Intervertebral disk degeneration is a common disorder in dogs. In the dog population in Sweden, IVD degeneration–related diseases are among the top 5 reasons for euthanasia of dogs <10 years old.1,4,5 Intervertebral disk degeneration–related diseases include Hansen type I and II IVD herniation, degenerative lumbosacral stenosis, and cervical spondylomyelopathy.6 Intervertebral disk degeneration is more common in chondrodystrophic breeds than in nonchondrodystrophic breeds and in older dogs than in younger dogs.2,7,8 It is important to mention that IVD degeneration is not synonymous with IVD disease. An IVD that causes clinical signs will invariably be degenerated, but degenerated IVDs are also common findings in dogs without clinical signs of disease.2,9–11

Many similarities exist between IVD degeneration in humans and dogs.12–14 For this reason, dogs are sometimes used in research settings as a means to study spontaneous IVD degeneration in humans.15 In humans and dogs, it can be difficult to diagnose IVD degeneration–related diseases, and the use of diagnostic imaging is essential for an accurate diagnosis. It is sometimes difficult to distinguish disks with pathological degeneration from disks that have normal age-related changes (known as normal senile remodeling).16 Moreover, not all degenerated disks cause clinical signs, irrespective of the cause of degeneration.17,18

It is believed that IVD herniation in dogs results from IVD degeneration.19 The current surgical treatment is laminectomy in combination with nucleotomy of the herniated disk. However, in humans with IVD degeneration, the Pfirrmann system for grading lumbar intervertebral disk (IVD) degeneration in humans can also be used in dogs.10–12

Objective—To assess whether the Pfirrmann system for grading lumbar intervertebral disk (IVD) degeneration in humans can also be used in dogs.

Animals—202 dogs.

Procedures—Magnetic resonance imaging was used to obtain images of vertebral segments from dogs, which were reviewed separately by 3 observers who graded the extent of degeneration in each visible IVD by use of the Pfirrmann classification system used for grading lumbar IVD degeneration in humans. Grading was validated against 2 factors associated with the extent of disk degeneration: type of dog (chondrodystrophic or nonchondrodystrophic breeds) and age.

Results—Interobserver and intraobserver agreement for Pfirrmann grading of IVD degeneration were good (κ scores, 0.81 to 0.93). An increase in the extent of disk degeneration was positively correlated with increases in age and with chondrodystrophic breed.

Conclusions and Clinical Relevance—The Pfirrmann system was reliably used to grade IVD degeneration in dogs of various breeds and ages. An increase in the extent of IVD degeneration was positively correlated with increases in age and with chondrodystrophic-type dogs. (Am J Vet Res 2011;72:893–898)

<table>
<thead>
<tr>
<th>Abbreviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVD</td>
</tr>
<tr>
<td>MRI</td>
</tr>
</tbody>
</table>

Received November 4, 2009.
Accepted April 22, 2010.
From the Department of Clinical Sciences of Companion Animals (Bergknut, Wijsman, Hazewinkel, Mei) and Division of Diagnostic Imaging (Auriemma, Voorhout), Faculty of Veterinary Medicine, Utrecht University, 3508 TC Utrecht, The Netherlands, and the Department of Clinical Sciences, Division of Small Animals, Faculty of Veterinary Medicine and Animal Sciences, Swedish University of Agricultural Sciences, 736 51 Uppsala, Sweden (Bergknut, Hagman, Lagerstedt).
Presented as a poster at the Dutch National Veterinary Conference, Amsterdam, April 2009.
The authors thank Hans Vernooij for assistance with the statistical analysis.
Address correspondence to Dr. Bergknut (n.bergknut@uu.nl).
Degeneration–related diseases, nucleotomy can further destabilize affected vertebrae, and it is likely that this is also true in dogs. Although vertebral instability in humans is currently treated by means of vertebrectomy, there is a trend to avoid this type of salvage technique and to focus on regenerative treatment strategies that can be used to achieve functional repair or even regeneration of degenerated IVDs. Such strategies require early identification of the degenerative process through the use of reliable and accurate diagnostic methods and grading systems. Although these types of regenerative treatments may soon be implemented in human medicine, difficulties in identifying early disease in dogs may delay the use of these treatments in veterinary medicine. However, with the rapid development of veterinary diagnostic methods, early treatment to prevent severe clinical signs may soon become available for dogs. When performing clinical research, attempting to stop IVD degeneration, or stimulating IVD regeneration, the Pfirrmann grading scheme is a valuable tool for use in monitoring the progression of degeneration or regeneration of IVDs in vivo.

In human medicine, the Pfirrmann system is the most widely used system for grading IVD degeneration on the basis of MRI findings. It is based on a system for grading gross pathological changes in IVDs proposed by Thompson et al, which is the most commonly used criterion-referenced standard in human medicine. Two systems have been proposed for the grading of IVD degeneration in dogs: use of histologic examination and use of gross morphology. A third grading system has been proposed for grading of thoracolumbar disks of dogs, but it has not yet been validated. None of these systems has been validated for interobserver and intraobserver agreement, nor is there a validated grading system available for IVDs in all locations of the vertebral column.

The purpose of the study reported here was to evaluate whether the MRI-based Pfirrmann system for the grading of IVD degeneration in lumbar disks of humans is applicable for use in all types of dogs and for IVDs in all locations of the vertebral column. Therefore, we determined the interobserver and intraobserver agreement between Pfirrmann scores and evaluated whether biological factors known to increase the extent of IVD degeneration were correlated with higher Pfirrmann scores.

Materials and Methods

Animals—The initial study population was 217 dogs examined at the Clinic of Companion Animals, Utrecht University, The Netherlands, between 2002 and 2008 for which clinicians obtained midsagittal T2-weighted MRI images of any portion of the vertebral column. Although most dogs included in the study were examined because of suspected IVD degeneration–related problems, many were examined for other reasons, such as suspected syringomyelia, neoplasia, fibrocorticalaginous emboli, and degenerative myelopathy. Poor image resolution led to the exclusion of data for 15 small-breed dogs. Thus, 994 IVDs from 202 dogs of various breeds and age and both sexes were used for assessment. Of the 202 dogs, 66 were considered (on the basis of breed) to be chondrodystrophic and the remaining 136 were considered to be nonchondrodystrophic. Although there is no comprehensive list of breeds regarded as chondrodystrophic, the Basenji, Beagle, Dachshund, Miniature Dachshund, English Bulldog, French Bulldog, Jack Russell Terrier, Pug, Miniature Poodle, Shih Tzu, and Welsh Corgi are typically considered chondrodystrophic breeds.

The term IVD degeneration was used to encompass all signs of disk degeneration on MRI images, regardless of the cause of the degeneration and whether the degeneration was associated with clinical signs.

MRI technique—Magnetic resonance imaging was performed by use of a 0.2-T open magnet. Only T2-weighted MRI images (repetition time, 3,835 to 4,450 milliseconds; echo time, 117 milliseconds) were used. The obtained slices had a thickness of 3 mm. A 21-cm-diameter multipurpose flex coil was used for smaller dogs (dogs fitting the coil), and a circular polarization flexible body and spine coil was used for larger dogs.

Image assessment—The 5-category system developed by Pfirrmann et al to grade IVD degeneration in lumbar disks of humans was used. Sagittal T2-weighted images of the cervical portion of the vertebral column of a dog. The C5-6 disk (arrow) has evidence of degeneration (Pfirrmann grade 3).

Figure 1—A midsagittal T2-weighted MRI image of the cervical portion of the vertebral column of a dog. The C5-6 disk (arrow) has evidence of degeneration (Pfirrmann grade 3).

Figure 2—A series of midsagittal T2-weighted MRI images of IVDs (arrows) obtained from representative dogs. The images depict Pfirrmann grades of 1 (healthy IVD; far left) to 5 (end-stage IVD degeneration; far right).
tal T2-weighted MRI images were graded separately by 3 observers: a veterinary medical student (observer 1), a student in a PhD program (observer 2), and a board-certified veterinary radiologist (observer 3). The MRI images were examined on standard computer screens by use of software. Each image depicted a portion of the vertebral column that contained several IVDs. The same disks and sagittal slices were viewed by all observers at the automatic settings provided by the software. All observers were familiar with the Pfirrmann grading procedure and had images with examples of the various grades of degeneration for humans available during the grading procedure. When criteria for >1 category were applicable to the same IVD, the higher grade was selected because the IVD was considered to have signs of progressive degeneration. The MRI images were graded twice with a 2-week interval between grading for observer 1 and a 6-month interval between grading for observers 2 and 3. Despite the shorter interval for observer 1, bias attributable to recollection of the grade assigned by that reviewer for each of the 994 images was deemed unlikely. Intraobserver agreement of the 3 observers was compared to determine whether the short interval affected grading.

Biological validation—To determine whether results for the Pfirrmann scoring correlated with variables known to be associated with increasing disk degeneration, we evaluated correlations between Pfirrmann scores and each dog’s age and predisposition to chondrodystrophy (ie, chondrodystrophic or nonchondrodystrophic breed). Investigators in other studies have reported that IVD degeneration is strongly correlated with older age and with chondrodystrophic breed. Thus, these variables were used to represent biological validation.

Table 1—Pfirrmann grades for 994 IVDs of 202 chondrodystrophic and nonchondrodystrophic dogs.

<table>
<thead>
<tr>
<th>Pfirrmann grade</th>
<th>Nonchondrodystrophic</th>
<th>Chondrodystrophic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>211 (33.7)</td>
<td>54 (14.7)</td>
<td>265 (26.6)</td>
</tr>
<tr>
<td>2</td>
<td>229 (36.6)</td>
<td>64 (17.4)</td>
<td>293 (29.5)</td>
</tr>
<tr>
<td>3</td>
<td>130 (20.8)</td>
<td>163 (44.3)</td>
<td>293 (29.5)</td>
</tr>
<tr>
<td>4</td>
<td>50 (8.0)</td>
<td>86 (23.3)</td>
<td>136 (13.7)</td>
</tr>
<tr>
<td>5</td>
<td>6 (0.9)</td>
<td>1 (0.3)</td>
<td>7 (0.7)</td>
</tr>
<tr>
<td>Total</td>
<td>626 (100)</td>
<td>368 (100)</td>
<td>994 (100)</td>
</tr>
</tbody>
</table>

Values reported are the number of IVDs (percentage of total for the column). Grading was performed on T2-weighted MRI images; Pfirrmann scores ranged from grade 1 (healthy IVD) to 5 (end-stage IVD degeneration).

Figure 3—Distribution of Pfirrmann grade (1 = light gray bars, 2 = black bars, 3 = white bars, 4 = dark gray bars, and 5 = vertical striped bars) in relation to age in 66 chondrodystrophic dogs (A) and 136 nonchondrodystrophic dogs (B). The number of IVDs per Pfirrmann grade are reported for each age group. In chondrodystrophic dogs, higher Pfirrmann grades predominate in dogs ≥ 2 years old. Grading was performed on T2-weighted MRI images; Pfirrmann scores range from grade 1 (healthy IVD) to 5 (end-stage IVD degeneration).

Figure 3—Distribution of Pfirrmann grade (1 = light gray bars, 2 = black bars, 3 = white bars, 4 = dark gray bars, and 5 = vertical striped bars) in relation to age in 66 chondrodystrophic dogs (A) and 136 nonchondrodystrophic dogs (B). The number of IVDs per Pfirrmann grade are reported for each age group. In chondrodystrophic dogs, higher Pfirrmann grades predominate in dogs ≥ 2 years old. Grading was performed on T2-weighted MRI images; Pfirrmann scores range from grade 1 (healthy IVD) to 5 (end-stage IVD degeneration).

Table 1—Pfirrmann grades for 994 IVDs of 202 chondrodystrophic and nonchondrodystrophic dogs.

<table>
<thead>
<tr>
<th>Pfirrmann grade</th>
<th>Nonchondrodystrophic</th>
<th>Chondrodystrophic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>211 (33.7)</td>
<td>54 (14.7)</td>
<td>265 (26.6)</td>
</tr>
<tr>
<td>2</td>
<td>229 (36.6)</td>
<td>64 (17.4)</td>
<td>293 (29.5)</td>
</tr>
<tr>
<td>3</td>
<td>130 (20.8)</td>
<td>163 (44.3)</td>
<td>293 (29.5)</td>
</tr>
<tr>
<td>4</td>
<td>50 (8.0)</td>
<td>86 (23.3)</td>
<td>136 (13.7)</td>
</tr>
<tr>
<td>5</td>
<td>6 (0.9)</td>
<td>1 (0.3)</td>
<td>7 (0.7)</td>
</tr>
<tr>
<td>Total</td>
<td>626 (100)</td>
<td>368 (100)</td>
<td>994 (100)</td>
</tr>
</tbody>
</table>

Values reported are the number of IVDs (percentage of total for the column). Grading was performed on T2-weighted MRI images; Pfirrmann scores ranged from grade 1 (healthy IVD) to 5 (end-stage IVD degeneration).

Statistical analysis—Cohen weighted κ analysis was used to evaluate interobserver and intraobserver agreement. Agreement was interpreted as follows: slight (κ, 0 to 0.20), fair (κ, 0.21 to 0.40), moderate (κ, 0.41 to 0.60), substantial (κ, 0.61 to 0.80), and almost perfect (κ, 0.81 to 1.00). The mean of the Pfirrmann grades assigned by the 3 observers for each MRI image was calculated, and this mean grade was used to determine whether the degree of degeneration detected on MRI images was correlated with the dog’s age or the type of dog (chondrodystrophic or nonchondrodystrophic). A Spearman rank test was used to test for the correlation between age and Pfirrmann grade, and a Mann-Whitney U test was used to test for the association between the type of dog and Pfirrmann grade. Values of P < 0.05 were considered significant.

Results

Animals—Of the 202 dogs, 109 were males and 93 were females. Dogs ranged from 1 to 13 years of age (mean, 5.4 years) and represented 58 breeds. There were IVDs in all regions of the vertebral column (C2-3 disk to L7-S1 disk) included in the study.

IVDs—A total of 994 disks (368 disks from chondrodystrophic dogs and 626 disks from nonchondrodystrophic dogs) were graded on T2-weighted MRI images that each depicted a portion of the vertebral column containing several IVDs...
Of the 994 disks evaluated, 265 (26.6%) were assigned a grade of 1, 293 (29.5%) were assigned a grade of 2, and 293 (29.5%) were assigned a grade of 3; only 136 (13.7%) and 7 (0.7%) disks were assigned a grade of 4 and 5, respectively (Figure 2; Table 1). Of the 368 disks from chondrodystrophic dogs, 250 (67.9%) were assigned a grade of 3, 4, or 5. Of the 626 disks from nonchondrodystrophic dogs, 440 (70.3%) were assigned a grade of 1 or 2. In the disks of the chondrodystrophic dogs, Pfirrmann grades 3 and 4 were predominant in dogs > 2 years old (Figure 3), whereas in the nonchondrodystrophic dogs, Pfirrmann grades 1 and 2 were predominant in all age groups. In contrast to other chondrodystrophic dogs, Jack Russell Terriers mainly had IVDs with low Pfirrmann grades, regardless of age of the dog.

**Intraobserver agreement**—Intraobserver agreement was almost perfect (mean ± SE \( \kappa = 0.93 \pm 0.032 \), \( 0.89 \pm 0.031 \), and \( 0.92 \pm 0.032 \) for observers 1, 2, and 3, respectively; Table 2). In the grades assigned for the 994 IVDs by each of the observers at the 2 grading sessions, there was a difference of 1 grade between scores assigned for 131 (13.2%), 221 (22.3%), and 156 (15.7%) IVDs for observers 1, 2, and 3, respectively. There was a difference of 2 or 3 grades for scores assigned for 20 (2%) of all intraobserver comparisons. No significant difference in intraobserver agreement was detected for the 3 observers.

**Interobserver agreement**—Interobserver agreement was almost perfect (mean ± SE \( \kappa = 0.93 \pm 0.032 \), \( 0.89 \pm 0.031 \), and \( 0.92 \pm 0.032 \) between observers 1 and 2, \( 0.87 \pm 0.031 \) between observers 1 and 3, and \( 0.85 \pm 0.031 \) between observers 2 and 3; Table 2). For 263 (26.5%) to 340 (34.2%) IVDs, there was a difference of 1 grade between 2 observers. There was a difference of 2 or 3 grades between observers for 3 (0.3%) to 35 (3.5%) IVDs, and a difference of 3 grades between observers for 1 (0.1%) image.

**Biological validation**—A significant association (\( Z = -10.8; P < 0.001 \)) was detected between Pfirrmann grade and type of dog (chondrodystrophic or nonchondrodystrophic). High grades (ie, more degeneration) were assigned more often for IVDs of chondrodystrophic dogs than for IVDs of nonchondrodystrophic dogs, and low grades (ie, less degeneration) were assigned more often for IVDs of nonchondrodystrophic dogs than for IVDs of chondrodystrophic dogs. In addition, a higher Pfirrmann grade was positively correlated (\( r = 0.169; P = 0.01 \)) with age of the dog.

**Discussion**

Use of the Pfirrmann system yielded highly reproducible results for grading IVD degeneration at all vertebral locations in dogs of various breeds and ages. Moreover, the Pfirrmann score was associated with chondrodystrophic breed and age, which are factors associated with IVD degeneration. Thus, the Pfirrmann system would appear to be suitable for use in grading IVD degeneration in dogs.

Of the 994 disks graded, there was perfect agreement between the observers for 618 (62.2%) to 720 (72.4%); for the remaining disks, there was a discrepancy of ≥ 1 grade between the observers. These discrepancies could have resulted from the fact that the Pfirrmann system is not a continuous scale and the cutoff point between 2 categories is not always clear. For example, the discriminating features between grade 1 and 2 (homogeneous vs nonhomogeneous bright signal of the nucleus) and grade 3 and 4 (ability to discriminate between annulus and nucleus) are to some extent subjective, which could explain the reason that 3 (0.3%) to 35 (3.5%) disks had a difference of 2 grades between observers and 1 (0.1%) disk had a difference of 3 grades between observers. In addition, human error, such as viewing the wrong sagittal image during grading or errors when entering IVD grades, is also a potential cause of discrepancies when disks had a difference of 2 or 3 grades.

Although most dogs included in the study were examined because of suspected IVD-related problems, some dogs were examined because of other reasons, which could explain the broad variation in healthy and degenerated disks. Only 7 disks were scored grade 5 (Table 1). It cannot be determined from the present study whether this was representative of end-stage disk degeneration in the canine population. We are not aware of any studies in which investigators examined the association between extent of disk degeneration and severity of clinical signs in dogs. Other studies have found that chondrodystrophic dogs are more prone to disk degeneration than are nonchondrodystrophic dogs and that older dogs are more likely to have degenerated disks than are younger dogs.

Because the breed standard for the Jack Russell Terrier states that these dogs should have abnormally short...
limbs in relation to their body length, they were classified as being chondrodystrophic. Surprisingly, most IVDs of the Jack Russell Terriers had low Pfirrmann grades, regardless of age of the dogs. This was in contrast to the higher Pfirrmann grades for the IVDs of the other chondrodystrophic breeds, which suggests that chondrodysplasia causing abnormally short limbs is not necessarily associated with IVD degeneration.

In veterinary medicine, low-field (0.2-T) MRI is generally more available than is high-field MRI. The low resolution of sagittal T2-weighted images with low-field MRI was a problem when viewing the IVDs of miniature breeds, and we had to discard the MRI images of 15 small-breed dogs because of poor image quality. Image resolution provided by use of the low-field MRI was sufficient to provide clear images of IVDs in most dogs, but because the clarity of the images decreased with a decrease in IVD size, it may have been relatively inaccurate for small-breed dogs (the breeds investigated in this study ranged from Bullmastiff to Miniature Dachshund). An additional potential problem was that the coil of the MRI machine provided a limited field of view, which resulted in a brighter signal in the focus area of the magnetic field than in areas of the vertebral column located farther from the coil. This is known as the coil effect, and it can lead to falsely higher Pfirrmann grades for disks located farther from the center of the coil on T2-weighted MRI images.41,42 The T2-weighted MRI images were used for Pfirrmann grading because they best depicted the glycosaminoglycan and water content of disks, which are negatively correlated with the extent of disk degeneration.43,44

The Pfirrmann grading system focuses on characteristic changes in the structure of a disk (T2-weighted signal intensity, disk structure, ability to discriminate between the nucleus and annulus, and disk height) and not on changes in the tissues surrounding each disk (end plate sclerosis, vertebral osteophytes, and disk herniation). However, these aspects in the surrounding tissues should be included to obtain a complete picture of the status of disks. Thus, to obtain accurate information about the extent of potential bulging, protrusion, or extrusion of a disk, transverse and midsagittal (for Pfirrmann grading) MRI images are needed. It has been proposed in 1 study27 that IVD degeneration in dogs is almost perfect, with κ scores ranging between 0.81 and 0.93. Moreover, the extent of disk degeneration was significantly associated with breed (chondrodystrophic breeds) and age. On the basis of these results, we conclude that the Pfirrmann grading system can be used to evaluate IVD degeneration in dogs.

References

Appendix

Description of the MRI-based grading system used to classify IVDs in dogs.

<table>
<thead>
<tr>
<th>Pfirrmann grade</th>
<th>Structure</th>
<th>Distinction between NP and AF</th>
<th>Signal intensity</th>
<th>Height of IVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Homogeneous and bright white</td>
<td>Clear</td>
<td>Hyperintense and isointense to CSF and AF</td>
<td>Normal</td>
</tr>
<tr>
<td>2</td>
<td>Nonhomogeneous with or without horizontal bands</td>
<td>Clear</td>
<td>Hyperintense and isointense to CSF</td>
<td>Normal</td>
</tr>
<tr>
<td>3</td>
<td>Nonhomogeneous and gray</td>
<td>Unclear</td>
<td>Intermediate to hypointense</td>
<td>Normal to slightly decreased</td>
</tr>
<tr>
<td>4</td>
<td>Nonhomogeneous and gray to black</td>
<td>Lost</td>
<td>Intermediate to hypointense</td>
<td>Normal to moderately decreased</td>
</tr>
<tr>
<td>5</td>
<td>Nonhomogeneous and black</td>
<td>Lost</td>
<td>Hypointense</td>
<td>Collapsed disk space</td>
</tr>
</tbody>
</table>

AF = Annulus fibrosus, NP = Nucleus pulposus, CSF = Cerebrospinal fluid