GENERAL ARTICLE

Clinical magnetic resonance image quality of the equine foot is significantly influenced by acquisition system

Christian A. Byrne [iD] | John F. Marshall | Lance C. Voute

Weipers Centre Equine Hospital, School of Veterinary Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK

Correspondence
Christian. A. Byrne, Weipers Centre Equine Hospital, School of Veterinary Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, 464 Bearsden Road, Glasgow, G61 1QH, UK.
Email: cbyrnevet@gmail.com

Abstract

Background: Investigation of image quality in clinical equine magnetic resonance (MR) imaging may optimise diagnostic value.

Objectives: To assess the influence of field strength and anaesthesia on image quality in MR imaging of the equine foot in a clinical context.

Study design: Analytical clinical study.

Methods: Fifteen equine foot studies (five studies per system) were randomly selected from the clinical databases of three MR imaging systems: low-field standing (LF St), low-field anaesthetised (LF GA) and high-field anaesthetised (HF GA). Ten experienced observers graded image quality for entire studies and seven clinically important anatomical structures within the foot (briefly, grade 1: textbook quality, grade 2: high diagnostic quality, grade 3: satisfactory diagnostic quality, grade 4: non-diagnostic). Statistical analysis assessed the effect of anaesthesia and field strength using a combination of the Pearson chi-square test or Fisher's exact test and Mann-Whitney test.

Results: There was no difference in the proportion of entire studies of diagnostic quality between LF St (90%, 95% CI 78%-97%) and LF GA (88%, 76-95%, $P = .7$). No differences were evident in the proportion of diagnostic studies or median image quality gradings between LF St and LF GA when assessing individual anatomical structures (both groups all median grades = 3). There was a statistically significant difference in the proportion of entire studies of diagnostic quality between LF GA and HF GA (100%, 95% CI lower bound 94%, $P = .03$). There were statistically significant differences in median image quality gradings between LF GA (all median grades = 3) and HF GA (median grades = 1 (5/7 structures) or 2 (2/7 structures) for all individual anatomical structures (all $P < .001$). The reasons reported for reduced image quality differed between systems.

Main limitations: Randomised selection of cases from clinical databases. Individual observer preferences may influence image quality assessment.

Conclusions: Field strength is a more important influencer of image quality than anaesthesia for magnetic resonance imaging of the equine foot in clinical patients.

KEYWORDS
horse, magnetic resonance imaging, foot, image quality

1 | INTRODUCTION

Foot pain is a common cause of lameness in the horse and magnetic resonance (MR) imaging of the foot is now a fundamental diagnostic method,[1,2] given the limitations of other imaging modalities in this anatomical region.[3-7] Understanding the factors that influence MR image quality in a clinical context can optimise the use of this modality. Increasing magnetic field strength effectively results in a linear increase in signal-to-noise ratio which significantly contributes to perceived image quality.[8,9] The effect of magnetic field strength on the identification of pathology in the equine cadaver foot has been investigated, demonstrating the diagnostic value of low-field MR imaging of this region.[10,11] Similar findings have been established in human orthopaedic MR imaging.[12] Motion is also a critical influencer of MR image quality and is reported to be more significant during imaging of the standing equine patient compared to a patient under general anaesthesia.[8,13] The position of the foot against the ground surface makes it less susceptible to pendulous sway motion than more proximal regions of the limb.[2,14] Careful patient management, optimal sequence selection, study planning and motion-correction techniques can minimise the impact of subtle motion.[8,14] However, other factors such as weight-bearing of the imaged limb can also influence the diagnostic value of MR images for structures such as the articular cartilage of the distal interphalangeal (DIP) joint.[15]

The application of previous literature to the clinical context is limited by the use of cadaver materials, investigation of particular anatomical regions, use of comparable sequences rather than those optimised for each system, deviation from the practical use of MR imaging systems and methods of image viewing.[10,11,14,15] There is little evidence describing the differences in perceived image quality between MR imaging systems for foot studies performed in a clinical context in live equine patients.[8,16] The aim of our study was to assess the influence of field strength and anaesthesia on subjective image quality assessment in studies of the equine foot performed in a clinical context. We hypothesised that: (a) there would be a difference in perceived image quality between MR images from low-field systems acquired with the patient standing compared to with the patient under general anaesthesia; (b) there would be a difference in perceived image quality between MR images acquired from low-field and high-field systems with the patient under general anaesthesia.

2 | MATERIALS AND METHODS

2.1 | MR imaging studies

MR imaging studies from clinical patients were selected from the databases of three referral institutions with different MR imaging systems:

1. LF St: Low-field (0.3T), weight-bearing (standing, sedated) open magnet system (EQ2 standing equine MRI, Hallmarq Veterinary Imaging Ltd)

2. LF GA: Low-field (0.3T), non-weight-bearing (anaesthetised) open magnet system (O-scan Equine, Esaote UK)

3. HF GA: High-field (1.5T) non-weight-bearing (anaesthetised) closed magnet system (Symphony, Siemens Medical Solutions USA, Inc.)

Using these three groups the study design was formulated to assess the effect of anaesthesia (LF St compared to LF GA) and field strength (LF GA compared to HF GA). Five foot studies were selected from the clinical databases from each of the three institutions using a random number generator (Random Integer Generator: www.random.org/integers/). MR imaging studies acquired between June 2015 and August 2018 were included in the search. The minimum sampling unit for the purposes of study selection was the individual foot. The shortlisted database studies included all the sequences that would typically be performed as a routine equine foot MR imaging study at that institution. These included MR imaging studies that were a repetition of previously imaged patients and studies of contralateral (ie non-lame) limbs. Targeted MR imaging studies with a reduced number of sequences (for example for pre-surgical planning or assessment of an acute foot penetrating injury) were not included. The patient details were anonymised by the acquiring organisation prior to submission to the primary author. Additional metadata attributes were modified to remove information that identified the acquiring institution using a DICOM anonymisation tool (DICOM Anonymizer: www.dicomanonymizer.com) and DICOM viewing software (Horos, Horos Project (version 2.2.0)). Metadata outlining the parameters used in the pulse sequences were unaltered. The studies were randomly ordered using a random sequence generator (Random Sequence Generator: www.random.org/sequences/) and labelled appropriately with their assigned case number in the Patient's Name and Patient ID metadata attributes. The finalised DICOM files were exported to individual study folders for distribution to observers.

2.2 | Image assessment platform

An online image assessment platform was developed to allow observers to evaluate MR imaging studies of the equine foot. The first component was a subjective assessment of image quality for the whole study, using a 4-point grading scale (briefly, grade 1: textbook quality, grade 2: high diagnostic quality, grade 3: satisfactory diagnostic quality, grade 4: non-diagnostic). Verbal descriptors were provided for each grade (see Data S1).

The second component focused on assessment of individual anatomical structures of the foot that are deemed to be clinically relevant in the investigation of lameness. Seven structures were selected based on previous literature reporting lesion distribution during MR imaging for the investigation of foot lameness.[1,17-21] The structures included were the deep digital flexor tendon, navicular bone, navicular bursa, DIP joint, collateral ligaments of the DIP joint, the third phalanx and the distal sesamoidean impar ligament. A version of the

4-point image quality grading scale was produced for each structure, with the verbal descriptors modified to reflect the structure's specific MR imaging features (see Data S1). Pathology within individual structures was assessed using a 4-point scale (grade 1: no pathology, grade 2: mild pathology, grade 3: moderate pathology, grade 4: severe pathology). Assessment of pathology was performed to quantify the relative degree of pathology in each group, rather than to test the pathology identification ability of the different acquisition systems. A field was also available at the end of each section for free text comments. The observers were not given specific instructions to dictate the content of free text responses. The image assessment platform was hosted using an online survey tool (Online Surveys, Jisc) and can be found in Data S2. The online assessment platform was distributed with the labelled DICOM files of the 15 studies for interpretation. Each observer accessed the image assessment platform using a unique link that allowed individual results to be recorded. The observers were able to assess the studies with their DICOM viewer of choice and manipulate the images as they would in a clinical context.

2.3 | Observers

The platform was distributed to 10 experienced observers. Selection criteria for observers included diploma or associate status in one or more relevant fields (Diagnostic Imaging: Diplomate or Associate Member of the European College of Veterinary Diagnostic Imaging or Diplomate of the American College of Veterinary Radiology, 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) and at least 5 years' experience interpreting equine MR images. Observers were recruited through direct individual contact or through a collaboration with an equine teleradiology service (VetCT: https://www.vet-ct.com/gb/). A short observer profile section was included at the outset of the image assessment platform. This documented the observers' qualifications, duration of equine MR imaging experience and their specific experience with different imaging systems. The observers also reported the DICOM viewing software they would use to participate in the study. The image assessment platform was formatted to allow observers to save their progress and return to complete the assessment of the 15 studies at their own convenience.

2.4 | Data analysis

The image assessment platform data were compiled into a database to allow graphical evaluation and descriptive statistical analysis (Microsoft Office Excel 2016, Microsoft Corporation). The data were ordered into frequency tables where appropriate, including for the observer profile data. For image quality assessment of whole studies and individual anatomical structures, contingency tables were formed with study gradings grouped as diagnostic (grades 1-3) and non-diagnostic (grade 4). Analysis of ordinal data output from the image assessment platform included calculation of the median and range.

Further statistical analysis was performed using statistical software (Minitab version 18.1, Minitab Ltd.). $P < .05$ were considered significant. Confidence intervals (95%) were determined for the proportions of diagnostic studies and better than high diagnostic quality studies using an exact method. Comparisons were made to assess the effect of anaesthesia (LF St vs LF GA) and field strength (LF GA vs HF GA). The proportion of diagnostic and non-diagnostic gradings was compared for whole studies and individual anatomical structures using a Pearson chi-square test or Fisher's exact test (if one or more expected cell values ≤5). The ranked gradings for image quality and pathology were compared using a Mann-Whitney test. Inter-observer agreement for image quality grading was assessed using Fleiss' Kappa and Kendall's coefficient of concordance for entire studies and individual anatomical structures. Agreement analysis included studies from all acquisition systems and gradings by all observers. Interpretation of kappa values was based on previously suggested descriptors (<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).[22] No arbitrary descriptors are reported for Kendall's coefficient of concordance but higher values indicate stronger agreement between gradings.[23]

Free text comments were collated for each study. The comments were categorised as containing information regarding image quality or pathology. Comments regarding image quality were categorised and combined into summary tables. Reasons for reduced image quality were categorised as: alignment or position, additional sequence desired, short tau inversion recovery (STIR) fat suppression, sequence parameter (other), motion artefact, magic angle effect, artefact (other), repeat sequence desired and other. Where a comment referred to multiple categories, this was noted as a count in each category. The pulse sequence data for MR imaging studies were also collected into a database and descriptive statistical analysis was performed including the median and range for number of sequences (including multiplanar reconstructions) per study for each group.

3 | RESULTS

3.1 | Observers and MR imaging studies

All observers had >5 years' experience interpreting equine MR images, with 6/10 having >10 years' experience. There were six diagnostic imaging Diplomates, one diagnostic imaging Associate Member, one diagnostic imaging and surgical Diplomate, one diagnostic imaging, surgical and sports medicine and rehabilitation Diplomate and one surgical and sports medicine and rehabilitation Diplomate. When reporting the frequency with which they

interpret images from different systems, observers were most familiar with low-field standing images (regularly: 6/10, frequently: 2/10, occasionally: 0/10, rarely or never: 2/10) than 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 median (range) number of sequences for each group was: LF St 10 (9-11), LF GA 12 (11-14) and HF GA 11 (11-12). These values do not include localiser or pilot sequences but include multiplanar reconstructions that formed part of the routine imaging protocol. The pulse sequence parameters varied between acquisition systems and studies. All studies contained sequences acquired in at least three planes. In addition, all studies contained T1 weighted, T2 weighted, proton density weighted and STIR images. The pulse sequence parameters of all studies are presented in Data S3.

3.2 | Low-field standing vs low-field under general anaesthesia

When assessing the image quality of entire studies, 90% (95% CI 78%, 97%) of the gradings were classified as diagnostic for LF St and 88% (95% CI 76%, 95%) for LF GA. Entire study gradings classified as better than "high diagnostic quality" (ie grades 1 and 2), accounted for 18% (95% CI 9%, 31%) of gradings for LF St and 34% (95% CI 21%, 49%) for LF GA. The distribution of image quality gradings is displayed in Figure 1. The median and range of image quality gradings for entire studies and individual anatomical structures of the LF St and LF GA groups are presented in Table 1. When comparing the proportion of diagnostic vs non-diagnostic studies between LF St and LF GA groups, there were no statistically significant differences

for entire studies or individual anatomical structures of the foot. There were no statistically significant differences in median image quality gradings between LF St and LF GA. The output of these comparisons is presented in Table 1. The distribution of pathology gradings is displayed in Figure 2. There were no significant differences in median pathology scores between LF St and LF GA. The LF St group had 19 free text comments that referred to image quality and 15 that referred to pathology. The LF GA group had 23 comments that referred to image quality and 19 that referred to pathology. The distribution of comments in the image quality categories is displayed in Figure 3.

3.3 | Low-field under general anaesthesia vs high-field under general anaesthesia

When assessing the image quality of entire studies, 88% (95% CI 76%, 95%) of the gradings were classified as diagnostic for LF GA and 100% (95% CI lower bound 94%) for HF GA. Entire study gradings classified as better than "high diagnostic quality" accounted for 34% (95% CI 21%, 49%) of gradings for LF GA and 96% (95% CI 86%, 100%) of gradings for HF GA. The distribution of image quality gradings is displayed in Figure 4. The median and range of image quality gradings for entire studies and individual anatomical structures of the LF GA and HF GA groups are presented in Table 2. There was a statistically significant difference in the proportion of diagnostic vs non-diagnostic studies between LF GA and HF GA for entire studies (P = .03), the deep digital flexor tendon (P = .003), navicular bursa (P = .01), collateral ligaments of the DIP joint (P = .006) and the distal sesamoidean impar ligament (P < .001). There were statistically significant differences in median image quality gradings between LF GA and HF GA for all individual anatomical structures. The output
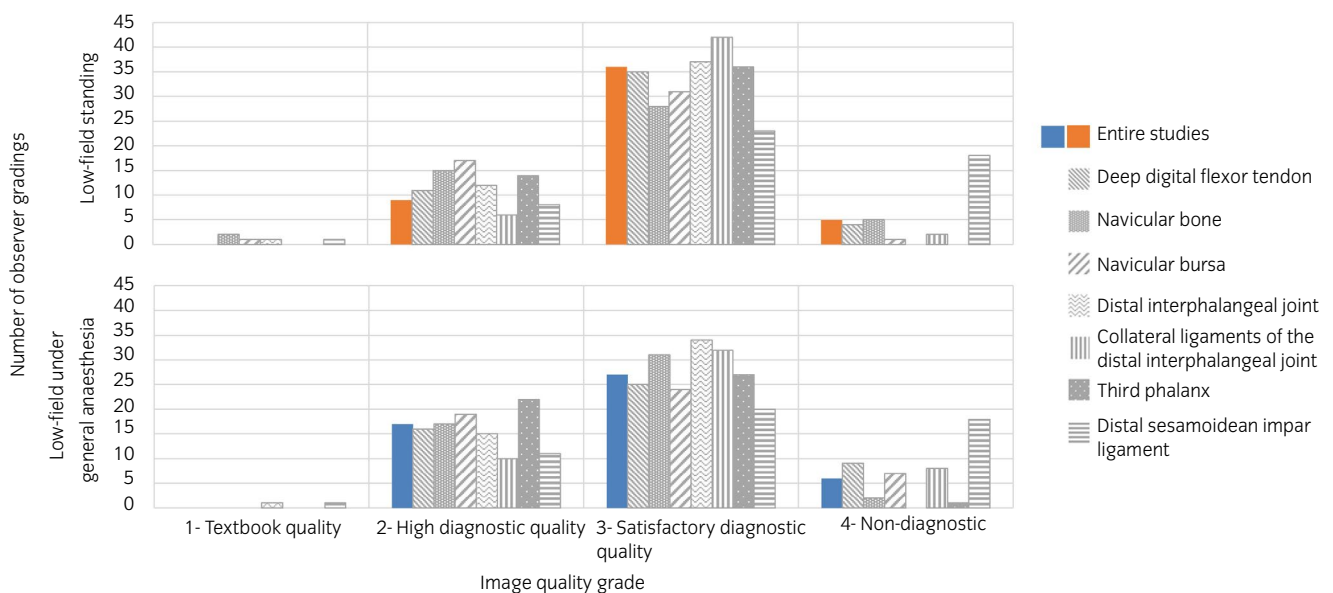

FIGURE 1 Bar chart displaying the image quality gradings for equine foot MR imaging studies from the low-field standing and low-field under general anaesthesia groups

TABLE 1 Key descriptive statistics, comparison of proportion of diagnostic gradings and comparison of the ranked gradings for magnetic resonance image quality for low-field standing and low-field under general anaesthesia groups

| Assessment category | LF St | | LF GA | | Comparison of LF St and LF GA groups | |
	Median grade	Range	Median grade	Range	Proportion of diagnostic gradings (*P* value)	Ranked gradings (*P* value)
Entire study	3	2-4	3	2-4	.7	.2
Deep digital flexor tendon	3	2-4	3	2-4	.1	.9
Navicular bone	3	1-4	3	2-4	.4[a]	.8
Navicular bursa	3	1-4	3	2-4	.06[a]	.5
DIP joint	3	1-3	3	1-3	Not performed[b]	.5
Collateral ligaments of the DIP joint	3	2-4	3	2-4	.09[a]	.8
Third phalanx	3	2-3	3	2-4	>.9[a]	.1
Distal sesamoidean impar ligament	3	1-4	3	1-4	>.9	.7

Note: Grading: Briefly, grade 1: textbook quality, grade 2: high diagnostic quality, grade 3: satisfactory diagnostic quality, grade 4: non-diagnostic. Please see Data S1 for complete grading scales and verbal descriptors.

Abbreviations: DIP, distal interphalangeal joint; LF GA, low-field under general anaesthesia; LF St, low-field standing.

[a]Indicates use of Fisher's exact test. All other for comparisons of proportion of diagnostic studies values are the result of chi-square tests.

[b]Analysis not performed due to 0 value(s) in contingency table (all assessments were graded as diagnostic).
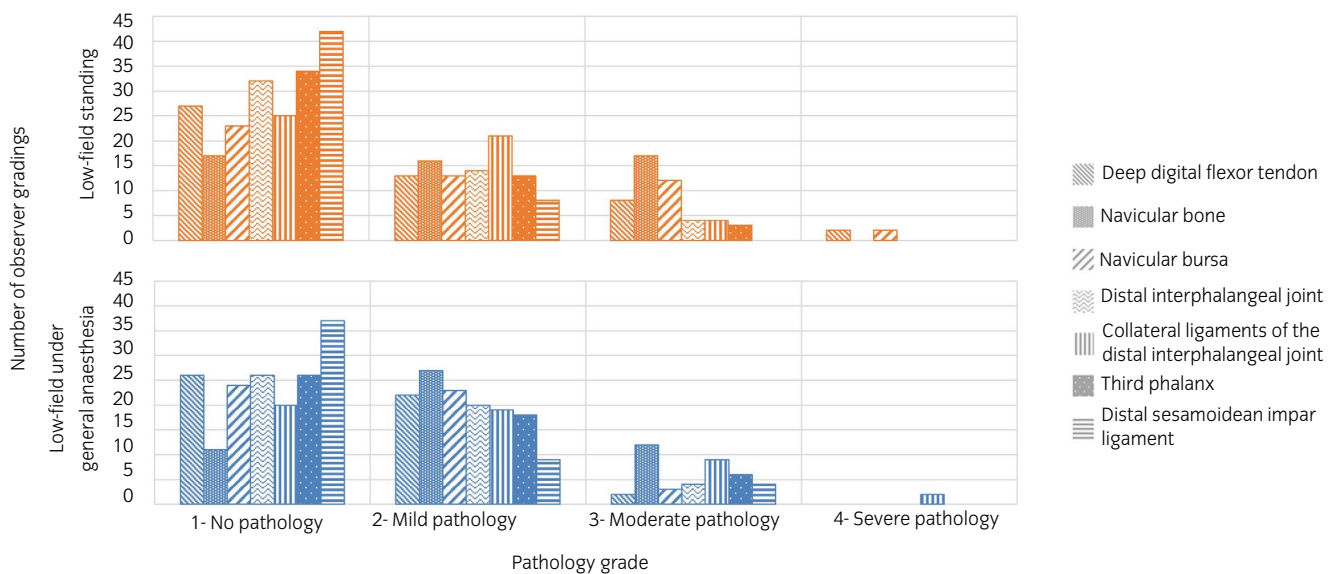
FIGURE 2 Bar chart displaying the pathology gradings for equine foot MR imaging studies from the low-field standing and low-field under general anaesthesia groups

of image quality comparisons between the LF GA and HF GA groups is presented in Table 2. Median pathology scores were significantly different for the deep digital flexor tendon (*P* < .001), navicular bursa (*P* < .001), DIP joint collateral ligaments (*P* .001) and distal sesamoidean impar ligament (*P* .002) for HF GA compared to LF GA. The distribution of pathology gradings is displayed in Figure 5. The LF GA group had 23 comments related to image quality and 19 related to pathology. The HF GA group had 8 comments related to image quality and 20 that related to pathology. The distribution of comments in the image quality categories is displayed in Figure 6.

3.4 | Inter-observer agreement

The results of inter-observer agreement analysis for image quality assessment are presented in Table 3. Absolute inter-observer agreement, as indicated by Fleiss' Kappa, was fair for entire studies. Absolute agreement varied between individual anatomical structures but was generally poor to fair. Inter-observer agreement accounting for relative of order of grading, as indicated by Kendall's coefficient of concordance, was moderate to high for entire studies and individual anatomical structures.
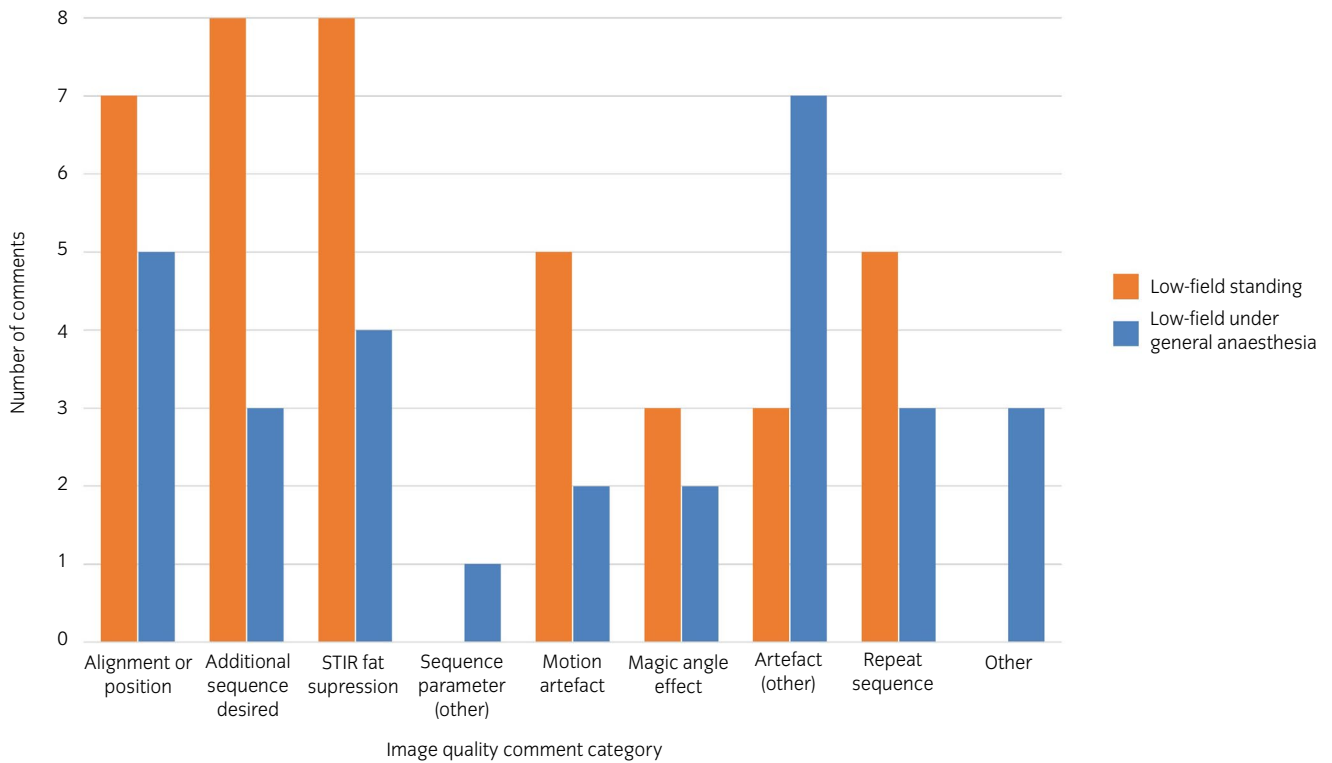
FIGURE 3 Bar chart displaying the categorisation of image quality comments for equine foot MR imaging studies from the low-field standing and low-field under general anaesthesia groups
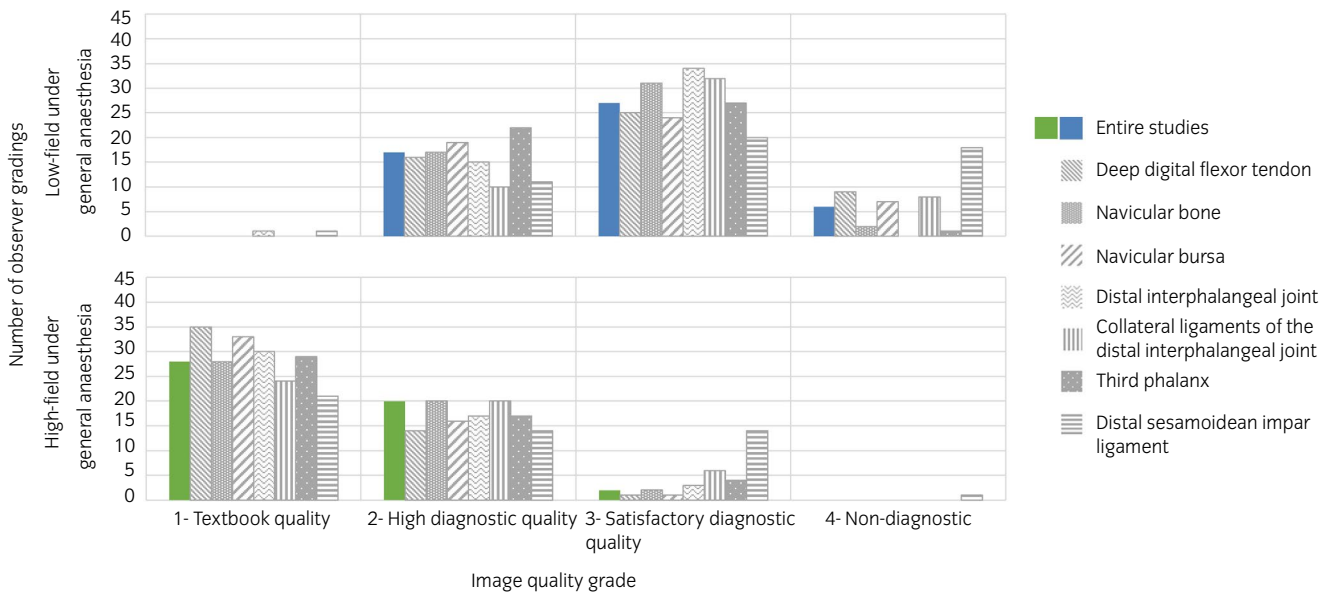

FIGURE 4 Bar chart displaying the image quality gradings for equine foot MR imaging studies from the low-field under general anaesthesia and high-field under general anaesthesia groups

4 | DISCUSSION

This study demonstrates that most MR imaging foot studies of live patients were deemed to be of diagnostic quality by experienced observers, regardless of acquisition system. There was no evidence of a significant difference in the perceived image quality between the LF St and LF GA groups, though the reasons described for reduced image quality appear to differ between groups. Field strength influenced perceived image quality for entire studies and individual anatomical structures of the foot. Many of the factors that reduced image quality could be influenced by the system operator.

4.1 | The effect of general anaesthesia

Our study indicates that there are minimal differences in perceived image quality between low-field MR imaging of the foot performed in the standing patient and with the patient under general anaesthesia. However, the reasons for reduced image quality may differ between groups. Motion is frequently cited as a reason for reduced image quality, particularly in the standing patient.[13] The image quality comments in the current study demonstrated that motion artefact was mentioned in reference to a small number of studies in the LF St and the LF GA groups. Our results indicate that impact of motion is effectively limited during standing imaging of the foot in clinical practice by patient management, repetition of movement affected sequences and motion-correction techniques (including motion insensitive sequences).[8,14,24] Given its position on the ground surface, the foot is less susceptible to patient sway movement than proximal regions of the limb.[14] It is likely that our findings cannot be extrapolated beyond the foot, therefore, other clinically important structures more susceptible to motion, such as the proximal metatarsal region, may warrant specific investigation in a similar manner.

Weight-bearing has been reported to influence the appearance of articular structures and the apparent sites of loading during MR imaging of human limbs.[25,26] An experimental equine study reported improved DIP joint articular cartilage visualisation in unloaded cadaver limbs when compared to images from live weight-bearing patients.[15] The LF St and LF GA groups of the current study are an analogous comparison, but there was no evidence of a statistically significant difference in the perceived image quality for the DIP joint between the two groups. This may reflect the broader image quality assessment involved in the current study, compared to the focused measurement of cartilage thickness and delineation in the previous report.[15] Other literature assessing identification of DIP articular

cartilage lesions primarily relates to non-weight-bearing limbs.[27,28] Further research is required to determine the significance of weight-bearing and limb positioning on the identification of DIP joint articular pathology. The current literature indicates that MR imaging under general anaesthesia (ideally high-field) may be warranted where subtle DIP joint articular cartilage lesions are suspected.[2]

4.2 | The effect of field strength

The impact of field strength on image quality has received attention in the equine veterinary literature, which has supported the use of low-field imaging in a clinical setting.[10,11] The results of our study indicated that high-field studies were more likely to be deemed diagnostic. The improved image quality is a result of the approximately linear relationship between field strength and the signal-to-noise ratio of images.[29] Previous studies assessing the influence of magnetic field strength have compared systems for standing and anaesthetised patients, which resulted in differences in the appearance of some (primarily soft tissue) structures and their anatomical positioning.[10,11] This confounding factor also may contribute to perceived image quality. The LF GA and HF GA groups of the current study had comparable positioning, reducing the influence of this factor.[30]

When considering individual anatomical structures there was no evidence of a statistically significant difference between LF GA and HF GA groups in the proportion of diagnostic studies for the navicular bone, DIP joint and the third phalanx. This may indicate that these structures are more readily assessed by observers and therefore a lower threshold of image quality is required to achieve a diagnostic study.[31,32] In addition, the importance of individual sequences differs between structures.[27,32–35] There may be inter-sequence variation in the relative disparity of image quality between comparable sequences from low-field and high-field systems. The findings

TABLE 2 Key descriptive statistics, comparison of proportion of diagnostic gradings and comparison of the ranked gradings for magnetic resonance image quality for low-field under general anaesthesia and high-field under general anaesthesia groups

Assessment category	LF GA Median grade	Range	HF GA Median grade	Range	Comparison of LF GA and HF GA groups Proportion of diagnostic gradings (P value)	Ranked gradings (P value)
Entire study	3	2-4	1	1-3	.03[a]	<.001
Deep digital flexor tendon	3	2-4	1	1-3	.003[a]	<.001
Navicular bone	3	2-4	1	1-3	.5[a]	<.001
Navicular bursa	3	2-4	1	1-3	.01[a]	<.001
DIP joint	3	1-3	1	1-3	Not performed[b]	<.001
Collateral ligaments of the DIP joint	3	2-4	2	1-3	.006[a]	<.001
Third phalanx	3	2-4	1	1-3	>.9[a]	<.001
Distal sesamoidean impar ligament	3	1-4	2	1-4	<.001	<.001

Note: Grading: Briefly, grade 1: textbook quality, grade 2: high diagnostic quality, grade 3: satisfactory diagnostic quality, grade 4: non-diagnostic. Please see Data S1 for complete grading scales and verbal descriptors.

Abbreviations: DIP, distal interphalangeal joint; HF GA, high-field under general anaesthesia; LF GA, low-field under general anaesthesia.

[a]Indicates use of Fisher's exact test. All other for comparisons of proportion of diagnostic studies values are the result of chi-square tests.

[b]Analysis not performed due to 0 value(s) in contingency table (all assessments were graded as diagnostic).
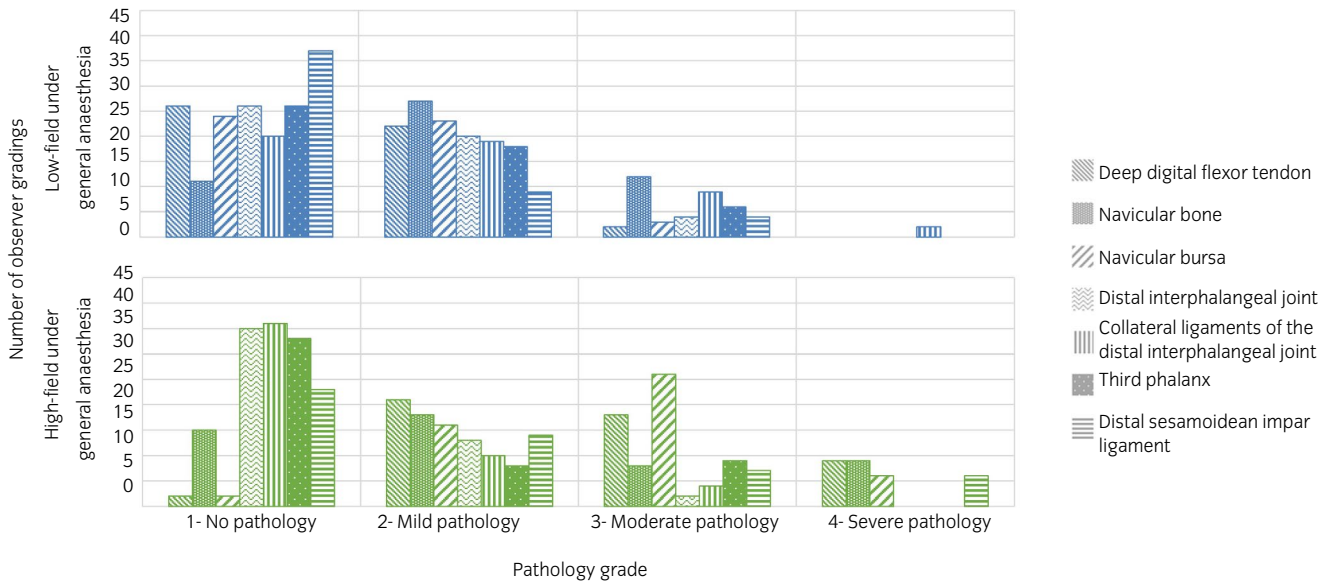
FIGURE 5 Bar chart displaying the pathology gradings for equine foot MR imaging studies from the low-field under general anaesthesia and high-field under general anaesthesia groups
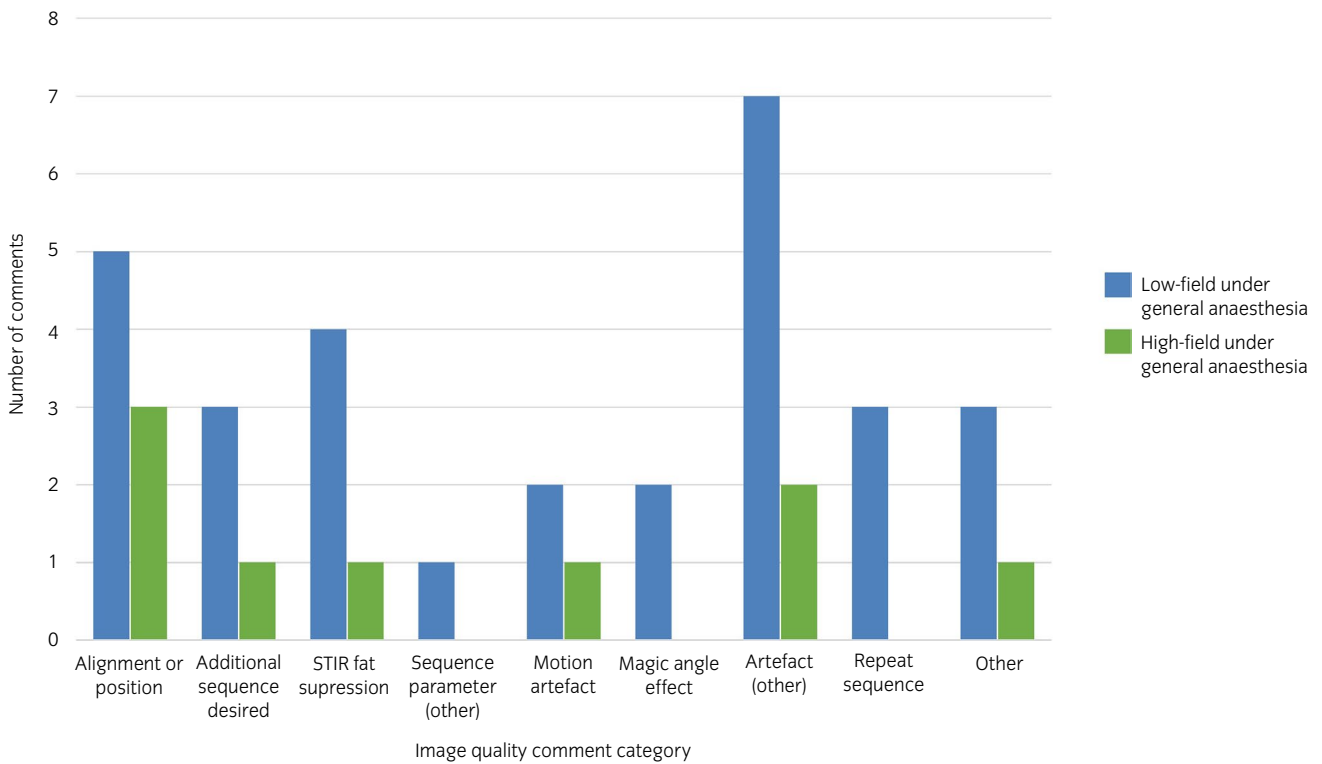

FIGURE 6 Bar chart displaying the categorisation of image quality comments for equine foot MR imaging studies from the low-field under general anaesthesia and high-field under general anaesthesia groups

are broadly consistent with a previous experimental image quality study using cadaver limbs.[11] However, the third phalanx received a relatively poor score in the low-field group (compared to high-field images) of the previous study. This was attributed to loss of signal in the distal aspect of the bone[11] due to positioning of the toe at the periphery of the magnet and radiofrequency coil in the low-field system.[6,11] This may not have been replicated by the magnet and

radiofrequency coil (a human knee coil) configuration used in the LF GA group of the current study.

The findings of this study relate to image quality, rather than pathology identification ability. Intuitively, these factors are intrinsically associated (though this is not necessarily a linear correlation). Comparison of ranked image quality gradings of the current study, which is a more refined indicator than the proportion of diagnostic

TABLE 3 Output of inter-observer agreement analysis for magnetic resonance image quality grading by all observers for all acquisition system groups

Assessment category	Fleiss' kappa		Kendall's coefficient of concordance
	Kappa statistic	Standard error	
Entire study	0.23	0.025	0.68
Deep digital flexor tendon	0.30	0.025	0.75
Navicular bone	0.18	0.026	0.62
Navicular bursa	0.21	0.025	0.73
DIP joint	0.23	0.028	0.62
Collateral ligaments of the DIP joint	0.24	0.025	0.74
Third phalanx	0.18	0.028	0.59
Distal sesamoidean impar ligament	0.06	0.023	0.56

Abbreviation: DIP, distal interphalangeal joint.

studies, demonstrated a significant difference in perceived image quality for all anatomical structures between the LF GA and HF GA groups. This supports the conclusions of previous literature that high-field and low-field images are generally considered diagnostic but high-field imaging affords more detailed assessment, particularly for small anatomical structures or lesions.[10,29] Similar findings have also been reported in human orthopaedic MR imaging.[12] For structures such as the distal sesamoidean impar ligament, evaluation of adjacent tissues may also contribute to the assessment of pathology in a clinical setting.[10] While studies comparing MR imaging findings with histology (as a gold standard) have confirmed the value of this modality for identification of pathology in the equine foot,[31,36–39] the relative diagnostic accuracy of MR imaging systems of different field strengths in clinical patients is yet to be elucidated.

4.3 | General considerations

The influence of other inherent features of the acquisition systems should also be considered. The availability of a radiofrequency coil which is specific to the anatomy of interest will assist in optimising signal-to-noise ratio of the images.[24,40] The equine foot presents a relatively unique shape, therefore, some systems use an equine specific coil rather than a human extremity coil for imaging of this region.[40] The orientation of the static magnetic field relative to the long axis of the limb is an intrinsic feature of the system and this can influence the susceptibility of structures (most commonly ligaments and tendons) to the magic angle effect.[8,10,16,41–45] The magic angle effect was reported to have reduced image quality in studies from both LF St and LF GA groups of the current study. While sequence selection can assist the observer in identifying magic angle effect,[46]

patient positioning has an influence on its occurrence. In the anaesthetised patient, positioning is partially dictated by practicalities of recumbency, the system design and by the skills of the system operator.[30] In the standing patient, leaning of the imaged limb is an additional factor to consider.[45,47,48] In some circumstances (particularly for high-field systems) the magic angle can assist in diagnostic assessment, such as for evaluation of the tendon-bone interface at the insertion of the deep digital flexor tendon.[8,11] Knowledge of the structures susceptible to magic angle effect for a specific MR acquisition system is important for observers interpreting equine MR imaging foot studies clinically.

Several other factors were also reported to reduce perceived image quality across groups in the current study including patient positioning, slice alignment, unsatisfactory fat suppression of STIR sequences and the presence of artefacts. The frequency of these varied between groups. Inferential statistical analysis was not performed on this component of the study because the comments were provided at the discretion of the observers. However, they indicate that operator (for example slice alignment) and institutional (for example standard imaging protocols) factors are important reasons for reduced image quality in a clinical context. The practical aspects of different acquisition systems will influence the relative importance of each factor. Further research with observers explicitly reporting the reasons for their assessment of image quality would be valuable to guide targeted training of operators for their specific acquisition system.

Previous image quality studies have reported incomplete fat suppression of STIR images from both low-field and high-field systems but have highlighted that this can be complicated by the temperature of the cadaver limbs used.[11,49,50] Unsatisfactory fat suppression of STIR images was reported in live patients from all groups of the current study, demonstrating that this does reduce image quality in the clinical scenario. In some circumstances incomplete fat suppression can be addressed by the system operator or its effect minimised by ensuring that the relevant region of anatomy is in the isocentre of the magnetic field where inversion recovery fat suppression is optimal.[24]

Factors other than image quality also influence whether MR imaging is performed with the patient standing or under general anaesthesia. These may include avoiding the risks of general anaesthesia, acquisition time, system availability, economic factors, suspected pathology, proposed concurrent treatment/surgery and patient compliance. Previous retrospective studies indicate that morbidity and mortality of general anaesthesia for MR imaging is very low, especially at institutions with experience performing this procedure.[30,51,52] There may be a greater risk of post-anaesthetic myopathy/neuropathy syndrome when imaging more challenging regions, such as the proximal hindlimb where limb traction may be required in some systems.[53] The regions imaged most commonly in a clinical setting (primarily the foot, pastern and fetlock) can usually be positioned in a straightforward manner.[30] Much of the previous literature relates to high-field MR imaging systems[51-53] which may present greater challenges for positioning (due to equipment

requirements and practicalities of the bore) and anaesthesia (due to reduced patient access and requirement for MRI compatible equipment) when compared to low-field imaging under general anaesthesia.[30] There is limited evidence quantifying the morbidity and mortality of low-field MR imaging in both the anaesthetised and standing, sedated equine patient.

When assessing inter-observer agreement with Kendall's coefficient of concordance, the inter-observer agreement was moderate to high, though absolute inter-observer agreement (indicated by Fleiss' kappa) was poor to fair.[22] Kendall's coefficient of concordance accounts for the relative of order of grading and is the most useful indicator of agreement in this study since image quality assessment is both subjective and complex, despite the use of a grading scale.[8,23,54] Inter-observer agreement was considered satisfactory for the purposes of image quality assessment. The variation in inter-observer agreement between structures may reflect the relative ease of MR image assessment for different anatomical regions.[31,32]

4.4 | Limitations

The studies contained the sequences of the standard foot imaging protocol of the acquisition institution and were deemed to be of diagnostic quality during their clinical acquisition. However, there is likely to be some variation in the quality of individual studies from the same acquisition system, which may be influenced by patient and system operator factors.

A grading scale guided image evaluation, however, individual observers may have individual preferences in imaging protocols for the equine foot. This may influence their judgment of MR studies and their value for assessment of individual structures. Nonetheless, observers were able to use their DICOM viewer of choice to allow their preferred image viewing and manipulation. Interventions were taken to blind the observers to the acquisition system of each study. However, given the experience of the observers in equine MR imaging, it is likely that they would be able to speculate on the acquisition system based on the appearance of the images alone. This is an inevitable consequence of observer-based assessment but could introduce bias due to individual observer experience and preferences. The observer profile demonstrated that the group most frequently interpreted images from low-field standing systems, which is likely a direct consequence of the preponderance of these systems in equine practice.[55,56] This familiarity with images acquired from low-field standing acquisition systems could influence the perception of the relative diagnostic quality of low-field images (particularly from standing acquisition systems).

Some practical considerations are inherent in a study of this nature. Observers reviewed 15 complete MR imaging studies and assessed 7 anatomical structures individually, which represents a significant time commitment. Previous studies in the equine literature using a similar methodology to investigate MR image quality were used as a guide for sample size. Considering a single observation to be one observer assessing one complete MR imaging study,

previous literature used 22-30 observations per acquisition system.[10,11] It was deemed that the clinically important effect size for MR image quality is large, since a small increase in image quality is unlikely to make a clinically relevant change in pathology identification or diagnostic confidence. Therefore, the 50 observations per acquisition system of our study was deemed to be sufficient to identify a clinically relevant difference in image quality, while still being achievable within the practical limitations. Previous literature used cadaver limbs and fewer observers (2-3), which increases the potential impact of observer bias. While there will be some intra-system variation in image quality, prior experience and literature indicates this is relatively small (especially for studies deemed to be sufficiently diagnostic in a clinical context at acquisition) compared to the more clinically relevant inter-system variation. Considering these factors, our study aimed to minimise the influence of individual observer bias by incorporating 10 experienced observers who evaluated a relatively smaller number of studies per acquisition system.

Assessment of pathology by the observers quantified the pathology of each study and demonstrated some statistically significant differences between the LF GA and HF GA groups. The clinical significance of this difference in pathology and its influence on our conclusions regarding image quality is deemed to be minimal.

4.5 | Conclusions

Field strength is a more important influencer of image quality than anaesthesia for magnetic resonance imaging of the equine foot in clinical patients. Observers deemed the majority of clinical MR imaging foot studies to be of diagnostic quality, regardless of acquisition system. The reasons described for reduced image quality appear to differ between acquisition systems. Many of the factors related to reduced image quality can be influenced by the system operator and individuals establishing imaging protocols. Further research focusing on specific factors that reduced image quality of individual studies would be valuable to guide targeted training of operators for their acquisition system.

ETHICAL ANIMAL RESEARCH

The study was approved by the School of Veterinary Medicine Research Ethics Committee, University of Glasgow (Ref 08a/18).

OWNER INFORMED CONSENT

Explicit owner consent was not requested for the retrospective use of anonymised magnetic resonance images from clinical cases. Images were anonymised at the acquiring institution prior to submission to the primary author.

DATA ACCESSIBILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

All authors contributed to study design. Data collection and analysis were performed by C. A. Byrne. C. A. Byrne was the lead author, but all authors contributed to and approved the final manuscript.

CONFLICT OF INTERESTS

No competing interests have been declared.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/evj.13330.

ORCID

Christian A. Byrne (iD) https://orcid.org/0000-0003-0798-0286

REFERENCES

1. Murray RC, Schramme MC, Dyson SJ, Branch MV, Blunden TS. Magnetic resonance imaging characteristics of the foot in horses with palmar foot pain and control horses. Vet Radiol Ultrasound. 2006;47:1–16.
2. Barrett MF, Frisbie DD, King MR, Werpy NM, Kawcak CE. A review of how magnetic resonance imaging can aid in case management of common pathological conditions of the equine foot. Equine Vet Educ. 2017;29:683–93.
3. Dyson S, Murray R, Schramme M, Branch M. Magnetic resonance imaging of the equine foot: 15 horses. Equine Vet J. 2003;35:18–26.
4. Dyson S, Murray R, Schramme M, Branch M. Lameness in 46 horses associated with deep digital flexor tendonitis in the digit: diagnosis confirmed with magnetic resonance imaging. Equine Vet J. 2003;35:681–90.
5. Whitton RC, Buckley C, Donovan T, Wales AD, Dennis R. The diagnosis of lameness associated with distal limb pathology in a horse: a comparison of radiography, computed tomography and magnetic resonance imaging. Vet J. 1998;155:223–9.
6. Vallance SA, Bell RJW, Spriet M, Kass PH, Puchalski SM. 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: anatomical visualisation scores. Equine Vet J. 2012;44:51–6.
7. Vallance SA, Bell RJW, Spriet M, Kass PH, Puchalski SM. 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 Vet J. 2012;44:149–56.
8. Murray RC, Werpy N. Image interpretation and artefacts. In: Murray RC, editor. Equine MRI. Chichester, UK: John Wiley & Sons Ltd, 2010; p. 101–45.
9. McRobbie DW, Moore EA, Graves MJ, Prince MR. Ghosts in the machine: quality control. In: McRobbie DW, Moore EA, Graves MJ, Prince MR, editors. MRI from picture to proton, 3rd ed. Cambridge: Cambridge University Press, 2017; p. 166–82.
10. Murray RC, Mair TS, Sherlock CE, Blunden AS. Comparison of high-field and low-field magnetic resonance images of cadaver limbs of horses. Vet Rec. 2009;165:281–8.
11. Bolen G, Audigié F, Spriet M, Vandenberghe F, Busoni V. Qualitative comparison of 0.27T, 1.5T, and 3T magnetic resonance images of the normal equine foot. J Equine Vet Sci. 2010;30:9–20.
12. Ghazinoor S, Crues JV 3rd, Crowley C. Low-field musculoskeletal MRI. J. Magn Reson Imaging. 2007;25:234–44.
13. Porter EG, Werpy NM. New concepts in standing advanced diagnostic equine imaging. Vet Clin North Am Equine Pract. 2014;30:239–68.
14. McKnight AL, Manduca A, Felmlee JP, Rossman PJ, McGee KP, Ehman RL. Motion-correction techniques for standing equine MRI. Vet Radiol Ultrasound. 2004;45:513–9.
15. Evrard L, Audigié F, Bertoni L, Jacquet S, Denoix J-M, Busoni V. Low field magnetic resonance imaging of the equine distal interphalangeal joint: comparison between weight-bearing and non-weight-bearing conditions. PLoS One. 2019;14:e0211101.
16. Gutierrez-Nibeyro SD, Werpy NM, White NA, Mccutcheon LJ, Weng HY, Christopher JM. Standing low-field magnetic resonance imaging appearance of normal collateral ligaments of the equine distal interphalangeal joint. Vet Radiol Ultrasound. 2011;52:521–33.
17. Mair TS, Kinns J, Jones RD, Bolas NM. Magnetic resonance imaging of the distal limb of the standing horse: technique and review of 40 cases of foot lameness. Proc Am Assoc Equine Practnrs. 2003;49:29–41.
18. Dyson S, Murray R, Schramme M. Lameness associated with foot pain: results of magnetic resonance imaging in 199 horses (January 2001-December 2003) and response to treatment. Equine Vet J. 2005;37:113–21.
19. Sampson SN, Schneider RK, Gavin PR, Ho CP, Tucker RL, Charles EM. Magnetic resonance imaging findings in horses with recent onset navicular syndrome but without radiographic abnormalities. Vet Radiol Ultrasound. 2009;50:339–46.
20. Dyson S, Murray RC. Chapter 12: the foot and pastern. In: Murray RC, editor. Equine MRI. Chichester, UK: John Wiley & Sons Ltd, 2010; p. 269–314.
21. Gutierrez-Nibeyro S, Werpy N, White N. Standing low-field magnetic resonance imaging in horses with chronic foot pain. Aust Vet J. 2012;90:75–83.
22. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
23. Howell DC. Statistical methods for psychology, International Edition. Belmont, USA: Cengage Learning; 2012.
24. Werpy N. Low-field MRI in horses: practicalities and image acquisition. In: Murray RC, editor. Equine MRI. Chichester, UK: John Wiley & Sons Ltd, 2010; p. 75–99.
25. Stehling C, Souza RB, Hellio Le Graverand M-P, Wyman BT, Li X, Majumdar S et al Loading of the knee during 3.0T MRI is associated with significantly increased medial meniscus extrusion in mild and moderate osteoarthritis. Eur J Radiol. 2012;81:1839–45.
26. Bruno F, Barile A, Arrigoni F, Laporta A, Russo A, Carotti M et al Weight-bearing MRI of the knee: a review of advantages and limits. Acta Biomed. 2018;89:78–88.

27. Olive J. Distal interphalangeal articular cartilage assessment using low-field magnetic resonance imaging. Vet Radiol Ultrasound. 2010;51:259–66.

28. McKnight AL, Posh J. Articular cartilage lesions seen with MRI at 0.25T in the distal equine limb. J Equine Vet Sci. 2012;32:667–71.

29. Werpy NM. Magnetic resonance imaging of the equine patient: a comparison of high- and low-field systems. Clin Tech Equine Pract. 2007;6:37–45.

30. Murray R, Leece E, Judy C. High-field MRI in horses: practicalities and image acquisition. In: Murray RC, editor. Equine MRI. Chichester, UK: John Wiley & Sons Ltd, 2010; p. 39–74.

31. Murray RC, Blunden TS, Schramme MC, Dyson SJ. How does magnetic resonance imaging represent histologic findings in the equine digit? Vet Radiol Ultrasound. 2006;47:17–31.

32. Murray RC, Dyson S, Branch M, Schramme M. Validation of magnetic resonance imaging use in equine limbs. Clin Tech Equine Pract. 2007;6:26–36.

33. Mair TS, Kinns J, Jones RD, Bolas NM. Magnetic resonance imaging of the distal limb of the standing horse. Equine Vet Educ. 2005;17:74–8.

34. Werpy NM, Ho CP, Pease AP, Kawcak CE. The effect of sequence selection and field strength on detection of osteochondral defects in the metacarpophalangeal joint. Vet Radiol Ultrasound. 2011;52:154–60.

35. Smith MA, Dyson SJ, Murray RC. Reliability of high- and low-field magnetic resonance imaging systems for detection of cartilage and bone lesions in the equine cadaver fetlock. Equine Vet J. 2012;44:684–91.

36. Busoni V, Heimann M, Trenteseaux J, Snaps F, Dondelinger RF. Magnetic resonance imaging findings in the equine deep digital flexor tendon and distal sesamoid bone in advanced navicular disease–an ex vivo study. Vet Radiol Ultrasound. 2005;46:279–86.

37. Dyson S, Blunden T, Murray R. The collateral ligaments of the distal interphalangeal joint: magnetic resonance imaging and post mortem observations in 25 lame and 12 control horses. Equine Vet J. 2008;40:538–44.

38. Dyson S, Pool R, Blunden T, Murray R. The distal sesamoidean impar ligament: comparison between its appearance on magnetic resonance imaging and histology of the axial third of the ligament. Equine Vet J. 2010;42:332–9.

39. Dyson S, Blunden T, Murray R. Comparison between magnetic resonance imaging and histological findings in the navicular bone of horses with foot pain. Equine Vet J. 2012;44:692–8.

40. Bolas N. Basic MRI principles. In: Murray RC, editor. Equine MRI. Chichester, UK: John Wiley & Sons Ltd, 2010; p. 1–37.

41. Busoni V, Snaps F. Effect of deep digital flexor tendon orientation on magnetic resonance imaging signal intensity in isolated equine limbs-the magic angle effect. Vet Radiol Ultrasound. 2002;43:428–30.

42. Spriet M, Mai W, McKnight A. 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. Vet Radiol Ultrasound. 2007;48:95–100.

43. Spriet M, McKnight A. Characterization of the magic angle effect in the equine deep digital flexor tendon using a low-field magnetic resonance system. Vet Radiol Ultrasound. 2009;50:32–6.

44. Werpy NM, Ho CP, Kawcak CE. Magic angle effect in normal collateral ligaments of the distal interphalangeal joint in horses imaged with a high-field magnetic resonance imaging system. Vet Radiol Ultrasound. 2010;51:2–10.

45. Smith MA, Dyson SJ, Murray RC. 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? Vet Radiol Ultrasound. 2008;49:509–15.

46. Werpy NM, Ho CP, Garcia EB, Kawcak CE. The effect of varying echo time using T2-weighted FSE sequences on the magic angle effect in the collateral ligaments of the distal interphalangeal joint in horses. Vet Radiol Ultrasound. 2013;54:31–5.

47. Sherlock CE, Mair TS. Magic angle effect on low field magnetic resonance images in the superficial digital flexor tendon in the equine proximal pastern region. Vet J. 2016;217:126–31.

48. Spriet M, Zwingenberger A. 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 Vet J. 2009;41:498–503.

49. Smith MA, Milmine R, Dyson SJ, Murray RC. Optimazation of STIR sequences for equine low field standing MRI. In: *Proceedings of the European Veterinary Conference Voorjaarsdagen*. 2008; p. 323.

50. Bolen G, Haye D, Dondelinger R, Busoni V. Magnetic resonance signal changes during time in equine limbs refrigerated at 4 degrees C. Vet Radiol Ultrasound. 2010;51:19–24.

51. Andersen MS, Clark L, Dyson SJ, Newton JR. Risk factors for colic in horses after general anaesthesia for MRI or nonabdominal surgery: absence of evidence of effect from perianaesthetic morphine. Equine Vet J. 2006;38:368–74.

52. Franci P, Leece EA, Brearley JC. Post anaesthetic myopathy/neuropathy in horses undergoing magnetic resonance imaging compared to horses undergoing surgery. Equine Vet J. 2006;38:497–501.

53. Moreno KL, Scallan EM, Friedeck WO, Simon BT. Transient pelvic limb neuropathy following proximal metatarsal and tarsal magnetic resonance imaging in seven horses. Equine Vet J. 2020;52:359–63.

54. Kendall MG, Babington Smith B. The problem of m rankings. Ann Math Stat. 1939;10:275–87.

55. Bladon B. Magnetic resonance imaging and foot lameness. Problem solved? Or do we know we know less now than we know more? Equine Vet J. 2014;46:264–6.

56. Hallmarq Veterinary Imaging Ltd. Above and beyond: over 100,000 horses scanned in over 100 Standing Equine MRI systems. 2019 [cited 2020 May 14]. https://hallmarq.net/wp-content/uploads/2019/07/100th-System-and-beyond-PR-310719.pdf.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.