



Byrne, Christian A. (2021) *Clinical magnetic resonance imaging of the equine foot: An investigation of factors influencing image quality and image interpretation*. MVM(R) thesis.

<http://theses.gla.ac.uk/82590/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>  
[research-enlighten@glasgow.ac.uk](mailto:research-enlighten@glasgow.ac.uk)

**Clinical magnetic resonance imaging of the equine  
foot: An investigation of factors influencing image  
quality and image interpretation**

**Christian A. Byrne  
BVM&S**

Submitted in fulfilment of the requirements for  
the Degree of Master of Veterinary Medicine

School of Veterinary Medicine  
College of Medical, Veterinary & Life Sciences  
University of Glasgow

December 2021  
© Christian A. Byrne 2021

## Abstract

Magnetic resonance imaging (MRI) is a fundamental imaging modality for evaluation of the equine foot. Optimising image quality and observer pathology identification is important to maximise the diagnostic value of MRI. There is limited information investigating factors that influence magnetic resonance (MR) image quality in live equine patients in a clinical setting. In addition, agreement between observers assessing pathology on clinical MRI studies of the equine foot has not been investigated. This project aimed to evaluate the influence of patient general anaesthesia (which encompasses the potential effects of motion and weight-bearing) and field strength on clinical MR image quality. In addition, the project aimed to determine the agreement between expert observers for pathology assessment of clinically important anatomical structures of the equine foot.

A total of fifteen routine equine MRI foot studies were acquired from the clinical databases of three different MRI systems: low-field standing, low-field under general anaesthesia and high-field under general anaesthesia. Ten experienced observers (diploma or associate level) assessed entire MRI studies and seven key individual anatomical structures of the equine foot. Observers used an online image assessment platform to grade subjective image quality (briefly, grade 1: textbook quality, grade 2: high diagnostic quality, grade 3: satisfactory diagnostic quality, grade 4: non-diagnostic), pathology, and their confidence in pathology assessment. Statistical analysis was performed to assess the influence of anaesthesia and field strength on image quality, and to document inter-observer agreement in pathology assessment.

Observers deemed most clinical MRI foot studies to be of diagnostic quality, regardless of acquisition system. There were no significant differences in image quality between low-field standing and low-field under general anaesthesia (for both groups all individual structure image quality median grades= 3). Conversely, high-field under general anaesthesia studies had significantly greater image quality for entire studies and all individual anatomical structures (median grades= 1 for 5/7 structures and 2 for 2/7 structures) compared to low-field under general anaesthesia (all individual structure median grades= 3). There was a general trend of agreement between observers for pathology assessment of

anatomical structures of the equine foot. Although absolute agreement for pathology assessment grading was generally low, relative agreement (accounting for the ranking of study pathology grading) was greater. Agreement was lowest for the distal interphalangeal joint (Kendall's coefficient of concordance= 0.19) and greatest for the navicular bone (Kendall's coefficient of concordance= 0.70). Importantly there were instances of marked variation in pathology assessment for individual MRI studies. In general, agreement was greater at the extremes of pathology.

The findings indicate that field strength is a more important influencer of image quality than general anaesthesia for MRI of the equine foot in clinical patients. However, the reasons described for reduced image quality appear to differ between MRI systems. There was a general tendency of agreement between observers for pathology assessment. However, there can be notable variation in pathology assessment for individual MRI studies, even when interpretation is performed by experienced observers. Future work is needed to evaluate the influence of image quality factors when imaging other regions of the equine limb and to investigate the processes of lesion identification and subsequent diagnostic decision making by those interpreting MRI images of the equine foot.

# Table of Contents

Abstract .....	2
List of Tables .....	6
List of Figures .....	8
List of Accompanying Material.....	9
Preface.....	10
Acknowledgement .....	11
Author's Declaration .....	12
Ethical Approval .....	13
Definitions/Abbreviations .....	14
Chapter 1 Introduction .....	15
1.1 Overview.....	15
1.2 Producing a magnetic resonance image.....	17
1.3 Magnetic resonance imaging of the equine foot .....	23
1.4 Magnetic resonance image quality in equine practice .....	32
1.5 Observer agreement during equine magnetic resonance imaging.....	41
1.6 Aims .....	43
Chapter 2 The influence of general anaesthesia on perceived image quality of clinical magnetic resonance imaging of the equine foot .....	44
2.1 Introduction .....	44
2.2 Aims .....	45
2.3 Study design.....	46
2.4 Methods .....	46
2.5 Results .....	63
2.6 Discussion .....	69
Chapter 3 The influence of field strength on perceived image quality of clinical magnetic resonance imaging of the equine foot.....	75
3.1 Introduction .....	75
3.2 Aims .....	76
3.3 Study design.....	77
3.4 Methods .....	77
3.5 Results .....	84
3.6 Discussion.....	91
Chapter 4 Inter-observer agreement for the assessment of pathology during clinical magnetic resonance imaging of the equine foot .....	96
4.1 Introduction .....	96
4.2 Aims .....	97
4.3 Study design.....	97
4.4 Methods .....	97

4.5	Results .....	104
4.6	Discussion .....	115
Chapter 5	General discussion.....	121
Appendices	.....	127
List of References	.....	161

## List of Tables

Table 2-1 Summary of magnetic resonance imaging lesion distribution in horses with foot pain from six studies. ....	48
Table 2-2 Grading system used for image quality assessment of entire magnetic resonance imaging foot studies. ....	49
Table 2-3 Interpretation of kappa statistic values based on Altman (1991). ....	62
Table 2-4 Median and range for image quality grading for low-field standing and low-field under general anaesthesia studies. ....	65
Table 2-5 Comparison of proportion of diagnostic image quality gradings for low-field standing and low-field under general anaesthesia studies.....	65
Table 2-6 Comparison of ranked gradings for image quality for low-field standing and low-field under general anaesthesia studies. ....	66
Table 2-7 Median pathology grades and rank comparison for low-field standing and low-field under general anaesthesia studies. ....	68
Table 2-8 Output of inter-observer agreement analysis for magnetic resonance image quality grading for all low-field standing and low-field under general anaesthesia studies. ....	68
Table 3-1 Interpretation of kappa statistic values based on Altman (1991) ....	83
Table 3-2 Median and range for image quality grading for low-field and high-field under general anaesthesia studies. ....	85
Table 3-3 Comparison of proportion of diagnostic image quality gradings for low-field under general anaesthesia and high-field under general anaesthesia studies. ....	87
Table 3-4 Comparison of ranked gradings for image quality for low-field under general anaesthesia and high-field under general anaesthesia studies. ....	87
Table 3-5 Median pathology grades and rank comparison for low-field and high-field under general anaesthesia studies. ....	89
Table 3-6 Output of inter-observer agreement analysis for magnetic resonance image quality grading for all low-field under general anaesthesia and high-field under general anaesthesia studies.....	90
Table 4-1 Frequency with which observers interpret images from different magnetic resonance imaging acquisition systems. ....	104
Table 4-2 Pathology gradings assigned to each grade for individual anatomical structures for 15 magnetic resonance imaging studies of the equine foot. ....	106
Table 4-3 Matched assessments and Fleiss' kappa values for observer agreement for the presence or absence of pathology on magnetic resonance imaging foot studies .....	108

Table 4-4 Inter-observer agreement analysis for pathology assessment of the equine foot .....	108
Table 4-5 Fleiss' kappa values for individual grades of pathology assessment of the equine foot. ....	110
Table 4-6 Pathology assessment confidence gradings assigned to each grade for individual anatomical structures for 15 magnetic resonance imaging studies of the equine foot. ....	111
Table 4-7 Inter-observer agreement analysis for pathology assessment confidence grading of the equine foot. ....	113
Table 4-8 Fleiss' kappa values for individual grades of pathology assessment confidence grading of the equine foot. ....	114



## List of Figures

Figure 2-1 Format of the entire study image quality assessment interface. ....	49
Figure 2-2 Format of the individual anatomical structure assessment interface.	54
Figure 2-3 Bar chart displaying the image quality gradings for low-field standing and low-field under general anaesthesia studies. ....	64
Figure 2-4 Bar chart displaying the categorisation of image quality comments for low-field standing and low-field under general anaesthesia studies. ....	66
Figure 2-5 Bar chart displaying pathology gradings for low-field standing and low-field under general anaesthesia studies. ....	67
Figure 3-1 Bar chart displaying the image quality gradings for low-field and high-field under general anaesthesia studies. ....	86
Figure 3-2 Bar chart displaying the categorisation of image quality comments for low-field under general anaesthesia and high-field under general anaesthesia studies. ....	88
Figure 3-3 Bar chart displaying pathology gradings for low-field and high-field under general anaesthesia studies. ....	89
Figure 4-1 Format of the pathology and pathology assessment confidence interface. ....	99
Figure 4-2 The observer profile section of the image assessment platform ....	101
Figure 4-3 Bar chart demonstrating the number of magnetic resonance imaging pathology gradings assigned to each grade for anatomical structures of the equine foot. ....	106
Figure 4-4 Linear dot plot displaying the Fleiss' kappa values for pathology assessment of seven anatomic structures of the equine foot. ....	109
Figure 4-5 Bar chart demonstrating the number of magnetic resonance imaging pathology assessment confidence gradings assigned to each grade for individual anatomical structures of the equine foot. ....	112
Figure 4-6 Linear dot plot displaying the Fleiss' kappa values for pathology assessment confidence grading for seven anatomic structures of the equine foot. ....	113

## List of Accompanying Material

Appendix 1: Representative anatomical and magnetic resonance images of the equine foot

Appendix 2: Pulse sequence parameters used in magnetic resonance imaging studies included in image assessment

Appendix 3: Inter-observer agreement analysis for image quality assessment across all systems

Appendix 4: Bubble charts for pathology assessment and pathology assessment confidence

## Preface

The project was conceived by the author and supervisors. Development of the methods, data collection and data analysis were performed by the author. The thesis and associated manuscripts were prepared by the author with mentorship from the supervisors.

The studies presented in Chapters 2 and 3 of the thesis have been published as a general article in a peer-reviewed journal (Byrne, C. A., Marshall, J. F. and Voute, L. C. 2021. "Clinical Magnetic resonance image quality of the equine foot is significantly influenced by acquisition system." *Equine Veterinary Journal* 53 (3): 469-80.). The study presented in Chapter 4 of the thesis has been presented as an abstract at a national meeting (British Equine Veterinary Association Congress 2021, "Pathology assessment during clinical magnetic resonance imaging of the equine foot: do observers agree?". Birmingham).

## Acknowledgement

Firstly, I would like to thank my supervisors Dr Lance Voute and Dr John Marshall for the opportunity to undertake this research project. The insights and support they have offered throughout the project have been invaluable.

This research project was generously funded by the Horserace Betting Levy Board (Grant Number: SPrj033). I am very grateful to the following individuals who kindly acted as observers for the study: Dagmar Berner, Marianna Biggi, Jonathon Dixon, Raphael Labens, Lucy Meehan, Julien Olive, Michael Schramme, Meredith Smith, Alex Young and Davide Zani. I would also like to thank the following organisations: Vet CT Specialists Ltd. (particularly Victoria Johnson for assistance in recruitment of observers), Hallmarq Veterinary Imaging Ltd. (particularly Nick Bolas for provision of images and feedback during the conception of the study), Liphook Equine Hospital Ltd. (for provision of images) and North Carolina State University-College of Veterinary Medicine (particularly Nathan Nelson for provision of images). I am also grateful to Gawain Hammond (University of Glasgow) for his feedback on study design and Tim Parkin (University of Glasgow) for his guidance on statistical aspects of the study.

I would like to thank Anna for her generous encouragement and advice during my research and residency. Finally, I would like to thank my parents for their unwavering support throughout my veterinary career.

## **Author's Declaration**

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Christian A. Byrne

## **Ethical Approval**

The project received ethical approval from the School of Veterinary Medicine Research Ethics Committee, University of Glasgow (Ref 08a/18).

All observer participants were advised that their assessments would be anonymised. By participating the observers consented to the collection of their responses, use of their responses in data analysis, contribution of these data to publications derived from the study and the long-term storage of data, which may be utilised for future research.

## Definitions/Abbreviations

ALDDFT: Accessory ligament of the deep digital flexor tendon

DDFT: Deep digital flexor tendon

DICOM: Digital imaging and communications in medicine

DIP: Distal interphalangeal

DSIL: Distal sesamoidean impar ligament

FLASH: Fast low angle shot

FOV: Field of view

FSE: Fast spin echo

HR: High resolution

MPR: Multiplanar reconstruction

MR: Magnetic resonance

MRI: Magnetic resonance imaging

PACS: Picture archiving and communication system

PD: Proton density

STIR: Short tau inversion recovery

T: Tesla

TE: Echo time

TI: Inversion time

TME: Turbo multi echo

TR: Repetition time

TSE: Turbo spin echo

URL: Uniform resource locator

# Chapter 1 Introduction

## 1.1 Overview

### 1.1.1 Lameness in the equine foot

The foot is a common and important origin of lameness in the horse (Ross, 2011). Localisation of the origin of pain is an important step in lameness investigation and typically consists of static assessment, dynamic evaluation, palpation, manipulation and diagnostic anaesthesia (Baxter, Stashak and Keegan, 2020). Following localisation of the origin of lameness, diagnostic imaging is typically a key step to characterise the inciting pathology. Advances in diagnostic imaging have provided the most significant recent developments in the clinician's ability to diagnose musculoskeletal disease and elucidate details of disease pathogenesis, the most appropriate treatment and likely prognosis.

The equine foot is a complex structure with an intricate combination of osseous and soft tissue structures enclosed within a specialised hoof capsule (Fails, 2020). An elaborate combination of epithelial and connective tissue forms the suspensory mechanism of the third phalanx which connects the bony column of the limb to the hoof and ultimately the weight-bearing surface (Budras *et al.*, 2009b). The osseous and cartilaginous structures are supported by an elaborate network of ligaments and closely associated tendinous structures (Fails, 2020). This diverse arrangement of anatomy results in an extensive list of anatomical structures that can be implicated as an origin of foot lameness. Therefore, diagnostic imaging of this region during lameness investigation is challenging but this has been revolutionised by the introduction of magnetic resonance imaging (MRI) (Barrett *et al.*, 2017; Schramme and Segard-Weisse, 2020).

### 1.1.2 Diagnostic imaging of the equine foot

In many clinical scenarios, radiography remains the primary diagnostic imaging modality for the evaluation of the equine foot. Radiographic imaging typically involves a series of five radiographic projections of each foot including a lateromedial, dorsopalmar (weight-bearing), dorsoproximal-palmarodistal oblique centred on the third phalanx, dorsoproximal-palmarodistal oblique centred on the navicular bone and a palmaroproximal-palmarodistal oblique



centred on the flexor surface of the navicular bone (Butler *et al.*, 2017). Radiography can provide key diagnostic information, particularly relating to the bones, joints and collateral cartilages of the foot (Barrett and Acutt, 2020). Valuable insight can also be yielded about the relationship of the foot with the shoe and ground surface, particularly in relation to static foot balance (Kummer *et al.*, 2009). Radiography provides limited diagnostic information about the soft tissues and the physiological activity of bone within the foot (Barrett and Acutt, 2020). Scintigraphy can also provide diagnostic information during the investigation of foot pain. Pool and bone phase images can demonstrate pathology within the deep digital flexor tendon (DDFT), the bony tissues (Dyson, 2002) and the collateral cartilages of the foot (Dyson and Nagy, 2011). However, this information is often insufficient to provide a definitive diagnosis and accurate assessment of prognosis, so may need to be combined with other modalities such as radiography. Ultrasonography also has significant limitations due to poor transmission of ultrasound waves through the hoof capsule (Carstens and Smith, 2014).

The first reported uses of MRI in equine medicine utilised cadaver limbs and demonstrated the advantages of this technique for the evaluation of distal limb structures (Park, Nelson and Hoopes, 1987; Denoix *et al.*, 1993). The clinical value of MRI was rapidly demonstrated, particularly for the evaluation of regions where traditional diagnostic imaging modalities had significant limitations (Whitton *et al.*, 1998; Dyson *et al.*, 2003). Given the relative importance of the equine foot as a source of lameness, the complex anatomy of this region and the imaging restrictions imposed by the enclosing hoof capsule, the foot received significant attention as an application for MRI (Kleiter *et al.*, 1999). MRI provides a beneficial combination of multiplanar imaging with the ability to thoroughly evaluate bony and soft tissue structures (Schramme and Segard-Weisse, 2020). However, the unique attributes of the technique also come with an array of considerations that need to be appreciated for those acquiring and interpreting MR images (Murray and Werpy, 2010). The key physics principles underlying MRI differ to those of more traditional diagnostic imaging modalities (Bolas, 2010; Collins, 2016). The multiplanar nature of imaging, ability to dramatically alter tissue contrast (weighting) and unfamiliar nature of artefacts present additional challenges to the implementation of this modality.

## 1.2 Producing a magnetic resonance image

### 1.2.1 Key components of the magnetic resonance imaging system

Several important components contribute to the operation of an MRI system. Key hardware includes the magnet, gradient coils, radiofrequency coils and their associated components. The system also requires a computer and appropriate software for data processing and operator interaction. Most systems require installation in a purpose built room to limit electromagnetic interference (McRobbie *et al.*, 2017c). In addition to dictating the practical applications of the MRI system, many of these design features have important implications for image quality and image interpretation.

### 1.2.2 Principle physics of magnetic resonance imaging

MRI utilises nuclear magnetic resonance to produce an image (McRobbie *et al.*, 2017d). This exploits the principle that protons within the nuclei of atoms have a positive charge and in atoms with an odd atomic number (isotopes where the number of protons and neutrons are not equal) there is a net spin (Westbrook, Roth and Talbot, 2011b). This applies to a variety of nuclei however, in a clinical context hydrogen nuclei are of greatest use given their abundance in the tissues of the body (Bolas, 2010).

Nuclear magnetic resonance imaging relies on the principle that a moving charged particle (for example the positively charged proton of the hydrogen nucleus) produces an associated magnetic field (McRobbie *et al.*, 2017d). When the relevant region of anatomy is placed within the main magnetic field of the MRI system, the  $^1\text{H}$  hydrogen nuclei precess in alignment, either parallel or anti-parallel with this strong external field (Westbrook, Roth and Talbot, 2011b). Each hydrogen nucleus precesses at a frequency that is proportional to the magnetic field it experiences. This specific resonant frequency is termed the Larmor frequency (as it is given by the Larmor equation). When a radiofrequency pulse is applied at the Larmor frequency of the precessing hydrogen nuclei, resonance causes the nuclei to temporarily transition to precession in a different orientation, for example perpendicular to the main magnetic field (McRobbie *et al.*, 2017d).

The magnetic resonance signal is the result of the relaxing, precessing hydrogen nuclei inducing a current in the receiver coil due to their charge (Westbrook, Roth and Talbot, 2011b). Spatial encoding and Fourier transformation are used to establish the origin and characteristics of the received signal and thus generate an MR image. There are two key components to the relaxation of the nuclei. As the nuclei gradually transfer energy to the surrounding lattice (of the body tissue) they progressively return to their equilibrium position parallel to the main magnetic field (McRobbie *et al.*, 2017d). The spin-lattice relaxation (T1 relaxation) is an exponential process and the T1 relaxation time is a time constant that indicates when approximately 63% of the T1 recovery has occurred (Westbrook, Roth and Talbot, 2011b). Simultaneously, the interaction of adjacent nuclei results in progressive desynchronisation of precession between different nuclei and a reduction in transverse magnetisation (McRobbie *et al.*, 2017d). This spin-spin relaxation (T2 relaxation) is also an exponential process and the T2 relaxation time is a time constant that indicates when approximately 37% of the coherent transverse magnetisation remains (Westbrook, Roth and Talbot, 2011b). T2 relaxation occurs relatively rapidly when compared to the longer process of T1 recovery (McRobbie *et al.*, 2017d).

MRI employs the intrinsic properties of different tissues, primarily the T1 relaxation, T2 relaxation and proton (hydrogen ion) density to produce tissue contrast that forms diagnostic images (Westbrook, Roth and Talbot, 2011b). Parameters associated with the pulse sequences of the MRI system can be used to change the relative contrast of different tissues.

### **1.2.3 Pulse sequences in magnetic resonance imaging**

#### **1.2.3.1 General considerations for pulse sequences**

A key step dictating the appearance of the MR images is the selection of the sequence parameters that form the pulse sequence (McRobbie *et al.*, 2017h). The pulse sequence consists of radiofrequency pulses, gradients (aligned relative to the axes of the MRI system) and predetermined time intervals between these components (Bolas, 2010; Westbrook, Roth and Talbot, 2011f). A wide array of different pulse sequences are available for different applications and terminology can vary between manufacturers and institutions (Bolas, 2010;

Westbrook, Roth and Talbot, 2011f). In many clinical situations imaging of a particular anatomical region will utilise a predetermined collection of pulse sequences with only minor modifications required to optimise the image (McRobbie *et al.*, 2017h).

It is important that MRI system operators, MR image interpreters and clinicians have a working knowledge of the fundamental categories of pulse sequences and the related parameters. Several key timing parameters are important in the formation of the pulse sequence and resultant appearance of the image, for example the degree of T1 weighting. The time from one excitation radiofrequency pulse to the next (i.e. the duration of the pulse sequence) is the repetition time (TR) (Westbrook, Roth and Talbot, 2011b). The echo time (TE) is the time from the excitation radiofrequency pulse to the peak of the magnetic resonance signal. Other parameters such as the inversion time (TI) are also relevant for some sequence types.

### **1.2.3.2 Pulse sequences and image weighting**

The primary purpose of different pulse sequences is to produce different tissue contrasts in the resulting images (McRobbie *et al.*, 2017i). Pulse sequences also have different predispositions to artefacts and this may also influence the decision making on the inclusion of pulse sequences in MRI protocols for different anatomical regions. Importantly, the information yielded from different sequences and resulting images is assimilated to provide an overall impression of the region (Murray and Werpy, 2010).

As outlined above some key tissue properties are important in dictating the contrast of an image. When the operator aims to use the differences in T2 properties of tissues as the dictator of contrast, acquisition is performed with a long echo time and repetition time (Westbrook, Roth and Talbot, 2011c). These T2 weighted images are particularly useful for the evaluation of fluid, especially within soft tissue structures (Schramme and Segard-Weisse, 2020). Conversely, if the T1 properties of the tissues are prioritised with a short echo time and repetition time then the resulting images will be T1 weighted (Westbrook, Roth and Talbot, 2011c). These images can be particularly valuable for the definition of tissue margins (Bolas, 2010; McRobbie *et al.*, 2017i). Proton density weighting

relies on the density of available hydrogen nuclei within the tissue and is produced with a short echo time and long repetition time. These images are particularly useful for the evaluation of soft tissue margins (McRobbie *et al.*, 2017i; Schramme and Segard-Weisse, 2020). These are not entirely distinct categories of weighting and depending on the pulse sequence parameters, images can have a combination of weightings. Similarly, other pulse sequence types, especially those with suppression techniques or those performed following administration of a contrast can have significantly different image contrast (Murray, Leece and Judy, 2010; McRobbie *et al.*, 2017i).

### 1.2.3.3 Spin echo and fast spin echo sequences

The spin echo (SE) sequence is commenced by a  $90^\circ$  excitation radiofrequency pulse which shifts the magnetisation vectors into the transverse plane (McRobbie *et al.*, 2017h). Once the pulse finishes a signal is produced as dephasing occurs. In spin echo sequences a second radiofrequency pulse is applied to generate a signal. A  $180^\circ$  radiofrequency pulse reverses the dephasing of the magnetisation vectors to form a coherent signal termed the spin echo (Westbrook, Roth and Talbot, 2011c). Following this the dephasing continues and the signal decays (McRobbie *et al.*, 2017h).

In a clinical scenario the traditional spin echo sequence can be prohibitively slow. Therefore, the synonymous turbo spin echo (TSE) or fast spin echo (FSE) sequences can be preferable (McRobbie *et al.*, 2017h). Following the initial  $90^\circ$  pulse, these sequences utilise a train of  $180^\circ$  radiofrequency pulses within a single repetition time (Bolas, 2010). This generates a corresponding echo train. The number of pulses (and thus the number of echoes) is dictated by the turbo factor (also called the echo train length) (McRobbie *et al.*, 2017h). The greater the turbo factor the shorter the scan time, for example a fast spin echo with a turbo factor of 4 will be approximately 4 times faster than the equivalent spin echo sequence. Given that the echoes of the fast spin echo sequence have different echo times, the image is composed around an effective echo time. Fast spin echo sequences utilise attributes of k-space order to optimise the final image in relation to the desired weighting and resolution (Bolas, 2010; McRobbie *et al.*, 2017a). Some systems utilise a similar approach where a pulse sequence produces images effectively collected at two different echo times to produce a

proton density-weighted image (with a shorter echo time) and a T2-weighted (with a longer echo time) (Bolas, 2010; McRobbie *et al.*, 2017a).

#### 1.2.3.4 Gradient echo sequences

Gradient echo sequences are also commenced by an excitation radiofrequency pulse. However, this pulse has a low flip angle (less than the 90° pulse used in spin echo sequences), therefore only part of the net magnetic vector is functionally within the transverse plane (Westbrook, Roth and Talbot, 2011c). After cessation of the radiofrequency pulse dephasing occurs. This is partially the result of inhomogeneity in the static magnetic field and interactions (with a time constant T2\*) (Bolas, 2010). In addition, a dephasing gradient is also applied across the field, which results in a linear spectrum of resonant frequencies across the tissue (McRobbie *et al.*, 2017h). The subsequent application of a rephasing gradient, the inverse of the dephasing gradient, causes the precessions to rephase and generate a signal, the gradient echo (Westbrook, Roth and Talbot, 2011f; McRobbie *et al.*, 2017h). Encoding gradients are applied as part of the pulse sequence and act in a summative manner with the static magnetic field. This causes continuous linear variation in the local effective magnetic field and the resultant local resonant frequencies of protons along the axis of the gradient (McRobbie *et al.*, 2017j). Therefore, when this slice selecting gradient is present, application of a narrow bandwidth radiofrequency pulse at the start of the pulse sequence will excite protons with a corresponding resonant frequency. This results in excitation of a targeted slice of the patient (McRobbie *et al.*, 2017j).

In addition to the repetition time and echo time mentioned previously, the flip angle is also relevant to the image weighting in gradient echo sequences (Bolas, 2010). This should also be considered in combination with the other parameters. For example, to generate a T1-weighted image a large flip angle and short repetition time are used to limit T1 relaxation allowing differentiation of tissues (Westbrook, Roth and Talbot, 2011f). As discussed previously, gradient echo sequences are influenced by the effect of magnetic field inhomogeneity and local tissue variation (McRobbie *et al.*, 2017b). This factor is combined with T2 decay and results in a shorter relaxation time which is termed apparent T2\* relaxation (McRobbie *et al.*, 2017b). Therefore, relevant images resulting from

this type of gradient echo sequence are termed to be T2\*-weighted. There are some similarities between T2\* images and T2 images, though images are not directly comparable and can provide different clinical information (Bolas 2010). Many variations of gradient echo sequences are available for different purposes with a range of acquisition times. Spoiled or incoherent gradient echo is an important variation that uses a gradient pulse or radiofrequency pulses to spoil remaining transverse magnetisation so that the signal from the most recent radiofrequency pulse contributes to image contrast (Westbrook, Roth and Talbot, 2011f; McRobbie *et al.*, 2017b). This sequence is frequently used to produce T1 weighted images in equine orthopaedic imaging though they are capable of various weightings (Bolas, 2010). Coherent gradient echo sequences use reversal of encoding gradients (rewinding) which allows transverse magnetisation from previous excitations to remain coherent and contribute to the signal (Westbrook, Roth and Talbot, 2011f; McRobbie *et al.*, 2017b).

#### **1.2.3.5 Fat suppression in pulse sequences**

Techniques to suppress the signal from fat are valuable in a clinical context to change contrast and allow differentiation of tissue constituents, particularly in the identification of fluid (Bolas, 2010). A number of methods are available to achieve fat suppressed images and these are primarily based on either the difference in T1 relaxation between fat and water or the difference in resonant frequency of protons in fat and water (de Kerviler *et al.*, 1998; Delfaut *et al.*, 1999). In the context of equine orthopaedic imaging, inversion-recovery sequences are the most frequently used, particularly short tau inversion recovery (STIR) (Bolas, 2010; Schramme and Segard-Weisse, 2020). This sequence commences with 180° radiofrequency pulse that inverts the magnetisation and following this pulse T1 relaxation commences (Bolas, 2010). The longitudinal magnetisation of fat relaxes more rapidly (than that of water) (Delfaut *et al.*, 1999). After a defined period, the inversion time, the magnetisation of fat is entirely in the transverse plane, the null point (Westbrook, Roth and Talbot, 2011f). This is not the case for the more slowly relaxing protons in high water content tissues, with a longer T1 relaxation. Application of a 90° radiofrequency pulse at this time point eliminates the transverse component of fat magnetisation so that it does not contribute to the signal (Westbrook, Roth and Talbot, 2011f). Conversely water still has a

transverse component before and after the 90° radiofrequency pulse and therefore contributes to the signal. This technique can be effectively applied in low-field and high-field MRI systems but is a relatively slow sequence.

Fat saturation is an alternative technique to achieve fat suppression (Bolas, 2010). Chemical shift results in subtle differences in the resonant frequencies of hydrogen nuclei. This difference in frequencies can be capitalised upon, with a spoiling gradient pulse utilised to dephase the fat signal, so that fat signal does not contribute to the image formed (Delfaut *et al.*, 1999).

#### **1.2.3.6 Other pulse sequences**

The key pulse sequence types for equine orthopaedic imaging have been outlined above but this is far from exhaustive. There is an expansive array of different pulse sequences and variations of familiar pulse sequences, with different applications across various specialities (Westbrook, Roth and Talbot, 2011f; McRobbie *et al.*, 2017h).

### **1.3 Magnetic resonance imaging of the equine foot**

#### **1.3.1 Magnetic resonance imaging systems used in equine orthopaedics**

The equine patient presented unique challenges in the practical acquisition of MR images. Early equine MR imaging systems required general anaesthesia of the patient (Dyson *et al.*, 2003; Schneider, Gavin and Tucker, 2003). The desire to avoid general anaesthesia and its potential complications lead to the development of a 0.3T open magnet system designed for use with the standing equine patient (Mair *et al.*, 2003, 2005). The 0.3T open magnet MRI system is now the most commonly used system in equine practice, with installations at over 100 sites internationally (Hallmarq Veterinary Imaging, 2019; Schramme and Segard-Weisse, 2020). Indeed, more than 100,000 horses have been imaged with this system (Hallmarq Veterinary Imaging, 2019). Several other systems have also been employed in equine clinical practice. These are primarily based on adaptations of MR imaging systems designed for use in human MR imaging, though some are marketed specifically to the equine veterinary market (Bolas, 2010). Both low-field and high-field systems have been used in this capacity



(Schramme and Segard-Weisse, 2020)(Budras *et al.*, 2009a). For these modified human systems, general anaesthesia is currently required to allow positioning of the patient within the magnet (Porter and Werpy, 2014)(Budras *et al.*, 2009a).

The preferred choice of system varies between institutions, but important considerations include the primary intention for use of the machine (for example, anatomic regions most commonly imaged), economic factors, expected imaging caseload, practical implications for magnet positioning and personnel experience. The choice of system also has a significant impact on the MR image quality and image interpretation (Werpy, 2010).

### **1.3.2 General considerations for magnetic resonance imaging of the foot**

MRI provides valuable anatomical and physiological information about bony and soft tissues within the foot. The anatomy and interactions of structures within the foot is complex and a thorough knowledge is fundamental to allow acquisition and interpretation of MR images of this area. The multisequence and multiplanar nature of MRI allows structures to be assessed from multiple perspectives, which can yield additional information (Mair *et al.*, 2005). This section will summarise the relevant anatomy of the equine foot and the normal appearance of these structures on MR images. Associated anatomical images and representative MR images for the following sections (1.3.3 to 1.3.8) are presented in Appendix 1. For the purposes of the subsequent discussion of the anatomy of the distal limb, the forelimb and hindlimb are considered comparable (Singh, 2018b). Descriptive terms for the forelimb (for example palmar) will be used in further discussion but are interchangeable with those of the hindlimb unless stated otherwise.

### **1.3.3 The deep digital flexor tendon**

The deep digital flexor tendon (DDFT) is the tendon of insertion of the deep digital flexor muscle. In the forelimb, the muscle is formed by the humeral, ulnar and radial heads (Budras *et al.*, 2009a). The tendon courses along the palmar aspect of the distal limb and is joined by the accessory ligament of the deep digital flexor tendon (ALDDFT) in the mid-metacarpus. In the hindlimb the deep digital flexor is formed by the lateral digital flexor, tibialis caudalis and

medial digital flexor muscles (Budras *et al.*, 2009b). An ALDDFT is also present in the hindlimb but this is less developed when compared to the forelimb (Eliashar *et al.*, 2010; Singh, 2018b). The DDFT is contained within the digital flexor tendon sheath and passes through the manica flexoria proximal to the proximal sesamoid bones (Denoix, 1994). The anatomy of the distal DDFT can be considered comparable between the forelimb and hindlimb (Denoix, 1994).

The DDFT courses from the digital flexor tendon sheath on the palmar/plantar aspect of the pastern, to insert on the facies flexoria of the third phalanx (Budras *et al.*, 2009b). The DDFT is intimately associated with the navicular bone, navicular bursa and their associated ligaments. Clinically, this grouping of structures is often termed the podotrochlear apparatus (Barrett *et al.*, 2017). Within the region of the foot the DDFT has a bilobed structure, which is typically symmetrical (Murray *et al.*, 2004). The DDFT flattens in profile as it courses over the flexor surface of the navicular bone prior to its crescent shaped insertion (Denoix, 1994).

The highly organised collagenous structure of the normal deep digital flexor tendon causes this structure to have low signal intensity on T1 weighted, T2 weighted and proton density weighted images (Busoni and Snaps, 2002). The fascicular structure is also evident due to the higher signal intensity of the loosely organised supportive connective tissue that encloses the fascicles (Dyson, Van Thielen and Murray, 2010).

#### **1.3.4 The navicular bone, associated ligaments and navicular bursa**

The functional unit of the podotrochlear apparatus is formed by the combination of the DDFT with the navicular bone, associated ligaments and navicular bursa. The navicular (distal sesamoid) bone derives its name from its resemblance to the hull of a boat (Singh, 2018a). The bone is positioned at the palmar aspect of the distal interphalangeal joint. Therefore, the navicular bone articulates with the palmarodistal aspect of the second phalanx and has a narrow articulation (the palmar facet) with the palmar aspect of the third phalanx (Gabriel *et al.*, 1998; Davies and Philip, 2007). The navicular bone has well defined cortical and spongiosa (medullary) bone structure (Wright, Kidd and Thorp, 1998; Dyson, Van

Thielen and Murray, 2010), though cortical thickness can vary significantly between individuals (Butler *et al.*, 2017). The palmar (flexor) surface of cortical bone is covered with fibrocartilage and is the surface over which the DDFT courses (Gabriel *et al.*, 1998). This surface has a prominent mid-sagittal ridge (Butler *et al.*, 2017). The dorsal and distal cortices of the bone that contribute to articular surfaces are covered with hyaline cartilage (Gabriel *et al.*, 1998). The concave distal surface of the bone contains a variable number of distal synovial invaginations (or fossae) (Butler *et al.*, 2017). The contour of the proximal border of the bone can vary between individuals but is typically symmetrical (Butler *et al.*, 2017). The spongiosa consists of organised trabecular bone (Gabriel *et al.*, 1998).

The navicular bursa is a subtendinous synovial cavity positioned between the palmar aspect of the navicular bone and the dorsal surface of the DDFT (Davies and Philip, 2007). The distal sesamoidean impar ligament contributes to the distal border of the bursa. The suspensory ligaments of the navicular bone and the 'T' ligament form the proximal border (McIlwraith, Nixon and Wright, 2015). The bursa contains synovial fluid, with the majority contained within the proximal recess and a smaller volume distally (Daniel *et al.*, 2016). The normal navicular bursa has an appearance typical of a synovial fluid filled structure on MRI, with high signal intensity on T2 weighted and fat suppressed sequences, in which the location of the bursa is clearly demarcated. The bursa has an intermediate intensity on proton density images and low signal intensity on T1 weighted images (Murray and Werpy, 2010).

A number of ligaments contribute to the suspensory apparatus of the navicular bone (Davies and Philip, 2007). The collateral distal sesamoidean ligament (also referred to as the suspensory ligament of the navicular bone) originates from the abaxial aspects of the distal first phalanx with additional attachments to the abaxial aspects of the second phalanx (Bowker, 2011). The branches of the ligament course distally and attach to the abaxial margin of the navicular bone (Fürst and Lischer, 2019). The ligaments merge to form a sagittal union along the proximal aspect of the navicular bone, contributing to the distal scutum (Denoix, 2000). The normal ligament is biaxially symmetrical. The ligament has a typical organised soft tissue appearance on MRI, with low intensity signal on both

T1 and T2 weighted sequences. The high intensity signal of the adjacent distal interphalangeal joint and navicular bursa outlines the margins of the collateral distal sesamoidean ligament on fat suppressed, T2, T2\* and proton density images (Dyson, Van Thielen and Murray, 2010).

The associated chondrosesamoidean ligaments course distally joining the abaxial aspects of the navicular bone to the axial aspects of the collateral cartilages of the foot and the palmar processes of the third phalanx (Fürst and Lischer, 2019). The ligaments have a low signal intensity on T1 and T2 weighted sequences, though a region of increased signal may be evident at the ligament origins on fat suppressed images (Dyson and Nagy, 2011).

The distal sesamoidean impar ligament is a short ligamentous structure that courses from the distal margin of the navicular bone to the opposing palmar aspect of the third phalanx adjacent to the insertion of the deep digital flexor tendon (Bowker, 2011). The fan like shape of the ligament spans the entire distal length of the navicular bone (Davies and Philip, 2007). The ligament has projections from the distal interphalangeal joint interspersed between the ligamentous fibres (Dyson, Van Thielen and Murray, 2010). As a result, the distal sesamoidean impar ligament can be difficult to evaluate on MR images, with a heterogeneous signal. Ligamentous fibres have low signal intensity on T1 and T2 weighted sequences though regions of synovial invaginations have high signal on T2-weighted and fat suppressed images (Dyson, Van Thielen and Murray, 2010).

### **1.3.5 The distal interphalangeal joint and associated soft tissues**

The distal interphalangeal is principally an articulation between the second phalanx and the distal phalanx, which both also articulate with the navicular bone (Dyson, 2011a). The articular surfaces are covered in smooth articular cartilage overlying the subchondral bone. A small, smooth depression in the subchondral bone covered by a compensatory increase in thickness of articular cartilage can be normal in the axial distal phalanx (Dyson, Van Thielen and Murray, 2010). The joint capsule extends proximally to form the dorsal and palmar pouches of the distal interphalangeal joint (Davies and Philip, 2007). The palmar pouch can also be divided into the larger proximal pouch and a less voluminous distal pouch, close to the distal sesamoidean impar ligament

(Bowker, 2011). The joint is supported by many soft tissues within the foot, including those associated with the navicular bone and the collateral cartilages of the foot. The joint capsule is closely associated with adjacent soft tissues (Bowker, 2011), though in the normal joint there is little synovial proliferation (Dyson, Van Thielen and Murray, 2010). The fluid filled synovial cavity of the distal interphalangeal joint is readily identified with high signal intensity on T2 weighted and fat suppressed sequences, though signal is intermediate on proton density images and low on T1 weighted sequences (Murray and Werpy, 2010). Articular cartilage has an intermediate signal intensity on most sequences, which may be accentuated by the contrasting signal from synovial fluid. Definition of the articular margins can be challenging (van Zadelhoff *et al.*, 2020). The underlying subchondral bone has low signal intensity in the normal horse (Dyson, Van Thielen and Murray, 2010).

The distal interphalangeal joint is supported by collateral ligaments biaxially. These, short, strap like ligaments originate from the collateral fossa of the second phalanx and course in an oblique distopalmar direction (perpendicular to the horizontal ground surface) to insert in the collateral fossa of the third phalanx (Denoix *et al.*, 2011). The ligaments are symmetrically located, though some asymmetry in cross-sectional area and a longer lateral collateral ligament may be normal (Murray *et al.*, 2007). The collateral ligaments have a homogenous low intensity signal on all MRI sequences. The cortex of the collateral fossae has a well-defined margin (Dyson, Van Thielen and Murray, 2010).

### **1.3.6 The phalanges and associated soft tissue structures**

The bony column of the digit is formed by the three aligned phalangeal bones. The first phalanx is the continuation of the bony column of the limb distal to the metacarpophalangeal joint, where it articulates with the third metacarpal bone (Budras *et al.*, 2009b). It is the longest of the phalanges with a tubular shape which flares slightly at the metaphyses to form articular surfaces. The first phalanx articulates with the second phalanx to form the proximal interphalangeal joint (Davies and Philip, 2007)(Budras *et al.*, 2009a). The first phalanx has well defined cortical and medullary components, with significant ligamentous attachments along its palmar surface. Only the distal most aspect of

the first phalanx is usually evident in MR imaging studies of the foot; therefore, it will not be discussed any further.

The second phalanx is shorter and more cuboidal in shape compared to the first phalanx (Budras *et al.*, 2009b). The second phalanx also has well defined cortices and a medulla. Proximally it has a concave shape to form its articular surface of the proximal interphalangeal joint (Davies and Philip, 2007). The proximopalmar aspect of the second phalanx is intimately associated with the fibrocartilaginous middle scutum, which acts as the insertion of the branches of the superficial digital flexor tendon and the straight distal sesamoidean ligament (Carnicer, Coudry and Denoix, 2013). Distally it has a convex shape and contributes to the distal interphalangeal joint, articulating with the third phalanx and the navicular bone. The cortical bone has a uniform low intensity on all sequences on MRI (Dyson, Van Thielen and Murray, 2010). The cancellous bone of the medulla has an intermediate intensity on T1 and T2 weighted sequences, with a low signal intensity on fat suppressed images (Dyson, Van Thielen and Murray, 2010). The MRI characteristics of the articular components of the proximal interphalangeal joint are comparable to those described for the distal interphalangeal joint and will not be described in any further detail.

The third phalanx has a wedge-shaped appearance in a sagittal plane and a crescent shaped distal margin (Davies and Philip, 2007). The crescent shape extends to form the two palmar processes. The bone has no medullary cavity but has an intricate network of vascular channels that give the bone a porous appearance (Dyson, 2011b). The largest of these are the solar foramina which form a crescent shaped channel, the solar canal (Schade, Arnoczky and Bowker, 2017). A parietal sulcus is present on the abaxial aspects of each palmar process. The bone has complex interactions with many other structures within the foot, these anatomic relationships are described in greater detail in their relevant section. The third phalanx articulates with the second phalanx and the navicular bone to form the distal interphalangeal joint. The bone is suspended from the hoof capsule by the laminae. The facies flexoria on the palmar aspect of the third phalanx is the site of insertion of the deep digital flexor tendon.

In the forelimb the common digital extensor muscle has three heads in the forelimb (humeral head, radial head and ulnar head) which contribute to the

common digital extensor tendon, which courses distally to insert on the extensor process of the dorsoproximal aspect of the third phalanx (Budras *et al.*, 2009a). The common digital extensor tendon is joined by the extensor branches of the suspensory ligament (third interosseous muscle) and has a smaller attachment to the dorsoproximal aspect of the second phalanx (Budras *et al.*, 2009a). In the hindlimb the long digital extensor tendon is the primary extensor tendon. This arises from the long digital extensor muscle which originates on the lateral condyle of the femur (Budras *et al.*, 2009b). The long digital extensor tendon is joined by the lateral digital extensor tendon in the proximal metatarsus. The anatomy of the extensor tendons in the distal limb can be considered comparable to that of the forelimb. The extensor tendon has a homogeneous low signal intensity on all MRI sequences in the normal horse (Dyson, Van Thielen and Murray, 2010).

### **1.3.7 The collateral cartilages of the foot and the digital cushion**

The collateral cartilages of the foot are positioned on the proximal aspect of the medial and lateral palmar processes of the third phalanx (Dyson and Nagy, 2011). The shape and size of the collateral cartilages of the foot can vary between individuals but they extend to the level of the coronary band or more proximal (Davies and Philip, 2007). The cartilages are formed of fibrocartilage in the adult horse, though they can be variably ossified, which may be an incidental finding. Ossification is frequently biaxial but may be more extensive within the lateral cartilage (Jones and Dyson, 2015). Unossified collateral cartilages have a mildly heterogeneous, intermediate signal intensity on T1 and T2 weighted images, with low signal intensity on fat suppressed images (Dyson and Nagy, 2011). Ossified regions of the collateral cartilages can have a variable appearance but are generally comparable to other bony structures with cortical bone of uniform low intensity on all sequences and cancellous bone with an intermediate intensity on T1 and T2 weighted sequences, with a low signal intensity on fat suppressed images (Dyson and Nagy, 2011).

The collateral cartilages are supported by several ligamentous structures including the chondrocoronal, chondrosesamoidean, chondroungular and chondrocompedal ligaments (Dyson and Nagy, 2011). The biaxial, paired chondrocoronal ligaments attach the dorsal aspect of the collateral cartilages of

the foot to the abaxial aspect of proximal collateral ligaments of the distal interphalangeal joint and the distal aspect of the second phalanx and have a sheet like appearance (Davies and Philip, 2007; Dyson, 2011b). The chondroungular attaches the distopalmar aspect of the collateral cartilages to the ipsi-axial palmar process of the third phalanx (Denoix, 2000). The chondrocompedal ligament attaches the palmaroproximal aspect of the collateral cartilage of the foot to the ipsi-axial palmar aspect of the centre of the first phalanx (Denoix, 2000; Dyson, 2011b). The chondrosesamoidean ligaments are biaxially paired, short ligaments that course between the abaxial margins of the navicular bone to the axial margin of the collateral cartilages and the palmar processes of the third phalanx (Norvall *et al.*, 2021). The ligaments have a generalised low signal intensity on T1 and T2 weighted sequences (Dyson and Nagy, 2011). Focal regions of mild increased signal intensity can be observed at the ligament origins in fat suppressed images (Dyson and Nagy, 2011).

The digital cushion is formed by two components, the toric part and the cunean part (Davies and Philip, 2007). The digital cushion is positioned in the palmar aspect of the foot and fills the region bound by the collateral cartilages of the foot the palmar aspect of the deep digital flexor tendon, the frog and the bulbs of the heel (Davies and Philip, 2007). The digital cushion is primarily composed of collagenous, elastic and adipose tissue, but also contains nervous and vascular structures (Faramarzi *et al.*, 2017). The digital cushion has a heterogenous low to intermediate signal intensity on all MRI sequences (Dyson, Van Thielen and Murray, 2010).

### **1.3.8 The laminae**

The laminae form the junction that suspend the third phalanx from the keratinised hoof wall. It has a complex gross and microscopic anatomy with interdigitating epidermal and dermal components and subsequent attachment to the periosteal surface of the third phalanx (Pollitt, 2010). The laminae extend along the entire hoof wall. The laminae appear hyperintense on MRI sequences, with a general pattern of increased intensity within the deeper layers, including the sublaminae dermis (Dyson, Van Thielen and Murray, 2010). Other regions of the corium are also evident on MRI of the foot, but these will not be discussed any further (Powder *et al.*, 2020). The keratinised components of the hoof wall



have low signal intensity and in some circumstances application of a fatty material to the external aspect of the hoof can be useful to delineate the capsule (Grundmann *et al.*, 2015).

## **1.4 Magnetic resonance image quality in equine practice**

### **1.4.1 Image quality and its influence on pathology assessment**

Image quality has an important influence on the diagnostic value of any imaging modality. There are multiple key components that determine image quality including resolution, contrast, noise and artefacts (Dowsett, Kenny and Johnston, 2006; Ding, 2018). Other factors related to practicalities of imaging such as positioning, and centring are also important, and these are primarily at the discretion of the system operator. The differences in the underlying physics and applications of different imaging modalities means that the relative importance of resolution, contrast and noise varies between modalities (Dowsett, Kenny and Johnston, 2006). Evaluation of MR image quality is complex due to its nature as a multiplanar modality and variable image weighting.

Assessment of medical image quality can be challenging (McRobbie *et al.*, 2017k). Objective factors that contribute to the image such as the signal-to-noise ratio, contrast-to-noise ratio and resolution can be useful markers of image quality and assist in quality assurance (McRobbie *et al.*, 2017k, 2017e). Diagnostic confidence, where observers grade their confidence in pathology assessment can also be useful and is an important outcome measure. However, comparisons of diagnostic confidence between different acquisition scenarios (for example different MRI systems, or different sequence combinations) requires careful matching of cases to prevent confounding by variation in pathology. In addition, diagnostic confidence is fundamentally susceptible to avoid confounding from incorrect diagnoses, i.e. it is possible for an observer to be very confident in an incorrect diagnosis (Ng and Palmer, 2007). Subjective assessment of image quality by experts is a valuable intermediary, giving a holistic impression of the diagnostic value of an MR image. This methodology is often used in clinical studies investigating factors that influence image quality in human and veterinary medicine (McKnight *et al.*, 2004; Stahl *et al.*, 2009; Bolen, Audigié, *et al.*, 2010; Chow, Rajagopal and Paramesran, 2016; Ding, 2018).

Maximising MR image quality ensures that observers are presented with high quality images for interpretation, which ultimately optimises patient outcomes in a clinical context. However, image quality is one of many considerations that dictates the practicalities of how MRI is performed in equine clinical practice, including costs, methods of restraint, patient factors, the nature of pathology under investigation and potential treatment methods (Werpy, 2007; Porter and Werpy, 2014).

### **1.4.2 Factors influencing diagnostic image quality**

Understanding the underlying elements that contribute to image quality in equine clinical MRI is essential to optimise the diagnostic value of this modality (Murray and Werpy, 2010). Previous research has demonstrated the significance of a number of factors.

#### **1.4.2.1 Field strength**

The magnet is the fundamental component of the MRI system and to an extent dictates many other aspects of system design. The magnetic field strength is an important dictator of MR image quality. Increasing magnetic field strength results in an associated increase in signal-to-noise ratio (Bolas, 2010). Other factors such as spatial resolution and imaging time will also influence signal-to-noise ratio, but these are typically optimised to accommodate the limits of a particular magnetic field strength (McRobbie *et al.*, 2017g). In a reciprocal manner, an increase in signal-to-noise ratio (as a result of increased magnetic field strength) can be substituted for benefits in other parameters such as faster acquisition time or resolution.

The magnetic field strength of a magnet is measured in tesla (T). Most systems used in a clinical setting in human orthopaedics utilise a magnet between 1.5-3T (McRobbie *et al.*, 2017g). Ultra-high-field systems (for example 7T) are increasingly common in human medical research and certain clinical settings (Ladd *et al.*, 2018). In the context of equine MRI, a system with magnetic field strength <1T would be considered low-field MRI and >1T would be considered high-field (Werpy, 2007). These definitions are arbitrary and can differ between clinical disciplines. In clinical equine imaging, most high-field systems are

between 1.5-3T and the majority of low-field systems have a field strength of approximately 0.3T (Bolas, 2010; Schramme and Segard-Weisse, 2020).

Magnets can typically be grouped into four categories depending on how the magnetic field is generated: air-cored resistive magnets, iron-cored electromagnets, permanent magnets and superconducting magnets (McRobbie *et al.*, 2017g). In the equine field most units are low-field and utilise a permanent magnet (Bolas, 2010; Hallmarq Veterinary Imaging, 2019; Schramme and Segard-Weisse, 2020). Permanent magnets have the magnetic field induced during the manufacturing phase and are limited to low-field applications. The systems often utilise an open 'U' shaped magnet (Mair *et al.*, 2005; Bolas, 2010). By the nature of their permanent magnetic field, these systems have relatively low purchase and running costs (McRobbie *et al.*, 2017g). Some systems utilise a magnet designed initially for human imaging. A system manufactured by Hallmarq Veterinary Imaging Ltd. specifically for imaging the standing equine patient uses a permanent 'U' shaped magnet that can be moved relative to the patient's limb (Mair *et al.*, 2005). The orientation of the magnet can also change the direction of the main magnetic field. For example, in the equine standing low-field MRI system the main magnetic field is horizontally orientated, which is perpendicular to the long axis of the limb but in most high-field systems the limb is positioned parallel to the static magnetic field (Murray and Werpy, 2010). Some low-field open magnets are extremely versatile and can be used to image the limb as far proximal as the stifle and can even accommodate the head (Schramme and Segard-Weisse, 2020).

High-field MRI systems use superconducting magnets (Werpy, 2007).

Superconduction is the property of a material that has no electrical resistance when cooled to approximately 0K (-273°C). MRI systems use the material in wire form, which is wound in large coils and the ends effectively joined with a superconducting switch to form a closed loop (Bolas, 2010). An initial current is applied to the loop and given that the loop is cooled, the superconducting properties allow the current to continue to flow without on-going loss (Bolas, 2010). The cooling required for superconduction is achieved using a cryostat chamber of liquid helium, which has a boiling point of 4.2K (McRobbie *et al.*, 2017g). There is a low level of continuous loss of liquid helium and MR systems

employ different methods to limit this. A typical design utilises an arrangement of vacuums and sequential temperature shields (McRobbie *et al.*, 2017g). However, the maintenance of this cryostat results in higher running costs than that of a permanent magnet (McRobbie *et al.*, 2017g). In addition, in the case of a fault with the magnet, re-energising the coil is an expensive procedure (McRobbie *et al.*, 2017g). Superconducting magnets are typically arranged as a cylinder, therefore the resultant MR system has a similar functional shape with a cylindrical bore (Bolas, 2010; Schramme and Segard-Weisse, 2020). Most of the systems have an approximately 50-70 cm internal bore diameter which can accommodate the distal limb of equine patients (Bolas, 2010; Schramme and Segard-Weisse, 2020), though some systems can accommodate the equine head and cervical region (Werpy, 2007). The formation of cylindrical shaped coils causes the main magnetic field to be orientated along the length of the coil, which is parallel to the long axis of the limb (Murray and Werpy, 2010).

Given the association between field strength and image quality, much research has been directed at evaluating this relationship (and its impact on diagnostic outcomes) in human medicine (Rutt and Lee, 1996; Ghazinoor, Crues and Crowley, 2007; Edelman, 2014). Investigations in equine orthopaedic imaging with cadaver materials have reported superior image quality with high-field MRI (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). High-field images often provided additional detail that was particularly beneficial for the evaluation of small anatomical structures or for the evaluation of small lesions (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). These cadaveric studies demonstrated that low-field imaging did have useful diagnostic value for many structures of the foot. Combining this with the practical benefits of low-field imaging the equine patient, standing low-field imaging has been extremely popular in clinical practice (Bladon, 2014).

However, previous literature assessing the influence of field strength on image quality in equine orthopaedics has confounders and limitations related to the use of cadaver materials, deviations from clinical positioning of limbs, use of pulse sequences not optimised for specific MRI systems and comparison of different system designs (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). There is

limited information to characterise the influence of field strength on MR image quality in clinical equine patients.

#### **1.4.2.2 Patient motion**

Motion is a well-recognised cause of reduced image quality in human medical imaging (Zaitsev, Maclaren and Herbst, 2015; McRobbie *et al.*, 2017f). Subtle motion often appears as a blurring or ghosting artefact in the phase encoding direction (McRobbie *et al.*, 2017f), but gross movement of the patient out of the field of view is also possible. This challenge is also common in equine veterinary MRI, where patient restraint requires additional consideration (Murray and Werpy, 2010).

Standing equine MRI is deemed to be most susceptible to motion artefact, compared to that performed with the patient under general anaesthesia (Porter and Werpy, 2014). In standing equine MRI, motion is not entirely under the control of the system operator and an important contribution is made by compliance of the standing patient (Murray and Werpy, 2010). Nonetheless, motion can still occur in the anaesthetised patient, particularly as a result of respiratory movements (McRobbie *et al.*, 2017f; Schramme and Segard-Weisse, 2020). In most situations gross patient motion is best addressed by repetition of the affected sequences (McRobbie *et al.*, 2017f). However, several other methods are used in practice to reduce the influence of motion on image quality. These methods can be divided into three categories: motion prevention, artefact reduction and motion correction (Zaitsev, Maclaren and Herbst, 2015). Patient management such as adequate sedation, careful handling and stable positioning are key in motion prevention in the standing equine patient (Werpy, 2010). Artefact reduction methods include use of fast sequences and careful study planning (Zaitsev, Maclaren and Herbst, 2015). Motion correction techniques, such as the use of navigator echoes to monitor patient motion are particularly helpful in the standing equine patient (McKnight *et al.*, 2004; Murray and Werpy, 2010). Some sequences are more challenging to acquire in the face of patient motion (for example STIR sequences) and this is often a manifestation of their relatively long acquisition time (Schramme and Segard-Weisse, 2020).

The equine foot is deemed to be the region of the limb least susceptible to patient motion due to its position against the ground during standing MRI (McKnight *et al.*, 2004). This is fortuitous given the clinical significance of equine foot lameness and the reliance on MRI for assessment of this region (Bladon, 2014; Barrett *et al.*, 2017). An equine cadaver study demonstrated the value of motion-correction techniques for standing MRI of the carpus and tarsus (McKnight *et al.*, 2004). However, patient motion is still reported as a significant disadvantage for standing equine MRI (Porter and Werpy, 2014; Schramme and Segard-Weisse, 2020). Despite this there is very limited evidence reporting the influence of patient motion on image quality for MRI of the equine foot and particularly the influence of general anaesthesia of the patient.

#### **1.4.2.3 Weight-bearing of the limb**

An additional consideration for standing equine MRI is the influence of weight-bearing, compared to the non-weight-bearing limb in the patient under general anaesthesia. The differences between non-weight-bearing and weight-bearing MRI has been utilised in human orthopaedic imaging, where it is recognised that imaging during physiologic positioning can be useful in assessment of pathology (Shapiro and Gold, 2012). It is predictable that positioning the equine limb in a weight-bearing manner can alter the appearance of some anatomical structures, principally as a result of changes in the orientation of the phalangeal bony column (Murray *et al.*, 2009). It is worth noting that due to the popularity of low-field standing MRI, many of those acquiring and interpreting MR images in equine clinical practice are more familiar with weight-bearing positioning (Hallmarq Veterinary Imaging, 2019).

Assessment of articular cartilage is recognised as a particular factor that is strongly influenced by positioning and weight-bearing. An early cadaver study demonstrated the value of low-field non-weight-bearing MRI for the evaluation of articular cartilage lesions, but highlighted that the findings may not translate directly to weight-bearing MRI (Olive, 2010). In the weight-bearing state the articular cartilage may be compressed into contact, making definition of the articular surface more challenging (Olive, 2010). This was subsequently confirmed in a cadaver study showing that delineation of the distal interphalangeal joint margins was more challenging in the weight-bearing limb

(Evrard *et al.*, 2019). A further study has demonstrated that low-field MRI was not a sensitive modality for the identification of naturally occurring cartilage defects, which may have less defined boundaries compared to those of experimental lesions (van Zadelhoff *et al.*, 2020). These findings indicate that non-weight-bearing MRI, such as that performed under general anaesthesia, may be preferable for the identification of distal interphalangeal joint articular cartilage pathology.

Much of the previous literature has focused on experimental models and utilised cadaver materials. There is little information characterising the influence of weight-bearing on image quality for MRI of the equine foot in live clinical patients.

#### **1.4.2.4 System components and other attributes**

The type of magnet and overall design of the system also influences other system components and features. Intrinsic properties of the magnet, such as closed or open in nature, can have important influences on factors influencing image quality, such as magnetic field homogeneity (McRobbie *et al.*, 2017g).

Radiofrequency coils have a fundamental function in transmitting and receiving radio waves (Bolas, 2010; McRobbie *et al.*, 2017g). Coils vary between systems and some may utilise separate transmit and receive coils, with others utilising a coil that combines these functions (Westbrook, Roth and Talbot, 2011d). It is important that the radiofrequency coil is suitable for the purpose and anatomical region of interest. To optimise signal, the coil should closely match the volume of the region of the patient to be imaged (McRobbie *et al.*, 2017g). Many systems using a magnet modified from human imaging also utilise a modified human coil. However, equine specific coils such as a coil shaped to conform to the shape of the equine foot, have been developed for use with the low-field standing MRI system (Bolas, 2010).

The MRI computer and software are also important considerations to ensure that the operator interface readily permits those acquiring images to do so with sufficient control to manipulate the study to suit patient requirements (Bolas, 2010). Magnet and system attributes (such as field strength and shielding) can

have important influences on patient and personnel safety (Bolas, 2010; McRobbie *et al.*, 2017c).

#### **1.4.2.5 Imaging protocols and the system operator**

Most clinical imaging utilises protocols that have been optimised for a particular system (McRobbie *et al.*, 2017k). However, many institutions may have a preferred protocol tailored to the population of patients they typically treat (Murray and Werpy, 2010). Nonetheless, a significant proportion of decision-making during image acquisition is at the discretion of the system operator. Ensuring that the patient is well positioned, that the limb is placed in the isocentre of the magnet, piloting is adequate and that sequence alignment is correct are just some of the examples of how the operator can influence ultimate image quality (Werpy, 2010). This is particularly the case in the equine foot given the complicated anatomy of this region, with soft tissue structures orientated in multiple planes (Dyson, Van Thielen and Murray, 2010).

#### **1.4.2.6 Other artefacts**

Motion artefact is regularly encountered in equine MRI but a multitude of other artefacts also occur regularly and can reduce image quality (Murray and Werpy, 2010). These can principally be divided as related to inhomogeneity or as a manifestation of digital imaging (McRobbie *et al.*, 2017f). In many situations, artefacts are an inherent feature of images and can be minimised but not eliminated. Therefore, their presence may need to be tolerated and considered during interpretation of the study (Westbrook, Roth and Talbot, 2011a).

Magnetic field inhomogeneity is an important origin of artefacts with a variety of primary causes (McRobbie *et al.*, 2017f). These can originate from heterogeneity of the static magnetic field due to inadequate shimming or fluctuations in temperature (Murray and Werpy, 2010). Endogenous patient factors such as haemoglobin and haemosiderin in the region of the imaged anatomy can cause susceptibility artefacts (Westbrook, Roth and Talbot, 2011a). However, the equine foot has some unique factors that can result in problematic magnetic susceptibility or metal artefacts. Remnants of shoeing nails can result in significant artefact, therefore complete removal and radiographic screening is



preferable (Murray and Werpy, 2010). Other materials commonly used in the equine foot such as farriery artificial hoof construction materials and materials used for packing during radiographic examination can also cause magnetic susceptibility artefacts (Murray and Werpy, 2010; Schramme and Segard-Weisse, 2020).

The magic angle effect is an increase in signal intensity that does not accurately reflect tissue signal. This occurs in tissues with collagen fibres that are orientated at approximately  $55^\circ$  to the static magnetic field and is commonly encountered in orthopaedic imaging (Bydder *et al.*, 2007; Murray and Werpy, 2010). Due to its origin this effect can manifest in different anatomical structures in different MRI systems, depending on the orientation of the static magnetic field relative to the imaged anatomy (Schramme and Segard-Weisse, 2020). The complex orientation of the soft tissues of the foot results in various structures having a propensity for this effect to occur. For example, in standing low-field systems the collateral ligaments of the distal interphalangeal joint (especially the lateral ligament), lateral aspect of the distal deep digital flexor tendon and the oblique distal sesamoidean ligament are common sites for the magic angle effect to occur (Smith, Dyson and Murray, 2008; Spriet and Zwingenberger, 2009; Gutierrez-Nibeyro *et al.*, 2011; Schramme and Segard-Weisse, 2020). In high-field systems the insertion of the deep digital flexor tendon is a common site for magic angle effect to occur (Murray and Werpy, 2010). Careful positioning and inclusion of sequences less susceptible to magic angle effect (such as long echo time T2 weighted fast spin echo sequences) can minimise the occurrence of this effect (Murray and Werpy, 2010).

Blood flow can also result in motion artefact (McRobbie *et al.*, 2017f). The distal location of the foot means that large vasculature is less commonly encountered during imaging of this region and flow is less problematic than in more proximal regions (Murray and Werpy, 2010). Nonetheless methods to reduce the impact of flow on image quality, such as encoding inversion, saturation bands and use of a tourniquet can be utilised if required (Murray and Werpy, 2010; McRobbie *et al.*, 2017f).

Chemical shift and phase cancellation artefacts are two closely associated artefacts that occur due to the inherently subtle differences in the precessional

frequencies of water and fat when these tissues are adjacent to each other (McRobbie *et al.*, 2017f). Chemical shift results in relative shifting of signals which can produce ghosting type manifestation on the image and can be particularly problematic in high-field systems (Murray and Werpy, 2010). Phase cancellation is specific event at certain echo times where the signal for water and fat effectively cancel resulting in a phase cancellation artefact (Murray and Werpy, 2010). This most commonly occurs on low-field standing system T2\* gradient echo sequences and can occur in the navicular bone (Schramme and Segard-Weisse, 2020). Evaluation of other sequences helps to characterise this cancellation, which can indeed be helpful to the interpreter in some situations (Murray and Werpy, 2010).

Partial volume averaging is an inherent property of converting a contiguous three dimensional structure into multiple slices, where a voxel may contain multiple tissues (McRobbie *et al.*, 2017f). This effect can be minimised but not entirely avoided and interpreters should be particularly aware of this artefact when assessing small structures or tissue margins (Murray and Werpy, 2010). Multiple other artefacts such as phase wrap artefacts and temperature artefacts (for STIR sequences) can also occur but are less commonly problematic (Schramme and Segard-Weisse, 2020).

## **1.5 Observer agreement during equine magnetic resonance imaging**

### **1.5.1 Image interpretation and inter-observer agreement**

The end-stage aim of increased image quality is to improve diagnostic accuracy in a clinical setting. However, the observer also performs a fundamental role in determining the diagnostic accuracy of MR image interpretation (Williams and Drew, 2019). Interpretation of diagnostic images is a complex process. In a clinical setting, observers are required to assess image quality, identify any image abnormalities, interpret the lesions, establish differential diagnoses and determine the clinical significance for the patient (Collins and Ehman, 2012; Thrall, 2018). In many clinical situations combinations of modalities are required to overcome their individual limitations and provide a more holistic imaging assessment.

The multiplanar and multisequence nature makes interpretation of MR images particularly challenging (Ganesan *et al.*, 2018; Williams and Drew, 2019). Observers may use different search patterns but a systematic and consistent approach is important to avoid search errors (Williams and Drew, 2019; Schramme and Segard-Weisse, 2020). Inter-observer agreement in pathology assessment is a useful output measure for image interpretation (Gwet, 2014). Correlation of MRI findings with those of histologic assessment has validated the use of this imaging modality (Kleiter *et al.*, 1999; Murray, Blunden, *et al.*, 2006; Murray *et al.*, 2007; Dyson, Blunden and Murray, 2008; Sherlock, Mair and Blunden, 2008; Dyson *et al.*, 2010; Kottmeier *et al.*, 2020). However, in a clinical context, consistency in MR image interpretation between observers is also important because it is a crucial determinant of case decision making including treatment and prognosis. In addition, consistency in interpretation is key at the level of the profession to ensure that research findings and updates in clinical practice are accurately disseminated.

### **1.5.2 Observer agreement in equine magnetic resonance imaging**

There is currently limited evidence to document inter- or intra-observer agreement for any aspect of equine MR image interpretation. Assessment of image quality of cadaver foot specimens from different MRI systems by three observers demonstrated some differences in perceived image quality between the observers (Bolen, Audigié, *et al.*, 2010). A further study documented that observer experience appeared to influence the ability of observers to recognise differences in positioning during image acquisition for MR imaging of the equine foot (Evrard *et al.*, 2019). A study focused on assessment of the distal sesamoidean impar ligament demonstrated that inter-observer agreement for the value of different sequences varied between sequences and following saline arthrography (Berner *et al.*, 2020). A study has also documented agreement between observers when assessing features of MR images of the foot that may be influenced by diagnostic anaesthesia procedures before imaging (Black *et al.*, 2013). Attributes assessed included features such as distal interphalangeal joint fluid volume and the presence of needle tracts (Black *et al.*, 2013).

The literature presents a significant body of detailed guidance on interpretation of equine foot MRI studies and a vast array of lesions have been well

characterised (Dyson and Murray, 2010; Dyson, Van Thielen and Murray, 2010; Murray and Werpy, 2010; Schramme and Segard-Weisse, 2020). However, there is very little evidence to document inter-observer agreement for pathology assessment of equine MR images.

## **1.6 Aims**

The first aspect of the project investigated key factors influencing image quality and image interpretation for MRI of the equine foot in clinical practice. Based on previous literature the focus of the image quality investigation was on the influence of general anaesthesia (incorporating motion and weight-bearing) and field strength on MR image quality in live equine patients. The image interpretation aspect of the project investigated the inter-observer agreement of expert observers for the assessment of pathology of the equine foot from clinical MRI of live patients.

The following project aims were therefore established:

1. To investigate the influence of general anaesthesia on image quality assessment for MRI studies of the equine foot acquired in a clinical context. This was achieved by comparing subjective quality of MR images of the equine foot from low-field systems acquired with the patient under standing sedation compared to those acquired with the patient under general anaesthesia.
2. To investigate the influence of field strength on image quality for MRI studies of the equine foot acquired in a clinical context. This was achieved by comparing subjective image quality between MR images from low-field and high-field systems with the patient under general anaesthesia.
3. To investigate inter-observer agreement for the assessment of pathology of clinically relevant anatomical structures of the equine foot during interpretation of MRI studies.

# Chapter 2 The influence of general anaesthesia on perceived image quality of clinical magnetic resonance imaging of the equine foot

## 2.1 Introduction

Magnetic resonance imaging (MRI) is a key diagnostic imaging modality for assessment of the equine foot (Smith, 2015; Barrett *et al.*, 2017). A significant proportion of clinical equine MRI is performed in the standing patient using a low-field MRI system custom designed for equine use (Bladon, 2014; Hallmarq Veterinary Imaging, 2019). Standing MRI has the important benefit of avoiding the requirement for general anaesthesia of the equine patient (Dyson *et al.*, 2003). However, other factors such as image quality may also need to be considered when deciding on the most suitable MRI system for a particular clinical scenario.

Patient motion is a common cause of reduced image quality during MRI in human patients (Havsteen *et al.*, 2017; McRobbie *et al.*, 2017f). This factor is exacerbated in veterinary patients, where patient compliance presents additional challenges. Intuitively, patient motion is considered to be most relevant to MRI of the standing sedated equine patient, compared to situations where the patient is under general anaesthesia (Murray and Werpy, 2010; Porter and Werpy, 2014). In both situations, a range of options are available to reduce the influence of mild motion on image quality including suitable patient restraint (for example a consistent level of adequate sedation), careful positioning, use of motion-insensitive sequences, study acquisition planning and post-processing motion-correction techniques (McKnight *et al.*, 2004; Murray and Werpy, 2010). In the case of significant patient motion, repetition of the affected sequence is required (McRobbie *et al.*, 2017f). Given the clinical importance of the equine foot and the value of MRI of this region, it is fortunate that the foot is perceived to be less susceptible to motion compared to other regions of the equine limb. This is principally due to the stability conferred by the position of the feet against the ground surface during standing imaging (McKnight *et al.*, 2004). However, there is relatively limited evidence to document the influence of patient motion on clinical image quality for different MRI systems in equine patients.

Furthermore, other important differences also exist between MR images acquired with the patient under standing sedation and general anaesthesia. Weight-bearing of the limb being imaged can influence the position and appearance of anatomic structures (Bruno *et al.*, 2018). For example, the definition of articular cartilage boundaries, which are typically separated by a thin layer of synovial fluid when non-weight-bearing but may be compressed into contact in the weight-bearing limb (Olive, 2010). A study simulating weight-bearing and non-weight-bearing imaging demonstrated thinning of the articular cartilage in the weight-bearing position, which could make identification of cartilage pathology more difficult (Evrard *et al.*, 2019). Differences in acquisition system design are also important to consider during MRI in weight-bearing and non-weight-bearing positions, for example the anatomic structures susceptible to the magic angle effect may vary with positioning (Busoni and Snaps, 2002; Spriet, Mai and McKnight, 2007; Spriet and McKnight, 2009; Werpy, Ho and Kawcak, 2010).

Despite these well recognised factors, there is limited literature documenting the differences in clinical image quality between low-field imaging performed with the live equine patient under standing sedation and under general anaesthesia.

## **2.2 Aims**

The aim of the study was to assess the influence of general anaesthesia on image quality assessment for MRI studies of the equine foot acquired in a clinical context. This was achieved by comparing perceived quality of MR images of the equine foot from low-field systems acquired with the patient under standing sedation compared to those acquired with the patient under general anaesthesia. It was hypothesised that there would be a difference in image quality between images acquired from a low-field system with the patient standing compared to those acquired from a low-field system with the patient under general anaesthesia.

## **2.3 Study design**

The study had an experimental study design including prospective observer assessment of MRI studies collected retrospectively from clinical cases acquired with different MRI systems.

## **2.4 Methods**

### **2.4.1 Clinically relevant structures of the equine foot**

Given the clinical importance of the foot as an origin of lameness and the significant attention which the foot has received following the introduction of MRI of the equine patient, this region was targeted as the focus of the project. Previous literature was reviewed to assess the distribution of MRI lesions of the foot in horses with foot lameness. Table 2-1 presents the distribution of MR foot lesions from 6 large, relevant studies. Based on these previous reports and overall clinical significance in the investigation of lameness, the following individual structures of the foot were selected for inclusion in the image assessment component of the study:

1. Deep digital flexor tendon
2. Navicular bone
3. Navicular bursa
4. Distal interphalangeal joint
5. Collateral ligaments of the distal interphalangeal joint
6. Third phalanx
7. Distal sesamoidean impar ligament

The phalanges were variably categorised in previous literature, either as separate structures (second phalanx and third phalanx) or grouped (phalanges). In some MRI systems, a study of the foot will not allow full assessment of the

second phalanx, which may require a second study to be performed (centring the isocentre of the magnet over the second phalanx). In addition, evidence suggests that lesions of the third phalanx are more frequent than the second phalanx in horses with foot pain (Murray, Schramme, *et al.*, 2006). Therefore, for the purposes of this study, analysis of the phalanges was limited to the third phalanx.

## **2.4.2 Image assessment platform**

An online image assessment platform was developed to allow observers to evaluate MRI studies of the equine foot. This utilised an online survey platform (Online Surveys, Jisc), which observers interacted with whilst viewing the MRI studies.

### **2.4.2.1 Entire study image quality assessment**

The first component addressed image quality for the whole study using a 1-4 verbal grading scale (Table 2-2). A four-tiered grading scale was deemed to provide sufficient scope for differentiation of image quality. Descriptors were provided for each grade to guide image assessment. These were carefully selected to ensure observers were not limited by overly prescriptive descriptions. This aimed to maximise the value of observer experience in image interpretation. For example, it was deemed important that the presence of artefact, which did not have a significant impact on interpretation, did not automatically result in a significant reduction in image quality grade. The statements regarding study repetition in a clinical setting (relevant for grades 2 to 4) were included to ensure the assessment was closely aligned to the clinical scenario. At the end of each study assessment there was a field for free text comments by the observers. Observers did not receive any specific guidance to dictate the content of free text comments. The entire study image quality assessment was presented as a multiple choice (single answer) question (Figure 2-1) for each MRI study



**Table 2-1 Summary of magnetic resonance imaging lesion distribution in horses with foot pain from six studies.**

Abbreviations: DIP- distal interphalangeal.

Study	Rank of individual anatomical structures by their reported frequency of lesion identification (proportion of cases)						
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>	7 <sup>th</sup>
Mair et al.(2003)	Deep digital flexor tendon (16/35)	Navicular bone/ bursa (11/35)	DIP joint (8/35)	Collateral sesamoidean ligament (5/35)	Distal sesamoidean impar ligament (3/35)	Collateral ligament of the DIP joint (2/35)	
Dyson, Murray, and Schramme (2005)	Deep digital flexor tendon (92/199)	Navicular bone (27/199)	Collateral ligament of the DIP joint (30/199)	Phalanges (14/199)	Distal sesamoidean impar ligament (11/199)	DIP joint (5/199)	
Murray et al. (2006)	Navicular bone (34/34)	Deep digital flexor tendon (31/34)	Third phalanx (27/34) Distal sesamoidean impar ligament (27/34)		DIP joint (26/34)	Navicular bursa (20/34)	Collateral sesamoidean ligament (16/34)
Sampson et al.(2009)	Navicular bone (62/72)	Collateral sesamoidean ligament (54/72)	DIP joint (36/72)	Deep digital flexor tendon (32/72) Navicular bursa (32/72)		Distal sesamoidean impar ligament (26/72)	
Dyson and Murray (2010)	Collateral ligament of DIP joint (179/584)	Deep digital flexor tendon (149/584)	Navicular bone (81/584)	Phalanges (25/584)	Distal sesamoidean impar ligament (20/584)	DIP joint (8/584)	Collateral sesamoidean ligament (3/584)
Gutierrez-Nibeyro, Werpy, and White (2012)	Navicular bone (62/79)	Navicular bursa (44/79)	Deep digital flexor tendon (43/79)	DIP joint (42/79)	Collateral ligament of the DIP joint (31/79)	Distal sesamoidean impar ligament (24/79)	Phalanges (19/79) Collateral sesamoidean ligament (19/79)

Please grade the overall image quality for the study as a whole.

- Grade 1- Textbook quality
- Grade 2- High diagnostic quality
- Grade 3- Satisfactory diagnostic quality
- Grade 4- Non-diagnostic

**Figure 2-1 Format of the entire study image quality assessment interface.**

**Table 2-2 Grading system used for image quality assessment of entire magnetic resonance imaging foot studies.**

Grade	Summary	Grading descriptor
1	Textbook quality	The study is well composed with optimal tissue definition. Images in the study are sufficient quality to be printed in a textbook as an anatomic image or demonstration of pathology. The study would not be repeated in a routine clinical context.
2	High diagnostic quality	The study is well composed, though mild loss of tissue definition may be evident. Artefact may be present but it does not limit interpretation. The study would not be repeated in a routine clinical context.
3	Satisfactory diagnostic quality	The study is well composed. Loss of tissue definition is readily evident, though structures can be assessed sufficiently. Artefact may be present but it does not significantly limit interpretation. The study would not be repeated in a routine clinical context.
4	Non-diagnostic	The study is poorly composed, is severely affected by artefact or has loss of tissue definition which prevents assessment of significant structures. The study would be repeated in a routine clinical context.

#### 2.4.2.2 Individual anatomical structure assessment

The second component of the platform assessed anatomical structures of the foot (section 1.4.1). The structures were selected based on previous studies describing the distribution of lesions identified on equine foot MRI. Evaluation included image quality assessment and pathology assessment for each structure. To maintain consistency, all grading systems used a 4-point scale. The individual structure image quality assessment was based on the descriptors for entire study assessment with specific adaptations to highlight the unique characteristics of the individual anatomical structure. Descriptors were formulated primarily around the Roentgen signs used in diagnostic radiology: size, shape, location, number, margination and opacity (Thrall, 2018). The descriptors also incorporated the terminology used in the standard descriptions of the individual anatomical

structures in reference texts for MRI of the equine foot (Dyson, Van Thielen and Murray, 2010) and descriptions used in previous literature evaluating image quality of equine MR images (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010).

Prior to finalisation the individual anatomic structure assessment grading system was presented to a number of parties, including diagnostic imaging diplomates and MRI industry experts for evaluation. Following feedback, the finalised version of the grading scale was produced. The components of the image quality grading system for each anatomical structure are presented below. The grading system guidelines were also assembled into a table for distribution to observers.

#### **2.4.2.3 Deep digital flexor tendon**

##### **Subjective image quality**

- Grade 1: Textbook quality- The study is well composed with optimal definition of the deep digital flexor tendon. Images are sufficient quality to be printed in a textbook.
- Grade 2: High diagnostic quality- Clearly defined tendon signal, shape and margins. Fascicular structure may be evident. Any artefact does not limit interpretation.
- Grade 3: Satisfactory diagnostic quality- Satisfactorily defined tendon signal, shape and margins. Fascicular structure may be visible. Any artefact does not limit interpretation.
- Grade 4: Non-diagnostic- Poorly defined tendon signal, shape and/or margin. Artefact may be present which significantly inhibits interpretation.

#### **2.4.2.4 Navicular bone**

##### **Subjective image quality**

- Grade 1: Textbook quality- The study is well composed with optimal definition of the navicular bone. Images are sufficient quality to be printed in a textbook.
- Grade 2: High diagnostic quality- The bone signal, shape and margins (including cortices, spongiosa and fibrocartilage) are clearly defined. Trabecular pattern evident. Any artefact does not limit interpretation.

- Grade 3: Satisfactory diagnostic quality- The bone signal, shape and margins (including cortices, spongiosa) are satisfactorily defined. Trabecular pattern may be evident. Any artefact does not limit interpretation.
- Grade 4: Non-diagnostic- Components of the bone signal, shape and/or margins (including cortices, spongiosa and fibrocartilage) are poorly defined. Trabecular pattern not visible. Artefact may be present which significantly inhibits interpretation.

#### **2.4.2.5 Navicular bursa**

##### **Subjective image quality**

- Grade 1: Textbook quality- The study is well composed with optimal definition of the navicular bursa. Images are sufficient quality to be printed in a textbook.
- Grade 2: High diagnostic quality- Bursa location, signal, shape and margins are clearly defined. Any artefact does not limit interpretation.
- Grade 3: Satisfactory diagnostic quality- Bursa location, signal, shape and margins are satisfactorily defined. Any artefact does not limit interpretation.
- Grade 4: Non-diagnostic- Bursa location, signal, shape and/or margins are poorly defined which limits interpretation. Artefact may be present which significantly inhibits interpretation.

#### **2.4.2.6 Distal interphalangeal joint**

##### **Subjective image quality**

- Grade 1: Textbook quality- The study is well composed with optimal tissue definition. Images are sufficient quality to be printed in a textbook.
- Grade 2: High diagnostic quality- Joint signal, shape and margins are clearly defined with demarcation of the articular surfaces. Any artefact does not limit interpretation.
- Grade 3: Satisfactory diagnostic quality- Joint signal, location, shape and margins are generally well defined with demarcation of the articular surfaces at the joint margins. Any artefact does not limit interpretation.

- Grade 4: Non-diagnostic- Joint location, shape and margins are poorly defined and/or insufficient demarcation of the articular surfaces which limits interpretation. Artefact may be present which significantly inhibits interpretation.

### **Collateral ligaments of the distal interphalangeal joint**

#### **Subjective image quality**

- Grade 1: Textbook quality- The study is well composed with optimal definition of the collateral ligaments. Images are sufficient quality to be printed in a textbook.
- Grade 2: High diagnostic quality- Clearly defined ligament location, signal, shape and margins. Any artefact does not limit interpretation.
- Grade 3: Satisfactory diagnostic quality- Satisfactory definition of ligament location, signal, shape and margins. Any artefact does not limit interpretation.
- Grade 4: Non-diagnostic- Poorly defined ligament location, signal, shape and/or margins. Artefact may be present which significantly inhibits interpretation.

### **2.4.2.7 Third phalanx**

#### **Subjective image quality**

- Grade 1: Textbook quality- The study is well composed with optimal definition of the third phalanx. Images are sufficient quality to be printed in a textbook.
- Grade 2: High diagnostic quality- The bone shape, signal and margins are clearly demarcated. Any artefact does not limit interpretation.
- Grade 3: Satisfactory diagnostic quality- The bone shape, signal and margins are satisfactorily demarcated. Any artefact does not limit interpretation.
- Grade 4: Non-diagnostic- Components of the bone shape, signal and/or margins are poorly demarcated. Artefact may be present which significantly inhibits interpretation.

#### **2.4.2.8 Distal sesamoidean impar ligament**

##### **Subjective image quality**

- Grade 1: Textbook quality- The study is well composed with optimal definition of the ligament. Images are sufficient quality to be printed in a textbook.
- Grade 2: High diagnostic quality- Clearly defined ligament location, shape and margins. Any artefact does not limit interpretation. Diagnostic information may be inferred from evaluation of adjacent structures.
- Grade 3: Satisfactory diagnostic quality- Ligament location, shape and margins may be defined. Any artefact does not limit interpretation. Diagnostic information may be inferred from evaluation of adjacent structures.
- Grade 4: Non-diagnostic- Poorly defined ligament location, shape and/or margin. Artefact may be present which significantly inhibits interpretation.

#### **2.4.2.9 Pathology assessment**

All anatomic structures were assessed for pathology. This was to quantify the degree of pathology in the studies from each system, rather than to compare pathology identification ability between systems. The pathology assessment was divided into four categories with brief verbal descriptors.

- Grade 1: No pathology
- Grade 2: Mild pathology
- Grade 3: Moderate pathology
- Grade 4: Severe pathology

#### **2.4.2.10 Individual structure assessment interface**

The individual anatomic structure assessment component of the platform was formatted as a grid with selection list format for each MRI study (Figure 2-2)

	Subjective image quality for structure	Extent of pathology within this structure
Deep digital flexor tendon	Please select ▾	Please select ▾
Navicular bone	Please select ▾	Please select ▾
Navicular bursa	Please select ▾	Please select ▾
DIP joint	Please select ▾	Please select ▾
Collateral ligaments of DIP joint	Please select ▾	Please select ▾
Third phalanx	Please select ▾	Please select ▾
Distal sesamoidean impar ligament	Please select ▾	Please select ▾

**Figure 2-2 Format of the individual anatomical structure assessment interface.**

## 2.4.3 Magnetic resonance imaging systems and studies

### 2.4.3.1 Magnetic resonance imaging systems used in the study

MRI studies of the equine foot were collected from a low-field (0.3T), weight-bearing (standing, sedated) open magnet system and a low-field (0.3T), non-weight-bearing (anaesthetised) open magnet system.

The MRI studies were collected from the clinical databases of two different institutions. Low-field standing MR images were derived from an MRI system manufactured by Hallmarq Veterinary Imaging Ltd. The manufacturer was contacted in the first instance to request the use of images and Liphook Equine Hospital was then selected as the site for image collection. The low-field under general anaesthesia MRI studies were derived from the Esaote O-scan equine MRI system at the Weipers Centre Equine Hospital, University of Glasgow, with which the author is familiar.

### 2.4.3.2 Selection of magnetic resonance imaging studies

To maximise the clinical relevance of the study the aim was to retrospectively utilise MR images that were collected in a clinical context for the investigation

of foot pathology (rather than the prospective, experimental nature of previous literature). Image collection was performed using databases of clinical MRI studies at each of the institutions. The timescale over which MR studies could be retrieved from the database was standardised to avoid significant discrepancies in the evidence base which could be used to guide MRI protocols at different time periods. The timescale for MR study acquisition was from June 2015 to August 2018. Five complete foot MRI studies were included in the image quality assessment for each of the MRI systems.

To avoid selection bias, there was no system for matching of case signalment, number of MR sequences or pathology between MRI systems. The study design ensured that differences in pathology between studies would be quantified during the observer image assessment. The MRI sequences were not matched between systems as protocols can vary significantly between MRI systems. Sequence selection is often tailored to provide maximum image quality in a particular system and scenario. The minimum sampling unit for the purposes of case selection was the individual foot. Several inclusion criteria ensured that MR studies were derived in a comparable manner. The shortlisted databases of studies included those with all the sequences that would typically be performed as a routine foot MRI study at that institution. This included MRI studies that were repetition of previously imaged patients and studies of contralateral (i.e. non-lame) limbs.

All studies were collected in a randomised manner using a random number generator (Random.org., Random Integer Generator, <https://www.random.org/integers/>). The format of the databases available (for example, using an institution Picture Archiving and Communication System (PACS), billing system or MRI system local computer database) dictated the exact manner of selection. The MRI studies from the low-field standing system were selected from a database which outlined the sequence count and exam type (standard, contralateral and rescan). Using previously derived statistics Hallmarq established that a routine equine foot MRI study using the low-field standing system had a mean of 10.44 sequences. Therefore, studies were included that contained greater than 11 sequences (range reported up to a maximum of 17). Single foot MRI studies were taken as the sampling unit. With the studies in an



arbitrary order with an assigned number, the random integer generator was used to randomly select six studies. Six studies were collected and forwarded for potential inclusion in the study, so that a single study could be excluded should it not meet all the inclusion criteria when reviewed by the author, without necessitating repetition of remote study collection. Prior to submission to the author, all Digital Imaging and Communications in Medicine (DICOM) files were anonymised by Hallmarq Veterinary Imaging Ltd. All studies were deemed to meet the inclusion criteria, therefore following arbitrary ordering of studies, the random integer generator was used to select the final five studies for inclusion in the study in a similar manner.

The MRI studies from the low-field under general anaesthesia system were identified from the MRI system local computer database. The case details were compiled into an external database (Microsoft Office Excel 2016, Microsoft Corporation). Single foot MRI studies were taken as the sampling unit. All foot studies were assigned a unique arbitrary identification number. The random integer generator was used to randomly select six studies for inclusion. The studies were then reviewed by the author to ensure they contained all the sequences that were routinely performed during MRI of the equine foot at the Weipers Centre Equine Hospital, University of Glasgow. This consists of at least seven primary sequences with a variable number of multiplanar reconstructions based on an isotropic T1 weighted sequence. A single study of the selected group did not contain all routine sequences and so was excluded. The remaining five studies were included. The DICOM files were anonymised using a DICOM modification tool (DICOM Anonymizer <https://dicomanonymizer.com/index.html>).

The initial DICOM files either received by transfer from Hallmarq or directly from the University of Glasgow veterinary PACS system were copied and a duplicate archived. All folders containing the DICOM files were reviewed. A database for studies and their related information was established. The organised subfolders containing the DICOM files were counted and their names recorded for each study. The pre-existing case number was also recorded for each study. All studies were reviewed to document the number of sequences presented in two DICOM viewers (ClearCanvas Workstation, Version 2.0 SP1, ClearCanvas Inc. and

Horos, Version 2.2.0, Horos Project). This was used as an opportunity to review the studies to ensure that all sequences were present and displayed correctly.

### **2.4.3.3 Presentation of the magnetic resonance imaging studies**

A random sequence generator (Random.org., Random Sequence Generator, <https://www.random.org/sequences/>) was used to generate the order of presentation of studies for the image assessment platform.

The DICOM files were systematically processed to prepare them for distribution. A modified version of the DICOM anonymisation tool (Modified DICOM Anonymizer Chest Radiologists Collaborative <http://chestradiologists.org/directory/dicom-anonymizer-osirix/>) was used to uniformly anonymise the studies. In addition, this allowed the studies to be renamed with the appropriate study order number in the 'Patient ID' metadata attribute. Metadata attributes that could provide useful information (relating to acquisition pulse sequence parameters) such as the echo time and repetition time were unchanged.

Given that the DICOM files to be distributed were derived from clinical databases the 'Study Instance UID' was updated to avoid duplication and streamline handling of the DICOM files by DICOM viewers. A unique UID chain was acquired (Medical Connections Ltd: [www.medicalconnections.co.uk/FreeUID/](http://www.medicalconnections.co.uk/FreeUID/)). This chain was used as a unique root for the 'Study Instance UID' metadata attribute in the format: 1.2.826.0.1.3680043.10.167.1.casenumber. The anonymised and labelled DICOM files for each study were saved in folders.

## **2.4.4 Observers**

### **2.4.4.1 Observer recruitment**

The online image assessment platform was distributed to 10 observers with significant experience in equine MRI. Criteria for inclusion of observers included diploma or associate status in at least one relevant field and at least 5 years' experience interpreting equine MR images. Relevant fields for diploma or associate status included:

- Diagnostic Imaging: Diplomate or Associate Member of the European College of Veterinary Diagnostic Imaging or Diplomate of the American College of Veterinary Radiology. The Associate Member status is awarded by the College to individuals who have made significant contributions to veterinary diagnostic imaging.
- Surgery: Diplomate of the European College of Veterinary Surgeons or Diplomate of the American College of Veterinary Surgeons.
- Sports Medicine and Rehabilitation: Diplomate of the European College of Veterinary Sports Medicine and Rehabilitation or Diplomate of the American College of Veterinary Sports Medicine and Rehabilitation.

Potential observers were contacted by email with a short summary of the project, a copy of the grading systems to be used in the project, the requirements for participation and an outline of the timescale for participation. Observers were recruited through direct individual contact and with the assistance of an equine teleradiology service (VetCT: <https://www.vet-ct.com/gb/>).

#### **2.4.4.2 Observer payment and acknowledgement**

At the recruitment stage it was outlined to observers that a stipend would be paid to each observer upon completion of the image assessment platform, within the required time constraints. It was also outlined that observer participation would be recognised as an acknowledgement in any material published following the project.

#### **2.4.4.3 Observer consent**

On the first page of the image assessment platform, a statement on observer participant consent was presented. Observers were advised that by continuing they consented to collection of their responses, use of their responses in data analysis, contribution of these data to publications derived from the study and the long-term storage of data, which may be utilised for future research. Contact details for the author were provided for any observers requiring further information.

#### **2.4.4.4 Observer orientation**

All observers were provided with a document describing the project aims and the format of their participation. At the outset of the image assessment platform a splash page gave a similar short summary and provided a link to additional guidance in the Observer Handbook, which also contained the grading scales to be used in the assessment.

#### **2.4.4.5 Observer profile**

An observer profile section was included at the beginning of the online survey platform. This documented the observers' current speciality status (i.e. Diplomate, Associate Member or Resident status in any of the relevant fields outlined above) and any other relevant postgraduate qualifications. Observers also reported the duration of their experience interpreting equine MR images by selecting the most appropriate category (<1 year experience, 1-5 years' experience, 5-10 years' experience, >10 years' experience). Observers were also asked to report the frequency with which they interpret images from relevant categories of equine MR image acquisition systems including low-field (0.3T) standing, low-field (0.3T) under general anaesthesia and high-field (>1.0T) under general anaesthesia. Categories for frequency of interpretation included regularly (daily), frequently (weekly), occasionally (monthly), and rarely or never. Observers were also asked to report which DICOM viewing software they would use to view the images in the project. Options proposed included Horos, OsiriX, Asteris, ClearCanvas, eFilm, Visbion and other. Multiple options of DICOM viewing software could be selected should this be appropriate.

#### **2.4.5 Image assessment platform and magnetic resonance imaging study distribution**

The online image assessment platform and organised DICOM files of the MRI studies were distributed to each observer by email with a unique link. The unique Uniform Resource Locator (URL) allowed individual responses to be tracked. A short description of how to access the DICOM files for use in combination with the image assessment platform was included in the Observer Handbook. Observers were advised that they were free to manipulate the images as they would in a clinical context using their DICOM viewer of choice. Observers

were able to save their progress during the image assessment platform and then return to complete the assessment at their own convenience.

## **2.4.6 Data analysis**

### **2.4.6.1 Initial data handling and general data analysis**

Raw data was exported from the image assessment platform and assembled into a database (Microsoft Office Excel 2016, Microsoft Corporation). Free text comments were removed for separate analysis. The database contents were coded for further assessment and were organised into frequency tables. Percentages were calculated for each category within frequency tables. With all anatomical structures combined the proportion image quality and pathology assessment assigned to each grade was also calculated to demonstrate the spectrum of grading scale use by observers. Initial graphical exploration included plotting of line graphs to assess individual observer grading and to assess grading trends for each MRI study, for image quality and pathology assessment.

Descriptive analysis of ordinal data included calculation of the median, mode and range grouped by assessment category. Data related to image quality assessment of whole studies and image quality assessment of individual anatomical structures was assembled into contingency tables with study gradings grouped as diagnostic (image quality grades 1-3) and non-diagnostic (image quality grade 4). Additional contingency tables were assembled with study gradings grouped as equal or greater than high diagnostic quality (image quality grades 1 and 2) and satisfactory diagnostic quality or less (image quality grades 3 and 4). Comparable contingency tables were made for pathology grading.

The pulse sequence data for MRI studies were assembled into a separate database. Pulse sequence data recorded for each MRI study included the pulse sequence, orientation, echo time, repetition time, field of view, slice thickness, interslice spacing, number of slices, flip angle and inversion time (where relevant). The pulse sequence data for each MRI study was exported as a table for reference. Descriptive statistical analysis was performed including the median and range for number of sequences (including multiplanar reconstructions where relevant) per study for each system.

#### **2.4.6.2 Data analysis for comparison of image quality between low-field standing and low-field under general anaesthesia studies**

Further statistical analysis was performed using additional statistical software Minitab (Minitab 18.1, Minitab Ltd.) and SPSS (IBM SPSS Statistics, IBM United Kingdom Ltd.). Throughout the analysis  $P < .05$  were considered significant.

Using image quality assessment contingency tables, confidence intervals (95%) were calculated for the proportions of diagnostic studies, and studies greater than or equal to high diagnostic quality using an exact method.

Statistical analysis was based on comparison to assess the effect of anaesthesia: low-field standing compared to low-field under general anaesthesia. The proportion of diagnostic and non-diagnostic gradings was compared (for entire studies and each individual anatomical structure) using a Pearson chi-square test or Fisher's exact test if one or more expected cell values  $\leq 5$ . The ranked gradings for image quality (for whole studies and each individual anatomical structure) and pathology (for each individual anatomic structure) were compared using a Mann-Whitney test. Graphical analysis included plotting of comparative clustered bar charts demonstrating the count of observer gradings for each assessment category (whole study or individual anatomical structure) with clustering by grade. Graphs were assembled for image quality assessment and pathology assessment.

#### **2.4.6.3 Free text responses**

The free text comments were collated for each study. The comments were grouped as either containing information regarding image quality or pathology assessment (or both where relevant). Comments relating to image quality were sub-categorised by reason reported for reduced image quality including:

- alignment or position
- additional sequence desired
- repeat sequence desired
- suboptimal short tau inversion recovery (STIR) fat suppression (for example incomplete fat suppression)
- sequence parameter (other)

- motion artefact
- magic angle effect
- artefact (other)
- other

Where a comment referred to multiple categories, this was documented as a count in each of the relevant categories. Clustered bar charts were plotted demonstrating the count of free text comments, with clustering by the category of reason for reduced image quality.

#### **2.4.6.4 Inter-observer agreement for image quality assessment**

Inter-observer agreement for image quality grading was evaluated with Fleiss' kappa and Kendall's coefficient of concordance. Agreement analysis was performed for entire studies and individual anatomical structures, including all studies from low-field standing and low-field under general anaesthesia with gradings by all observers. Interpretation of kappa values was based on previously suggested generic descriptors (Altman, 1991), which are presented in Table 2-3. Established descriptors for interpretation of Kendall's coefficient of concordance are not available. Values range from 0 to 1 and higher values were interpreted as demonstrating greater concordance (Gwet, 2014).

**Table 2-3 Interpretation of kappa statistic values based on Altman (1991).**

<b>Kappa statistic</b>	<b>Interpretation</b>
<0.20	poor agreement
0.21-0.40	fair agreement
0.41-0.60	moderate agreement
0.61-0.80	good agreement
0.81-1.00	very good agreement

## **2.5 Results**

### **2.5.1 General results**

#### **2.5.1.1 Observer characteristics**

All observers completed the observer profile. When observers reported their relevant qualifications in MRI associated fields there were: six diagnostic imaging Diplomates, one diagnostic imaging Associate Member, one diagnostic imaging Associate Member and surgical Diplomate, one surgical and sports medicine and rehabilitation Diplomate and diagnostic imaging Associate Member, and one surgical and sports medicine and rehabilitation Diplomate. Two observers also held postgraduate certificate qualifications (equivalent to the Royal College of Veterinary Surgeons Certificate in Advanced Veterinary Practice).

Four observers reported to have 5-10 years' experience interpreting equine MR images with the remaining 6/10 having >10 years' experience. Observers had greater experience interpreting images from low-field standing systems (interpreting from this system type regularly: 6/10, frequently: 2/10, occasionally: 0/10 and rarely/never: 2/10) compared to low-field under general anaesthesia systems (interpreting from this system type regularly: 1/10, frequently: 0/10, occasionally: 5/10 and rarely/never: 4/10).

#### **2.5.1.2 Magnetic resonance imaging study acquisition features**

All the MRI studies contained T1 weighted, T2 weighted, proton density weighted, and fat suppressed (STIR) sequences. All MRI studies contained sequences acquired in at least three planes. Low-field standing studies had a median of 10 sequences (range 9-11) and low-field under general anaesthesia had a median of 12 sequences (range 11-14). These counts did not include initial orientation sequences (variably referred to as localiser or pilot sequences by manufacturers) but included multiplanar reconstructions that formed part of the routine imaging protocol of the acquisition institution (routinely performed for the low-field under general anaesthesia studies). The pulse sequence parameters for all MRI studies, grouped by MRI acquisition system are presented in Appendix 2.



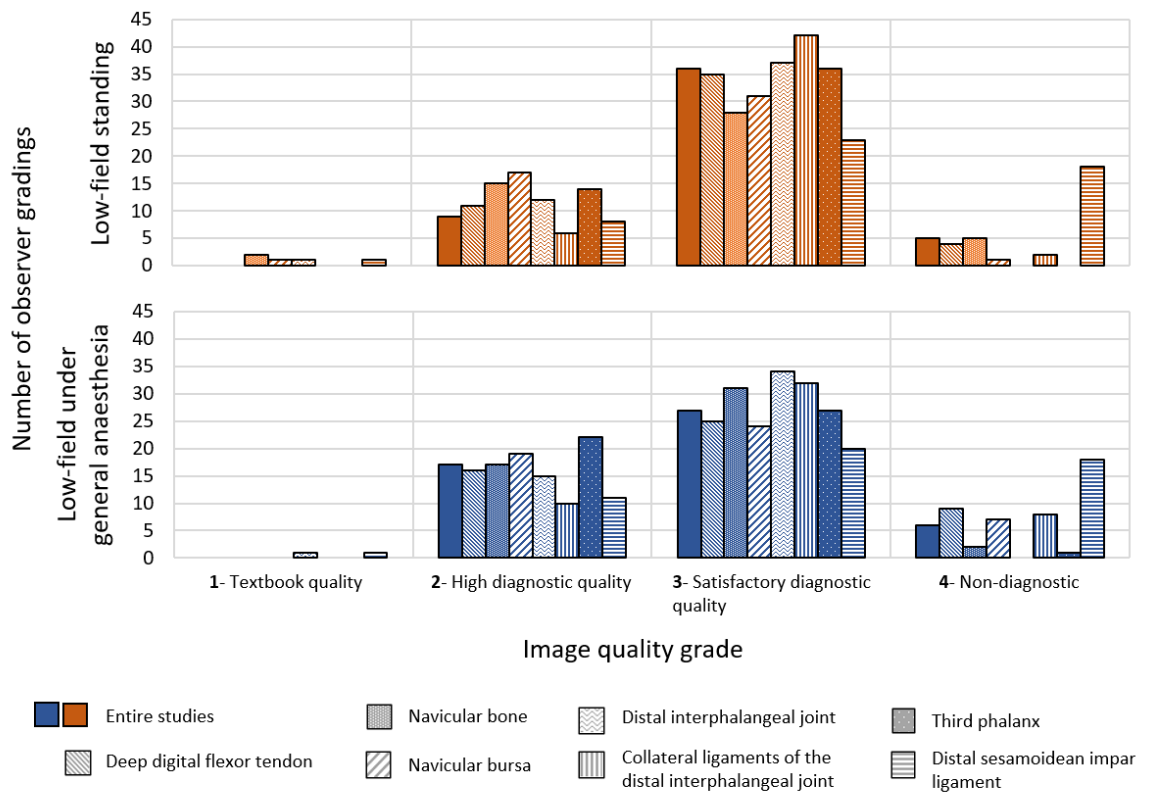
## 2.5.2 Comparison of image quality

### 2.5.2.1 Key statistics for image quality assessment

All observers completed the image quality assessment of all the low-field standing and low-field under general anaesthesia studies.

When observers assessed the image quality of entire studies, 90% (95% CI 78%, 97%) were graded as diagnostic for low-field standing studies and 88% (95% CI 76%, 95%) were graded as diagnostic for the low-field under general anaesthesia studies. Grouping entire study image quality gradings as high-diagnostic quality or better (those graded 1 or 2) accounted for 18% (95% CI 9%, 31%) of gradings for low-field standing studies and 34% (95% CI 21%, 49%) for low-field under general anaesthesia studies. The distribution of image quality gradings for entire MRI studies and individual anatomical structures is displayed in Figure 2-3.

For entire MRI studies and individual anatomical structures, the median grade for image quality was grade 3 for both low-field standing and low-field under general anaesthesia studies (Table 2-4).



**Figure 2-3** Bar chart displaying the image quality gradings for low-field standing and low-field under general anaesthesia studies.

**Table 2-4 Median and range for image quality grading for low-field standing and low-field under general anaesthesia studies.**

Abbreviations: DIP- distal interphalangeal.

Assessment category	Low-field standing		Low-field under general anaesthesia	
	Median grade	Range	Median grade	Range
Entire study	3	2-4	3	2-4
Deep digital flexor tendon	3	2-4	3	2-4
Navicular bone	3	1-4	3	2-4
Navicular bursa	3	1-4	3	2-4
DIP joint	3	1-3	3	1-3
Collateral ligaments of the DIP joint	3	2-4	3	2-4
Third phalanx	3	2-3	3	2-4
Distal sesamoidean impar ligament	3	1-4	3	1-4

### 2.5.2.2 Comparison of the proportion of diagnostic studies

There were no significant differences in the proportion of diagnostic versus non-diagnostic studies between low-field standing and low-field under general anaesthesia for entire studies. The P values generated from chi-square and Fisher's exact tests comparing the proportion of diagnostic gradings are presented in Table 2-5.

**Table 2-5 Comparison of proportion of diagnostic image quality gradings for low-field standing and low-field under general anaesthesia studies.**

Abbreviations: DIP- distal interphalangeal.

<sup>a</sup> Indicates use of Fisher's exact test. Other comparisons used chi-square tests.

<sup>b</sup> Analysis not performed due to 0 value(s) in the contingency table.

Assessment category	Comparison of proportion of diagnostic gradings (P value)
Entire study	0.7
Deep digital flexor tendon	0.1
Navicular bone	0.4 <sup>a</sup>
Navicular bursa	0.06 <sup>a</sup>
DIP joint	Not performed <sup>b</sup>
Collateral ligaments of the DIP joint	0.09 <sup>a</sup>
Third phalanx	>0.9 <sup>a</sup>
Distal sesamoidean impar ligament	>0.9

### 2.5.2.3 Comparisons of the ranked gradings for image quality

When comparing the ranked image quality gradings for entire studies and individual anatomical structures with a Mann-Whitney test, there were no significant differences between low-field standing and low-field under general anaesthesia studies (Table 2-6).

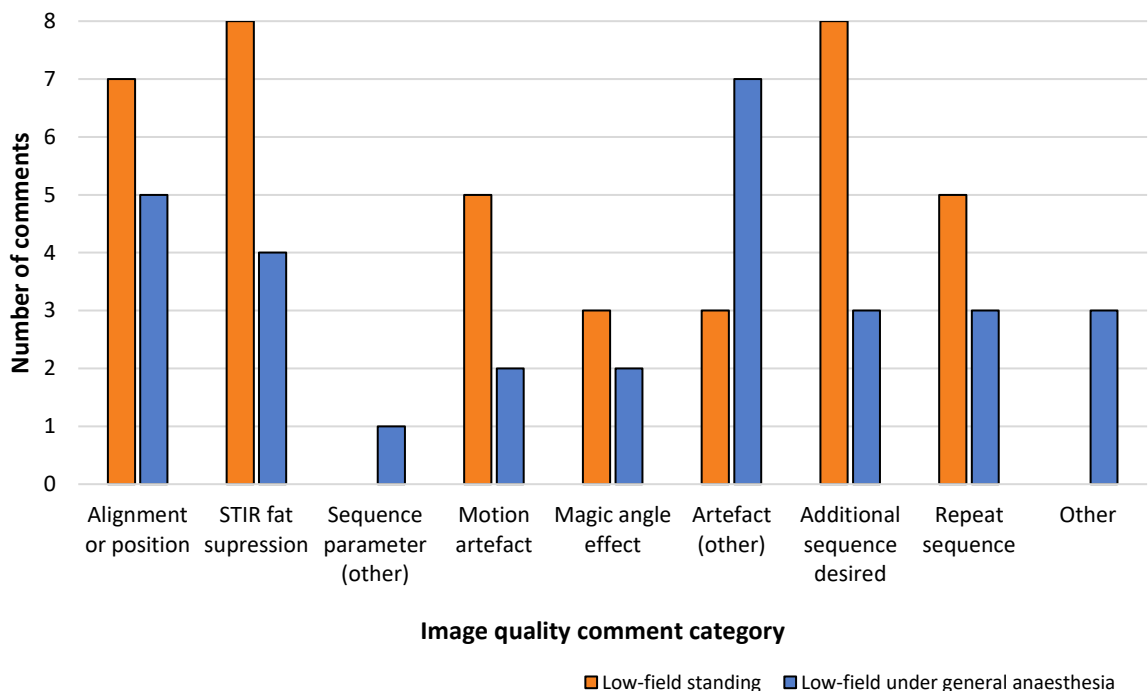
**Table 2-6 Comparison of ranked gradings for image quality for low-field standing and low-field under general anaesthesia studies.**

Abbreviations: DIP- distal interphalangeal.

Assessment category	Comparison of ranked gradings for image quality (P value)
Entire study	0.2
Deep digital flexor tendon	0.9
Navicular bone	0.8
Navicular bursa	0.5
DIP joint	0.5
Collateral ligaments of the DIP joint	0.8
Third phalanx	0.1
Distal sesamoidean impar ligament	0.7

#### 2.5.2.4 Free text comments regarding image quality

There were 19 free text comments relating to image quality for the low-field standing studies and 23 free text comments relating to image quality for the low-field under general anaesthesia studies. The distribution of image quality comments across the categories for reduced image quality differed between systems and are displayed in Figure 2-4.

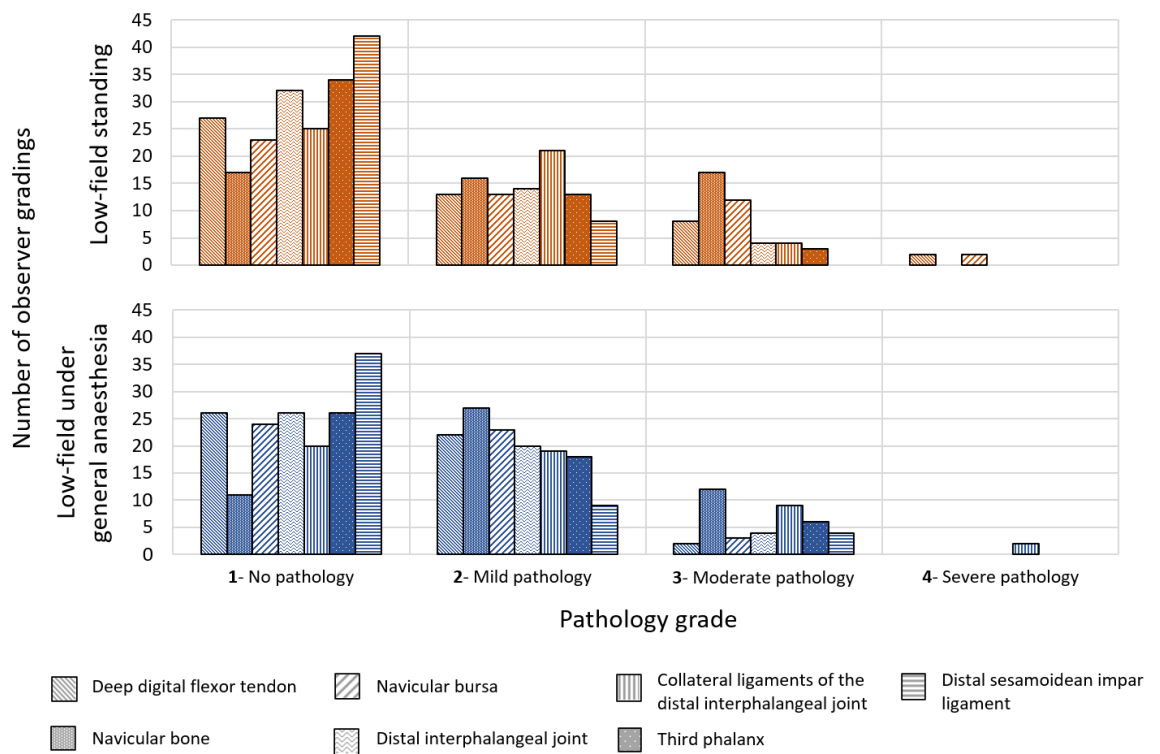


**Figure 2-4 Bar chart displaying the categorisation of image quality comments for low-field standing and low-field under general anaesthesia studies.**

### 2.5.3 Pathology considerations for image quality assessment

All observers completed the pathology assessment of all the low-field standing and low-field under general anaesthesia studies. The distribution of pathology gradings is demonstrated in Figure 2-5.

When comparing the ranked pathology gradings for individual anatomical structures with a Mann-Whitney test, there were no significant differences between low-field standing and low-field under general anaesthesia studies (Table 2-7).



**Figure 2-5 Bar chart displaying pathology gradings for low-field standing and low-field under general anaesthesia studies.**

**Table 2-7 Median pathology grades and rank comparison for low-field standing and low-field under general anaesthesia studies.**

Abbreviations: DIP- distal interphalangeal.

Assessment category	Median pathology grade		Comparison of ranked gradings for pathology (P value)
	Low-field standing	Low-field under general anaesthesia	
Deep digital flexor tendon	1	1	0.6
Navicular bone	2	2	0.9
Navicular bursa	2	2	0.2
DIP joint	1	1	0.3
Collateral ligaments of the DIP joint	1.5	2	0.1
Third phalanx	1	1	0.09
Distal sesamoidean impar ligament	1	1	0.2

### 2.5.4 Inter-observer agreement for image quality

The absolute inter-observer agreement determined by Fleiss' kappa for image quality grading of entire studies was poor. Similarly, the absolute inter-observer agreement was poor for all individual anatomic structures. Assessment of inter-observer agreement accounting for the ranking of gradings, as determined by Kendall's coefficient of concordance was also low. The results of inter-observer agreement analysis are presented in Table 2-8.

**Table 2-8 Output of inter-observer agreement analysis for magnetic resonance image quality grading for all low-field standing and low-field under general anaesthesia studies.**

Abbreviations: DIP- distal interphalangeal.

Assessment category	Fleiss' kappa		Kendall's coefficient of concordance
	Kappa statistic	Standard error	
Entire studies	0.04	0.04	0.18
Deep digital flexor tendon	0.03	0.04	0.08
Navicular bone	-0.01	0.04	0.20
Navicular bursa	-0.04	0.04	0.07
DIP joint	-0.07	0.04	0.08
Collateral ligaments of the DIP joint	0.02	0.04	0.19
Third phalanx	-0.05	0.05	0.09
Distal sesamoidean impar ligament	-0.03	0.03	0.22

## 2.6 Discussion

Most clinical low-field MRI studies of the equine foot were deemed to be of diagnostic quality, irrespective of whether they were performed with the patient under standing sedation or general anaesthesia. Perceived image quality was broadly comparable between the systems. Approximately 10% of the low-field imaging studies were reported by observers as non-diagnostic. The median image quality was grade 3 (diagnostic quality) for all anatomic structures for both system types, with relatively few assessments assigned grade 1 image quality (textbook quality).

When observers graded entire MRI studies, there were no significant differences in image quality between low-field standing and low-field under general anaesthesia systems (in the proportion of diagnostic studies or in the ranked gradings). Grading entire studies is a relatively broad assessment of image quality. However, comparison of image quality for assessment of the seven key anatomical structures of the foot also demonstrated that there were no significant differences in image quality between low-field imaging in the standing and anaesthetised equine patient. This finding supports the assumption that systems with similar field-strengths are thought to have broadly comparable image quality (Werpy, 2010). There was no evidence to support the hypothesis that there would be a difference in image quality between images acquired from a low-field system with the patient standing compared to those acquired from a low-field system with the patient under general anaesthesia.

Patient motion is commonly reported as a reason for reduced image quality during equine MRI and is deemed to be particularly problematic for imaging the standing patient (Porter and Werpy, 2014). Patient motion was reported as a reason for reduced image quality in observer free text comments for both low-field standing (5 comments) and low-field under general anaesthesia (2 comments). However, motion accounted for a relatively small proportion of the free text comments related to image quality, indicating that motion was not a major limitation for either system. This finding indicates that the influence of patient motion on image quality is practically limited during low-field MRI of the equine foot in a clinical context. Limiting the influence of patient movement may be a result of patient selection, suitable patient management, use of

appropriate motion-insensitive sequences, motion-correction techniques and careful study planning (McKnight *et al.*, 2004; Werpy, 2010). This study is primarily outcome based (i.e. focused on image quality assessment of the MRI studies). Therefore, it does not account for the possibility that multiple repeat sequences (due to patient motion) may be required to obtain images deemed satisfactory at the time of original acquisition, especially if rejected sequences were not included in the final study. Whilst this may not alter the final image quality of the study, repetition of sequences is an important factor to consider in the clinical setting, where time constraints make repetition undesirable. The position of the equine foot against the ground surface during imaging of the standing patient is thought to make this region less susceptible to sway motion of the sedated patient compared to more proximal regions where pendulous sway motion can occur readily (McKnight *et al.*, 2004). Therefore, the conclusions of the current study are restricted to the equine foot, as the effect of patient motion may have a greater influence on image quality (particularly in the standing sedated patient) in other anatomical locations.

Imaging the patient under standing sedation or general anaesthesia also has important influences on limb positioning (Werpy, 2010). In human medical imaging the influence of patient positioning and manipulation has explored variations including weight-bearing, stressed and real-time MRI (Shapiro and Gold, 2012). For example, upright MRI of the human knee provides a physiologic representation of the joint which can be useful for the evaluation of structures such as the load-bearing articular cartilage and the menisci (Gold *et al.*, 2004; Shapiro and Gold, 2012; Stehling *et al.*, 2012; Barile *et al.*, 2013; Bruno *et al.*, 2018). The ability of patients to replicate painful actions during MRI is also considered valuable for detection of pathology (Bruno *et al.*, 2018).

Given the relative size of equine patients, the influence of weight-bearing may not be directly comparable to that of human patients. In the equine literature, patient positioning is primarily binary—effectively either standing or recumbent. During low-field standing imaging the limb is in a weight-bearing position. For low-field imaging under general anaesthesia, the limb is in a non-weight-bearing position. Assessment of articular cartilage during weight-bearing has been a particular focus in the equine literature. A cadaveric study indicated that low-

field MRI of the non-weightbearing equine foot could be useful for the detection of full and partial-thickness articular cartilage lesions in the distal interphalangeal joint (Olive, 2010). However, it was highlighted that this conclusion may not directly extrapolate into the standing patient where the adjacent articular cartilage margins may be compressed into contact (Olive, 2010). A subsequent cadaveric study supported this suspicion, by demonstrating that definition of distal interphalangeal joint cartilage margins is more challenging in a weight-bearing position (Evrard *et al.*, 2019). Similarly, a study assessing the identification of naturally occurring distal interphalangeal joint articular cartilage damage indicated that low-field MRI and MR arthrography were not sensitive methods for the detection of pathology (van Zadelhoff *et al.*, 2020). Importantly, this study also utilised cadaver limbs in a non-weight-bearing position and the authors outlined that sensitivity may be further reduced during weight-bearing imaging in a live patient (van Zadelhoff *et al.*, 2020). Given the findings of the previous literature, it is interesting that in the current study there were no significant differences in observer perceived image quality between low-field standing and low-field under general anaesthesia imaging for assessment of the distal interphalangeal joint. This may reflect the differences in assessment of image quality (as in the current study) compared to identification of articular pathology or the more focused measurement of cartilage thickness and definition of cartilage margins. The latter methods are likely to be more sensitive in identifying small differences in cartilage assessment ability between positioning types. Nonetheless, the apparent lack of difference between weight-bearing and non-weight-bearing low-field MR image quality for the distal interphalangeal in the current study may indicate that small differences to cartilage delineation are overcome by experienced observers in a clinical setting. At present, the literature indicates that if a distal interphalangeal joint cartilage lesion is suspected, then MRI under general anaesthesia (non-weight-bearing) may be preferable and that even subtle alterations in cartilage or subchondral bone signal should be considered suspicious of pathology (Olive, 2010; Barrett *et al.*, 2017; Evrard *et al.*, 2019; van Zadelhoff *et al.*, 2020).

The skill of the operator positioning the limb within the magnetic field is also important (Werpy, 2010). As with other imaging modalities, poor positioning can



make image acquisition and subsequent interpretation challenging. Unilateral or uneven limb positioning can influence the appearance of the equine foot during radiography and similar challenges are present during MRI (Contino, Barrett and Werpy, 2014; Evrard *et al.*, 2019; Joostens, Evrard and Busoni, 2019). Standing MRI is more likely to be negatively influenced by static foot balance (lateromedial balance in particular) of the patient compared to non-weight-bearing imaging under general anaesthesia, which negates this factor (Evrard *et al.*, 2019). The design of the standing low-field MRI system can require some patients to stand with a relatively base wide stance (Werpy, 2010). Uneven limb loading can cause anatomic structures to appear asymmetric on MR images and make interpretation more challenging. Unlike other modalities, the magic angle effect also needs to be considered during positioning for MRI (Murray and Werpy, 2010). Operators typically aim to reduce the occurrence of the magic angle effect or to at least make its occurrence as predictable as possible (Bydder *et al.*, 2007). Differences in the orientation of the static magnetic field ( $B_0$ ) between systems can change the anatomic structures that are susceptible to the magic angle effect but the collateral ligaments of the distal interphalangeal joint and the distal aspect of the deep digital flexor tendon are common locations for this effect to occur (Spriet, Mai and McKnight, 2007; Smith, Dyson and Murray, 2008). The low-field system designed for imaging standing equine patients has a static magnetic field orientated perpendicular to the sagittal plane of the limb (i.e. perpendicular to the ground surface) (Smith, Dyson and Murray, 2008; Spriet and Zwingerberger, 2009). The orientation of the static magnetic field can vary between low-field imaging systems designed for use with the patient under general anaesthesia (Spriet, Mai and McKnight, 2007; Spriet and McKnight, 2009). In the current study the magic angle effect was only noted by observers in a small proportion of studies from both systems, with 5 comments from the total of 100 MRI study interpretations. This indicates that techniques utilised in a clinical setting, including meticulous limb positioning and inclusion of sequences with a long echo time (that are less susceptible to this effect) can successfully reduce deleterious occurrences of the magic angle effect (Bydder *et al.*, 2007; Murray and Werpy, 2010). In addition, observers should have a thorough knowledge of the specific MRI systems which they interpret images from, to ensure they can accurately identify the magic angle effect and avoid mistaking this for pathology (Spriet, Mai and McKnight, 2007).

The observer profile demonstrated that observer groups were more experienced interpreting low-field standing images compared to those acquired under general anaesthesia. This is unsurprising given the global popularity of the low-field standing system for clinical imaging in equine practice (Hallmarq Veterinary Imaging, 2019). The use of a defined grading system for image quality assessment provided a framework for decision making, but this inherently relied on the experience of the observers. Familiarity with images acquired from low-field standing systems could have influenced the observers' perception and expectations of a high-quality image. It is plausible to expect that this familiarity could increase the confidence of observers in assessing image quality and pathology on images from the low-field standing system, which could have altered any apparent difference in perceived image quality between systems. This effect is challenging to avoid and is essentially inherent in the general observer population given the predominance of low-field standing systems used in equine practice. Pathology was comparable between studies from the two systems, so it is unlikely that this will have influenced image quality assessment.

Inter-observer agreement analysis for image quality assessment demonstrated poor absolute agreement (as indicated by Fleiss' kappa) and low agreement accounting for the order of grading (as indicated by the Kendall's coefficient of concordance). The distribution of image quality gradings assigned in the current study is primarily within grades 2 and 3 (as demonstrated in Figure 2-1), with very few studies assigned grade 1 image quality. Agreement analysis statistics like Fleiss' kappa can be paradoxically small when counts in some assessed categories are low, without necessarily representing poor agreement (Wongpakaran *et al.*, 2013; Bland, 2015). Evaluation of the inter-observer agreement for image quality grading including assessments from this study and those from Chapter 3 (The influence of field strength on perceived image quality of clinical magnetic resonance imaging of the equine foot) is presented in Appendix 3. This dataset has a broader distribution of assigned image quality gradings and is deemed to be a better representation of the agreement for image quality assessment. Whilst absolute observer agreement (indicated by Fleiss' kappa) was still poor to fair, agreement accounting for the relative order of image quality grading (indicated by Kendall's coefficient of concordance) was moderate to high. Given the subjectivity of image quality assessment and the

overall complexity of evaluating MRI studies, Kendall's coefficient of concordance is the most useful indicator of agreement. Therefore, in the context of the current study, inter-observer agreement was deemed to be sufficient for image quality assessment.

In conclusion, perceived image quality of the equine foot appears to be broadly comparable for studies acquired from a low-field system with the patient under standing sedation or under general anaesthesia. Most studies were deemed to be of diagnostic quality regardless of the imaging system, though reasons for reduced image quality may differ between systems and the system operator has an important role in minimising these. Other factors beyond image quality also need to be considered when deciding on the type of system used for MR image acquisition in clinical equine practice.

## Chapter 3 The influence of field strength on perceived image quality of clinical magnetic resonance imaging of the equine foot

### 3.1 Introduction

Pathology of the foot is commonly implicated as an origin of lameness in equine clinical orthopaedics (Ross, 2011). Magnetic resonance imaging (MRI) has significant advantages as a diagnostic imaging modality for this region due its ability to allow evaluation of bony and soft tissue structures within the hoof capsule in a multiplanar manner (Vallance *et al.*, 2012a, 2012b; Barrett *et al.*, 2017). As a result, MRI has become a key modality for imaging of this region and the foot is the most commonly magnetic resonance (MR) imaged region in equine practice (Barrett *et al.*, 2017). However, MR image acquisition has inherent intricacies which much be addressed to optimise image quality and diagnostic value in clinical practice (Werpy, 2007).

The strength of the static magnetic field of the MRI system (i.e. the field strength) has a proportional relationship with image quality (Werpy, 2007; Bolas, 2010). This arises due to the generally linear relationship between field strength and the signal-to-noise ratio of the resultant images (McRobbie *et al.*, 2017g). The signal-to-noise ratio is intimately associated with other acquisition parameters such as resolution, slice dimensions and imaging time (Bolas, 2010). Therefore, increases in signal-to-noise ratio can be translated into other beneficial factors (McRobbie *et al.*, 2017g). These considerations have long been recognised in human medicine and have driven a gradual trend of increasing magnetic field strength in some disciplines (Kladny *et al.*, 1995; Wong *et al.*, 2009; Ladd *et al.*, 2018; Sarracanie and Salameh, 2020).

In the context of equine imaging, a simple definition of a low-field magnet would be a system with a field strength <1 tesla (T) with high field magnets having a field strength >1T. These definitions are arbitrary, differ between user groups and tend to change over time. Most low-field magnets utilised in equine practice have a field strength of approximately 0.3T (Bolas, 2010). Indeed, by far the most commonly used system in equine practice globally is a bespoke low-field 0.3T system designed specifically for imaging of the equine distal limb

(Bladon, 2014; Hallmarq Veterinary Imaging, 2019). Other low-field systems are also marketed for use in equine practice designed for use in anaesthetised patients (Bolas, 2010; Esaote, 2021). High-field systems are less common in equine clinical practice but typically utilise a 1.5T or 3.0T magnet which also require the patient to be under general anaesthesia (Schramme and Segard-Weisse, 2020).

Increasing field strength does introduce other considerations including altered patient access, magnetic field safety, system location and increases system cost (Bolas, 2010). These issues are exacerbated in the equine patient, compared to the human. The desire to avoid general anaesthesia and its associated risks has been a strong driver for the success of the low-field standing system in equine practice (Bolas, 2010).

Previous studies comparing image quality of low-field and high-field MR images have demonstrated the diagnostic value of low-field imaging for the equine foot but indicated that high-field imaging could offer increased anatomical detail and characterisation of pathology (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). However, literature does not accurately reflect the clinical scenario due to the use of cadaver material, variable pathology inclusion, limb positioning different to that used in live horses and the use of comparable sequence parameters between systems, as opposed to those optimised for a particular system (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). Importantly, previous investigations have compared the low-field standing system with high-field systems, which introduces confounding factors that may also influence perceived image quality, beyond the influence of field strength alone (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). Therefore, there is limited literature to document the influence of field strength on image quality for MRI of the foot in live equine patients in a clinical context.

### **3.2 Aims**

The aim of the study was to assess the influence of field strength on image quality for MRI studies of the equine foot acquired in a clinical context. This was performed by comparing perceived image quality between MR images from low-field and high-field systems with the patient under general anaesthesia. It was

hypothesised that there would be a difference in image quality between studies acquired from low-field and high-field systems with the patient under general anaesthesia.

### **3.3 Study design**

The study had an experimental study design with prospective observer assessment of MRI studies derived from the clinical databases of two different MRI systems.

### **3.4 Methods**

The methods used in this study were broadly comparable to those reported in Chapter 2. A summary of the methods is presented in the following section. Any differences between the previously reported methods are highlighted.

#### **3.4.1 Clinically relevant structures of the equine foot**

The selection process used to identify key clinically relevant structures of the equine foot is presented in Chapter 2 (section 2.4.1). Briefly, the following anatomical structures were included in image assessment based on their clinical relevance and the findings of studies reporting the pathology identified during clinical MRI of the equine foot:

1. Deep digital flexor tendon
2. Navicular bone
3. Navicular bursa
4. Distal interphalangeal joint
5. Collateral ligaments of the distal interphalangeal joint
6. Third phalanx
7. Distal sesamoidean impar ligament

#### **3.4.2 Image assessment platform**

The online image assessment platform described in Chapter 2 (2.4.2) was used for this study (Online Surveys, Jisc). This allowed observers to review an MRI foot study and simultaneously grade image quality. The image assessment interface

consisted of two main components: whole study image quality assessment and individual anatomical structure assessment.

The whole study image quality assessment used a 1-4 grading scale. Briefly, this consisted of, grade 1: textbook image quality, grade 2: high diagnostic quality, grade 3: satisfactory diagnostic quality and grade 4: non-diagnostic with more extensive verbal descriptors included for observers (Table 2-2).

Individual anatomical structure assessment utilised modified versions of the 1-4 image quality gradings scale adapted to incorporate the expected MRI characteristics of each structure. The descriptors were based on the radiological Roentgen signs including the structure size, shape, location, number, margination and signal intensity (Thrall, 2018). The full descriptions for the seven anatomical structures are presented in Chapter 2, sections 2.4.2.3 to 2.4.2.8. Pathology was also assessed using a simple grade 1-4 grading scale (grade 1: no pathology, grade 2: mild pathology, grade 3: moderate pathology, grade 4: severe pathology). Inclusion of pathology assessment was for the purpose of quantification of pathology in the studies from each system, rather than for comparison of the pathology identification ability of the two systems. The observers were provided with a handbook that provided guidance on the use of the grading scales. At the end of the image assessment section for each MRI study there was a field for free text comments by the observers, though no overt instructions were given to guide these comments.

### **3.4.3 Magnetic resonance imaging systems and studies**

#### **3.4.3.1 Magnetic resonance imaging systems used in the study**

To assess the significance of field strength (low-field under general anaesthesia versus high-field under general anaesthesia), MRI studies of the equine foot were collected from the following system configurations:

1. Low-field (0.3T), non-weight-bearing (anaesthetised) open magnet system.
2. High-field (1.5T) non-weight-bearing (anaesthetised) closed magnet system.

The systems were collected from the clinical databases of two institutions. The low-field MR images were derived from the Esaote O-scan equine (O-scan Equine, Esaote UK) MRI system at the Weipers Centre Equine Hospital, University of Glasgow. The high-field images were retrieved from the 1.5T Siemens Symphony unit (Symphony, Siemens Medical Solutions USA, Inc.) at the North Carolina State University Veterinary Hospital.

#### **3.4.3.2 Selection of magnetic resonance imaging studies**

The principles and criteria for selection of MRI studies was broadly comparable to that presented in Chapter 2 (section 2.4.3.2). The precise methods of study selection for the low-field under general anaesthesia system are described in in Chapter 2 (section 2.4.3.2).

To acquire the high-field MRI studies, a database was assembled that contained all the equine foot MRI studies performed during the defined period. A random integer generator (Random.org., Random Integer Generator, <https://www.random.org/integers/>) was used to randomly select six studies. The studies were reviewed by the supervising Diplomate of the American College of Veterinary Radiology at the institution to ensure they contained all the sequences routinely included in a foot MRI study at North Carolina State University Veterinary Hospital. Six studies were collected, so that a single study could be excluded should it not meet all the inclusion criteria when assessed by the author. The Digital Imaging and Communications in Medicine (DICOM) files were anonymised by the supervising Diplomate of the American College of Veterinary Radiology using a DICOM modification tool (DICOM Anonymizer <https://dicomanonymizer.com/index.html>). All studies were deemed to meet the inclusion criteria, so following arbitrary ordering of studies, the random integer generator was used to select the final five studies for inclusion in the study. Some MRI studies from this group contained two imaged limbs. The limb included from each patient was randomly selected using a binary modification of the random integer generator.



### **3.4.3.3 Presentation of the magnetic resonance imaging studies**

The methods used to process the DICOM files and then adapt their metadata attributes for the purposes of the study was comparable to those presented in Chapter 2 (section 2.4.3.3).

The studies listed in the MRI study database were arbitrarily numbered and a random sequence generator (Random.org., Random Sequence Generator, <https://www.random.org/sequences/>) produced the order for the studies to be presented in the image assessment platform. The anonymised and labelled DICOM files for each study were saved in folders named corresponding to their assigned 'Case Number'.

### **3.4.4 Observers**

Features of observer recruitment, payment, acknowledgement, and consent were comparable those described in Chapter 2 (section 2.4.4). The observers were provided with material to assist with orientation and use of the online image assessment platform.

An observer profile section was included at the start of the online platform. This recorded the observers' current speciality status and any other relevant postgraduate qualifications. The observers also reported the duration of their experience interpreting equine MR images and their relative experience interpreting images from different MRI systems (including low-field and high-field). Observers selected which DICOM viewing software they would use to view the images in the project.

### **3.4.5 Distribution of the image assessment platform and magnetic resonance imaging studies**

The online image assessment platform was circulated with the DICOM files of the relevant MRI studies by email. Observers received a unique Uniform Resource Locator (URL) to allow tracking of individual responses and to permit observers to save their progress and return to the platform later. A copy of the Observer Handbook was also included in contact with the observers, to provide guidance on the use of the image assessment platform, the grading scales and the DICOM

files. Observers were advised that once downloaded the DICOM files could be viewed in their DICOM viewer of choice and could be manipulated as they would for clinical interpretation.

### **3.4.6 Data analysis**

#### **3.4.6.1 Initial data handling and key statistics**

Raw data was exported from the image assessment platform into a database and coded for further analysis (Microsoft Office Excel 2016, Microsoft Corporation). Frequency tables were assembled for relevant data including from the image quality assessment, pathology assessment and observer profile. Percentages were calculated for each cell of the frequency tables. With all anatomical structures combined the proportion image quality and pathology assessment gradings assigned to each grade (from 1 to 4) was also calculated to assess the distribution of grading scale use by observers. Primary graphical evaluation utilised line graphs to visualise patterns for individual observers and each MRI study for image quality and pathology assessment.

The median, mode and range were calculated as appropriate for ordinal data. Contingency tables were constructed for image quality assessment of entire studies and individual anatomic structures with gradings grouped as diagnostic (image quality grades 1-3) and non-diagnostic (image quality grade 4). Comparable tables were also constructed for equal to or greater than high diagnostic quality (image quality grades 1 and 2) and satisfactory diagnostic quality or less (image quality grades 3 and 4). Similar contingency tables were also assembled for pathology grading for individual structures.

A separate database was established for pulse sequence data for the MRI studies (Microsoft Office Excel 2016, Microsoft Corporation). For each MRI study pulse sequence data documented included the pulse sequence, orientation, echo time, repetition time, field of view, slice thickness, interslice spacing, number of slices, flip angle and inversion time. Descriptive statistical analysis was performed including the median and range for number of sequences (including multiplanar reconstructions if these were routinely performed) for low-field and high-field studies.

### **3.4.6.2 Data analysis for image quality assessment and pathology assessment**

Statistical analysis was performed using the statistical software Minitab (Minitab 18.1, Minitab Ltd.) and SPSS (IBM SPSS Statistics, IBM United Kingdom Ltd.).  $P < .05$  were considered significant for further analysis.

Confidence intervals (95%) were calculated for the proportions of diagnostic MRI studies and greater than or equal to high diagnostic quality MRI studies using an exact method for low-field and high-field studies. The proportions used for this analysis were derived from the contingency tables described in the initial data handling.

Statistical analysis was based on comparison to assess the effect of field strength (low-field under general anaesthesia compared to high-field under general anaesthesia). For entire studies and individual anatomical structures, the proportion of diagnostic and non-diagnostic gradings was compared between low-field and high-field groups using a Pearson chi-square test or Fisher's exact test (if one or more expected cell values  $\leq 5$ ). The ranked gradings for image quality and pathology, were compared between groups for each assessment category using a Mann-Whitney test. Graphical visualisation included plotting of clustered bar charts for image quality and pathology assessment demonstrating the count of observer gradings for each assessment category (entire study or individual anatomical structure) clustered by grade.

### **3.4.6.3 Free text comments**

The free text comments were removed from the initial database and grouped for each study into a separate database (Microsoft Office Excel 2016, Microsoft Corporation). The comments were categorised as related to image quality or pathology assessment (or both where relevant). Comments relating to reduced image quality were grouped into the following topics:

- alignment or position
- additional sequence desired
- repeat sequence desired

- suboptimal short tau inversion recovery (STIR) fat suppression (for example incomplete fat suppression)
- sequence parameter (other)
- motion artefact
- magic angle effect
- artefact (other)
- other

Comments that referred to multiple categories were documented as a count in each category. Clustered bar charts were plotted demonstrating the count of free text comments, with clustering by the category.

#### 3.4.6.4 Inter-observer agreement for image quality assessment

The inter-observer agreement for image quality assessment grading was determined using Fleiss' kappa and Kendall's coefficient of concordance. Agreement analysis was performed for entire studies and individual anatomical structures, encompassing all studies from the low-field and high-field groups. Gradings from all observers were included in the analysis. Interpretation of kappa values was performed using previously reported generic descriptors presented in Table 3-1 (Altman, 1991). There are no established descriptors for interpretation of Kendall's coefficient of concordance though higher values (from 0 to 1) demonstrate greater agreement (Gwet, 2014).

**Table 3-1 Interpretation of kappa statistic values based on Altman (1991)**

Kappa statistic	Interpretation
<0.20	poor agreement
0.21-0.40	fair agreement
0.41-0.60	moderate agreement
0.61-0.80	good agreement
0.81-1.00	very good agreement

## **3.5 Results**

### **3.5.1 General results**

#### **3.5.1.1 Observer profile**

All observers had >5 years' experience interpreting equine MR images, with 4/10 having 5-10 years' experience and 6/10 having >10 years' experience. There were six Diagnostic imaging Diplomates, one diagnostic imaging Associate Member, one diagnostic imaging Associate Member and surgical Diplomat, one surgical, sports medicine and rehabilitation Diplomat and diagnostic imaging Associate Member, and one surgical and sports medicine and rehabilitation Diplomat.

Overall, observers had relatively similar experience interpreting images from low-field under general anaesthesia (regularly: 1/10, frequently: 0/10, occasionally: 5/10, rarely or never: 4/10) and high-field under general anaesthesia (regularly: 0/10, frequently: 2/10, occasionally: 3/10, rarely or never: 5/10). The majority of observers interpreted images from low-field standing systems most frequently.

#### **3.5.1.2 Magnetic resonance imaging studies**

All the MRI studies contained T1 weighted, T2 weighted, proton density weighted, and fat suppressed (STIR) sequences. Sequences were acquired in at least three planes in all studies. The median number of sequences per study for low-field under general anaesthesia was 12 (range 11-14) and for high-field under general anaesthesia was 11 (range 11-12). Multiplanar reconstructions were routinely performed from isotropic T1 sequences from the low-field MRI system and were included in these sequence counts. Orientation sequences (localisers or pilots) were not included in these counts. The pulse sequence parameters for all MRI studies are presented in Appendix 2.

### **3.5.2 Key statistics for image quality assessment**

All observers completed the image quality assessment of the low-field and high-field studies. When observers assessed the image quality of entire studies 88%

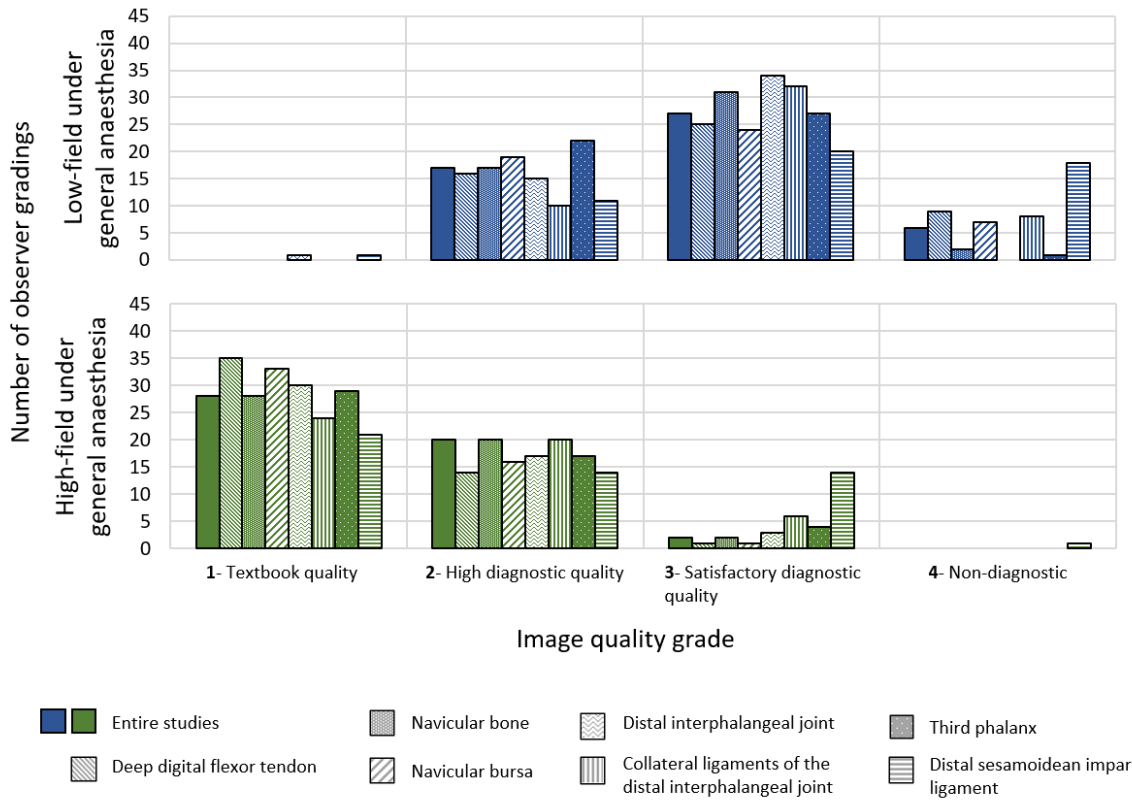
(95% CI 76%, 95%) were graded as diagnostic for the low-field under general anaesthesia studies and 100% (95% CI lower limit 94%) were graded as diagnostic for high-field under general anaesthesia studies. Grouping entire study image quality gradings as high-diagnostic quality or better (those graded 1 or 2) accounted for 34% (95% CI 21%, 49%) of low-field under general anaesthesia studies and 96% (95% CI 86%, 100%) for high-field under general anaesthesia studies. The distribution of image quality gradings for entire MRI studies and individual anatomical structures is displayed in Figure 3-1.

For entire MRI studies, the median grade for image quality was grade 3 for low-field studies and grade 1 for high-field studies. The median image quality grade was grade 3 for all individual anatomical structures in the low-field under general anaesthesia studies. For the high-field under general anaesthesia studies the median image quality grade was grade 1 for 5/7 individual anatomical structures and grade 2 for 2/7 structures. The median and range of image quality gradings for entire MRI studies and individual anatomical structures for the low-field under general anaesthesia and high-field under general anaesthesia studies are presented in Table 3-2.

**Table 3-2 Median and range for image quality grading for low-field and high-field under general anaesthesia studies.**

Abbreviations: DIP- distal interphalangeal.

Assessment category	Low-field under general anaesthesia		High-field under general anaesthesia	
	Median grade	Range	Median grade	Range
Entire study	3	2-4	1	1-3
Deep digital flexor tendon	3	2-4	1	1-3
Navicular bone	3	2-4	1	1-3
Navicular bursa	3	2-4	1	1-3
DIP joint	3	1-3	1	1-3
Collateral ligaments of the DIP joint	3	2-4	2	1-3
Third phalanx	3	2-4	1	1-3
Distal sesamoidean impar ligament	3	1-4	2	1-4



**Figure 3-1** Bar chart displaying the image quality gradings for low-field and high-field under general anaesthesia studies.

### 3.5.3 Comparison of the proportion of diagnostic studies

There was a statistically significant difference in the proportion of diagnostic versus non-diagnostic entire studies between low-field and high-field under general anaesthesia studies. There was also a statistically significant difference in the proportion of diagnostic versus non-diagnostic studies between low-field under general anaesthesia for the deep digital flexor tendon, navicular bursa, collateral ligaments of the distal interphalangeal joint and the distal sesamoidean impar ligament. The P values generated from chi-square and Fisher’s exact tests comparing the proportion of diagnostic gradings are presented in Table 3-3.

**Table 3-3 Comparison of proportion of diagnostic image quality gradings for low-field under general anaesthesia and high-field under general anaesthesia studies.**

Abbreviations: DIP- distal interphalangeal.

<sup>a</sup> Indicates use of Fisher's exact test. Other comparisons used chi-square tests.

<sup>b</sup> Analysis not performed due to 0 value(s) in the contingency table.

Assessment category	Comparison of proportion of diagnostic gradings (P value)
Entire study	0.03 <sup>a</sup>
Deep digital flexor tendon	0.003 <sup>a</sup>
Navicular bone	0.5 <sup>a</sup>
Navicular bursa	0.01 <sup>a</sup>
DIP joint	Not performed <sup>b</sup>
Collateral ligaments of the DIP joint	0.006 <sup>a</sup>
Third phalanx	>0.9 <sup>a</sup>
Distal sesamoidean impar ligament	<0.001

### 3.5.4 Comparisons of the ranked gradings for image quality

Comparing the ranked image quality gradings for entire studies with a Mann-Whitney test, there was a statistically significant difference between low-field and high-field studies. There were also statistically significant differences in the ranked image quality gradings between low-field and high field studies for all individual anatomical structures (Table 3-4).

**Table 3-4 Comparison of ranked gradings for image quality for low-field under general anaesthesia and high-field under general anaesthesia studies.**

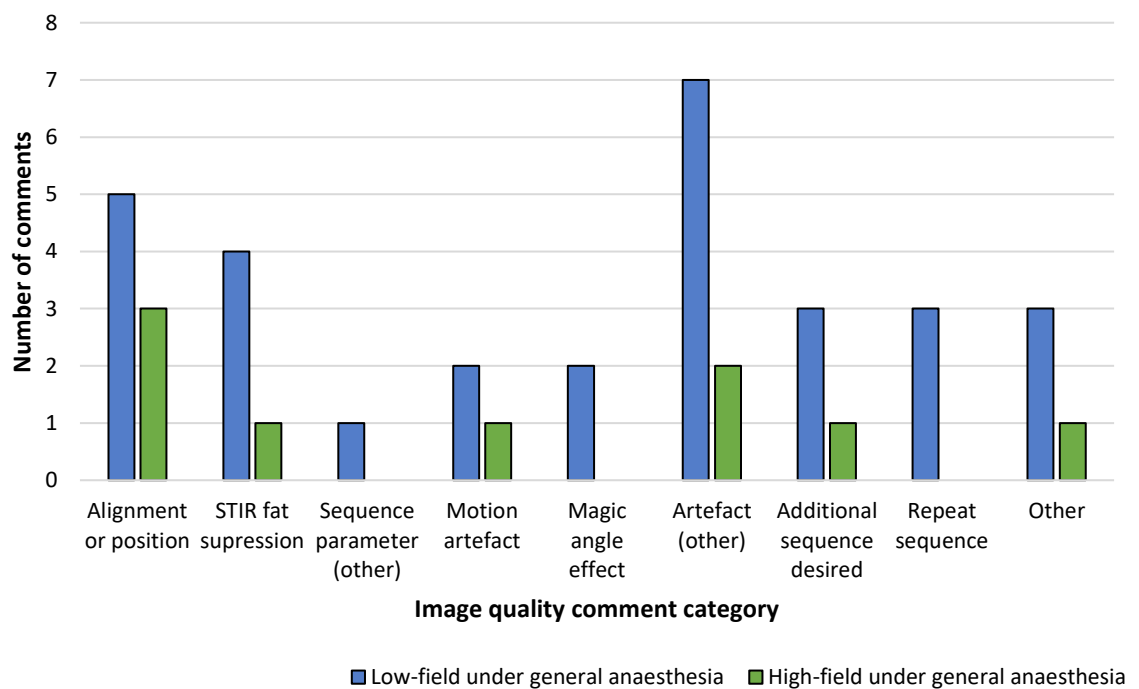
Abbreviations: DIP- distal interphalangeal.

Assessment category	Comparison of ranked gradings for image quality (P value)
Entire study	<0.001
Deep digital flexor tendon	<0.001
Navicular bone	<0.001
Navicular bursa	<0.001
DIP joint	<0.001
Collateral ligaments of the DIP joint	<0.001
Third phalanx	<0.001
Distal sesamoidean impar ligament	<0.001



### 3.5.5 Free text comments regarding image quality

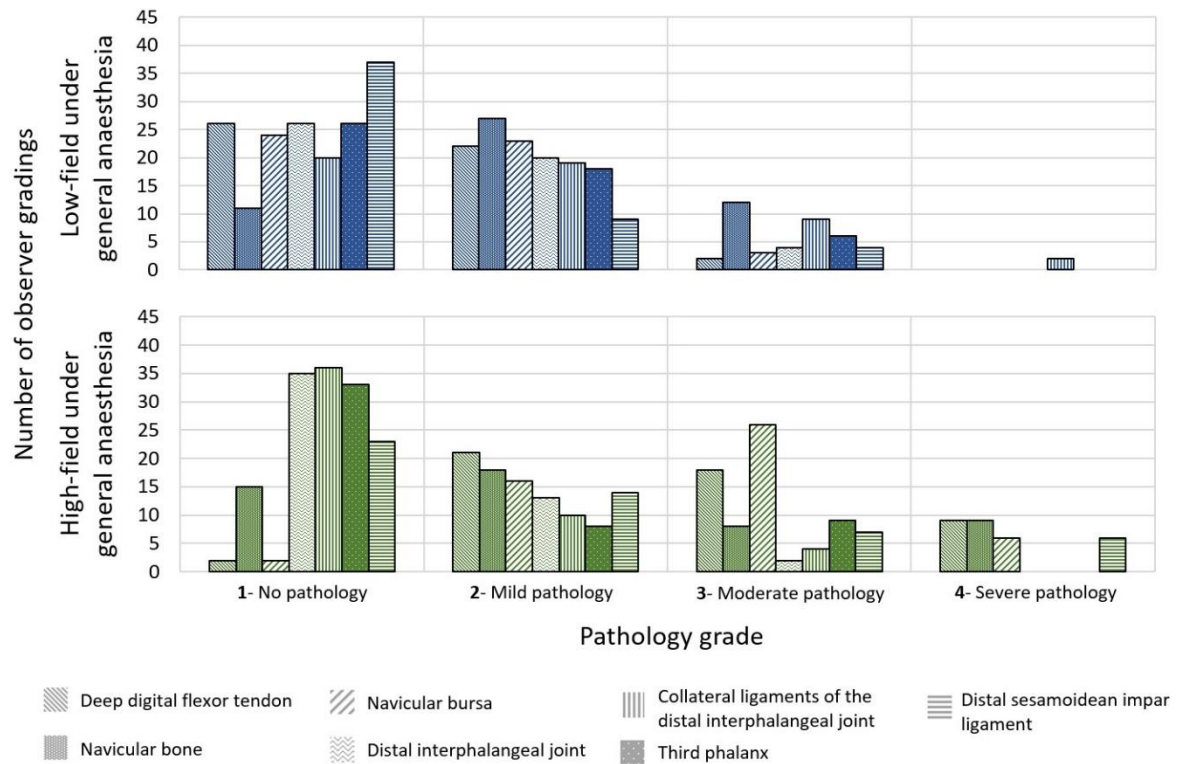
There were 23 free text comments relating to image quality for the low-field studies and 8 comments for the high-field studies. The distribution of image quality comments across the categories for reduced image quality are displayed in Figure 3-2.



**Figure 3-2** Bar chart displaying the categorisation of image quality comments for low-field under general anaesthesia and high-field under general anaesthesia studies.

### 3.5.6 Pathology considerations for image quality assessment

All observers completed the pathology assessment. The distribution of pathology gradings was broadly similar between groups (Figure 3-3). However, when comparing the ranked pathology gradings for individual anatomical structures with a Mann-Whitney test, there were statistically significant differences in pathology between low-field and high-field under general anaesthesia studies for the deep digital flexor tendon, navicular bursa, collateral ligaments of the distal interphalangeal joint and distal sesamoidean impar ligament (Table 3-5).



**Figure 3-3** Bar chart displaying pathology gradings for low-field and high-field under general anaesthesia studies.

**Table 3-5** Median pathology grades and rank comparison for low-field and high-field under general anaesthesia studies.

Abbreviations: DIP- distal interphalangeal.

Assessment category	Median pathology grade		Comparison of ranked gradings for pathology (P value)
	Low-field under general anaesthesia	High-field under general anaesthesia	
Deep digital flexor tendon	1	3	<0.001
Navicular bone	2	2	0.6
Navicular bursa	2	3	<0.001
DIP joint	1	1	0.06
Collateral ligaments of the DIP joint	2	1	0.001
Third phalanx	1	1	0.4
Distal sesamoidean impar ligament	1	2	0.002

### 3.5.7 Inter-observer agreement for image quality

The absolute inter-observer agreement determined by Fleiss' Kappa for image quality grading of entire studies was fair. Absolute inter-observer agreement was fair for the deep digital flexor tendon, navicular bone, navicular bursa and the distal interphalangeal joint. The collateral ligaments of the distal interphalangeal joint, the third phalanx and the distal sesamoidean impar ligament had poor absolute inter-observer agreement. Assessment of inter-observer agreement accounting for the ranking of gradings, as determined by Kendall's coefficient of concordance, ranged from moderate to high for entire studies and individual anatomical structures (Table 3-6).

**Table 3-6 Output of inter-observer agreement analysis for magnetic resonance image quality grading for all low-field under general anaesthesia and high-field under general anaesthesia studies**

Abbreviations: DIP- distal interphalangeal.

Assessment category	Fleiss' kappa		Kendall's coefficient of concordance
	Kappa statistic	Standard error	
Entire studies	0.22	0.03	0.70
Deep digital flexor tendon	0.31	0.03	0.84
Navicular bone	0.21	0.03	0.74
Navicular bursa	0.21	0.03	0.82
DIP joint	0.23	0.03	0.66
Collateral ligaments of the DIP joint	0.20	0.03	0.74
Third phalanx	0.15	0.03	0.61
Distal sesamoidean impar ligament	0.08	0.03	0.61

### 3.6 Discussion

The results of this study demonstrate that field strength has a significant influence on perceived image quality for MRI of the equine foot. More high-field studies were deemed to be of diagnostic quality. Furthermore, increased field strength resulted in increased image quality at the level of individual anatomical structures of the foot.

Field strength is well recognised as a key factor dictating image quality in MRI. This is principally due to the increased signal-to-noise ratio which effectively increases in correlation with increasing magnetic field strength (McRobbie *et al.*, 2017g). Therefore, magnetic field strength has been a topic of discussion and investigation in the equine MRI literature (Werpy, 2007). Previous cadaver studies comparing high-field and low-field images of the equine distal limb demonstrated that both images had considerable anatomic/diagnostic value, but additional detail may be afforded by high-field images (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). Other factors, in addition to field strength, differed between systems in these studies particularly given that they compared a low-field system designed for use with standing sedated patients with high-field systems used with the patient under general anaesthesia (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). This disparity alters the appearance of some anatomical structures, changes the radiofrequency coils available and alters the occurrence of magic angle effect (Spriet, Mai and McKnight, 2007; Smith, Dyson and Murray, 2008; Murray *et al.*, 2009; Spriet and Zwingenberger, 2009; Bolas, 2010; Werpy, Ho and Kawcak, 2010). These factors can also have important implications for image quality and could act as confounders for the apparent influence of field-strength (Murray *et al.*, 2009). The current study compared low-field and high-field systems designed for equine patients under general anaesthesia. These systems have more comparable positioning and therefore comparison provides a more accurate assessment of the influence of field strength on image quality. In the current study magic angle effect was not reported in the high-field group and was only described in two comments for the low-field studies.

For assessment at the level of entire MRI studies, high-field images of the equine foot were of greater diagnostic quality compared to low-field studies, based on

the proportion of diagnostic studies and ranked gradings. When comparing the proportion of diagnostic studies between high and low-field studies for individual structures, there were no significant differences between groups for the navicular bone, the distal interphalangeal joint and the third phalanx. This apparent lack of difference (when comparing the proportion of diagnostic studies) may reflect disparities in the ease at which certain anatomical structures are assessed on MRI compared to others. Such a disparity would effectively result in a different threshold of image quality to be deemed diagnostic for each anatomic structure. Similarly, the diagnostic value of MRI sequences can vary between tissues and anatomical structures (Mair *et al.*, 2005; Murray *et al.*, 2007; Olive, 2010; Werpy *et al.*, 2011; Smith, Dyson and Murray, 2012; van Zadelhoff *et al.*, 2020). Consequently, the relative disparity in image quality between high-field and low-field images may not be consistent across all sequence types. Therefore, evaluation of structures reliant on sequences where the disparity in image quality between field strengths is greatest, may be the most likely to demonstrate challenges to low-field diagnostic quality.

A previous study comparing image quality of cadaver limbs from a low-field standing system with high-field images highlighted a relatively poor score for evaluation of the third phalanx on low-field images (Bolen, Audigié, *et al.*, 2010). This feature was due to poor signal from the distal dorsal aspect of the third phalanx which is placed at the periphery of the magnet and the radiofrequency coil of the low-field standing unit. In the current study the low-field system also featured an open, permanent magnet but used a cylindrical human knee coil. Poor signal at the periphery of the third phalanx was not a feature of the current study, indicating that the magnet and coil configuration of the system may provide a more homogenous signal or a field of view that encompasses the toe (including the distal third phalanx). It is worth noting that the low-field system utilised in the previous study has now been superseded by a new iteration the Hallmarq EQ2 (Bolen, Audigié, *et al.*, 2010; Schramme and Segard-Weisse, 2020).

Identification of articular cartilage pathology on MRI is challenging (Smith, Dyson and Murray, 2012). Studies have highlighted the limitations of identifying

articular cartilage pathology on low field images, particularly in the standing patient (Murray *et al.*, 2009; van Zadelhoff *et al.*, 2020). However, positioning and weight-bearing are known to influence assessment of articular cartilage on MRI. The comparable non-weight-bearing positioning in the low-field and high-field systems of the current study effectively reduced the effect of these confounding factors. In the current study all MRI studies (from both high-field and low-field systems) were deemed to be diagnostic for assessment of the distal interphalangeal joint. However, the high-field group had significantly higher image quality for assessment of the distal interphalangeal joint when the ranked grades were compared. This finding is consistent with other studies that have reported improved cartilage visualisation on high-field images (Murray *et al.*, 2009). Relatively subtle MRI changes can be significant in this structure (van Zadelhoff *et al.*, 2020). Currently, evidence indicates that high-field imaging may be preferable for the identification of articular cartilage pathology (including the distal interphalangeal joint).

Although there were no significant differences in the proportion of diagnostic studies between the system types for 3/7 anatomical structures assessed, when the ranked gradings were compared, high-field images were deemed to have significantly greater image quality for all structures. The first comparison demonstrates that images from both systems were sufficient quality to be considered diagnostic, but the second comparison reveals the significant benefit high-field imaging has on image quality. This broad assessment is consistent with previous cadaver studies that found that images from both system types have diagnostic value but high-field imaging provides greater characterisation of small anatomical structures or lesions (Werpy, 2007; Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). The increased image quality is primarily associated with the improved signal-to-noise which increases with field strength, though this is not a directly linear relationship (Maubon *et al.*, 1999; Werpy, 2007; Collins, 2016). Higher field strengths afford the operator with opportunities to substitute the improved signal-to-noise ratio with other parameters, such as reduced scan times or improved spatial resolution, without compromising image quality (Westbrook, Roth and Talbot, 2011e; McRobbie *et al.*, 2017g).

It is expected that an increase in image quality results in improved pathology identification and diagnostic confidence (Bolen, Audigié, *et al.*, 2010). However, this is a complex interaction and whilst these characteristics are clearly intimately associated, correlation is not necessarily linear. Experienced observers may be able to deduce diagnostic information even when a structure is difficult to visualise, for example assessment of the associated navicular bursa and osseous margins may infer valuable information about the distal sesamoidean impar ligament (Murray *et al.*, 2009; Dyson *et al.*, 2010). The balance of progressive improvement in image quality (particularly by increasing field strength) with what is required to obtain an accurate diagnosis is a long-standing discourse in human orthopaedic imaging (Ghazinoor, Crues and Crowley, 2007; Edelman, 2014; Sarracanie and Salameh, 2020). Similar discussions in the equine literature have also emphasised the need to define the boundary of image quality that is sufficient for diagnosis (Werpy, 2007). Comparative histological studies have demonstrated the value of MRI for identification of pathology of the equine foot (Busoni *et al.*, 2005; Murray, Blunden, *et al.*, 2006; Murray *et al.*, 2007; Dyson, Blunden and Murray, 2008, 2012; Blunden, Murray and Dyson, 2009; Blunden *et al.*, 2010a, 2010b; Kottmeier *et al.*, 2020; van Zadelhoff *et al.*, 2020). However, further research is required to determine the relative diagnostic accuracy (and diagnostic confidence of observers) of different field strengths during MRI of clinical equine patients.

There were few free text comments relating to image quality for the high-field MRI studies, which is predictable given the excellent image quality reported by observers. The spread of image quality comments was relatively broad across the comment categories for both systems, but with a trend to a greater number of comments for the low-field system. There were limited comments related to motion artefact, which is unsurprising given that imaging was performed under general anaesthesia (Porter and Werpy, 2014). There were more comments relating to artefacts for the low-field group. Both groups had comments related to factors such as positioning, alignment and sequence selection, indicating that operator (or institutional protocol) factors are important considerations for image quality that transcend field strength.

Agreement analysis demonstrated poor to fair absolute inter-observer agreement for image quality assessment of entire studies and individual structures. However, agreement accounting for the ranking of gradings, as indicated by Kendall's coefficient of concordance, was moderate to high for entire studies and individual anatomical structures. An ordinal grading system was used to guide image quality assessment, but image quality assessment is still an inherently subjective and multifactorial process. Therefore, Kendall's coefficient of concordance is a more useful indicator of agreement. Inter-observer agreement was acceptable for the purposes of image quality assessment. There were some statistically significant differences in pathology between studies in the low-field and high-field groups (for the deep digital flexor tendon, navicular bursa, collateral ligaments of the distal interphalangeal joint and the distal sesamoidean impar ligament). These differences in pathology were not deemed to be clinically significant for the purposes of image quality assessment in the current study.

In conclusion, field strength has an important influence on perceived image quality for MRI of the equine foot in a clinical context. Whilst most studies were deemed to be diagnostic regardless of field strength, high-field MRI studies had greater image quality, including at the level of individual anatomical structures. These findings indicate that high-field MRI is desirable in situations where image quality is prioritised, but other factors including cost, availability and patient risk must also be considered when selecting an MRI system in a clinical context.



# Chapter 4 Inter-observer agreement for the assessment of pathology during clinical magnetic resonance imaging of the equine foot

## 4.1 Introduction

Magnetic resonance imaging (MRI) is a fundamental diagnostic imaging modality used during the investigation of equine orthopaedics (Kleiter *et al.*, 1999; Dyson *et al.*, 2003; Barrett *et al.*, 2017). MRI has developed a particularly important role in imaging of the equine foot, where the limitations imposed by the hoof capsule and the complex anatomy have restrictions for some other modalities. Interpretation of magnetic resonance (MR) images is complex, particularly when utilised for regions with challenging anatomy (Dyson and Murray, 2010). From a clinical perspective, the primary focus of MRI inter-observer agreement investigation is determining whether observers produce consistent interpretation and diagnosis. Ultimately MRI interpretation influences decisions about case prognosis and treatment so consistency in pathology assessment is important to optimise patient outcomes. In a broader context, inter-observer agreement is important to ensure that equine MRI clinical research disseminates accurately between parties, for example between research groups and those undertaking MRI interpretation in a clinical setting.

There is limited literature demonstrating the level of agreement between observers for any form of assessment of equine MR images (Murray *et al.*, 2007). A cadaver study reported variable inter-observer agreement for assessment of sequence quality during evaluation of the distal impar sesamoidean ligament (Berner *et al.*, 2020). A study using three observers to assess image quality for MRI of the equine foot at different field strengths also demonstrated differences in scores between observers (Bolen, Audigié, *et al.*, 2010). A further cadaver study has investigated the ability of observers to recognise weightbearing versus non-weightbearing positioning of the equine foot during MRI, with a particular focus on the distal interphalangeal joint (Evrard *et al.*, 2019). There was variation in ability of individual observers to identify positioning of the foot with improving ability correlating with experience of equine MR image interpretation (Evrard *et al.*, 2019). A study investigating the effects of diagnostic anaesthesia prior to MRI of the foot reported variable inter-observer agreement for

evaluation of features such as the volume of synovial structures (Black *et al.*, 2013).

These studies report inter-observer agreement and variation for some aspects of MRI interpretation, such as image quality or identification of individual anatomical structures. However, there is little information to characterise the inter-observer agreement for the assessment of pathology or how study interpretation is performed during MRI in the equine patient. This is surprising, given the ubiquity of MRI in equine orthopaedics, particularly in relation to imaging of the foot.

## **4.2 Aims**

The aim of this study was to investigate inter-observer agreement for the assessment of pathology of key anatomical structures of the equine foot during interpretation of clinical MRI studies.

## **4.3 Study design**

This study had an experimental study design involving prospective observer assessment of MRI studies collected retrospectively from clinical imaging databases.

## **4.4 Methods**

The methods utilised in this study are related to those described in Chapter 2. Where similar methods were used a summary is provided here. Detailed descriptions are provided where different methods were used.

### **4.4.1 Clinically relevant structures of the equine foot**

Anatomical structures of the equine foot assessed in the study were selected based review of relevant literature reporting the distribution of lesions during MRI and their clinical significance. The selection process for anatomical structures was comparable to that outlined previously and can be found in Chapter 2 (section 2.4.1). Following the selection process, the following seven structures were included in the image assessment:

1. Deep digital flexor tendon
2. Navicular bone
3. Navicular bursa
4. Distal interphalangeal joint
5. Collateral ligaments of the distal interphalangeal joint
6. Third phalanx
7. Distal sesamoidean impar ligament

#### **4.4.2 Image assessment platform**

An MR image assessment platform was developed to allow observers to assess clinical MRI studies of the equine foot, using an online tool (Online Surveys, Jisc). Observer's graded pathology of individual anatomic structures for each MRI study. The pathology assessment used a 1-4 grading system for each anatomical structure:

- Grade 1: no pathology
- Grade 2: mild pathology
- Grade 3: moderate pathology
- Grade 4: severe pathology

After assigning a pathology assessment grade, the image assessment required observers to provide a pathology assessment confidence grade for each anatomical structure of each MRI study. The pathology assessment confidence was also graded on a 1-4 grading system:

- Grade 1: high confidence
- Grade 2: moderate confidence
- Grade 3: limited confidence
- Grade 4: no confidence

For each MRI study the individual anatomical structure assessment component of the platform was formatted as a grid with a selection list for pathology assessment and pathology assessment confidence (Figure 4-1). The observers were also provided with an Observer Handbook that outlined details of the study and the grading systems utilised.

	Extent of pathology within this structure	How confident are you with this assessment of pathology?
Deep digital flexor tendon	Please select ▾	Please select ▾
Navicular bone	Please select ▾	Please select ▾
Navicular bursa	Please select ▾	Please select ▾
DIP joint	Please select ▾	Please select ▾
Collateral ligaments of DIP joint	Please select ▾	Please select ▾
Third phalanx	Please select ▾	Please select ▾
Distal sesamoidean impar ligament	Please select ▾	Please select ▾

**Figure 4-1 Format of the pathology and pathology assessment confidence interface.**

### 4.4.3 Magnetic resonance imaging studies

A total of fifteen MRI studies of the equine foot were used in combination with the image assessment platform. Five studies were taken from each of the following MRI systems:

- Low-field (0.3T), weight-bearing (patient standing, sedated) open magnet system (EQ2 standing equine MRI, Hallmarq Veterinary Imaging Ltd).
- Low-field (0.3T), non-weight-bearing (patient under general anaesthesia) open magnet system (O-scan Equine, Esaote).
- High-field (1.5T) non-weight-bearing (patient under general anaesthesia) closed magnet system (Symphony, Siemens Medical Solutions USA, Inc.).

The MRI studies were randomly selected from the clinical databases of live patients from each acquiring institution. MRI studies needed to include all the sequences that would typically be performed as a routine equine foot study at the acquiring institution. All studies were anonymised and presented in a randomised order. Additional information regarding the selection and processing of study Digital Imaging and Communications in Medicine (DICOM) files is presented in Chapter 2 (section 2.4.3).

#### **4.4.4 Observers**

The experienced observers involved in the project were the same as those recruited for Chapters 2 and 3. Additional information regarding observer recruitment, acknowledgement, payment and consent is presented in Chapter 2 (section 2.4.4).

##### **4.4.4.1 Observer profile**

An observer profile section was included at the start of the image quality assessment platform (Figure 4-2). This recorded the qualifications of the observers, the duration of their experience interpreting equine MR images and the frequency with which observers interpret images from different MRI systems. Given that observers were permitted to use their preferred DICOM viewing software, they were asked to report which software they would use during the study.

##### **4.4.5 Distribution of the image assessment platform and magnetic resonance imaging studies**

The image assessment platform and the folders containing the DICOM files for the MRI studies were distributed to observers by email. Each observer received a unique Uniform Resource Locator (URL) to access the image assessment platform. This permitted observers to save their progress through the platform and return to complete the assessment at their convenience. The unique URL also allowed individual observer responses to be tracked. A copy of the Observer Handbook was also included in the materials. Observers were informed that they were free to view the MRI studies using their DICOM viewer of choice and could manipulate the studies as they would in a clinical context.

## Page 2: Observer Details

What is your current speciality status? (select all that apply)

- Diagnostic imaging- Diplomat
- Diagnostic imaging- Resident
- Diagnostic imaging- Associate ECVDI
- Surgery- Diplomat
- Surgery- Resident
- Sports Medicine- Resident
- Sports Medicine- Diplomat
- Certificate holder (CertAVP or similar)
- Other

If you selected Other, please specify:

Please indicate the duration of your experience interpreting equine MR images.

- < 1 year
- 1-5 years
- 5-10 years
- >10 years

This part of the survey uses a table of questions, [view as separate questions instead?](#)

Please rate the frequency with which you interpret images acquired from the following MR systems  
Abbreviations (GA- general anaesthesia)

Please don't select more than 1 answer(s) per row.

Please select at least 3 answer(s).

	Regularly (daily)	Frequently (weekly)	Occasionally (monthly)	Rarely or never
Low field (0.3T) standing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Low field (0.3T)under GA	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
High field (>1T)under GA	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Which DICOM viewer will you be using to view the images?

- Horos
- OsiriX
- Asteris
- ClearCanvas
- eFilm
- Visbion
- Other

**Figure 4-2** The observer profile section of the image assessment platform

## **4.4.6 Data analysis**

### **4.4.6.1 Initial data handling and key descriptive analysis**

The data was exported from the image assessment platform and assembled into a database (Microsoft Office Excel 2016, Microsoft Corporation). The database was coded for further analysis. Relevant data from the observer profile, pathology assessment and pathology assessment components were assembled into frequency tables. Percentages were calculated for each category of the frequency tables. Clustered bar charts were plotted to visualise the distribution of pathology assessment and pathology assessment confidence counts at each grade, with clustering by anatomical structure. The pathology gradings for the individual anatomical structures of each MRI study were grouped into contingency tables, with assessments dichotomously categorised as pathology deemed present (pathology grades 2, 3 and 4) or absent (pathology grade 1).

Pulse sequence data for the MRI studies were compiled into a separate database. The documented parameters included the pulse sequence type, orientation, echo time, repetition time, field of view, slice thickness, interslice spacing, number of slices, flip angle and for relevant sequences the inversion time.

### **4.4.6.2 Inter-observer agreement for pathology assessment**

Inter-observer agreement for the dichotomous presence or absence of pathology in individual anatomical structures was assessed with the percentage of matched responses (where all observers agreed on the presence or absence of pathology in a study) and Fleiss' kappa ( $k$ ) across all MR imaging studies. Inter-observer agreement for pathology assessment grading and pathology assessment confidence grading was assessed using percentage of matched assessments, Fleiss' kappa and Kendall's coefficient of concordance for individual anatomical structures. This agreement analysis included studies from all acquisition systems and gradings by all observers. The standard error and 95% confidence intervals (by an asymptotic method) were also calculated for Fleiss' kappa. Interpretation of Fleiss' kappa used previously described arbitrary values (Altman, 1991): <0.20 poor agreement, 0.21-0.40 fair agreement, 0.41-0.60 moderate agreement, 0.61-0.80 good agreement and 0.81-1.00 very good agreement. There are no arbitrary values described for interpretation of Kendall's coefficient of

concordance but values can range from 0 to 1 with larger values indicating greater concordance (Gwet, 2014; Minitab, 2020)

Bubble charts were also produced to demonstrate the distribution of gradings by observers for pathology assessment and pathology assessment confidence. Linear dot plots were assembled to demonstrate the Fleiss' kappa values for individual anatomic structures when considering individual pathology assessment grades and the combined Fleiss' kappa values for all grades.

#### **4.4.6.3 Inter-observer agreement for pathology assessment confidence**

The pathology assessment confidence data was handled in a similar manner to that of pathology assessment. Inter-observer agreement for pathology assessment confidence grading was assessed using the percentage of matched assessments, Fleiss' kappa and Kendall's coefficient of concordance for individual anatomical structures. This agreement analysis included studies from all acquisition systems and gradings by all observers. The standard error and 95% confidence intervals (by an asymptotic method) were also calculated for Fleiss' kappa. Interpretation of Fleiss' kappa and Kendall's coefficient of concordance was comparable to that described for pathology assessment (section 4.4.6.2).

Bubble charts were also produced to demonstrate the distribution of gradings by observers for pathology assessment confidence. Linear dot plots were assembled to demonstrate the Fleiss' kappa values for individual anatomic structures when considering individual pathology assessment grades and the combined Fleiss' kappa values for all grades.



## 4.5 Results

### 4.5.1 General results

#### 4.5.1.1 Observer profile

All observers completed the observer profile and the image assessment platform. Each observer held Diplomate or Associate Member status in at least one relevant field with: six diagnostic imaging Diplomates, one diagnostic imaging Associate Member, one diagnostic imaging Associate Member and surgical Diplomate, one surgical and sports medicine and rehabilitation Diplomate and diagnostic imaging Associate Member, and one surgical and sports medicine and rehabilitation Diplomate. Postgraduate certifications (equivalent to the Royal College of Veterinary Surgeons Certificate in Advanced Veterinary Practice) were held by two observers.

Six observers had >10 years' and four had 5-10 years' experience interpreting equine MR images. Observers interpreted images from low-field standing systems most frequently (Table 4-1). OsiriX was the most used DICOM viewer, accounting for 6/10 observers. Horos was used by 3/10 observers and a single observer reported use of both Horos and eFilm.

**Table 4-1 Frequency with which observers interpret images from different magnetic resonance imaging acquisition systems.**

Frequency interpreting images from system	Type of magnetic resonance imaging acquisition system		
	Low-field standing	Low-field under general anaesthesia	High-field under general anaesthesia
Regularly (daily)	6	1	0
Frequently (weekly)	2	0	2
Occasionally (monthly)	0	5	3
Rarely or never	2	4	5

#### **4.5.1.2 Magnetic resonance imaging studies**

Each MRI study contained T1 weighted, T2 weighted, proton density weighted and fat suppressed (STIR) sequences. All MRI studies contained sequences in at least three orthogonal planes. The pulse sequence parameters of the MRI studies are presented in Appendix 2.

#### **4.5.2 Observer pathology assessment**

##### **4.5.2.1 Key descriptive analysis**

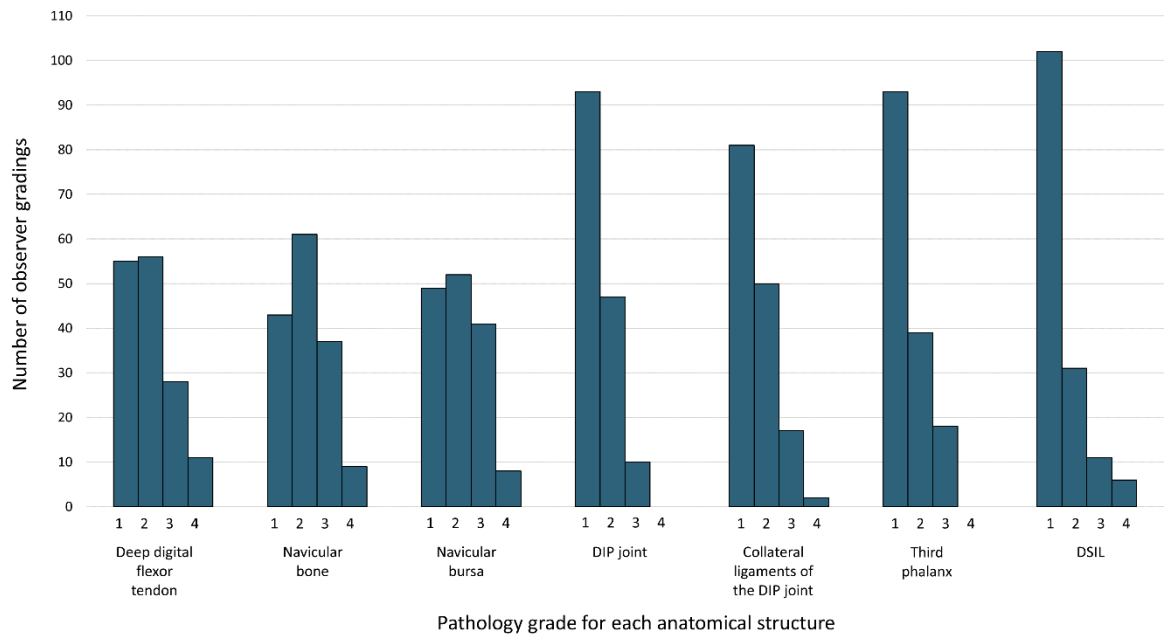
The distribution of pathology gradings varied across individual structures. All anatomic structures had MRI studies where observers assigned pathology grade 1 (no pathology), grade 2 (mild pathology) and grade 3 (moderate pathology). Other than the distal interphalangeal joint and the third phalanx all structures also had studies where observers assigned pathology grade 4 (severe pathology). The count of pathology gradings assigned to each grade across all MR imaging studies for individual anatomical structures is presented in Table 4-2 and Figure 4-3. Bubble charts were deemed to provide a more valuable visualisation of the data compared to radar charts. Bubble charts for each anatomical structure, demonstrating the spread of observer pathology assessment grades for each MRI study are presented in Appendix 4.

Median pathology gradings for individual MRI studies ranged from grade 1 to 4 for the deep digital flexor tendon, navicular bone and distal sesamoid impar ligament, from grade 1 to 3.5 for the navicular bursa, grade 1 to 3 for collateral ligaments of the distal interphalangeal joint and the third phalanx, and from grade 1 to 2 for the distal interphalangeal joint.

**Table 4-2 Pathology gradings assigned to each grade for individual anatomical structures for 15 magnetic resonance imaging studies of the equine foot.**

Abbreviations: DIP- distal interphalangeal.

Assessment Category	Count of pathology assessment gradings assigned to each grade across all magnetic resonance imaging studies (%)			
	Grade 1- no pathology	Grade 2- mild pathology	Grade 3- moderate pathology	Grade 4- severe pathology
Deep digital flexor tendon	55 (36.7%)	56 (37.3%)	28 (18.7%)	11 (7.3%)
Navicular bone	43 (28.7%)	61 (40.7%)	37 (24.7%)	9 (6.0%)
Navicular bursa	49 (32.7%)	52 (34.7%)	41 (27.3%)	8 (5.3%)
DIP joint	93 (62.0%)	47 (31.3%)	10 (6.7%)	0 (0.0%)
Collateral ligaments of the DIP joint	81 (54.0%)	50 (33.3%)	17 (11.3%)	2 (1.3%)
Third phalanx	93 (62.0%)	39 (26.0%)	18 (12.0%)	0 (0.0%)
Distal sesamoidean impar ligament	102 (68.0%)	31 (20.7%)	11 (7.3%)	6 (4.0%)
All anatomical structures combined	516 (49.1%)	336 (32.0%)	162 (15.4%)	36 (3.4%)



**Figure 4-3 Bar chart demonstrating the number of magnetic resonance imaging pathology gradings assigned to each grade for anatomical structures of the equine foot.**

Abbreviations: DIP- distal interphalangeal, DSIL- distal sesamoidean impar ligament.

#### **4.5.2.2 Inter-observer agreement for the presence or absence of pathology**

Absolute inter-observer agreement for the dichotomous presence (pathology assessment grades 2 to 4) or absence (pathology assessment grade 1) of pathology varied between anatomic structures from poor to moderate based on assessment of Fleiss' kappa values (Table 4-3). This ranged from  $k=0.03$  for the distal interphalangeal joint to  $k=0.52$  for the navicular bone. The proportion of MRI studies with matched observer pathology assessments also varied but this mirrored the pattern of Fleiss' kappa values Table 4-3.

#### **4.5.2.3 Overall inter-observer agreement for pathology assessment grading**

The agreement analysis across all grades for pathology assessment of individual anatomical structures is presented in Table 4-4. This includes data from all observers across all acquisition systems. Absolute observer agreement, as determined by interpretation of Fleiss' kappa values, indicated poor agreement for the distal interphalangeal joint, collateral ligaments of the distal interphalangeal joint, the third phalanx and the distal sesamoidean impar ligament. Absolute observer agreement was interpreted as fair for the deep digital flexor tendon, navicular bone and the navicular bursa. The relative agreement (analysis that accounts for the relative ranking of grading) as determined by Kendall's coefficient of concordance was deemed to be low for the distal interphalangeal joint and moderate for the collateral ligaments of the distal interphalangeal joint, the third phalanx and the distal sesamoidean impar ligament. Relative agreement was moderate to high for the navicular bursa and the deep digital flexor tendon and was deemed to be high for the navicular bone.

**Table 4-3 Matched assessments and Fleiss' kappa values for observer agreement for the presence or absence of pathology on magnetic resonance imaging foot studies**

Abbreviations: DIP- distal interphalangeal, CI- confidence interval.

Structure	Percentage matched assessments (%)			Fleiss' kappa		
	Value	95% CI lower bound	95% CI upper bound	Kappa statistic	95% CI lower bound	95% CI upper bound
Deep digital flexor tendon	27	7.8	55	0.36	0.36	0.36
Navicular bone	47	21	73	0.52	0.52	0.52
Navicular bursa	27	7.8	55	0.37	0.37	0.37
DIP joint	0.0	0.0	18	0.03	0.03	0.04
Collateral ligaments of the DIP joint	6.7	0.17	32	0.16	0.16	0.16
Third phalanx	6.7	0.17	32	0.18	0.18	0.18
Distal sesamoidean impar ligament	20	4.3	48	0.22	0.22	0.23

**Table 4-4 Inter-observer agreement analysis for pathology assessment of the equine foot**

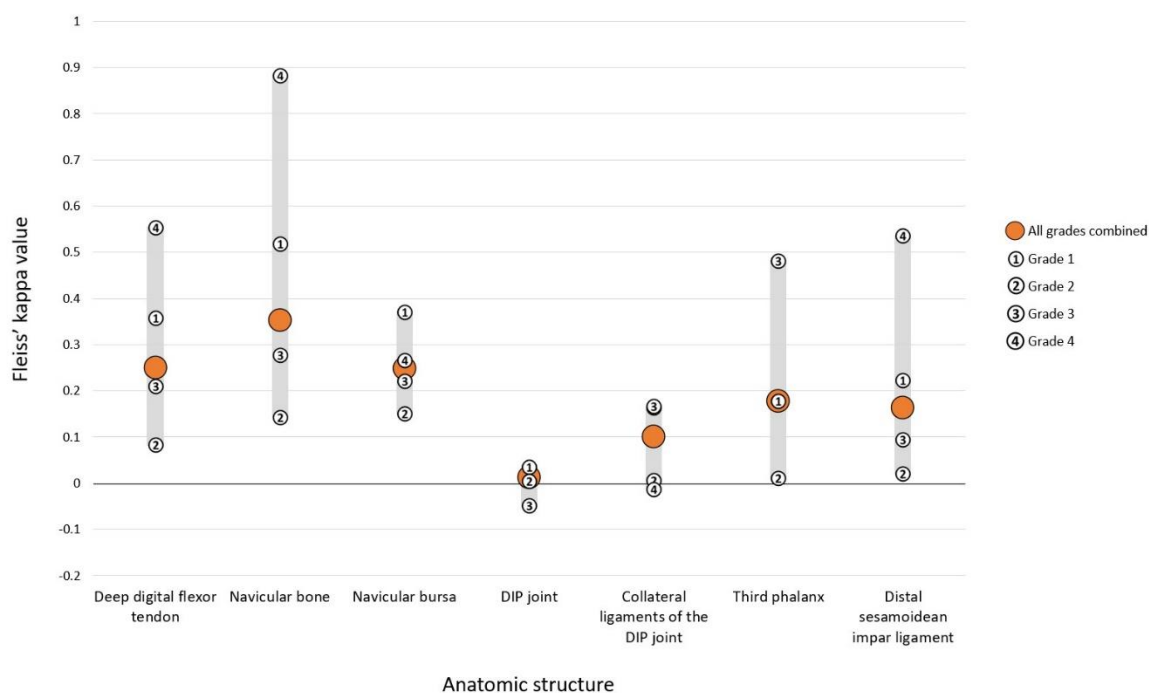
Abbreviations: DIP- distal interphalangeal, CI- confidence interval.

Assessment category	Percentage matched assessments		Fleiss' kappa			Kendall's coefficient of concordance
	Value (%)	95% CI	Kappa statistic	95% CI lower bound	95% CI upper bound	
Deep digital flexor tendon	0.0	0.0, 18	0.25	0.25	0.25	0.68
Navicular bone	6.7	0.17, 32	0.35	0.35	0.35	0.70
Navicular bursa	0.0	0.0, 18	0.25	0.25	0.25	0.67
DIP joint	0.0	0.0, 18	0.01	0.01	0.01	0.19
Collateral ligaments of the DIP joint	6.7	0.17, 32	0.10	0.10	0.10	0.33
Third phalanx	6.7	0.17, 32	0.18	0.18	0.18	0.44
Distal sesamoidean impar ligament	13	1.7, 40	0.16	0.16	0.16	0.45

#### 4.5.2.4 Inter-observer agreement for individual grades of pathology assessment grading

Fleiss' kappa values were also calculated for individual grades of pathology assessment for each individual structure (Table 4-5 and Figure 4-4). There was notable variation in absolute agreement between grades of the same anatomic structure and agreement for each grade differed between anatomic structures. In general agreement was greatest for grades at the extremes of the pathology spectrum, with grade 4 (severe) pathology accounting for the greatest agreement for the deep digital flexor tendon, navicular bone and the distal

sesamoidean impar ligament. No observers assigned a grade 4 assessment to the third phalanx, therefore the next highest grade (grade 3- moderate pathology) accounted for the greatest agreement for this structure. The opposite end of the pathology spectrum, grade 1 (no pathology), accounted for the greatest agreement for the navicular bursa and the distal interphalangeal joint.



**Figure 4-4 Linear dot plot displaying the Fleiss' kappa values for pathology assessment of seven anatomic structures of the equine foot.**  
Abbreviations: DIP- distal interphalangeal.

**Table 4-5 Fleiss' kappa values for individual grades of pathology assessment of the equine foot.**

Abbreviations: DIP- distal interphalangeal, CI- confidence interval.

Structure	Pathology assessment grade	Fleiss' kappa		
		Kappa statistic	95% CI lower bound	95% CI upper bound
Deep digital flexor tendon	Overall	0.25	0.25	0.25
	Grade 1	0.36	0.36	0.36
	Grade 2	0.08	0.08	0.08
	Grade 3	0.21	0.21	0.21
	Grade 4	0.55	0.55	0.56
Navicular bone	Overall	0.35	0.35	0.35
	Grade 1	0.52	0.52	0.52
	Grade 2	0.14	0.14	0.15
	Grade 3	0.28	0.28	0.28
	Grade 4	0.88	0.88	0.88
Navicular bursa	Overall	0.25	0.25	0.25
	Grade 1	0.37	0.37	0.37
	Grade 2	0.15	0.15	0.15
	Grade 3	0.22	0.22	0.22
	Grade 4	0.27	0.26	0.27
DIP joint	Overall	0.01	0.01	0.01
	Grade 1	0.03	0.03	0.04
	Grade 2	0.01	0.00	0.01
	Grade 3	-0.05	-0.05	-0.05
	Grade 4	N/A	N/A	N/A
Collateral ligaments of the DIP joint	Overall	0.10	0.10	0.10
	Grade 1	0.16	0.16	0.16
	Grade 2	0.01	0.00	0.01
	Grade 3	0.17	0.16	0.17
	Grade 4	-0.01	-0.02	-0.01
Third phalanx	Overall	0.18	0.18	0.18
	Grade 1	0.18	0.18	0.18
	Grade 2	0.01	0.01	0.01
	Grade 3	0.48	0.48	0.48
	Grade 4	N/A	N/A	N/A
Distal sesamoidean impar ligament	Overall	0.16	0.16	0.16
	Grade 1	0.22	0.22	0.23
	Grade 2	0.02	0.02	0.02
	Grade 3	0.10	0.09	0.10
	Grade 4	0.54	0.53	0.54

### 4.5.3 Observer pathology assessment confidence

#### 4.5.3.1 Key descriptive analysis

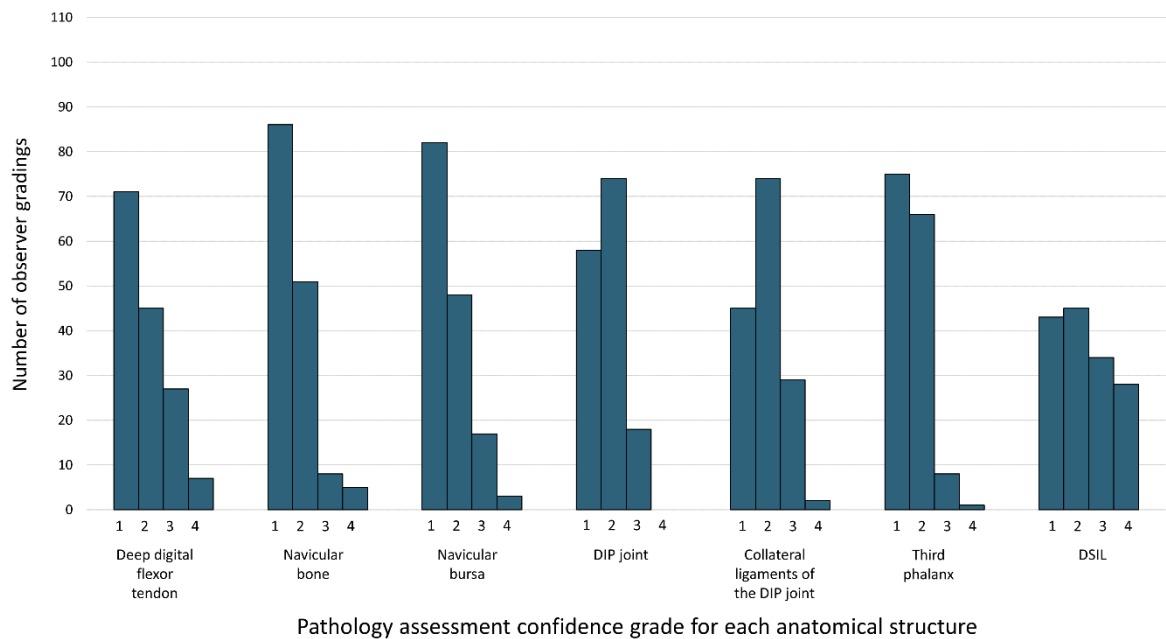
The distribution of pathology assessment confidence gradings differed between individual anatomical structures. Observers used pathology assessment confidence grade 1 (high confidence), grade 2 (moderate confidence) and grade 3 (limited confidence) for all individual anatomical structures. Grade 4 pathology assessment confidence (no confidence) was assigned by at least one observer for all structures except for the distal interphalangeal joint.

Median pathology assessment confidence gradings for individual MRI studies ranged from grade 1 to 2 for the navicular bone, navicular bursa, distal interphalangeal joint and third phalanx, from 1 to 2.5 for the deep digital flexor tendon and the collateral ligaments of the distal interphalangeal joint, and from grade 1 to 3 for the distal sesamoidean impar ligament. The count of pathology assessment confidence gradings assigned to each grade across all MRI studies for individual anatomical structures is presented in Table 4-6 and Figure 4-5. Bubble charts were a more useful method of visualisation than radar charts. Bubble charts for each anatomical structure, demonstrating the observer pathology assessment confidence grading are presented in Appendix 4.

**Table 4-6 Pathology assessment confidence gradings assigned to each grade for individual anatomical structures for 15 magnetic resonance imaging studies of the equine foot.**  
Abbreviations: DIP- distal interphalangeal.

Assessment Category	Count of pathology assessment confidence gradings assigned to each grade across all magnetic resonance imaging studies (%)			
	Grade 1- high confidence	Grade 2- moderate confidence	Grade 3- limited confidence	Grade 4- no confidence
Deep digital flexor tendon	71 (47.3%)	45 (30.0%)	27 (18.0%)	7 (4.7%)
Navicular bone	86 (57.3%)	51 (34.0%)	8 (5.3%)	5 (3.3%)
Navicular bursa	82 (54.7%)	48 (32.0%)	17 (11.3%)	3 (2.0%)
DIP joint	58 (38.7%)	74 (49.3%)	18 (12.0%)	0 (0.0%)
Collateral ligaments of the DIP joint	45 (30.0%)	74 (49.3%)	29 (19.3%)	2 (1.3%)
Third phalanx	75 (50.0%)	66 (44.0%)	8 (5.3%)	1 (0.7%)
Distal sesamoidean impar ligament	43 (28.7%)	45 (30.0%)	34 (22.7%)	28 (18.7%)
All anatomical structures combined	460 (43.8%)	403 (38.4%)	141 (13.4%)	46 (4.4%)





**Figure 4-5 Bar chart demonstrating the number of magnetic resonance imaging pathology assessment confidence gradings assigned to each grade for individual anatomical structures of the equine foot.**

Abbreviations: DIP- distal interphalangeal, DSIL- distal sesamoidean impar ligament.

#### 4.5.3.2 Overall inter-observer agreement for pathology assessment confidence grading

The agreement analysis across all grades for pathology assessment confidence for individual anatomic structures is presented in Table 4-7. This includes data from all observers across all acquisition systems. For all anatomical structures, the absolute observer agreement was interpreted as poor based on Fleiss' kappa values. The relative agreement (analysis that accounts for the relative ranking of grading) as determined by Kendall's coefficient of concordance, was deemed to be low for the navicular bone and the third phalanx, and moderate for the deep digital flexor tendon, navicular bursa, distal interphalangeal joint, collateral ligaments of the distal interphalangeal joint and the distal sesamoidean impar ligament.

#### 4.5.3.3 Inter-observer agreement for individual grades of pathology assessment confidence grading

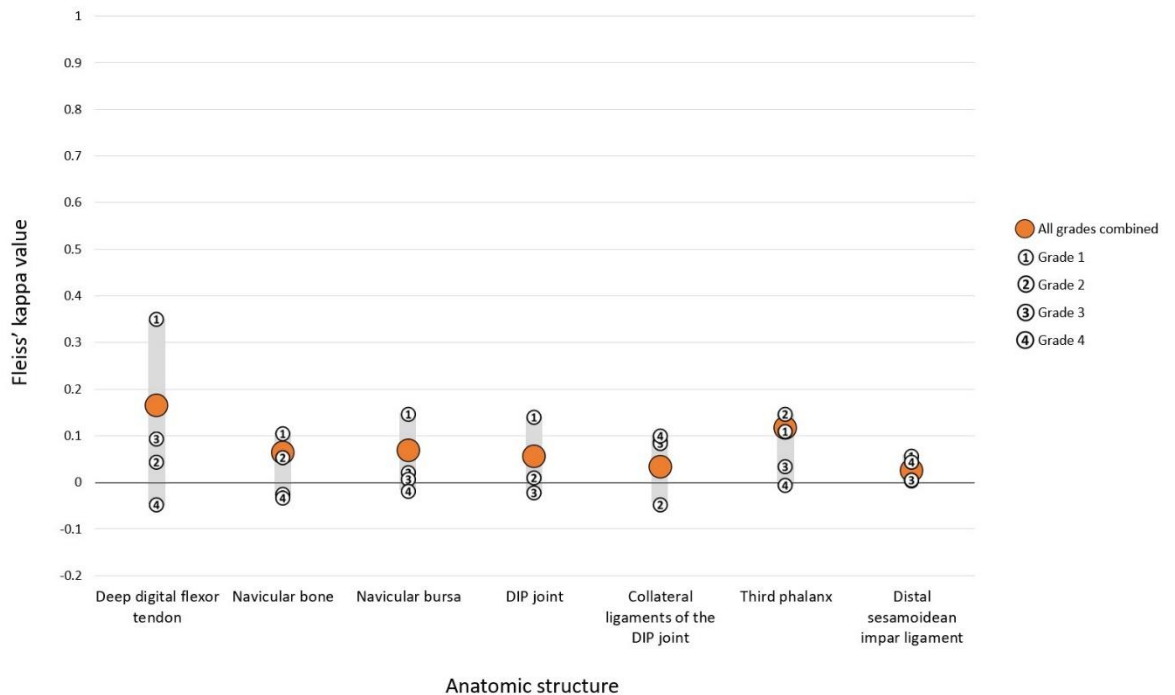
Fleiss' kappa values were also calculated for individual grades of pathology assessment confidence grading for each individual structure (Table 4-8 and Figure 4-6). There was some variation between anatomic structures in the

agreement for each grade of pathology assessment confidence. Agreement was greatest for grade 1 pathology assessment confidence (high confidence) for the deep digital flexor tendon, navicular bone, navicular bursa, distal interphalangeal joint and distal sesamoidean impar ligament.

**Table 4-7 Inter-observer agreement analysis for pathology assessment confidence grading of the equine foot.**

Abbreviations: DIP- distal interphalangeal, CI- confidence interval.

Assessment category	Percentage matched assessments		Fleiss' kappa			Kendall's coefficient of concordance
	Value (%)	95% CI	Kappa statistic	95% CI lower bound	95% CI upper bound	
Deep digital flexor tendon	6.7	0.17, 32	0.17	0.16	0.17	0.51
Navicular bone	6.7	0.17, 32	0.06	0.06	0.06	0.22
Navicular bursa	0.0	0.0, 18	0.07	0.07	0.07	0.33
DIP joint	0.0	0.0, 18	0.06	0.05	0.06	0.38
Collateral ligaments of the DIP joint	0.0	0.0, 18	0.03	0.03	0.03	0.35
Third phalanx	6.7	0.17, 32	0.12	0.11	0.12	0.22
Distal sesamoidean impar ligament	0.0	0.0, 18	0.03	0.02	0.03	0.32



**Figure 4-6 Linear dot plot displaying the Fleiss' kappa values for pathology assessment confidence grading for seven anatomic structures of the equine foot.**

Abbreviations: DIP- distal interphalangeal.

**Table 4-8 Fleiss' kappa values for individual grades of pathology assessment confidence grading of the equine foot.**

Abbreviations: DIP, distal interphalangeal joint.

Structure	Pathology assessment confidence grade	Fleiss' kappa		
		Kappa statistic	95% CI lower bound	95% CI upper bound
Deep digital flexor tendon	Overall	0.17	0.16	0.17
	Grade 1	0.35	0.35	0.35
	Grade 2	0.04	0.04	0.05
	Grade 3	0.09	0.09	0.09
	Grade 4	-0.05	-0.05	-0.05
Navicular bone	Overall	0.06	0.06	0.06
	Grade 1	0.10	0.10	0.11
	Grade 2	0.05	0.05	0.06
	Grade 3	-0.03	-0.03	-0.02
	Grade 4	-0.03	-0.04	-0.03
Navicular bursa	Overall	0.07	0.07	0.07
	Grade 1	0.15	0.14	0.15
	Grade 2	0.02	0.02	0.02
	Grade 3	0.00	0.00	0.01
	Grade 4	-0.02	-0.02	-0.02
DIP joint	Overall	0.06	0.05	0.06
	Grade 1	0.14	0.14	0.14
	Grade 2	0.01	0.01	0.01
	Grade 3	-0.02	-0.03	-0.02
	Grade 4	N/A	N/A	N/A
Collateral ligaments of the DIP joint	Overall	0.03	0.03	0.03
	Grade 1	0.09	0.08	0.09
	Grade 2	-0.05	-0.05	-0.05
	Grade 3	0.08	0.08	0.09
	Grade 4	0.10	0.10	0.10
Third phalanx	Overall	0.12	0.11	0.12
	Grade 1	0.11	0.11	0.11
	Grade 2	0.15	0.14	0.15
	Grade 3	0.03	0.03	0.03
	Grade 4	-0.01	-0.01	0.00
Distal sesamoidean impar ligament	Overall	0.03	0.02	0.03
	Grade 1	0.05	0.05	0.06
	Grade 2	0.00	0.00	0.00
	Grade 3	0.00	0.00	0.01
	Grade 4	0.04	0.04	0.05

## 4.6 Discussion

The results of this study indicate that there was a general trend of agreement between observers during interpretation of pathology on equine foot MRI studies. However, there can be notable variation in observer pathology assessment for anatomical structures at the individual case level. In the MRI studies presented in this study, agreement in pathology assessment varied between anatomical structures. In general, agreement was greatest at the extremes of pathology. Observer confidence in their assessment of pathology also varied widely.

There is currently little literature evaluating agreement between observers for evaluation of MRI equine foot studies. Researchers have assessed interobserver variation for factors such as ability to recognise limb positioning (Evrard *et al.*, 2019), rating of sequence diagnostic value (Berner *et al.*, 2020), definition of soft tissue cross-sectional area (Murray *et al.*, 2007) and assessment of image quality (Bolen, Audigié, *et al.*, 2010). A cadaver study has reported variation in observer perception of distal interphalangeal joint articular cartilage defects and demonstrated the difficulty of articular cartilage assessment during MRI (van Zadelhoff *et al.*, 2020). A study has demonstrated repeatability of positioning and image acquisition of the cadaver equine tarsus (Murray *et al.*, 2007). A study investigating the effects of diagnostic anaesthesia prior to MRI of the equine foot also reported inter-observer agreement for some assessment components (Black *et al.*, 2013). Agreement was fair for assessment of distal interphalangeal joint volume, moderate for navicular bursa volume and navicular bone signal, substantial for digital flexor tendon sheath volume and 'almost perfect' for the presence of needle tracts (Black *et al.*, 2013). However, no studies have focused on observer agreement of pathology assessment in MRI of the equine foot.

In the current study absolute agreement across all grades (as indicated by Fleiss' kappa) ranged from poor to fair for individual anatomical structures. However, relative agreement is a more useful form of assessment in the context of pathology assessment given that assessment is a complex and subjective process (Murray and Werpy, 2010). Relative agreement accounts for the relative order of grading, whereby a small difference in grades between observers (for example grades 1 and 2) indicates less disagreement than a bigger disparity (for example

grades 1 and 4) (Gwet, 2014). Absolute agreement, for example with Fleiss' kappa does not discriminate on the level of disagreement (Minitab, 2020). Relative agreement accounting for the ranking of grading (indicated by Kendall's coefficient of concordance) was low for the distal interphalangeal joint and moderate for the collateral ligaments of the distal interphalangeal joint, the third phalanx and the distal sesamoidean impar ligament. The navicular bursa, deep digital flexor tendon and navicular bone had the highest relative agreement. There is inherently some variation in how easily different anatomical structures are assessed during evaluation of MR images but other factors like susceptibility to artefacts can also influence this (Murray and Werpy, 2010). Studies have demonstrated good agreement between MRI and histologic findings for pathology of the deep digital flexor tendon, navicular bursa, collateral sesamoidean ligament and medulla of the navicular bone (Murray, Blunden, *et al.*, 2006). Lower agreement between MRI and histology was noted for the margins of the navicular bone and the distal sesamoidean impar ligament (Murray, Blunden, *et al.*, 2006). Disparities in how well MRI reflects true pathology may be a factor which contributes to variations in detection of pathology and ultimately observer agreement for different anatomical structures of the foot.

Agreement was generally greatest at the extremes of the pathology spectrum, as indicated by the Fleiss' kappa values for individual grades. This matches intuitive expectation that studies that have severe pathology or no pathology are more readily agreed upon by observers. For example, there was very good absolute agreement between observers on what constituted grade 4 (severe) pathology of the navicular bone. However, differentiation between more borderline decisions, for example between grade 2 (mild) and grade 3 (moderate) was more challenging, which is common in a clinical scenario (Nelson and Pepe, 2000).

This situation reflects one of the challenges of agreement analysis and demonstrates how analysis of this type is complex and influenced by multiple factors including the observers, the assessment subjects (MRI studies in this case) and the grading scale used for assessment (Watson and Petrie, 2010; Gwet, 2014). Kappa values are commonly used in agreement analysis in the equine

veterinary literature (Mejdell *et al.*, 2010; Menzies-Gow *et al.*, 2010; McLellan and Plevin, 2019; Berner *et al.*, 2020; Dyson *et al.*, 2020). Interpretation of kappa values often relies on previously published guidelines (Altman, 1991; Bowers, 2019). These are beneficial for interpretation but are inherently arbitrary and may not apply adequately in all clinical studies. A number of kappa statistics are available and suit different purposes (Gwet, 2014). These are useful indicators of inter-observer agreement but do have important limitations and can be susceptible to paradoxes where the agreement indicated by a kappa value differs to what would be expected based on apparent agreement (Shankar and Bangdiwala, 2014). In situations where the distribution of assessments across an assessment scale is asymmetric the expected chance agreement can be high, resulting in a surprisingly low kappa statistic (Gwet, 2014). Therefore, it is important that agreement analysis output is presented alongside the primary data to assess the distribution of the data, so that these can be considered together. Figure 4-3 and Table 4-2 give a useful overview of the pathology assessment data and demonstrate a general trend toward lower pathology grades for all structures in the current study. The distribution pattern of pathology grades was relatively similar for the deep digital flexor tendon, navicular bone and navicular bursa. The distal interphalangeal joint, collateral ligaments of the distal interphalangeal joint, third phalanx and distal sesamoidean impar ligament also had relatively similar pathology grading distributions. The method of randomised MRI study selection from different MRI systems aimed to replicate potential interpretation situations encountered in clinical practice from a broad population. However, whilst this reflects the distribution of pathology from a clinical population it does result in some differences in the relative severities of pathology between structures. This difference in distribution could influence the agreement analysis outcomes, for example compared to those that might be generated from a formulated group of cases with even distribution of pathology across all grades.

These points do not detract from the value of agreement (particularly kappa) statistics but highlight the importance of combining these with other methods. Graphical methods of assessment, such as those presented in Appendix 4 can be extremely valuable in demonstrating the pattern of agreement. The bubble charts give a clear visualisation of the distribution of gradings across the

pathology spectrum that can easily be interpreted in relation to clinical situations. Observer pathology gradings for individual MRI studies centred around the median pathology grade with tapering counts toward the extremes of the grading scale. Overall, this is an encouraging finding which indicates that there is a trend of observer agreement with most pathology assessments falling within 1 grade of one another. Nonetheless, for many anatomical structures there are MRI studies where observers assigned grades at opposing ends of the pathology spectrum. This is clearly a crucial finding for clinical practice and indicates that at the level of the individual MRI study, there can be important variation in the pathology perceived, even by experienced observers. This finding raises questions about whether disparities in the foot pathology observers report are due to differences in the perception of lesions (i.e. do observers see the same features of the image) or whether this reflects a difference in subsequent decision making after lesion identification (i.e. observers see the same lesion but make a different decision on its severity or significance). There is little information to characterise observer recognition of pathology and subsequent decision making in equine MRI. Gaining further insight into these factors will be key steps in further investigation of equine MR image interpretation and observer agreement.

The current study used a 1-4 grading scale with brief verbal descriptors for pathology assessment, broadly comparable to that used in a study comparing MRI with histologic findings of foot pathology (Murray, Blunden, *et al.*, 2006). The aim of this scale was to balance the guidance provided to observers in how to quantify pathology, without imposing overly specific classification descriptors that remove decision making (and then don't reflect the clinical scenario). Use of a grading scale of this type inherently requires observers to be proficient at equine MR image interpretation, relying on their significant previous experience and knowledge of the spectrum of pathology for different anatomical structures. During clinical MRI reporting most observers document their image evaluation findings and then produce a conclusion/ summary providing an impression of the clinical significance of the findings (Murray and Werpy, 2010; Schramme and Segard-Weisse, 2020). In the current study observers did not receive a clinical history to interpret in conjunction with the MR images. This is an important feature of clinical MRI reporting, but this study was primarily focused on the

step of recognising foot pathology (rather than complete clinical interpretation). Therefore, although the approach of the current study using pathology grading systems does not exactly replicate clinical reporting, it encouraged observers to follow their typical image evaluation process and then document their impressions in a straightforward manner. The process of integrating the clinical context, history and MRI findings into a clinical impression is another challenging step that may warrant further dedicated investigation in the future (Murray and Werpy, 2010).

There was also notable variation in the confidence of observers in their evaluation of pathology. The distal sesamoidean impar ligament appeared to be particularly susceptible to variation in observer pathology assessment confidence. Assessment confidence may be partly influenced by the spectrum of pathology for individual structures in the MRI studies presented in the platform and the inherent variation in how easily different structures are evaluated. The distal sesamoidean impar ligament is recognised as a challenging structure to assess on MR images (Dyson *et al.*, 2010). A study investigating methods to improve visualisation of the distal sesamoidean impar ligament during MRI demonstrated slight to fair interobserver agreement when selecting sagittal sequences that provided optimal visualisation (Berner *et al.*, 2020). Agreement was greater following saline arthrography of the distal interphalangeal joint (Berner *et al.*, 2020). In a clinical context, evaluation of the distal sesamoidean impar ligament also relies on assessment of adjacent structures including the distal aspect of the navicular bone and the solar surface of the third phalanx (Murray *et al.*, 2009; Dyson *et al.*, 2010). Therefore, it is predictable that observers had relatively lower confidence in their pathology assessments for this structure. Evaluation of individual observer confidence gradings demonstrated that there was some intra-observer variation in confidence between MRI studies, which may be due to differences in the severity in pathology or image quality. However, an important source of the variation in confidence gradings was due to inter-observer differences, with some observers being consistently confident in their pathology assessment and others expressing a tendency to be less confident in their assessments. It is important to acknowledge that a less confident assessment does not necessarily equate to a poorer quality interpretation. In a clinical context reduced confidence in assessment may



warrant the acquisition of repeated or additional sequences (Murray and Werpy, 2010). All the observers involved in the current study had significant experience in MR image interpretation but their backgrounds, roles, qualifications, locations (which may impact their typical case population) and experience with different MRI systems did vary. The assessment confidence findings give further insight into how differences in observer perception in pathology may arise. The underlying variation in observer confidence may be due to multiple factors including differences in observer experience, contrasting approaches to MR image assessment, and disparities in observer recognition and perceptions of pathology. Understanding this process of decision making is important to promote consistency in MRI interpretation, optimising patient outcomes and encourage translation of MRI clinical research amongst the profession.

In conclusion, there was a general trend of agreement between observers for pathology assessment of anatomical structures of the equine foot. However, there can be notable variation in pathology assessment at the level of individual MRI studies, even when interpretation is performed by experienced observers. Observer confidence in their assessment of pathology also varied widely. These findings highlight the need to understand the processes of lesion identification and subsequent diagnostic decision making by those interpreting images in MRI of the equine foot.

## Chapter 5 General discussion

The findings of these studies demonstrate that most equine foot MRI studies acquired from live clinical patients were deemed to be of diagnostic quality, regardless of the acquisition system. However, field strength appears to have more influence than general anaesthesia on perceived image quality for MRI studies of the equine foot. Furthermore, whilst there is a general trend of agreement in pathology assessment between experienced observers, there can be notable variation in perceived pathology at the level of individual patients.

There were no significant differences in perceived image quality between low-field MRI studies of the equine foot acquired with the patient standing or under general anaesthesia. Though motion is commonly reported as a reason for reduced image quality in standing equine MRI (Porter and Werpy, 2014), the results presented in Chapter 2 indicated that this is effectively controlled in a clinical context to result in diagnostic quality MRI studies. The equine foot is deemed to be less susceptible to motion understanding sedation than other regions of the limb (McKnight *et al.*, 2004) so the findings of the current study cannot be assumed to apply to other regions of the limb. Conversely, Chapter 3 demonstrated that greater field strength resulted in an increase in perceived image quality of entire MRI studies and individual anatomical structures of the equine foot. This finding concurs with those of previous cadaver studies, indicating that whilst low-field MRI can provide significant diagnostic information, high-field MRI can offer improved visualisation, particularly of small anatomical structures or small lesions (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). Therefore, in clinical situations where image quality is prioritised then high-field MRI of the equine foot is preferable, especially when small structures or subtle pathologic changes are implicated. Whilst the current study did not demonstrate a significant influence of anaesthesia on image quality, previous literature focusing on evaluation of the distal interphalangeal joint has indicated that assessment of the articular cartilage is likely to be improved in the non-weightbearing limb (Olive, 2010; Evrard *et al.*, 2019; van Zadelhoff *et al.*, 2020). Therefore, when a distal interphalangeal joint articular lesion is suspected non-weight-bearing MRI (i.e. performed under general anaesthesia)

may be beneficial, particularly if high-field MRI is available (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010).

The reasons reported for reduction in perceived image quality appeared to vary between MRI systems. Many reasons for reduced image quality were attributable to system operator or institutional factors. For example, positioning, slice alignment, requirement for sequence repetition or desire to include an additional sequence were all common reasons for reduced image quality. Similarly, preventable artefacts that could reduce image quality may also be influenced by the system operator in some circumstances. Clear communication between observers interpreting the resultant images and those operating the MRI system is vital to ensure that there is harmonious understanding of what is required for each patient's MRI study. Given that comments were provided at the discretion of observers, inferential analysis of these data was not pursued in the current studies. However, further investigation or audit of clinical images would appear to be valuable to help guide training of operators for their specific system and to highlight any areas where the interaction between operators and image interpreters could be improved. A report of suboptimal fat suppression of STIR images was recorded at least once for all systems but appears to be more common on low-field systems. Previous cadaver studies have reported the influence of temperature on STIR fat suppression (Smith *et al.*, 2008; Bolen, Audigié, *et al.*, 2010; Bolen, Haye, *et al.*, 2010). However, the current studies indicate that this is a factor that also reduces image quality in a clinical context in live patients. In some systems operators can optimise the appearance of STIR images during study acquisition, but in all situations aiming to position the region of interest at the isocentre of the magnetic field will provide the most reliable fat suppression (Werpy, 2010).

The findings of the current studies are focused upon the equine foot and cannot be directly extrapolated to other regions of the limb, where factors influencing image quality may differ in importance. For example, motion is deemed to be more problematic for imaging of the proximal limb when compared to the foot, particularly in the standing horse (McKnight *et al.*, 2004). Therefore, further research could utilise a similar methodology to that reported in the current studies to investigate the relative importance of factors influencing clinical

image quality in live patients for other commonly imaged anatomical regions such as the fetlock, proximal metatarsus and carpus.

Image quality is only one of the factors that contribute to the decision making on the type of MRI acquisition system used to image a patient in a clinical context. Other important factors may include the desire to avoid general anaesthesia of the patient, MRI system availability, acquisition time, proposed region(s) to be imaged, suspected pathology, potential treatment options, patient compliance, clinician preference, interpretation availability and cost of imaging. The risks of general anaesthesia in equine patients is a widely investigated topic, having implications beyond MRI (Johnston *et al.*, 2002; Bidwell, Bramlage and Rood, 2007; Dugdale, Obhrai and Cripps, 2016; Laurenza, Ansart and Portier, 2019; Gozalo-Marcilla *et al.*, 2020). Avoidance of general anaesthesia has been a common theme in equine MRI and a strong driver for the success of the low-field standing MRI system (Bladon, 2014; Porter and Werpy, 2014; Hallmarq Veterinary Imaging, 2019). This benefit is particularly relevant to cases where recovery from anaesthesia could exacerbate a pre-existing lesion. However, studies have demonstrated that the risk for peri-anaesthetic morbidity and mortality for equine MRI under general anaesthesia is very low (Andersen *et al.*, 2006; Franci, Leece and Brearley, 2006). This is particularly true for systems designed for MRI under general anaesthesia and at institutions familiar with the procedure (Murray, Leece and Judy, 2010). There is evidence that MRI of proximal regions, where positioning the limb within the bore is difficult, may have an increased risk of post-anaesthetic myopathy/ neuropathy syndrome (Moreno *et al.*, 2020). However, at present a significant proportion of equine MRI studies in a clinical setting focus upon the distal limb, which is typically positioned within the system in a straightforward manner (Murray, Leece and Judy, 2010; Barrett *et al.*, 2017). At present the literature investigating morbidity and mortality of equine MRI under general anaesthesia is primarily based upon high-field imaging (Andersen *et al.*, 2006; Franci, Leece and Brearley, 2006; Moreno *et al.*, 2020). There are additional considerations relating to the MRI system (such as bore size and shape), patient positioning and anaesthesia that could increase the likelihood of adverse events during high-field MRI (Murray, Leece and Judy, 2010). Therefore, it is possible that the risk of general anaesthesia for low-field MRI is lower than that of high-field MRI.

Similarly, MRI of the standing equine patient may also have associated morbidity and mortality. This may occur due to the gastrointestinal effects of the alpha-2-agonists and opioid agents commonly used for sedation, that can predispose to post-procedure colic (Koenig and Cote, 2006; Thibault *et al.*, 2019). As with other standing procedures, patient compliance is also important to avoid injury. Unpredictable reactions or recalcitrant behaviour of the sedated patient when the limb is within the relatively narrow confines of the bore can be problematic. At present there is little evidence reporting the morbidity and mortality of low-field equine MRI under standing sedation or general anaesthesia.

A key aim of improvements in diagnostic imaging modalities is to optimise diagnostic accuracy. Whilst pathology identification is closely associated with image quality, this relationship is not necessarily proportionate. The observer also has a fundamental contribution to the ultimate diagnostic accuracy of a diagnostic imaging intervention (Williams and Drew, 2019). The findings presented in Chapter 4 demonstrate a general trend of agreement between experienced observers for identification of pathology in MRI studies of the equine foot. However, there were instances of notably variable pathology assessment at the level of individual cases, which could be important in a clinical context. Similarly, there was prominent variation in pathology assessment confidence between experienced observers. This may originate from differences in experience or disparities in decision making criteria. At present there is very limited information to characterise patterns of equine MR image assessment, pathology identification and subsequent decision making. This is a complex area that has been explored more thoroughly in human medical imaging (Krupinski, 2010; Kagadis *et al.*, 2013; Tourassi *et al.*, 2013; Brunyé *et al.*, 2019; Williams and Drew, 2019). Further investigations could focus on components of this process such as approaches taken by individual observers to assess equine MRI studies and how observers evaluate lesions and determine their clinical significance. This work could assist in developing consistency in clinical image interpretation and support training of those interpreting images. This consistency is likely to become increasingly important with developments in teleradiology (Johnson, 2011), where international interpreters may assess images from diverse equine populations with which they may be unfamiliar.

There were practical limitations to consider in the design of the current studies that warrant discussion. The 10 experienced observers involved were requested to interpret a total of 15 MRI equine foot studies using the image assessment platform. Though observers were offered a stipend, this represents a significant time commitment to the project. In the context of MR image quality, the effect size of clinical significance is deemed to be relatively large, for there to be any influence on pathology identification and diagnostic confidence. Previous studies with similar methodologies used 22-30 interpretations per MRI system (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010). Therefore, the 50 interpretations per MRI system of the current studies was deemed to be adequate to identify any clinically relevant differences in image quality, without imposing onerous demands on the observers. In addition, previous studies involving equine MR image interpretation have utilised relatively few experienced observers (typically 1-3) (Murray *et al.*, 2009; Bolen, Audigié, *et al.*, 2010; Evrard *et al.*, 2019; Berner *et al.*, 2020; van Zadelhoff *et al.*, 2020). This results in individual observers having an important influence on reported image perception and overall study outcomes. Routine quality assurance methods should result in relatively little inter-system variation in image quality, especially for MRI studies deemed to be of diagnostic quality at the time of acquisition (McRobbie *et al.*, 2017e). However, there is very limited evidence to document the influence of the observer (and their inherent biases) on outcomes. Consequently, the approach taken in the current studies was to utilise a relatively broad panel of experienced observers (10) to assess a more moderate number of MRI studies. This approach also provided a larger number of observers to be involved in agreement analysis. As expected, most observers were more familiar with the interpretation of images from low-field standing MRI systems. This is deemed to be a general reflection of equine MR image interpretation, where the standing low-field unit predominates in a clinical context (Bladon, 2014; Hallmarq Veterinary Imaging, 2019). This familiarity with low-field and weight-bearing images may have influenced the perception of image quality. This familiarity is not necessarily problematic if this reflects the realistic preferences of the population of observers interpreting equine MR images in a clinical context.

The study presented in Chapter 4 documents agreement of observer pathology assessment for MRI of the foot. However, there is still much for us to learn about

the interpretation of equine MR images. MRI characteristics of anatomical structures and pathology, and how to interpret them are well defined (Murray and Werpy, 2010; Winter, 2012). However, the practical approach of how observers interpret (or “read”) an equine MRI study is not well described (Schramme and Segard-Weisse, 2020). It is likely that individual observers will have preferred methods to multi-sequence display and routines of sequence assessment when interpreting MRI studies. Prospective studies recording observer approaches to MRI interpretation for different anatomical regions could provide useful evidence to characterise these techniques. Indeed, identifying key steps of interpretation could be valuable to improve diagnostic outcomes and act as a guide to those learning to interpret equine MRI studies. More detailed assessment of interpretation behaviour, for example using eye-tracking methods, could improve our understanding of how observers (particularly those with significant experience) survey MR images (Tourassi *et al.*, 2013; Brunyé *et al.*, 2019; Wu and Wolfe, 2019). In addition, eye-tracking may also yield further information about how differences in pathology identification and interpretation might arise. However, studies using eye-tracking methods are challenging, particularly in relation to multiplanar imaging where interpretation has a volumetric component (Williams and Drew, 2019).

In conclusion, the studies presented in this thesis demonstrate that most studies of the equine foot acquired in a clinical context are deemed to be of diagnostic quality. Field-strength has a significant impact on perceived image quality of equine foot MRI studies, but general anaesthesia of the patient has limited influence. In situations where image quality is prioritised over other factors, then the high-field MRI of the foot under general anaesthesia would be the preferred method of image acquisition. In addition, observer related factors may also have an important influence on image interpretation and diagnostic decision making. There was a trend of agreement for pathology identification between experienced observers during MRI of the equine foot. However, there were variations in perceptions of pathology at the level of individual MRI foot studies. These differences in perception could have important consequences for diagnosis and subsequent case management in a clinical context.

## Appendices

### Appendix 1: Representative anatomical and magnetic resonance images of the equine foot

#### Deep digital flexor tendon

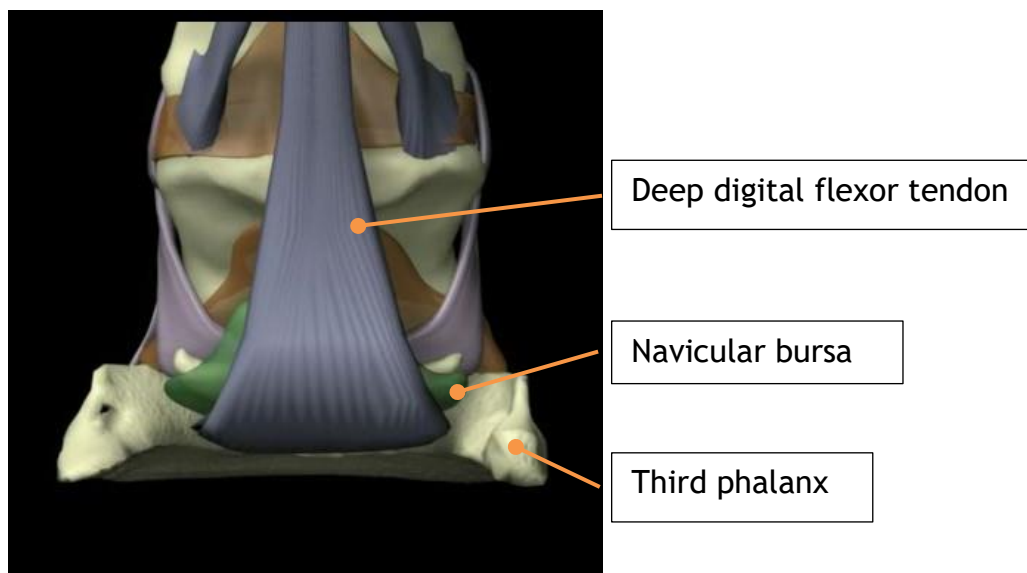


Figure A1-1 Anatomic diagram of the palmar aspect of the distal limb to depict the deep digital flexor tendon. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

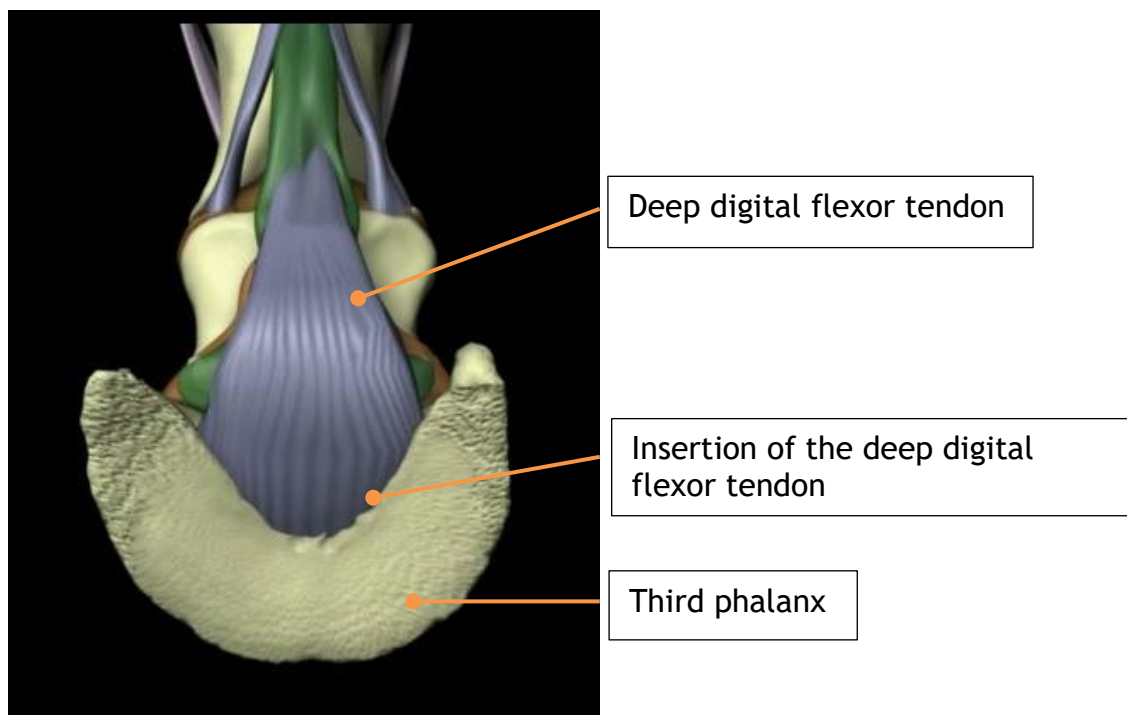


Figure A1-2 Anatomic diagram of the palmar aspect of the insertion of the deep digital flexor tendon. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.



## Navicular bone, associated ligaments and the navicular bursa

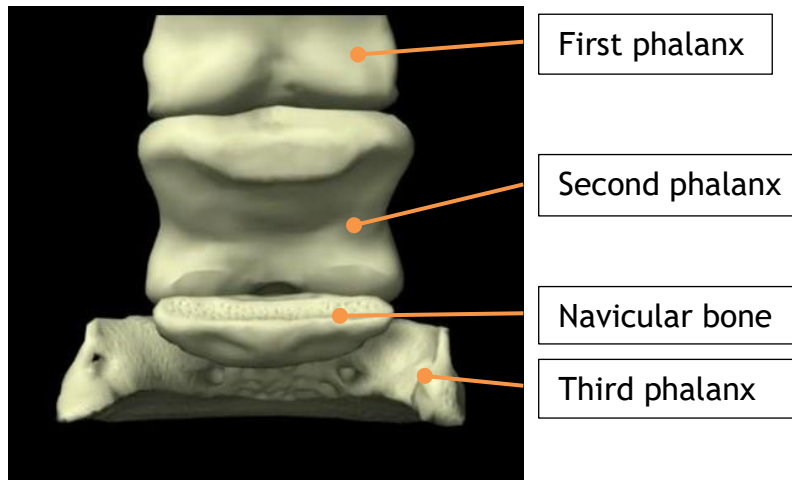


Figure A1-3 Anatomic diagram of the palmar aspect of the phalanges and navicular bone. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

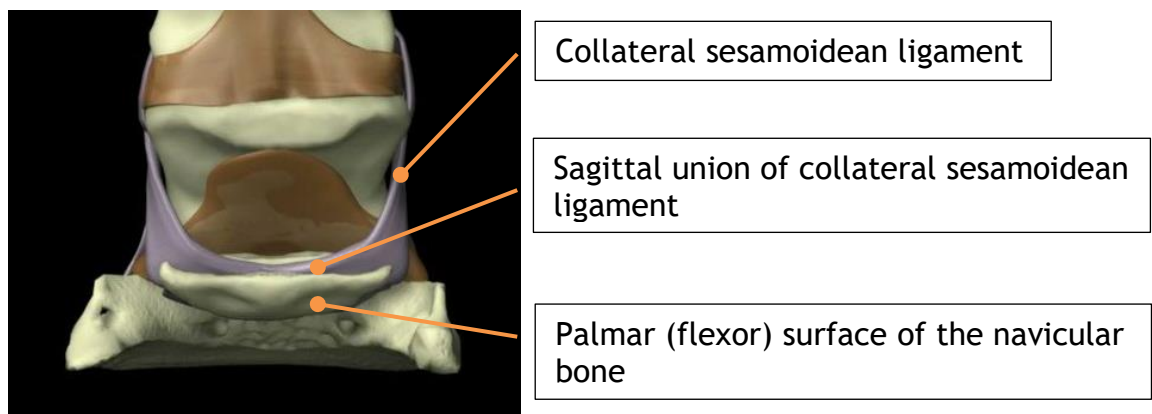


Figure A1-4 Anatomic diagram of the palmar aspect of the navicular bone and associated ligaments. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

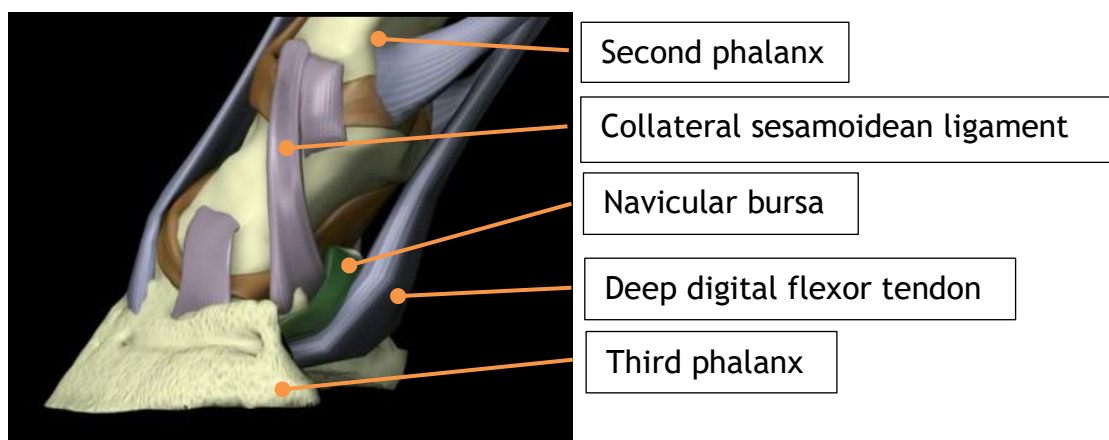


Figure A1-5 Anatomic diagram of the navicular bursal region from the palmarolateral oblique aspect. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

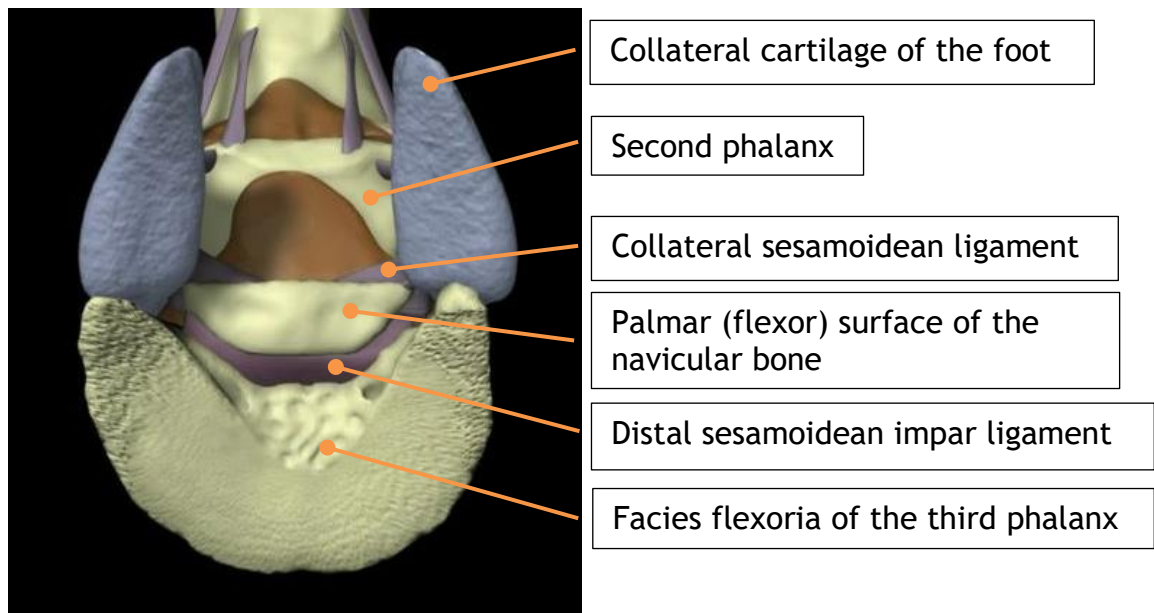


Figure A1-6 Anatomic diagram of the distal aspect of the navicular bone and associated ligaments. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

## Distal interphalangeal joint and associated soft tissues

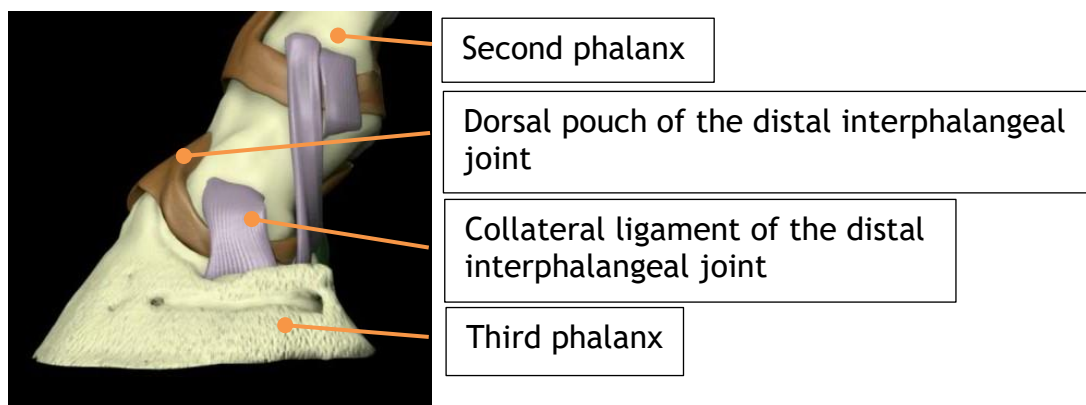


Figure A1-7 Anatomic diagram of the lateral aspect of the distal interphalangeal joint region. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

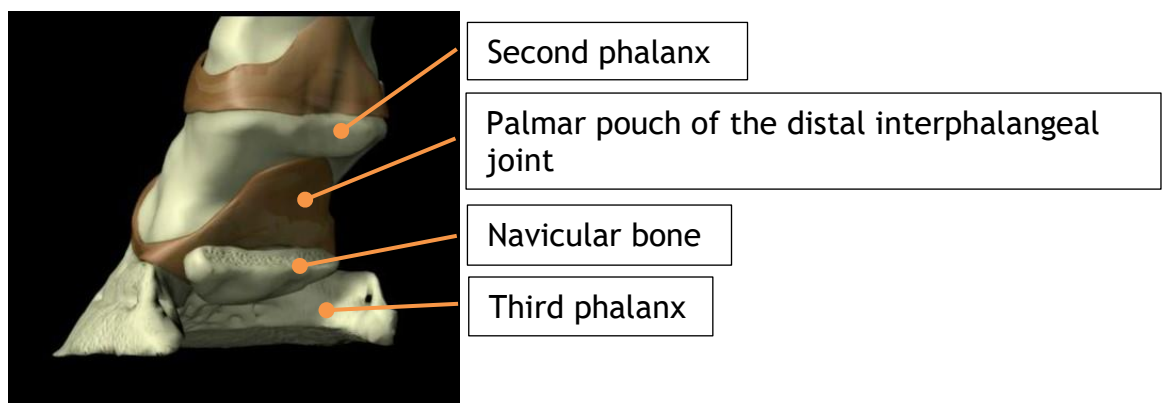


Figure A1-8 Anatomic diagram of the distal interphalangeal joint region from the palmarolateral oblique aspect. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

## Phalanges and associated soft tissues

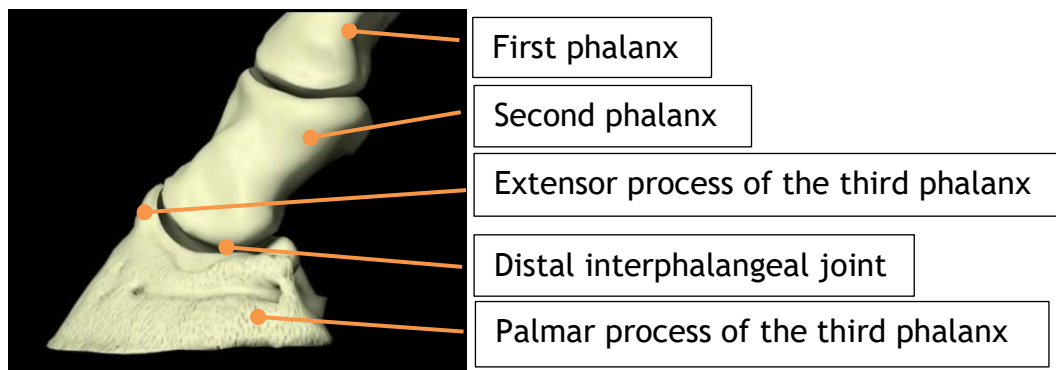


Figure A1-9 Anatomic diagram of the phalanges from the lateral aspect. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

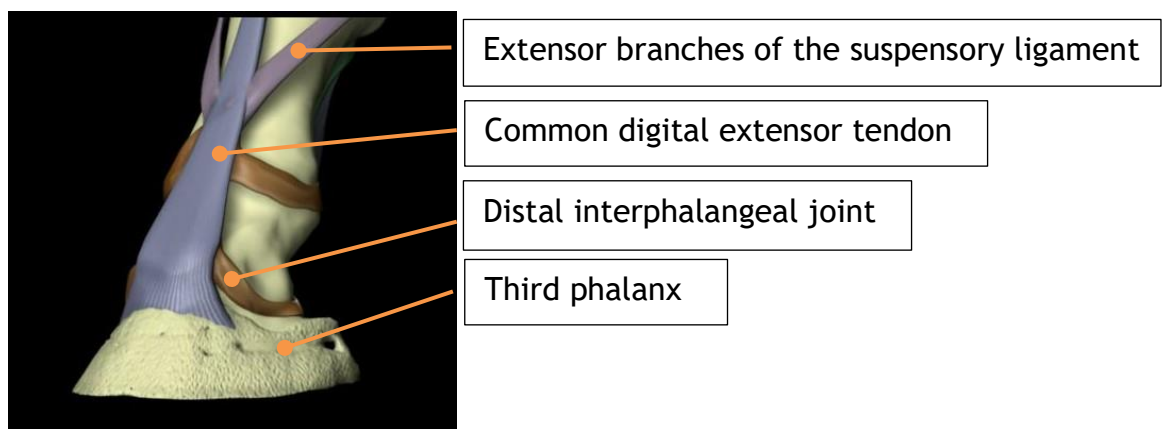


Figure A1-10 Anatomic diagram of the common digital extensor tendon from the dorsolateral aspect. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

## Collateral cartilages of the foot and the digital cushion

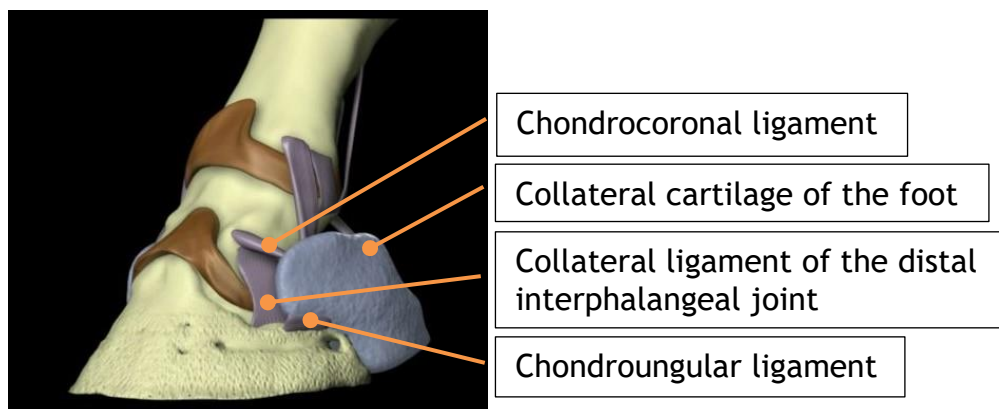
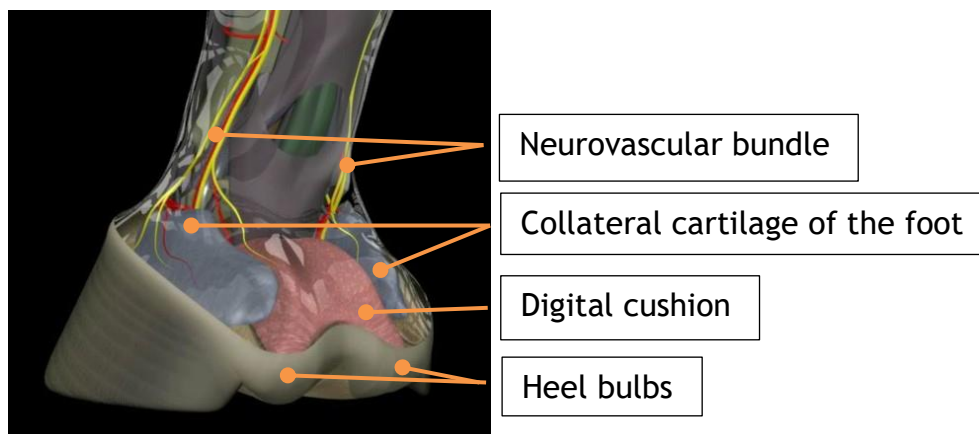
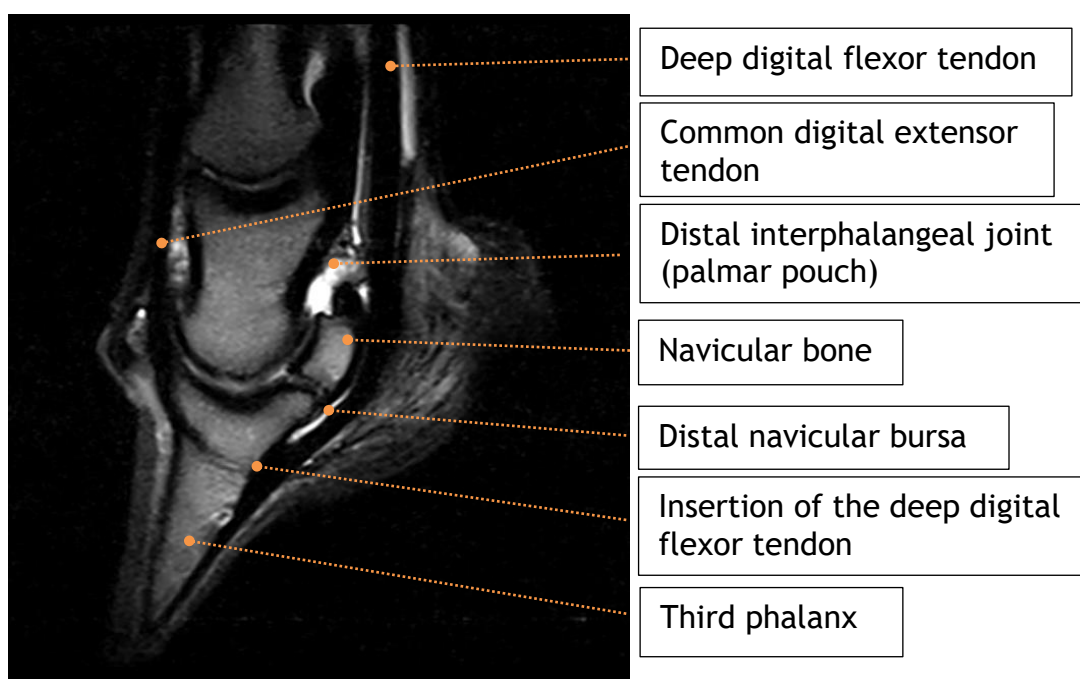


Figure A1-11 Anatomic diagram of the collateral cartilages of the foot from the dorsolateral aspect. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.



**Figure A1-12** Anatomic diagram of the digital cushion and the collateral cartilages of the foot from the palmarolateral aspect. Permission to reproduce this image from Elements of the Equine Distal Limb has been granted by Science in 3D, Inc.

## Representative magnetic resonance images



**Figure A1-13** Low-field sagittal T2 weighted fast spin echo magnetic resonance image of a fore foot (slice just abaxial to mid-sagittal).

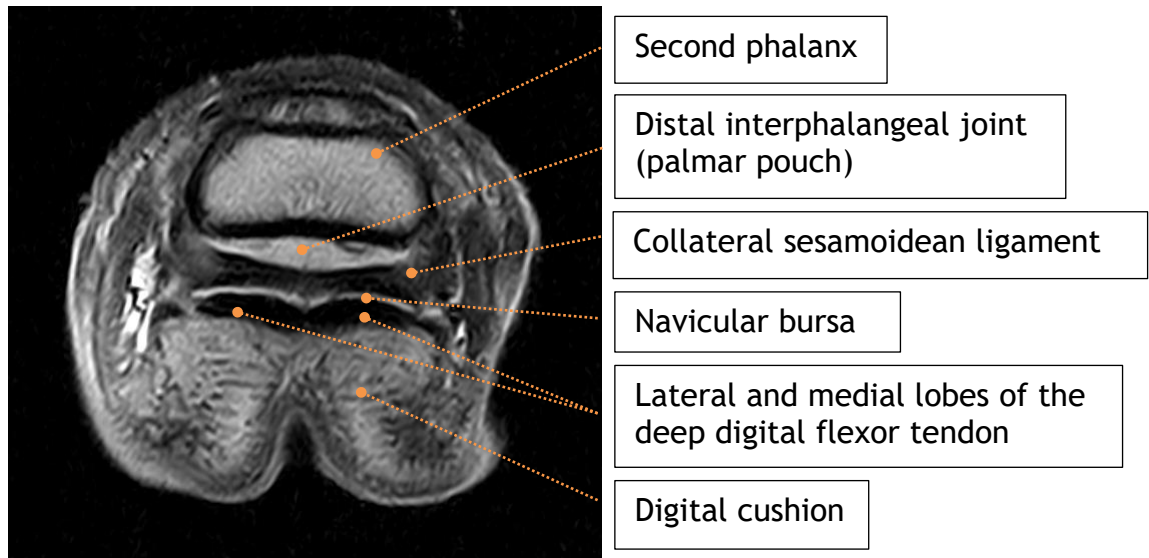


Figure A1-14 Low-field transverse proton density weighted turbo multi-echo magnetic resonance image of a fore foot (slice at the level of the proximal navicular bursa).

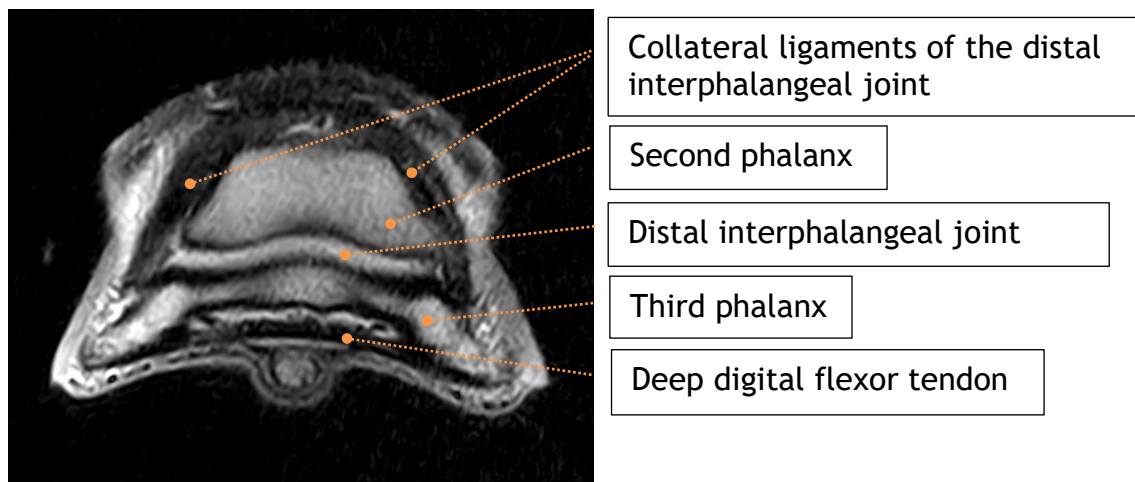


Figure A1-15 Low-field transverse (relative to the distal deep digital flexor tendon) T2 weighted turbo multi-echo magnetic resonance image of a fore foot (slice at the level of the collateral ligaments of the distal interphalangeal joint).

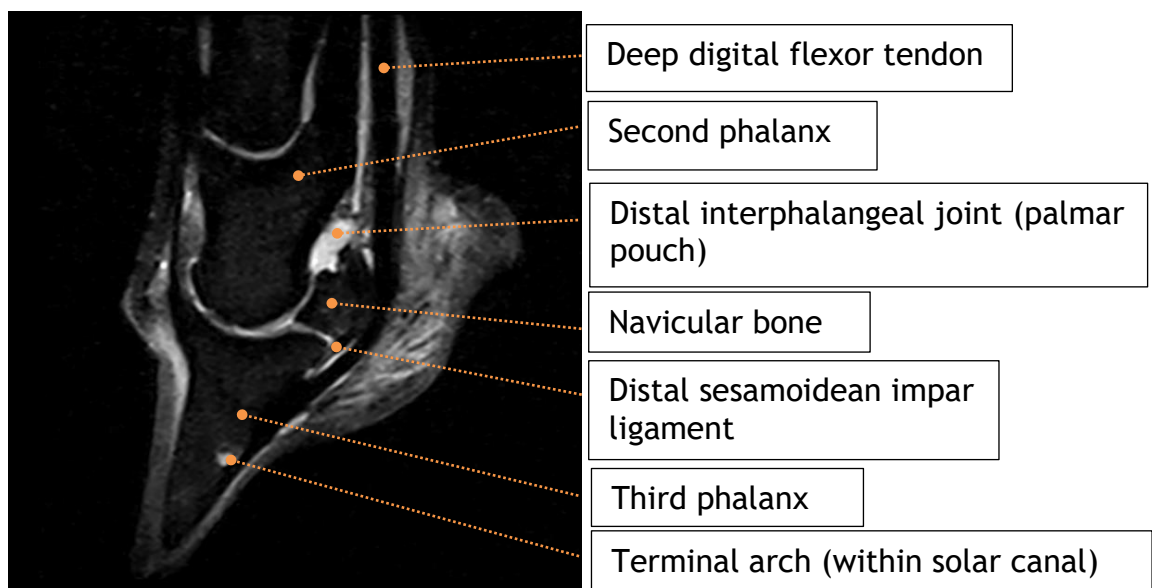
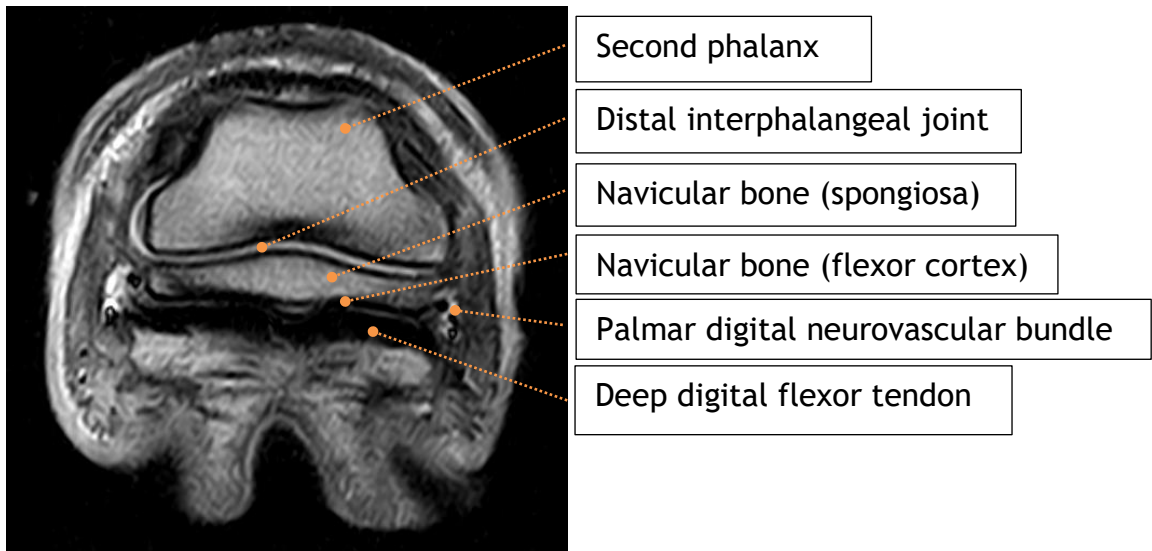
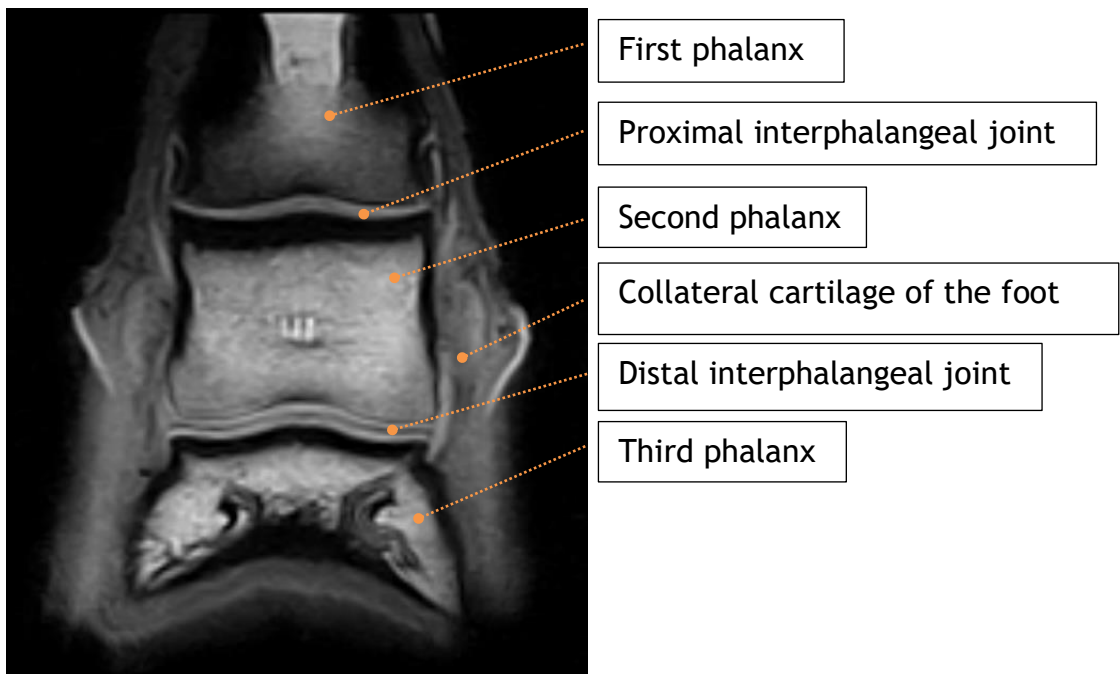


Figure A1-16 Low-field sagittal short tau inversion recovery magnetic resonance image of a fore foot (slice just abaxial to mid-sagittal).





**Figure A1-17** Low-field transverse proton density weighted turbo multi-echo magnetic resonance image of a fore foot (slice at the level of the proximal navicular bone).



**Figure A1-18** Low-field frontal T1 weighted magnetic resonance image of a fore foot (slice at the level of the centre of the bony column).

## **Appendix 2: Pulse sequence parameters used in magnetic resonance imaging studies included in image assessment**

### **Background**

This appendix contains the pulse sequence parameters used in clinical magnetic resonance imaging studies included in image assessment from three different magnetic resonance imaging acquisition systems.

The magnetic resonance imaging studies are ordered by magnetic resonance imaging acquisition system.

### **Terminology**

Please note that the terminology of the orientation has been kept consistent with that of the acquiring institution (i.e. some sequences described as transverse when aligned with the distal deep digital flexor tendon may closely resemble images described as dorsal in other sequences). Also note that 'Interslice spacing' refers to spacing from the centre-centre of adjacent slices.

## Low-field standing

### Case 5

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
PD SE	Transverse	24	1000	174 x 174	5	6	12	90	
STIR FSE	Sagittal	27	3803	175 x 175	5	6	15	90	95
STIR FSE	Sagittal	22	2811	185 x 185	5	6	18	90	95
STIR FSE	Transverse	27	3042	175 x 175	5	6	12	90	95
T1 3D HR	Dorsal	8	24	175 x 175	1.5	1.5	48	43	
T1 3D	Sagittal	7	24	175 x 175	3.8	3.8	26	43	
T1 3D	Transverse	7	24	175 x 175	3	3	26	43	
T2 FSE	Transverse	81	1848	175 x 175	5	6	12	90	
T2 FSE	Transverse	81	1500	175 x 175	5	6	9	90	
T2 FSE HR	Transverse	87	1980	175 x 175	3.5	4.2	12	90	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** PD- Proton density weighted, SE- Spin echo, STIR- Short tau inversion recovery, FSE- Fast spin echo, T1- T1 weighted, HR- High resolution, T2- T2 weighted.



## Case 6

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
PD SE	Transverse	24	1000	179 x 179	5	6	12	90	
STIR FSE	Sagittal	22	2543	180 x 180	5	6	15	90	135
STIR FSE	Sagittal	27	4410	180 x 180	5	6	15	90	135
STIR FSE	Transverse	27	3468	180 x 180	5	6	12	90	130
T1 3D HR	Dorsal	8	24	180 x 180	1.5	1.5	48	43	
T1 3D	Sagittal	7	24	180 x 180	3.6	3.6	26	43	
T1 3D	Transverse	7	24	180 x 180	3	3	26	43	
T2 FSE	Transverse	81	1848	180 x 180	5	6	12	90	
T2 FSE	Transverse	81	1848	180 x 180	5	6	12	90	
T2 FSE HR	Transverse	87	1980	180 x 180	4.6	5	12	90	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** PD- Proton density weighted, SE- Spin echo, STIR- Short tau inversion recovery, FSE- Fast spin echo, T1- T1 weighted, HR- High resolution, T2- T2 weighted.

## Case 10

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
STIR FSE	Sagittal	27	4464	175 x 175	5	6	16	90	120
PD SE	Transverse	24	1000	174 x 174	5	6	12	90	
STIR FSE	Sagittal	22	2811	175 x 175	5	6	18	90	95
STIR FSE	Sagittal	27	4464	175 x 175	5	6	16	90	120
STIR FSE	Transverse	27	3296	175 x 175	5	6	13	90	95
T1 3D HR	Dorsal	8	24	175 x 175	1.5	1.5	48	43	
T1 3D	Sagittal	7	24	175 x 175	4.1	4.1	26	43	
T1 3D	Transverse	7	24	175 x 175	3.4	3.4	26	43	
T2 FSE	Transverse	81	2002	175 x 175	5	6	13	90	
T2 FSE	Transverse	81	1848	175 x 175	5	6	12	90	
T2 FSE HR	Transverse	87	2145	175 x 175	3.5	4.2	13	90	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** STIR- Short tau inversion recovery, FSE- Fast spin echo, PD- Proton density weighted, SE- Spin echo, T1- T1 weighted, HR- High resolution, T2- T2 weighted.

## Case 12

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
PD SE	Transverse	24	1000	174 x 174	5	6	12	90	
STIR FSE	Sagittal	27	4743	175 x 175	5	6	17	90	120
STIR FSE	Transverse	27	3627	175 x 175	5	6	13	90	120
T1 3D HR	Dorsal	8	24	175 x 175	1.5	1.5	48	43	
T1 3D	Sagittal	7	24	175 x 175	3.9	3.9	26	43	
T1 3D	Transverse	7	24	175 x 175	3.2	3.2	26	43	
T2 FSE	Transverse	81	1848	175 x 175	5	6	12	90	
T2 FSE	Transverse	81	2002	175 x 175	5	6	13	90	
T2 FSE HR	Transverse	87	2145	175 x 175	3.5	4.2	13	90	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** PD- Proton density weighted, SE- Spin echo, STIR- Short tau inversion recovery, FSE- Fast spin echo, T1- T1 weighted, HR- High resolution, T2- T2 weighted.

## Case 14

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
PD SE	Transverse	24	1000	174 x 174	5	6	12	90	
STIR FSE	Sagittal	27	2282	175 x 175	5	6	9	90	95
STIR FSE	Sagittal	27	2282	175 x 175	5	6	9	90	95
STIR FSE	Transverse	27	3042	175 x 175	5	6	12	90	95
T1 3D HR	Dorsal	8	24	175 x 175	1.5	1.5	48	43	
T1 3D	Sagittal	7	24	175 x 175	3.8	3.8	26	43	
T1 3D	Transverse	7	24	175 x 175	3	3	26	43	
T2 FSE	Transverse	81	1848	175 x 175	5	6	12	90	
T2 FSE	Transverse	81	1540	175 x 175	5	6	10	90	
T2 FSE	Transverse	132	2316	175 x 175	5	6	12	90	
T2 FSE PSAT	Transverse	81	1848	175 x 175	5	6	12	90	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** PD- Proton density weighted, SE- Spin echo, STIR- Short tau inversion recovery, FSE- Fast spin echo, T1- T1 weighted, HR- High resolution, T2- T2 weighted, PSAT- pre-saturation sequence.

## Low-field under general anaesthesia

### Case 1

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
T2 FSE	Sagittal	75	4540	140 x 140	4	4.4	23	90	
T2 FSE	Transverse	75	3350	140 x 140	4.5	4.9	17	90	
STIR	Sagittal	26	3500	140 x 140	4	4.4	23	90	85
STIR	Transverse	26	4110	140 x 140	4.5	4.9	27	90	85
TME (PD)	Transverse	28	2930	140 x 140	4	4.5	9	90	
TME (T2)	Transverse	90	2930	140 x 140	4	4.5	9	90	
TME (PD)	Transverse	28	4650	140 x 140	4	4.5	9	90	
TME (T2)	Transverse	90	4650	140 x 140	4	4.5	9	90	
T1 Turbo 3D	Dorsal	16	38	140 x 140	0.8	0.8	104	65	
T1 MPR	Transverse				0.6	0.6	96		
T1 MPR	Transverse				0.6	0.6	96		
T1 MPR	Transverse				0.6	0.6	96		
T1 MPR	Sagittal				0.6	0.6	96		
T1 MPR	Sagittal				0.6	0.6	96		

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** T2- T2 weighted, FSE- Fast spin echo, STIR- Short tau inversion recovery, TME- Turbo multi echo, PD- Proton density weighted, T1- T1 weighted, MPR- multiplanar reconstruction.

### Case 3

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
T2 FSE	Sagittal	75	5320	140 x 140	4	4.4	27	90	
T2 FSE	Transverse	75	2960	140 x 140	4.5	4.9	15	90	
STIR	Sagittal	26	3200	140 x 140	4	4.4	21	90	82
STIR	Transverse	26	3200	140 x 140	4.5	4.9	21	90	85
TME (PD)	Transverse	28	2580	140 x 140	4	4.5	15	90	
TME (T2)	Transverse	90	2580	140 x 140	4	4.5	15	90	
TME (PD)	Transverse	28	2580	140 x 140	4	4.5	15	90	
TME (T2)	Transverse	90	2580	140 x 140	4	4.5	15	90	
T1 Turbo 3D	Dorsal	16	38	140 x 140	0.8	0.8	104	65	
T1 MPR	Transverse				0.6	0.6	96		
T1 MPR	Transverse				0.6	0.6	96		
T1 MPR	Sagittal				0.6	0.9	96		

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** T2- T2 weighted, FSE- Fast spin echo, STIR- Short tau inversion recovery, TME- Turbo multi echo, PD- Proton density weighted, T1- T1 weighted, MPR- multiplanar reconstruction.

## Case 9

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
T2 FSE	Sagittal	75	4140	140 x 140	4	4.4	21	90	
T2 FSE	Transverse	75	2960	140 x 140	4.5	4.9	15	90	
STIR	Sagittal	26	3200	140 x 140	4	4.4	21	90	85
STIR	Transverse	26	3500	140 x 140	4.5	4.9	23	90	85
TME (PD)	Transverse	28	2930	140 x 140	4	4.5	17	90	
TME (T2)	Transverse	90	2930	140 x 140	4	4.5	17	90	
TME (PD)	Transverse	28	2800	140 x 140	4	4.5	13	90	
TME (T2)	Transverse	90	2800	140 x 140	4	4.5	13	90	
T1 Turbo 3D	Dorsal	16	38	140 x 140	0.8	0.8	104	65	
T1 MPR	Sagittal				0.6	0.6	96		
T1 MPR	Transverse				0.6	0.6	96		

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** T2- T2 weighted, FSE- Fast spin echo, STIR- Short tau inversion recovery, TME- Turbo multi echo, PD- Proton density weighted, T1- T1 weighted, MPR- multiplanar reconstruction.

### Case 13

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
T2 FSE	Sagittal	75	4140	140 x 140	4	4.4	21	90	
T2 FSE	Transverse	75	2960	140 x 140	4.5	4.9	15	90	
STIR	Sagittal	26	2890	140 x 140	4	4.4	19	90	85
STIR	Transverse	26	3200	140 x 140	4.5	4.9	21	90	85
TME (PD)	Transverse	28	2930	140 x 140	4	4.5	17	90	
TME (T2)	Transverse	90	2930	140 x 140	4	4.5	17	90	
TME (PD)	Transverse	28	2800	140 x 140	4	4.5	13	90	
TME (T2)	Transverse	90	2800	140 x 140	4	4.5	13	90	
T1 3D Turbo	Dorsal	16	38	140 x 140	0.8	0.8	104	65	
T1 MPR	Sagittal				0.6	0.6	96		
T1 MPR	Transverse				0.6	0.6	96		

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** T2- T2 weighted, FSE- Fast spin echo, STIR- Short tau inversion recovery, TME- Turbo multi echo, PD- Proton density weighted, T1- T1 weighted, MPR- multiplanar reconstruction.



## Case 15

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
T2 FSE	Sagittal	75	4140	140 x 140	4	4.4	21	90	
T2 FSE	Transverse	75	2960	140 x 140	4.5	4.9	15	90	
STIR	Sagittal	26	3200	140 x 140	4	4.4	21	90	85
STIR	Transverse	26	3500	140 x 140	4.5	4.9	23	90	85
TME (PD)	Transverse	28	3960	140 x 140	4	4.5	23	90	
TME (T2)	Transverse	90	3960	140 x 140	4	4.5	23	90	
TME (PD)	Transverse	28	2800	140 x 140	4	4.5	26	90	
TME (T2)	Transverse	90	2800	140 x 140	4	4.5	26	90	
T1 3D Turbo	Dorsal	16	38	140 x 140	0.8	0.8	104	65	
T1 MPR	Transverse				0.6	0.6	96		
T1 MPR	Transverse				0.6	0.6	96		
T1 MPR	Transverse				0.6	0.6	96		
T1 MPR	Transverse				0.6	0.6	96		
T1 MPR	Sagittal				0.6	0.6	96		

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** T2- T2 weighted, FSE- Fast spin echo, STIR- Short tau inversion recovery, TME- Turbo multi echo, PD- Proton density weighted, T1- T1 weighted, MPR- multiplanar reconstruction.

## High-field under general anaesthesia

### Case 2

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
PD FS	Dorsal	14	3990	150 x 150	2.5	3	25	180	
PD TSE	Sagittal	15	3970	140 x 140	2.5	2.8	27	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
STIR	Sagittal	26	5840	140 x 140	3.5	4	23	180	155
STIR	Transverse	26	7610	150 x 136	4	4.5	30	180	155
T1 3D FLASH FS	Sagittal	10	36	150 x 150	2		48	40	
T1 3D FLASH FS	Transverse	10	36	140 x 140	2		40	40	
T2 FS	Transverse	81	4840	150 x 150	3	3.6	26	180	
T2 FS	Dorsal	81	5590	150 x 150	3	3.6	30	180	
T2 TSE	Transverse	81	4750	150 x 150	4	4.5	30	180	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** PD- Proton density weighted, FS- Fat saturated, TSE- Turbo spin echo, STIR- Short tau inversion recovery, T1- T1 weighted, FLASH- fast low angle shot.

## Case 4

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
PD TSE	Sagittal	15	3970	140 x 140	2.5	2.8	27	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
PD TSE	Transverse	14	2000	150 x 150	4	4.5	12	180	
STIR	Sagittal	26	5840	140 x 140	3.5	4	23	180	155
STIR	Transverse	26	7610	150 x 136	4	4.5	30	180	155
T1 3D FLASH	Dorsal	10	18	140 x 140	2		32	40	
T1 3D FLASH FS	Sagittal	10	36	150 x 150	2		48	40	
T1 3D FLASH FS	Transverse	10	36	140 x 140	2		40	40	
T2 FS	Transverse	81	4840	150 x 150	3	3.6	26	180	
T2 FS	Dorsal	81	5590	150 x 150	3	3.6	30	180	
T2 TSE	Transverse	81	4750	150 x 150	4	4.5	30	180	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** PD- Proton density weighted, TSE- Turbo spin echo, STIR- Short tau inversion recovery, T1- T1 weighted, FLASH- fast low angle shot, FS- Fat saturated, T2- T2 weighted.

## Case 7

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
PD TSE	Dorsal	14	3950	150 x 150	2.5	2.8	30	180	
PD TSE	Sagittal	15	3970	140 x 140	2.5	2.8	27	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
STIR	Sagittal	26	5840	140 x 140	3.5	4	23	180	155
STIR	Transverse	26	7610	150 x 136	4	4.5	22	180	155
T1 3D FLASH FS	Sagittal	10	36	150 x 150	2		48	40	
T1 3D FLASH FS	Transverse	10	36	140 x 140	2		40	40	
T2 FS	Transverse	81	4840	150 x 150	3	3.6	26	180	
T2 FS	Dorsal	81	5590	150 x 150	3	3.6	30	180	
T2 TSE	Transverse	81	4750	150 x 150	4	4.5	30	180	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** PD- Proton density weighted, TSE- Turbo spin echo, STIR- Short tau inversion recovery, T1- T1 weighted, FLASH- fast low angle shot, FS- Fat saturated, T2- T2 weighted.

**Case 8**

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
PD TSE	Sagittal	15	3970	130 x 130	2.5	2.8	27	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
STIR	Sagittal	26	5840	140 x 140	3.5	4	23	180	155
STIR	Transverse	26	7610	150 x 136	4	4.5	30	180	155
T1 3D FLASH	Dorsal	10	18	140 x 140	2		32	40	
T1 3D FLASH FS	Sagittal	10	36	150 x 150	2		48	40	
T1 3D FLASH FS	Transverse	10	36	140 x 140	2		40	40	
T2 FS	Transverse	81	4840	150 x 150	3	3.6	26	180	
T2 FS	Dorsal	81	5590	150 x 150	3	3.6	30	180	
T2 TSE	Transverse	84	4930	150 x 138	4	4.5	30	180	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** PD- Proton density weighted, TSE- Turbo spin echo, STIR- Short tau inversion recovery, T1- T1 weighted, FLASH- fast low angle shot, FS- Fat saturated, T2- T2 weighted.

## Case 11

Pulse sequence	Orientation	TE (ms)	TR (ms)	FOV (mm)	Slice thickness (mm)	Interslice spacing (mm)	Number of slices	Flip angle	Inversion time (ms)
PD TSE	Dorsal	14	3950	150 x 150	2.5	2.8	30	180	
PD TSE	Sagittal	15	3970	140 x 140	2.5	2.8	27	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
PD TSE	Transverse	14	3940	150 x 150	4	4.5	30	180	
STIR	Sagittal	26	5840	140 x 140	3.5	4	23	180	155
STIR	Transverse	26	7610	150 x 136	4	4.5	30	180	155
T1 3D FLASH FS	Sagittal	10	36	150 x 150	2		48	40	
T1 3D FLASH FS	Transverse	10	36	140 x 140	2		40	40	
T2 FS	Transverse	81	4840	170 x 170	3	3.6	26	180	
T2 FS	Dorsal	81	5590	170 x 170	3	3.6	30	180	
T2 FSE	Transverse	81	4750	150 x 150	4	4.5	30	180	

**Header abbreviations:** TE- Echo time, TR- Repetition time, FOV- Field of view.

**Pulse sequence abbreviations:** PD- Proton density weighted, TSE- Turbo spin echo, STIR- Short tau inversion recovery, T1- T1 weighted, FLASH- fast low angle shot, FS- Fat saturated, T2- T2 weighted.

## **Appendix 3: Inter-observer agreement analysis for image quality assessment across all systems**

### **Part 1: Inter-observer agreement for the diagnostic or non-diagnostic quality of magnetic resonance imaging studies of the foot**

The Fleiss' kappa values for all assessment categories were interpreted to demonstrate poor agreement for the dichotomous diagnostic versus non-diagnostic quality of studies (Table A3-1).

### **Part 2: Overall inter-observer agreement for image quality assessment grading**

The agreement analysis across all grades for image quality assessment of entire studies and individual anatomic structures included data from all observers across all acquisition systems (Table A3-2). Interpretation of the kappa values demonstrated poor agreement for assessment of diagnostic quality for the navicular bone, third phalanx and distal sesamoidean impar ligament. The kappa values were interpreted to indicate fair agreement for assessment of diagnostic quality for entire studies, the deep digital flexor tendon, navicular bursa, distal interphalangeal joint and the collateral ligaments of the distal interphalangeal joint. The relative agreement (analysis that accounts for the relative ranking of grading) as determined by Kendall's coefficient of concordance was deemed to be moderate to high for all assessment categories.

### **Part 3: Inter-observer agreement for individual grades of image quality assessment grading**

Fleiss' kappa values were determined for individual grades of the image quality grading scale for all assessment categories. Except for the collateral ligaments of the distal interphalangeal joint, agreement was greatest for grade 1 image quality, "Textbook image quality" (Figure A3-1).

**Table A3-1 Fleiss' kappa values for dichotomous diagnostic versus non-diagnostic quality of all magnetic resonance imaging studies**

Abbreviations: DIP- distal interphalangeal, CI- confidence interval.

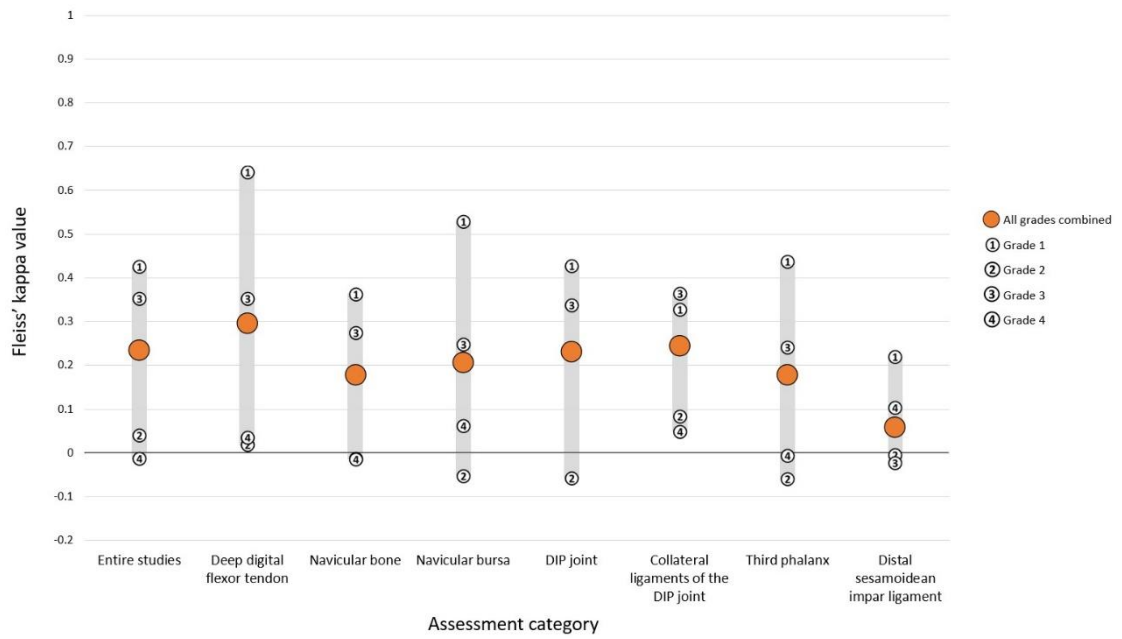
Assessment category	Fleiss' kappa		
	Kappa statistic	95% CI lower bound	95% CI upper bound
Entire studies	-0.01	-0.02	-0.01
Deep digital flexor tendon	0.04	0.03	0.04
Navicular bone	-0.02	-0.02	-0.01
Navicular bursa	0.06	0.06	0.06
DIP joint	N/A	N/A	N/A
Collateral ligaments of the DIP joint	0.05	0.05	0.05
Third phalanx	-0.01	-0.01	0.00
Distal sesamoidean impar ligament	0.10	0.10	0.11

**Table A3-2 Output of inter-observer agreement analysis for magnetic resonance image quality grading for all studies**

Abbreviations: DIP- distal interphalangeal, CI- confidence interval.

Assessment category	Percentage matched assessments		Fleiss' Kappa			Kendall's coefficient of concordance
	Value (%)	95% CI	Kappa statistic	95% CI lower bound	95% CI upper bound	
Entire studies	0.0	0.0, 18	0.23	0.23	0.23	0.68
Deep digital flexor tendon	6.7	0.17, 32	0.30	0.30	0.30	0.75
Navicular bone	0.0	0.0, 18	0.18	0.18	0.18	0.62
Navicular bursa	0.0	0.0, 18	0.21	0.20	0.21	0.73
DIP joint	0.0	0.0, 18	0.23	0.23	0.23	0.62
Collateral ligaments of the DIP joint	0.0	0.0, 18	0.24	0.24	0.24	0.73
Third phalanx	0.0	0.0, 18	0.18	0.18	0.18	0.59
Distal sesamoidean impar ligament	0.0	0.0, 18	0.06	0.06	0.06	0.56







**Figure A3-1 Linear dot plot displaying the Fleiss' kappa values for image quality assessment of the equine foot for seven anatomic structures.**

## Appendix 4: Bubble charts for pathology assessment and pathology assessment confidence

### General information

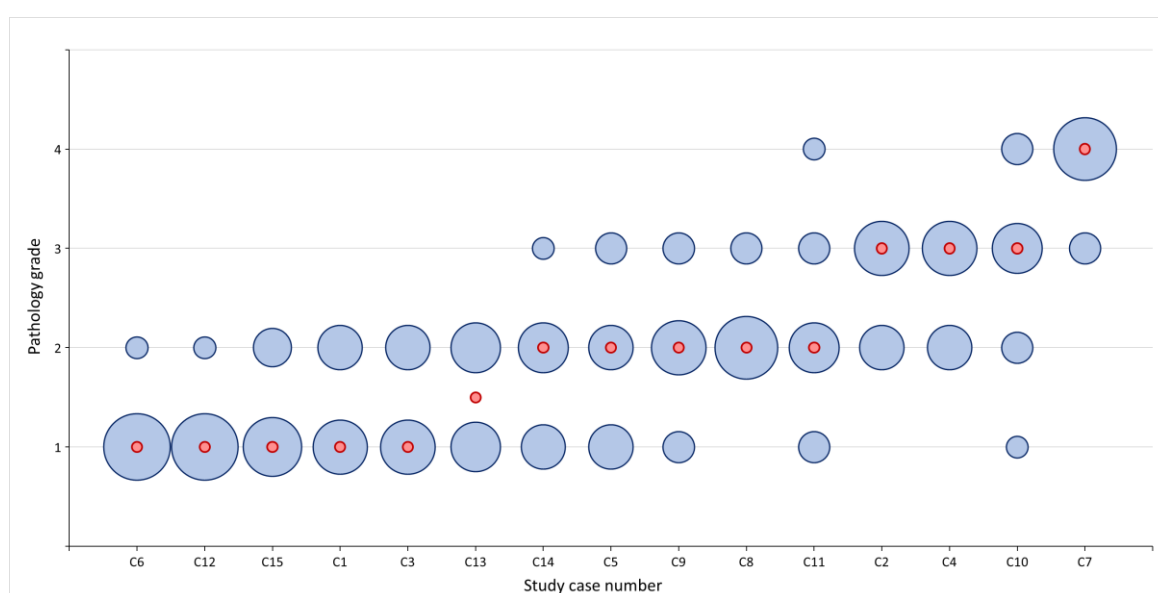
The bubble charts enclosed in this appendix demonstrate the distribution of observer gradings for pathology assessment (part 1) and pathology assessment confidence (part 2). For each individual anatomical structure, a chart plots the gradings for each MRI study. The MRI studies on the x-axes are organised by median grade from lowest grade to highest grade for the respective structure. The size (area) of the bubble at each site is proportional to the number of observers selecting that grade.

### Legend

-  Bubble area proportional to the count of observers assigning grade for an MRI study
-  Median grade for MRI study

### Part 1: Pathology assessment

#### Deep digital flexor tendon



**Figure A4-1** Bubble chart demonstrating deep digital flexor tendon pathology assessment gradings for 15 MRI studies.

### Navicular bone

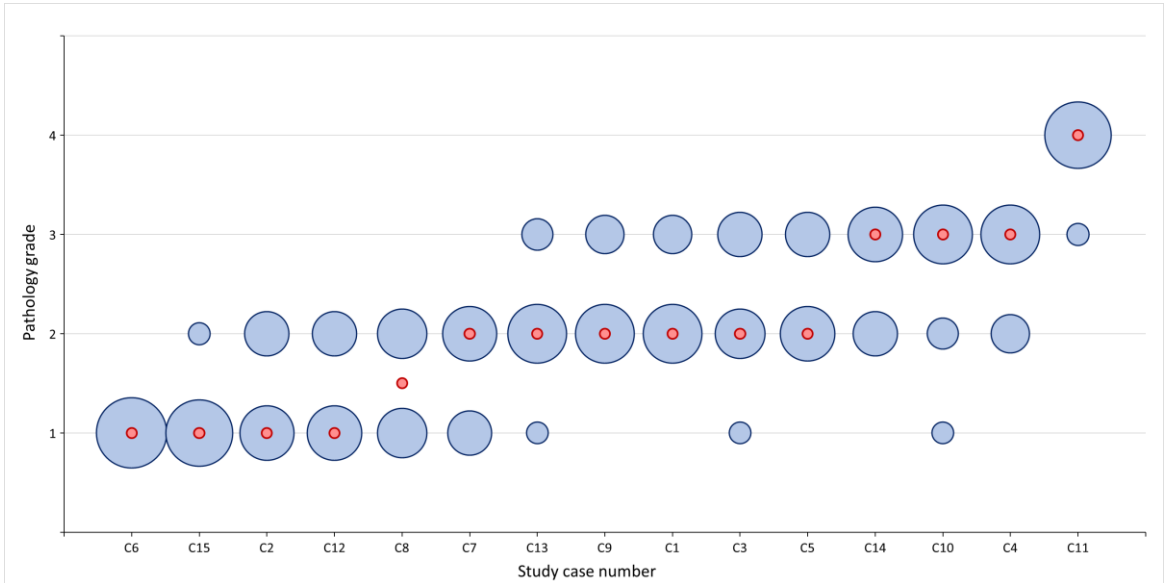


Figure A4-2 Bubble chart demonstrating navicular bone pathology assessment gradings for 15 MRI studies.

### Navicular bursa

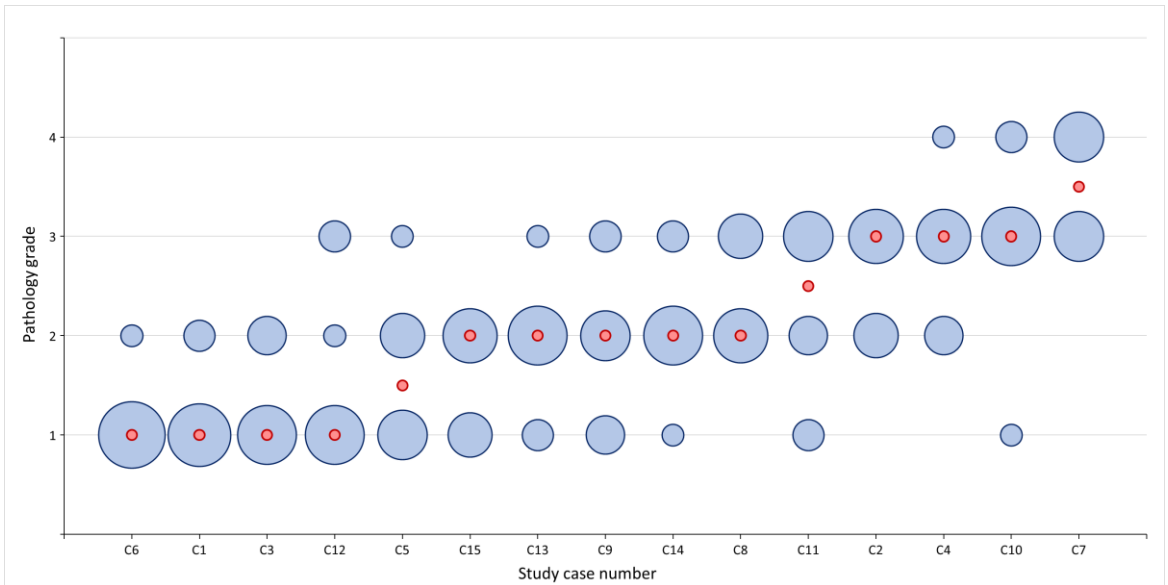


Figure A4-3 Bubble chart demonstrating navicular bursa pathology assessment gradings for 15 MRI studies.

## Distal interphalangeal joint

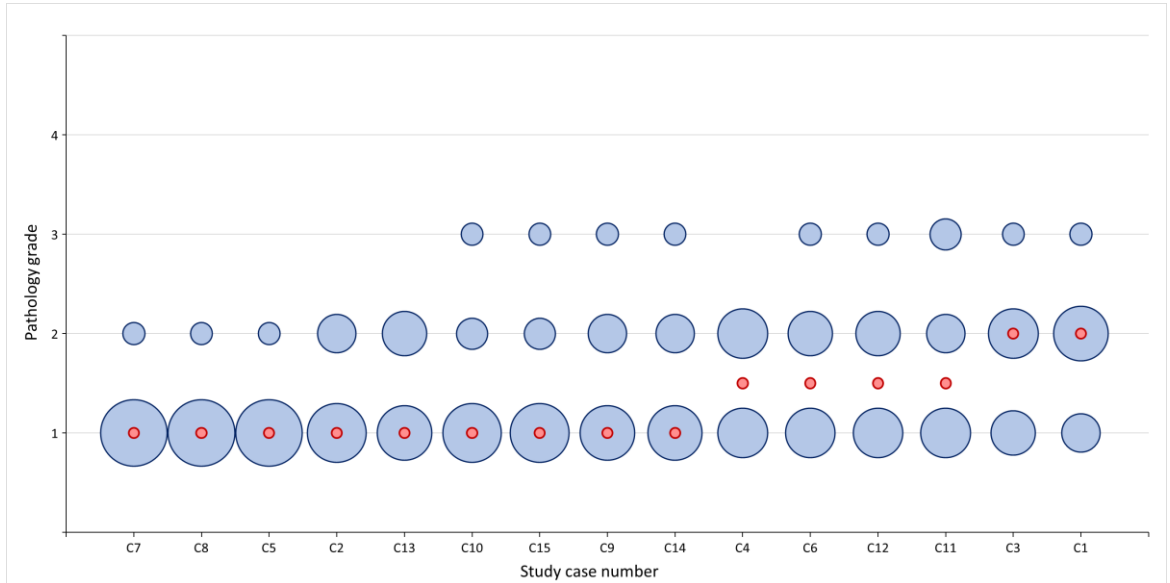


Figure A4-4 Bubble chart demonstrating distal interphalangeal joint pathology assessment gradings for 15 MRI studies.

## Collateral ligaments of the distal interphalangeal joint

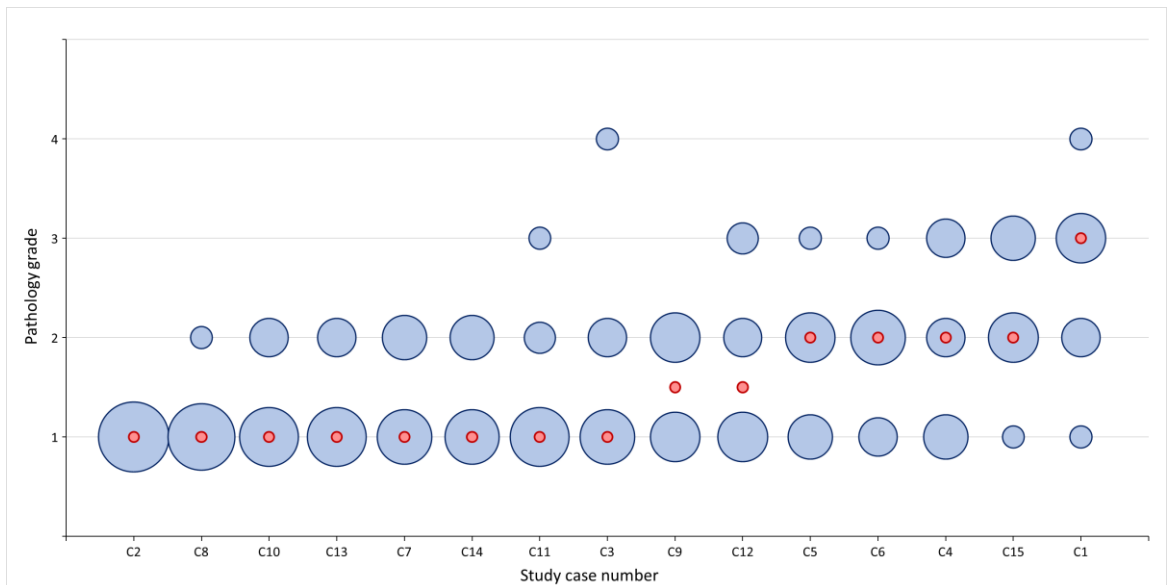


Figure A4-5 Bubble chart demonstrating collateral ligament of the distal interphalangeal joint pathology assessment gradings for 15 MRI studies.

### Third phalanx

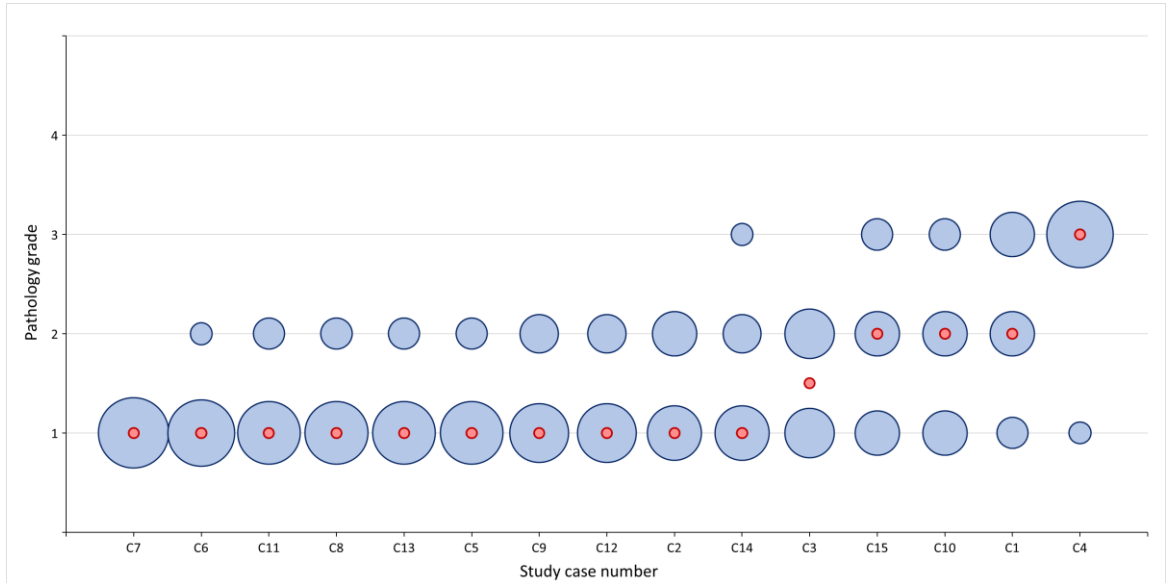


Figure A4-6 Bubble chart demonstrating third phalanx pathology assessment gradings for 15 MRI studies.

### Distal sesamoidean impar ligament

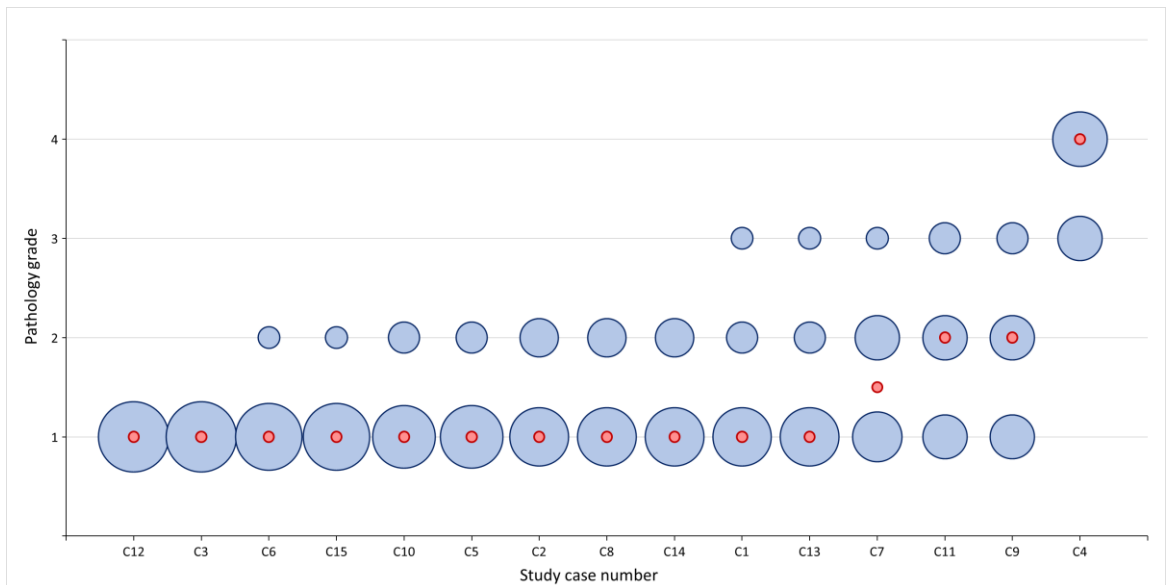


Figure A4-7 Bubble chart demonstrating distal sesamoidean impar ligament pathology assessment gradings for 15 MRI studies.

## Part 2: Pathology assessment confidence

### Deep digital flexor tendon

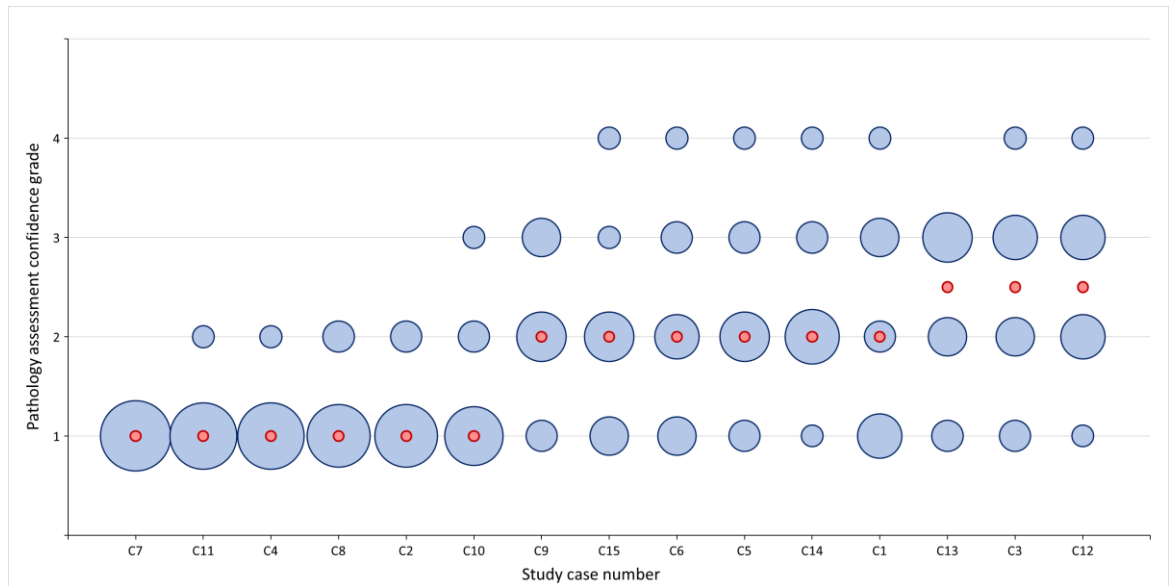


Figure A4-8 Bubble chart demonstrating deep digital flexor tendon pathology assessment confidence gradings for 15 MRI studies.

### Navicular bone

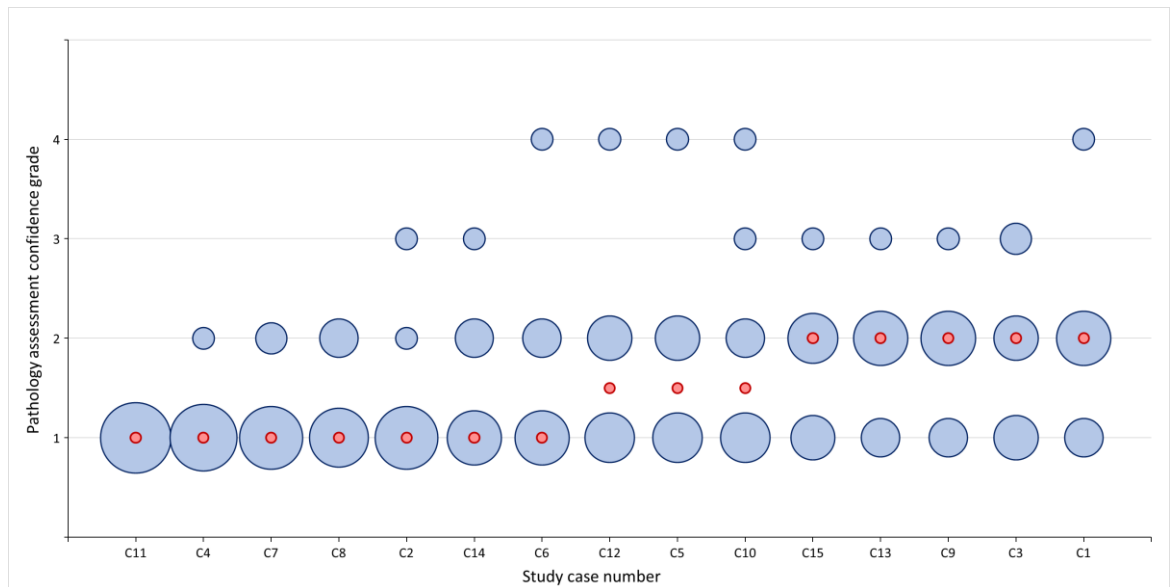


Figure A4-9 Bubble chart demonstrating navicular bone pathology assessment confidence gradings for 15 MRI studies.

### Navicular bursa

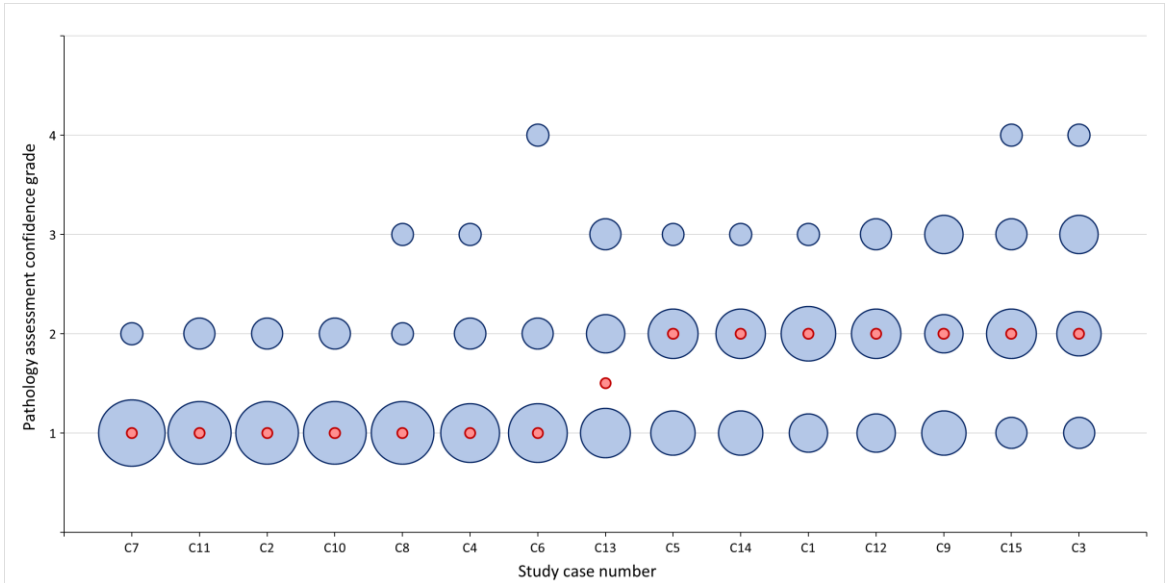


Figure A4-10 Bubble chart demonstrating navicular bursa pathology assessment confidence gradings for 15 MRI studies.

### Distal interphalangeal joint

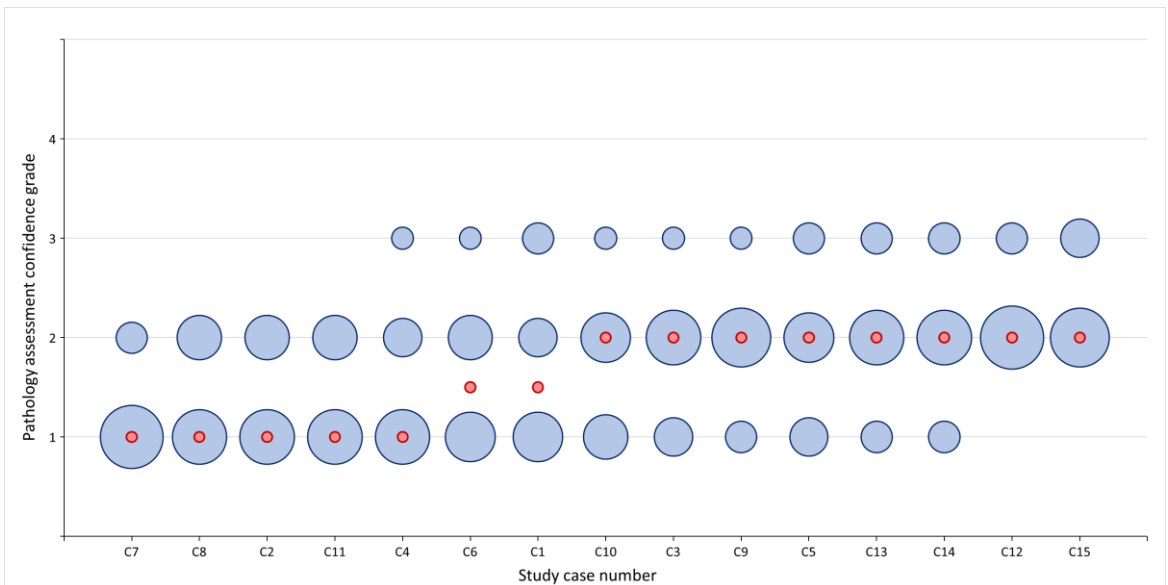


Figure A4-11 Bubble chart demonstrating distal interphalangeal joint pathology assessment confidence gradings for 15 MRI studies.

## Collateral ligaments of the distal interphalangeal joint

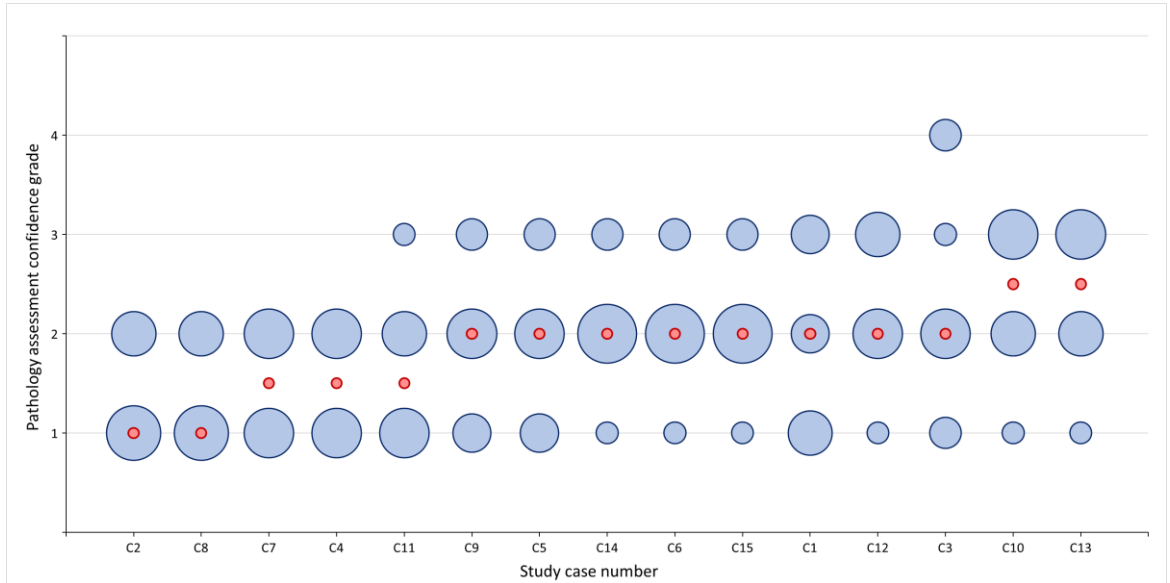


Figure A4-12 Bubble chart demonstrating collateral ligament of the distal interphalangeal joint pathology assessment confidence gradings for 15 MRI studies.

## Third phalanx

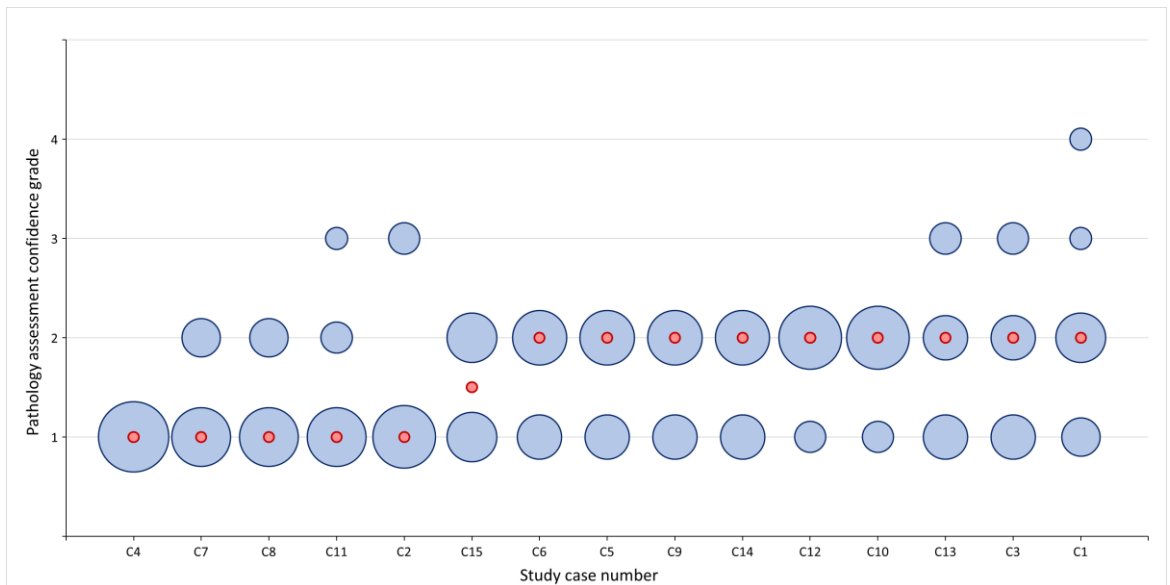


Figure A4-13 Bubble chart demonstrating third phalanx pathology assessment confidence gradings for 15 MRI studies.



### Distal sesamoidean impar ligament

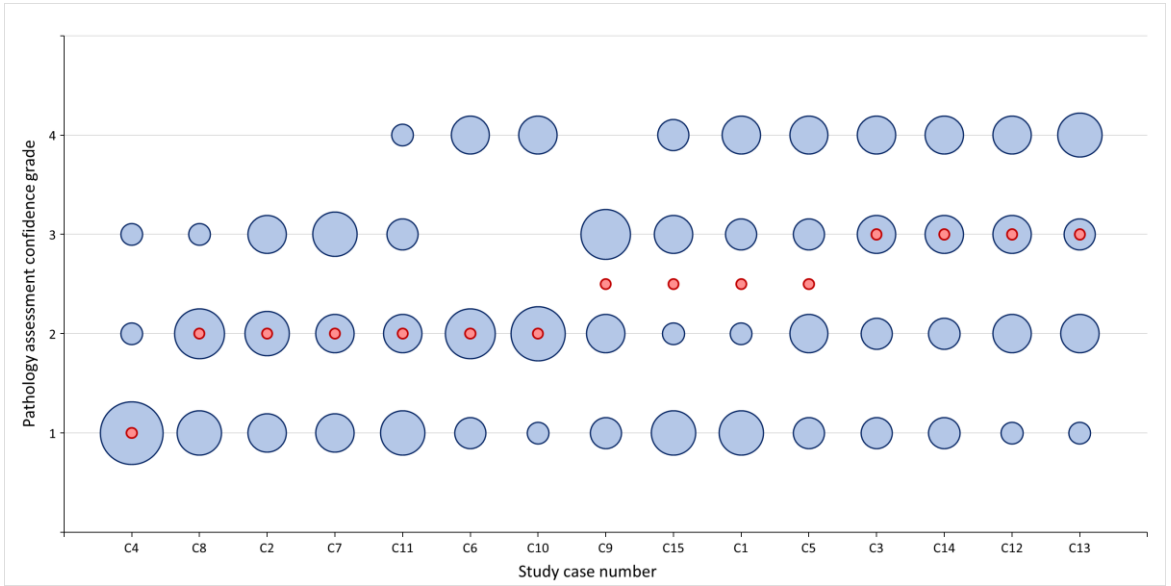


Figure A4-14 Bubble chart demonstrating distal sesamoidean impar ligament pathology assessment confidence gradings for 15 MRI studies.

## List of References

- Altman, D. G. (1991) *Practical Statistics for Medical Research*. London: Chapman and Hall.
- Andersen, M. S., Clark, L., Dyson, S. J. and Newton, J. R. (2006) 'Risk factors for colic in horses after general anaesthesia for MRI or nonabdominal surgery: absence of evidence of effect from perianaesthetic morphine', *Equine Veterinary Journal*, 38(4), pp. 368-374.
- Barile, A., Conti, L., Lanni, G., Calvisi, V. and Masciocchi, C. (2013) 'Evaluation of medial meniscus tears and meniscal stability: weight-bearing MRI vs arthroscopy', *European Journal of Radiology*, 82(4), pp. 633-639.
- Barrett, M. and Acutt, E. (2020) 'Radiography', in Baxter, G. M. (ed.) *Adams and Stashak's Lameness in Horses*. Wiley, pp. 189-300.
- Barrett, M. F., Frisbie, D. D., King, M. R., Werpy, N. M. and Kawcak, C. E. (2017) 'A review of how magnetic resonance imaging can aid in case management of common pathological conditions of the equine foot', *Equine Veterinary Education*, 29(12), pp. 683-693.
- Baxter, G. M., Stashak, T. S. and Keegan, K. G. (2020) 'Examination for lameness', in Baxter, G. M. (ed.) *Adams and Stashak's Lameness in Horses*. Wiley, pp. 67-188.
- Berner, D., Mader, D., Groß, C. and Gerlach, K. (2020) 'Effect of Scan Plane and Arthrography on Visibility and Interobserver Agreement of the Equine Distal Sesamoidean Impar Ligament on Magnetic Resonance Images', *Journal of Equine Veterinary Science*, 94, p. 103227.
- Bidwell, L. A., Bramlage, L. R. and Rood, W. A. (2007) 'Equine perioperative fatalities associated with general anaesthesia at a private practice--a retrospective case series', *Veterinary Anaesthesia and Analgesia*, 34(1), pp. 23-30.
- Black, B., Cribb, N. C., Nykamp, S. G., Thomason, J. J. and Trout, D. R. (2013) 'The effects of perineural and intrasynovial anaesthesia of the equine foot on subsequent magnetic resonance images', *Equine Veterinary Journal*, 45(3), pp. 320-325.
- Bladon, B. (2014) 'Magnetic resonance imaging and foot lameness. Problem solved? Or do we know we know less now that we know more?', *Equine Veterinary Journal*, 46(3), pp. 264-266.
- Bland, M. (2015) *An Introduction to Medical Statistics*. Oxford University Press.
- Blunden, A., Dyson, S., Murray, R. and Schramme, M. (2010a) 'Histopathology in horses with chronic palmar foot pain and age-matched controls. Part 1: Navicular bone and related structures', *Equine Veterinary Journal*, 38(1), pp. 15-22.

- Blunden, A., Dyson, S., Murray, R. and Schramme, M. (2010b) 'Histopathology in horses with chronic palmar foot pain and age-matched controls. Part 2: The deep digital flexor tendon', *Equine Veterinary Journal*, 38(1), pp. 23-27.
- Blunden, A., Murray, R. and Dyson, S. (2009) 'Lesions of the deep digital flexor tendon in the digit: A correlative MRI and post mortem study in control and lame horses', *Equine Veterinary Journal*, 41(1), pp. 25-33.
- Bolas, N. (2010) 'Basic MRI principles', in Murray, R. C. (ed.) *Equine MRI*. Chichester, UK: John Wiley & Sons, Ltd, pp. 1-37.
- Bolen, G., Audigié, F., Spriet, M., Vandenberghe, F. and Busoni, V. (2010) 'Qualitative Comparison of 0.27T, 1.5T, and 3T Magnetic Resonance Images of the Normal Equine Foot', *Journal of Equine Veterinary Science*, 30(1), pp. 9-20.
- Bolen, G., Haye, D., Dondelinger, R. and Busoni, V. (2010) 'Magnetic resonance signal changes during time in equine limbs refrigerated at 4 degrees C', *Veterinary Radiology & Ultrasound*, 51(1), pp. 19-24.
- Bowers, D. (2019) *Medical Statistics from Scratch: An Introduction for Health Professionals*. Wiley.
- Bowker, R. M. (2011) 'Chapter 29 - Functional Anatomy of the Palmar Aspect of the Foot', in Ross, M. W. and Dyson, S. J. (eds) *Diagnosis and Management of Lameness in the Horse*. 2nd edn. Saint Louis: W.B. Saunders, pp. 320-323.
- Bruno, F., Barile, A., Arrigoni, F., Laporta, A., Russo, A., Carotti, M., Splendiani, A., Di Cesare, E. and Masciocchi, C. (2018) 'Weight-bearing MRI of the knee: a review of advantages and limits', *Acta bio-medica: Atenei Parmensis*, 89(1-S), pp. 78-88.
- Brunyé, T. T., Drew, T., Weaver, D. L. and Elmore, J. G. (2019) 'A review of eye tracking for understanding and improving diagnostic interpretation', *Cognitive Research: Principles and Implications*, 4(1), p. 7.
- Budras, K.-D., Sack, W. O., Rock, S., Horowitz, A. and Berg, R. (2009a) 'Chapter 2: Thoracic Limb', in *Anatomy of the Horse*. Schluetersche, Germany, pp. 4-15.
- Budras, K.-D., Sack, W. O., Rock, S., Horowitz, A. and Berg, R. (2009b) 'Chapter 3: Pelvic Limb', in *Anatomy of the Horse*. Schluetersche, Germany, pp. 16-31.
- Busoni, V., Heimann, M., Trenteseaux, J., Snaps, F. and Dondelinger, R. F. (2005) 'Magnetic resonance imaging findings in the equine deep digital flexor tendon and distal sesamoid bone in advanced navicular disease--an ex vivo study', *Veterinary Radiology & Ultrasound*, 46(4), pp. 279-286.
- Busoni, V. and Snaps, F. (2002) 'Effect of deep digital flexor tendon orientation on magnetic resonance imaging signal intensity in isolated equine limbs-the magic angle effect', *Veterinary Radiology & Ultrasound*, 43(5), pp. 428-430.
- Butler, J. A., Colles, C. M., Dyson, S. J., Kold, S. E. and Poulos, P. W. (2017) 'Chapter 3- The foot', in Butler, J. A., Colles, C. M., Dyson, S. J., Kold, S. E., and Poulos, P. W. (eds) *Clinical Radiology of the Horse*. John Wiley & Sons, pp. 55-148.

- Bydder, M., Rahal, A., Fullerton, G. D. and Bydder, G. M. (2007) 'The magic angle effect: a source of artifact, determinant of image contrast, and technique for imaging', *Journal of Magnetic Resonance Imaging*, 25(2), pp. 290-300.
- Carnicer, D., Coudry, V. and Denoix, J.-M. (2013) 'Ultrasonographic examination of the palmar aspect of the pastern of the horse: Sesamoidean ligaments', *Equine Veterinary Education*, 25(5), pp. 256-263.
- Carstens, A. and Smith, R. K. W. (2014) 'Ultrasonography of the Foot and Pastern', in Kidd, J. A., Lu, K. G., and Frazer, M. L. (eds) *Atlas of Equine Ultrasonography*. Oxford, UK: John Wiley & Sons, Ltd, pp. 23-44.
- Chow, L. S., Rajagopal, H. and Paramesran, R. (2016) 'Correlation between subjective and objective assessment of magnetic resonance (MR) images', *Magnetic Resonance Imaging*, 34(6), pp. 820-831.
- Collins, C. M. (2016) *Electromagnetics in Magnetic Resonance Imaging*. Morgan & Claypool Publishers.
- Collins, M. S. and Ehman, R. L. (2012) 'Interpretation of Magnetic Resonance Images', in Berquist, T. H. (ed.) *MRI of the Musculoskeletal System*. Wolters Kluwer Health, pp. 34-63.
- Contino, E. K., Barrett, M. F. and Werpy, N. M. (2014) 'Effect of limb positioning on the radiographic appearance of the distal and proximal interphalangeal joint spaces of the forelimbs of horses during evaluation of dorsopalmar radiographs', *Journal of the American Veterinary Medical Association*, 244(10), pp. 1186-1190.
- Daniel, A. J., Goodrich, L. R., Barrett, M. F., Werpy, N. M., Morley, P. S. and McIlwraith, C. W. (2016) 'An optimised injection technique for the navicular bursa that avoids the deep digital flexor tendon', *Equine Veterinary Journal*, 48(2), pp. 159-164.
- Davies, H. M. S. and Philip, C. (2007) 'Chapter 1 - Gross Anatomy of the Equine Digit', in Floyd, A. E. and Mansmann, R. A. (eds) *Equine Podiatry*. Saint Louis: W.B. Saunders, pp. 1-24.
- Delfaut, E. M., Beltran, J., Johnson, G., Rousseau, J., Marchandise, X. and Cotten, A. (1999) 'Fat suppression in MR imaging: techniques and pitfalls', *RadioGraphics*, 19(2), pp. 373-382.
- Denoix, J. M. (1994) 'Functional anatomy of tendons and ligaments in the distal limbs (manus and pes)', *The Veterinary Clinics of North America: Equine Practice*. W.B. Saunders Company, 10(2), pp. 273-322.
- Denoix, J.-M. (2000) 'Dissections of the equine foot', in Denoix, J.-M. (ed.) *The Equine Distal Limb: An Atlas of Clinical Anatomy and Comparative Imaging*. Taylor & Francis, pp. 1-34.
- Denoix, J.-M., Bertoni, L., Heitzmann, A.-G., Werpy, N. and Audigié, F. (2011) 'Ultrasonographic examination of the collateral ligaments of the distal interphalangeal joint in horses: Part A: Technique and normal images', *Equine Veterinary Education*, 23(11), pp. 574-580.

- Denoix, J.-M., Crevier, N., Roger, B. and Lebas, J.-F. (1993) 'Magnetic resonance imaging of the equine foot', *Veterinary Radiology & Ultrasound*, 34(6), pp. 405-411.
- Ding, Y. (2018) 'Medical Image Quality Assessment', in Ding, Y. (ed.) *Visual Quality Assessment for Natural and Medical Image*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 215-264.
- Dowsett, D., Kenny, P. A. and Johnston, R. E. (2006) *The Physics of Diagnostic Imaging*. 2nd edn. Taylor & Francis (A Hodder Arnold Publication).
- Dugdale, A. H. A., Obhrai, J. and Cripps, P. J. (2016) 'Twenty years later: a single-centre, repeat retrospective analysis of equine perioperative mortality and investigation of recovery quality', *Veterinary Anaesthesia and Analgesia*, 43(2), pp. 171-178.
- Dyson, S., Blunden, T. and Murray, R. (2012) 'Comparison between magnetic resonance imaging and histological findings in the navicular bone of horses with foot pain', *Equine Veterinary Journal*, 44(6), pp. 692-698.
- Dyson, S. J. (2002) 'Subjective and quantitative scintigraphic assessment of the equine foot and its relationship with foot pain', *Equine Veterinary Journal*, 34(2), pp. 164-170.
- Dyson, S. J. (2011a) 'Chapter 33 - The Distal Phalanx and Distal Interphalangeal Joint', in Ross, M. W. and Dyson, S. J. (eds) *Diagnosis and Management of Lameness in the Horse*. 2nd edn. Saint Louis: W.B. Saunders, pp. 349-366.
- Dyson, S. J. (2011b) 'Nonseptic osteitis of the distal phalanx and its palmar processes', *Equine Veterinary Education*, 23(9), pp. 472-485.
- Dyson, S. J., Blunden, T. and Murray, R. (2008) 'The collateral ligaments of the distal interphalangeal joint: Magnetic resonance imaging and post mortem observations in 25 lame and 12 control horses', *Equine Veterinary Journal*, 40(6), pp. 538-544.
- Dyson, S. J. and Murray, R. C. (2010) 'Chapter 12: The foot and pastern', in Murray, Rachel C. (ed.) *Equine MRI*. Chichester, UK: John Wiley & Sons, Ltd, pp. 269-314.
- Dyson, S. J., Murray, R. C. and Schramme, M. (2005) 'Lameness associated with foot pain: results of magnetic resonance imaging in 199 horses (January 2001-December 2003) and response to treatment', *Equine Veterinary Journal*, 37(2), pp. 113-121.
- Dyson, S. J., Murray, R. C., Schramme, M. and Branch, M. (2003) 'Magnetic resonance imaging of the equine foot: 15 horses', *Equine Veterinary Journal*, 35(1), pp. 18-26.
- Dyson, S. J. and Nagy, A. (2011) 'Injuries associated with the cartilages of the foot', *Equine Veterinary Education*, 23(11), pp. 581-593.
- Dyson, S. J., Pool, R., Blunden, T. and Murray, R. C. (2010) 'The distal sesamoidean impar ligament: comparison between its appearance on magnetic

resonance imaging and histology of the axial third of the ligament', *Equine Veterinary Journal*, 42(4), pp. 332-339.

Dyson, S. J., Thomson, K., Quiney, L., Bondi, A. and Ellis, A. D. (2020) 'Can veterinarians reliably apply a whole horse ridden ethogram to differentiate nonlame and lame horses based on live horse assessment of behaviour?', *Equine Veterinary Education*, 32(S10), pp. 112-120.

Dyson, S. J., Van Thielen, B. and Murray, R. C. (2010) 'Chapter 5: The foot and pastern', in Murray, Rachel C. (ed.) *Equine MRI*. Chichester, UK: John Wiley & Sons, Ltd, pp. 147-172.

Edelman, R. R. (2014) 'The history of MR imaging as seen through the pages of radiology', *Radiology*, 273(2 Suppl), pp. S181-200.

Eliashar, E., Dyson, S. J., Archer, R. M., Singer, E. R. and Smith, R. K. W. (2010) 'Two clinical manifestations of desmopathy of the accessory ligament of the deep digital flexor tendon in the hindlimb of 23 horses', *Equine Veterinary Journal*, 37(6), pp. 495-500.

Esaote (2021) *VET MRI Applications for Equine*. Available at: <https://www.esaote.com/veterinary-systems-and-applications/vet-mri-applications/equine/> (Accessed: 1 April 2021).

Evrard, L., Audigié, F., Bertoni, L., Jacquet, S., Denoix, J.-M. and Busoni, V. (2019) 'Low field magnetic resonance imaging of the equine distal interphalangeal joint: Comparison between weight-bearing and non-weight-bearing conditions', *PLOS ONE*, 14(1), p. e0211101.

Fails, A. D. (2020) 'Functional anatomy of the equine musculoskeletal system', in Baxter, G. M. (ed.) *Adams and Stashak's Lameness in Horses*. Wiley, pp. 1-65.

Faramarzi, B., Lantz, L., Lee, D. and Khamas, W. (2017) 'Histological and functional characterizations of the digital cushion in Quarter horses', *Canadian Journal of Veterinary Research*, 81(4), pp. 285-291.

Franci, P., Leece, E. A. and Brearley, J. C. (2006) 'Post anaesthetic myopathy/neuropathy in horses undergoing magnetic resonance imaging compared to horses undergoing surgery', *Equine Veterinary Journal*, 38(6), pp. 497-501.

Fürst, A. E. and Lischer, C. J. (2019) 'Chapter 91 - Foot', in Auer, J. A., Stick, J. A., Kümmerle, J. M., and Prange, T. (eds) *Equine Surgery*. 5th edn. W.B. Saunders, pp. 1543-1587.

Gabriel, A., Jolly, S., Detilleux, J., Dessy-Doize, C., Collin, B. and Reginster, J. Y. (1998) 'Morphometric study of the equine navicular bone: variations with breeds and types of horse and influence of exercise', *Journal of Anatomy*, 193 (Pt 4), pp. 535-549.

Ganesan, A., Alakhras, M., Brennan, P. C. and Mello-Thoms, C. (2018) 'A review of factors influencing radiologists' visual search behaviour', *Journal of Medical Imaging and Radiation Oncology*, 62(6), pp. 747-757.

- Ghazinoor, S., Crues, J. V., 3rd and Crowley, C. (2007) 'Low-field musculoskeletal MRI', *Journal of Magnetic Resonance Imaging*, 25(2), pp. 234-244.
- Gold, G. E., Besier, T. F., Draper, C. E., Asakawa, D. S., Delp, S. L. and Beaupre, G. S. (2004) 'Weight-bearing MRI of patellofemoral joint cartilage contact area', *Journal of Magnetic Resonance Imaging*, 20(3), pp. 526-530.
- Gozalo-Marcilla, M., Redondo, J. I., Johnston, M., Taylor, P. and Bettschart-Wolfensberger, R. (2020) 'CEPEF4: update and plan', *Veterinary Anaesthesia and Analgesia*, 47(5), pp. 724-725.
- Grundmann, I. N. M., Drost, W. T., Zekas, L. J., Belknap, J. K., Garabed, R. B., Weisbrode, S. E., Parks, A. H., Knopp, M. V. and Maierl, J. (2015) 'Quantitative assessment of the equine hoof using digital radiography and magnetic resonance imaging', *Equine Veterinary Journal*, 47(5), pp. 542-547.
- Gutierrez-Nibeyro, S. D., Werpy, N. M. and White, N. A. (2012) 'Standing low-field magnetic resonance imaging in horses with chronic foot pain', *Australian Veterinary Journal*, 90(3), pp. 75-83.
- Gutierrez-Nibeyro, S. D., Werpy, N. M., White, N. A., Mccutcheon, L. J., Weng, H. Y. and Christopher, J. M. (2011) 'Standing low-field magnetic resonance imaging appearance of normal collateral ligaments of the equine distal interphalangeal joint', *Veterinary Radiology & Ultrasound*, 52(5), pp. 521-533.
- Gwet, K. L. (2014) *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring The Extent of Agreement Among Raters*. 4th edn. Advanced Analytics, LLC.
- Hallmarq Veterinary Imaging (2019) 'Above and beyond: over 100,000 horses scanned in over 100 Standing Equine MRI systems'. Available at: <http://hallmarq.net/wp-content/uploads/2019/07/100th-System-and-beyond-PR-310719.pdf>.
- Havsteen, I., Ohlhues, A., Madsen, K. H., Nybing, J. D., Christensen, H. and Christensen, A. (2017) 'Are Movement Artifacts in Magnetic Resonance Imaging a Real Problem?-A Narrative Review', *Frontiers in Neurology*, 8, p. 232.
- Johnson, V. (2011) 'Teleradiology: practicalities and implications', *In Practice*, 33(4), pp. 180-185.
- Johnston, G. M., Eastment, J. K., Wood, J. L. N. and Taylor, P. M. (2002) 'The confidential enquiry into perioperative equine fatalities (CEPEF): mortality results of Phases 1 and 2', *Veterinary Anaesthesia and Analgesia*, 29(4), pp. 159-170.
- Jones, L. E. and Dyson, S. J. (2015) 'Radiographic characterization of ossification of the ungular cartilages in horses: 271 cases (2005-2012)', *Journal of the American Veterinary Medical Association*, 247(7), pp. 801-811.
- Joostens, Z., Evrard, L. and Busoni, V. (2019) 'Unipodal stance influences radiographic evaluation of foot balance in horses', *Veterinary Radiology & Ultrasound*, 60(3), pp. 273-279.

- Kagadis, G. C., Walz-Flannigan, A., Krupinski, E. A., Nagy, P. G., Katsanos, K., Diamantopoulos, A. and Langer, S. G. (2013) 'Medical imaging displays and their use in image interpretation', *RadioGraphics*, 33(1), pp. 275-290.
- de Kerviler, E., Leroy-Willig, A., Clément, O. and Frija, J. (1998) 'Fat suppression techniques in MRI: an update', *Biomedicine & Pharmacotherapy*, 52(2), pp. 69-75.
- Kladny, B., Glückert, K., Swoboda, B., Beyer, W. and Weseloh, G. (1995) 'Comparison of low-field (0.2 Tesla) and high-field (1.5 Tesla) magnetic resonance imaging of the knee joint', *Archives of Orthopaedic and Trauma Surgery*, 114(5), pp. 281-286.
- Kleiter, M., Kneissl, S., Stanek, C., Mayrhofer, E., Baulain, U. and Deegen, E. (1999) 'Evaluation of magnetic resonance imaging techniques in the equine digit', *Veterinary Radiology & Ultrasound*, 40, pp. 15-22.
- Koenig, J. and Cote, N. (2006) 'Equine gastrointestinal motility--ileus and pharmacological modification', *The Canadian Veterinary Journal*, 47(6), pp. 551-559.
- Kottmeier, L. K., Seehusen, F., Helweg, M., Rohn, K., Stadler, P. and Hellige, M. (2020) 'High-field (3 Tesla) MRI of the navicular apparatus of sound horses shows good agreement to histopathology', *Veterinary Radiology & Ultrasound*, 61(1), pp. 48-57.
- Krupinski, E. A. (2010) 'Current perspectives in medical image perception', *Attention, Perception & Psychophysics*, 72(5), pp. 1205-1217.
- Kummer, M., Gygax, D., Lischer, C. and Auer, J. (2009) 'Comparison of the trimming procedure of six different farriers by quantitative evaluation of hoof radiographs', *The Veterinary Journal*, 179(3), pp. 401-406.
- Ladd, M. E., Bachert, P., Meyerspeer, M., Moser, E., Nagel, A. M., Norris, D. G., Schmitter, S., Speck, O., Straub, S. and Zaiss, M. (2018) 'Pros and cons of ultra-high-field MRI/MRS for human application', *Progress in Nuclear Magnetic Resonance Spectroscopy*, 109, pp. 1-50.
- Laurenza, C., Ansart, L. and Portier, K. (2019) 'Risk Factors of Anesthesia-Related Mortality and Morbidity in One Equine Hospital: A Retrospective Study on 1,161 Cases Undergoing Elective or Emergency Surgeries', *Frontiers in Veterinary Science*, 6, p. 514.
- Mair, T. S., Kinns, J., Jones, R. D. and Bolas, N. M. (2003) 'Magnetic resonance imaging of the distal limb of the standing horse: technique and review of 40 cases of foot lameness', *Proceedings of the 49th Annual Convention of the American Association of Equine Practitioners, New Orleans, Louisiana, USA, 21-25 November 2003*, pp. 29-41.
- Mair, T. S., Kinns, J., Jones, R. D. and Bolas, N. M. (2005) 'Magnetic resonance imaging of the distal limb of the standing horse', *Equine Veterinary Education*, 17(2), pp. 74-78.



- Maubon, A. J., Ferru, J.-M., Berger, V., Soulage, M. C., DeGraef, M., Aubas, P., Coupeau, P., Dumont, E. and Rouanet, J.-P. (1999) 'Effect of Field Strength on MR Images: Comparison of the Same Subject at 0.5, 1.0, and 1.5 T', *RadioGraphics*, 19(4), pp. 1057-1067.
- McIlwraith, C. W., Nixon, A. J. and Wright, I. M. (2015) 'Chapter 13 - Bursoscopy', in McIlwraith, C. W., Nixon, A. J., and Wright, I. M. (eds) *Diagnostic and Surgical Arthroscopy in the Horse*. 4th edn. Mosby, pp. 387-406.
- McKnight, A. L., Manduca, A., Felmlee, J. P., Rossman, P. J., McGee, K. P. and Ehman, R. L. (2004) 'Motion-correction techniques for standing equine MRI', *Veterinary Radiology & Ultrasound*, 45(6), pp. 513-519.
- McLellan, J. and Plevin, S. (2019) 'Evaluation of videoendoscopic examinations of arytenoid function in the 2-year-old Thoroughbred: Can we all agree?', *Equine Veterinary Journal*, 51(3), pp. 364-369.
- McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (2017a) 'Acronyms Anonymous I: Spin Echo', in *MRI from Picture to Proton*. Cambridge University Press, pp. 185-206.
- McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (2017b) 'Acronyms Anonymous II: Gradient Echo', in *MRI from Picture to Proton*. Cambridge University Press, pp. 207-224.
- McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (2017c) 'Early Daze: Your First Week in MR', in *MRI from Picture to Proton*. Cambridge University Press, pp. 11-25.
- McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (eds) (2017d) 'Getting in Tune: Resonance and Relaxation', in *MRI from Picture to Proton*. 3rd edn. Cambridge: Cambridge University Press, pp. 124-143.
- McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (eds) (2017e) 'Ghosts in the Machine: Quality Control', in *MRI from Picture to Proton*. 3rd edn. Cambridge: Cambridge University Press, pp. 166-182.
- McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (eds) (2017f) 'Improving Your Image: How to Avoid Artefacts', in *MRI from Picture to Proton*. 3rd edn. Cambridge: Cambridge University Press, pp. 81-101.
- McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (2017g) 'Let's Talk Technical: MR Equipment', in *MRI from Picture to Proton*. Cambridge University Press, pp. 144-165.
- McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (2017h) 'Lost in the Pulse Sequence Jungle?', in *MRI from Picture to Proton*. Cambridge University Press, pp. 41-54.
- McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (2017i) 'Seeing is Believing: Introduction to Image Contrast', in *MRI from Picture to Proton*. Cambridge University Press, pp. 26-40.

McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (2017j) 'Spaced Out: Spatial Encoding', in *MRI from Picture to Proton*. Cambridge University Press, pp. 102-123.

McRobbie, D. W., Moore, E. A., Graves, M. J. and Prince, M. R. (eds) (2017k) 'What You Set is What You Get: Basic Image Optimization', in *MRI from Picture to Proton*. 3rd edn. Cambridge: Cambridge University Press, pp. 67-80.

Mejdell, C. M., Jørgensen, G. H. M., Rehn, T., Fremstad, K., Keeling, L. and Bøe, K. E. (2010) 'Reliability of an injury scoring system for horses', *Acta Veterinaria Scandinavica*, 52, p. 68.

Menzies-Gow, N. J., Stevens, K. B., Sepulveda, M. F., Jarvis, N. and Marr, C. M. (2010) 'Repeatability and reproducibility of the Obel grading system for equine laminitis', *The Veterinary Record*, 167(2), pp. 52-55.

Minitab (2020) *Kappa statistics and Kendall's coefficients*. Available at: <https://support.minitab.com/en-us/minitab/19/help-and-how-to/quality-and-process-improvement/measurement-system-analysis/supporting-topics/attribute-agreement-analysis/kappa-statistics-and-kendall-s-coefficients/> (Accessed: 5 April 2021).

Moreno, K. L., Scallan, E. M., Friedeck, W. O. and Simon, B. T. (2020) 'Transient pelvic limb neuropathy following proximal metatarsal and tarsal magnetic resonance imaging in seven horses', *Equine Veterinary Journal*, 52(3), pp. 359-363.

Murray, R. C., Blunden, T. S., Schramme, M. C. and Dyson, S. J. (2006) 'How does magnetic resonance imaging represent histologic findings in the equine digit?', *Veterinary Radiology & Ultrasound*, 47(1), pp. 17-31.

Murray, R. C., Dyson, S., Branch, M. and Schramme, M. (2007) 'Validation of Magnetic Resonance Imaging Use in Equine Limbs', *Clinical Techniques in Equine Practice*, 6(1), pp. 26-36.

Murray, R. C., Leece, E. and Judy, C. (2010) 'High-field MRI in horses: practicalities and image acquisition', in Murray, R. C. (ed.) *Equine MRI*. Chichester, UK: John Wiley & Sons, Ltd, pp. 39-74.

Murray, R. C., Mair, T. S., Sherlock, C. E. and Blunden, A. S. (2009) 'Comparison of high-field and low-field magnetic resonance images of cadaver limbs of horses', *The Veterinary Record*, 165(10), pp. 281-288.

Murray, R. C., Roberts, B. L., Schramme, M. C., Dyson, S. J. and Branch, M. (2004) 'Quantitative evaluation of equine deep digital flexor tendon morphology using magnetic resonance imaging', *Veterinary Radiology & Ultrasound*, 45(2), pp. 103-111.

Murray, R. C., Schramme, M. C., Dyson, S. J., Branch, M. V. and Blunden, T. S. (2006) 'Magnetic resonance imaging characteristics of the foot in horses with palmar foot pain and control horses', *Veterinary Radiology & Ultrasound*, 47(1), pp. 1-16.

- Murray, R. C. and Werpy, N. (2010) 'Image interpretation and artefacts', in Murray, Rachel C. (ed.) *Equine MRI*. Chichester, UK: John Wiley & Sons, Ltd, pp. 101-145.
- Nelson, J. C. and Pepe, M. S. (2000) 'Statistical description of interrater variability in ordinal ratings', *Statistical methods in Medical Research*, 9(5), pp. 475-496.
- Ng, C. S. and Palmer, C. R. (2007) 'Analysis of diagnostic confidence and diagnostic accuracy: A unified framework', *The British Journal of Radiology*, 80(951), pp. 152-160.
- Norvall, A., Spriet, M., Espinosa, P., Ariño-Estrada, G., Murphy, B. G., Katzman, S. A. and Galuppo, L. D. (2021) 'Chondrosesamoidean ligament enthesopathy: Prevalence and findings in a population of lame horses imaged with positron emission tomography', *Equine Veterinary Journal*, 53(3), pp. 451-459.
- Olive, J. (2010) 'Distal interphalangeal articular cartilage assessment using low-field magnetic resonance imaging', *Veterinary Radiology & Ultrasound*, 51(3), pp. 259-266.
- Park, R. D., Nelson, T. R. and Hoopes, P. J. (1987) 'Magnetic resonance imaging of the normal equine digit and metacarpophalangeal joint', *Veterinary Radiology & Ultrasound*, 28(4), pp. 105-116.
- Pollitt, C. C. (2010) 'The anatomy and physiology of the suspensory apparatus of the distal phalanx', *The Veterinary Clinics of North America: Equine Practice*, 26(1), pp. 29-49.
- Porter, E. G. and Werpy, N. M. (2014) 'New Concepts in Standing Advanced Diagnostic Equine Imaging', *The Veterinary Clinics of North America: Equine Practice*, 30(1), pp. 239-268.
- Powder, S. L., Caserto, B. G., Bowker, R. M., Lin, B., Potter, H. G. and Koff, M. F. (2020) 'Quantitative magnetic resonance imaging and histological hoof wall assessment of 3-year-old Quarter Horses', *Equine Veterinary Journal*, 52(3), pp. 435-440.
- Ross, M. W. (2011) 'Chapter 2 - Lameness in Horses: Basic Facts Before Starting', in Ross, M. W. and Dyson, S. J. (eds) *Diagnosis and Management of Lameness in the Horse*. 2nd edn. Saint Louis: W.B. Saunders, pp. 3-8.
- Rutt, B. K. and Lee, D. H. (1996) 'The impact of field strength on image quality in MRI', *Journal of Magnetic Resonance Imaging*, A Wiley Company, 6(1), pp. 57-62.
- Sampson, S. N., Schneider, R. K., Gavin, P. R., Ho, C. P., Tucker, R. L. and Charles, E. M. (2009) 'Magnetic resonance imaging findings in horses with recent onset navicular syndrome but without radiographic abnormalities', *Veterinary Radiology & Ultrasound*, 50(4), pp. 339-346.
- Sarracanie, M. and Salameh, N. (2020) 'Low-Field MRI: How Low Can We Go? A Fresh View on an Old Debate', *Frontiers in Physics*, 8, p. 172.

- Schade, S. M., Arnoczky, S. P. and Bowker, R. M. (2017) 'The microvasculature in the equine distal phalanx: Implications for fracture healing', *Veterinary and Comparative Orthopaedics and Traumatology*, 27(02), pp. 102-106.
- Schneider, R. K., Gavin, P. R. and Tucker, R. L. (2003) 'What MRI is teaching us about navicular disease', *Proceedings of the 49th Annual Convention of the American Association of Equine Practitioners, New Orleans, Louisiana, USA, 21-25 November 2003*, pp. 210-219.
- Schramme, M. and Segard-Weisse, E. (2020) 'Magnetic resonance imaging', in Baxter, G. M. (ed.) *Adams and Stashak's Lameness in Horses*. Wiley, pp. 387-430.
- Shankar, V. and Bangdiwala, S. I. (2014) 'Observer agreement paradoxes in 2x2 tables: comparison of agreement measures', *BMC Medical Research Methodology*, 14, p. 100.
- Shapiro, L. M. and Gold, G. E. (2012) 'MRI of weight bearing and movement', *Osteoarthritis and Cartilage*, 20(2), pp. 69-78.
- Sherlock, C., Mair, T. and Blunden, T. (2008) 'Deep erosions of the palmar aspect of the navicular bone diagnosed by standing magnetic resonance imaging', *Equine Veterinary Journal*, 40(7), pp. 684-692.
- Singh, B. (ed.) (2018a) 'Chapter 23: The forelimb of the horse', in *Dyce, Sack, and Wensing's Textbook of Veterinary Anatomy*. St. Louis, Missouri: Elsevier, pp. 574-611.
- Singh, B. (ed.) (2018b) 'Chapter 24: The hindlimb of the horse', in *Dyce, Sack, and Wensing's Textbook of Veterinary Anatomy*. St. Louis, Missouri: Elsevier, pp. 612-631.
- Smith, M. A. (2015) 'Optimising the use of MRI in the management of foot lameness in the horse', *Livestock*, 20(5), pp. 290-293.
- Smith, M. A., Dyson, S. J. and Murray, R. C. (2008) 'Is a magic angle effect observed in the collateral ligaments of the distal interphalangeal joint or the oblique sesamoidean ligaments during standing magnetic resonance imaging?', *Veterinary Radiology & Ultrasound*, 49(6), pp. 509-515.
- Smith, M. A., Dyson, S. J. and Murray, R. C. (2012) 'Reliability of high- and low-field magnetic resonance imaging systems for detection of cartilage and bone lesions in the equine cadaver fetlock', *Equine Veterinary Journal*, 44(6), pp. 684-691.
- Smith, M. A., Milmine, R., Dyson, S. J. and Murray, R. C. (2008) 'Optimization of STIR sequences for equine low field standing MRI', in *Proceedings of the European Veterinary Conference Voorjaarsdagen. European Veterinary Conference - Voorjaarsdagen*, p. 323.
- Spriet, M., Mai, W. and McKnight, A. (2007) 'Asymmetric signal intensity in normal collateral ligaments of the distal interphalangeal joint in horses with a low-field MRI system due to the magic angle effect', *Veterinary Radiology & Ultrasound*, 48(2), pp. 95-100.

Spriet, M. and McKnight, A. (2009) 'Characterization of the magic angle effect in the equine deep digital flexor tendon using a low-field magnetic resonance system', *Veterinary Radiology & Ultrasound*, 50(1), pp. 32-36.

Spriet, M. and Zwingenberger, A. (2009) 'Influence of the position of the foot on MRI signal in the deep digital flexor tendon and collateral ligaments of the distal interphalangeal joint in the standing horse', *Equine Veterinary Journal*, 41(5), pp. 498-503.

Stahl, R., Krug, R., Kelley, D. A. C., Zuo, J., Ma, C. B., Majumdar, S. and Link, T. M. (2009) 'Assessment of cartilage-dedicated sequences at ultra-high-field MRI: Comparison of imaging performance and diagnostic confidence between 3.0 and 7.0 T with respect to osteoarthritis-induced changes at the knee joint', *Skeletal Radiology*, 38(8), pp. 771-783.

Stehling, C., Souza, R. B., Hellio Le Graverand, M.-P., Wyman, B. T., Li, X., Majumdar, S. and Link, T. M. (2012) 'Loading of the knee during 3.0T MRI is associated with significantly increased medial meniscus extrusion in mild and moderate osteoarthritis', *European Journal of Radiology*, 81(8), pp. 1839-1845.

Thibault, C. J., Wilson, D. V., Robertson, S. A., Sharma, D. and Kinsley, M. A. (2019) 'A retrospective study of fecal output and postprocedure colic in 246 horses undergoing standing sedation with detomidine, or general anesthesia with or without detomidine', *Veterinary Anaesthesia and Analgesia*, 46(4), pp. 458-465.

Thrall, D. E. (2018) 'Chapter 7 - Introduction to Radiographic Interpretation', in Thrall, D. E. (ed.) *Textbook of Veterinary Diagnostic Radiology*. 7th edn. W.B. Saunders, pp. 110-122.

Tourassi, G., Voisin, S., Paquit, V. and Krupinski, E. (2013) 'Investigating the link between radiologists' gaze, diagnostic decision, and image content', *Journal of the American Medical Informatics Association*, 20(6), pp. 1067-1075.

Vallance, S. A., Bell, R. J. W., Spriet, M., Kass, P. H. and Puchalski, S. M. (2012a) 'Comparisons of computed tomography, contrast enhanced computed tomography and standing low-field magnetic resonance imaging in horses with lameness localised to the foot. Part 1: anatomic visualisation scores', *Equine Veterinary Journal*, 44(1), pp. 51-56.

Vallance, S. A., Bell, R. J. W., Spriet, M., Kass, P. H. and Puchalski, S. M. (2012b) 'Comparisons of computed tomography, contrast-enhanced computed tomography and standing low-field magnetic resonance imaging in horses with lameness localised to the foot. Part 2: Lesion identification', *Equine Veterinary Journal*, 44(2), pp. 149-156.

Watson, P. F. and Petrie, A. (2010) 'Method agreement analysis: a review of correct methodology', *Theriogenology*, 73(9), pp. 1167-1179.

Werpy, N. M. (2007) 'Magnetic Resonance Imaging of the Equine Patient: A Comparison of High- and Low-Field Systems', *Clinical Techniques in Equine Practice*. W.B. Saunders, 6(1), pp. 37-45.

- Werpy, N. M. (2010) 'Low-field MRI in horses: practicalities and image acquisition', in Murray, R. C. (ed.) *Equine MRI*. Chichester, UK: John Wiley & Sons, Ltd, pp. 75-99.
- Werpy, N. M., Ho, C. P. and Kawcak, C. E. (2010) 'Magic angle effect in normal collateral ligaments of the distal interphalangeal joint in horses imaged with a high-field magnetic resonance imaging system', *Veterinary Radiology & Ultrasound*, 51(1), pp. 2-10.
- Werpy, N. M., Ho, C. P., Pease, A. P. and Kawcak, C. E. (2011) 'The effect of sequence selection and field strength on detection of osteochondral defects in the metacarpophalangeal joint', *Veterinary Radiology & Ultrasound*, 52(2), pp. 154-160.
- Westbrook, C., Roth, C. K. and Talbot, J. (eds) (2011a) 'Artefacts and their compensation', in *MRI in Practice*. 4th edn. John Wiley & Sons, pp. 225-260.
- Westbrook, C., Roth, C. K. and Talbot, J. (eds) (2011b) 'Basic principles', in *MRI in Practice*. 4th edn. John Wiley & Sons, pp. 1-20.
- Westbrook, C., Roth, C. K. and Talbot, J. (eds) (2011c) 'Image weighting and contrast', in *MRI in Practice*. 4th edn. John Wiley & Sons, pp. 21-58.
- Westbrook, C., Roth, C. K. and Talbot, J. (eds) (2011d) 'Instrumentation and equipment', in *MRI in Practice*. 4th edn. John Wiley & Sons, pp. 307-340.
- Westbrook, C., Roth, C. K. and Talbot, J. (eds) (2011e) 'Parameters and trade-offs', in *MRI in Practice*. 4th edn. John Wiley & Sons, pp. 103-139.
- Westbrook, C., Roth, C. K. and Talbot, J. (eds) (2011f) 'Pulse sequences', in *MRI in Practice*. 4th edn. John Wiley & Sons, pp. 140-197.
- Whitton, R. C., Buckley, C., Donovan, T., Wales, A. D. and Dennis, R. (1998) 'The diagnosis of lameness associated with distal limb pathology in a horse: A comparison of radiography, computed tomography and magnetic resonance imaging', *The Veterinary Journal*, 155(3), pp. 223-229.
- Williams, L. H. and Drew, T. (2019) 'What do we know about volumetric medical image interpretation?: a review of the basic science and medical image perception literatures', *Cognitive Research: Principles and Implications*, 4(1), p. 21.
- Winter, M. D. (2012) 'The basics of musculoskeletal magnetic resonance imaging: terminology, imaging sequences, image planes, and descriptions of basic pathologic change', *The Veterinary Clinics of North America: Equine Practice*, 28(3), pp. 599-616.
- Wong, S., Steinbach, L., Zhao, J., Stehling, C., Ma, C. B. and Link, T. M. (2009) 'Comparative study of imaging at 3.0 T versus 1.5 T of the knee', *Skeletal Radiology*, 38(8), pp. 761-769.
- Wongpakaran, N., Wongpakaran, T., Wedding, D. and Gwet, K. L. (2013) 'A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater

reliability coefficients: a study conducted with personality disorder samples', *BMC Medical Research Methodology*, 13, p. 61.

Wright, I. M., Kidd, L. and Thorp, B. H. (1998) 'Gross, histological and histomorphometric features of the navicular bone and related structures in the horse', *Equine Veterinary Journal*, 30(3), pp. 220-234.

Wu, C.-C. and Wolfe, J. M. (2019) 'Eye Movements in Medical Image Perception: A Selective Review of Past, Present and Future', *Vision*, 3(2). doi: 10.3390/vision3020032.

van Zadelhoff, C., Schwarz, T., Smith, S., Engerand, A. and Taylor, S. (2020) 'Identification of Naturally Occurring Cartilage Damage in the Equine Distal Interphalangeal Joint Using Low-Field Magnetic Resonance Imaging and Magnetic Resonance Arthrography', *Frontiers in Veterinary Science*, 6, p. 508.

Zaitsev, M., Maclaren, J. and Herbst, M. (2015) 'Motion artifacts in MRI: A complex problem with many partial solutions', *Journal of Magnetic Resonance Imaging*, 42(4), pp. 887-901.