

RVC OPEN ACCESS REPOSITORY – COPYRIGHT NOTICE

This is the peer reviewed version of the following article:

De Decker, S., Gomes, S. A., Packer, R. M., Kenny, P. J., Beltran, E., Parzefall, B., Fenn, J., Nair, D., Nye, G. and Volk, H. A. (2016), EVALUATION OF MAGNETIC RESONANCE IMAGING GUIDELINES FOR DIFFERENTIATION BETWEEN THORACOLUMBAR INTERVERTEBRAL DISK EXTRUSIONS AND INTERVERTEBRAL DISK PROTRUSIONS IN DOGS. *Veterinary Radiology & Ultrasound*. doi: 10.1111/vru.12394

which has been published in final form at <http://dx.doi.org/10.1111/vru.12394>.

This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#).

The full details of the published version of the article are as follows:

TITLE: EVALUATION OF MAGNETIC RESONANCE IMAGING GUIDELINES FOR DIFFERENTIATION BETWEEN THORACOLUMBAR INTERVERTEBRAL DISK EXTRUSIONS AND INTERVERTEBRAL DISK PROTRUSIONS IN DOGS

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

JOURNAL TITLE: *Veterinary Radiology & Ultrasound*

PUBLISHER: Wiley

PUBLICATION DATE: 4 July 2016 (online)

DOI: 10.1111/vru.12394

1 **Evaluation of Magnetic Resonance Imaging Guidelines for Differentiation**
2 **Between Thoracolumbar Intervertebral Disk Extrusions And Intervertebral**
3 **Disk Protrusions**

4

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

7

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

10

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

13

14 Running head: MRI guidelines for intervertebral disk disease

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

17

18

19 Name, address, and e-mail address of the corresponding author:

20 Steven De Decker, sdedecker@rvc.ac.uk, Department of Veterinary Clinical Science
21 and Services, Royal Veterinary College, University of London, North Mymms,
22 Hertfordshire, AL97TA, England

23

24 Funding Sources: None

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

26 Annual Symposium of the European College of Veterinary Neurology – European
27 Society of Veterinary Neurology, 18-19 September 2015, Amsterdam, The
28 Netherlands.
29
30 Acknowledgments: None
31
32
33

34 **Abstract**

35 Four MRI variables have recently been suggested to be independently associated with
36 a diagnosis of thoracolumbar intervertebral disk extrusion or protrusion. Midline
37 intervertebral disk herniation, and partial intervertebral disk degeneration were
38 associated with intervertebral disk protrusion, while presence of a single intervertebral
39 disk herniation and disk material dispersed beyond the boundaries of the
40 intervertebral disk space were associated with intervertebral disk extrusion. The aim
41 of this retrospective, cross sectional study was to evaluate if using these MRI
42 variables improves differentiation between thoracolumbar intervertebral disk
43 extrusions and protrusions. Eighty large breed dogs with surgically confirmed
44 thoracolumbar intervertebral disk extrusions or protrusions were included.
45 Randomized MRI studies were presented on two occasions to six blinded observers,
46 which were divided into three experience categories. During the first assessment,
47 observers made a presumptive diagnosis of thoracolumbar intervertebral disk
48 extrusion or protrusion without guidelines. During the second assessment they were
49 asked to make a presumptive diagnosis with the aid of guidelines. Agreement was
50 evaluated by Kappa-statistics. Diagnostic accuracy significantly improved from
51 70.8% to 79.6% and inter-observer agreement for making a diagnosis of intervertebral
52 disk extrusion or intervertebral disk protrusion improved from fair ($\kappa = 0.27$) to
53 moderate ($\kappa = 0.41$) after using the proposed guidelines. Diagnostic accuracy was
54 significantly influenced by degree of observer experience. Intra-observer agreement
55 for the assessed variables ranged from fair to excellent and inter-observer agreement
56 ranged from fair to moderate. The results of this study suggest that the proposed
57 imaging guidelines can aid in differentiating thoracolumbar intervertebral disk
58 extrusions from protrusions.

59

60 Two types of degenerative thoracolumbar intervertebral disk disease have been
61 recognized; intervertebral disk extrusion or Hansen Type-I, and intervertebral disk
62 protrusion or Hansen Type-II intervertebral disk disease.¹⁻³ Both types of
63 intervertebral disk disease are associated with different pathological and clinical
64 characteristics.^{4,5} Thoracolumbar intervertebral disk extrusion is characterized by a
65 sudden herniation of degenerated and calcified nucleus pulposus through a fully
66 ruptured annulus fibrosus into the vertebral canal.^{1,5,6} Affected dogs can be young and
67 present often with an acute onset of clinical signs, which can vary from spinal
68 hyperesthesia to paraplegia with loss of nociception.^{4,5} Intervertebral disk protrusion
69 is characterized by slowly progressive and focal extension of the annulus fibrosus and
70 dorsal longitudinal ligament into the vertebral canal.^{1,5-8} Affected animals are
71 typically older, present with a more insidious clinical history and milder clinical
72 signs, such as ambulatory paresis and ataxia of the pelvic limbs.^{4,5,9,10} Apart from the
73 above mentioned differences in clinical presentation, thoracolumbar intervertebral
74 disk extrusion and protrusion are also associated with different suggested surgical
75 techniques¹¹⁻¹⁵, postoperative complications, and possibly also a different prognosis.⁴
76 Although outcomes after medical and surgical treatment have been extensively
77 reported for dogs with extrusions, this information is only scarce for dogs with
78 thoracolumbar intervertebral disk protrusions. It seems therefore important to
79 accurately differentiate extrusions from protrusions before treatment options and
80 associated outcomes are discussed with owners of affected dogs. Although magnetic
81 resonance imaging (MRI) is considered the imaging modality of choice to diagnose
82 intervertebral disk disease in dogs¹⁶, it is currently unknown how well this technique
83 can be used to differentiate between both types of intervertebral disk herniation. A
84 recent study⁹ identified four MRI variables, which were suggested to be

85 independently associated with a diagnosis of thoracolumbar intervertebral disk
86 extrusion or protrusion in large breed dogs (Figure 1): (1) midline instead of
87 lateralized intervertebral disk herniation, and (2) partial instead of complete
88 intervertebral disk degeneration were associated with a diagnosis of intervertebral
89 disk protrusion, while (3) the presence of a single instead of multiple intervertebral
90 disk herniations and (4) dispersed intervertebral disk material beyond the borders of
91 the intervertebral disk space were associated with a diagnosis of intervertebral disk
92 extrusion.⁹ Before these MRI variables can be suggested as diagnostic guidelines, it is
93 however important to evaluate if applying them would be clinically useful and
94 reliable. In other words, it should be assessed how well thoracolumbar intervertebral
95 disk extrusion can be differentiated from intervertebral disk protrusion without using
96 these guidelines, if using these guidelines improves indeed accuracy and reliability to
97 differentiate both types of intervertebral disk herniation, and if assessment of the
98 proposed MRI variables is associated with acceptable intra –and inter-observer
99 agreements. The aim of this retrospective, cross sectional study was therefore to
100 evaluate if implementation of the above mentioned MRI variables would facilitate
101 obtaining an accurate and reliable diagnosis of thoracolumbar intervertebral disk
102 extrusion or protrusion. An additional aim was to evaluate how levels of experience
103 would influence MRI assessments with and without the proposed imaging guidelines.
104 It was hypothesized that application of the proposed MRI guidelines would improve
105 diagnostic accuracy and reliability of diagnosing thoracolumbar intervertebral disk
106 extrusions and protrusions, that the proposed variables could be assessed with
107 acceptable intra and interobserver agreements, and that observers with limited MRI
108 experience would benefit the strongest from using the proposed MRI guidelines.
109

110 **Methods**

111 Included dogs were presented to the Royal Veterinary College (RVC),
112 University of London, between July 2002 and January 2014 for further assessment of
113 a suspected spinal condition. The studied animals were also included in a previous
114 study, evaluating MRI characteristics of 95 large breed dogs with surgically
115 confirmed thoracolumbar intervertebral disk extrusion (n=52) or intervertebral disk
116 protrusion (n=43).⁹ Of these 95 animals, 40 dogs with thoracolumbar intervertebral
117 disk extrusion and 40 dogs with intervertebral disk protrusion were randomly selected
118 using a random number generator. This selection aimed to obtain an equal number of
119 dogs in each group. To assess intraobserver agreement, the MRI studies of 20 dogs of
120 each group were randomly selected and duplicated. A total of 80 dogs and 120 MRI
121 studies were therefore included (Figure 2). All dogs included in this study received
122 general physical and complete neurological examinations. Dogs were included if (1)
123 they were large breed dogs, defined as a body weight exceeding 20kg⁴, (2) underwent
124 an MRI study of the thoracolumbar or lumbar vertebral column, (3) following a
125 diagnosis of intervertebral disk herniation underwent spinal surgery consisting of a
126 hemilaminectomy or hemilaminectomy combined with a partial discectomy and (4)
127 the type of intervertebral disk herniation (extrusion or protrusion) was clearly noted in
128 the surgical reports. Dogs were excluded if the medical records or imaging studies
129 were incomplete, if they were not available in a digital format, or if the type of
130 intervertebral disk herniation (extrusion or protrusion) was not clearly noted in the
131 surgical reports, if more than one type of intervertebral disk herniation (both extrusion
132 and protrusion present), or acute herniations of flaps of anulus were observed during
133 surgery. Information retrieved from the medical records included signalment,
134 duration, type, and severity of clinical signs, general physical and neurological

135 examinations findings and type of surgery with surgical confirmation of the type of
136 intervertebral disk herniation (extrusion or protrusion). Severity of neurological
137 deficits was graded by the modified Frankel score, which was defined as paraplegia
138 with no deep nociception (grade 0), paraplegia with no superficial nociception (grade
139 1), paraplegia with nociception (grade 2), non-ambulatory paraparesis (grade 3),
140 ambulatory paraparesis and ataxia (grade 4), spinal hyperesthesia only (grade 5), or
141 no dysfunction.¹⁷

142

143 All dogs underwent MRI under general anesthesia. MRI was performed with a
144 1.5T magnet (Intera, Philips Medical Systems, Eindhoven, the Netherlands). The
145 selection of variables was based on the results of a previous study evaluating MRI
146 characteristics of thoracolumbar intervertebral disk disease in large breed dogs and
147 were suggested to be independently associated with a diagnosis of intervertebral disk
148 extrusion or intervertebral disk protrusion.⁹ Assessed variables included (1)
149 lateralization of herniated disk material, (2) degree of intervertebral disk
150 degeneration, (3) the presence of multiple intervertebral disk herniations, and (4)
151 location of herniated disk material relative to the intervertebral disk space (Figure 1).
152 Lateralization of herniated intervertebral disk material was assessed on transverse
153 images and was described as being exclusively in the midline ventral to the spinal
154 cord or lateralized. Assessment of intervertebral disk degeneration was based on
155 nucleus pulposus signal intensity on sagittal T2-weighted images. A non-degenerate
156 intervertebral disk (grade 0) had a homogenous hyperintense signal, a partially
157 degenerate intervertebral disk (grade 1) had heterogeneous loss of hyperintense
158 signal, and a completely degenerate intervertebral disk (grade 2) had complete loss of
159 hyperintense signal.¹⁸⁻²⁰ Presence of a single or multiple intervertebral disk

160 herniations was evaluated on sagittal and transverse T2-weighted images. Location of
161 herniated disk material relative to the affected intervertebral disk space was assessed
162 on T1 –and T2-weighted sagittal images and was described as dispersed or confined
163 to the intervertebral disk space. Dispersed intervertebral disk material was defined as
164 intervertebral disk material beyond the borders of the affected intervertebral disk
165 space and associated vertebral endplates.²¹ Disk material confined to the intervertebral
166 disk space was defined as herniated disk material not exceeding the limits of the
167 intervertebral disk space or associated vertebral endplates.²¹

168 The 120 MRI studies were presented twice in a randomized order to six
169 blinded observers (Figure 2). The sequence of MRI studies and the identity of
170 duplicate studies differed among observers. The observers were not informed about
171 the clinical history, type of clinical signs, the number of dogs with thoracolumbar
172 intervertebral disk extrusions or protrusions, were not informed about the inclusion of
173 duplicate MRI studies, were not aware of the results of our previous study⁹, and were
174 also not involved in assessment of MRI studies for our previous study⁹. Standard
175 image archiving and communication system software (Osirix Foundation, V.5.5.2
176 Geneva, Switzerland) was used to view the imaging studies. During the first round of
177 assessments, the observers were asked for each MRI study to make a presumptive
178 diagnosis of thoracolumbar intervertebral disk extrusion or protrusion without any
179 guidelines. They were however informed about the location of the affected and
180 surgically confirmed intervertebral disk space. After returning their answers, they
181 received again 120 MRI studies (different sequence and different identity of
182 duplicates) with at least a 1-month interval between the two rounds of assessments.
183 During the second round of assessments the observers were again asked for each MRI
184 study to make a presumptive diagnosis of intervertebral disk extrusion or protrusion.

185 This time they were however provided with the proposed MRI guidelines.
186 Additionally, they were asked to record the presence or absence of each of the
187 proposed MRI variables (Figure 2). They were again informed about the location of
188 the affected intervertebral disk space. To evaluate the role of experience, the six
189 observers were divided into three groups of different experience levels; expert,
190 moderate, and no experience. Two board-certified neurologists represented the group
191 of expert observers (PJK and EB). Two final year residents in neurology represented
192 the group of observers with moderate experience (BP and JF). The group of observers
193 without experience was represented by two veterinary surgeons completing a small
194 animal rotating internship (DN and GN). Because of their lack of experience and
195 unfamiliarity with the imaging software, the two observers without experience
196 received a training session by the first author of this study. During this session, they
197 learnt how to use the imaging viewing software, were instructed about the
198 pathophysiological differences between intervertebral disk extrusions and protrusions,
199 and were handed a recent review manuscript discussing intervertebral disk disease in
200 chondrodystrophic and non-chondrodystrophic dog breeds.⁵ After the first round of
201 assessments they received illustrated guidelines with examples of the proposed MRI
202 variables. While these training sessions were deemed necessary to facilitate study
203 enrollment of observers without experience, it was considered unnecessary for the
204 more experienced observers.

205 Wilcoxon signed ranks tests were used to detect the influence of diagnostic
206 guidelines on diagnostic accuracy and agreement. Chi-squared tests were used to
207 detect differences in accuracy between the three experience groups, and the effect of
208 intervertebral disk herniation type (intervertebral disk extrusion vs. protrusion) on
209 diagnostic accuracy with and without use of guidelines. These statistics were carried

210 out in IBM SPSS Statistics v21 (IBM SPSS Statistics, New York) by one of the
211 authors (RMAP). Agreement statistics were calculated using Minitab v17 (Minitab
212 version 17, Pennsylvania). Fleiss' kappa (κ) for more than two observers was
213 calculated²² for diagnosis and each assessed MRI variable. The strength of agreement
214 was interpreted on the basis of the κ values suggested by Altman²³, as adapted from
215 the method of Landis and Koch²⁴: κ -values of 0.81 – 1.00 indicated very good
216 agreement; 0.61 – 0.80, good agreement; 0.41 – 0.60, moderate agreement; 0.21 –
217 0.40, fair agreement; and 0.20 or lower, poor agreement. As a guide, the minimum
218 threshold for κ is often arbitrarily set at $\kappa \leq 0.4$, below which variables are considered
219 unreliable.²⁵ Calculation of interobserver agreement and accuracy was based on the
220 assessment of the 80 original MRI studies. Calculation of intraobserver agreement
221 was based on assessment of 40 duplicate MRI studies (Figure 2). Results were
222 considered significant if $P < 0.05$.

223

224 **Results**

225 A total of 80 large breed dogs with surgically confirmed thoracolumbar
226 intervertebral disk extrusion (n=40) or protrusion (40) were included in this study.
227 The group of dogs with intervertebral disk extrusion included German Shepherd Dogs
228 (n=10), Staffordshire Bull Terriers (five), Clumber Spaniels (four), Labrador
229 Retrievers (four), Basset Hounds (three), Rottweilers (three); there were six breeds
230 represented by one dog each and five crossbreeds. This group included 21 males and
231 19 females aged between three and 12 years (mean, 7.2 years). Duration of clinical
232 signs ranged from 12 hours to six months (median, three days) and affected dogs
233 presented with neurological grades 0 (n=three dogs), 1 (one), 2 (10), 3 (nine), and 4
234 (17). Affected intervertebral disk spaces in order of occurrence were T12-T13, L1-L2

235 (both n=nine), T13-L1 (seven), L2-L3 (six), L3-L4 (five), T11-T12 (two), T3-T4 and
236 L4-L5 (both one). The group of dogs with thoracolumbar intervertebral disk
237 protrusion included German Shepherd Dogs (n=20), English Staffordshire Bull
238 terriers (eight), Basset Hounds (two), Labrador Retrievers (two); there were four
239 breeds represented by one dog each and four crossbreeds. This group included 32
240 males and eight females aged between four and 12.2 years (mean, 8.6 years). Duration
241 of clinical signs ranged from 12 hours to two years (median, 29 days) and affected
242 dogs presented with neurological grades 2 (n=one dog), 3 (seven), and 4 (32).
243 Affected intervertebral disk spaces in order of occurrence were T13-L1 (n=17), T12-
244 T13, L1-L2 (both nine) and L2-L3 (five).

245 MRI studies were performed with dogs in dorsal recumbency, using a
246 dedicated spinal coil. Imaging studies included a minimum of T2-weighted (repetition
247 time (ms) (TR)/ echo time (ms) (TE); 3000/120) and T1-weighted (TR/TE, 400/8)
248 sagittal and transverse images. Slice thickness for sagittal and transverse images were
249 respectively 1.75 and 2.5mm with an interslice gap of 0.3mm in both planes. The
250 transverse images were aligned perpendicular to the spinal cord. Diagnostic accuracy
251 was calculated using the assessments of the 80 original MRI studies (Table 1).
252 Overall diagnostic accuracy without using the proposed guidelines for the six
253 observers combined was 70.8% (340/480) and ranged from 45% (36/80) to 88.8%
254 (71/80) for individual observers (Table 1). There was a significant influence of
255 observer experience with observers without experience being least accurate (58%)
256 and observers with moderate experience being most accurate (78%) ($X^2 = 18.78$, $P <$
257 0.001). There was no significant difference between the number of accurately
258 diagnosed thoracolumbar intervertebral disk extrusions or protrusions before
259 guidelines were included (74% vs. 75%, respectively, $P > 0.05$). Overall diagnostic

260 accuracy for the six observers combined improved significantly ($P = 0.028$) from
261 70.8% to 79.6% (382/480) after using the proposed diagnostic guidelines and ranged
262 from 52.5% (42/80) to 92.5% (74/80) after using the MRI guidelines. Additionally, all
263 six observers diagnosed a higher number of thoracolumbar intervertebral disk
264 extrusions or protrusions correctly after using the guidelines (mean improvement;
265 13.4%). There was again a significant influence of observer experience ($X^2 = 31.72$, P
266 < 0.001) with observers without experience being least accurate (65%) and observers
267 with expert experience being most accurate (88%). After using the guidelines
268 thoracolumbar intervertebral disk protrusions (84%) were significantly more often (P
269 $= 0.024$) correctly diagnosed than intervertebral disk extrusions (75%) ($X^2 = 5.13$, $P =$
270 0.02)

271

272 Intraobserver agreement was calculated using the assessments of the 40
273 duplicate MRI studies (Table 2). Intraobserver agreement for making a diagnosis of
274 thoracolumbar intervertebral disk extrusion or protrusion ranged from 0.4 to 0.95
275 before, and from 0.51 to 0.85 after applying the proposed guidelines. This difference
276 was not significant and there was no significant influence of observer experience.
277 Although a high degree of variability was seen among individual observers, most
278 variables were associated with moderate or good intraobserver agreement (Table 2).
279 Assessing if multiple intervertebral disk herniations were present was associated with
280 the highest intraobserver agreement ($\kappa = 0.43 - 0.79$) and assessing if an
281 intervertebral disk herniation was midline or lateralized was associated with the
282 lowest intraobserver agreement ($\kappa = 0.33 - 0.75$). There was no significant influence
283 of level of experience on intraobserver agreement of the assessed MRI variables.

284

285 Interobserver agreement was calculated using the assessments of the 80
286 original MRI studies. Overall interobserver agreement for making a diagnosis of
287 thoracolumbar intervertebral disk extrusion or protrusion significantly improved ($P <$
288 0.001) from fair ($\kappa = 0.27$) to moderate ($\kappa = 0.41$) after using the proposed guidelines.
289 Assessing if an intervertebral disk herniation was midline or lateralized and if and
290 intervertebral disk herniation was confined to or exceeded the borders of the affected
291 intervertebral disk space were associated with moderate interobserver agreement
292 (both $\kappa = 0.43$). Assessing the presence of partial intervertebral disk degeneration ($\kappa =$
293 0.35), complete disk degeneration ($\kappa = 0.40$), and multiple intervertebral disk
294 herniations ($\kappa = 0.35$) were associated with fair interobserver agreement.

295

296 **Discussion**

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

306

307 Differentiating thoracolumbar intervertebral disk extrusions from protrusions is
308 clinically important. Although several studies have characterized the clinical
309 presentation and outcomes after medical or surgical treatment for dogs with

310 thoracolumbar intervertebral disk extrusion²⁶⁻³⁰, this information is less well
311 characterized for dogs with intervertebral disk protrusion. While surgical treatment of
312 thoracolumbar intervertebral disk extrusion typically consists of a hemilaminectomy
313 with fenestration of the affected intervertebral disk²⁸⁻³⁰, it is suggested that a lateral
314 corpectomy or vertebral stabilization should be considered for dogs with
315 intervertebral disk protrusion.¹¹⁻¹⁵ While the former can be considered a basic spinal
316 surgical technique, the latter are probably more technically demanding. Furthermore,
317 little is known about results of medical management in dogs with thoracolumbar
318 intervertebral disk protrusions⁴, it has been suggested that dogs with thoracolumbar
319 intervertebral disk extrusion have a better prognosis after surgery than dogs with
320 intervertebral disk protrusion⁴, and that dogs with thoracolumbar intervertebral disk
321 protrusion are at increased risk of early postoperative neurological deterioration.^{4,31}
322 This illustrates that reaching an accurate and reliable imaging diagnosis of
323 thoracolumbar intervertebral disk extrusion or protrusion is not only important for
324 surgical planning, but also for managing the expectations of owners and clinical staff.
325 The overall accuracy to differentiate thoracolumbar intervertebral disk extrusion from
326 intervertebral disk protrusion improved significantly from 70.8 to 79.6%, and the
327 interobserver agreement for making a diagnosis improved significantly from fair to
328 moderate after applying the proposed MRI guidelines. Although these findings
329 indicate clinical usefulness, a proportion of cases were still diagnosed incorrectly after
330 using the proposed guidelines. This suggests that MRI in isolation cannot be
331 completely relied on to differentiate between thoracolumbar extrusions and
332 protrusions and confirms that imaging findings should always be interpreted in light
333 of a thorough clinical history and results of a complete clinical examination.³² In
334 combination to the evaluated MRI variables, a clinical variable has been suggested to

335 be independently associated with a diagnosis of thoracolumbar intervertebral disk
336 extrusion or protrusion. Increased duration of clinical signs, with a threshold value of
337 21 days, was significantly associated with a diagnosis of intervertebral disk
338 protrusion.⁹ The observers in this study were however unaware of the clinical history
339 and clinical signs of included dogs. It is therefore possible that better diagnostic
340 accuracy and reliability would have been reached if this information had been
341 available.

342

343 Intra –and interobserver agreement for the assessed MRI variables was evaluated with
344 kappa statistics. This is a useful and widespread statistical technique to evaluate
345 reliability and reproducibility.^{25,33} Obtained kappa coefficients are commonly
346 translated into poor, fair, moderate, good or very good agreement.^{23,24} Although this
347 can provide useful information, results of kappa statistics do not directly determine if
348 a certain variable can be assessed with acceptable reliability and reproducibility.^{25,33}
349 Although such a decision is dependent on the specific variable and clinical context,
350 studies in musculoskeletal research have suggested that kappa coefficients lower than
351 0.4 should be considered clinically unacceptable.^{25,33} When applying this threshold
352 value to the study presented here, intraobserver agreement for the assessed MRI
353 variables could be considered clinically acceptable. Only one observer did not reach
354 intraobserver agreement coefficients ≥ 0.4 for all assessed MRI variables.

355 Interobserver agreement for the assessed variables ranged from fair to moderate with
356 3 of the 5 assessed variables reaching the arbitrary threshold value (≥ 0.4) for
357 acceptable agreement. Presence of multiple intervertebral disk herniations and degree
358 of intervertebral disk degeneration were associated with only fair interobserver
359 agreement. The lower interobserver agreement for the presence of multiple

360 intervertebral disk herniations is in agreement with a previous study evaluating the
361 number of spinal cord compressions in dogs with disk-associated cervical
362 spondylomyelopathy.³⁴ In the current study, assessment of multiple intervertebral disk
363 herniations was among the variables with the highest intraobserver agreement,
364 indicating consistent rating within an individual observer. Different degrees of
365 intervertebral disk herniation can occur in clinically normal dogs^{20,35} and it is not
366 always possible to clearly differentiate clinically relevant from irrelevant
367 intervertebral disk herniations.³² It is therefore possible that different observers will
368 have different, but individually consistent, criteria for evaluating if an intervertebral
369 disk herniation is present or not. This variable is therefore likely associated with
370 inherent subjectivity and it should be considered to develop a more objectively
371 definition of intervertebral disk herniation. Although a grading system based on
372 accepted terminology^{19,20,36} with easily recognizable characteristics was used,
373 evaluation of degree of intervertebral disk degeneration was also associated with
374 lower values for interobserver agreement. This finding is however in contrast with
375 several human and veterinary studies.^{19,32,37-40} A previous veterinary study
376 demonstrated good agreement between experienced observers for rating intervertebral
377 disk degeneration using the same criteria as this study.³² Although other reasons
378 cannot be excluded, it is possible that the inclusion of observers with varying degrees
379 of training and experience, including observers without experience, contributed to the
380 lower value for interobserver agreement in this study.^{39,41}

381

382 The influence of observer experience was further highlighted by the fact that ‘low
383 observer experience’ negatively influenced the diagnostic accuracy of the assessed
384 MRI studies. Even after using the diagnostic guidelines, the group of observers

385 without experience had significantly lower accuracy for diagnosing thoracolumbar
386 intervertebral disk extrusion and intervertebral disk protrusion. This finding is in
387 agreement with several human studies^{39,42,43} and indicates that assessment and
388 clinical interpretation of MRI studies should be preserved for people with sufficient
389 experience and training. It is therefore possible that the overall results of this study are
390 influenced by the inclusion of observers with different experience levels and it can be
391 debated if our results therefore appropriately reflect referral practice. For the purpose
392 of this study, observers with different experience categories were included to evaluate
393 if differentiation between thoracolumbar intervertebral disk extrusion and disk
394 protrusion would be more problematic for observers with little experience, if using the
395 proposed guidelines would be easy and objective enough to be used by novices, and if
396 using these imaging guidelines could compensate for lack of training and experience.

397

398 A potential limitation of this study was the limited number of observers in each
399 ‘experience category’. Although care was taken to select observers of similar
400 experience levels. In addition, there was no standardization across observers in terms
401 of how the assessments were performed (i.e. all in one sitting or a few at a time) or
402 the monitor on which they were viewed. Another limitation was that all patients had
403 surgically confirmed intervertebral disk extrusion or protrusion. The combination of
404 imaging and clinical findings was therefore severe enough to warrant surgical
405 intervention. These factors could have influenced the assessment of the observers;
406 however surgical confirmation was necessary to provide a “golden standard”
407 diagnosis with which to assess diagnostic accuracy. It should further be emphasized
408 that MRI studies were evaluated by general clinicians or veterinary surgeons trained
409 in veterinary neurology and neurosurgery. It can however not be excluded that

410 interpretation of MRI studies by observers with a different training background, such
411 as veterinary radiologists, would have influenced our results. Veterinary neurologists
412 were however preferred in this study, because in our institution, these clinicians
413 perform MRI interpretation and spinal surgery on a daily basis and were therefore
414 considered to have most experience in taking clinical decisions after assessing spinal
415 MRI studies. Although only limited data is available on direct comparisons between
416 veterinary neurosurgeons and radiologists for assessing spinal MRI studies^{34,44,45},
417 human studies indicate acceptable agreement between specialities.⁴⁶⁻⁴⁹ Although
418 comparing the accuracy and agreement between veterinary neurologists and
419 radiologists was considered beyond the aims of this study, further studies are
420 necessary to not only evaluate the role of experience, but also the type of training on
421 the reliability and accuracy of evaluating spinal MRI studies

422

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

431

432

433 **Acknowledgments:** None

434

435

436

437 **References**

438

- 439 1. Hansen HJ. A pathologic-anatomical study on disc degeneration in dog, with
440 special reference to the so-called enchondrosis intervertebralis. Acta Orthop Scand
441 Suppl 1952;11:1–117.
- 442 2. Berknut N, Rutges JP, Kranenburg HJ, et al. The dog as an animal model for
443 intervertebral disc degeneration? Spine 2012;37:351-358.
- 444 3. Bergknut N, Meij BP, Hagman R, et al. Intervertebral disc disease in dogs – Part 1:
445 A new histological grading scheme for classification of intervertebral disc
446 degeneration in dogs. Vet J 2013;195:156–163.
- 447 4. Macias C, McKee WM, May C, Innes JF. Thoracolumbar disc disease in large
448 dogs: a study of 99 cases. J Small Anim Pract 2002;43:439–446.
- 449 5. Smolders LA, Bergknut N, Grinwis GCM, et al. Intervertebral disc degeneration in
450 the dog. Part 2: Chondrodystrophic and non-chondrodystrophic breeds. Vet J
451 2013;195:292–299.
- 452 6. Brisson BA. Intervertebral disc disease in dogs. Vet Clin North Am Small Anim
453 Pract 2010;40:829–858.
- 454 7. Griffiths IR. Some aspects of the pathogenesis and diagnosis of lumbar disc
455 protrusion in the dog. J Small Anim Pract 1972;13:439–447.
- 456 8. Bray JP, Burbidge HM. The canine intervertebral disk. Part Two: Degenerative
457 changes--nonchondrodystrophoid versus chondrodystrophoid disks. J Am Anim Hosp
458 Assoc 1998;34:135–144.
- 459 9. Gomes SA, Volk HA, Packer RMA, et al. Clinical and magnetic resonance imaging
460 characteristics of thoracolumbar intervertebral disk extrusions and protrusions in large
461 breed dogs. Vet Radiol Ultrasound 2016; doi: 10.1111/vru.12359..

- 462 10. Cudia SP, Duval JM. Thoracolumbar intervertebral disk disease in large,
463 nonchondrodystrophic dogs: a retrospective study. *J Am Anim Hosp Assoc*
464 1997;33:456–460.
- 465 11. Moissonnier P, Meheust P, Carozzo C. Thoracolumbar Lateral Corpectomy for
466 Treatment of Chronic Disk Herniation: Technique Description and Use in 15 Dogs.
467 *Vet Surg* 2004;33:620–628.
- 468 12. McKee WM, Downes CJ. Vertebral stabilisation and selective decompression for
469 the management of triple thoracolumbar disc protrusions. *J Small Anim Pract*
470 2008;49:536–539.
- 471 13. Downes CJ, Gemmill TJ, Gibbons SE, McKee WM. Hemilaminectomy and
472 vertebral stabilisation for the treatment of thoracolumbar disc protrusion in 28 dogs. *J*
473 *Small Anim Pract* 2009;50:525–535.
- 474 14. Flegel T, Boettcher IC, Ludewig E, et al. Partial Lateral Corpectomy of the
475 Thoracolumbar Spine in 51 Dogs: Assessment of Slot Morphometry and Spinal Cord
476 Decompression. *Vet Surg* 2010;40:14–21.
- 477 15. Salger F, Ziegler L, Böttcher IC, et al. Neurologic outcome after thoracolumbar
478 partial lateral corpectomy for intervertebral disc disease in 72 dogs. *Vet Surg*
479 2014;43:581-588.
- 480 16. Robertson I, Thrall DE. Imaging dogs with suspected disc herniation: pros and
481 cons of myelography, computed tomography, and magnetic resonance. *Vet Radiol*
482 *Ultrasound* 2011;52:81–84.
- 483 17. Van Wie EY, Fosgate GT, Mankin JM, et al. Prospectively recorded versus
484 medical record-derived spinal cord injury scores in dogs with intervertebral disk
485 herniation. *J Vet Intern Med* 2013;27:1273-1277.

- 486 18. Modic MT, Masaryk TJ, Ross JS, Carter JR. Imaging of degenerative disk
487 disease. *Radiology* 1988;168:177–186.
- 488 19. Bergknut N, Auriemma E, Wijsman S, et al. Evaluation of intervertebral disk
489 degeneration in chondrodystrophic and nonchondrodystrophic dogs by use of
490 Pfirrmann grading of images obtained with low-field magnetic resonance imaging.
491 *Am J Vet Res* 2011;72:893–898.
- 492 20. da Costa RC, Parent JM, Partlow G, et al. Morphologic and morphometric
493 magnetic resonance imaging features of Doberman Pinschers with and without
494 clinical signs of cervical spondylomyelopathy. *Am J Vet Res* 2006;67:1601-1612.
- 495 21. Besalti O, Pekcan Z, Sirin YS, Erbas G. Magnetic resonance imaging findings in
496 dogs with thoracolumbar intervertebral disk disease: 69 cases (1997-2005). *J Am Vet*
497 *Med Assoc* 2006;228:902–908.
- 498 22. Fleis JL. Measuring nominal scale agreement among many raters. *Psychol Bull*
499 1971;76:378-382.
- 500 23. Landis JR, Koch GG. Measurement of observer agreement for categorical data.
501 *Biometrics* 1977;33:159-174.
- 502 24. Altman DG. In: *Practical statistics for medical research*, London England:
503 Chapman and Hall CRC 1991;404.
- 504 25. Sim J, Wright CC. The Kappa statistic in reliability studies: use, interpretation,
505 and sample size requirements. *Phys Ther* 2005;85:257-268.
- 506
- 507 26. Olby N, Levine J, Harris T, et al. Long-term functional outcome of dogs with
508 severe injuries of the thoracolumbar spinal cord: 87 cases (1996-2001). *J Am Vet*
509 *Med Assoc* 2003;222:762–769.

- 510 27. Ruddle TL, Allen DA, Schertel ER, et al. Outcome and prognostic factors in non-
511 ambulatory Hansen type I intervertebral disc extrusions: 308 cases. *Vet Comp Orthop*
512 *Traumatol* 2006;19:29–34.
- 513 28. Brisson BA, Holmberg DA, Parent J, et al. Comparison of the effect of single-site
514 and multiple-site disk fenestration on the rate of recurrence of thoracolumbar
515 intervertebral disk herniation in dogs. *J Am Vet Med Assoc* 2011;238:1593–1600.
- 516 29. Aikawa T, Fujita H, Shibata M, et al. Recurrent thoracolumbar intervertebral disc
517 extrusion after hemilaminectomy and concomitant prophylactic fenestration in 662
518 chondrodystrophic dogs. *Vet Surg* 2012;41:381–390.
- 519 30. Aikawa T, Fujita H, Kanazono S, et al. Long-term neurologic outcome of
520 hemilaminectomy and disk fenestration for treatment of dogs with thoracolumbar
521 intervertebral disk herniation: 831 cases (2000-2007). *J Am Vet Med Assoc*
522 2012;241:1617–1626.
- 523 31. Olby NJ, Jeffery N. Pathogenesis and physiology of central nervous system
524 disease and injury. In: Tobias, KM, Johnston SA, ed. *Veterinary surgery. Small*
525 *animal*. Missouri: Elsevier Saunders; 2012:374-387.
- 526 32. De Decker S, Gielen I, Duchateau L, et al. Intraobserver and interobserver
527 agreement for results of low-field magnetic resonance imaging in dogs with and
528 without clinical signs of disk-associated wobbler syndrome. *J Am Vet Med Assoc*
529 2011;238:74–80.
- 530 33. Petersen IS. using the kappa coefficient as a measure of reliability and
531 reproducibility. *Chest* 1998;114:946-947.
- 532 34. De Decker S, Gielen IM, Duchateau L, et al. Intraobserver, interobserver, and
533 intermethod agreement for results of myelography, computed tomography-

534 myelography, and low-field magnetic resonance imaging in dogs with disk-associated
535 wobbler syndrome. *J Am Vet Med Assoc* 2011;238:1601-1608.

536 35. De Decker S, Gielen I, Duchateau L, et al. Low-field magnetic resonance imaging
537 findings of the caudal portion of the cervical region in clinically normal Doberman
538 Pinschers and Foxhounds. *Am J Vet Res* 2010;71:428–434.

539 36. Seiler G, Häni H, Scheidegger J, Busato A, Lang J. Staging of lumbar
540 intervertebral disc degeneration in nonchondrodystrophic dogs using low-field
541 magnetic resonance imaging. *Vet Radiol Ultrasound* 2003;44:179–184.

542 37. Cihangiroglu M, Yildirim H, Bozgeyik Z, et al. Observer variability based on the
543 strength of MR scanners in the assessment of lumbar degenerative disc disease. *Eur J*
544 *Radiol* 2004;51:202-208.

545 38. Solgaard Sorensen J, Kjaer P, Jensen ST, Andersen P. Low-field magnetic
546 resonance imaging of the lumbar spine: reliability of qualitative evaluation of disc and
547 muscle parameters. *Acta Radiol* 2006;47:947-953.

548 39. Truszczyńska A, Brychcy A, Rapala K, Walczak P, Truszczyński O, Tarnowski
549 A. Interrater and intrarater reliability of lumbar spine magnetic resonance assessment
550 of military aircraft pilot candidates. *Ortop Traumatol Rehabil* 2011;13:343-351.

551 40. Sander AL, Laurer H, Lehnert T, et al. A clinically useful classification of
552 traumatic intervertebral disk lesions. *Am J Roentgenol* 2013;200:618-623.

553 41. Knight K, Hiller ML, Simpson DD, Broome KM. the validity of self-reported
554 cocaine use in a criminal justice treatment sample. *Am J Drug Alcohol Abuse*
555 1998;24:647-660.

556 42. Song KJ, Choi BW, Kim GH, Kim JR. Clinical usefulness of CT-myelogram
557 comparing with the MRI in degenerative cervical spinal disorders; is CTM still useful
558 for primary diagnostic tool? *J Spinal Disord Tech* 2009;22:353-357.

559 43. Baltzer PAT, Kaiser WA, Dietzel M. Lesion type and reader experience affect the
560 diagnostic accuracy of breast MRI: a multiple reader ROC study. *Eur J Radiol*
561 2015;84:86-91.

562 44. Bos AS, Brisson BA, Nykamp SG, et al. Accuracy, intermethod agreement, and
563 inter-reviewer agreement for use of magnetic resonance imaging and myelography in
564 small-breed dogs with naturally occurring first-time intervertebral disk extrusion. *J*
565 *Am Vet Med Assoc* 2012;240:969-977.

566 45. Fenn J, Drees R, Volk HA, De Decker S. Inter- and intraobserver agreement for
567 diagnosing presumptive ischemic myelopathy and acute noncompressive nucleus
568 pulposus extrusion in dogs using magnetic resonance imaging. *Vet Radiol Ultrasound*
569 2016;57:33-40.

570 46. Rihn JA, Yang N, Fisher C, et al. Using magnetic resonance imaging to accurately
571 assess injury to the posterior ligamentous complex of the spine: a prospective
572 comparison of the surgeon and radiologist. *J Neurosurg Spine* 2010;12:391-396.

573 47. Lee WJ, Yoon SH, Kim JY, et al. Interobserver and intraobserver reliability of
574 sub-axial injury classification and severity scale between radiologist, resident and
575 spine surgeon. *J Korean Neurosurg Soc* 2012;52:200-203.

576 48. Imaad-ur-Rehman. Hamid RS, Akhtar W, et al. Observer variation in MRI
577 evaluation of patients with suspected lumbar disc herniation and nerve root
578 compression: comparison of neuroradiologist and neurosurgeon's interpretations. *J*
579 *Pak Med Assoc* 2012;62:826-829.

580 49. Fisher CG, Versteeg AL, Schouten R, et al. Reliability of the spinal instability
581 neoplastic scale among radiologists: an assessment of instability secondary to spinal
582 metastases. *Am J Roentgenol* 2014;2013:869-874.

583

584

585

586

587

588

589

590 **Legends**

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

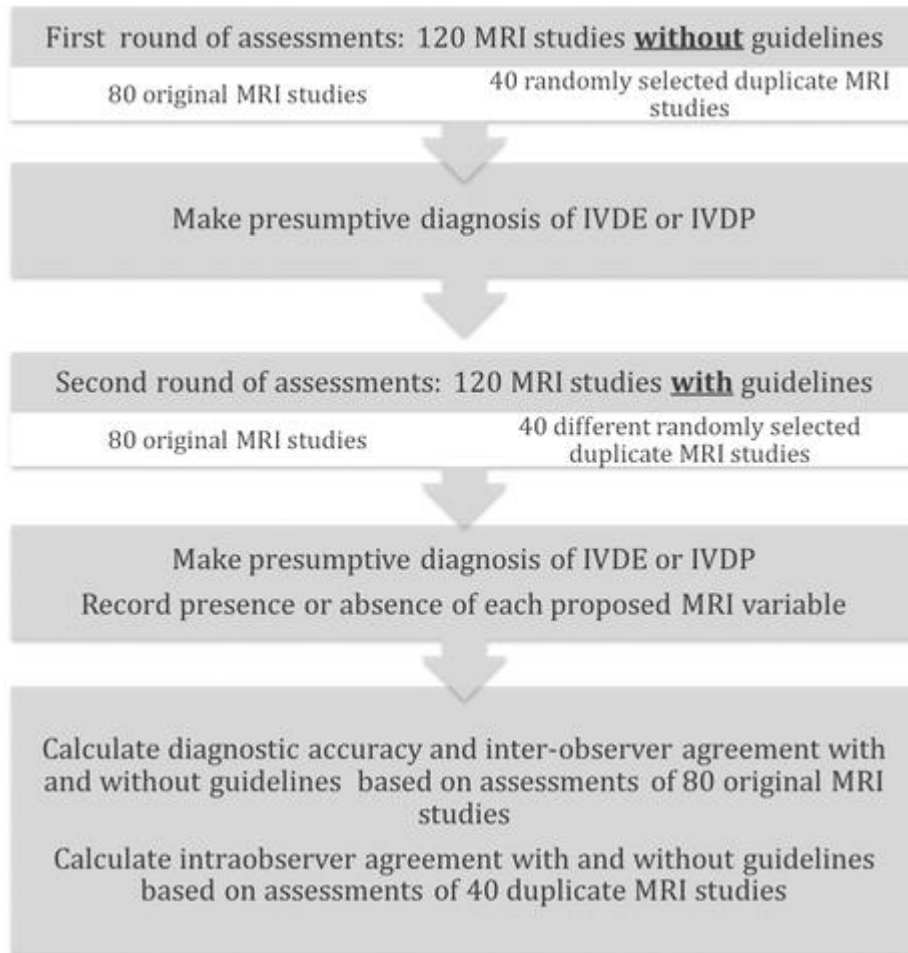
MRI variable	Intervertebral disk Extrusion	Intervertebral disk Protrusion
Lateralized intervertebral disk herniation	More likely	Less likely
Partial instead of complete intervertebral disk degeneration	Less likely	More likely
Multiple intervertebral disk herniations	Less likely	More likely
Herniated disk material confined to intervertebral disk space	Less likely	More likely

594

595

596

597 **Figure 2.** Flow-chart illustrating different MRI assessments in order to calculate the
598 diagnostic accuracy, inter-observer agreement and intra-observer agreement for
599 differentiating between thoracolumbar intervertebral disk extrusions and protrusions
600 with and without the use of the proposed imaging guidelines.



601

602 **Table 1.** Number of correctly diagnosed thoracolumbar intervertebral disk extrusions
 603 or protrusions before and after using the proposed MR imaging guidelines.
 604 Calculations were based on 80 original MRI studies
 605

Observer	Degree of experience	Correct diagnoses without guidelines (out of 80)	Correct diagnoses with guidelines (out of 80)
Observer 1	Expert	69 (86.2%)	73 (91.2%)
Observer 2	Moderate	53 (66.2%)	63 (78.8%)
Observer 3	None	36 (45%)	42 (52.5%)
Observer 4	Expert	54 (67.5%)	68 (85%)
Observer 5	Moderate	71 (88.8%)	74 (92.5%)
Observer 6	None	57 (71.2%)	62 (77.5%)
Overall	NA	340 (70.8%)	382 (79.6%)

606

607 **Table 2.** Intraobserver agreement κ -values for making a diagnosis of thoracolumbar
608 intervertebral disk extrusion or protrusion and for assessment of each of the proposed
609 MRI variables. Calculations were based on 40 duplicate MRI studies

610

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

611

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

615

616

617