-3 Both types of intervertebral disk disease are associated with different pathological and clinical characteristics.4,5 Thoracolumbar intervertebral disk extrusion is characterized by a sudden herniation of degenerated and calcified nucleus pulposus through a fully ruptured anulus fibrosus into the vertebral canal.1,5,6 Affected dogs can be young and present often with an acute onset of clinical signs, which can vary from spinal hyperesthesia to paraplegia with loss of nociception.4,5 Intervertebral disk protrusion is characterized by slowly progressive and focal extension of the anulus fibrosus and dorsal longitudinal ligament into the vertebral canal.1,5-8 Affected animals are typically older, present with a more insidious clinical history and milder clinical signs, such as ambulatory paresis and ataxia of the pelvic limbs.4,5,9,10 Apart from the above mentioned differences in clinical presentation, thoracolumbar intervertebral disk extrusion and protrusion are also associated with different suggested surgical techniques11-15, postoperative complications, and possibly also a different prognosis.4 Although outcomes after medical and surgical treatment have been extensively reported for dogs with extrusions, this information is only scarce for dogs with thoracolumbar intervertebral disk protrusions. It seems therefore important to accurately differentiate extrusions from protrusions before treatment options and associated outcomes are discussed with owners of affected dogs. Although magnetic resonance imaging (MRI) is considered the imaging modality of choice to diagnose intervertebral disk disease in dogs 16, it is currently unknown how well this technique can be used to differentiate between both types of intervertebral disk herniation. A recent study9 identified four MRI variables, which were suggested to be
independently associated with a diagnosis of thoracolumbar intervertebral disk extrusion or protrusion in large breed dogs (Figure 1): (1) midline instead of lateralized intervertebral disk herniation, and (2) partial instead of complete intervertebral disk degeneration were associated with a diagnosis of intervertebral disk protrusion, while (3) the presence of a single instead of multiple intervertebral disk herniations and (4) dispersed intervertebral disk material beyond the borders of the intervertebral disk space were associated with a diagnosis of intervertebral disk extrusion.  

Before these MRI variables can be suggested as diagnostic guidelines, it is however important to evaluate if applying them would be clinically useful and reliable. In other words, it should be assessed how well thoracolumbar intervertebral disk extrusion can be differentiated from intervertebral disk protrusion without using these guidelines, if using these guidelines improves indeed accuracy and reliability to differentiate both types of intervertebral disk herniation, and if assessment of the proposed MRI variables is associated with acceptable intra–and inter-observer agreements. The aim of this retrospective, cross sectional study was therefore to evaluate if implementation of the above mentioned MRI variables would facilitate obtaining an accurate and reliable diagnosis of thoracolumbar intervertebral disk extrusion or protrusion. An additional aim was to evaluate how levels of experience would influence MRI assessments with and without the proposed imaging guidelines. It was hypothesized that application of the proposed MRI guidelines would improve diagnostic accuracy and reliability of diagnosing thoracolumbar intervertebral disk extrusions and protrusions, that the proposed variables could be assessed with acceptable intra and interobserver agreements, and that observers with limited MRI experience would benefit the strongest from using the proposed MRI guidelines.
Methods

Included dogs were presented to the Royal Veterinary College (RVC), University of London, between July 2002 and January 2014 for further assessment of a suspected spinal condition. The studied animals were also included in a previous study, evaluating MRI characteristics of 95 large breed dogs with surgically confirmed thoracolumbar intervertebral disk extrusion (n=52) or intervertebral disk protrusion (n=43). Of these 95 animals, 40 dogs with thoracolumbar intervertebral disk extrusion and 40 dogs with intervertebral disk protrusion were randomly selected using a random number generator. This selection aimed to obtain an equal number of dogs in each group. To assess intraobserver agreement, the MRI studies of 20 dogs of each group were randomly selected and duplicated. A total of 80 dogs and 120 MRI studies were therefore included (Figure 2). All dogs included in this study received general physical and complete neurological examinations. Dogs were included if (1) they were large breed dogs, defined as a body weight exceeding 20kg, (2) underwent an MRI study of the thoracolumbar or lumbar vertebral column, (3) following a diagnosis of intervertebral disk herniation underwent spinal surgery consisting of a hemilaminectomy or hemilaminectomy combined with a partial discectomy and (4) the type of intervertebral disk herniation (extrusion or protrusion) was clearly noted in the surgical reports. Dogs were excluded if the medical records or imaging studies were incomplete, if they were not available in a digital format, or if the type of intervertebral disk herniation (extrusion or protrusion) was not clearly noted in the surgical reports, if more than one type of intervertebral disk herniation (both extrusion and protrusion present), or acute herniations of flaps of anulus were observed during surgery. Information retrieved from the medical records included signalment, duration, type, and severity of clinical signs, general physical and neurological
examinations findings and type of surgery with surgical confirmation of the type of
intervertebral disk herniation (extrusion or protrusion). Severity of neurological
deficits was graded by the modified Frankel score, which was defined as paraplegia
with no deep nociception (grade 0), paraplegia with no superficial nociception (grade
1), paraplegia with nociception (grade 2), non-ambulatory paraparesis (grade 3),
ambulatory paraparesis and ataxia (grade 4), spinal hyperesthesia only (grade 5), or
no dysfunction. 17

All dogs underwent MRI under general anesthesia. MRI was performed with a
1.5T magnet (Intera, Philips Medical Systems, Eindhoven, the Netherlands). The
selection of variables was based on the results of a previous study evaluating MRI
characteristics of thoracolumbar intervertebral disk disease in large breed dogs and
were suggested to be independently associated with a diagnosis of intervertebral disk
extrusion or intervertebral disk protrusion. 9 Assessed variables included (1)
lateralization of herniated disk material, (2) degree of intervertebral disk
degeneration, (3) the presence of multiple intervertebral disk herniations, and (4)
location of herniated disk material relative to the intervertebral disk space (Figure 1).
Lateralization of herniated intervertebral disk material was assessed on transverse
images and was described as being exclusively in the midline ventral to the spinal
cord or lateralized. Assessment of intervertebral disk degeneration was based on
nucleus pulposus signal intensity on sagittal T2-weighted images. A non-degenerate
intervertebral disk (grade 0) had a homogenous hyperintense signal, a partially
degenerate intervertebral disk (grade 1) had heterogeneous loss of hyperintense
signal, and a completely degenerate intervertebral disk (grade 2) had complete loss of
hyperintense signal. 18-20 Presence of a single or multiple intervertebral disk
herniations was evaluated on sagittal and transverse T2-weighted images. Location of herniated disk material relative to the affected intervertebral disk space was assessed on T1–and T2-weighted sagittal images and was described as dispersed or confined to the intervertebral disk space. Dispersed intervertebral disk material was defined as intervertebral disk material beyond the borders of the affected intervertebral disk space and associated vertebral endplates. Disk material confined to the intervertebral disk space was defined as herniated disk material not exceeding the limits of the intervertebral disk space or associated vertebral endplates.

The 120 MRI studies were presented twice in a randomized order to six blinded observers (Figure 2). The sequence of MRI studies and the identity of duplicate studies differed among observers. The observers were not informed about the clinical history, type of clinical signs, the number of dogs with thoracolumbar intervertebral disk extrusions or protrusions, were not informed about the inclusion of duplicate MRI studies, were not aware of the results of our previous study, and were also not involved in assessment of MRI studies for our previous study. Standard image archiving and communication system software (Osirix Foundation, V.5.5.2 Geneva, Switzerland) was used to view the imaging studies. During the first round of assessments, the observers were asked for each MRI study to make a presumptive diagnosis of thoracolumbar intervertebral disk extrusion or protrusion without any guidelines. They were however informed about the location of the affected and surgically confirmed intervertebral disk space. After returning their answers, they received again 120 MRI studies (different sequence and different identity of duplicates) with at least a 1-month interval between the two rounds of assessments. During the second round of assessments the observers were again asked for each MRI study to make a presumptive diagnosis of intervertebral disk extrusion or protrusion.
This time they were however provided with the proposed MRI guidelines. Additionally, they were asked to record the presence or absence of each of the proposed MRI variables (Figure 2). They were again informed about the location of the affected intervertebral disk space. To evaluate the role of experience, the six observers were divided into three groups of different experience levels; expert, moderate, and no experience. Two board-certified neurologists represented the group of expert observers (PJK and EB). Two final year residents in neurology represented the group of observers with moderate experience (BP and JF). The group of observers without experience was represented by two veterinary surgeons completing a small animal rotating internship (DN and GN). Because of their lack of experience and unfamiliarity with the imaging software, the two observers without experience received a training session by the first author of this study. During this session, they learnt how to use the imaging viewing software, were instructed about the pathophysiological differences between intervertebral disk extrusions and protrusions, and were handed a recent review manuscript discussing intervertebral disk disease in chondrodystrophic and non-chondrodystrophic dog breeds. After the first round of assessments they received illustrated guidelines with examples of the proposed MRI variables. While these training sessions were deemed necessary to facilitate study enrollment of observers without experience, it was considered unnecessary for the more experienced observers.

Wilcoxon signed ranks tests were used to detect the influence of diagnostic guidelines on diagnostic accuracy and agreement. Chi-squared tests were used to detect differences in accuracy between the three experience groups, and the effect of intervertebral disk herniation type (intervertebral disk extrusion vs. protrusion) on diagnostic accuracy with and without use of guidelines. These statistics were carried
out in IBM SPSS Statistics v21 (IBM SPSS Statistics, New York) by one of the
authors (RMAP). Agreement statistics were calculated using Minitab v17 (Minitab
version 17, Pennsylvania). Fleiss’ kappa (κ) for more than two observers was
calculated for diagnosis and each assessed MRI variable. The strength of agreement
was interpreted on the basis of the κ values suggested by Altman, as adapted from
the method of Landis and Koch: κ-values of 0.81 – 1.00 indicated very good
agreement; 0.61 – 0.80, good agreement; 0.41 – 0.60, moderate agreement; 0.21 –
0.40, fair agreement; and 0.20 or lower, poor agreement. As a guide, the minimum
threshold for κ is often arbitrarily set at κ ≤ 0.4, below which variables are considered
unreliable. Calculation of interobserver agreement and accuracy was based on the
assessment of the 80 original MRI studies. Calculation of intraobserver agreement
was based on assessment of 40 duplicate MRI studies (Figure 2). Results were
considered significant if \( P < 0.05 \).

**Results**

A total of 80 large breed dogs with surgically confirmed thoracolumbar
intervertebral disk extrusion (n=40) or protrusion (40) were included in this study.
The group of dogs with intervertebral disk extrusion included German Shepherd Dogs
(n=10), Staffordshire Bull Terriers (five), Clumber Spaniels (four), Labrador
Retrievers (four), Basset Hounds (three), Rottweilers (three); there were six breeds
represented by one dog each and five crossbreeds. This group included 21 males and
19 females aged between three and 12 years (mean, 7.2 years). Duration of clinical
signs ranged from 12 hours to six months (median, three days) and affected dogs
presented with neurological grades 0 (n=three dogs), 1 (one), 2 (10), 3 (nine), and 4
(17). Affected intervertebral disk spaces in order of occurrence were T12-T13, L1-L2
The group of dogs with thoracolumbar intervertebral disk protrusion included German Shepherd Dogs (n=20), English Staffordshire Bull terriers (eight), Basset Hounds (two), Labrador Retrievers (two); there were four breeds represented by one dog each and four crossbreeds. This group included 32 males and eight females aged between four and 12.2 years (mean, 8.6 years). Duration of clinical signs ranged from 12 hours to two years (median, 29 days) and affected dogs presented with neurological grades 2 (n=one dog), 3 (seven), and 4 (32). Affected intervertebral disk spaces in order of occurrence were T13-L1 (n=17), T12-T13, L1-L2 (both nine) and L2-L3 (five).

MRI studies were performed with dogs in dorsal recumbency, using a dedicated spinal coil. Imaging studies included a minimum of T2-weighted (repetition time (ms) (TR)/echo time (ms) (TE); 3000/120) and T1-weighted (TR/TE, 400/8) sagittal and transverse images. Slice thickness for sagittal and transverse images were respectively 1.75 and 2.5mm with an interslice gap of 0.3mm in both planes. The transverse images were aligned perpendicular to the spinal cord. Diagnostic accuracy was calculated using the assessments of the 80 original MRI studies (Table 1). Overall diagnostic accuracy without using the proposed guidelines for the six observers combined was 70.8% (340/480) and ranged from 45% (36/80) to 88.8% (71/80) for individual observers (Table 1). There was a significant influence of observer experience with observers without experience being least accurate (58%) and observers with moderate experience being most accurate (78%) ($X^2 = 18.78$, $P < 0.001$). There was no significant difference between the number of accurately diagnosed thoracolumbar intervertebral disk extrusions or protrusions before guidelines were included (74% vs. 75%, respectively, $P > 0.05$). Overall diagnostic
accuracy for the six observers combined improved significantly \((P = 0.028)\) from 70.8% to 79.6% \((382/480)\) after using the proposed diagnostic guidelines and ranged from 52.5% \((42/80)\) to 92.5% \((74/80)\) after using the MRI guidelines. Additionally, all six observers diagnosed a higher number of thoracolumbar intervertebral disk extrusions or protrusions correctly after using the guidelines (mean improvement; 13.4%). There was again a significant influence of observer experience \((X^2 = 31.72, P < 0.001)\) with observers without experience being least accurate (65%) and observers with expert experience being most accurate (88%). After using the guidelines thoracolumbar intervertebral disk protrusions (84%) were significantly more often \((P = 0.024)\) correctly diagnosed than intervertebral disk extrusions (75%) \((X^2 = 5.13, P = 0.02)\).

Intraobserver agreement was calculated using the assessments of the 40 duplicate MRI studies (Table 2). Intraobserver agreement for making a diagnosis of thoracolumbar intervertebral disk extrusion or protrusion ranged from 0.4 to 0.95 before, and from 0.51 to 0.85 after applying the proposed guidelines. This difference was not significant and there was no significant influence of observer experience. Although a high degree of variability was seen among individual observers, most variables were associated with moderate or good intraobserver agreement (Table 2). Assessing if multiple intervertebral disk herniations were present was associated with the highest intraobserver agreement \((\kappa = 0.43 – 0.79)\) and assessing if an intervertebral disk herniation was midline or lateralized was associated with the lowest intraobserver agreement \((\kappa = 0.33 – 0.75)\). There was no significant influence of level of experience on intraobserver agreement of the assessed MRI variables.
Interobserver agreement was calculated using the assessments of the 80 original MRI studies. Overall interobserver agreement for making a diagnosis of thoracolumbar intervertebral disk extrusion or protrusion significantly improved \((P < 0.001)\) from fair \((\kappa = 0.27)\) to moderate \((\kappa = 0.41)\) after using the proposed guidelines. Assessing if an intervertebral disk herniation was midline or lateralized and if and intervertebral disk herniation was confined to or exceeded the borders of the affected intervertebral disk space were associated with moderate interobserver agreement \((\kappa = 0.43)\). Assessing the presence of partial intervertebral disk degeneration \((\kappa = 0.35)\), complete disk degeneration \((\kappa = 0.40)\), and multiple intervertebral disk herniations \((\kappa = 0.35)\) were associated with fair interobserver agreement.

**Discussion**

This study evaluated the accuracy, reliability and reproducibility of MRI guidelines to improve the differentiation of thoracolumbar intervertebral disk extrusion and disk protrusion in large breed dogs. The results of this study suggest that applying the proposed guidelines improves the accuracy and inter-observer agreement of diagnosing thoracolumbar intervertebral disk extrusion and intervertebral disk protrusion. The assessed MRI variables were associated with fair to excellent intraobserver and fair to moderate interobserver agreements. The results of this study indicate further that accuracy of assessing spinal MRI studies is influenced by the observer’s degree of training and experience.

Differentiating thoracolumbar intervertebral disk extrusions from protrusions is clinically important. Although several studies have characterized the clinical presentation and outcomes after medical or surgical treatment for dogs with
thoracolumbar intervertebral disk extrusion \textsuperscript{26-30}, this information is less well characterized for dogs with intervertebral disk protrusion. While surgical treatment of thoracolumbar intervertebral disk extrusion typically consists of a hemilaminectomy with fenestration of the affected intervertebral disk \textsuperscript{28-30}, it is suggested that a lateral corpectomy or vertebral stabilization should be considered for dogs with intervertebral disk protrusion.\textsuperscript{11-15} While the former can be considered a basic spinal surgical technique, the latter are probably more technically demanding. Furthermore, little is known about results of medical management in dogs with thoracolumbar intervertebral disk protrusions\textsuperscript{4}, it has been suggested that dogs with thoracolumbar intervertebral disk extrusion have a better prognosis after surgery than dogs with intervertebral disk protrusion \textsuperscript{4}, and that dogs with thoracolumbar intervertebral disk protrusion are at increased risk of early postoperative neurological deterioration.\textsuperscript{4,31} This illustrates that reaching an accurate and reliable imaging diagnosis of thoracolumbar intervertebral disk extrusion or protrusion is not only important for surgical planning, but also for managing the expectations of owners and clinical staff. The overall accuracy to differentiate thoracolumbar intervertebral disk extrusion from intervertebral disk protrusion improved significantly from 70.8 to 79.6\%, and the interobserver agreement for making a diagnosis improved significantly from fair to moderate after applying the proposed MRI guidelines. Although these findings indicate clinical usefulness, a proportion of cases were still diagnosed incorrectly after using the proposed guidelines. This suggests that MRI in isolation cannot be completely relied on to differentiate between thoracolumbar extrusions and protrusions and confirms that imaging findings should always be interpreted in light of a thorough clinical history and results of a complete clinical examination.\textsuperscript{32} In combination to the evaluated MRI variables, a clinical variable has been suggested to
be independently associated with a diagnosis of thoracolumbar intervertebral disk extrusion or protrusion. Increased duration of clinical signs, with a threshold value of 21 days, was significantly associated with a diagnosis of intervertebral disk protrusion. The observers in this study were however unaware of the clinical history and clinical signs of included dogs. It is therefore possible that better diagnostic accuracy and reliability would have been reached if this information had been available.

Intra- and interobserver agreement for the assessed MRI variables was evaluated with kappa statistics. This is a useful and widespread statistical technique to evaluate reliability and reproducibility. Obtained kappa coefficients are commonly translated into poor, fair, moderate, good or very good agreement. Although this can provide useful information, results of kappa statistics do not directly determine if a certain variable can be assessed with acceptable reliability and reproducibility. Although such a decision is dependent on the specific variable and clinical context, studies in musculoskeletal research have suggested that kappa coefficients lower than 0.4 should be considered clinically unacceptable. When applying this threshold value to the study presented here, intraobserver agreement for the assessed MRI variables could be considered clinically acceptable. Only one observer did not reach intraobserver agreement coefficients ≥ 0.4 for all assessed MRI variables. Interobserver agreement for the assessed variables ranged from fair to moderate with 3 of the 5 assessed variables reaching the arbitrary threshold value (≥ 0.4) for acceptable agreement. Presence of multiple intervertebral disk herniations and degree of intervertebral disk degeneration were associated with only fair interobserver agreement. The lower interobserver agreement for the presence of multiple
intervertebral disk herniations is in agreement with a previous study evaluating the number of spinal cord compressions in dogs with disk-associated cervical spondylomyelopathy. In the current study, assessment of multiple intervertebral disk herniations was among the variables with the highest intraobserver agreement, indicating consistent rating within an individual observer. Different degrees of intervertebral disk herniation can occur in clinically normal dogs and it is not always possible to clearly differentiate clinically relevant from irrelevant intervertebral disk herniations. It is therefore possible that different observers will have different, but individually consistent, criteria for evaluating if an intervertebral disk herniation is present or not. This variable is therefore likely associated with inherent subjectivity and it should be considered to develop a more objectively definition of intervertebral disk herniation. Although a grading system based on accepted terminology with easily recognizable characteristics was used, evaluation of degree of intervertebral disk degeneration was also associated with lower values for interobserver agreement. This finding is however in contrast with several human and veterinary studies. A previous veterinary study demonstrated good agreement between experienced observers for rating intervertebral disk degeneration using the same criteria as this study. Although other reasons cannot be excluded, it is possible that the inclusion of observers with varying degrees of training and experience, including observers without experience, contributed to the lower value for interobserver agreement in this study.

The influence of observer experience was further highlighted by the fact that ‘low observer experience’ negatively influenced the diagnostic accuracy of the assessed MRI studies. Even after using the diagnostic guidelines, the group of observers
without experience had significantly lower accuracy for diagnosing thoracolumbar intervertebral disk extrusion and intervertebral disk protrusion. This finding is in agreement with several human studies and indicates that assessment and clinical interpretation of MRI studies should be preserved for people with sufficient experience and training. It is therefore possible that the overall results of this study are influenced by the inclusion of observers with different experience levels and it can be debated if our results therefore appropriately reflect referral practice. For the purpose of this study, observers with different experience categories were included to evaluate if differentiation between thoracolumbar intervertebral disk extrusion and disk protrusion would be more problematic for observers with little experience, if using the proposed guidelines would be easy and objective enough to be used by novices, and if using these imaging guidelines could compensate for lack of training and experience.

A potential limitation of this study was the limited number of observers in each ‘experience category’. Although care was taken to select observers of similar experience levels. In addition, there was no standardization across observers in terms of how the assessments were performed (i.e. all in one sitting or a few at a time) or the monitor on which they were viewed. Another limitation was that all patients had surgically confirmed intervertebral disk extrusion or protrusion. The combination of imaging and clinical findings was therefore severe enough to warrant surgical intervention. These factors could have influenced the assessment of the observers; however surgical confirmation was necessary to provide a “golden standard” diagnosis with which to assess diagnostic accuracy. It should further be emphasized that MRI studies were evaluated by general clinicians or veterinary surgeons trained in veterinary neurology and neurosurgery. It can however not be excluded that
interpretation of MRI studies by observers with a different training background, such as veterinary radiologists, would have influenced our results. Veterinary neurologists were however preferred in this study, because in our institution, these clinicians perform MRI interpretation and spinal surgery on a daily basis and were therefore considered to have most experience in taking clinical decisions after assessing spinal MRI studies. Although only limited data is available on direct comparisons between veterinary neurosurgeons and radiologists for assessing spinal MRI studies \(^{34,44,45}\), human studies indicate acceptable agreement between specialities.\(^{46-49}\) Although comparing the accuracy and agreement between veterinary neurologists and radiologists was considered beyond the aims of this study, further studies are necessary to not only evaluate the role of experience, but also the type of training on the reliability and accuracy of evaluating spinal MRI studies.

In summary, the results of this study suggest that differentiation between thoracolumbar intervertebral disk extrusion and protrusion becomes more accurate and reliable after applying the proposed MRI guidelines, that the proposed MRI variables can be assessed with, in general, clinically acceptable agreement, and that diagnostic guidelines cannot replace thorough clinical training and experience. Further studies are necessary to evaluate the influence of observer experience and availability of clinical history and clinical signs on the evaluation of spinal MRI studies.

**Acknowledgments:** None
References


34. De Decker S, Gielen IM, Duchateau L, et al. Intraobserver, interobserver, and intermethod agreement for results of myelography, computed tomography-


Legends

Figure 1. Proposed MRI guidelines to differentiate between thoracolumbar intervertebral disk extrusion (IVDE) and intervertebral disk protrusion (IVDP) in large breed dogs.

<table>
<thead>
<tr>
<th>MRI variable</th>
<th>Intervertebral disk Extrusion</th>
<th>Intervertebral disk Protrusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lateralized intervertebral disk herniation</td>
<td>More likely</td>
<td>Less likely</td>
</tr>
<tr>
<td>Partial instead of complete intervertebral disk degeneration</td>
<td>Less likely</td>
<td>More likely</td>
</tr>
<tr>
<td>Multiple intervertebral disk herniations</td>
<td>Less likely</td>
<td>More likely</td>
</tr>
<tr>
<td>Herniated disk material confined to intervertebral disk space</td>
<td>Less likely</td>
<td>More likely</td>
</tr>
</tbody>
</table>
**Figure 2.** Flow-chart illustrating different MRI assessments in order to calculate the diagnostic accuracy, inter-observer agreement and intra-observer agreement for differentiating between thoracolumbar intervertebral disk extrusions and protrusions with and without the use of the proposed imaging guidelines.

<table>
<thead>
<tr>
<th>First round of assessments: 120 MRI studies <strong>without</strong> guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 original MRI studies</td>
</tr>
<tr>
<td>40 randomly selected duplicate MRI studies</td>
</tr>
<tr>
<td>Make presumptive diagnosis of IVDE or IVDP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Second round of assessments: 120 MRI studies <strong>with</strong> guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 original MRI studies</td>
</tr>
<tr>
<td>40 different randomly selected duplicate MRI studies</td>
</tr>
<tr>
<td>Make presumptive diagnosis of IVDE or IVDP</td>
</tr>
<tr>
<td>Record presence or absence of each proposed MRI variable</td>
</tr>
<tr>
<td>Calculate diagnostic accuracy and inter-observer agreement with and without guidelines based on assessments of 80 original MRI studies</td>
</tr>
<tr>
<td>Calculate intraobserver agreement with and without guidelines based on assessments of 40 duplicate MRI studies</td>
</tr>
</tbody>
</table>
Table 1. Number of correctly diagnosed thoracolumbar intervertebral disk extrusions or protrusions before and after using the proposed MR imaging guidelines. Calculations were based on 80 original MRI studies.

<table>
<thead>
<tr>
<th>Observer</th>
<th>Degree of experience</th>
<th>Correct diagnoses without guidelines (out of 80)</th>
<th>Correct diagnoses with guidelines (out of 80)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1</td>
<td>Expert</td>
<td>69 (86.2%)</td>
<td>73 (91.2%)</td>
</tr>
<tr>
<td>Observer 2</td>
<td>Moderate</td>
<td>53 (66.2%)</td>
<td>63 (78.8%)</td>
</tr>
<tr>
<td>Observer 3</td>
<td>None</td>
<td>36 (45%)</td>
<td>42 (52.5%)</td>
</tr>
<tr>
<td>Observer 4</td>
<td>Expert</td>
<td>54 (67.5%)</td>
<td>68 (85%)</td>
</tr>
<tr>
<td>Observer 5</td>
<td>Moderate</td>
<td>71 (88.8%)</td>
<td>74 (92.5%)</td>
</tr>
<tr>
<td>Observer 6</td>
<td>None</td>
<td>57 (71.2%)</td>
<td>62 (77.5%)</td>
</tr>
<tr>
<td>Overall</td>
<td>NA</td>
<td><strong>340 (70.8%)</strong></td>
<td><strong>382 (79.6%)</strong></td>
</tr>
</tbody>
</table>
Intraobserver agreement $\kappa$-values for making a diagnosis of thoracolumbar intervertebral disk extrusion or protrusion and for assessment of each of the proposed MRI variables. Calculations were based on 40 duplicate MRI studies.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observer 1- Expert</th>
<th>Observer 2- Moderate experience</th>
<th>Observer 3- No experience</th>
<th>Observer 4- Expert</th>
<th>Observer 5- Moderate experience</th>
<th>Observer 6 – No experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis without guidelines</td>
<td>0.75</td>
<td>0.81</td>
<td>0.42</td>
<td>0.80</td>
<td>0.95</td>
<td>0.40</td>
</tr>
<tr>
<td>Diagnosis with guidelines</td>
<td>0.75</td>
<td>0.85</td>
<td>0.51</td>
<td>0.59</td>
<td>0.70</td>
<td>0.75</td>
</tr>
<tr>
<td>IVD herniation midline?</td>
<td>0.75</td>
<td>0.75</td>
<td>0.33</td>
<td>0.69</td>
<td>0.55</td>
<td>0.45</td>
</tr>
<tr>
<td>Multiple IVD herniations present?</td>
<td>0.79</td>
<td>0.73</td>
<td>0.46</td>
<td>0.43</td>
<td>0.85</td>
<td>0.75</td>
</tr>
<tr>
<td>IVD herniation confined to IVDS?</td>
<td>0.95</td>
<td>0.89</td>
<td>-0.14</td>
<td>0.84</td>
<td>0.69</td>
<td>0.56</td>
</tr>
<tr>
<td>Partial IVD degeneration present?</td>
<td>0.68</td>
<td>0.85</td>
<td>0.55</td>
<td>0.65</td>
<td>0.49</td>
<td>0.47</td>
</tr>
<tr>
<td>Complete IVD degeneration present?</td>
<td>0.68</td>
<td>0.90</td>
<td>0.44</td>
<td>0.62</td>
<td>0.49</td>
<td>0.75</td>
</tr>
</tbody>
</table>

IVD = intervertebral disk, IVDS = intervertebral disk space on and for assessment of each of the proposed MRI variables. Calculations were based on 80 original MRI studies.