Agreement between transverse T2W and 3D-CISS sequences in the evaluation of spinal cord disease in dogs

Maria Oliveira DVM MRCVS[1], James Fraser Mcconnell BVM&S DVR DipECVDI CertSAM MRCVS[2], Thomas W. Maddox BVSc PhD CertVDI DipECVDI MRCVS[2], Daniel Sanchez-Masian DVM Dip ECVN MRCVS[2], Rita Gonçalves DVM MVM DipECVN FHEA MRCVS[2]

[1]Department of Neurology and Neurosurgery, Pride Veterinary Centre, Riverside Road, Derby, DE24 8HX, UK

[2]Institute of Veterinary Science, Small Animal Teaching Hospital, University of Liverpool, Leahurst, Chester High Road, Neston, CH64 7TE, UK

Email for correspondence: r.goncalves@liv.ac.uk

This work was performed at the Small Animal Teaching Hospital, University of Liverpool, Leahurst, Chester High Road, Neston CH64 7TE, UK.

This work was not supported by any grant, and none of the authors of this paper have conflict of interest that could influence the content of the paper.

ABSTRACT

The constructive interference in steady state (CISS) sequence has been widely used in human neuroimaging. It has been shown to be advantageous in the evaluation of intra- and extra-axial cystic abnormalities, arteriovenous and dysraphic malformations and disturbances of cerebrospinal fluid circulation. To assess the utility of this technique in small animals, interpretations based on this sequence were compared with those based on T2-weighted (T2W) sequences in 145 dogs that underwent MRI of the spine for suspected spinal cord disease. Two sets of images (T2W and CISS) were reviewed separately by three observers in random order and intra and inter observer agreements between both sequences were evaluated for several categorical variables. The overall agreement between T2W and CISS sequences was good. The highest agreement was observed for lesion diagnosis ($0.739<k<0.928$), treatment recommendation ($0.715<k<0.833$) and degree of spinal cord compression ($0.772<k<0.952$). The agreement for intramedullary intensity change ($0.192<k<0.332$) was lower compared to the other variables. Lesions that were predominantly characterised by focal hyperintense parenchymal changes on T2W were in some instances undetected on the CISS sequence whilst lesions consistent with spinal arachnoid diverticula on CISS sequences were occasionally missed on T2W. CISS enabled demonstration that lesions were directly affecting associated spinal nerves in some cases where T2W sequence was equivocal. Although CISS does not replace standard spin echo sequences, the results support inclusion of this sequence in small animal spinal MRI studies when subarachnoid diverticula or spinal nerve compression are suspected.

INTRODUCTION

Three-dimensional (3D) constructive interference in steady state (CISS) is a fully refocused steady-state gradient-echo magnetic resonance imaging (MRI) sequence. In human medicine, this sequence is now widely available and frequently used to investigate a varied range of pathologies, particularly when routine MRI sequences do not provide the desired anatomic information (Ramli 2001, Kulkarni 2011). The image contrast in CISS is determined by the T2/T1 ratio of the tissue with tissues that have long T2 relaxation times and short T1 relaxation times showing increased signal intensity. Because of high T2/T1 ratio, water has high signal on this sequence resulting in excellent contrast between cerebrospinal fluid (CSF) and other structures (Chavhan and others 2008). The other tissues have poor contrast though, and the gray-white differentiation is also not well visualized (Chavhan and others 2008). Other advantages are the high signal-to-noise ratio (Scheffler 2003) and better contrast-to-noise ratio (Roser and others 2008) relative to other routine imaging sequences (Scheffler and others 2003). Furthermore, CISS images do not present significant susceptibility, motion, or flow-related artifacts (Gonçalves and Amaral 2011).

In people CISS is mainly used to image the central nervous system, but it is also used to assess the abdomen, the musculoskeletal system and the breast. It has also become the sequence of choice for evaluating the cranial nerves in human patients due to its high spatial and contrast resolution between the cerebrospinal fluid and the nerve (Gonçalves and Amaral 2011, Kulkarni 2011, Besta and others 2016, Tsutsumi and others 2016). The use of CISS for visualisation of spinal abnormalities is a more recent development. It has been shown to be advantageous in the evaluation of intra- and extra-axial cystic abnormalities, arteriovenous and dysraphic malformations and disturbances of CSF circulation (Chavhan and others 2008, Gonçalves and Amaral 2011, Kulkarni 2011). It has also been shown to better characterise syringomyelia in humans, as it delineated the extension of the syrinx better, had higher contrast-to-noise ratio and less CSF flow artifact and allowed better visualisation of septation and communication between the syringomyelia cavities (Hirai and others 2000).

To our knowledge, there have been no clinical reports of the use of CISS sequences in dogs. The purpose of this retrospective cross-sectional study was to evaluate the usefulness of CISS in various spinal pathologies and to discuss its potential role in the canine spinal imaging protocol.

Materials and Methods

The imaging database of the Small Animal Teaching Hospital of the University of Liverpool was searched from October 2013 to June 2015. All MRI studies of the spine performed in dogs during this period were included if both T2-weighted (T2W) and 3D-CISS sequences in the transverse plane were available for review. All studies were performed with a 1.0 Tesla MRI scanner (Siemens Magnetom, Erlangen, Germany), using an integrated spine coil or an extremity coil in small patients. Sequences were variable but included sagittal, dorsal (both used exclusively for localisation of the lesion) and transverse T2W and transverse CISS images for all patients. Imaging parameters varied based on patient size (T2W sequence: TR = 2750-6200ms, TE = 90-110ms, slice thickness 2.5–3.0 mm; CISS sequence: TR = ~6-6.8ms, TE = ~12-13.6ms, slice thickness 1-1.2mm).

Studies meeting the inclusion criteria were exported to a picture archiving and communication system for evaluation (Visbion PACS 4.1, Visbion Limited, Surrey, UK). The T2W and CISS images were randomized and reviewed by a neurology resident in training (MO), a board-certified veterinary neurologist (RG) and a board-certified veterinary radiologist (FM) who were unaware of the patient information. Evaluation was performed independently for each sequence. All images were interpreted with the same software (OsiriX 7.5, Pixmeo, Geneva, Switzerland), which allowed adjustment for brightness, contrast, window width, and magnification.

Images were evaluated for spinal cord lesions and the following was determined for each sequence: presence or absence of a notable lesion (the largest was considered if multiple lesions were present), extradural, intradural extramedullary or intradural intramedullary location, degree of spinal cord compression (1 = none, 2 = mild, 3 = moderate, 4 = severe), degree of spinal cord signal changes (1 = none, 2 = mild, 3 = moderate, 4 = severe) and whether the lesion was directly affecting an associated spinal nerve (i.e. if the spinal nerve was enlarged or directly compressed by the lesion). The number of slices in which the lesion could be visualised was recorded for each sequence. The reviewers recorded the most likely diagnosis for each sequence and the two reviewers with neurosurgical experience (MO and RG) also recorded their treatment recommendation based on the imaging findings: 0 = no surgery, 1 = equivocal for surgery, 2 = surgery. Image quality was evaluated for each study and graded as follows: 1 = poor, 2 = fair, 3 = good, 4 = excellent. Artifacts (motion artifacts and those induced by pulsatile CSF flow) were also documented for each study and ranked as 1= absent, 2 = mild, 3 = moderate and 4 = severe. The different categorical variables used in image evaluation are summarised in Table 1.

The medical records were reviewed and data regarding breed, sex, age, bodyweight and duration of the clinical signs (recorded as the number of days between manifestation of the initial clinical sign and MRI) was collected for each patient.

Statistical analysis

All statistical analyses were performed with SPSS (SPSS 22.0 for Windows, SPSS Inc, Chicago, Illinois, USA). Independent variables were generated from the signalment data and clinical records (breed, sex, age, weight and duration of clinical signs) and the results from the observers on reviewing the MRI studies. Descriptive statistics were calculated for each variable; normally distributed continuous data were summarised as means and standard deviation and non-normally distributed as medians and interquartile ranges. The distribution of continuous variables was assessed through graphical analysis and the Kolmogorov-Smirnov test. Categorical data for breed and diagnosis were amalgamated into appropriate groups and expressed as frequencies with 95% confidence intervals (95% CI) as applicable as there were too many categories with small numbers to allow analysis; groupings were made on the basis of clinical judgement. A total of 46 different dog breeds were grouped into Crossbreed, Dachshund, French Bulldog, Jack Russell Terrier, Labrador Retriever, Pug, ShihTzu, Spaniel breed, Staffordshire Bull Terrier, Toy breed and Other. Fifteen different diagnoses were grouped into 11: acute non-compressive nucleus pulposus extrusion (ANNPE), spinal arachnoid diverticula, bony malformations (including articular process malformation, osseous-associated cervical spondylomyelopathy, atlantoaaxial instability and vertebral stenosis secondary to hemivertebra), foraminal stenosis causing nerve root compression, ischaemic myelopathy, degenerative intervertebral disc disease (including intervertebral disc extrusion and protrusion), neoplasia, neuritis, Chiari-like malformation and syringomyelia or no abnormalities identified.

Intra- and inter-observer agreement of categorical variables was assessed through calculation of Cohen's Kappa (κ) for intra-observer (comparing T2W and CISS sequences) for two observers (comparing treatment recommendation) agreements or Fleiss' Kappa for three observer agreement (presence of lesion, lesion diagnosis, anatomical location and nerve root compression). The intra- and inter-observer agreement of ordinal variables (degree of spinal cord compression, spinal cord intensity change, image quality and image artifacts) was assessed through calculation of intraclass correlation coefficients (ICC, two-way single measure for absolute agreement) with 95% confidence intervals. An adaptation of the κ evaluation grid by Landis and Koch (Landis and Koch 1997) was used to qualify the level of inter- and intra-observer agreement (Appendix).

Responses of the three observers for image quality and image artifacts were summarised as medians and associations between these and independent variables (breed, sex, age, weight and duration of clinical signs and lesion location) were assessed through ordinal logistic regression analysis. Validity of the ordinal logistic regression assumption of proportional odds was tested with the test of parallel lines (Brant 1990). Any independent variable demonstrating some association with image quality or artifacts on preliminary univariable analysis (a P-value <0.25) was considered for inclusion in a multivariable model. For any variables having a pair-wise correlation coefficient of >0.70, only the variable with the smallest univariable P-value was selected for incorporation in the multivariable analysis. All multivariable models were constructed with a manual backwards stepwise removal approach; variables with Wald and likelihood ratio test $P<0.05$ were retained. Confounding factors were detected by assessing parameter estimates for substantial changes (>20%) following removal from the models. Model goodness of fit was assessed using the Pearson's chi-square statistic for the model.

Results

A total of 145 dogs were included in the study. Breeds included dachshunds (14), spaniel breeds (14), crossbreeds (14), pugs (11), toy breeds (9), French bulldogs (8), Jack Russell terriers (8), Staffordshire bull terriers (7), Labrador retrievers (6), Shih Tzus (6) and five or fewer of 31 other pedigree breeds (48). Neutered male dogs were most frequent (37.2%), followed by entire males (33.8%), neutered females (23.4%) and entire females (5.5%). The median age of dogs was 78 months (interquartile range [IQR]: 49.5-101.4) and median weight 12.0kg (IQR: 8.0-24.5). The median duration of clinical signs was 10 days (IQR: 3-42). The areas imaged included the cervical region in 52 cases, the thoracic region in 64 cases and the lumbar region in 29 cases. The diagnosis was confirmed through surgery on 70 cases (62 with degenerative intervertebral disease, 5 spinal arachnoid diverticula, 1 ANNPE, 1 atlantoaxial instability and 1 osseous-associated cervical spondylomyelopathy).

Intra- and inter-observer agreement for presence of lesion, lesion diagnosis and treatment recommendation

Kappa agreement and intraclass correlations coefficients for agreement are presented in Tables 2. There was variability in the intra-observer agreement between T2W and CISS sequences for presence of a lesion: substantial agreement was seen in observer 1, moderate agreement was seen in observer 2 and slight agreement was seen in observer 3. The inter-observer agreement for this categorical variable was fair for T2W and moderate for CISS sequence. Both T2W and CISS sequences were considered unremarkable in two studies interpreted by observer 1, in five studies evaluated by observer 2 and in one study assessed by observer 3. A lesion was detected on T2W images that was not perceived on corresponding CISS in two studies evaluated by observer 1, in five studies assessed by observer 2 and in eight studies interpreted by observer 3. The lesions that were undetected on the CISS sequences were predominantly characterised by focal hyperintense parenchymal changes on T2W that in some cases were associated with reduced volume of nucleus pulposus of the intervertebral disc ventral to the area of hyperintensity, consistent with ANNPE or ischaemic myelopathy (2/2 in observer 1, 3/5 in observer 2 and 6/8 in observer 3) (Fig 1). In the remaining studies, the abnormalities detected on T2W were interpreted as intramedullary neoplasia (1) by observer 2, foraminal stenosis (1) by observer 2 and 3 and intervertebral disc protrusion (1) by both observer 2 and 3. A lesion was detected on CISS that was unrecognized on corresponding T2W in two studies evaluated by observer 2 and in four studies interpreted by observer 3. On CISS sequences these lesions were mainly described as focal dilatations in the subarachnoid space filled with CSF compatible with arachnoid diverticulum (2/2 in observer 2 and 3/4 in observer 3) (Fig 2). In the remaining study evaluated by observer 3, the abnormalities in CISS images were consistent with foraminal stenosis. There was substantial to almost perfect intra-observer agreement between T2W and CISS sequences concerning imaging diagnosis and treatment recommendation for a given patient. The inter-observer agreement between these sequences was substantial for imaging diagnosis and moderate for treatment recommendation.

Intra- and inter-observer agreement for parenchymal change and spinal corf compression

Kappa agreement and intraclass correlations coefficients for agreement are presented in Table 3. There was slight to fair intra-observer agreement between both sequences for detection of spinal cord intensity changes. The inter-observer agreement for this variable was moderate to substantial for each sequence. A degree of spinal cord intensity changes was identified on T2W but undetected on CISS sequences in 47% (68/145) of the studies evaluated by observer 1, in 44% (64/145) of the studies assessed by observer 2 and in 31% (46/145) of the studies interpreted by observer 3.

There was substantial to almost perfect intra-observer agreement between T2W and CISS sequences regarding degree of spinal cord compression. The inter-observer agreement between sequences was substantial for this variable.

Intra- and inter-observer agreement for anatomical location and nerve root involvement

Kappa agreement and intraclass correlations coefficients for agreement are presented in Table 4. The intra-observer agreement concerning anatomical location was moderate to almost perfect between sequences and the inter-observer agreement for this categorical variable was substantial for both sequences. The intra-observer agreement for nerve root compression was moderate to substantial between T2W and CISS sequences and the inter-observer agreement was fair for each of them. CISS images clearly demonstrated that the lesion was directly affecting an associated nerve root while no root involvement was visible on corresponding T2W sequence in 21 studies evaluated by observer 1, in 6 studies assessed by observer 2 and in 24 studies interpreted by observer 3. This finding was predominantly observed in cases of disc herniation. With exception of one study interpreted by observer 3, all the remaining cases where nerve root involvement was detected on T2W images were also visible on corresponding CISS image.

Association of image quality and image artifacts with independent variables

The inter-observer agreement of image quality and artifact severity recorded by three observers reviewing the two MRI sequences are presented in table 5. On univariable analysis, age, duration of clinical signs, weight, breed and sex and lesion location showed some evidence of association ($P < 0.25$) with both image quality and image artifacts for the T2W and CISS sequences. However, on multivariable analysis only weight and lesion location remained in the final models (Table 6). Associations for weight were similar, with increasing weight being associated with higher image quality for T2W and CISS sequences ($P = 0.001$ and 0.03 respectively) and with decreasing severity of artifacts in the T2W and CISS sequences ($P = 0.01$ and 0.048 respectively). For the CISS sequences, location other than the thoracolumbar spine was associated with higher image quality ($P = 0.005$ for cervical and $P < 0.001$ for lumbar) and decreasing severity of artifacts ($P = 0.011$ for cervical and $P = 0.009$ for lumbar), whilst for the T2W sequence only location within the lumbar spine was associated with higher image quality ($P = 0.021$).

Discussion

Recent developments in MRI technology and pulse sequences have refined this imaging technique further. Newer sequences are being developed constantly and the CISS, with its high spatial and contrast resolution between the cerebrospinal fluid and the nervous tissue offers certain advantages in terms of anatomical detail (Ramli 2001, Gonçalves and Amaral 2011). Pulsatile cerebrospinal fluid flow is minimized by acquiring the sequence with flow compensation applied over each TR cycle, rather than over each echo as in the case of conventional compensation techniques (Ramli and others 2001). Turbulent flow, however, is not suppressed, with phase dispersion resulting in signal loss (Ramli and others 2001). Two data sets are acquired successively with alternating and non-alternating radio frequency pulses, which are subsequently combined in a maximum intensity projection to produce an image with excellent CSF-to-spinal cord (Casselman and others 1993, Ramli and others 2001). To our knowledge, there have been no clinical reports of CISS images in dogs although its use has been described for imaging of the cranial nerves in the horse (Gonçalves and others 2015). Here, we describe the use of this MR sequence, with respect to its utility for the diagnosis of different spinal cord pathologies.

In our study, the overall agreement between T2W and CISS sequences was good. The highest agreement was observed for lesion diagnosis, treatment recommendation and degree of spinal cord compression for both inter- and intra-observer assessments. The intra-observer agreement for spinal cord intensity changes was lower compared to the other variables and the inter-observer agreement for this variable was substantial to almost perfect for each sequence. In our study, spinal cord intensity changes were more likely to be detected on T2W images than on CISS images. Lesions that were predominantly characterised by focal hyperintense parenchymal changes on T2W consistent with ANNPE/ischaemic myelopathy were in some instances undetected on the CISS sequence by all three observers. The morphological characteristics of spinal cord abnormalities with ill-defined margins and variable signal characteristics on T2W images typical of these conditions can be difficult to evaluate on CISS sequences, due to the intrinsic lack of contrast between such structures.

On the other hand, a lesion consistent with spinal arachnoid diverticula on CISS sequences was in some occasions missed on T2W images by two observers. In people, the use of CISS was shown to be useful for characterising both intra- and extramedullary cystic abnormalities (Gonçalves and Amaral 2011). The increased sensitivity of this sequence is a consequence of the accentuation of the T2 values between CSF and abnormal structures due to the higher intrinsic resolution between neural structures, CSF and lesions surrounded by CSF. Consequently, the extremely thin wall of diverticulum or cystic structures, which may easily evade detection with conventional sequences can be resolved using CISS. In people, the anatomical information provided by CISS is of particular value in planning surgical interventions, most notably determining the spatial relationship between intradural extramedullary masses and affected spinal nerve roots (Zhu and others 2015) as well as in the management of intraaxial and extraaxial cystic abnormalities, dysraphic malformations and disturbances of cerebrospinal fluid circulation (Ramli and others 2001).

The inter-observer agreement for presence of a lesion varied between observers. This range of results is likely representative of the different levels of radiological experience and different backgrounds between the participating observers. Investigations on inter-observer variability, which have become more common in recent years, demonstrated that reading agreement can vary among tests in veterinary medicine, including MRI (Leclerc and others 2013).

In people, the CISS sequence has been shown to be highly sensitive and specific in the evaluation of lumbar disc herniations, especially for the extruded and sequestrated discs (Aydin and others 2011). Although this sequence cannot differentiate properly disc material from osteophytes, it easily revealed the site of injury and relationship of the herniated fragment to the dural sac, nerve roots and bulged discs (Ramli and others 2001). In our study, although the intra-observer agreement between T2W and CISS sequences for nerve root compression was moderate to substantial, in some studies the T2W was insufficient for detailed analysis of nerve roots, especially in cases of disc herniation. In contrast, the slice thickness of the CISS enabled the nerve roots to be identified in cases where conventional T2W sequence was equivocal (although in most cases this could not be further confirmed by surgery or post-mortem examination and false positive findings on the CISS sequence may have occurred). Due to the thin slices used by the CISS, neurographic images can be obtained, improving the spatial resolution and enabling good definition of the spinal and neuroforaminal canal dimensions and nerve root compression. Therefore, CISS sequence should be considered in cases of suspected radiculopathy not explained by routine sequences.

Our study did not investigate sensitivity or specificity of T2W or CISS sequence in the diagnosis of different pathologies. Our purpose was solely to evaluate agreement between those two sequences and identify possible discrepancies between them. An accuracy study was not possible because not all patients underwent surgery or had histopathology performed.

Weight and lesion location were the only factors affecting imaging quality in both sequences. Increasing weight was associated with higher image quality and with reduced severity of artifacts for both T2W and CISS sequences. For both sequences the thoracolumbar spine was associated with worse image quality. This likely reflects motion related artifact associated with respiration (particularly for the CISS sequence) and possibly because some parts of the spine in this region may be further from the coil, due to longer spinous processes.

Artifacts in CISS images are considered to be banding artifacts caused by susceptibility variation and/or motion artifact (Hashiguchi and others 2007, Ramli and others 2001). Additional reported disadvantages of CISS images include a long acquisition time and their lack of contrast between soft tissues and neuronal tissues. Generally, there is no marked difference in acquisition times between CISS and T2W sequences, with both typically in the order of 4-5 minutes; therefore, no significant time-saving is achieved using either sequence.

In conclusion, overall there was a good agreement between T2W and CISS images for evaluating spinal cord diseases in dogs. Acquiring additional CISS images after normal T2W images is likely to result in detection of otherwise occult spinal lesions in only a small proportion of patients, such as those with small lesions close to the CSF pathways. As for humans, this sequence showed to be beneficial in identification of arachnoid diverticulum, which may evade detection with conventional sequences. CISS can also allow visualisation in fine detail of the spinal nerves and may be able to identify structural abnormalities affecting these or compression of nerve roots. Although CISS does not replace standard spin echo sequences, the results support inclusion of this sequence in small animal spinal MRI studies, mainly when imaging cystic lesions or those thought to affect nerve roots.

References

AYDIN, H., KIZILGOZ, V. & HEKIMOGLU, B. (2011) Compared with the conventional MR Imaging, do the Constructive interference steady state sequence and diffusion weighted imaging aid in the diagnosis of lumbar disc hernias? Eurasian Journal of Medicine 43, 152-161

BESTA, R., SHANKAR, Y. U., KUMAR, A., RAJASEKHAR, E. & PRAKASH S. B. (2016) MRI 3D CISS - A Novel Imaging Modality in Diagnosing Trigeminal Neuralgia - A Review. Journal of Clinical and Diagnostic Research 10, ZE01-ZE03

BRANT, R. (1990) Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression Biometrics 46, 1171-1178

CASSELMAN, J. W., KUHWEIDE, R., DEIMLING, M., AMPE, W., DEHAENE, I. & MEEUS, L. (1993) Constructive interference in steady state-3DFT MR imaging of the inner ear and cerebellopontine angle. American Journal of Neuroradiology 14, 47–57

CHAVHAN, G. B, BABYN, P. S., JANKHARIA, B. G., CHENG, H. L. & SHROFF, M. M. (2008) Steady-state MR imaging sequences: physics, classification, and clinical applications. Radiographics 28, 1147-1160

GONÇALVES, F.G. & AMARAL, L. L. F. (2011) Constructive Interference in Steady State Imaging in the Central Nervous System. European Neurological Review 6, 138-142

GONÇALVES, R., MALALANA, F., MCCONNELL, J. F. & MADDOX, T. (2015) Anatomical study of cranial nerve emergence and skull foramina in the horse using magnetic resonance imaging and computed tomography. Veterinary Radiology & Ultrasound 56, 391-397

HASHIGUCHI, K., MORIOKA T., YOSHIDA, F., MIYAGI, Y., MIHARA, F., YOSHIURA, T., NAGATA, S. & SASAKI, T. (2007) Feasibility and limitation of constructive interference in steady-state (CISS) MR imaging in neonates with lumbosacral myeloschisis. Neuroradiology 49, 579-585

HIRAI, T., KOROGI, Y., SHIGEMATSU, Y., SUGAHARA, T., TAKAHASHI, M., USHIO, Y. & UEMURA, S. (2000) Evaluation of syringomyelia with three-dimensional constructive interference in a steady state (CISS) sequence. Journal of Magnetic Resonance Imaging 11, 120-126

JIA, J. M., GUO H., HUO, W. J., HU S. W., HE, F., SUN, X. D. & LIN, G. J. (2016) Preoperative Evaluation of Patients with Hemifacial Spasm by Three-dimensional Time-of-Flight (3D-TOF) and Three-dimensional Constructive Interference in Steady State (3D-CISS) Sequence. Clinical Neuroradiology 26, 431-438

KULKARNI, M. (2011) Constructive interference in steady-state/FIESTA-C clinical applications in neuroimaging. Journal of Medical Imaging and Radiation Oncology 55, 183-190

LANDIS, J. & KOCH, G. (1997) An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 33, 363–374

LECLERC, M. K., D'ANJOU, M. A., BLOND, L., CARMEL É. N., DENNIS, R., KRAFT, S. L., MATTHEWS, A. R. & PARENT, J.M. (2013) Interobserver agreement and diagnostic accuracy of brain magnetic resonance imaging in dogs. Journal of American Veterinary Medical Association 15, 1688-1695

RAMLI, N., COOPER, A. & JASPAN, T. (2001) High resolution CISS imaging of the spine. British Journal Radiology 74, 862–873

ROSER, F., EBNER, F.H., DANZ, S., RIETHER, F., RITZ, R., DIETZ, K., NAEGELE, T., TATAGIBA, M.S. (2008) Three-dimensional constructive interference in steady-state magnetic resonance imaging in syringomyelia: advantages over conventional imaging. Journal of Neurosurgery Spine 8, 429-435

SCHEFFLER, K., LEHNHARDT, S. (2003) Principles and applications of balanced SSFP technique. European Radiology 13, 2409–2418

TSUTSUMI, S., ONO, H. & YASUMOTO, Y. (2016) Visualization of the olfactory nerve using constructive interference in steady state magnetic resonance imaging. Surgical Radiological Anatomy 39, 315-321

ZHU, Y. J., YING, G. Y., CHEN, A. Q., WANG, L. L., YU, D. F., ZHU, L. L., REN, Y. C., WANG, C., WU, P. C., YAO, Y., SHEN, F. & ZHANG, JM. (2015) Minimally invasive removal of lumbar intradural extramedullary lesions using the interlaminar approach. Neurosurgical focus 39, E10

Table 1. Summary of categorical variables used in evaluation of each set of images

Presence of lesion	Present/absent
Anatomical location	Extradural, intradural extramedullary or intramedullary and corresponding intervertebral disc space or vertebra
Nerve root compression	Present/absent
Lesion diagnosis	e.g. no abnormalities, nucleus pulposus extrusion, spinal arachnoid diverticula, bony malformations, foraminal stenosis, Chiari-like malformation and syringomyelia.
Treatment recommendation	0 (no surgery), 1 (equivocal), 2 (surgery)

Table 2. Kappa values (with 95% confidence intervals in parenthesis) for intra- and inter-observer agreement for presence of lesion, lesion diagnosis recorded and treatment recommendation by three observers reviewing two MRI sequences from 145 dogs. Cohen's kappa was calculated for intra-observer agreement and inter-observer agreement where there were only two observers (treatment recommendation), with Fleiss' kappa calculated for all remaining agreements with three observers.

	Intra-observer agreement for T2W versus CISS sequences			**Inter-observer agreement**	
	Observer 1	**Observer 2**	**Observer 3**	**T2W sequence**	**CISS sequence**
Presence of lesion	0.66 (0.221-1.000)	0.493 (0.205-0.777)	0.131 (0.001-0.423)	0.351 (0.257-0.445)	0.449 (0.354-0.544)
Lesion diagnosis	0.928 (0.867-0.989)	0.739 (0.649-0.829)	0.782 (0.698-0.866)	0.772 (0.724-0.821)	0.794 (0.743-0.845)
Treatment recommendation	-	0.833 (0.755-0.911)	0.715 (0.617-0.813)	0.477 (0.359-0.595)	0.500 (0.378-0.622)

Table 3. Intraclass correlation coefficients (with 95% confidence intervals in parenthesis) for intra- and inter-observer agreement for intramedullary change and spinal cord compression recorded by three observers reviewing two MRI sequences of 145 dogs.

	Intra-observer agreement for T2W versus CISS sequences			**Inter-observer agreement**	
	Observer 1	**Observer 2**	**Observer 3**	**T2W sequence**	**CISS sequence**
Spinal cord compression	0.952 (0.933-0.965)	0.862 (0.774-0.911)	0.772 (0.691-0.833)	0.694 (0.522-0.799)	0.711 (0.639-0.778)
Intramedullary intensity change	0.192 (-0.040-0.396)	0.390 (0.104-0.588)	0.332 (-0.036-0.586)	0.658 (0.548-0.744)	0.471 (0.372-0.567)

Table 4. Kappa values (with 95% confidence intervals in parenthesis) for intra- and inter-observer agreement for anatomical location and nerve root compression recorded by three observers reviewing two MRI sequences of 145 dogs.

	Intra-observer agreement for T2W versus CISS sequences			**Inter-observer agreement**	
	Observer 1	**Observer 2**	**Observer 3**	**T2W sequence**	**CISS sequence**
Anatomical location	0.952 (0.889-1.015)	0.621 (0.498-0.744)	0.596 (0.463-0.729)	0.609 (0.544-0.674)	0.745 (0.675-0.815)
Nerve root compression	0.491 (0.052-0.930)	0.794 (0.649-0.939)	0.608 (0.481-0.735)	0.349 (0.251-0.441)	0.342 (0.248-0.436)

Table 5. Intraclass correlation coefficients (with 95% confidence intervals in parenthesis) for inter-observer agreement of image quality and artifact severity recorded by three observers reviewing two MRI sequences of 145 dogs.

	Intra-observer agreement for T2W versus CISS sequences			**Inter-observer agreement**	
	Observer 1	**Observer 2**	**Observer 3**	**T2W sequence**	**CISS sequence**
Image quality	-	-	-	0.291 (0.185-0.400)	0.317 (0.201-0.432)
Artifact severity	-	-	-	0.095 (-0.010-0.214)	0.111 (-0.013-0.254)

Table 6. Multivariable ordinal logistic regression analysis of variables affecting image quality and severity of artifacts on review of two MRI sequences (T2W and CISS) of 145 dogs.

Outcome	Variable	Category	Odds Ratio	95% CI	P-value
T2W artifact severity	Weight	(kg)	0.96	(0.93-0.98)	0.01
CISS artifact severity	Weight	(kg)	0.97	(0.94-1.00)	0.048
	Lesion location	Thoracolumbar	(ref)	-	-
		Cervical	0.34	(0.15-0.78)	0.011
		Lumbar	0.27	0.10-0.72)	0.009
T2W image quality	Weight	(kg)	1.07	(1.03-1.11)	0.001
	Lesion location	Thoracolumbar	(ref)	-	-
		Cervical	1.44	(0.60-3.46)	0.42
		Lumbar	3.34	(1,20-9.31)	0.021
CISS image quality	Weight	(kg)	1.03	(1.00-1.06)	0.03
	Lesion location	Thoracolumbar	(ref)	-	-
		Cervical	3.38	(1.44-7.95)	0.005
		Lumbar	7.47	(2.78-20.1)	<0.001

Ref = reference category, CI = confidence intervals, *P*-Values are from the Wald Chi-squared test

Appendix Scale used to qualify the level of inter- and intra-observer agreement determined on the basis of the κ statistic.

κ	**Agreement**
>0.80	Almost perfect
>0.60 to 0.80	Substantial
>0.40 to 0.60	Moderate
>0.20 to 0.40	Fair
>0.00 to 0.20	Slight
0	Poor