Pferdeheilkunde – Equine Medicine 38 (2022) 6 (November/December) 500–514 DOI 10.21836/PEM20220601

Findings and interobserver agreement in radiography and ultrasonography of the vertebral column of a large population of normally performing horses

Brice Donati[1], Virginie Coudry[2], Jean-Marie Denoix[2], Stefanie Ohlerth[1], Marie Dittmann[3], Henning Richter[1], Michael Weishaupt[3] and José Suárez Sánchez-Andrade[1]

[1] Clinic of Diagnostic Imaging, Department of Clinical Diagnostics and Services, Vetsuisse Faculty, University of Zurich, Switzerland
[2] Centre d'Imagerie et de Recherche sur les Affections Locomotrices Equines (CIRALE), Ecole Nationale Vétérinaire d'Alfort, Université Paris-Est, Goustranville, France
[3] Equine Sports Medicine Unit, Equine Department, Vetsuisse Faculty, University of Zurich, Switzerland

Summary: Diagnostic imaging is a mainstay in the investigation of equine neck and back pain, but interpretation of radiographic and ultrasonographic findings in the spinal column of horses is not straightforward for a variety of reasons including individual anatomical variations, progressive degenerative nature of most pathologies and superimposition as well as technical limits. These issues are even more relevant in the context of examinations of apparently sound horses with absent, mild or unclear complaints, like it is often the case in pre-purchase examinations or cases of failure to meet expected performance. The first aim was to report on the spectrum, degree and location of first-line imaging findings in the spine of a large population of normally performing horses. Limited data is available about agreement of interpretation of equine vertebral column imaging by radiologists. The second aim of this prospective study was to determine interobserver agreement on radiographic and ultrasonographic diagnostic imaging findings in the vertebral column of the same population between multiple observers with longstanding experience in equine diagnostic imaging at two different institutions. Seventy-one horses randomly selected from a larger population of 250 normally performing horses participating in a swiss project on equine back health were examined at one referral center. Radiographic and ultrasonographic examinations were performed in a standardized fashion and images graded separately by two experienced radiologists at two different institutions. Focus was placed on osteoarthrosis of the synovial intervertebral articulations (SIVAs), impinging and overriding of the spinous processes (SPs) and spondylosis. Cohen's weighted kappa was calculated for each pathology, location and segment in each modality. Interobserver agreement was calculated for findings at specific locations, single pathologies and single grades. Most horses showed no changes and were allocated grades 0 at most locations. Few abnormalities were found throughout the vertebral column, with clusters of abnormalities of the SIVAs in the caudal cervical segment and cranial lumbar segment as well as a cumulation of changes at the dorsal spinous processes in the caudal thoracic segment. These coincide with previously reported predilection sites of imaging findings in symptomatic populations. Overall mean value of agreement for imaging findings was moderate (k= 0.7). Agreement was moderate for imaging findings regarding the synovial intervertebral articulations of the thoracolumbar spine in radiography (k=0.66) but weak in ultrasonography (k=0.58). There was moderate agreement in the imaging findings of the cervical spine in ultrasonography (k=0.61) as well as radiography (k=0.62). Strong agreement was found in the radiographic assessment of changes of the thoracolumbar spinous processes (k=0.80). Almost perfect agreement was found in the radiographic assessment of thoracolumbar spondylosis (k=0.95). Agreement between radiologists in detection and grading of pathologies of the equine vertebral column is weak to almost perfect depending on pathology and location. Our results confirm findings about distribution of specific pathologies found in other studies and support the importance of interpreting imaging findings along with clinical findings for definitive case management and decision making. Additional studies are needed for determination of the correlation of imaging findings among different modalities and correlation of diagnostic imaging with clinical findings.

Keywords: horse, back pain, ultrasonography, radiography, imaging findings, interobserver-agreement

Citation: Donati B., Coudry V., Denoix J.-M., Ohlerth S., Dittmann M., Richter H., Weishaupt M., Suárez Sánchez-Andrade J. (2022) Findings and interobserver agreement in radiography and ultrasonography of the vertebral column of a large population of normally performing horses. Pferdeheilkunde 38, 500–514; DOI 10.21836/PEM20220601

Correspondence: Dr. Brice Donati, Vetsuisse-Faculty University of Zurich, Winterthurerstrasse 260, 8057 Zurich, Switzerland; bdonati@vetclinics.uzh.ch

Submitted: May 15, 2022 | **Accepted:** July 27, 2022

Introduction

Pain originating from the vertebral column, more specifically the cervical spine and the thoracolumbar spine, is a common and a clinically relevant complaint in horses. Particularly sport horses' performance and wellbeing are affected by this type of pain (*Riccio* et al. 2018, *Haussler* and *Jeffcott* 2014). Beside a thorough clinical and orthopedic examination, radiographic and ultrasonographic examinations of the vertebral column are the mainstays in the diagnostic workup of pain originating at the vertebral column (*Burns* et al. 2018, *Gillen* et al. 2009, *Jeffcott* 1979, *Jeffcott* 1993, *Graf von Schweinitz* 1999, *Martin* and *Klide* 1999). There is a wide range of symptoms, clinical findings and pathologies present in horses affected by this kind of pain. The most commonly encountered pathologies include osteoarthrosis (OA) of the facet joints, from here on referred to as synovial intervertebral articulations (SIVAs), narrowed/overriding dorsal spinous processes (DSP) (kissing-spines-syndrome) and spondylosis (vertebral bodies changes, VB) (*Clayton* and *Stubbs* 2016, *Haussler* 1999, *Jeffcott* 1980, *Mayaki* et al. 2020, *Meehan* et al. 2009, *Haussler* et al. 2019, *Zimmerman* et al. 2011, *Dyson* 2011, *Wilsmann* et al. 2020).

For standardization of the interpretation of imaging findings, guidelines and scoring systems for the imaging features of the most common pathologies of the cervical and thoracolumbar spine have been described (*Meehan* et al. 2009, *Zimmerman* et al. 2011, *Crijns* and *Broeckx* 2021, *Espinosa-Mur* et al. 2020). Despite published scoring systems, the interpretation of radiography and ultrasonography remains challenging because of possibly irrelevant natural anatomic variations, the subtlety of lesions, their progressive and degenerative nature, the superimposition of neighboring structures, possible incidental projectional obliquity as well as technical limitations (*Clayton* and *Stubbs* 2016, *Erichsen* 2003, *Erichsen* et al. 2004, *Veraa* et al. 2020). Moreover, imaging findings have been found to not perfectly correlate with the clinical findings, which add to the difficulty of interpretation (*Haussler* et al. 1999, *Gundel* et al. 1997, *Gerlach* et al. 2018, *Lischer* et al. 2010, *Geiger* and *Gerhards* 2015, *Holmer* et al. 2007).

Ordinal grading systems are commonly used in equine medicine, for example for the assessment of lameness, heart murmurs and ataxia in horses, and for these systems moderate to substantial interobserver reliability and agreement have been reported (*Olsen* et al. 2014, *Fuller* et al. 2006, *Menzies-Gow* et al. 2010).

Regarding equine diagnostic imaging, a few studies have described the interobserver agreement for common pathologies or changes relevant to prepurchase examinations. (*Groth* et al. 2009, *Beccati* et al. 2013, *Hellige* et al. 2018). One study found interobserver agreement to be moderate to excellent for detecting and grading femoral trochlear ridge osteochondrosis (*Beccati* et al. 2013), whereas another study found the interobserver agreement to range from poor to excellent for evaluating a standard set of projections for prepurchase examinations (*Hellige* et al. 2018). The wide range of kappa values and variability in interobserver agreement was explained by the wide range of findings: large and easy to categorize findings resulted in higher agreement, whereas smaller and more subtle lesions resulted in poorer agreement.

To the authors knowledge, there are no studies determining the interobserver agreement of radiographic and ultrasonographic findings of vertebral columns of a large population of normally performing horses. Previous studies determined the interobserver agreement of the interpretation of vertebral column findings as a secondary result of the study without it being the focus. The authors reported for example fair agreement for presence/absence of articular process joint OA on radiographs of cadaver necks, or fair to substantial agreement for the grading of OA by board-certified radiologists using a modified three-grade system (*Wood* et al. 2014, *Espinosa-Mur* et al. 2020). Determination of intra- and interobserver agreement and reliability of each system is important, as reliable and reproducible interpretation as well as consistent classification when applying a specific scoring system are desirable as they are valuable tools for choosing therapies, defining prognosis and performing further research.

Therefore, the aims of the present study were 1) to describe the prevalent imaging findings in the vertebral column of a normally performing equine population and 2) to evaluate the interobserver agreement between two experienced observers when interpreting radiographic and ultrasonographic images of the vertebral column of normally performing horses. We hypothesized that 1) the distribution of pathologies would be similar as encountered in symptomatic populations, although of lower degree and 2) the interobserver agreement would be moderate for osteoarthritic changes in the SIVAs in both modalities and strong for spondylotic and DSPs radiographic changes.

Material and methods

Population

Horses were randomly selected out of an ongoing field study population for which recruitment has been previously described in detail (*Dittmann* et al. 2020). In brief, inclusion was published in the Swiss equitation press and announced to Swiss horse stable owners with the aim to investigate the back health of normally performing riding horses. Interested riders applied to the Section of Sports Medicine at the Equine Department, University of Zurich via an online questionnaire. Riders and their horses were included in the field study if the age of the rider was > 18 years, horses were exercised 2 hours per week at minimum, and riders as well as horses provided a good physical health according to the riders and/or horse owners. Riders had to confirm that they were riding their horse themselves at least two-thirds of the time. Seventy-one horses of this population were randomly selected for a 1-day work-up at the Equine Hospital of the University of Zurich where they underwent a clinical and lameness examination on hard ground at the walk and trot and where standard radiographic examination as well as ultrasonography of SIVAs was performed. Horses were excluded if a lameness grade > 2 (AAEP lameness grading scale) was diagnosed upon lameness examination on hard ground.

Image acquisition

All horses underwent the same standard imaging protocol and with the same dedicated equipment throughout the study, as previously described (*Audigié* et al. 2015). Radiographic examination was performed in the standing and sedated horse, and consisted of at least three laterolateral projections of the cervical spine with the neck supported in a standard neutral position as this has shown to be important for comparing images (*Berner*, *Brehm* and *Gerlach* 2012). Additionally, at least one laterolateral projection of the caudal segment of the thoracic spine and one laterolateral projection of the cranial segment of the lumbar spine were obtained. These last two projections were repeated with strong collimation centered at the level of the facets to increase conspicuity of the SIVAs. An aluminum wedge was used for compensation of large difference in tissue thickness at these sites. A high-power generator (Optimus 65, Philips Medical System, Switzerland) in combination with a digital radiography system (FDR D-EVO I, Fujifilm (Switzerland) AG) was used. Exposure factors are summarized in table 1 and were adjusted according to the anatomic region, the projection, and the size of the horse (exposure factor ranges: 75–115 kV, and 80–180 mAs). The cassette was placed in a rolling stand-alone holder perpendicular to the floor and as close as possible to the horse, at a focal distance of 115 cm from the x-ray tube. For the thoracolumbar spine a metal opaque marker was placed dorsal to the last thoracic spinous process, identifying the last thoracic vertebra by palpating the most caudal rib and following it to its dorsal aspect.

Ultrasonographic examination was performed transcutaneously as previously described (*Berg* et al. 2003, *Chope* 2008, *Head* 2014, *Denoix* 1999) by one board certified large animal radiologist using a macroconvex probe (2–6 MHz, Aloka ProSound Alpha 7, Hitachi, Switzerland). One representative transverse image and/or cineloop per location and laterality was stored. For the cervical region only the three most caudal articulations C5-T1 were examined ultrasonographically. Articulations C7-T1 were identified based on the position at the base of the neck medial to the musculature of the scapula and shoulder, as well as based on the dorsal location of the vertebral artery, as previously described (*Berg* et al. 2003, *Denoix* 2017). From there, the two more cranially located joints were examined. For examination of the thoracic and lumbar SIVAs, rough orientation was given by palpating the most caudal rib and following it proximally. Ultrasonographic examination of the SIVAs provided detailed orientation by identifying the difference of transition from thoracic vertebrae articulating with ribs and lumbar vertebrae showing transverse processes. Overlying soft tissue thickness and depth of the SIVAs allowed consistent identification of these structures in the T18-L4 segment.

Image interpretation

Images were reviewed at a later stage by two board certified radiologists, one from the Clinic for Diagnostic Imaging, Vetsuisse Faculty, University of Zurich (JS) and one from CIRALE (JMD). In order to keep radiologists' individual unaltered judgement, no training set was created nor were training sessions performed but consensus was previously reached for the use of grading systems based on published studies as well as personal experience and preferences. Evaluation of images and grading of changes was carried out separately by the two radiologists. Where interpretation was not possible because of missing images or due to insufficient image quality, no grade was assigned. The evaluation of ultrasonographic images by the in-house board-certified radiologist was performed blindly months later.

Interpretation of radiography

SIVAs were evaluated for OA changes based on four criteria: joint size, orientation and shape of the joint space, bone opacity and periarticular bone remodelling. Considering previously described grading systems as well as personal experience and preference, authors agreed on a grading from 0 to 4 (0 = no changes, 1 = mild, 2 = moderate, 3 = marked, 4 = severe changes) for each joint (*Down* and *Henson* 2009, *Girodroux* et al. 2009, *Crijns* and *Broeckx* 2021). Due to superimposition of the left and right SIVAs on laterolateral radiographs, the total grade referred to the most severe changes of each segment.

Radiographic findings at the spinous processes were graded 0 to 4 based on established systems taking into account interspinous space width, bone opacity and radiolucencies of the spinous processes as well as remodelling of the spinous processes (1 = narrowing of the interspinous space with mild sclerosis of the cortical margins of the spinous processes, 2 = showing loss of the interspinous space with moderate sclerosis of the cortical margins of the spinous processes, grade 3 = severe sclerosis of the cortical margins of the spinous processes, caused in part by transverse thickening, or radiolucent areas and 4 = severe sclerosis of the cortical margins, osteolysis and change in shape of the spinous processes) (*Zimmerman* et al. 2012 and 2011, *Denoix* and *Dyson* 2003).

Smooth new bone formation visible at the ventral aspect of the vertebral endplates and ventral to the intervertebral spaces were graded from 0 to 4 (0 = no changes, 1 = ostephyte from one vertebral body, 2 = ostephytes from 2 adjacent vertebral bodies, not meeting, 3 = osteophytes from two adjacent vertebral bodies, meeting, 4 = complete bridging of intervertebral space), a modification of a previously described grading system for spondylosis lesions (*Meehan* et al. 2009).

Interpretation of ultrasonography

Ultrasonographic loops and still images were reviewed for OA at the SIVAs based on 4 criteria: presence of osteophytes/enthesophytes, synovial effusion, echogenic foci in the synovial fluid, and visibility of the joint space. Similarly, to radiography, a grade ranging from 0 (normal) to 4 (severe changes) was allotted to each joint (*Blättler* 2020).

Table 1 Exposure factors used for radiographic examination of the vertebral column. | *Belichtungstabelle für Röntgenuntersuchung der Wirbelsäule.*

Region	kV	mAs	Grid	Filter
Cervical				
C1-C3	77	20	Yes	No
C3-C5	81	22	Yes	No
C5-C7	85	28	Yes	No
Thoracolumbar overview				
T13-T18	100	100	Yes	Yes
T18-L5	115	125	Yes	Yes
Thoracolumbar focused				
T13-T18	100	100	Yes	No
T18-L5	115	125	Yes	No

Statistical analysis

Data were stored in Microsoft Excel. Results were reported as the mean ± standard deviation (SD) for variables with parametric distribution and median (range) for variables with non-parametric distributions. The median grades given by the two radiologists were used for further descriptive analysis of severity and distribution of findings. Interobserver agreement between the two board certified radiologists was assessed by calculating Cohen's weighted kappa for their gradings at each location separately. Kappa values were interpreted as described previously (0–0.2 no agreement, 0.21–0.39 minimal, 0.4–0.59 weak, 0.6–0.79 moderate, 0.8–0.9 strong, > 0.9 almost perfect) (Koo and *Li* 2016, *McHugh* 2012). All analyses were performed in R Studio (version 1.1.442, 2019).

Results

Study population

Thirty-seven geldings (52.1 %), 32 mares (45.1 %) and 2 stallions (2.8 %) with a mean age of 10.2 ± 2.7 years were included in the study. The horses were primarily used for show jumping (33.8 %), dressage (29.6 %), leisure (16.8 %) and eventing (9.8 %). The remaining horses (9.8 %) were used for western or endurance riding. Warmbloods accounted for 74.6 % of the study population. The mean withers height was 164 ± 7.5 cm, and the mean body weight was 548.7 ± 61.3 kg.

Cervical segment

Radiography

Laterolateral radiographs of the cervical spine of 71 horses were acquired and evaluated. In 53 horses, all of the cervical SIVAs were evaluated. In 18 horses the most caudal SIVAs (C7-T1) could not be evaluated due to underexposure and/or other technical limits. This equates to a total number of 408 evaluated SIVAs, consisting of 71 × C2-3, 71 × C3-4, 71 × C4-5, 71 × C5-6, 711 × C6-7 and 53 × C7-T1. The majority of locations were graded as having no or mild pathologic changes (363/408 mean grades ≤ 1). Of the remaining 45 locations with higher degree changes, 32 were at C6-7 and 7 at C5–6. Detailed numbers of allocated grades at the different SIVAs are depicted in figure 1A and summarized in Table 2.

Interobserver agreement for radiographic assessment of the SIVAs in the cervical segment (C5-T1) was moderate (mean kappa value: 0.62, range: 0.41–0.75). There were 342/408 cases of agreement, 54/408 cases of disagreement by one grade, 11/408 cases of disagreement by 2 grades and 1/408 out cases of disagreement by two grades; an overview is shown in table 3A.

Ultrasonography

Ultrasonographic examination of the caudal cervical SIVAs yielded images of insufficient quality at all locations in 7 Horses. Of 64 horses which generated transverse images and/or loops suitable for evaluation, 60 yielded quality images at all examined locations. One of 64 showed insufficient quality at the left C7-T1 SIVA and 4/64 at the right C7-T1 SIVA. This equates to a total number of 379 evaluated SIVAs, consisting of 64 left C5-C6, 64 right C5-C6, 64 left C6-C7, 64 right C6-C7, 63 left C7-T1 and 60 right C7-T1. The majority of locations were graded as having no or mild pathologic changes (341/379 mean grades ≤ 1). Of the remaining 36 locations with higher degree changes, only 2 involved the C5-C6 articulations, whereas 21/36 involved C6-C7 SIVAs and 13/36 C7-T1 SIVAs. Detailed numbers of allocated grades at the different SIVAs are depicted in figure 1B and summarized in table 2.

The overall interobserver agreement was moderate (mean kappa value: 0.61, range: 0.50–0.77). There were 284/379 cases of agreement, 88/379 cases of disagreement by 1 grade and 7/379 cases of disagreement by two grades; an overview is shown in table 3B.

Thoracolumbar segment

Radiography

Laterolateral radiographs of the thoracolumbar (T12–L5) spine of 71 horses were acquired and evaluated. Regarding

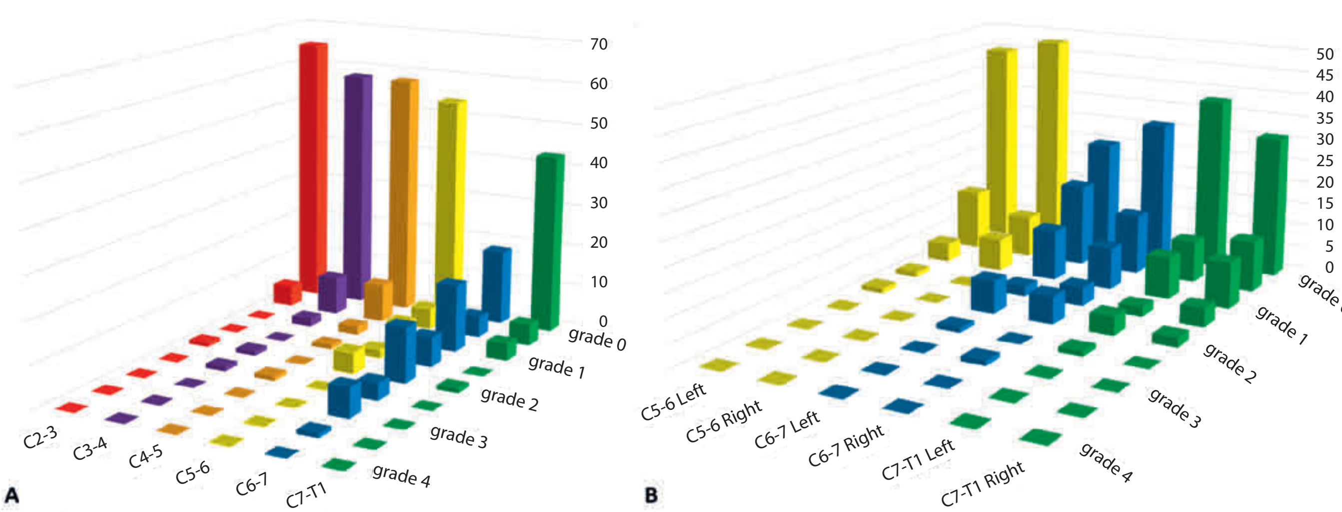

Fig. 1 Distribution of graded for the cervical SIVAs, in radiography (A) and ultrasonography (B). | *Verteilung von Befunden an den synovialen Intervertebralgelenke, in Röntgen (A) und Ultraschall (B).*

Table 2 Mean grades for findings at single locations by two observers. | *Mittelwert der Bewertung von Befunden an einzelnen Lokalisationen zwischen zwei Untersucher.*

Region	Modality	Structure	Location	N	grade 0	grade 0.5	grade 1	grade 1.5	grade 2	grade 2.5	grade 3	grade 3.5	grade 4
Cervical	Rx	SIVA	C2-3	71	65	5	0	0	1	0	0	0	0
Cervical	Rx	SIVA	C3-4	71	58	9	2	0	1	1	0	0	0
Cervical	Rx	SIVA	C4-5	71	58	9	2	1	0	1	0	0	0
Cervical	Rx	SIVA	C5-6	71	54	5	5	2	5	0	0	0	0
Cervical	Rx	SIVA	C6-7	71	18	5	16	7	13	4	7	1	0
Cervical	Rx	SIVA	C7-T1	53	43	5	4	0	1	0	0	0	0
Cervical	Us	SIVA	C5-6 LEFT	64	45	13	4	1	1	0	0	0	0
Cervical	Us	SIVA	C5-6 RIGHT	64	48	9	7	0	0	0	0	0	0
Cervical	Us	SIVA	C6-7 LEFT	64	25	18	11	2	7	1	0	0	0
Cervical	Us	SIVA	C6-7 RIGHT	64	31	13	9	4	6	0	1	0	0
Cervical	Us	SIVA	C7-T1 LEFT	63	38	9	9	2	4	1	0	0	0
Cervical	Us	SIVA	C7-T1 RIGHT	60	31	11	10	4	2	0	0	0	0
Thoracolumbar	Rx	SIVA	T12-13	71	70	1	0	0	0	0	0	0	0
Thoracolumbar	Rx	SIVA	T13-14	71	68	0	1	0	1	0	0	0	0
Thoracolumbar	Rx	SIVA	T14-15	71	69	0	1	0	0	1	0	0	0
Thoracolumbar	Rx	SIVA	T15-16	71	63	2	4	0	1	1	0	0	0
Thoracolumbar	Rx	SIVA	T16-17	71	53	6	7	1	3	0	1	0	0
Thoracolumbar	Rx	SIVA	T17-18	71	49	9	4	1	5	2	1	0	0
Thoracolumbar	Rx	SIVA	T18-L1	71	41	14	5	2	5	2	2	0	0
Thoracolumbar	Rx	SIVA	L1-L2	71	37	13	10	2	6	3	0	0	0
Thoracolumbar	Rx	SIVA	L2-L3	71	39	13	5	1	10	2	1	0	0
Thoracolumbar	Rx	SIVA	L3-4	71	56	4	4	0	4	2	0	0	1
Thoracolumbar	Rx	SIVA	L4-5	71	69	1	1	0	0	0	0	0	0
Thoracolumbar	Us	SIVA	T18-L1 LEFT	67	46	13	2	3	2	0	1	0	0
Thoracolumbar	Us	SIVA	T18-L1 IGHT	67	45	9	8	3	2	0	0	0	0
Thoracolumbar	Us	SIVA	L1-2 LEFT	69	45	20	1	1	2	0	0	0	0

Region	Modality	Structure	Location	N	grade 0	grade 0.5	grade 1	grade 1.5	grade 2	grade 2.5	grade 3	grade 3.5	grade 4
Thoracolumbar	Us	SIVA	L1-2 RIGHT	69	39	15	8	2	3	2	0	0	0
Thoracolumbar	Us	SIVA	L2-3 LEFT	69	43	15	7	1	1	2	0	0	0
Thoracolumbar	Us	SIVA	L2-3 RIGHT	68	47	10	5	3	3	0	0	0	0
Thoracolumbar	Us	SIVA	L3-4 LEFT	68	44	9	8	4	2	0	1	0	0
Thoracolumbar	Us	SIVA	L3-4 RIGHT	67	42	13	4	1	4	3	0	0	0
Thoracolumbar	Us	SIVA	L4-5 LEFT	4	2	1	0	1	0	0	0	0	0
Thoracolumbar	Us	SIVA	L4-5 RIGHT	4	1	2	0	1	0	0	0	0	0
Thoracolumbar	Rx	SP	T12-13	71	66	1	0	0	1	0	2	0	1
Thoracolumbar	Rx	SP	T13-14	71	58	3	3	1	3	0	2	0	1
Thoracolumbar	Rx	SP	T14-15	71	45	7	7	1	7	0	3	0	1
Thoracolumbar	Rx	SP	T15-16	71	38	7	12	3	6	1	3	0	1
Thoracolumbar	Rx	SP	T16-17	71	40	2	11	4	7	2	4	0	1
Thoracolumbar	Rx	SP	T17-18	71	43	9	6	1	5	1	5	0	1
Thoracolumbar	Rx	SP	T18-L1	71	58	2	2	1	5	1	2	0	0
Thoracolumbar	Rx	SP	L1-L2	71	64	2	2	0	3	0	0	0	0
Thoracolumbar	Rx	SP	L2-L3	71	67	1	1	0	2	0	0	0	0
Thoracolumbar	Rx	SP	L3-L4	71	66	1	3	0	1	0	0	0	0
Thoracolumbar	Rx	SP	L4-L5	71	53	8	4	0	3	0	3	0	0
Thoracolumbar	Rx	VB	T12_13	71	64	2	0	0	1	0	3	0	1
Thoracolumbar	Rx	VB	T13_14	71	67	1	0	1	0	0	1	0	1
Thoracolumbar	Rx	VB	T14_15	70	68	0	0	0	0	0	1	0	1
Thoracolumbar	Rx	VB	T15_16	71	70	0	0	0	0	0	0	0	1
Thoracolumbar	Rx	VB	T16_17	71	0	0	0	0	0	0	0	0	0
Thoracolumbar	Rx	VB	T17_18	71	0	0	0	0	0	0	0	0	0
Thoracolumbar	Rx	VB	T18_L1	71	0	0	0	0	0	0	0	0	0
Thoracolumbar	Rx	VB	L1_2	71	0	0	0	0	0	0	0	0	0
Thoracolumbar	Rx	VB	L2_3	71	0	0	0	0	0	0	0	0	0
Thoracolumbar	Rx	VB	L3_4	71	0	0	0	0	0	0	0	0	0
Thoracolumbar	Rx	VB	L4_5	71	0	0	0	0	0	0	0	0	0

SIVAs, all depicted SIVAs in the segment (T12-T13 through L4-L5) were evaluated in all horses. This equates to a total number of 781 evaluated SIVAs. The majority of locations were graded as having no or mild pathologic changes (720/781 mean grades ≤ 1). Of the remaining 61 locations, higher degree changes (mean grades > 1) concentrated in the cranial lumbar area; 14 were at L2-L3, 11 at L1-L2, 11 at T18-L1, 9 at T17–18, 7 at L3-L4, 5 at T16-T17, 2 at T15-T16, 1 at T14–15 and 1 at T13-T14. Detailed numbers of allocated grades at the different SIVAs are depicted in figure 2A and summarized in table 2.

Interobserver agreement for radiographic assessment of the thoracolumbar SIVAs was moderate (mean kappa value: 0.66, range: 0.49–0.81), with 669/781 cases of agreement, 82/781 cases of disagreement by 1 grade, 29/781 cases of disagreement by 2 grades and 1/781 case of disagreement by 3 grades; an overview is shown in table 3C.

Regarding spinous processes, all depicted locations in the segment were assessed in all horses (T12-T13 through L4-L5), resulting in a total number of 781 evaluated locations. The majority of locations were graded as having no or mild pathologic changes (692/781 mean grades ≤ 1). Of the remaining 89 locations, higher degree changes (mean grades > 1) concentrated in the caudal thoracic area, 18 at T16-T17, 14 at T15-T16, 13 at T17-T18, 12 at T14-T15, 9 at T18-L1, 7 at T13-T14, 6 at L4-L5, 4 at T12-T13, 3 at L1-L2, 2 at L2-L3 and 1 at L3-L4. Detailed numbers of allocated grades at the different locations are depicted in figure 2B and summarized in table 2.

Interobserver agreement for radiographic assessment of the spinous processes in the thoracolumbar region was strong (mean kappa value: 0.80, range: 0.36–0.97) with 698/781 cases of agreement, 59/781 cases of disagreement by 1 grade, 24 cases of disagreement by 2 grades and no cases of disagreement by 3 grades or more; an overview is shown in table 3D.

Regarding spondylotic changes in the thoracolumbar segment (T12-L5), only 283 locations between T12 and T16 were evaluated: T12-T13 in 71 horses, T13-T14 in 71, T14-T15 in 70 and T15-T16 in 71. The majority of locations were graded as having no or mild changes (272/283 mean grades ≤ 1). Of the remaining 11 locations with higher degree changes (>1), 5 were at T12-T13, 3 at T13-T14, 2 at T14-T15 and 1 at T15-T16. Detailed numbers of allocated grades at the different locations are summarized in table 2.

Interobserver agreement for radiographic assessment of spondylotic changes at the vertebral bodies in the thoracolumbar segment was almost perfect (mean kappa value: 0.95, range:

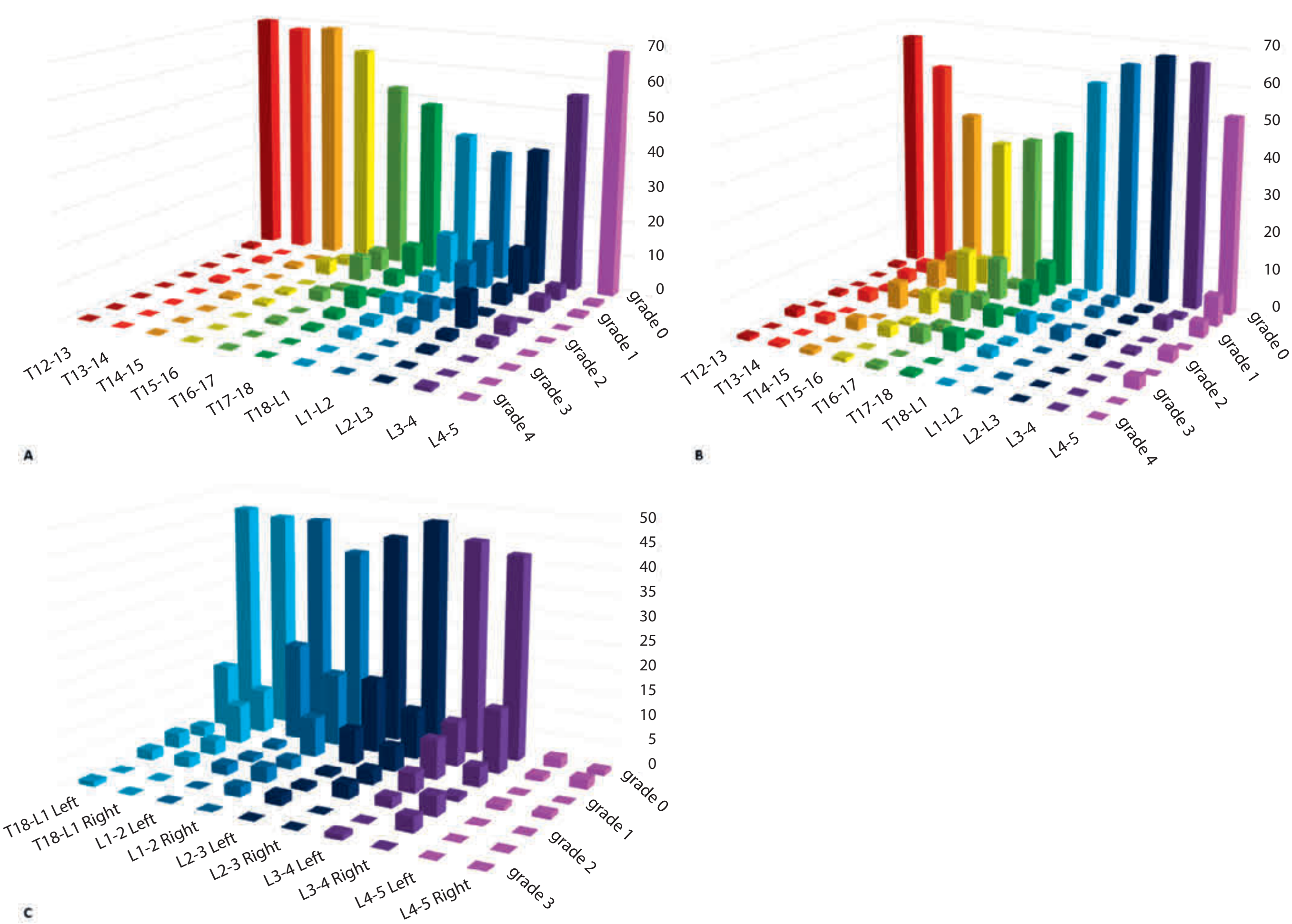

Fig. 2 Distribution of gradings for thoracolumbar pathology; SIVAs in radiography (A), SP in radiography (B) and SIVAs in ultrasonography (C). | *Verteilung von Befunden an der Brust- und Lendenwirbelsäule; Röntgenbefunde an den synovialen Intervertebralgelenke (A), Röntgenbefunde an den Dornfortsätze (B) und Ultraschallbefunde an den synovialen Intervertebralgelenke (C).*

0.83–1), with 279/283 cases of agreement, 3/283 cases of disagreement by one grade, one case of disagreement by 3 grades; an overview is shown in table 3E.

Ultrasonography

Ultrasonographic examination of thoracolumbar SIVAs in the T18-L5 segment was performed in 71 horses. In two horses the images were of insufficient quality at all locations, whereas in the remaining 69 horses the examination generated transverse images and/or loops suitable for evaluation at least at one location; 67 at both left and right T18-L1, 69 at both left and right L1-L2, 69 at the left L2-L3 and 68 at the right L2-L3, 68 at the left L3-L4 and 67 at the right L3-L4 and 4 at both the right and left L4-L5. This equates to a total number of 552 evaluated locations. The majority of locations were graded as

Table 3 Crosstables for overview of agreement between two observers for finding in the spinal column of horses. | *Übersicht der Übereinstimmung der Bewertung von Befunden an der Wirbelsäule zwischen zwei Untersucher.*

Cervical SIVA Rx		Cirale				
N = 408		0	1	2	3	4
Zurich	0	**296**	7	2	0	0
	1	31	**19**	3	0	0
	2	8	6	**20**	4	0
	3	1	1	2	**7**	0
	4	0	0	0	1	**0**
				cases of agreement:		342
				cases of disagreement by 1 grade:		54
				cases of disagreement by 2 grades:		11
A				cases of disagreement by 3 grades:		1

Cervical SIVA Us		Cirale				
N = 379		0	1	2	3	4
Zurich	0	**218**	23	6	0	0
	1	50	**43**	9	0	0
	2	1	4	**22**	2	0
	3	0	0	0	**1**	0
	4	0	0	0	0	**0**
				cases of agreement:		284
				cases of disagreement by 1 grade:		88
				cases of disagreement by 2 grades:		7
B				cases of disagreement by 3 grades:		0

Thoracolumbar SIVA Rx		Cirale				
N = 781		0	1	2	3	4
Zurich	0	**615**	31	11	1	0
	1	31	**16**	2	2	0
	2	16	5	**33**	8	0
	3	0	0	5	**5**	0
	4	0	0	0	0	**0**
				cases of agreement:		669
				cases of disagreement by 1 grade:		82
				cases of disagreement by 2 grades:		29
C				cases of disagreement by 3 grades:		1

Thoracolumbar DSP RX		Cirale				
N = 781		0	1	2	3	4
Zurich	0	**598**	24	15	0	0
	1	19	**28**	6	0	0
	2	8	5	**43**	1	1
	3	0	0	4	**23**	0
	4	0	0	0	0	**6**
				cases of agreement:		698
				cases of disagreement by 1 grade:		59
				cases of disagreement by 2 grades:		24
D				cases of disagreement by 3 grades:		0

Thoracolumbar VB RX		Cirale				
N = 283		0	1	2	3	4
Zurich	0	**269**	2	0	0	0
	1	1	**0**	0	0	0
	2	0	0	**1**	0	0
	3	1	0	0	**5**	0
	4	0	0	0	0	**4**
				cases of agreement:		776
				cases of disagreement by 1 grade:		3
				cases of disagreement by 2 grades:		0
E				cases of disagreement by 3 grades:		1

Thoracolumbar SIVA US		Cirale				
N = 552		0	1	2	3	4
Zurich	0	**357**	13	1	0	0
	1	92	**30**	5	0	0
	2	12	13	**19**	0	0
	3	1	0	7	**2**	0
	4	0	0	0	0	**0**
				cases of agreement:		408
				cases of disagreement by 1 grade:		130
				cases of disagreement by 2 grades:		13
F				cases of disagreement by 3 grades:		1

having no or mild changes (504/552 mean grades ≤ 1). Detailed numbers of allocated grades at the different locations are depicted in figure 2C and summarized in table 2.

At the left L4-L5 SIVA both observers agreed there were no changes in all horses.

Interobserver agreement for ultrasonographic assessment of thoracolumbar SIVAs was moderate (mean kappa value: 0.58, range: 0.39–0.77). There were 408/552 cases of agreement, 130/552 cases of disagreement by 1 grade, 13/552 cases of disagreement by 2 grades and 1/552 case of disagreement by 3 grades; an overview is shown in table 3F.

Discussion

We report on the distribution and degree of radiographic and ultrasonographic changes of the vertebral column in a normally performing equine population. Distribution of specific pathologies coincides with the reported occurrence of specific changes at some predilections sites especially regarding OA of the SIVAs and DSP changes with most of the cervical articular changes consistently being described clustered at the neck base, in particular at C6-C7, and most dorsal spinous process changes at the thoracolumbar transition and cranial lumbar spine (*Holmer* 2005, *Audigié* et al. 2015, *Espinosa-Mur* et al. 2021, *Dämmrich* et al. 1993, *Fritsche* et al. 2020). Despite the paucity of spondylotic changes found in this population the distribution matches the published data with majority of spondylosis being found at the T11-T13 segment (*Meehan* et al. 2009).

Regarding severity, most of the changes were of low degree, while some horses showed moderate or even severe findings at random locations without an apparent negative effect on performance. Beside showing a lameness grade ≤ 2 at clinical examination, the definition of soundness of the included horses was also relying on the owner's judgement, which has previously been shown to be incongruent with veterinary examinations (*Müller-Quirin* et al. 2020). Therefore, some symptoms or clinical complaints, possibly originating from the vertebral column, could have been overlooked, in particular if only arising under some circumstances like after exercise, while doing complex movement sequences or on specific soil conditions. It was not the purpose of the study to trivialize imaging findings described in this paper as not being clinically relevant, however clinicians should be aware about the risk of a satisfaction-of-search-bias. This should prompt the clinician to investigate further if there is a suspicion that the presenting clinical complaint can't be totally attributed to the imaging finding. Also, findings such as those found in this study population may be regarded as concurrent to other complaints or secondary to other pathologies.

Comparing ultrasonographic and radiographic changes at the SIVAs shows that changes seem to be milder or graded less severely in ultrasonography. The relatively less pronounced ultrasonographic findings at the SIVAs compared to radiographic changes are in agreement with published data and the opinion of the authors, which is that, only if changes are diffuse or superficially located, they can be assessed by ultrasonography (*Audigié* et al. 2015). As previously described, osteoarthritic changes in the caudal cervical segment often affect the medial aspect of the joints (*Lindgren* et al. 2020, *Gough* et al. 2020), which cannot be assessed with ultrasound, possibly resulting in a radiographically severely abnormal SIVAs being graded as only mildly affected by changes in ultrasonography.

The second aim of this study was to report on the interobserver agreement in evaluating radiographic and ultrasonographic findings in the vertebral column of a large normally performing equine population. We hypothesized that the agreement between two board certified experienced large animal radiologists for grading imaging findings of the equine vertebral column would be moderate for OA changes at the SIVAs and strong for spondylotic and DSPs changes. Our results predominantly support the hypotheses, although agreement was found to be lower at some single locations and for certain modalities. Interpretation of imaging findings has been found to be far from consistent between radiologists and subjective to a certain degree, in particular for mild, small or equivocal changes, as well as for changes related to chronic, progressive and/or degenerative pathologies (*Hellige* et al. 2018, *Jackson* et al. 2014). Allocating a continuous variable such as progressive degenerative changes to the discrete values of an imaging grading system represents a source of variability in interpretation. In non-straightforward cases the experience and subjective perception of the assessor has an important influence. These issues are even more pronounced when a horse is presented and examined for reasons unrelated to back pain or with unspecific clinical signs. This population's characteristics in this study exacerbate difficulties in interpretation given the low incidence and severity of changes in a normally performing equine population. A particular difficulty in these cases is to decide whether mild changes are indeed abnormal changes, just still within normal limits or anatomical variants.

Subjective interpretation is the result of training as well as experience and is an intrinsic feature of many medical disciplines. In radiology, this subjective interpretation particularly applies to pathologies such as those that are investigated in this paper which are chronic, progressive and degenerative in nature, and abnormal imaging features cannot always be defined by reference ranges and yes/no criteria. Detailed grading systems based on specific features, rather than just categories of severity grades, often are developed to partially overcome the limitations dictated by subjective interpretation. However also this type of assessment can have limitations when changes vary along a continuum as it is the case in degenerative pathologies, because ordinal grading systems require categorization of severity according to predetermined criteria or definitions (*Wise* et al. 2021).

Regarding the SIVAs, the overall interobserver agreement was higher in the cervical region than in the thoracic and lumbar regions for both modalities. This can be explained by the different thickness of the region of interest. This affects both ultrasonography and radiography, by increasing depth and thus reducing the number of echoes returning to the transducer due to attenuation of the ultrasound beam and by increasing superimposition respectively.

Observers performed with minimally higher agreement when interpreting radiography compared to ultrasonography of the SIVAs independent of the segment. It has already been suggested that ultrasonography bears the risk of overinterpretation of a normal status or non-clinical remodelling, as well as not being an ideal survey tool, that is best interpreted in conjunction with other modalities and together with clinical signs (*Chope* 2008). One explanation for the superiority in radiographic agreement of SIVAs could be that on radiographs the structure is depicted in its entireness, providing a more comprehensive image, and a more confident and reproducible opinion about the degree of changes. One ultrasonographic image or loop may both underestimate or overestimate the degree of pathology, if the ultrasonographically depictable region of the structure is the part less affected or more affected by changes, respectively. Moreover, ultrasound examination is operator dependent. One observer was not involved in image acquisition but only graded images remotely. This adds to the difficulty and possibly affected agreement as one might be used to other technical settings or other projections. Ultrasonographic interpretation starts during the examination itself, as the operator will adjust technique to achieve the best images possible, and builds an opinion not just based on images being stored but also on the overall impression he perceives during the examination. Storing of relevant and illustrative images is obviously useful, and these should be representative of findings and of good quality for allowing others to review images at a later stage and be able to relate findings with the interpretation provided by the first observer. We acknowledge the limited usefulness of assessing interobserver agreement on interpretation of stored ultrasound images as this is rarely performed in practice. But, at the same time, along with the telemedicine and teleradiology trend, we experience increased interest by referring veterinarians to have their ultrasonographic examinations remotely evaluated by more experienced veterinarians or radiologists. In the case of chronic progressive degenerative pathology such as OA, our results show that even when images of very good quality are interpreted by experienced observers, agreement can be weak. This reinforces our skepticism about remote interpretation of ultrasonographic images and our opinion that interpretation of ultrasound is most valuable when done by the ultrasound operator who has been examining the entire region in a dynamic way, and not only still pictures or recorded loops.

Agreement on interpretation of radiographic findings regarding dorsal spinous process changes showed moderate to almost perfect agreement, with no significant variation depending on location. The high degree of agreement reflects the relative low depth of the structures to be examined and clear grading systems, leaving a limited amount of space for subjectivity.

Spondylotic changes at the vertebral bodies visible on radiographs were rare and observer graded with very high agreement. This is due to the very low prevalence of this type of changes, resulting in a high number of grades 0 allocated by both observers. Moreover, when present, changes were also easily spotted and assessed, due to their conspicuity being contrasted by radiolucent lung parenchyma at their predilection site in the thoracic segment (*Meehan* et al. 2009).

Result of our kappa statistics are in line with most reported values determined by other studies on interobserver agreement for the assessment of radiography in horses, which has shown to range from poor to almost perfect, especially for chronic progressive degenerative changes, such as navicular bone pathology (*Groth* et al. 2009, *Hellige* et al. 2018). Interobserver agreement studies on ultrasonographic examination in horses are rare, and some report very high agreements, mainly due to the assessed pathologies, which are simpler in nature such as the presence of osteochondral fragments, or the presence of free pleural gas (*Beccati* et al. 2013, *Partlow* et al. 2017, *Plevin* and *McLellan* 2021). We attempted to reach highest possible kappa values, first by using board certified equine radiologists with longstanding experience, as well as by standardizing evaluation through consensus on the grading systems applied. Despite this, the results show modest agreement at some locations. To further improve kappa values, one could have created a training session for the observers, use more observers or use grading systems with less grades (*Koo* and *Li* 2016). Choice of grading system influences consequent agreement values, as a system with few grades will result in lower variation in interpretation and thus in higher agreements (*Wise* et al. 2021, *Bell* et al. 2007). Depending on the aim of the study, authors can adjust grading systems, for example merging normal and low grades as there will be no therapeutic implication in case management, or merge moderate and severe grades as both will result in same therapeutic management.

We acknowledge the fact that the chosen population is suboptimal to perform kappa statistics on, given the high prevalence of normal status. Kappa statistics yield most information with a more even distribution of variables than we found at some locations, or in some modalities in our population. This is also the reason why at some locations no kappa value could be calculated, especially where there was complete absence of abnormalities. At these locations, observers perfectly agreed on the absence of changes in all horses. However, given the lack of variability in abnormalities and evaluations no kappa value could be calculated.

It is important that radiologists and clinicians are aware of the limited agreements in imaging evaluation of certain regions of the equine vertebral column, in particular the SIVAs. This should be kept in mind when interpreting images, and caution should be applied to try and avoid excessive subjective interpretation and keep in line with more objective criteria for example in prepurchase situations.

In conclusion this paper shows that ultrasonographic and radiographic changes at the vertebral column in normally performing horses occur at previously reported predilections sites found in symptomatic populations. Interobserver agreement on imaging findings in the equine vertebral column of this population is moderate and the results underline the importance of interpreting imaging findings with clinical findings for case management, and decision making. Further research should aim at assessing intermodality agreements, as well as agreement between imaging findings and clinical symptoms to improve our understanding of the clinical significance of imaging findings in the equine vertebral column.

References

Audigié F., Coudry V., Jacquet S., Bertoni L., Denoix J.-M. (2015) Diagnostic imaging of the equine back. EAVCDI Yearbook 2015, 25–46

Beccati F., Chalmers H. J., Dante S., Lotto E., Pepe M. (2013) Diagnostic sensitivity and interobserver agreement of radiography and ultrasonography for detecting trochlear ridge osteochondrosis lesions in the equine stifle. Vet. Radiol. Ultrasound 54 , 176–184; DOI 10.1111/vru.12004

Bell R. J., Kingston J. K., Mogg T. D. (2007) A comparison of two scoring systems for endoscopic grading of gastric ulceration in horses. New Zeel. Vet. J. 55, 19–22; DOI 10.1080/00480169.2007.36730

Berg L. C., Nielsen J. V., Thoefner M. B., Thomsen P. D. (2003) Ultrasonography of the equine cervical region: a descriptive study in eight horses. Equine Vet. J. 35, 647–655; DOI 10.2746/042516403775696311

Berner D., Brehm W., Gerlach K. (2012) Die Bedeutung der Kopf-Hals-Haltung bei der röntgenologischen Darstellung der Foramina intervertebralia des Pferdehalses in der seitlichen Projektion. Pferdeheilkunde 28, 39–45; DOI 10.21836/PEM20180405

Blättler C. (2020) Back health of Swiss riding horses. Masterthesis, University of Zurich.

Burns G., Dart A., Jeffcott L. (2018) Clinical progress in the diagnosis of thoracolumbar problems in horses. Equine Vet. Educ. 30, 477–485

Chope K. (2008) How to Perform Sonographic Examination and Ultrasound-Guided Injection of the Cervical Vertebral Facet Joints in Horses. AEEP Proceedings.

Clayton H. M., Stubbs N. C. (2016) Enthesophytosis and Impingement of the Dorsal Spinous Processes in the Equine Thoracolumbar Spine. J. Equine Vet. Sci. 47, 9–15

Crijns C. P., Broeckx B. J. G. (2021) Evaluation of cervical radiographs in Dutch Warmblood horses, using a novel radiographic grading system for the cervical articular process joints. Equine Vet. Educ. 33, 593–601; DOI 10.1111/eve.13375

Denoix J.-M. (2017) Ataxia is not the only clinical manifestation of neck injuries. Proceeding of the American College of Veterinary Internal Medicine ACVIM Forum Scientific Sessions, National Harbour, MD

Denoix J.-M. (1999) Ultrasonographic evaluation of back lesions. Vet. Clin. North Am. Equine Pract. 15l, 131–159; DOI 10.1016/s0749-0739(17)30169-4

Denoix J.-M., Dyson S. J. (2003) Thoracolumbar Spine. In Diagnosis and Management of Lameness in the Horse, edited by *Ross M. W. and Dyson S. J.*, 509–521. W. B. Saunders

Dittmann M. T., Latif S. N., Hefti R., Hartnack S., Hungerbühler V., Weishaupt M. A. (2020) Husbandry, Use, and Orthopedic Health of Horses Owned by Competitive and Leisure Riders in Switzerland. J. Equine Vet. Sci. 91, 103–107; DOI 10.1016/j.jevs.2020.103107

Down S. S., Henson F. M. (2009) Radiographic retrospective study of the caudal cervical articular process joints in the horse. Equine Vet. J. 41, 518–24; DOI 10.2746/042516409x391015

Dyson S. J. (2011) Lesions of the equine neck resulting in lameness or poor performance. Vet. Clin. North Am. Equine Pract. 27, 417–437

Dämmrich K., Randelhoff A., Weber B. (1993) Ein morphologischer Beitrag zur Biomechanik der thorakolumbalen Wirbelsäule und Pathogenese des Syndroms sich berührender Dornfortsätze (Kissing-Spines-Syndrom) bei Pferden. Pferdeheilkunde 5, 267–263, 276–281; DOI 10.21836/PEM19930501

Erichsen C. (2003) Diagnostic Imaging of the Equine Thoracolumbar Spine and Sacroiliac Joint Region. Doctor Medicinae Veterinariae, Department of Large Animal Clinical Sciences, The Norwegian School of Veterinary Science.

Erichsen C., Eksell P., Holm K. R., Lord P., Johnston C. (2004) Relationship between scintigraphic and radiographic evaluations of spinous processes in the thoracolumbar spine in riding horses without clinical signs of back problems. Equine Vet. J. 36 , 458–65; DOI 10.2746/0425164044877341

Espinosa-Mur P., Phillips K. L., Galuppo L. D., DeRouen A., Benoit P., Anderson E., Shaw K., Puchalski S., Peters D., Kass P. H., Spriet M. (2020) Radiological prevalence of osteoarthritis of the cervical region in 104 performing Warmblood jumpers. Equine Vet. J. 53, 972–978; DOI 10.1111/evj.13383

Fritsche B., Lorenz I., Busch-Tenter B., Gerlach K. (2020) Cone Beam-Computertomographie der Halswirbelsäule am stehenden Pferd – Teil 1: Befunde und Klinik. Pferdeheilkunde 36, 430–437; DOI 10.21836/PEM20200506

Fuller C. J., Bladon B. M., Driver A. J., Barr A. R. (2006) The intra- and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses. Vet. J. 171, 281–286; DOI 10.1016/j.tvjl.2004.10.012

Geiger C. P., Gerhards H. (2015) Radiologische Befunderhebung an der Brustwirbelsäule des Pferdes gemäß des Röntgenleitfadens 2007 unter Berücksichtigung der klinischen Relevanz. Pferdeheilkunde 31, 39–48; DOI 10.21836/PEM20150106

Gerlach K., Winter K., Zeller S., Kafka U. C. M. (2018) Nuklearmedizinische retrospektive Untersuchungen der Facettengelenke C 6/7 der Halswirbelsäule beim Pferd. Pferdeheilkunde – Equine Medicine 34, 347–352; DOI 10.21836/PEM20180405

Gillen A., Dyson S., Murray R. (2009) Nuclear scintigraphic assessment of the thoracolumbar synovial intervertebral articulations. Equine Vet. J. 41, 534–540; DOI 10.2746/042516409x376940.

Girodroux M., Dyson S., Murray R. (2009) Osteoarthritis of the thoracolumbar synovial intervertebral articulations: clinical and radiographic features in 77 horses with poor performance and back pain. Equine Vet. J. 41, 130–138; DOI 10.2746/042516408x345099.

Gough S. L., Anderson J. D. C., Dixon J. J. (2020) Computed tomographic cervical myelography in horses: Technique and findings in 51 clinical cases. J. Vet. Intern Med 34, 2142–2151; DOI 10.1111/jvim.15848

Graf von Schweinitz D. (1999) Thermographic diagnostics in equine back pain. Vet. Clin North Am Equine Pract 15, 161–77, viii; DOI 10.1016/s0749-0739(17)30170-0

Groth A. M., May S. A., Weaver M. P., Weller R. (2009) Intra- and interobserver agreement in the interpretation of navicular bones on radiographs and computed tomography scans. Equine Vet. J. 41, 124–9; DOI 10.2746/042516408x345125

Gundel M., Schatzmann U., Ueltschi G. (1997) Rückenprobleme beim Pferd: Ein Vergleich der bildgebenden Verfahren mit der klinischen Untersuchung. Pferdeheilkunde Equine Medicine 13, 659–666; DOI 10.21836/PEM19970605

Haussler K. K. (1999) Osseous spinal pathology. Vet. Clin. North Am Equine Pract 15, 103–12, vii; DOI 10.1016/s0749-0739(17)30167-0

Haussler K. K., Pool R. R., Clayton H. M. (2019) Characterization of bony changes localized to the cervical articular processes in a mixed population of horses. PLoS One 14, e0222989; DOI 10.1371/journal.pone.0222989.

Haussler K. K., Stover S. M., Willits N. H. (1999) Pathologic changes in the lumbosacral vertebrae and pelvis in Thoroughbred racehorses. Am. J. Vet. Res. 60, 143–53

Haussler K. K., Jeffcott L. B. (2014) Back and pelvis. In Equine Sports Medicine and Surgery, edited by *Hinchcliff K. W., Kaneps A. J.* and *Georg R. J.*, 419–456. W.B. Saunders.

Head M. (2014) Ultrasonography of the neck and back. In Atlas of Equine Ultrasonography, edited by Kidd J. A., Lu K. G., Frazer M. L., John Wiley & Sons, Ltd.

Hellige M., Rohn K., Buschkamp L., Stadler P. (2018) Interuntersucher-Varianz bei der Beurteilung von Röntgenaufnahmen von Pferden nach dem Röntgenleitfaden 2007. Pferdeheilkunde 34, 212–222; DOI 10.21836/PEM20180301

Holmer M., Wollanke B., Stadtbäumer G. (2007) Röntgenveränderungen an den Dornfortsätzen von 295 klinisch rückengesunden Warmblutpferden. Pferdeheilkunde 23, 507–511; DOI 10.21836/PEM20070506

Holmer M. (2005) Röntgenbefunde an den Dornforstätzen klinisch rückengesunder Warmblutpferde. Dissertationsarbeit, Ludwig-Maximilians-Universität München.

Jackson M. A., Vizard A. L., Anderson G. A., Mattoon J. S., Lavelle R. B., Smithenson B. T., Lester N. V., Clarke A. F., Whitton R. C. (2014) An assessment of intra- and interobserver agreement of reporting orthopaedic findings on presale radiographs of Thoroughbred yearlings. Equine Vet. J.46, 567–74; DOI 10.1111/evj.12150

Jeffcott L B. (1993) Rückenprobleme des Athleten Pferd - I. Ein Bericht über das Erkennen und die Möglichkeiten der Diagnose. Pferdeheilkunde 9, 143–150; DOI 10.21836/PEM19930301

Jeffcott L. B. (1979) Radiographic Features of the Normal Equine Thoracolumbar Spine. Veterinary Radiology 20:140–147

Jeffcott L. B. (1980) Disorders of the thoracolumbar spine of the horse - a survey of 443 cases. Equine Vet. J. 12, 197–210; DOI 10.1111/j.2042-3306.1980.tb03427.x

Koo T. K., Li M. Y. (2016) A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J.Chiropr Med 15, 155–63; DOI 10.1016/j.jcm.2016.02.012

Lindgren C. M., Wright L., Kristoffersen M., Puchalski S. M. (2020) Computed tomography and myelography of the equine cervical spine: 180 cases. (2013–2018). Equine Vet. Educ. 33; 475–483

Lischer C. J., Withers J. M., Parkin T. (2010) Die Genauigkeit von radiologischen Messungen an Fazettengelenken der Halswirbelsäule beim Pferd. Pferdeheilkunde 26, 553–558; DOI 10.21836/PEM20100409

Martin B. B., Klide A. M. (1999) Physical examination of horses with back pain. Vet. Clin North Am Equine Pract 15, 61–70, vi; DOI 10.1016/s0749-0739(17)30163-3

Mayaki A. M., Abdul Razak I. S., Adzahan N. M., Mazlan M., Rasedee A. (2020) Clinical assessment and grading of back pain in horses. J. Vet. Sci. 21, e82; DOI 10.4142/jvs.2020.21.e82.

McHugh M. L. (2012) Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 22, 276–82

Meehan L., Dyson S., Murray R. (2009) Radiographic and scintigraphic evaluation of spondylosis in the equine thoracolumbar spine: a retrospective study. Equine Vet. J. 41, 800–7; DOI 10.2746/042516409x436592

Menzies-Gow N. J., Stevens K. B., Sepulveda M. F., Jarvis N., Marr C. M. (2010) Repeatability and reproducibility of the Obel grading system for equine laminitis. Vet. Rec 167, 52–5; DOI 10.1136/vr.c3668

Müller-Quirin J., Dittmann M. T., Roepstorff C., Arpagaus S., Latif S. N., Weishaupt M. A. (2020) Riding Soundness-Comparison of Subjective With Objective Lameness Assessments of Owner-Sound Horses at Trot on a Treadmill. J. Equine Vet. Sci. 95, 103314; DOI 10.1016/j.jevs.2020.103314

Olsen E., Dunkel B., Barker W. H., Finding E. J., Perkins J. D., Witte T. H., Yates L. J., Andersen P. H., Baiker K., Piercy R. J. (2014) Rater agreement on gait assessment during neurologic examination of horses. J. Vet. Intern. Med. 28, 630–8; DOI 10.1111/jvim.12320

Partlow J., David F., Hunt L. M., Relave F., Blond L., Pinilla M., Lavoie J. P. (2017) Comparison of thoracic ultrasonography and radiography for the detection of induced small volume pneumothorax in the horse. Vet. Radiol. Ultrasound 58, 354–360; DOI 10.1111/vru.12480

Plevin S., McLellan J. (2021) Comparison of ultrasonography and radiography with arthroscopy for diagnosis of dorsoproximal osteochondral fragmentation of the proximal phalanx in 56 Thoroughbred racehorses. Equine Vet. J.; DOI 10.1111/evj.13497

Riccio B., Fraschetto C., Villanueva J., Cantatore F., Bertuglia A. (2018) Two Multicenter Surveys on Equine Back-Pain 10 Years a Part. Front Vet. Sci. 5, 195; DOI 10.3389/fvets.2018.00195.

Veraa S., de Graaf K., Wijnberg I. D., Back W., Vernooij H., Nielen M., Belt A. J. M. (2020) Caudal cervical vertebral morphological variation is not associated with clinical signs in Warmblood horses. Equine Vet. J. 52, 219–224; DOI 10.1111/evj.13140

Wilsmann F., Gerhards H., Pichon S. (2020) Statistische Erhebungen zur Prävalenz röntgenologisch erfassbarer Befunde an der Halswirbelsäule des Pferdes. Pferdeheilkunde – Equine Medicine 36, 128–136; DOI 10.21836/PEM20200205

Wise J. C., Wilkes E. J. A., Raidal S. L., Xie G., Crosby D. E., Hale J. N., Hughes K. J. (2021) Interobserver and intraobserver reliability for 2 grading systems for gastric ulcer syndrome in horses. J. Vet. Intern. Med. 35, 571–579; DOI 10.1111/jvim.15987

Wood R., Mullard J., Tambaschi M., Weller R. (2014) Diagnostic value of laterolateral radiographs in equine cervical articular process joint osteoarthritis. European veterinary diagnostic imaging annual conference, Utrecht, The Netherlands.

Zimmerman M., Dyson S., Murray R. (2011) Comparison of radiographic and scintigraphic findings of the spinous processes in the equine thoracolumbar region. Vet. Radiol. Ultrasound 52, 661–671; DOI 10.1111/j.1740-8261.2011.01845.x

Zimmerman M., Dyson S., Murray R. (2012) Close, impinging and overriding spinous processes in the thoracolumbar spine: the relationship between radiological and scintigraphic findings and clinical signs. Equine Vet. J. 44, 178–184; DOI 10.1111/j.2042-3306.2011.00373.x