

Article

Strain Elastography of Injured Equine Superficial Digital Flexor Tendons: A Reliability Study of Manual Measurements

Valentina Secchi ^{1,2}, Gerolamo Masala ^{1,2,3}, Andrea Corda ^{1,2,*} , Francesca Corda ^{1,2}, Enrica Potop ¹, Alicia Barbero Fernandez ⁴ , Maria Luisa Pinna Parpaglia ^{1,2}  and Eraldo Sanna Passino ^{1,2,3}

¹ Department of Veterinary Medicine, University of Sassari, 07100 Sassari, Italy; valentinasecchi.vet@gmail.com (V.S.); gemasala72@gmail.com (G.M.); francescacorda91@tiscali.it (F.C.); enrica.potop@yahoo.it (E.P.); pinnapar@uniss.it (M.L.P.P.); esp@uniss.it (E.S.P.)

² Veterinary Teaching Hospital, University of Sassari, 07100 Sassari, Italy

³ Comparative Surgery Research Laboratory, University of Sassari, 07100 Sassari, Italy

⁴ Department of Veterinary Medicine, University Alfonso X el Sabio, 28691 Villanueva de la Cañada, Madrid, Spain; aliciabarbero.vet@gmail.com

* Correspondence: andreacorda@uniss.it

Simple Summary: Strain elastography is an ultrasound-based technique that assesses the mechanical properties of tissues and gives a relative representation of elasticity. Early diagnosis of tendon injuries and long-term monitoring of the healing process are key to equine practice; thus, an accurate method is needed for analyzing and interpreting the images obtained with strain elastography. The first aim of the study was to demonstrate the intraoperator repeatability and interoperator reproducibility of manual measurements of elastograms obtained from injured superficial digital flexor tendons of horses; the second aim was to perform a standardization of the manual measurement method by comparing it with external software. Despite their subjectivity, manual measurements proved to be repeatable and reproducible. In addition, the results obtained with the manual method matched those obtained with the external software.

Abstract: Early diagnosis of tendon injuries and accurate long-term monitoring of the healing process are key for equine veterinarians that use conventional ultrasonography. The development of strain elastography could improve the management of clinical cases. The aim of the study was to assess the intraobserver repeatability and interobserver reproducibility of manual measurements of the colored areas of the tendons within elastograms and to standardize this manual modality by comparing the analysis of the images with ImageJ. Twenty elastograms of the injured superficial digital flexor tendons (SDFTs) of horses were analyzed by two different operators after an acute injury was diagnosed with ultrasonography. Statistical analysis demonstrated excellent intraobserver repeatability (intraclass correlation coefficient, ICC = 0.949) and good interobserver reproducibility (ICC = 0.855) for manual measurements performed with tools available on the ultrasound unit. A good agreement between manual measurements and measurements performed with ImageJ (ICC = 0.849) was then demonstrated. Despite its subjectivity, the manual modality proved to be a valid method for analyzing images obtained with strain elastography.

Keywords: horse; strain elastography; SDFT; manual measurement; ImageJ; reliability



Citation: Secchi, V.; Masala, G.; Corda, A.; Corda, F.; Potop, E.; Barbero Fernandez, A.; Pinna Parpaglia, M.L.; Sanna Passino, E. Strain Elastography of Injured Equine Superficial Digital Flexor Tendons: A Reliability Study of Manual Measurements. *Animals* **2021**, *11*, 795. <https://doi.org/10.3390/ani11030795>

Academic Editor: Brian Nielsen

Received: 10 February 2021

Accepted: 8 March 2021

Published: 12 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sports horses are particularly subject to tendon injuries. This is a huge issue due to the high recurrence rate and the time needed for complete healing [1,2]. Two-dimensional ultrasonography (2D-US) is commonly used to diagnose and monitor equine tendinopathies. In fact, due to the use of portable ultrasound units, 2D-US is low-risk and easy to use in the field. However, it is not able to assess the mechanical properties of tissues, or to predict

injuries or detect abnormal signs in tendons over five months post-injury, even though tendons may need up to 18 months for a total recovery [1,3,4].

On the other hand, ultrasound elastography (USE) is able to assess the mechanical properties of tissues and provide information on their elasticity. The two main types of USE are shear wave elastography (SWE) and strain elastography (SE).

Shear wave elastography is a quantitative technique based on the propagation of shear waves, thanks to a force applied on the tissues and created by the ultrasound unit.

Strain elastography, also called real-time elastography or compression elastography, is a qualitative or semi-quantitative USE technique. It is performed by applying a gentle manual rhythmic compression on tissues with the transducer, causing an axial displacement, higher for more elastic (softer) structures and lower for less elastic (harder) ones. Dedicated ultrasound elastography software calculates the axial displacement by comparing 2D-US images pre- and post-compression, thus obtaining a color-coded strain map (elastogram), superimposed on the greyscale image.

Unlike SWE, SE does not produce numerical elasticity parameters and provides a relative representation of the elasticity of tissues within the selected window, field of view (Fov). Colors in the elastogram thus depend on the structures in the Fov, which must be kept constant throughout ultrasound examinations for a correct comparison of images [5].

Despite being highly operator-dependent and its relativity, SE is a feasible, repeatable and reproducible method for assessing normal tendons and ligaments of equine distal limbs [6]. It can also be used to detect early tendinopathies and to evaluate tendon healing, if always associated with 2D-US [7–11].

Over five months post-injury, during the healing process, 2D-US would seem unable to recognize subtle abnormalities in tendon structures [3,12].

In contrast, SE is able to identify small differences in the mechanical properties of the affected area compared to the surrounding healthy tendon. It could thus be useful in evaluating tendon healing, even during the late phase of rehabilitation, because of its higher sensitivity than 2D-US. This would help clinicians to establish a more accurate prognosis for clinical cases [12].

Another important aspect to be considered for the correct clinical application of SE is an accurate evaluation and interpretation of images after the acquisition. A comprehensive and objective analysis of the color patterns identified by SE could be key in assessing the healing process of tendons.

However, due to the lack of objective and numeric parameters, SE elastograms are difficult to interpret, and only a subjective and categorical classification is usually performed [6,11,13,14].

An alternative method for image analysis using data from external software could help to make manual modality more objective.

Although image analysis with external software is more objective, it is also a complex and long procedure, and thus is impracticable in daily clinical activities. On the other hand, taking manual measurements with the ultrasound unit is a simpler method and applicable in the field, but it is also highly operator-dependent and subjective.

This problem could be avoided by standardizing the procedure with external software. Consequently, the first aim of our study was to assess the intraobserver repeatability and interobserver reproducibility of manual measurements of colored areas (red, green and blue) in the elastograms of injured tendons, obtained with an ultrasound unit equipped with SE software.

The second aim was to standardize the manual measurement modality, using external software for image analysis.

2. Materials and Methods

2.1. Horses and Preparation

Twenty Anglo-Arabian racehorses (11 mares and 9 geldings, aged between 4 and 9 years) with a forelimb superficial digital flexor tendon (SDFT) core lesion, which had

occurred during racing or training, were included in the study. Each horse was first given a clinical evaluation, and then assessed with both 2D-US and SE within two weeks from the onset of injury.

Before performing the ultrasonographic examination, the hair was cut on the palmar metacarpus, which was then washed thoroughly with water. A coupling gel was finally applied in order to improve probe contact and obtain a better resolution.

Horses were not sedated because they were all sufficiently calm.

The study was approved by the ethics committee (Organismo Preposto al Benessere Animale) of the University of Sassari (protocol code no. 128528, approved on 15 November 2019).

2.2. Ultrasound Technique and Devices

Ultrasonographic examinations were performed by one veterinarian (VS) using an ultrasound unit (My Lab Alpha, Esaote, Florence, Italy) equipped with the SE software “ElaXto” (Esaote, Florence, Italy) and a linear 3–13-MHz transducer.

Horses were evaluated in a square stance, under weight-bearing conditions, and forelimbs were assessed from the carpometacarpal joint to the distal sesamoid bones, dividing the palmar aspect of the metacarpus into seven levels (1–7) for transverse planes, and into three levels (1–3) for longitudinal planes. This system differentiates the metacarpal regions according to specific anatomic features and was essential for a better comparison of the images of the various horses involved in the study [15].

The examinations were performed from proximal to distal, first in transverse then in longitudinal planes, at each level of the metacarpus. The probe was positioned in order to obtain lateral structures on the left and medial structures on the right side of the screen, in transverse orientation, and distal parts on the left and proximal parts on the right, in longitudinal orientation.

In order to obtain correct 2D-US images and to avoid anisotropy artefacts, the transducer was held perpendicularly to the surface analyzed [16]. A standoff pad was also used for the 2D-US examinations.

In the trials, power (100%), frequency (intermediate), number and position of foci and depth (~4 cm) were kept stable.

2.3. Elastographic Evaluation

Strain elastography was performed by the same trained veterinarian (VS) by applying a gentle manual rhythmic pressure with the linear transducer. As recommended by the manufacturer, the manual compression was minimal, i.e., a vibration with a movement equal or less than 1 mm [5]. A standoff pad was not used in the SE examination.

A color-coded map was selected, with a range of red (softer tissues) through yellow/green (intermediate stiffness) to blue (stiffer tissues) colors.

A wide Fov was set and kept constant (depth of ~4 cm) to obtain as detailed a chromatic scale as possible [5]. In these experimental trials, the Fov included SDFT, deep digital flexor tendon, accessory ligament of deep digital flexor tendon, suspensory ligament, and palmar border of the cannon bone.

A 2D-US image was placed simultaneously beside the elastogram to ensure that the image was kept constant and the probe was not moved laterally, and to assess the movement of tissue during compression–release cycles (1 mm maximum). A minimal precompression was applied.

The accuracy of the elastograms obtained was evaluated during the examination by checking a visual indicator. As the correct compression was applied, the indicator color changed from grey to green, and a videotape was recorded for each level of the metacarpus (Figure 1).

Only images in long-axis view were selected for the following analysis.

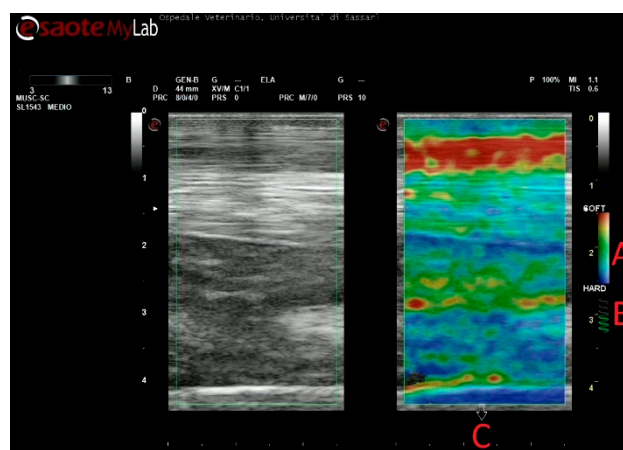


Figure 1. Example of strain elastography. (A) Color range selected, (B) Indicator for correct compression, (C) Field of view.

2.4. Image Analysis

Twenty images, one of each horse, were randomly selected from the most representative videotapes. Elastograms were considered selectable if they enclosed an SDFT lesion and if the indicator for correct compression was green. The same images were then analyzed on the same ultrasound unit by two different veterinarians (VS and AC), who were blinded to each other's results, in order to test interobserver reproducibility. Image analysis was repeated by one vet (VS) two weeks apart, in order to test intraobserver repeatability.

Image analysis consisted of a measurement of the total area of the SDFT on the longitudinal scan selected and of the area of each predominant color within the tendon (red, green and blue). All measurements were performed three times, using the manual tools available on the ultrasound unit, and the mean values were used for further statistical evaluation. Areas were expressed in cm^2 .

The same images were then objectively evaluated with ImageJ (Version 1.44) [17].

The evaluation was made on a laptop (Ideapad 310-15IKB, Lenovo, Hong Kong, China). Original images (in PNG format), including all the metacarpal structures, were imported into the laptop from the ultrasound unit and cropped along the SDFT perimeter with a basic image editing program; no other manipulation was performed.

Cropped elastograms showing the injured SDFTs were imported in PNG format and analyzed with ImageJ. In each image, the tendon total area and areas of the main colors (red, green and blue) were measured with the "Threshold Color" plugin.

For the selection of colors, the Hue bar was the main tool used, with total hues ranging from 0 (red hues) to 255 (red hues). For the selection of green and blue, a range of hues from 30 to 100 and from 100 to 220 was chosen, respectively. On the other hand, the red area was obtained by subtracting green and blue areas from the total areas. Areas were expressed in numbers of pixels (Figure 2).

All the results were recorded in an ad hoc database, and the percentage areas of each color was calculated from the mean values derived from the three consecutive measurements.

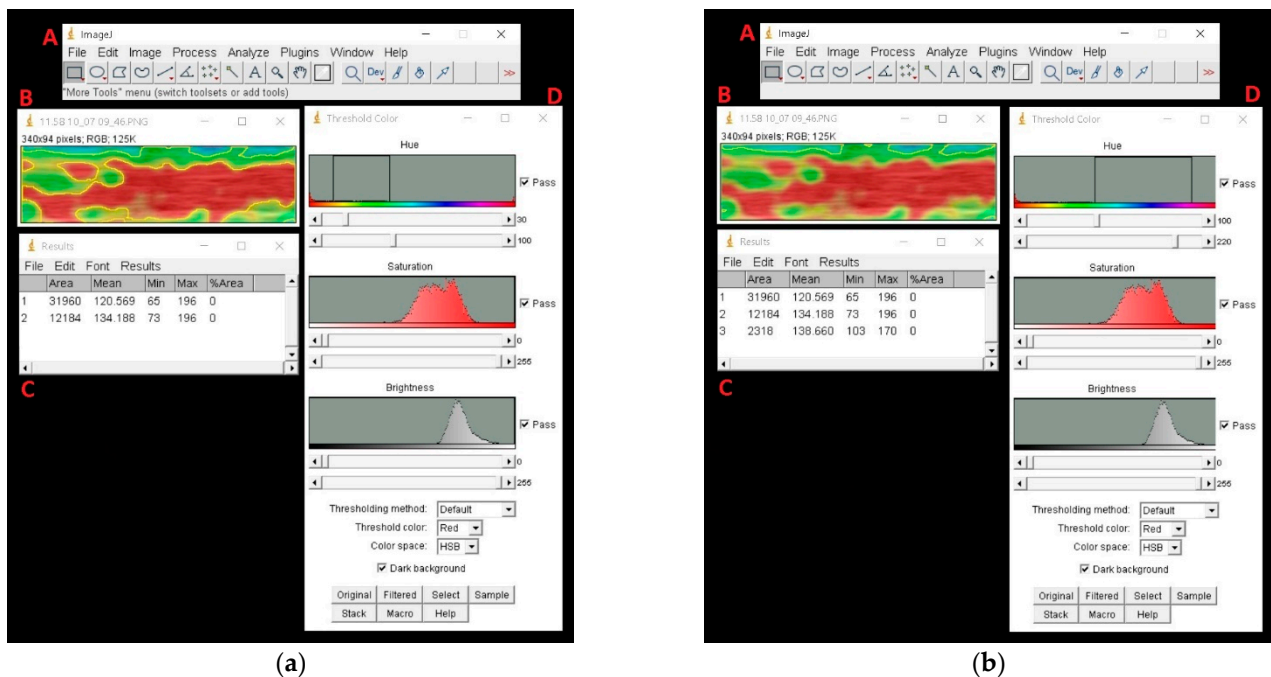


Figure 2. Example of analysis on ImageJ: (a) Selection of green area; (b) Selection of blue area. The selection of colors is mainly regulated with the Hue bar, and areas (expressed in pixels) of the selected part are measured. (A = ImageJ main tool bar, B = superficial digital flexor tendon elastogram image imported, C = Measurement results, D = “Threshold color” plugin).

2.5. Statistical Analysis

The statistical analysis was performed using *R* (Version 3.6.1) [18].

The observed data were triplets $(x_{red}, x_{green}, x_{blue})$ such that $x_{red} + x_{green} + x_{blue} = 100\%$. For example, x_{red} represented the ratio between the red area and the total area of the image, expressed as a percentage. A similar interpretation was given to x_{green} and x_{blue} . Such data structures describe the parts of a whole and are known in the statistical literature as *compositional data* [19,20]. Mathematically, the observations lie on a 3-part simplex, since x_{red} , x_{green} and x_{blue} are constrained to add up to 100%. This constraint violates the assumptions of many standard statistical methods for data analysis. The observations should therefore be projected from the simplex to a 2D real space for further processing. An appropriate transformation is the *isometric log ratio transform* (ILR) [21]. For each observation $(x_{red}, x_{green}, x_{blue})$, the ILR transformation returns a point (y_1, y_2) in the real plane \mathbb{R}^2 which is suitable for further statistical analysis:

$$(y_1, y_2) = \text{ILR}(x_{red}, x_{green}, x_{blue}). \quad (1)$$

The ILR-transformed values (y_1, y_2) of the measurements $(x_{red}, x_{green}, x_{blue})$ were used in this work as response variables within the statistical models used for analysis.

A more detailed description of the ILR transform is given in [21]. Practical details regarding the analysis of compositional data in *R* are given in [22], and the associated package compositions in [23]. In this work, the default settings of the package were used to calculate the ILR transformation.

Some of the observations for x_{blue} were exactly zero, which must be substituted before applying ILR. We assumed that the zeros arose for percentages of blue below the detection limit of 1%. The zeros were then replaced using a simple rule of two-thirds of the detection limit. A more detailed description of the replacement strategies in compositional data is given in [24].

Since the response variable was bivariate, two-way repeated measures MANOVA was used to assess intraobserver repeatability, grouped by time of measurement and image, restricted to the observations recorded by the veterinarian (VS). The different images were

considered as a random effect, and the resulting mixed model was fitted using the R package *lme4* [25]. The model assumptions were tested using standard statistical tests for multivariate observations: normality was tested using Mardia's skewness and kurtosis tests (implemented in the R package *MVN* [26]), and homogeneity of the covariance matrices was tested using Box's M-test (implemented in the R package *heplots* [27]). The models with and without effects for the time of measurement were fitted via maximum likelihood, and the likelihood ratio test was used to assess the difference between them. The adjusted intraclass correlation coefficient (ICC, [28]) was also calculated, with confidence intervals estimated via bootstrapping.

In order to further strengthen the results, intraobserver repeatability was assessed using a one-sample Hotelling's T^2 -test (implemented in the R package *ICSNP* [29]), applied on the difference between the two measurements of VS. The Mahalanobis distance calculated for the T^2 -test was used as a measure of effect size. Bland–Altman plots [30] were also used to visually assess the agreement between the two measurements, separately for each dimension, after the ILR transformation. In order to provide an intuitive interpretation of the results, the boxplots of the differences between the two measurements were also plotted for each of the three color channels.

Similarly, two-way mixed-effects MANOVA was used to evaluate interobserver repeatability, with grouping factors represented by the veterinarian and the image. The image was considered as a random effect, whereas the veterinarian was assumed to be a fixed effect. Model assumptions were tested using the same procedures used for the study on intraobserver repeatability. The likelihood ratio test was used to assess whether the difference between the measurements of the two veterinarians was statistically significant, and the ICC was calculated with bootstrapped confidence intervals. Bland–Altman plots were constructed from an average of the two measurements of VS, compared to the results obtained by AC. Boxplots for the three color channels were used to provide an intuitive interpretation of the results.

Finally, the difference between the methodologies (manual and ImageJ) was tested using two-way mixed-effects MANOVA, fitted on the entire dataset using maximum likelihood, assuming methodology as the fixed effect and image as the random effect. The underlying model assumptions were again tested using the same procedures described for the two previous studies. The difference between the two methodologies was tested using the likelihood ratio test, which tested the significance of the model coefficients associated with the manual method and ImageJ. In addition, Bland–Altman plots were used to visually assess the agreement between the two methodologies; the plots were constructed from an average of the three measurements from the manual method, compared to the single measurement obtained from ImageJ. As before, boxplots of the three color channels were also used to visually assess the difference between the two methodologies. The threshold for significance of the results of each statistical test was assumed to be $\alpha = 0.05$. The criteria of [31] were used to evaluate the ICC: poor, (0–0.50); moderate, (0.5–0.75); good, (0.75–0.9); excellent, (0.9–1).

3. Results

Figure 3 shows the scatterplots of the measurements before and after the ILR transformation. Table 1 reports the summary statistics of the ILR-transformed data, and Table 2 the corresponding tests for validation of the MANOVA assumptions. The Mardia and Box's M-test showed no statistical evidence of violations of any of the model assumptions. The mean and confidence interval for the mean of each of the two dimensions (y_1, y_2) suggested that the two measurements of VS were in agreement, whereas, as expected, more substantial differences were observed in the measurements taken by the different observers (VS and AC) and in the methods (manual and ImageJ).

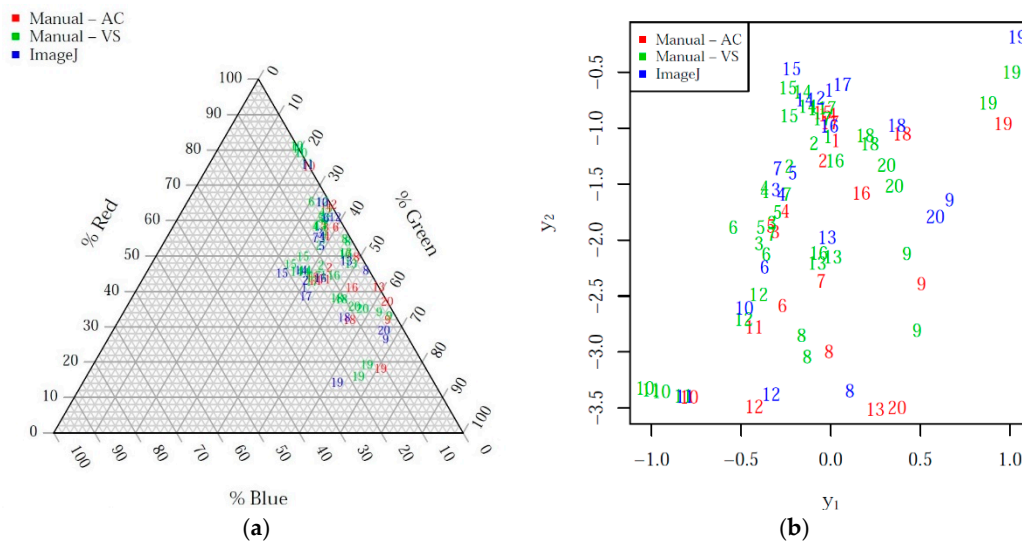


Figure 3. Scatterplot of the measurements before (a) and after (b) the isometric log ratio (ILR) transformation, labeled by image and colored by observer and method.

Table 1. Summary (mean and 95% Student’s t confidence interval for the mean) of the isometric log ratio (ILR)-transformed data.

Method	Observer	y_1 (1st Dimension of ILR)		y_2 (2nd Dimension of ILR)	
		Mean	95% C.I.	Mean	95% C.I.
Manual	VS (Measurement 1)	−0.174	(−0.371, 0.023)	−1.812	(−2.195, −1.429)
	VS (Measurement 2)	−0.134	(−0.346, 0.078)	−1.728	(−2.146, −1.309)
	AC	−0.015	(−0.198, 0.168)	−2.062	(−2.519, −1.606)
ImageJ	VS	−0.037	(−0.238, 0.164)	−1.582	(−2.044, −1.120)

Table 2. Statistical tests for validation of the mixed-effects model assumptions.

Method	Observer	Mardia’s p -Values		Box’s M p -Value
		Skewness	Kurtosis	(~Grouping)
Manual	VS (Measurement 1)	0.314	0.559	0.940
	VS (Measurement 2)	0.246	0.538	(~measurement)
	AC	0.615	0.546	0.429
ImageJ	VS	0.299	0.711	(~observer)
				0.957 (~method)

Intraobserver repeatability, interobserver reproducibility and the difference between the manual method and ImageJ were more precisely quantified by the likelihood ratio tests reported in Table 3. The two measurements by VS on the same images did not significantly differ (p -value 0.157), demonstrating the intraobserver repeatability of the methodology. This was confirmed by Hotelling’s T^2 -test on the differences between the measurements on individual images (p -value 0.097). The ICC further demonstrated that most of the variability was explained by the differences between the images (adjusted ICC 0.949, with the lower bound of the 95% confidence interval exceeding 0.9), showing

excellent intraobserver repeatability. The agreement between the two measurements is also confirmed by Figure 4, which shows the Bland–Altman plots on each of the two dimensions of the ILR-transformed data, and the boxplot of the difference between the measurements for each color channel. In the Bland–Altman plots, the observations do not significantly deviate from the horizontal line at zero. In addition, in the boxplots, the differences for each color channel appear to have a symmetric distribution around zero, further demonstrating the intraobserver repeatability of the methodology.

Table 3. Likelihood ratio test, intraclass correlation coefficient and Hotelling’s T^2 -test for the intraobserver repeatability study. ICC, Intraclass Correlation Coefficient; C.I., Confidence Interval.

Likelihood Ratio Test		Adjusted ICC		Hotelling’s T^2 -Test	
χ^2 Score	p -Value	Estimate	95% C.I.	Mahalanobis Distance	p -Value
3.698	0.157	0.949	(0.905, 0.972)	2.711	0.094

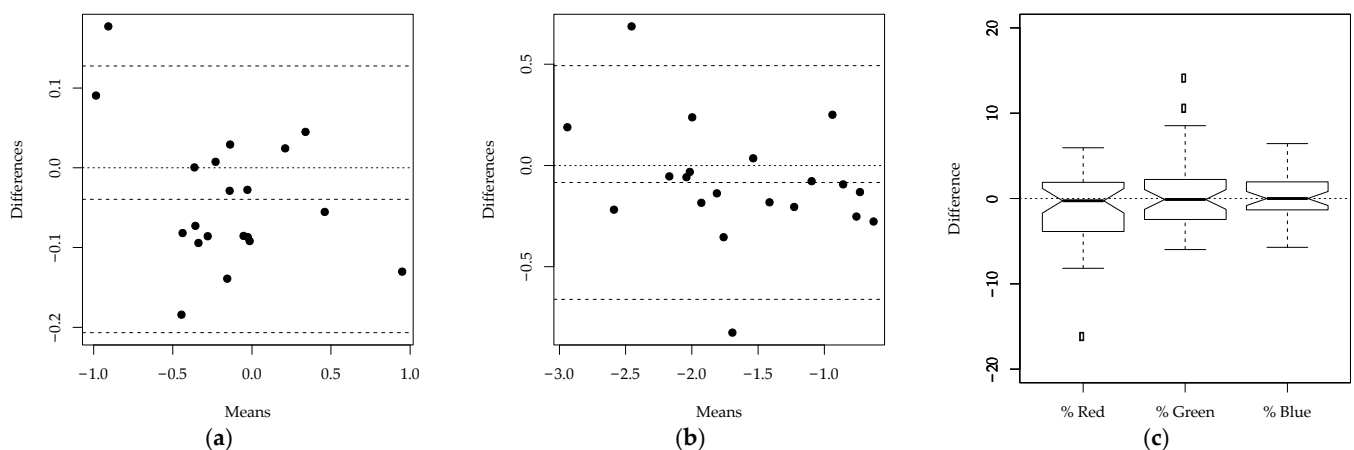


Figure 4. Intraobserver reproducibility study: (a) Bland–Altman plot—First dimension of isometric log ratio transformed data; (b) Bland–Altman plot—Second dimension of isometric log ratio transformed data; (c) Boxplots of the differences between the first and second measurement by VS for each color channel.

Similar statistical tests were carried out to assess interobserver reproducibility. From the likelihood ratio test shown in Table 4, it appeared that the measurements taken by the veterinarians, AC and VS, were significantly different (p -value < 0.001). This is confirmed by Figure 5, as the observations in the Bland–Altman plots are not scattered around the horizontal line at zero, and the boxplots of the differences for the three color channels are not centered around zero. A significant difference was observed between the two veterinarians in the assessment of the percentages of the colors red and green. Despite the significant difference between the two observers, the adjusted ICC (0.855, with the lower bound of the 95% confidence interval exceeding 0.75) reported in Table 4 still suggests that there is good agreement between the two veterinarians’ measurements, since most of the variability is due to the differences between the images. The images refer to lesions with different gravity levels and are therefore highly heterogeneous. The variability between the measurements on different images is thus much larger than the variability of measurements on the same image, which leads to large values of the ICC.

Table 4. Likelihood ratio test and intraclass correlation coefficient for the interobserver reproducibility study. ICC, Intraclass Correlation Coefficient; C.I., Confidence Interval.

Likelihood Ratio Test		Adjusted ICC	
χ^2 Score	<i>p</i> -Value	Estimate	95% C.I.
18.250	<0.001	0.855	(0.753, 0.907)

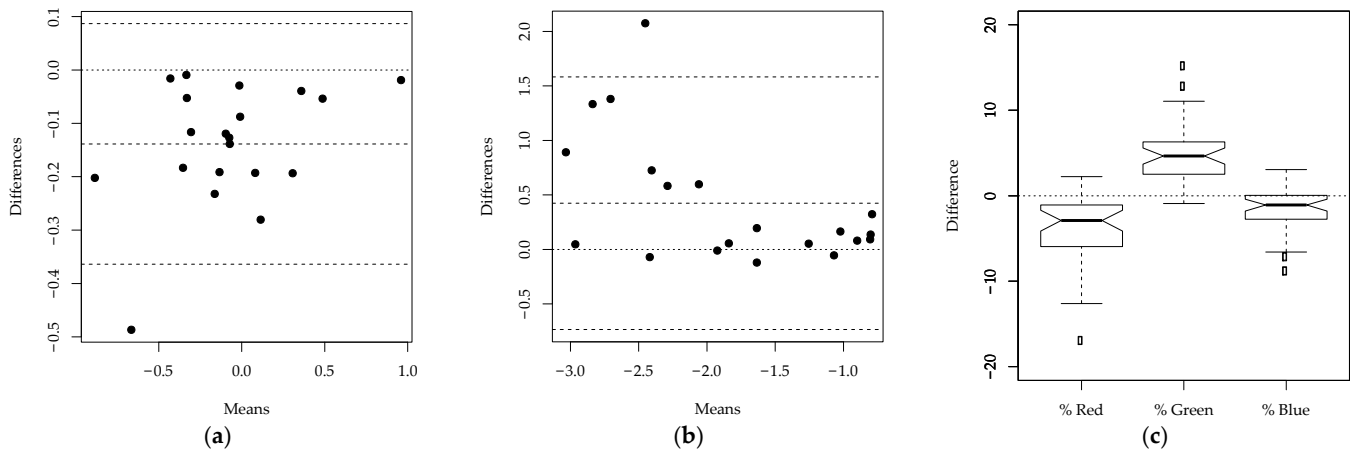


Figure 5. Interobserver reproducibility study: (a) Bland-Altman plot—First dimension of isometric log ratio transformed data; (b) Bland-Altman plot—Second dimension of isometric log ratio transformed data; (c) Boxplots of the differences between the measurements of AC and VS for each color channel.

Finally, the difference between the measurements obtained using the manual method and ImageJ was assessed. The likelihood ratio test in Table 5 suggests that there is a statistically significant difference between the measurements obtained using the two methodologies (*p*-value < 0.001). Again, this is confirmed by the Bland–Altman plots and boxplots reported in Figure 6: the observations in the Bland–Altman plots are not randomly distributed around zero, and the boxplots of the differences for the three color channels show that the measurements of the colors red and blue significantly deviate from zero. The adjusted ICC (0.849, with the lower bound of the 95% confidence interval exceeding 0.75) in Table 5 demonstrates that, despite the significant differences between the two methodologies, there is good agreement between the measurements. Similarly to the study on interobserver reproducibility, the ICC is large because most of the variability is due to the differences between images.

Table 5. Likelihood ratio test and intraclass correlation coefficient for the study on the differences between the manual method and ImageJ. ICC, Intraclass Correlation Coefficient; C.I., Confidence Interval.

Likelihood Ratio Test		Adjusted ICC	
χ^2 Score	<i>p</i> -Value	Estimate	95% C.I.
16.076	<0.001	0.849	(0.751, 0.899)

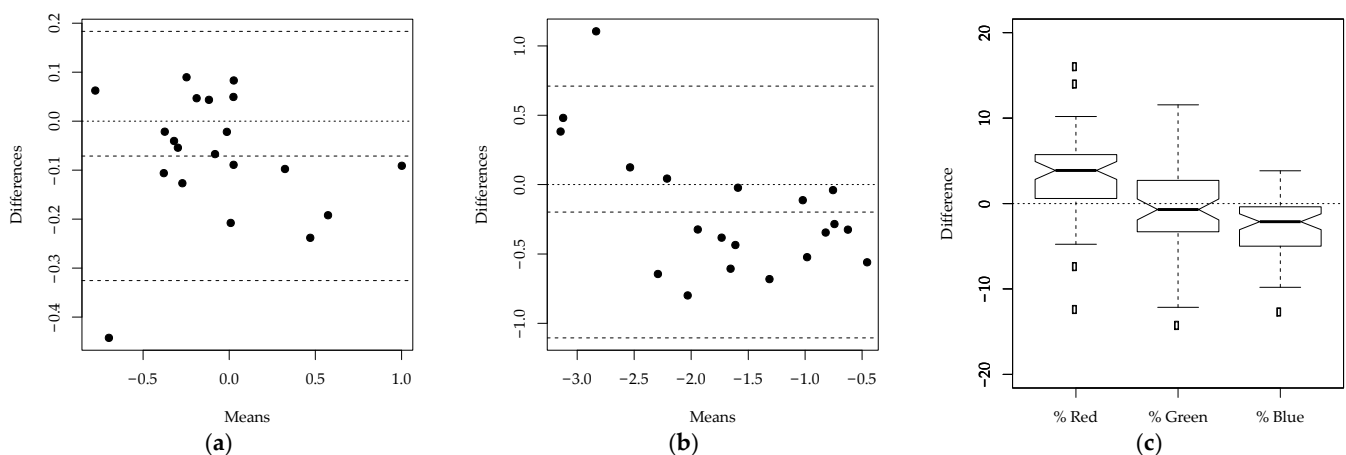


Figure 6. Agreement between manual method and ImageJ: (a) Bland-Altman plot—First dimension of ILR-transformed data; (b) Bland-Altman plot—Second dimension of ILR-transformed data; (c) Boxplots of the differences between the measurements obtained using manual method and ImageJ, for each color channel.

4. Discussion

The manual measurements of colors in the elastograms of acutely injured SDFTs in horses showed excellent intraobserver repeatability (ICC 0.949) and good interobserver reproducibility (ICC 0.855). Good agreement was also obtained between the manual method and the ImageJ methodologies of the color area measurements (ICC 0.849).

The manual measurement of elastograms was thus shown to be a repeatable and reproducible method which could be considered as interchangeable with the objective measurement obtained with the software image analysis.

Our results confirmed previous findings showing that SE image analysis had an almost perfect interobserver agreement when imaging injured tendons [11]. Moreover, our results showed a higher level of agreement compared to a previous study conducted on healthy tendons, in which the qualitative evaluation had moderate interobserver agreement and good intraobserver agreement [6]. Only one study reported the variability associated both with acquiring and analyzing images, which was reported as low [6]. To the best of our knowledge, in equine medicine, no study has reported the standardization of the manual analysis of SE-derived elastograms using image analysis software.

Strain elastography is a qualitative USE technique and is performed by applying a gentle manual rhythmic compression on tissues. It does not produce numerical elasticity parameters, but a relative representation of the elasticity of tissues within the selected window. It is highly operator-dependent, requiring well-trained sonographers. In addition, accurate acquisitions involve careful preparation [5,7,32–34].

Healthy equine SDFTs normally appear in elastograms as predominantly blue structures, with a mostly green peritendinous tissue, and surrounded by soft tissue structures, which appear as red [6]. When a lesion occurs, during the acute inflammatory phase, the injured area of the tendons becomes softer (red), due to hemorrhage, edema and fibrin clot organization [35]. As the granulation tissue forms, tendons become more stable, and the color changes to yellow/green, representing intermediate stiffness [36,37]. During the remodeling phase, tendons become progressively harder, returning to their original mechanical properties (mainly blue) [38].

In order to obtain accurate elastograms of SDFTs in horses, animals should be evaluated under weight-bearing conditions, because there are usually fewer artefacts in the elastographic evaluation compared to a non-weight-bearing position [6,11]. Minimal precompression should be used in order to prevent precompression artefacts. In fact, when excessive precompression is exerted, elastograms are not representative of the real mechanical characteristics of tissues [5]. The visual indicator for correct compression does not consider the precompression degree and thus could appear green despite an

insufficiently accurate evaluation. In our study, elastographic examinations of the SDFT were relatively easy to perform, due to the superficial location of the examined structure, thus obtaining a qualitative representation of the mechanical properties of pathologic tendons. We only analyzed longitudinal images due to problems encountered in obtaining elastograms in the short-axis view, especially in distal levels of the metacarpus. Images of transverse planes showed more artefacts than longitudinal ones, especially on medial and lateral parts, due to the convex margins of the tendons. Both in human medicine, for the evaluation of Achilles tendons, and in veterinary medicine, for the evaluation of SDFTs in horses and calcaneal and patellar tendons in dogs, longitudinal images are preferred because they are considered to be of better quality, while transverse images show more artefacts and lower reproducibility [12,39–41].

The SE images were obtained without a standoff pad, which can be problematic for a correct evaluation of the elasticity pattern. The lack of accurate contact between the probe and standoff pad can produce reverberation artifacts, which are visible as parallel red areas in the elastogram, which can be confused with areas of softness [6,42]. In addition, the inclusion of the pad in the Fov can alter the color pattern of SE-derived elastograms, because of its well-known relativity. The standoff pad can represent the softest part within the elastographic window, and, as a consequence, SDFT lesions can be misinterpreted or underestimated, no longer appearing as red but yellow/green—in other words, with intermediate stiffness. Finally, the use of the pad can alter the elasticity estimation since the deepest structures, which we decided to include (suspensory ligament and cannon bone), are probably more difficult to reach. This can subsequently lead to a lack of signal, without any color (Figure 7).

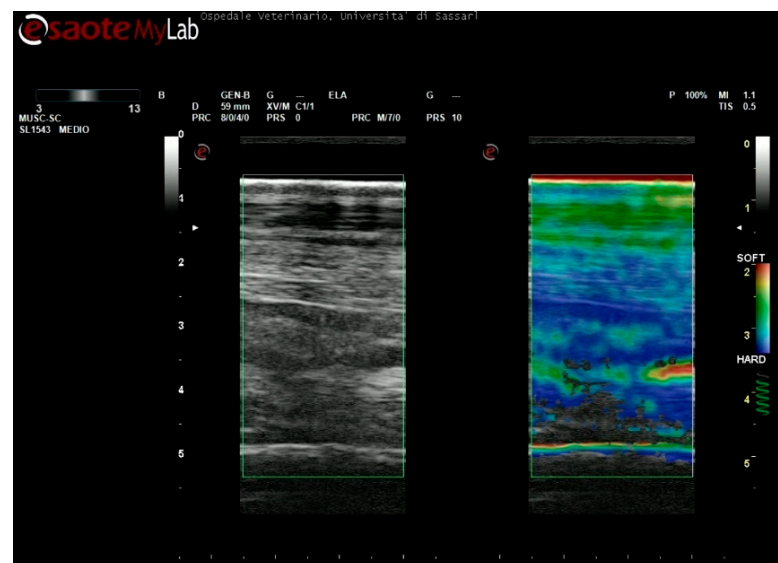


Figure 7. At the bottom of the field of view (Fov), an artifact is present as a lack of signal, without any colors. This could be explained by the use of the pad, as the deepest structures, where the artifact appears, are probably more difficult to reach, compared with images obtained without a pad. Secondly, the inclusion of a standoff pad within Fov can alter the color pattern of strain elastography-derived elastograms, because of its relativity; the standoff pad can represent the softest part within elastographic window, and as a consequence, the superficial digital flexor tendon lesions can be misinterpreted or underestimated, no longer appearing red but yellow/green.

In our trials, it was not necessary to sedate the horses; however, the use of a sedative is recommended if needed. The influence of sedation on USE results is still controversial, with some authors stating that the results are not compromised [6,12] and others declaring that images can be potentially influenced by sedation [43].

For the SE image analysis, categorical color-grading is normally the most common system used. It is a qualitative method which was first used in human medicine for the evaluation of Achilles tendons [13,14] and subsequently adapted for equine SDFTs [6,11]. This simple categorical assessment facilitates a semi-quantitative analysis of data, with high interobserver agreement and applicability in clinical practice [11,12].

However, due to the subjective nature of color-grading, we used ImageJ for a more objective analysis, which we compared with a manual assessment using an ultrasound unit. As recommended [44,45], the evaluation consisted in calculating the percentage of the three main colors present in the elastograms: red (softest structures), green (intermediate stiffness) and blue (stiffest areas).

The software analysis was performed with the “Threshold Color” plugin, because it was considered easier to interpret clinically and more reproducible than qualitative or quantitative methods, which measure the mean echo-intensity for each color considered [44,46,47].

A significant difference was initially observed between the two veterinarians’ (VS and AC) measurements and the measurements taken using the two methodologies. However, the subsequent analyses established that most of the variability was due to differences between images; thus, in reality, good ICCs were obtained.

Differences between images could be explained considering different sizes, distributions and shapes of the colored areas which characterize injured tendon elastograms. For example, measuring an area with a regular and larger shape is simpler than measuring an area with an irregular and smaller shape. This can be explained clinically by the heterogeneity of the SDFT lesions included in the sample. As a consequence, some elastograms were very different than others, in terms of the proportion, morphology and margins of colors within Fov.

Moreover, the differences in tracing could also be a product of the differences in the visual perception of colors, caused by the wide range of hues in the elastograms. Different shapes and irregular distributions of colors could also make the visual perception problematic.

The study has several limitations.

Firstly, a small sample size was recruited, with a relative heterogeneity regarding the region and gravity of disease. However, patients were homogenous concerning breed, activity and the time of diagnosis from the injury onset.

Secondly, two-dimensional ultrasonography and SE are both considered highly operator-dependent methods and, especially for SE, several precautions had to be taken, as previously explained. For this reason, all the evaluations were performed by the same trained veterinarian (VS). In addition, strain elastography is a type of USE that shows a relative categorization of tissue elasticity. Other methods, such as SWE, which provide objective numerical parameters, could be more reliable, and the data obtained may be easier to analyze. Moreover, USE is still not commonly used in veterinary medicine; thus, more studies and comparisons are needed in order to standardize the method and subsequent analysis. However, this is still difficult due to the high variability of procedures and analysis techniques present in the literature.

No histopathologic assessment was performed, since all cases consisted of racehorses undergoing rehabilitation in order to return to competition.

Finally, image analysis using ImageJ was a long and complex procedure and several parameters needed to be considered. Despite the easy automatic selection and measurement of colored areas by the software, ranges of hue, saturation and brightness (Figure 1) had to be previously selected and tested by the operator, since a minimal change in range selection of each parameter can influence the results. To prove this, an evaluation was performed, considering different configurations of hue, brightness and saturation, showing significant differences between image categorizations. For this reason, a standardization of the selection of ranges is needed in order to make the assessment as accurate as possible, which should correspond as much as possible to the manual assessment. Furthermore, the low use of ImageJ and especially of the “Threshold Color” plugin in veterinary medicine provided few models for comparison.

5. Conclusions

This study demonstrates that the manual measurements of colors within elastograms of pathologic tendons showed good and excellent inter- and intraobserver reliability, respectively.

In addition, good agreement was also obtained between the manual and ImageJ analyses, for the evaluation of injured SDFTs in horses. We believe that this demonstrates that manual measurement is a feasible and accurate method and that it could be considered as valid as the objective software assessment. As an accurate interpretation and measurement of lesions is fundamental, it could thus be used to evaluate tendon healing.

Future studies are recommended to better evaluate the potential of SE for musculoskeletal structures and different imaging analysis techniques in horses.

Author Contributions: Conceptualization, V.S., G.M., A.C., A.B.F., M.L.P.P. and E.S.P.; methodology, V.S., G.M., A.C., F.C. and E.S.P.; software, V.S., A.C. and A.B.F.; formal analysis, E.S.P., V.S. and A.C.; investigation, V.S., A.C., F.C. and E.P.; writing—original draft preparation, V.S., A.C., E.P., E.S.P.; writing—review and editing, V.S., G.M., A.C., F.C., E.P., A.B.F., M.L.P.P., and E.S.P.; supervision, G.M., E.S.P. and M.L.P.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee (Organismo Preposto al Benessere Animale) of the University of Sassari (protocol code no. 128528, approved on 15 November 2019).

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available on request to the authors.

Acknowledgments: The authors thank Francesco Sanna Passino for helpful comments and discussion on the statistical analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. O'meara, B.; Bladon, B.; Parkin, T.D.H.; Fraser, B.; Lischer, C.J. An investigation of the relationship between race performance and superficial digital flexor tendonitis in the Thoroughbred racehorse. *Equine Vet. J.* **2010**, *42*, 322–326. [[CrossRef](#)]
2. Fortier, L.A.; Smith, R.K. Regenerative Medicine for Tendinous and Ligamentous Injuries of Sport Horses. *Vet. Clin. N. Am. Equine Pract.* **2008**, *24*, 191–201. [[CrossRef](#)]
3. Marr, C.M.; McMillan, I.; Boyd, J.S.; Wright, N.G.; Murray, M. Ultrasonographic and histopathological findings in equine superficial digital flexor tendon injury. *Equine Vet. J.* **1993**, *25*, 23–29. [[CrossRef](#)]
4. Avella, C.S.; Ely, E.R.; Verheyen, K.L.P.; Price, J.S.; Wood, J.L.N.; Smith, R.K.W. Ultrasonographic assessment of the superficial digital flexor tendons of National Hunt racehorses in training over two racing seasons. *Equine Vet. J.* **2009**, *41*, 449–454. [[CrossRef](#)]
5. Barr, R.G. Strain elastography. In *Breast Elastography*; Thieme Medical Publishers, Inc.: New York, NY, USA, 2015; pp. 11–30.
6. Lustgarten, M.; Redding, W.R.; Labens, R.; Morgan, M.; Davis, W.; Seiler, G.S. Elastographic characteristics of the metacarpal tendons in horses without clinical evidence of tendon injury. *Vet. Radiol. Ultrasound* **2014**, *55*, 92–101. [[CrossRef](#)]
7. Ooi, C.C.; Malliaras, P.; Schneider, M.E.; Connell, D.A. "Soft, hard, or just right?" Applications and limitations of axial-strain sonoelastography and shear-wave elastography in the assessment of tendon injuries. *Skelet. Radiol.* **2014**, *43*, 1–12. [[CrossRef](#)] [[PubMed](#)]
8. De Zordo, T.; Chhem, R.; Smekal, V.; Feuchtner, G.; Reindl, M.; Fink, C.; Faschingbauer, R.; Jaschke, W.; Klauser, A.S. Real-time sonoelastography: Findings in patients with symptomatic achilles tendons and comparison to healthy volunteers. *Ultraschall Med. Eur. J. Ultrasound* **2010**, *31*, 394–400. [[CrossRef](#)] [[PubMed](#)]
9. Winn, N.; Lalam, R.; Cassar-Pullicino, V. Sonoelastography in the musculoskeletal system: Current role and future directions. *World J. Radiol.* **2016**, *8*, 868. [[CrossRef](#)] [[PubMed](#)]
10. Busilacchi, A.; Olivieri, M.; Ulisse, S.; Gesuita, R.; Skrami, E.; Lording, T.; Fusini, F.; Gigante, A. Real-time sonoelastography as novel follow-up method in Achilles tendon surgery. *Knee Surg. Sports Traumatol. Arthrosc.* **2016**, *24*, 2124–2132. [[CrossRef](#)]
11. Tamura, N.; Kuroda, T.; Kotoyori, Y.; Fukuda, K.; Nukada, T.; Kato, T.; Kuwano, A.; Kasashima, Y. Application of sonoelastography for evaluating the stiffness of equine superficial digital flexor tendon during healing. *Vet. Rec.* **2017**, *180*, 120. [[CrossRef](#)]
12. Tamura, N.; Nukada, T.; Kato, T.; Kuroda, T.; Kotoyori, Y.; Fukuda, K.; Kasashima, Y. The use of sonoelastography to assess the recovery of stiffness after equine superficial digital flexor tendon injuries: A preliminary prospective longitudinal study of the healing process. *Equine Vet. J.* **2017**, *49*, 590–595. [[CrossRef](#)]
13. De Zordo, T.; Fink, C.; Feuchtner, G.M.; Smekal, V.; Reindl, M.; Klauser, A.S. Real-time sonoelastography findings in healthy Achilles tendons. *Am. J. Roentgenol.* **2009**, *193*, W134–W138. [[CrossRef](#)]

14. Klauser, A.S.; Miyamoto, H.; Tamegger, M.; Faschingbauer, R.; Moriggl, B.; Klima, G.; Feuchtner, G.M.; Kastlunger, M.; Jaschke, W.R. Achilles tendon assessed with sonoelastography: Histologic agreement. *Radiology* **2013**, *267*, 837–842. [[CrossRef](#)]
15. Smith, R.K.W.; Cauvin, E.R.J. Ultrasonography of the Metacarpus and Metatarsus. In *Atlas of Equine Ultrasonography*; Kidd, J.A., Lu, K.G., Frazer, M.L., Eds.; John Wiley & Sons, Ltd.: Chichester, UK, 2014; pp. 73–105.
16. Palgrave, K.; Kidd, J.A. Introduction. In *Atlas of Equine Ultrasonography*; Kidd, J.A., Lu, K.G., Frazer, M.L., Eds.; John Wiley & Sons, Ltd.: Chichester, UK, 2014; pp. 1–22.
17. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675. [[CrossRef](#)]
18. R Core Team. *R: A Language and Environment for Statistical Computing*; R Core Team: Vienna, Austria, 2019.
19. Aitchison, J. The Statistical Analysis of Compositional Data. *J. R. Stat. Soc. Ser. B.* **1982**, *44*, 139–177. [[CrossRef](#)]
20. Aitchison, J. *The Statistical Analysis of Compositional Data*; Chapman & Hall, Ltd.: London, UK, 1986.
21. Egozcue, J.J.; Pawlowsky-Glahn, V.; Mateu-Figueras, G.; Barceló-Vidal, C. Isometric Logratio Transformations for Compositional Data Analysis. *Math. Geol.* **2003**, *35*, 279–300. [[CrossRef](#)]
22. Van den Boogaart, K.G.; Tolosana-Delgado, R. *Analyzing Compositional Data with R*; Use R! Springer: Berlin/Heidelberg, Germany, 2013.
23. Van den Boogaart, K.G.; Tolosana-Delgado, R.; Bren, M. Compositions: Compositional Data Analysis. R Package Version 2.0-0. 2020. Available online: <https://CRAN.R-project.org/package=compositions> (accessed on 10 November 2020).
24. Martín-Fernández, J.A.; Barceló-Vidal, C.; Pawlowsky-Glahn, V. Dealing with Zeros and Missing Values in Compositional Data Sets Using Nonparametric Imputation. *Math. Geol.* **2003**, *35*, 253–278. [[CrossRef](#)]
25. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting Linear Mixed-Effects Models Using {lme4}. *J. Stat. Softw.* **2015**, *67*, 1–48. [[CrossRef](#)]
26. Korkmaz, S.; Goksuluk, D.; Zararsiz, G. MVN: An R Package for Assessing Multivariate Normality. *R J.* **2014**, *6*, 151–162. [[CrossRef](#)]
27. Fox, J.; Friendly, M.; Monette, G. Hplots: Visualizing Tests in Multivariate Linear Models. R Package Version 1.3-8. 2020. Available online: <https://CRAN.R-project.org/package=hplots> (accessed on 10 November 2020).
28. Nakagawa, S.; Johnson, P.C.D.; Schielzeth, H. The coefficient of determination R² and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J. R. Soc. Interface* **2017**, *14*, 20170213. [[CrossRef](#)]
29. Nordhausen, K.; Sirkia, S.; Oja, H.; Tyler, D.E. ICSNP: Tools for Multivariate Nonparametrics. R Package Version 1.1-1. 2018. Available online: <https://CRAN.R-project.org/package=ICSNP> (accessed on 10 November 2020).
30. Bland, J.M.; Altman, D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *1*, 307–310. [[CrossRef](#)]
31. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [[CrossRef](#)]
32. Prado-Costa, R.; Rebelo, J.; Monteiro-Barroso, J.; Preto, A.S. Ultrasound elastography: Compression elastography and shear-wave elastography in the assessment of tendon injury. *Insights Imaging* **2018**, *9*, 791–814. [[CrossRef](#)]
33. Fusini, F.; Langella, F.; Busilacchi, A.; Tudisco, C.; Gigante, A.; Massé, A.; Bisicchia, S. Real-time sonoelastography: Principles and clinical applications in tendon disorders. a systematic review. *Muscles. Ligaments Tendons J.* **2017**, *7*, 467–477. [[CrossRef](#)]
34. Klauser, A.; Faschingbauer, R.; Jaschke, W. Is Sonoelastography of Value in Assessing Tendons? *Semin. Musculoskelet. Radiol.* **2010**, *14*, 323–333. [[CrossRef](#)] [[PubMed](#)]
35. Patterson-Kane, J.C.; Firth, E.C. The pathobiology of exercise-induced superficial digital flexor tendon injury in Thoroughbred racehorses. *Vet. J.* **2009**, *181*, 79–89. [[CrossRef](#)]
36. Williams, I.F.; Heaton, A.; McCullagh, K.G. Cell morphology and collagen types in equine tendon scar. *Res. Vet. Sci.* **1980**, *28*, 302–310. [[CrossRef](#)]
37. Cheung, D.T.; Di Cesare, P.; Benya, P.D.; Libaw, E.; Nimni, M.E. The presence of intermolecular disulfide cross-links in type III collagen. *J. Biol. Chem.* **1983**, *258*, 7774–7778. [[CrossRef](#)]
38. Sharma, P.; Maffulli, N. Tendon injury and tendinopathy: Healing and repair. *J. Bone Jt. Surg. Ser. A* **2005**, *87*, 187–202.
39. Drakonaki, E.E.; Allen, G.M.; Wilson, D.J. Real-time ultrasound elastography of the normal Achilles tendon: Reproducibility and pattern description. *Clin. Radiol.* **2009**, *64*, 1196–1202. [[CrossRef](#)]
40. Piccionello, A.P.; Serrani, D.; Busoni, V.; Salvaggio, A.; Bonazzi, M.; Bergamino, C.; Volta, A. Sonoelastographic Features of the Patellar Ligament in Clinically Normal Dogs. *Vet. Comp. Orthop. Traumatol.* **2018**, *31*, 279–284. [[CrossRef](#)]
41. Del Signore, F.; De Dominicis, S.; Mastromatteo, G.; Simeoni, F.; Scapolo, P.A.; Tamburro, R.; Vignoli, M. Sonoelastography of Normal Canine Common Calcaneal Tendon: Preliminary Results. *Vet. Comp. Orthop. Traumatol.* **2020**, in press. [[CrossRef](#)]
42. Lustgarten, M.; Redding, W.R.; Labens, R.; Davis, W.; Daniel, T.M.; Griffith, E.; Seiler, G.S. Elastographic evaluation of naturally occurring tendon and ligament injuries of the equine distal limb. *Vet. Radiol. Ultrasound* **2015**, *56*, 670–679. [[CrossRef](#)]
43. De Gasperi, D.; Dzierzak, S.L.; Muir, P.; Vanderby, R.; Brounts, S.H. In vivo evaluation of effects of sedation on results of acoustoelastography of the superficial digital flexor tendons in clinically normal horses. *Am. J. Vet. Res.* **2017**, *78*, 1421–1425. [[CrossRef](#)]
44. Martínez-Payá, J.J.; Del Baño-Aledo, M.E.; Ríos-Díaz, J.; Fornés-Ferrer, V.; Vázquez-Costa, J.F. Sonoelastography for the Assessment of Muscle Changes in Amyotrophic Lateral Sclerosis: Results of a Pilot Study. *Ultrasound Med. Biol.* **2018**, *44*, 2540–2547. [[CrossRef](#)] [[PubMed](#)]

-
45. Teng, P.P.C.; Lo, Y.L. A Comparison Study of Conventional Ultrasound and Ultrasound Strain Elastography in the Evaluation of Myopathy. *Ultrasound Q.* **2020**, *36*, 32–37. [[CrossRef](#)] [[PubMed](#)]
 46. Ríos-Díaz, J.; Martínez-Payá, J.J.; Del Baño-Aledo, M.E.; De Groot-Ferrando, A.; Botía-Castillo, P.; Fernández-Rodríguez, D. Sonoelastography of Plantar Fascia: Reproducibility and Pattern Description in Healthy Subjects and Symptomatic Subjects. *Ultrasound Med. Biol.* **2015**, *41*, 2605–2613. [[CrossRef](#)] [[PubMed](#)]
 47. Magarelli, N.; Carducci, C.; Bucalo, C.; Filograna, L.; Rapisarda, S.; De Waure, C.; Dell’Atti, C.; Maccauro, G.; Leone, A.; Bonomo, L. Sonoelastography for qualitative and quantitative evaluation of superficial soft tissue lesions: A feasibility study. *Eur. Radiol.* **2014**, *24*, 566–573. [[CrossRef](#)] [[PubMed](#)]