

Objective evaluation versus subjective evaluation of flexion tests in the pelvic limb of horses

Avaliação objetiva versus avaliação subjetiva dos testes de flexão no membro pélvico de cavalos

Marcos da Silva Azevedo¹ ; Flávio Desessards De La Côte²; Ricardo Pozzobon²; Stefano Leite Dau²; Miguel Gallio²

¹ Universidade Federal do Pampa, Medicina Veterinária, Departamento de Clínica de Grandes Animais, Uruguaiana – RS, Brazil

² Universidade Federal de Santa Maria, Centro de Ciências Rurais, Departamento de Clínica de Grandes Animais, Santa Maria – RS, Brazil

ABSTRACT

The agreement between subjective and objective evaluation methods was studied to identify claudication in the pelvic limbs of horses before and after flexion tests were performed. Twenty-nine horses were equipped with a wireless system of inertial sensors and evaluated during seven times while trotting. Videos were recorded to be evaluated by three veterinarians, with different levels of experience, to evaluate the agreement between the two different methods and between the evaluators. The evaluators and the objective evaluation had a low rate of agreement, with the exception of moderate agreement between the objective evaluation and evaluator 1 to identify lameness after the left total flexion and moderate agreement in evaluating the response to the tests, between objective evaluation and evaluator 2, after right distal flexion. This shows that there was a low agreement among the evaluators and between them and the objective evaluation for identifying lameness, measuring the degree of lameness and the response to the flexion tests.

Keywords: Equine. Flexion test. Inertial sensors. Lameness Locator[®].

RESUMO

A concordância entre métodos de avaliação subjetiva e objetiva foi estudada para identificar claudicação nos membros pélvicos antes e após a realização dos testes de flexão. Vinte e nove cavalos foram equipados com um sistema de sensores inerciais sem fio e avaliados durante sete momentos enquanto troteavam. Vídeos foram gravados para serem avaliados por três veterinários com diferentes níveis de experiência, a fim de avaliar a concordância entre os dois métodos diferentes e entre os avaliadores. Os avaliadores e a avaliação objetiva tiveram uma baixa taxa de concordância, com exceção de uma moderada concordância entre avaliação objetiva e avaliador 1 para identificar claudicação após o teste de flexão total esquerdo e uma moderada concordância na avaliação da resposta aos testes, entre avaliação objetiva e avaliador 2, após o teste de flexão distal direito. Isso demonstra que houve uma baixa concordância entre os avaliadores e entre estes e a avaliação objetiva para identificação da claudicação, mensuração do grau de claudicação e resposta aos testes de flexão.

Palavras-chave: Equino. Teste de flexão. Sensores inerciais. Lameness Locator[®].

Correspondence to:

Marcos da Silva Azevedo
Universidade Federal do Pampa, Departamento de Clínica de
Grandes Animais
BR 472, Km 585
CEP: 97501-970, Uruguaiana – RS, Brazil
e-mail: socram_vet@yahoo.com.br

Received: April 23, 2019

Approved: August, 26, 2019

How to cite: Azevedo MS, De La Côte FD, Pozzobon R, Dau SL, Gallio M. Objective evaluation versus subjective evaluation of flexion tests in the pelvic limb of horses. *Braz J Vet Res Anim Sci.* 2019;56(4):e157248. <https://doi.org/10.11606/issn.1678-4456.bjvras.2019.157248>

Introduction

Unquestionably, lameness is the most important medical problem in athletic horses, frequently being responsible for determining the discontinuation of training or the

removal of these animals from competition, generating significant financial losses to the owner (Keegan et al., 1998; McIlwraith, 2002; Ross, 2011b).

The early identification of the lame limb and the painful area responsible for this alteration is important to quickly establish effective treatment. Among the procedures carried out in the physical exam of the locomotor system are flexion tests, which are routinely used to apply pressure to one joint or a group of joints, frequently provoking or accentuating lameness that is not easily observable before performing the test (Keegan et al., 1998; Ross, 2011b).

Pelvic limb flexion tests are very useful for exacerbating initial lameness or detecting potential sources of lameness. However, positive responses after distal or proximal flexion tests of the pelvic limb are not considered specific for distinguishing the source of the pain causing lameness, since it is almost impossible to flex a single joint without flexing other joints or affecting adjacent structures (Ross, 2011a).

It is known that the interpretation of the results of flexion tests is inherently subjective and, therefore, subject to a considerable amount of variation between observers. This is why objective methods of analyzing lameness are highly relevant in clinical investigations, justifying the emergence of several computerized systems to identify and quantify the lameness (Armentrout et al., 2012; Keegan et al., 2004). Among them, systems that use inertial sensors have demonstrated the ability to identify alterations of asymmetry in the movement of the pelvis before and after performing flexion tests and would be very useful in documenting small differences between the tests carried out (Marshall et al., 2012).

The objective of this study was to determine if there was agreement between the objective method of evaluating lameness using wireless inertial sensors and three veterinarians with different levels of experience in examining lameness in horses, regarding the identification of lameness after performing flexion tests in the hindlimb.

Materials and Methods

Committee on Animal Research and Ethics approval for this project was obtained from Federal University of Santa Maria (protocol 042/2014).

Animals

Twenty-nine horses (15 male and 14 female), between 4-14 years of age, and an average weight of 420 ± 73 Kg, of different breeds and sporting activities, were used in this study. The criterion for inclusion of the horses was that at least one of the three evaluators or the objective evaluation

identified lameness in one of the pelvic limbs during the baseline evaluation.

Data collection and objective analysis

The objective lameness evaluation was based on a system of wireless inertial sensors (Lameness Locator[®]), while the instrumentation of the horses, as well as the collection and analysis of the data were conducted as recommended (Keegan et al., 2011). Each horse was equipped with an accelerometer on the dorsal aspect of the head and another on the sacral tuberosity, on the dorsal midline. Finally, a gyroscope was attached to the dorsal aspect of the pastern of the right thoracic limb (Figure 1).

The data collection was completed on a concrete floor with the animals being pulled by the halter to a trot in a straight line, for a distance of 25-30 m, so that at least 10-15 steps could be obtained. First, a pre-flexion evaluation was done, also called the baseline. Next, the horses were submitted to the flexion tests and evaluated again as previously described. The tests always started with the pelvic limb that did not have noticeable lameness, taking precautions to wait at least two minutes before performing a new test on the contralateral limb. The same professional performed all the flexion tests so as to minimize the changes in strength during performing the tests. The sequence of the flexion test was done to minimize the flexion of the other joints of the pelvic limb as much as possible. The first test performed was the left or right flexion test of the stifle joints for one minute, then, the left or right distal flexion test of the pelvic limb for thirty seconds and, finally, the left or right full flexion test of the pelvic limb, also called the "spavin test" for one minute (Baxter & Stashak, 2011). These tests were performed as shown in Figure 1.

The data collected were transmitted in real time, via Bluetooth, to a tablet where a specific software analyzed the data (Keegan et al., 2011). This analysis comprehends the calculation of the following variables: PDmax for the pelvic limbs (mean and standard deviation) – the difference between the highest point of the pelvis after support of the right pelvic limb and the highest point of the pelvis after support of the left pelvic limb; PDmin for the pelvic limbs (mean and standard deviation) – the difference between the lowest point of the pelvis during support of the right pelvic limb and the lowest point of the pelvis during support of the left pelvic limb. The data obtained by the gyroscope positioned on the digit of the right thoracic limb served to detect the position of the digit of the right thoracic limb and to infer the position of each one of the limbs, since the trot is a symmetrical gait in which the limbs move in diagonal pairs.



Figure 1 – Horse equipped with a system of wireless inertial sensors (Lameness Locator[®]) and flexion test performed in the pelvic limb. (A) Head sensor; (B) Pelvic sensor; (C) Pastern sensor; (D) Stifle flexion; (E) Distal flexion; (F) Total flexion.

To evaluate the pre-flexion and after each of the flexion tests, the horses were classified as follows: positive lameness (2) or negative lameness (1), considering that the horses were considered to have positive lameness when at least one of the variables (PDmax or PDmin) was higher than the reference value (3 mm) and the standard deviation was lower than the mean. The identification of the limb with lameness was established based on the PDmax and PDmin signal, where a negative signal corresponded to the left pelvic limb (1) and a positive signal corresponded to the right pelvic limb (2). When the animal had bilateral lameness, we took as reference for the evaluations only the limb with a greater intensity of lameness. After performing the flexion tests, the response was classified as negative (1) and positive (2), when there was an increase in the PDmax or PDmin values higher than the reference value of 3mm.

Subjective evaluation

Simultaneous to the objective evaluation, the exams were recorded and video clips (.MPG from Windows Media Player) were generated for each of the evaluations completed in a straight line with the horses trotting away from the camera. The videos could be reviewed by the evaluator as many times as necessary, although once the evaluation was completed, the videos could no longer be consulted. The order of evaluation started with the pre-flexion video, followed by the left flexion test of the stifle, right flexion test of the stifle, left distal flexion test, right distal flexion test, left full flexion test and right full flexion test evaluation videos.

Three evaluators with different levels of experience were asked to classify the horses as having positive (2) or negative (1) lameness, to identify the limb with lameness: left (1) or right (2) and the response for the flexion test as negative (1) and positive (2) for each test performed. Evaluator 1 has more than 20 years of experience, evaluator 2 has seven years and evaluator 3 has six years of experience. For evaluation of the pre-flexion video, the evaluators were also asked to estimate the lameness intensity on the AAEP scale (American Association of Equine Practitioners, 1999).

Statistical analysis

We used the Kappa agreement test at a 95% level of significance. The agreement was evaluated between the objective evaluation and each of the evaluators, as well as the agreement between the different evaluators in identifying lameness, the limb with lameness and a response of each one of flexion test. The level of agreement was based on the following interpretation, where the kappa values between 0-0.20 translated into a slight agreement; 0.21-0.40 a considerable agreement, 0.41-0.60 moderate agreement; 0.61-0.80 substantial agreement and over 0.81 exceptional agreement (Landis & Koch, 1977).

Results

At the baseline, the objective evaluation showed that 19 (65.52%) of the horses showed lameness, while on the subjective evaluation for evaluator 1; 22 (75.86%), evaluator 2; 14 (48.28%) and evaluator 3; 26 (89.66%) had lameness (Table 1).

The identification of the lameness degree for the 29 horses by the evaluators, in the pre-flexion baseline evaluation, demonstrated considerable agreement ($k=0.22$), with $P<0.05$ for evaluator 1 and evaluator 2. The agreement between the other evaluators was slight, although the kappa values were insignificant (Table 2). During the baseline evaluation, the identification of whether the horse was lame or not in one of the pelvic limbs was of considerable agreement ($k=0.32$), with $P<0.05$ between evaluator 1 and evaluator 2, as well as between objective evaluation and evaluator 3 ($k=0.36$), with $P<0.05$. While the agreement between objective evaluation and evaluator 2 was considerable agreement ($k=0.32$), it was not significant. Evaluator 1 presented a failure in agreement in relation to the objective evaluation (Table 3). The correct identification of the lame pelvic limb was of considerable agreement ($k=0.28$), with $P<0.05$ between evaluator 1 and evaluator 3, as well as between objective evaluation and evaluator 3 ($k=0.36$), with $P<0.05$ (Table 4).

After performing the left flexion test of the stifle, none of the levels of agreement were significant in identifying lameness in the pelvic limbs, while the identification of the lame limb was of considerable agreement ($k=0.29$), with $P<0.05$ between objective evaluation and evaluator 3 (Table 4). The response rate for the left flexion test of the stifle presented no significant agreement.

The evaluations after the right flexion test of the stifle test demonstrated considerable agreement between the objective evaluation and evaluator 2 ($k=0.32$), with $P<0.05$ and between evaluator 1 and evaluator 2 ($k=0.34$ and $P<0.05$) for identifying lameness (Table 3). Considerable agreement for the lame limb was determined between evaluator 1 and evaluator 2 ($k=0.32$) and $P<0.05$ (Table 4). The response rate after performing the right flexion test of the stifle test also had considerable agreement between the objective evaluation and evaluator 2 ($k=0.37$ and $P<0.05$) and between evaluator 1 and evaluator 2 ($k=0.34$ and $P<0.05$) (Table 5).

Table 1 – Number and percentage of animals with lameness identified by the objective evaluation and subjective evaluators after each flexion test performed, Uruguaiiana, Apr. 2019

Test/Lameness	OE		EV 1		EV 2		EV 3	
	No	Yes	No	Yes	No	Yes	No	Yes
BL	10 (34.48)	19 (65.52)	7 (24.14)	22 (75.86)	15 (51.72)	14 (48.28)	3 (10.34)	26 (89.66)
LSF	4 (13.79)	25 (86.21)	11 (37.93)	18 (62.07)	13 (44.83)	16 (55.17)	7 (24.14)	22 (75.86)
RSF	13 (44.83)	16 (55.17)	11 (37.93)	18 (62.07)	17 (58.62)	12 (41.38)	14 (48.28)	15 (51.72)
LDF	8 (27.59)	21 (72.41)	25 (86.21)	4 (13.79)	16 (55.17)	13 (44.83)	15 (51.72)	14 (48.28)
RDF	8 (27.59)	21 (72.41)	22 (75.86)	7 (24.14)	23 (79.31)	6 (20.69)	14 (48.28)	15 (51.72)
LTF	7 (24.14)	22 (75.86)	7 (24.14)	22 (75.86)	16 (55.17)	13 (44.83)	1 (3.45)	28 (96.55)
RTF	12 (41.38)	17 (58.62)	4 (13.79)	25 (86.21)	9(31.03)	20 (68.97)	5 (17.24)	24 (82.76)

OE: Objective evaluation; EV 1: Evaluator 1; EV 2: Evaluator 2; EV 3: Evaluator 3; BL: Baseline evaluation; LSF: Left stifle flexion; RSF: Right stifle flexion; LDF: Left distal flexion; RDF: Right distal flexion; LTF: Left total flexion; RTF: Right total flexion.

Table 2 – Description of the kappa values, P and level of agreement for the degree of lameness of the baseline evaluation (BL), Uruguaiiana, Apr. 2019

	EV 1 vs. EV 2		EV 1 vs. EV 3		EV 2 vs. EV 3	
	k	Level of agreement	k	Level of agreement	k	Level of agreement
BL	0.22*	C	0.07	S	0.08	S

EV 1: Evaluator 1; EV 2: Evaluator 2; EV 3: Evaluator 3; k: Kappa index; C: Considerable; S: Slight; BL: Baseline evaluation; * $P<0.05$.

Table 3 – Description of the kappa values, P and level of agreement for identifying lameness, Uruguaiiana, Apr. 2019

	OE vs. EV 1		OE vs. EV 2		OE vs. EV 3		EV 1 vs. EV 2		EV 1 vs. EV 3		EV 2 vs. EV 3	
	k	Level of agreement	k	Level of agreement	k	Level of agreement	k	Level of agreement	k	Level of agreement	k	Level of agreement
BL	-0.68	E	0.25	C	0.36*	C	0.32*	C	0.07	S	0.06	S
LSF	0.25	C	0.03	S	0.23	C	0.29	C	-0.10	E	-0.17	E
RSF	0.15	S	0.32*	C	-0.04	E	0.34*	C	0.1	S	-0.03	E
LDF	0.12	S	0.21	C	-0.02	E	0.33*	C	0.01	S	0.24	C
RDF	0.07	S	0.15	S	0.16	S	-0.01	E	0.12	S	0.19	S
LTF	0.44*	M	0.28*	C	-0.06	E	0.15	S	-0.06	E	-0.07	E
RTF	-0.04	E	0.04	S	0.3	C	0.01	S	0.14	S	-0.10	E

OE: Objective evaluation; EV 1: Evaluator 1; EV 2: Evaluator 2; EV 3: Evaluator 3; BL: Baseline evaluation; LSF: Left stifle flexion; RSF: Right stifle flexion; LDF: Left distal flexion; RDF: Right distal flexion; LTF: Left total flexion; RTF: Right total flexion; k: Kappa index; C: Considerable; E: Error; S: Slight; M: Moderate; * $P<0.05$.

Table 4 – Description of the kappa values, *P* and level of agreement for identifying the lameness limb, Uruguaiiana, Apr. 2019

	OE vs. EV 1		OE vs. EV 2		OE vs. EV 3		EV 1 vs. EV 2		EV 1 vs. EV 3		EV 2 vs. EV 3	
	k	Level of agreement	k	Level of agreement	k	Level of agreement	k	Level of agreement	k	Level of agreement	k	Level of agreement
BL	0.18	S	0.11	S	0.36*	C	0.17	S	0.28*	C	0.00	S
LSF	0.17	S	-0.03	E	0.29*	C	0.29	C	-0.13	E	-0.12	E
RSF	0.11	S	0.28	C	0.02	S	0.32*	C	0.22	C	0.07	S
LDF	0.05	S	0.08	S	0.13	S	0.33*	C	0.01	S	0.24	C
RDF	0.06	S	0.16*	S	0.07	S	-0.01	E	0.10	S	0.23	C
LTF	0.18	S	0.29*	C	-0.05	E	0.15	S	-0.06	E	-0.07	E
RTF	-0.02	E	0.00	S	0.18*	S	0.01	S	0.14	S	-0.10	E

OE: Objective evaluation; EV 1: Evaluator 1; EV 2: Evaluator 2; EV 3: Evaluator 3; BL: Baseline evaluation; LSF: Left stifle flexion; RSF: Right stifle flexion; LDF: Left distal flexion; RDF: Right distal flexion; LTF: Left total flexion; RTF: Right total flexion; k: Kappa index; C: Considerable; E: Error; S: Slight; * $P < 0.05$.

Table 5 – Description of the kappa values, *P* and level of agreement for response to flexion test, Uruguaiiana, Apr. 2019

	OE vs. EV 1		OE vs. EV 2		OE vs. EV 3		EV 1 vs. EV 2		EV 1 vs. EV 3		EV 2 vs. EV 3	
	k	Level of agreement	k	Level of agreement	k	Level of agreement	k	Level of agreement	k	Level of agreement	k	Level of agreement
LSF	0.29	C	0.19	S	0.11	S	0.31	C	-0.08	E	-0.17	E
RSF	0.13	S	0.37*	C	0.04	S	0.34*	C	0.01	S	-0.03	E
LDF	0.10	S	0.25	C	0.18	S	0.33*	C	0.01	S	0.24	C
RDF	0.05	S	0.41*	M	0.04	S	-0.01	E	0.12	S	0.19	S
LTF	0.13	S	0.25	C	-0.07	E	0.15	S	-0.06	E	-0.07	E
RTF	0.06	S	0.05	C	0.19	S	0.01	C	0.14	S	-0.10	E

OE: Objective evaluation; EV 1: Evaluator 1; EV 2: Evaluator 2; EV 3: Evaluator 3; LSF: Left stifle flexion; RSF: Right stifle flexion; LDF: Left distal flexion; RDF: Right distal flexion; LTF: Left total flexion; RTF: Right total flexion; k: Kappa index; C: Considerable; E: Error; S: Slight; M: Moderate; * $P < 0.05$.

For the left distal flexion test, there was considerable agreement between evaluator 1 and evaluator 2 ($k=0.33$ and $P < 0.05$ regarding the identification of lameness (Table 3), the lame limb (Table 4) and response rate (Table 5).

There was no considerable agreement for identifying lameness or the limb after performing the right distal flexion test. However, this test presented a moderate agreement between objective evaluation and evaluator 2 ($k=0.41$) and $P < 0.05$ (Table 5) to the response rate after performing the test.

After performing the left full flexion test, moderate agreement was determined, regarding the identification of lameness between objective evaluation and evaluator 1 ($k=0.44$ and $P < 0.05$) and considerable agreement between objective evaluation and evaluator 2 ($k=0.28$) and $P < 0.05$ (Table 3). Next, identification of the limb was only considerable between objective evaluation and evaluator 2 ($k=0.29$ and $P < 0.05$) (Table 3). On the other hand, the response rate after performing the test did not present any significant agreement between the objective and subjective evaluations. Performance of the right full flexion test did not demonstrate any considerable agreement between the objective evaluation and subjective evaluators regarding the identification of lameness, identification of the limb or the response rate after performing the test.

Discussion

The main goal of this study was to determine the degree of agreement between three evaluators with different levels of experience, and a wireless system of inertial sensors to identify lameness after performing three different flexion tests in the pelvic limb. The subjective evaluation, while still being the most frequently used method to evaluate lameness, has demonstrated an important variability of results that relies on the intensity of lameness and the experience of the evaluator (Keegan et al., 2010; McCracken et al., 2012). On the other hand, the use of wireless inertial sensors has shown to be effective in identifying subtle changes in asymmetry after performing flexion tests in the pelvic limb (Marshall et al., 2012).

Lameness in the pelvic limbs can be difficult to identify through subjective evaluations, since asymmetry in movement between the coxal tuberosities is considered the key point for this identification (Baxter & Stashak, 2011). However, some factors may complicate this evaluation process, such as: the horse moving away from the evaluator, intermittent lameness, and asymmetries between contralateral limbs may not be visible due to the low range and, also, even healthy horses may present asymmetries in the movement of coxal tuberosities (Parkes et al., 2009). The relatively low agreement found between the evaluators of this study in relation to identifying the lame limb, and between them

and the objective evaluation may be in part explained by these factors, as well as by the experience of the evaluators. A study demonstrated that the ability to correctly recognize asymmetrical movement seems to improve with experience, as qualitatively demonstrated by perception tests using data from horses with visible lameness and horses with subtle or hard to visualize lameness, where the equine practitioner with the most experience made significantly better identifications than the physician with less experience (Parkes et al., 2009). The authors also compared the wireless system based on inertial sensors with the subjective evaluation and found a slight to exact agreement for identifying lameness in the pelvic limbs between the different evaluation methods and between the subjective evaluators, with the level of experience having a positive influence on the identification of lameness (Keegan et al., 2010). In the present study, there was no correlation between the level of experience and the identification of lameness, as well as identification of the lame limb and degree of lameness.

This can be seen, for example, in the baseline evaluation, where evaluator 3 had a considerable correlation with objective evaluation in identifying lameness and later in identifying the lameness limb and evaluator 1 had only a slight correlation for the same evaluations. One of the possible explanations for this is that equine practitioners with more experience may take into consideration other parameters, such as joint flexion and step length, during the evaluation and identification of lameness (McCracken et al., 2012).

The performance of different flexion tests for the pelvic limb did not make it easier for the evaluators to agree with the objective, to identify lameness, given that most of the evaluations had a slight to considerable correlation, with exception of the right distal flexion test, which had a moderate correlation between evaluator 2 and the objective evaluation. Our results are not that different from another study that found great variability between the evaluators in the interpretation of the flexion tests in the pelvic limb, when performed for 5 or 60 seconds, given that individually these evaluators classified as positive responses from 21 to 81% of all the tests performed (Armentrout et al., 2012). According to this study, the existing variability between the evaluation results after the flexion tests is due to the fact that many clinicians are excessively critical and consider any small change in the gait to be a positive response, while other clinicians may disregard these small changes (such as a positive response for only a few steps) (Armentrout et al., 2012).

Though the evaluation through videos has demonstrated a good agreement between the scores given during the lameness exam with those attributed after looking at the videos, in a blind study there is a risk of the evaluator's ability being affected when the interpretation of the exam is done only through videos (Fuller et al., 2006). Among the factors that may compromise the evaluation is the fact that the horse is frequently recorded in only one direction and under adverse conditions in terms of lighting and the quality of the videos, as well as the audio and image (Keegan et al., 1998).

One study determined that the human ability to perceive lameness, based on asymmetry, is limited. It also demonstrated that the use of other techniques, such as diagnostic blocks and analysis by a second clinician, would be interesting (Parkes et al., 2009). However, several studies have shown that insignificant asymmetries are difficult to detect, and there is disagreement between observers for subtle lameness (Keegan, 2007). These data corroborate those seen in our study. The use of other techniques, such as kinetics or kinematics, may be useful in the diagnosis of subtle lameness (Parkes et al., 2009).

Conclusions

In conclusion, regardless of the level of experience, we found a low agreement between evaluators and between them and the objective evaluation method for identifying lameness, measuring the degree of lameness and the response to the flexion tests. These results demonstrate the need for improvement regarding the subjective evaluation of lameness and observation of other parameters besides just the asymmetry of pelvic movement, to assist in this identification.

Conflict of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Ethics Statement

The investigation was approved by the Ethical Commission of use of animals in experimentation from the Universidade de Santa Maria, protocol: 042/2014.

Acknowledgements

The authors thank Lucas Campara, DVM, and Guilherme Freitas, DVM, at the Military Circle, Uruguaiana, for their contribution to this study.

References

- AAEP: American Association of Equine Practitioners. Guide to veterinary services for horse shows. 7th ed. Lexington, Kentucky: AAEP; 1999.
- Armentrout AR, Beard WL, White BJ, Lillich J. D. A comparative study of proximal hindlimb flexion in horses: 5 versus 60 seconds. *Equine Vet J*. 2012;44(4):420-4. <http://dx.doi.org/10.1111/j.2042-3306.2011.00474.x>. PMID:21923880.
- Baxter GM, Stashak TS. Examination for lameness. In: Baxter GM Adams and Stashak's lameness in horses. 6th ed. West Sussex: Blackwell Publishing; 2011. p. 128-233.
- Fuller CJ, Bladon BM, Driver AJ, Barr AR. The intra- and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses. *Vet J*. 2006;171(2):281-6. <https://10.1016/j.tvjl.2004.10.012>.
- Keegan KG, Dent EV, Wilson DA, Janicek J, Kramer J, Lacarrubba A, Walsh DM, Cassells MW, Esther TM, Schiltz P, Frees KE, Wilhite CL, Clark JM, Pollitt CC, Shaw R, Norris T. Repeatability of subjective evaluation of lameness in horses. *Equine Vet J*. 2010;42(2):92-7. <https://10.2746/042516409X479568>.
- Keegan KG, Kramer J, Yonezawa Y, Maki H, Pai PF, Dent EV, Kellerman TE, Wilson DA, Reed SK. Assessment of repeatability of a wireless, inertial sensor-based system for horses. *Am J Vet Res*. 2011;72(9):1156-63. <https://10.2460/ajvr.72.9.1156>.
- Keegan KG, Wilson DA, Wilson DJ, Smith B, Gaughan EM, Pleasant RS, Lillich JD, Kramer J, Howard RD, Bacon-Miller C, Davis EG, May KA, Cheramie HS, Valentino WL, Van Harreveld PD. Evaluation of mild lameness in horses trotting on a treadmill by clinicians and interns or residents and correlation of their assessments with kinematic gait analysis. *Am J Vet Res*. 1998;59(11):1370-7. PMID:9829392.
- Keegan KG, Yonezawa Y, Pai PF, Wilson DA, Kramer J. Evaluation of a sensor-based system of motion analysis for detection and quantification of forelimb and hind limb lameness in horses. *Am J Vet Res*. 2004;65(5):665-70. <https://doi:10.2460/ajvr.2004.65.665>
- Keegan KG. Evidence-based lameness detection and quantification. *Vet Clin North Am Equine Pract*. 2007;23(2):403-23. <https://10.1016/j.cveq.2007.04.008>.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74. <http://dx.doi.org/10.2307/2529310>. PMID:843571.
- Marshall JF, Lund DG, Voute LC. Use of a wireless, inertial sensor-based system to objectively evaluate flexion tests in the horse. *Equine Vet J Suppl*. 2012;44(43):8-11. <https://10.1111/j.2042-3306.2012.00611.x>.
- McCracken MJ, Kramer J, Keegan KG, Lopes M, Wilson DA, Reed SK, Lacarrubba A, Rasch M. Comparison of an inertial sensor system of lameness quantification with subjective lameness evaluation. *Equine Vet J*. 2012;44(6):652-6. <https://10.1111/j.2042-3306.2012.00571.x>.
- McIlwraith CW. Diseases of joints, tendons, ligaments and related structures. In: Stashak TS. Adams' lameness in horses. 5th ed. Philadelphia: Lippincott Williams and Wilkins; 2002. p. 459- 644.
- Parkes RS, Weller R, Groth AM, May S, Pfau T. Evidence of the development of 'domain-restricted' expertise in the recognition of asymmetric motion characteristics of hindlimb lameness in the horse. *Equine Vet J*. 2009;41(2):112-7. <https://10.2746/042516408X343000>.
- Ross MW. Manipulation. In: Ross MW, Dyson SJ. Diagnosis and management of lameness in the horse. 2nd ed. Philadelphia: WB Saunders; 2011a. p. 80-8. <http://dx.doi.org/10.1016/B978-1-4160-6069-7.00008-0>.
- Ross MW. The lameness examination. In: Ross MW, Dyson SJ. Diagnosis and management of lameness in the horse. 2nd ed. Philadelphia: WB Saunders; 2011b. p. 1-79.

Financial Support: None.

Authors Contributions

Marcos da Silva Azevedo: Lead author, project design, data collection and writing of the manuscript. Flávio Desessards De La Côte: Assistance in project design and manuscript writing. Ricardo Pozzobon, Stefano Leite Dau and Miguel Gallio: Assistance in data collection.