

## Article

# Linear Discriminant Analysis for Investigating Differences in Upper Body Movement Symmetry in Horses before/after Diagnostic Analgesia in Relation to Expert Judgement

Thilo Pfau <sup>1,2,3,\*</sup> , David M. Bolt <sup>1</sup>, Andrew Fiske-Jackson <sup>1</sup> , Carolin Gerdes <sup>4</sup>, Karl Hoenecke <sup>1</sup>, Lucy Lynch <sup>1</sup>, Melanie Perrier <sup>1</sup> and Roger K. W. Smith <sup>1</sup> 

<sup>1</sup> Department of Clinical Science and Services, The Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield AL9 7TA, UK; dbolt@rvc.ac.uk (D.M.B.); afiskejackson@rvc.ac.uk (A.F.-J.); khoenecke7@rvc.ac.uk (K.H.); llynch6@rvc.ac.uk (L.L.); mperrier@rvc.ac.uk (M.P.); rksmith@rvc.ac.uk (R.K.W.S.)

<sup>2</sup> Faculty of Veterinary Medicine, University of Calgary, 2500 University Dr NW, Calgary, AB T2N 1N4, Canada

<sup>3</sup> Faculty of Kinesiology, University of Calgary, 2500 University Dr NW, Calgary, AB T2N 1N4, Canada

<sup>4</sup> Pferdeklinik Hochmoor GmbH, D-48712 Gescher, Germany; c.gerdes@pferdeklinik-hochmoor.de

\* Correspondence: thilo.pfau@ucalgary.ca

**Simple Summary:** Identifying the anatomical structures involved in causing pain and an associated lameness in a horse typically requires assessment in straight lines and circles and using regional administration of local anesthetic drugs (diagnostic analgesia). Visual assessment of changes in movement are affected by bias, i.e., expected changes influence decisions. Quantitative measurements with inertial sensors aim at removing this bias. The current study is aimed at investigating how a specific data-driven method, linear discriminant analysis (LDA), may be useful for aiding veterinary decision making about perceived changes in lameness. Changes in movement data after diagnostic analgesia and expert judgements from 53 lame horses were used to study (a) the accuracy of LDA-based decision making, (b) differences between straight-line and circular movement and (c) which commonly used movement features are most useful in this context. Accuracy was comparatively low and varied considerably between 36% and 57%, indicating considerable overlap between movement symmetry data of the diagnostic analgesia categories. The best data-driven separation between categories was observed when the limb in which perineural anaesthesia had been performed was on the inside of the circle (on hard ground for forelimb and on soft ground for hindlimb diagnostic analgesia). Movement features of all three landmarks (head, withers, pelvis) were important for data-driven classification, emphasizing the complexity of the movement pattern changes after diagnostic analgesia observed in lame horses.

**Abstract:** Diagnostic analgesia and lunging are parts of the equine lameness examination, aiding veterinarians in localizing the anatomical region(s) causing pain-related movement deficits. Expectation bias of visual assessment and complex movement asymmetry changes in lame horses on the lunge highlight the need to investigate data-driven approaches for optimally integrating quantitative gait data into veterinary decision-making to remove bias. A retrospective analysis was conducted with inertial sensor movement symmetry data before/after diagnostic analgesia relative to subjective judgement of efficacy of diagnostic analgesia in 53 horses. Horses were trotted on the straight and on the lunge. Linear discriminant analysis (LDA) applied to ten movement asymmetry features quantified the accuracy of classifying negative, partial and complete responses to diagnostic analgesia and investigated the influence of movement direction and surface type on the quality of the data-driven separation between diagnostic analgesia categories. The contribution of movement asymmetry features to decision-making was also studied. Leave-one-out classification accuracy varied considerably (38.3–57.4% for forelimb and 36.1–56.1% for hindlimb diagnostic analgesia). The highest inter-category distances (best separation) were found with the blocked limb on the inside of the circle, on hard ground for forelimb diagnostic analgesia and on soft ground for hindlimb diagnostic analgesia. These exercises deserve special attention when consulting quantitative gait



**Citation:** Pfau, T.; Bolt, D.M.; Fiske-Jackson, A.; Gerdes, C.; Hoenecke, K.; Lynch, L.; Perrier, M.; Smith, R.K.W. Linear Discriminant Analysis for Investigating Differences in Upper Body Movement Symmetry in Horses before/after Diagnostic Analgesia in Relation to Expert Judgement. *Animals* **2022**, *12*, 762. <https://doi.org/10.3390/ani12060762>

Academic Editors: Andrea Bertuglia and Sybille Molle

Received: 5 February 2022

Accepted: 16 March 2022

Published: 17 March 2022

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



**Copyright:** © 2022 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/>).

data in lame horses. Head and pelvic upward movement and withers minimum differences were the features with the highest weighting within the first canonical LDA function across exercises and forelimb and hindlimb diagnostic analgesia. This highlights that movement changes after diagnostic analgesia affect the whole upper body. Classification accuracies based on quantitative movement asymmetry changes indicate considerable overlap between subjective diagnostic analgesia categories.

**Keywords:** horse; lameness; diagnostic analgesia; straight-line; circular movement; linear discriminant analysis

---

## 1. Introduction

A major component of the equine veterinary lameness examination is the use of diagnostic analgesia, i.e., perineural (“nerve blocking”) and/or intra-synovial (“joint blocking”) anesthesia, with the aim of providing information that will aid the decision making about which part(s) of a limb is (are) experiencing discomfort during locomotion [1]. In non-verbal patients (animals) this is particularly important for guiding diagnostic imaging in terms of which anatomical region to focus on and which technology to choose (e.g., focusing on soft tissue versus bony structures) and thus influencing treatment options. Relating the visual appearance of a lameness to findings from palpation or diagnostic imaging is crucial in order to be able to judge the ‘functional’ relevance of any identified abnormalities [2].

In addition to diagnostic analgesia, diagnosis of the source of pain causing a lameness is assisted by exercising the horse under different conditions, e.g., in circles on hard and soft ground [3,4]. Clinical experience indicates that the origin of a lameness may result in more or less severe changes in relation to the affected limb positioned on the inside or outside of the circle and/or to the surface characteristics [3]. The first may be related to the difference in extra-sagittal angulation and/or force between the inside and the outside limb [5,6] and/or the amount of body lean [7]. The second may be related to differences in impact accelerations/vibrations [8–10], in stance time, stride time or duty factor [10] and consequently peak force [11,12] and/or changes in breakover [13].

In the lameness examination, both techniques—diagnostic analgesia and circular exercise on the lunge—are often combined. Studies with quantitative gait analysis have established guideline values for the amount of head (6 mm) and pelvic (3 mm) movement asymmetry [14], which, when exceeded in a horse presented to a veterinarian, can be used to identify the most likely limb causing the lameness. Importantly, these movement asymmetries have also been associated with force imbalances [15,16]. Adding quantitative assessments relating to the effects of diagnostic analgesia may help with reducing the amount of expectation bias that has been documented for visual decision making [17] and assist in determining partial improvements confidently, overcoming the limited sensitivity of the human visual system when assessing movement symmetry [18]. Previous quantitative data from lunge exercise have provided information about variations between repeat measurements [19]; the expected changes as a function of speed, surface and circle radius [20,21]; changes after induction of lameness [22] and associations between asymmetries on the lunge with pre-existing movement asymmetries measured during straight line locomotion [23,24]. However, relevant studies with quantitative gait data combining diagnostic analgesia with lunge exercise are lacking.

The aim of this study was to investigate the discriminative power of upper body movement symmetry differences calculated before and after administration of diagnostic analgesia for differentiating between blocks judged subjectively according to their efficacy by experienced veterinarians during a lameness exam. The study objectives were to quantify differences in upper body movement symmetry before and after diagnostic analgesia for different exercise conditions typically encountered during the lameness examination, such as straight-line trot and trot in circles on soft and hard ground and to categorize block efficacy as ‘negative’ (subjectively no improvement: 0–30% [25]), ‘partially positive’

(subjectively some improvement but with residual lameness: >30–70% [25]) and ‘positive’ (subjectively lameness resolved or switched to different limb: >70–100% [25]). This was achieved by consulting the case records and applying a data-driven feature transformation and classification method—linear discriminant analysis, which has been successfully applied to quantitative data collected in horses with hindlimb lameness [26]—to provide evidence relevant for investigating three hypotheses:

1. The accuracy of a data-driven categorization of the efficacy of diagnostic analgesia from quantitative gait changes would reflect that increased uncertainties exist when scoring hindlimb lameness (increased intra- and inter-observer variability for scoring hindlimb than for scoring forelimb lameness [27,28]) and would hence provide higher accuracy for categorizing forelimb and lower accuracy for hindlimb diagnostic analgesia.
2. Quantitative distances between data clusters, as a measure of the data-driven ability to differentiate between categories of block efficacy, would vary across exercise conditions. For example, we expect lunge exercises with the blocked limb on the inside of the circle to provide higher distances between diagnostic analgesia categories, i.e., a better classification, due to the documented increase in movement asymmetry for that exercise condition in horses with pre-existing straight-line movement asymmetries [21,24].
3. In analogy to visual observation, quantitative features derived from head movement would be most useful for differentiating between forelimb diagnostic analgesia efficacy categories, and pelvic movement asymmetry features would be most useful for differentiating between hindlimb categories.

## 2. Materials and Methods

This retrospective study was approved by the Royal Veterinary College Social Sciences Research Ethics Board with the unique reference number (SR2020-0250).

### 2.1. Horses and Facilities

Inertial measurement unit (IMU) gait analysis data of 53 horses were evaluated retrospectively from hospital records of clinical gait analysis examinations at two veterinary referral hospitals (Royal Veterinary College (RVC) Equine Referral Hospital, United Kingdom, N = 39; Pferdeklinik Hochmoor GmbH, Germany, N = 14). Measurements had been obtained for each horse in trot, on a straight line on a hard surface and during circular exercise to the left and right on a soft surface, a hard surface, or both. The hard surface at one location was a non-slip coated tarmac and at the other a paving stone surface. The soft surface consisted of a sand/fiber mixture at one location and of a sand-based arena surface at the other. Circular exercise was performed with radii during lunging (or circular movement in-hand where deemed appropriate) ranging between 5 m and 10 m.

All horses had been referred for clinical investigation of lameness and/or poor performance by their primary care veterinarian. There were 26 Warmblood or Warmblood X, seven Thoroughbred or Thoroughbred X, five Irish Sport horses, four riding ponies, three cobs, two Quarter horses, and one each of British Sport horse, Shire X, Irish Draft X and Spanish, as well as two horses of unknown breed. Average age of the horses was 10.4 years (median 11, range 3–20). Thirty-one of the horses were geldings, 21 were mares and there was one stallion.

### 2.2. Instrumentation

Vertical displacement data were collected by attaching five to eight MTw (Xsens, Enschede, The Netherlands) IMUs (dimensions 47 × 30 × 13 mm, mass 16 g) placed at the: poll, withers, sacrum, left and right tuber coxae (5-sensors) and additionally at the thirteenth thoracic vertebra, first lumbar vertebra and above the tail root (8-sensors). In this study, only data from the 5-sensor setup were evaluated. Raw sensor data (tri-axial acceleration, tri-axial rate of turn and tri-axial magnetic field strength) were transmitted wirelessly at

update rates of 60 to 100 Hz (see [29] for influence of sample rate) per individual data channel via a proprietary protocol (Xsens) to a receiver station (Awinda, Xsens) connected to a laptop computer running operating system Windows 10 (Pro, 64 bit, Microsoft, Redmond, WA, USA) communicating with MTManager (Xsens) software providing calibrated sensor data and orientation data.

### 2.3. Movement Symmetry Parameters

Using MATLAB user interface scripts (R2015b, The Mathworks Inc., Natick, MA, USA), data were further processed by calculating movement asymmetry parameters for each trot stride based on vertical displacement calculated based on rotating calibrated tri-axial acceleration data into a horse- and gravity-based reference frame via sensor orientation and knowledge of alignment of the sensor x-axis with the horse's sagittal axis following published protocols [30,31]. Stride segmentation was defined by the sacrum sensor using published methods [32], and median values were calculated for previously described movement asymmetry parameters quantifying differences in head displacement (HD), withers displacement (WD) and pelvic displacement (PD) minima (min), maxima (max) and upward movement amplitudes (up) between stride halves: HDmin, HDmax, HDup, WDmin, WDmax, WDup, PDmin, PDmax, PDup. Additionally, the hip hike difference was calculated comparing left and right tuber coxae movement amplitudes during contralateral stance [33].

#### 2.3.1. Sign Convention

For all movement parameters, asymmetries typically encountered in left forelimb (head, withers) or left hindlimb (pelvis) lame horses were assigned negative values and asymmetries typically encountered in right forelimb (head, withers) or right hindlimb (pelvis) lame horses were assigned positive values.

#### 2.3.2. Normalization

Differences were then calculated between the condition before and after diagnostic analgesia. In order to be able to combine data from diagnostic analgesia administered to left limbs with data from diagnostic analgesia administered to right limbs without the opposite signs masking potential effects, data were normalized to reflect administering diagnostic analgesia to right limbs, by multiplying by  $-1$  (inverting) movement asymmetry differences calculated from left limb diagnostic analgesia.

### 2.4. Exercise Conditions

In addition to the type of surface (soft or hard), normalized movement asymmetry differences before/after diagnostic analgesia were labelled according to the exercise from which they had been calculated as follows: straight line, inside rein, outside rein, inside/outside average rein. Lunge direction was expressed in relation to the positioning of the limb to which the diagnostic analgesia had been applied as "inside" for horses on the left rein with left limb diagnostic analgesia (or on the right rein with right limb diagnostic analgesia) and as "outside" for horses on the left rein after right limb diagnostic analgesia or on the right rein after left limb diagnostic analgesia.

### 2.5. Block Efficacy

Perceived outcome of diagnostic analgesia "efficacy" was subjectively determined by clinician observation during the lameness examination. Where lameness examinations had been conducted over multiple days, new "baseline" asymmetry values were collected at the beginning of each day to which differences were calculated for the first block on that day, and each subsequent block was assessed against its preceding condition, i.e., the first block compared to baseline gait data, the second block compared to the first block, and so on.

Block outcomes were categorized retrospectively based on consulting electronic case records. These included statements about the efficacy of the performed blocks in three subjective categories: “negative”, “partially positive” or “positive”, or statements about the subjective ‘percentage change’ after diagnostic analgesia which were in consultation with the clinicians mapped into the three categories as follows: “negative”, 0–30% [25]; “partially positive”, >30–70% [25]; “positive”, >70–100% [25]. The chosen retrospective procedure means that the subjective categorization does not specifically differentiate the efficacy of a block for each exercise (for each surface/direction combination) but is a generalized, subjective judgement about the efficacy across exercises. This general block efficacy category was assigned ‘as a global subjective judgement’ to all exercise conditions under which the horse had been observed for the specific diagnostic block and for which quantitative gait data was available.

Straight line data of a total of 86 instances of forelimb and 88 instances of hindlimb diagnostic analgesia were included in the dataset. Blocks performed included perineural nerve blocks, as well as intra-articular and intrathecal intra-synovial techniques. A frequency distribution of the forelimb and hindlimb datasets with respect to the specific blocks administered can be found in Figure 1, with additional information about the corresponding exercise conditions in Table S1. In the context of the LDA method and the normalized features based on difference values calculated from conditions directly before and after diagnostic analgesia, one of the focal points of this first study is the identification of features that are useful across a number of potential sources of variation (clinicians, horses, source(s) of lameness etc.) rather than a detailed analysis of all potential sources of variation.

## 2.6. Linear Discriminant Analysis

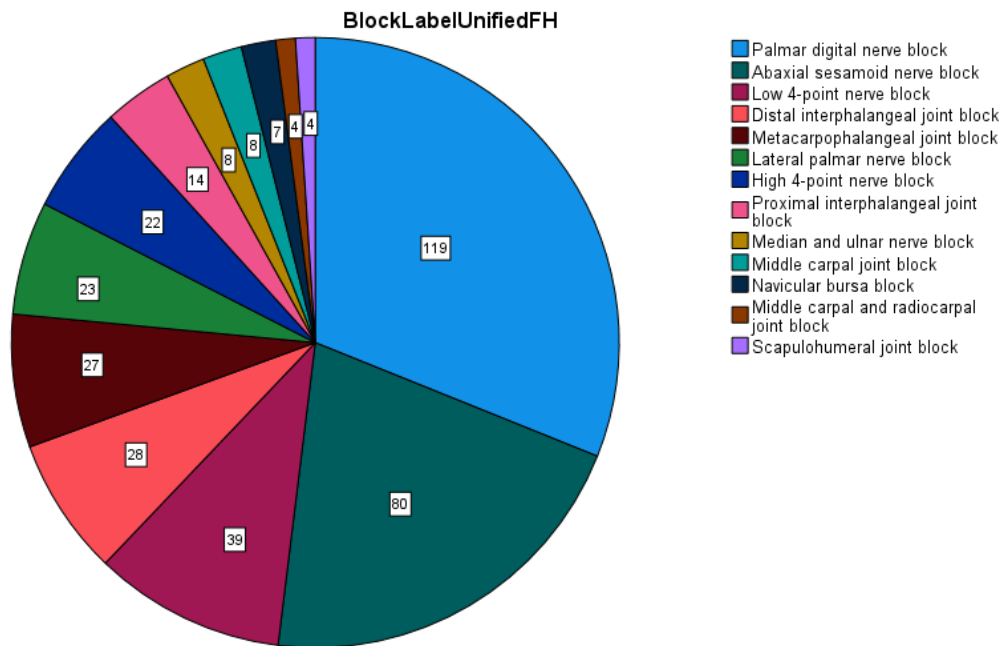
Linear discriminant analysis (LDA), a method that has been successfully employed previously to categorize data from horses with hindlimb lameness [26], was employed in the present study to investigate how well the ten quantified movement asymmetry parameters were able to differentiate between the three subjective block efficacy categories.

LDA is a method that has been investigated and used in the context of multidimensional pattern recognition applications, such as speech recognition [34,35], and is an extension to principal component analysis with particular focus on differentiation between discrete categories in the face of sources of variation. In speech recognition, the categories are typically small units of speech (phonemes), and the sources of variation are, for example, different speakers or different accents or dialects. Here, the categories are the subjective judgements of the outcome of the block (block efficacy) and the sources of variation are the different clinicians, different horses with different cause(s) of lameness, different blocks or different exercise characteristics such as, for example, lunge radii.

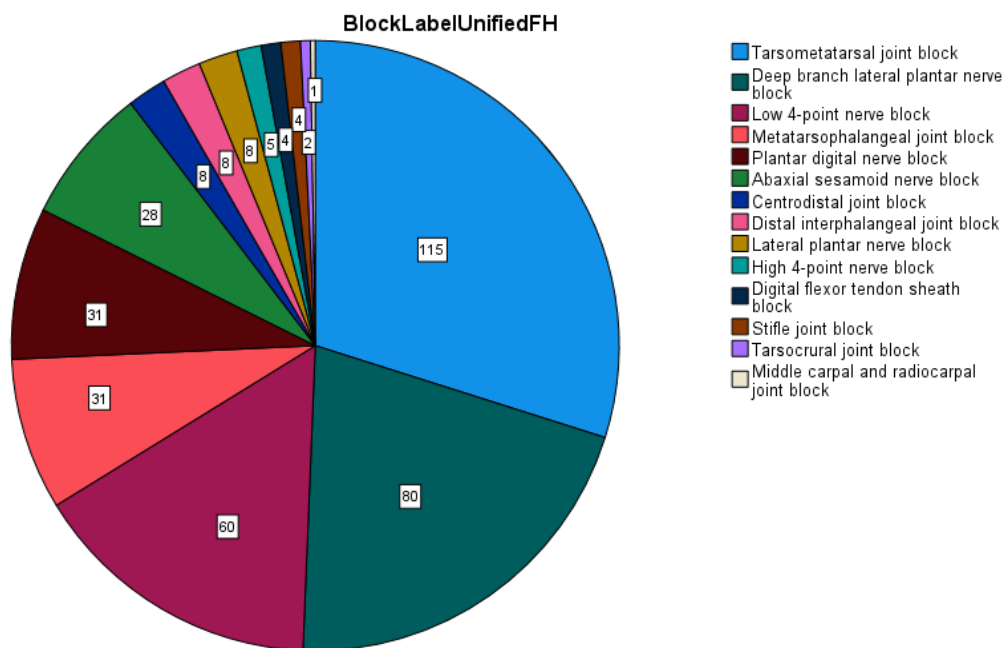
In this study, the complete dataset was first subdivided into two subsets: data for forelimb and data for hindlimb diagnostic analgesia. Furthermore, separate LDA feature transformations were implemented for straight-line data, for inside rein, outside rein and average values calculated between the two reins [36]. A further subdivision was implemented by surface type (hard/soft). This resulted in twenty LDA transformations; ten for forelimb and ten for hindlimb diagnostic analgesia: straight line (hard surface), average rein (overall, soft, hard), inside rein (overall, soft, hard) and outside rein (overall, soft, hard). SPSS (version 28.0.0.0 (190)) was used for implementing all LDA transformations.



a



b



**Figure 1.** Frequency distribution of forelimb and hindlimb datasets of movement asymmetry differences calculated before/after diagnostic analgesia (DA), subdivided by the specific diagnostic blocks. (a) forelimb DA; (b) hindlimb DA. Further details about blocks: abaxial sesamoid nerve block: palmar/plantar nerves blocked at the level of the base of the proximal sesamoid bone; Low 4-point nerve block: palmar and palmar metacarpal nerves (forelimb), plantar and plantar metatarsal nerves (hindlimb) blocked just above the metacarpophalangeal joint (forelimb) or just above the metatarsophalangeal joint (hindlimb); High 4-point nerve block: palmar and palmar metacarpal nerves (forelimb), plantar and plantar metatarsal nerves (hindlimb) blocked just below the carpometacarpal joint (forelimb) or just below the tarsometatarsal joint (hindlimb); centrodistal joint block, also referred to as distal intertarsal joint block.

### 2.6.1. Classification Accuracy

The performance of each LDA transformation was assessed by calculating reclassification accuracy and leave-one-out cross-validation accuracy (each case classified by the functions derived from all cases other than that case) from confusion matrices for the three block efficacy categories.

### 2.6.2. Variation between Exercises

In order to evaluate how specific exercise conditions might influence LDA classification, centroid distances between the three block efficacy categories were calculated in the LDA-transformed feature space along the two canonical discriminant functions. The larger the differences between block efficacy category centroids, the better the data-driven separation. Individual distances between all three two-category pairings were calculated (negative versus partially positive, negative versus positive, partially positive versus positive) as well as the sum of the distances as an overall indicator of across-category separation. Scatter plots in the LDA-transformed feature space were created, illustrating the location of each individual diagnostic analgesia data point (i.e., of each investigated diagnostic block) as well as category centroids for the three efficacy categories in the transformed feature space.

### 2.6.3. Movement Asymmetry Parameters

Standardized canonical discriminant function coefficients for the first and second function were tabulated according to the exercise condition (combination of movement direction and surface). The contribution of each LDA input feature, i.e., the weighting of each movement asymmetry parameter within each discriminant function, was quantified in terms of the absolute value. This contribution was illustrated via bar graphs (created in Microsoft Excel) for each discriminant function of each exercise condition. In order to allow reproducing the LDA transformation from the absolute values given in the results tables, originally positive values within the discriminant function are labelled “+ve”, and originally negative values are labelled “-ve”.

## 3. Results

In total, 768 feature sets containing all ten movement asymmetry parameter differences measured before and after diagnostic analgesia in 53 horses investigated for lameness/poor performance were analysed in this retrospective study. The case notes of five experienced veterinary specialists were accessed retrospectively and consulted for notes about the efficacy of individual diagnostic blocks categorized as ‘negative’, ‘partially positive’ or ‘positive’.

### 3.1. Classification Accuracy

LDA reclassification and leave-one-out classification confusion matrices are reported in Table 1. Figure 2 (forelimb diagnostic analgesia) and Figure 3 (hindlimb diagnostic analgesia) present scatter plots of the two standardized canonical discriminant functions as the x- and y-axis with individual data points (individual blocks) and diagnostic analgesia efficacy category centroids.

**Table 1.** Reclassification and leave-one-out classification accuracy after linear discriminant analysis (LDA) transformation. Presented are results for two data subsets (forelimb diagnostic analgesia DA: F; hindlimb DA: H) of movement asymmetry differences before/after administration of diagnostic analgesia in N = 53 horses. Classification results are given for straight-line, inside rein, outside rein and the average rein conditions (avg) as well as a further subdivision by surface (soft or hard), where applicable. Weighted averages are calculated taking into account the sample sizes for each condition. Highest values, the best separation between subjectively judged diagnostic analgesia efficacy categories (from retrospective consultation of case records), are indicated in red, and the lowest values, worst separation, are indicated in green.

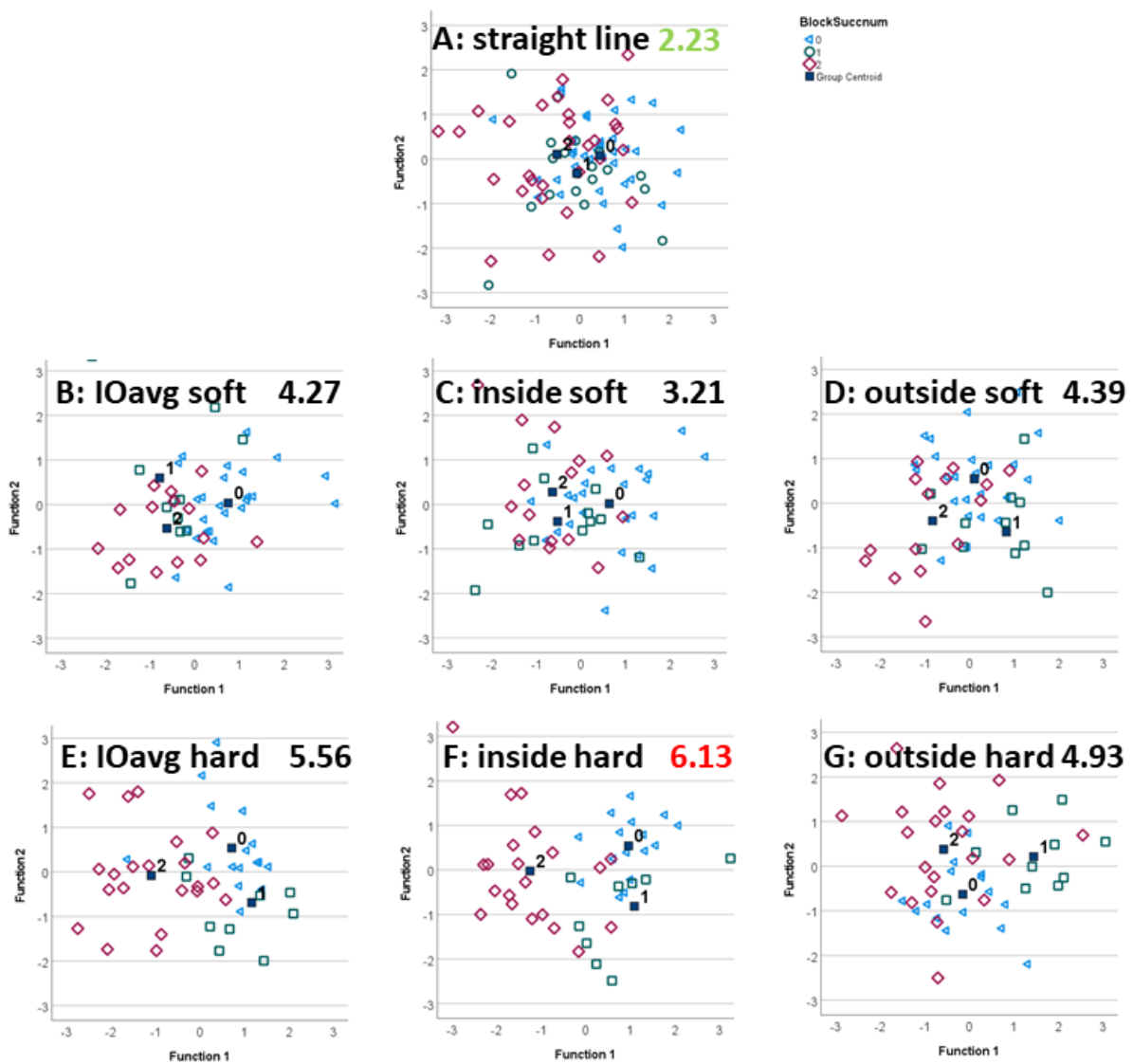
Fore/ Hind	Direction	Surface	Reclass. (%)	Leave-1-Out (%)	Abs Diff. (%)	N
F	str	Hard	53.5	39.5	14	86
F	Avg	All	63.6	52.5	11.1	99
		Soft	63.5	46.2	17.3	52
		Hard	80.9	51.1	29.8	47
F	Inside	All	63.6	50.5	13.1	99
		Soft	57.7	40.4	17.3	52
		Hard	68.1	57.4	10.7	47
F	Outside	All	61.6	52.5	9.1	99
		Soft	63.5	55.8	7.7	52
		Hard	72.3	38.3	34	47
F	<i>weighted total</i>		60.82	49.06		383
H	Str	Hard	60.2	51.1	9.1	88
H	Avg	All	60.6	49.5	11.1	99
		Soft	70.7	46.3	24.4	41
		hard	60.3	48.3	12	58
H	Inside	all	61.6	44.4	17.2	99
		soft	78	56.1	21.9	41
		hard	62.1	44.8	17.3	58
H	Outside	all	51.5	38.4	13.1	99
		soft	68.3	34.1	34.2	41
		hard	58.6	39.7	18.9	58
H	<i>weighted total</i>		58.43	45.70		385

Table 1 provides an overview of the reclassification and leave-one-out classification accuracies for LDA classification based on different subsets of the data subdivided by diagnostic analgesia location (front or hindlimb), exercise (straight-line, inside rein, outside rein, average) and surface (soft or hard).

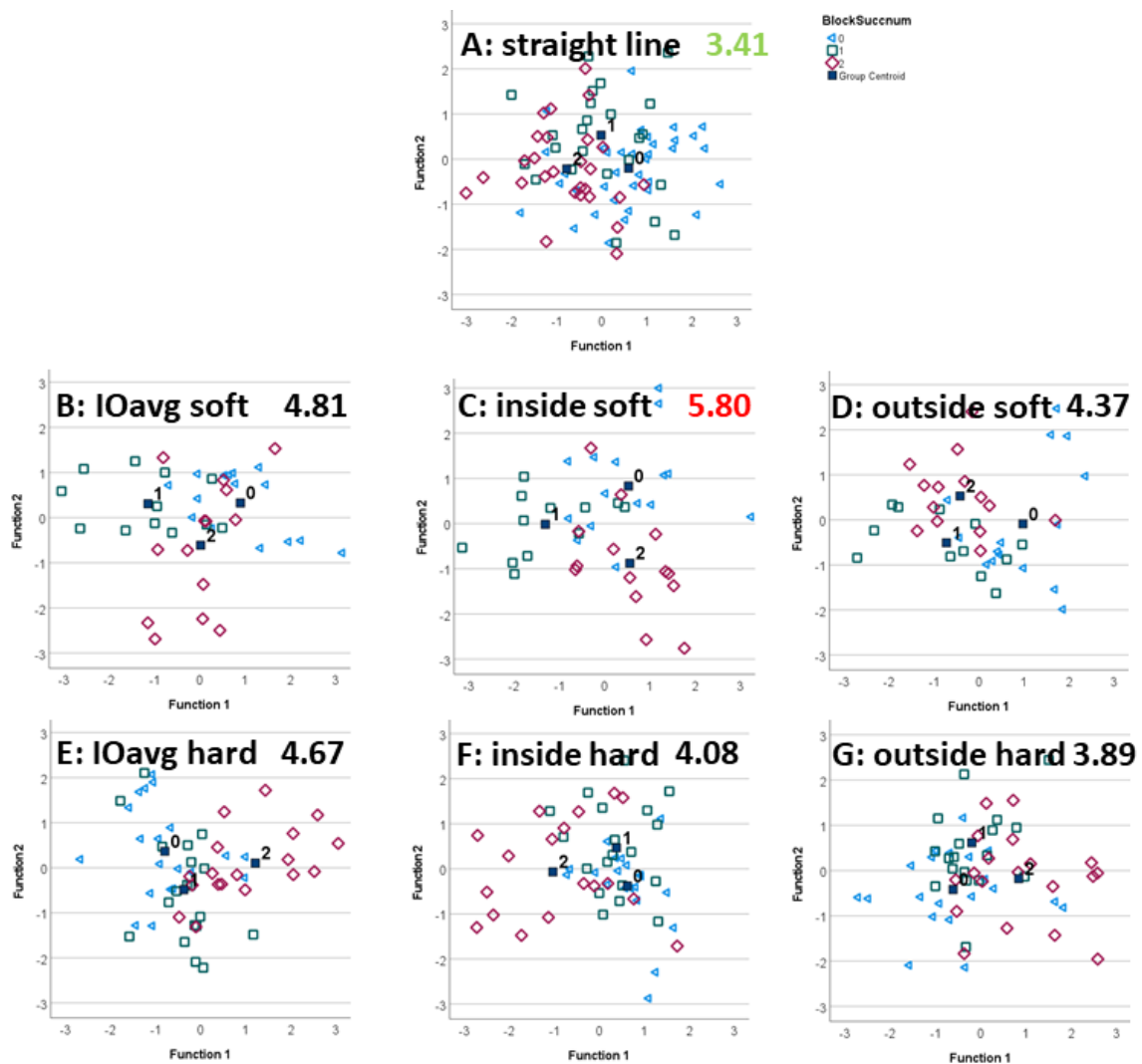
For forelimb diagnostic analgesia, the highest leave-one-out LDA-classification accuracy (57.4%) was achieved with the data from the lunge and the horse trotting with the blocked limb on the inside of the circle on hard ground, followed by exercise with the blocked limb on the outside of the circle on soft ground (55.8%). The lowest leave-one-out classification with 38.3% was found for the outside rein data on the hard surface followed by the straight-line data with 39.5%.

For hindlimb diagnostic analgesia, the highest leave-one-out LDA-classification accuracy (56.1%) was found with the blocked limb on the inside of the circle on soft ground, followed by straight-line trot (51.1%). The lowest leave-one-out classification accuracy was with 34.1%, found with the blocked limb on the outside of the circle when trotting on a soft surface.





**Figure 2.** Scatter plots of the first (Function1, x-axis) and second (Function2, y-axis) standardized canonical discriminant function calculated from linear discriminant analysis (LDA) applied to movement symmetry differences before/after forelimb diagnostic analgesia (negative (0, blue triangle), partially positive (1, green square) and positive (2, red diamond)). Black filled squares indicate group centroids of the three category clusters. The highest value, the best separation between subjectively judged diagnostic analgesia efficacy categories is indicated in red, and the lowest values, worst separation, is indicated in green.



**Figure 3.** Scatter plots of the first and second standardized canonical discriminant function calculated from LDA applied to movement symmetry differences before/after hindlimb diagnostic analgesia (negative (0, blue triangles), partially positive (1, green square) and positive (2, red diamond)). Black filled squares indicate group centroids of the three category clusters. The highest value, the best separation between subjectively judged diagnostic analgesia efficacy categories is indicated in red, and the lowest values, worst separation, is indicated in green.

### 3.2. Exercise Condition

Figures 2 and 3 illustrate the ability of the LDA-transformed feature space to differentiate between the efficacy categories. Distances between the category centroids (Table 2) provide quantitative insights into the overall distance between the three clusters as well as the distances between pairs of categories.

In agreement with the highest leave-one-out classification accuracy for forelimb diagnostic analgesia, the highest overall distance between efficacy categories (6.13) was achieved for trot on a hard surface with the blocked limb on the inside of the circle. Figure 2F shows generally good separation between the three category clusters, with little overlap. The lowest overall distance (2.23) was found for straight-line trot, almost a factor of three less than for the condition with the highest inter-category distance. Further analysis, shown in Table 2, indicates that trotting with the blocked limb on the inside of the circle on a hard surface provides the highest pairwise inter-category distances for differentiating between partially positive and positive blocks as well as between negative and positive blocks. The

highest pairwise inter-category distance between negative and partially positive forelimb DA was found with the blocked limb on the outside of the circle on a hard surface.

In agreement with the highest leave-one-out classification accuracy for hindlimb diagnostic analgesia, the highest overall distance between hindlimb block efficacy categories (5.80) was achieved for trot on a soft surface with the blocked limb on the inside of the circle. Figure 3C shows good separation between the three category clusters. In analogy to the forelimb data, the lowest overall distance (3.41) was found for the straight-line trot. Further analysis, see Table 2, indicates that trotting with the blocked limb on the inside of the circle on a soft surface provides the highest pairwise inter-category distance for differentiating between partially positive and positive blocks. The highest pairwise inter-category distance between negative and partially positive forelimb blocks was found for the average lunge assessment on the soft surface and for the average lunge assessment on the hard surface between negative and positive forelimb blocks.

**Table 2.** Distances between category centroids of ‘negative’, ‘partially positive’ and ‘positive’ forelimb diagnostic analgesia (DA) based on canonical discriminant functions calculated through linear discriminant analysis for seven different exercise conditions for features quantifying differences in movement asymmetry before/after forelimb and hindlimb diagnostic analgesia: neg2part, distance between centroids of negative and partially positive blocks; part2pos, distance between centroids of partially positive and positive blocks; neg2pos, distance between centroids of negative and positive blocks; Highest Individ., highest distance between block categories with respect to which exercise condition it was recorded for. Highest values, the best separation between subjectively judged diagnostic analgesia efficacy categories (from retrospective consultation of case records), are indicated in red, and the lowest values, worst separation, are indicated in green; separately for forelimb and hind limb diagnostic analgesia.

Blocked Limb	Direction	Surface	neg2part	part2pos	neg2pos	Total	Highest Individ.
F	str	hard	0.65	0.62	0.96	2.23	
	avg	Soft	1.64	1.14	1.49	4.27	
		Hard	1.30	2.34	1.91	5.56	
	inside	Soft	1.23	0.67	1.31	3.21	part2pos, neg2pos
		Hard	1.36	2.48	2.29	6.13	
	outside	Soft	1.39	1.68	1.33	4.39	neg2part
Hard		1.80	2.03	1.10	4.93		
H	str	Hard	0.96	1.07	1.38	3.41	
	avg	Soft	2.05	1.48	1.28	4.81	neg2part
		Hard	0.95	1.69	2.03	4.67	neg2pos
	inside	Soft	2.03	2.06	1.71	5.80	part2pos
		Hard	0.88	1.51	1.69	4.08	
	outside	Soft	1.75	1.08	1.53	4.37	
Hard		1.11	1.30	1.47	3.89		

### 3.3. Movement Asymmetry Parameters

#### 3.3.1. Forelimb Diagnostic Analgesia

Generally, there is a high level of variation in the distribution of the absolute values of the first canonical discriminant function across the ten features across the seven exercise conditions calculated for forelimb diagnostic analgesia (Figure 4).

absolute values	canon1 straight	canon1 ioavg soft	canon1 ins. soft	canon1 outs. soft	canon1 ioavg hard	canon1 ins. hard	canon1 outs. hard	canon1 average	rank average
PDmin	+ve 0.447	-ve 1.479	+ve 2.694	-ve 1.234	-ve 0.139	+ve 1.105	-ve 0.112	1.030	7th
PDmax	+ve 1.339	-ve 0.806	+ve 1.399	-ve 0.967	+ve 1.837	+ve 0.277	+ve 0.585	1.030	7th
PDup	-ve 1.125	+ve 2.145	-ve 2.338	+ve 0.87	-ve 1.376	-ve 0.824	-ve 0.444	1.303	4th
HHD	-ve 0.994	-ve 0.459	-ve 0.165	+ve 0.296	+ve 0.173	+ve 1.34	+ve 0.936	0.623	10th
WDmin	+ve 0.501	+ve 0.914	+ve 1.572	+ve 1.642	+ve 0.825	+ve 0.099	+ve 3.899	1.350	3rd
WDmax	+ve 0.332	+ve 1.022	+ve 1.372	+ve 0.957	+ve 1.13	+ve 0.561	+ve 2.103	1.068	5th
WDup	-ve 0.573	-ve 0.631	-ve 0.644	-ve 1.243	-ve 0.933	-ve 0.348	-ve 3.037	1.058	6th
HDmin	-ve 0.663	-ve 0.371	+ve 1.116	-ve 1.159	+ve 2.259	-ve 2.377	-ve 3.309	1.608	2nd
HDmax	+ve 0.059	-ve 0.231	+ve 0.454	-ve 1.435	+ve 1.113	-ve 1.277	-ve 1.164	0.819	9th
HDup	+ve 1.31	+ve 1.211	-ve 1.912	+ve 2.668	-ve 1.715	+ve 1.864	+ve 3.382	2.009	1st

absolute values	canon2 straight	canon2 ioavg soft	canon2 ins. soft	canon2 outs. soft	canon2 ioavg hard	canon2 ins. hard	canon2 outs. hard	canon2 average	rank average
PDmin	-ve 0.064	+ve 1.295	+ve 0.011	-ve 0.877	+ve 3.551	-ve 3.726	-ve 1.501	1.575	5th
PDmax	+ve 0.929	+ve 0.792	+ve 0.532	+ve 0.584	+ve 1.859	-ve 1.731	-ve 0.682	1.016	8th
PDup	+ve 1.112	-ve 0.67	-ve 1.239	-ve 2	-ve 5.222	+ve 4.677	+ve 2.49	2.487	1st
HHD	-ve 1.184	+ve 0.348	+ve 1.176	+ve 0.883	+ve 0.989	-ve 1.375	-ve 1.072	1.004	9th
WDmin	+ve 0.779	-ve 1.481	-ve 4.588	+ve 0.401	+ve 1.824	-ve 2.549	+ve 2.132	1.965	2nd
WDmax	+ve 0.562	-ve 0.456	-ve 3.109	+ve 0.057	+ve 0.354	+ve 2.285	+ve 1.754	1.225	6th
WDup	-ve 0.606	+ve 1.78	+ve 2.799	-ve 1.063	-ve 1.851	+ve 2.576	-ve 2.561	1.891	4th
HDmin	+ve 0.765	-ve 1.31	+ve 0.129	+ve 0.108	+ve 3.923	-ve 1.597	-ve 0.164	1.142	7th
HDmax	+ve 0.873	-ve 1.024	+ve 0.326	+ve 0.013	+ve 1.933	-ve 0.915	+ve 0.125	0.744	10th
HDup	-ve 1.61	+ve 2.05	-ve 0.347	-ve 0.46	-ve 5.47	+ve 3.071	-ve 0.447	1.922	3rd

**Figure 4.** Absolute values of standardized canonical discriminant function coefficients for LDA applied to differences in ten upper body movement asymmetry parameters before/after forelimb diagnostic analgesia. Absolute values are used to create bar graphs illustrating the weighting of all features within the first (canon1) and within the second (canon2) canonical discriminate function. Labels “+ve” and “-ve” indicate the sign of the original values before taking absolute values for illustrative purposes. Ranking of movement symmetry features within the average absolute values (column average) from highest (1st) to lowest (10th) is given in the last column (ranking). The highest total value (one fore forelimb blocks, one for hindlimb blocks), the best separation between subjectively judged diagnostic analgesia efficacy categories is indicated in red, and the lowest values, worst separation, is indicated in green.

Along the first canonical function, two head asymmetry parameters, HDmin and HDup, overall show the highest weighting, followed by a withers (WDmin) and then a pelvic (PDup) feature. For straight-line assessment, the first, third and fourth highest weightings are attributed to pelvic movement asymmetry parameters (PDmax, PDup, HHD), with the second highest weighting attributed to a head movement asymmetry parameter (HDup). For lunge assessments, head movement asymmetry parameters appear to receive higher weightings on the hard surface, and pelvic movement asymmetry parameters on the soft surface.

Along the second canonical function, HDup is the only head movement feature in the top 6 in the average weighting across exercises. HDup also gets the highest weighting on the straight, where again three of the four pelvic parameters are in the top four (HHD, PDup, PDmax).

### 3.3.2. Hindlimb Diagnostic Analgesia

Generally, there is a high level of variation in the distribution of the absolute values of the first canonical discriminant function across the ten movement asymmetry parameters across the seven exercise conditions calculated for hindlimb diagnostic analgesia (Figure 5).

absolute values	canon1 straight	canon1 ioavg soft	canon1 ins. soft	canon1 outs. soft	canon1 ioavg hard	canon1 ins. hard	canon1 outs. hard	canon1 average	rank average
PDmin	+ve 1.513	+ve 0.117	+ve 1.375	-ve 0.008	-ve 0.355	+ve 0.301	+ve 0.208	0.554	9th
PDmax	+ve 1.822	+ve 0.461	+ve 0.868	-ve 0.104	+ve 0.535	+ve 0.244	-ve 0.189	0.603	8th
PDup	-ve 2.31	+ve 0.763	-ve 3.469	+ve 0.612	+ve 0.347	+ve 0.951	-ve 0.076	1.218	2nd
HHD	+ve 0.691	-ve 0.46	-ve 1.782	-ve 0.068	-ve 1.155	-ve 0.144	-ve 0.975	0.754	5th
WDmin	-ve 0.142	+ve 0.854	+ve 3.79	+ve 0.225	-ve 0.623	-ve 0.984	+ve 1.6	1.174	3rd
WDmax	-ve 0.226	-ve 0.001	+ve 1.957	+ve 0.028	-ve 0.417	+ve 0.545	+ve 1.3	0.639	6th
WDup	+ve 0.982	-ve 1.014	-ve 2.976	-ve 0.223	+ve 0.521	-ve 1.325	-ve 1.735	1.254	1st
HDmin	+ve 0.069	+ve 0.064	+ve 0.941	-ve 0.108	-ve 0.584	-ve 0.723	+ve 0.405	0.413	10th
HDmax	+ve 0.072	+ve 0.485	+ve 1.221	-ve 0.493	-ve 0.84	+ve 1.109	+ve 0.067	0.612	7th
HDup	-ve 0.003	+ve 0.128	-ve 2.004	+ve 0.375	+ve 1.026	+ve 1.538	-ve 0.411	0.784	4th

absolute values	canon2 straight	canon2 ioavg soft	canon2 ins. soft	canon2 outs. soft	canon2 ioavg hard	canon2 ins. hard	canon2 outs. hard	canon2 average	rank average
PDmin	-ve 0.249	+ve 2.563	-ve 3.752	+ve 0.075	+ve 1.219	-ve 1.48	+ve 1.383	1.532	2nd
PDmax	+ve 1.379	+ve 1.834	-ve 2.841	+ve 0.408	+ve 0.168	+ve 0.296	+ve 0.969	1.128	4th
PDup	-ve 1.058	-ve 2.171	+ve 3.635	-ve 0.363	-ve 3.394	-ve 1.723	-ve 5.073	2.488	1st
HHD	+ve 0.312	-ve 0.281	-ve 0.217	+ve 0.581	+ve 2.123	+ve 1.435	+ve 2.902	1.122	5th
WDmin	+ve 0.005	-ve 0.618	+ve 0.52	+ve 0.113	-ve 1.039	-ve 0.333	+ve 1.731	0.623	8th
WDmax	+ve 1.291	-ve 0.578	+ve 0.031	-ve 0.109	+ve 0.229	-ve 0.085	+ve 2.862	0.741	7th
WDup	-ve 0.69	+ve 1.399	-ve 0.351	-ve 0.018	+ve 1.225	-ve 1.823	-ve 3.099	1.229	3rd
HDmin	+ve 0.664	-ve 0.14	+ve 0.041	-ve 1.941	+ve 0.417	+ve 0.281	+ve 0.226	0.530	9th
HDmax	+ve 0.239	-ve 0.011	+ve 0.177	+ve 1.024	-ve 0.935	+ve 0.413	-ve 0.136	0.419	10th
HDup	-ve 1.342	-ve 0.024	-ve 0.666	-ve 2.638	+ve 0.474	-ve 0.301	-ve 0.961	0.915	6th

**Figure 5.** Absolute values of standardized canonical discriminant function coefficients for linear discriminant analysis (LDA) applied to differences in ten upper body movement asymmetry parameters before/after hindlimb diagnostic analgesia. These values are used to create bar graphs (presented in the figure above) illustrating the weighting of asymmetry values within the first (canon1) and the second (canon2) canonical discriminant function. Labels “+ve” and “-ve” are indicating the sign of the values. Ranking of movement symmetry features within the average absolute values (column average) from highest (1st) to lowest (10th) is given in the last column (ranking).

Along the first canonical function, averaged across all exercises, two withers (WDmin and WDup) and one pelvic movement asymmetry parameter (PDup) are assigned the highest weighting. For differentiation between block efficacy categories based on straight-line data, three pelvic movement asymmetry parameters (PDup, PDmax, PDmin) followed by a withers movement asymmetry parameter (WDup) are contributing, with the highest weighting. Except for the hard surface circle assessments, PDup is consistently within the top three weightings.

Along the second canonical function, averaged across all exercises, the four pelvic movement asymmetry parameters and one withers movement asymmetry parameter (WDup) are assigned the highest weightings. With the exception of straight-line trot and lunge exercise on the soft surface with the blocked limb on the outside (where HDmin and HDup are in the top five weightings), head movement asymmetry parameters are assigned low weightings.

#### 4. Discussion

In this study, we have investigated the association between the subjectively judged efficacy of diagnostic analgesia and quantitative changes in movement asymmetry.

Linear discriminant analysis (LDA) provides the ability to reduce the dimensionality of a data-driven classification by transforming the original feature space into a set of transformed features that maximize the distance between a set of defined classes. It also provides a framework for expressing the classification performance, for calculating



distances between classes in a comparable manner between datasets and is a general method used in pattern recognition for differentiating between multiple categories in the presence of sources of variation [34,35]. It has also been applied previously to scoring hindlimb lameness from IMU data [26]. This has allowed us to compare different exercise conditions and investigate the importance of the original features—derived from head, withers and pelvic movement—in the transformation and classification process. In this first study, we were not interested in further analyzing the detailed influence of potential sources of variation (individual clinician, horse, source of lameness or specific diagnostic block) but rather in investigating the discriminative power of LDA for common exercises and how the exercises highlight which movement asymmetry features are essential in this data-driven process. With increasing amounts of quantitative data collected from many more clinical cases, additional sources of variation will become more important for decision-making about individual patients, and will then need detailed analysis.

#### 4.1. Classification Accuracy

For forelimb diagnostic analgesia, both the highest and lowest accuracies for the leave-one-out classification tasks were found for lunge assessment on the hard surface, with 57.4% on the inside rein and 38.3% on the outside rein. In analogy, for hindlimb diagnostic analgesia, both the highest and lowest accuracies for the leave-one-out classification tasks were found for lunge assessment, however, on the soft surface, with 56.1% on the inside rein and 34.1% on the outside rein. The low accuracy, in particular when exercising with the blocked limb on the outside of the circle, may be related to the effect that circle-induced and lameness-induced movement asymmetries cancel each other out [24,37] and that, as such, asymmetries may be small and likely more difficult to perceive [18]. However, this aspect needs to be investigated further because, in the present study, diagnostic analgesia efficacy was not assigned individually to each exercise condition, but rather an overall efficacy value was assigned to all exercises assessed for the evaluated block and subjective ‘labels’ used to reflect the expert opinion of the veterinarians working up the lameness case. Additionally, the variations in accuracy might be related to specific lameness types appearing to be exacerbated depending on the surface or the positioning of the affected limb on the inside or the outside of the circle. For example, lameness in horses with suspensory desmitis is exacerbated on deep, soft surfaces and is often worse with the affected limb on the outside of the circle, and this injury is seen much more commonly in hindlimbs compared to forelimbs [3].

Generally, the present study provided no supporting evidence for the hypothesis that LDA-based classification accuracy for differentiating between different subjective efficacy categories would be higher for forelimb than for hindlimb diagnostic analgesia. The ranges of reclassification and leave-one-out classification accuracies across exercises given in Table 1 are similar between forelimb and hindlimb data. This is surprising in as much as that higher levels of observer agreement have been reported for forelimb lameness scoring compared to the scoring of hindlimb lameness [28]. In the context of a data-driven classification, the high variability in asymmetry measures found for head movement among horses and between reins [19,38] contributes to higher levels of ‘noise’ in the data likely impeding automated classification. It should also be noted that the comparatively high differences between reclassification and leave-one-out classification accuracy reported in the present study (Table 1) indicate that some of the LDA transformations may not generalize well across the data subsets. In particular, for the individual-rein data with smaller sample sizes ( $N = 41$  to  $N = 58$ ), differences of up to 34% are found between reclassification and leave-one-out classification. In contrast, for straight-line and combined-surface data, smaller differences of 9–17% are reported with larger sample sizes between  $N = 86$  and  $N = 99$ .



#### 4.2. Exercise Condition

Scatter plots and distances between block efficacy category centroids confirmed that lunge data are contributing in a useful manner to the data-driven differentiation between subjective block efficacies (Figures 2 and 3).

##### 4.2.1. Forelimb Diagnostic Analgesia

Exercising with the blocked limb on the inside of the circle on a hard surface appears particularly advantageous and provides the highest inter-centroid difference (Figure 2, Table 2). This exercise condition also shows the highest values for two of the distances between pairs of centroids (between partially positive and positive as well as between negative and positive forelimb blocks; Table 2). That circling on a hard surface is beneficial should not come as a surprise, based on clinical experience [3] but also because head movement asymmetry increases on a hard surface compared to straight-line exercise [21]. However, there is a considerable amount of variation between horses and reins in relation to changes in head movement on the lunge [23,24]. Investigating potential confounding variables such as the specific block administered within a limb may aid in better understanding the source of these variations. Specific LDA transformations for specific types of diagnostic analgesia may improve classification performance in the future.

##### 4.2.2. Hindlimb Diagnostic Analgesia

The highest inter-centroid difference was reported for lunge exercise with the blocked limb on the inside of the circle (Table 2, Figure 3). However, in contrast to forelimb diagnostic analgesia, distances between subjective efficacy centroids were highest for soft surface exercise with the blocked limb on the inside of the circle. This exercise condition also provides the highest distance between partially positive and positive hindlimb blocks (Table 2).

##### 4.2.3. Average across Reins

For both forelimb and hindlimb data, inter-centroid distances were smaller for straight-line exercise than for any of the lunge exercises. In agreement with clinical practice [3], this highlights the usefulness of exercising horses in circles from a data-driven decision-making perspective.

For hindlimb data, the highest inter-centroid distances between negative and partially positive as well as between negative and positive blocks were found for average-rein movement asymmetry data. This is particularly interesting in light of a previous study (without diagnostic analgesia) which reported smaller average-rein asymmetry values across reins compared to straight-line asymmetry values [36]. Despite that, the higher inter-centroid differences between efficacy categories for average-rein data indicate that differences in movement asymmetry before/after hindlimb diagnostic analgesia may increase data consistency and remove 'noise' from LDA-based decision making. It should also be noted that the circle radius was not tightly controlled in this multi-center study and that differences in circle radius and speed have been shown to influence movement symmetry [20,39]. This should be kept in mind when combining data from left and right reins.

For forelimb data, on the other hand, average-rein data appeared less beneficial, with the highest inter-centroid distances all reported for individual-rein data. This may be related to the higher variation between horses and reins for head movement [23,24] and indicates, in contrast to hindlimb data, that average-rein data cannot effectively remove this 'noise'.

#### 4.3. Asymmetry Parameters

When studying the ranking of the weightings attributed to each of the ten asymmetry parameters across exercise conditions and across differentiating between efficacy categories for both forelimb and hindlimb diagnostic analgesia (Figures 4 and 5), the top three parameters are HDup (1st and 4th), PDup (4th and 2nd) and WDmin (3rd and 3rd). This

highlights the fact that movements across the whole body from the head, via the withers, to the sacrum need to be considered when judging the effect of diagnostic analgesia and reinforces recent studies reporting that withers movement is important for quantifying lameness [40–42].

In agreement with our third hypothesis, two (of three) head movement parameters receive the first and second highest weighting across forelimb LDAs, emphasizing the importance of the head nod in the context of forelimb lameness [43]. One pelvic movement parameter (PDup) is among the top three highest weightings for hindlimb LDAs, providing some supporting evidence for our third hypothesis in relation to hindlimb diagnostic analgesia. However, withers movement asymmetry (WDup and WMin) appears at least equally important for data-driven categorization of the success of hindlimb DA, emphasizing the cross-body complexity of movement patterns [40,41] beyond the documented hip or pelvic hike [44].

Movement asymmetry changes after diagnostic analgesia confirm the veterinary “law-of-sides”, indicating that horses with a true hindlimb lameness show head movement mimicking an ipsilateral forelimb lameness and that horses with a true forelimb lameness show pelvic movement mimicking a contralateral hindlimb lameness [45,46]. Some supporting evidence for the “law-of-sides” can be found in the present study: for example, the weightings for PDup and HDup for the first canonical function of all seven hindlimb LDAs (Figure 5) agree, indicating ipsilateral changes in these asymmetry parameters, as expected for changes after hindlimb DA [45]. However, in contrast to this finding, movement asymmetry changes can be more complex on the circle than what the “law-of-sides” suggests [37].

The compensatory effect is also less clear for forelimb DA (Figure 4): Three exercises (straight line, outside soft, outside hard) show evidence of the expected contralateral effect, i.e., a positive weighting for HDup paired with a negative weighting for PDup in agreement with the “law-of-sides”. The other four exercise conditions are showing evidence of an ipsilateral effect between changes in HDup and PDup. There is previous conflicting evidence that compensatory movements in forelimb lameness are inconsistent between horses [47], that there are both ipsilateral and contralateral compensatory components during circular movement [37] and that specifically the pushoff phase of the contralateral hindlimb is affected [46].

Withers movement has been shown to provide relevant information for differentiating between a true forelimb lameness and a compensatory head nod in reaction to a hindlimb lameness [40,41]. The fact that one withers parameter (WMin) receives the third highest weighting within the first canonical discriminant function for both forelimb and hindlimb LDAs (averaged across exercises) highlights the general importance of withers movement asymmetry for data-driven categorization of changes of diagnostic analgesia. However, the relationships between the signs of corresponding head and withers parameters, for example between HDmin and WMin or between HDup and WDup (Figures 4 and 5), do not provide evidence for consistent differentiation between forelimb and hindlimb diagnostic analgesia.

Similar to studies conducted for straight-line movement [15,16,48,49], it appears beneficial to quantify the contralateral force distribution between pairs of limbs in lame horses on the lunge. This might help uncover fundamental mechanisms governing compensatory movement patterns across exercises.

#### 4.4. Limitations

In addition to the documented expectation bias [17] when subjectively judging the effect of diagnostic analgesia, the present study was conducted at two veterinary hospitals routinely employing quantitative gait analysis during lameness examinations. As a consequence, the clinicians judging the efficacy of diagnostic analgesia have had access to quantitative information in the form of asymmetry parameters during the workup. This will likely have influenced their subjective judgement; a subjective assessment blinded

to the outcomes of gait analysis would have provided two truly independent sources of information. There are some considerations that appear particularly relevant in this context.

First, the gait analysis system used in this study provides data with a delay of 30 s or more after the horse has finished a straight-line exercise and of around 60 s for longer exercises such as lunging. Additional time is then needed to create left-right rein average data or comparison plots. This general delay until the gait analysis results are available provides the clinicians time to consider their subjective opinion about the efficacy of the block. This may reduce, but does not completely eliminate, the influence of gait analysis results in the decision-making process about block efficacy.

Second, the current guideline values of 6 mm for head movement and 3 mm for pelvic movement asymmetry [14] (adjusted to 8 mm and 5 mm, respectively, for the system used in this study [50]), are only directly applicable to straight-line data because inter-measurement variability has been documented to be higher on the lunge [19]. The lack of relevant guideline values for anything but straight-line data limits (but again does not eliminate) the influence of quantitative gait data for the subjective judgement.

Third, contrary to clinical practice whereby individual exercise conditions are assigned different lameness grades [3], our approach here was to assign one 'global' efficacy category across all the exercises based on retrospective consultation of the clinical case records where terms such as 'negative response', 'partially positive response' or 'positive response' were mentioned for specific blocks or subjective impressions about the amount of improvement after a block were given in percentage numbers which were mapped to the categories as described. The process of using a 'global' efficacy score was implemented to limit the 'noise' (uncertainty) in the subjective data, i.e., the documented variabilities in inter-observer agreement for lameness grading [28,51,52], and specific difficulties when visually assessing lameness on the lunge [37]. There are currently no quantitative studies on the effect of inter-observer variation in evaluating the efficacy of diagnostic analgesia for lunge exercise.

It appears crucial that future studies with larger datasets investigate inter-observer variation as an important source of variation and consider the subjective judgement of each exercise condition in relation to, for example, specific origins of lameness, the number and order of blocks applied and the number of limbs involved in a lameness. This appears crucial in the context of documented challenges encountered in human decision making based on multiple parameters [53]. Put simply, the aim of the current study was to provide some limited, first insights into the fundamental 'mechanisms' by employing a robust, basic data-driven method to 'real-life' data including a number of sources of variation. The reader is firmly encouraged to take into account the limitations of the chosen approach when drawing conclusions about the individual exercise conditions.

## 5. Conclusions

Accuracy of data-driven classification of changes in gait asymmetry after diagnostic analgesia varied considerably for both forelimb and hindlimb data, likely reflecting the variable responses of lame horses across exercise conditions.

The best separation between categories was measured with the blocked limb on the inside of the circle, highlighting the usefulness of lunge exercise and in particular consulting quantitative gait data of this exercise condition before/after diagnostic analgesia, on hard ground for horses with forelimb lameness and on soft ground for horses with hindlimb lameness.

Movement asymmetry parameters from across the body—derived from vertical movement of head, withers and pelvis—received high weightings within the first canonical LDA functions across exercises. This highlights that movement changes after diagnostic analgesia affect the whole upper body and that compensatory movement changes should be considered when quantitative gait data for changes before/after diagnostic analgesia.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ani12060762/s1>, Table S1: Difference in normalized vertical head,

withers and pelvic movement symmetry before/after diagnostic analgesia in 53 horses undergoing veterinary lameness examinations during straight-line trot and trot on the lunge on hard and/or soft ground.

**Author Contributions:** Conceptualization, T.P., D.M.B., A.F.-J., C.G., M.P., R.K.W.S.; methodology, T.P.; software, T.P.; formal analysis, K.H., L.L., T.P.; data curation, D.M.B., A.F.-J., C.G., M.P., R.K.W.S.; writing—original draft preparation, K.H., L.L., T.P.; writing—review and editing, all authors; supervision, T.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:** This retrospective study was approved by the Royal Veterinary College Social Sciences Research Ethics Board with the unique reference number SR2020-0250.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are available via figshare at <https://doi.org/10.6084/m9.figshare.19105016.v2> (accessed on 4 February 2022).

**Conflicts of Interest:** T.P. is owner of EquiGait Ltd., a company offering gait analysis products and services. This does not alter our adherence to policies on sharing data and materials.

## References

1. Bassage, L., II; Ross, M.W. Diagnostic Analgesia. In *Diagnosis and Management of Lameness in the Horse*; Ross, M.W., Dyson, S.J., Eds.; Elsevier Health Sciences: London, UK, 2011; pp. 100–134.
2. Ross, M.W.; Durham, M. Applied Anatomy of the Musculoskeletal System. In *Diagnosis and Management of Lameness in the Horse*; Ross, M.W., Dyson, S.J., Eds.; Elsevier Health Sciences: London, UK, 2011; pp. 88–99.
3. Ross, M.W. Movement. In *Diagnosis and Management of Lameness in the Horse*; Ross, M.W., Dyson, S.J., Eds.; Elsevier Health Sciences: London, UK, 2011; pp. 64–79.
4. Ross, M.W. Coexistent Lameness. In *Diagnosis and Management of Lameness in the Horse*; Ross, M.W., Dyson, S.J., Eds.; Elsevier Health Sciences: London, UK, 2011; ISBN 978-1-4160-6069-7.
5. Hobbs, S.J.; Licka, T.; Polman, R. The Difference in Kinematics of Horses Walking, Trotting and Cantering on a Flat and Banked 10 m Circle. *Equine Vet. J.* **2011**, *43*, 686–694. [[CrossRef](#)] [[PubMed](#)]
6. Chateau, H.; Camus, M.; Holden-Douilly, L.; Falala, S.; Ravary, B.; Vergari, C.; Lepley, J.; Denoix, J.-M.; Pourcelot, P.; Crevier-Denoix, N. Kinetics of the Forelimb in Horses Circling on Different Ground Surfaces at the Trot. *Vet. J.* **2013**, *198*, e20–e26. [[CrossRef](#)] [[PubMed](#)]
7. Greve, L.; Pfau, T.; Dyson, S. Alterations in Body Lean Angle in Lameness before and after Diagnostic Analgesia in Straight Lines in Hand and on the Lunge. *Vet. J.* **2018**, *239*, 1–6. [[CrossRef](#)] [[PubMed](#)]
8. Barstow, A.; Bailey, J.; Campbell, J.; Harris, C.; Weller, R.; Pfau, T. Does ‘Hacking’ Surface Type Affect Equine Forelimb Foot Placement, Movement Symmetry or Hoof Impact Deceleration during Ridden Walk and Trot Exercise. *Equine Vet. J.* **2018**, *51*, 108–114. [[CrossRef](#)]
9. Setterbo, J.J.; Garcia, T.C.; Campbell, I.P.; Reese, J.L.; Morgan, J.M.; Kim, S.Y.; Hubbard, M.; Stover, S.M. Hoof Accelerations and Ground Reaction Forces of Thoroughbred Racehorses Measured on Dirt, Synthetic, and Turf Track Surfaces. *Am. J. Vet. Res.* **2009**, *70*, 1220–1229. [[CrossRef](#)]
10. Chateau, H.; Holden, L.; Robin, D.; Falala, S.; Pourcelot, P.; Estoup, P.; Denoix, J.-M.; Crevier-Denoix, N. Biomechanical Analysis of Hoof Landing and Stride Parameters in Harness Trotter Horses Running on Different Tracks of a Sand Beach (from Wet to Dry) and on an Asphalt Road. *Equine Vet. J.* **2010**, *42* (Suppl. 3), 488–495. [[CrossRef](#)]
11. Alexander, R.M.; Maloiy, G.M.O.; Hunter, B.; Jayes, A.S.; Nturibi, J. Mechanical Stresses during Fast Locomotion of Buffalo (*Syncerus Caffer*) and Elephant (*Loxodonta Africana*). *J. Zool. Lond.* **1979**, *189*, 135–144. [[CrossRef](#)]
12. Witte, T.H.; Knill, K.; Wilson, A.M. Determination of Peak Vertical Ground Reaction Force from Duty Factor in the Horse (*Equus Caballus*). *J. Exp. Biol.* **2004**, *207*, 3639–3648. [[CrossRef](#)]
13. Horan, K.; Coburn, J.; Kourdache, K.; Day, P.; Harborne, D.; Brinkley, L.; Carnall, H.; Hammond, L.; Peterson, M.; Millard, S.; et al. Influence of Speed, Ground Surface and Shoeing Condition on Hoof Breakover Duration in Galloping Thoroughbred Racehorses. *Animals* **2021**, *11*, 2588. [[CrossRef](#)]
14. McCracken, M.J.; Kramer, J.; Keegan, K.G.; Lopes, M.; Wilson, D.A.; Reed, S.K.; LaCarrubba, A.; Rasch, M. Comparison of an Inertial Sensor System of Lameness Quantification with Subjective Lameness Evaluation. *Equine Vet. J.* **2012**, *44*, 652–656. [[CrossRef](#)]
15. Keegan, K.G.; MacAllister, C.G.; Wilson, D.A.; Gedon, C.A.; Kramer, J.; Yonezawa, Y.; Maki, H.; Pai, P.F. Comparison of an Inertial Sensor System with a Stationary Force Plate for Evaluation of Horses with Bilateral Forelimb Lameness. *Am. J. Vet. Res.* **2012**, *73*, 368–374. [[CrossRef](#)] [[PubMed](#)]

16. Bell, R.P.; Reed, S.K.; Schoonover, M.J.; Whitfield, C.T.; Yonezawa, Y.; Maki, H.; Pai, P.F.; Keegan, K.G. Associations of Force Plate and Body-Mounted Inertial Sensor Measurements for Identification of Hind Limb Lameness in Horses. *Am. J. Vet. Res.* **2016**, *77*, 337–345. [[CrossRef](#)] [[PubMed](#)]
17. Arkell, M.; Archer, R.M.; Guitian, F.J.; May, S.A. Evidence of Bias Affecting the Interpretation of the Results of Local Anaesthetic Nerve Blocks When Assessing Lameness in Horses. *Vet Rec.* **2006**, *159*, 346–349. [[CrossRef](#)]
18. Parkes, R.S.V.; Weller, R.; Groth, A.M.; 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*, 112–117. [[CrossRef](#)]
19. Hardeman, A.M.; Serra Bragança, F.M.; Swagemakers, J.H.; Weeren, P.R.; Roepstorff, L. Variation in Gait Parameters Used for Objective Lameness Assessment in Sound Horses at the Trot on the Straight Line and the Lunge. *Equine Vet. J.* **2019**, *51*, 831–839. [[CrossRef](#)] [[PubMed](#)]
20. Pfau, T.; Stubbs, N.C.; Kaiser, L.J.; Brown, L.E.A.; Clayton, H.M. Effect of Trotting Speed and Circle Radius on Movement Symmetry in Horses during Lunging on a Soft Surface. *Am. J. Vet. Res.* **2012**, *73*, 1890–1899. [[CrossRef](#)]
21. Pfau, T.; Jennings, C.; Mitchell, H.; Olsen, E.; Walker, A.; Egenvall, A.; Tröster, S.; Weller, R.; Rhodin, M. Lunging on Hard and Soft Surfaces: Movement Symmetry of Trotting Horses Considered Sound by Their Owners. *Equine Vet. J.* **2016**, *48*, 83–89. [[CrossRef](#)]
22. Rhodin, M.; Pfau, T.; Roepstorff, L. Egenvall, a Effect of Lunging on Head and Pelvic Movement Asymmetry in Horses with Induced Lameness. *Vet. J.* **2013**, *198*, e39–e45. [[CrossRef](#)]
23. Rhodin, M.; Roepstorff, L.; French, A.; Keegan, K.G.; Pfau, T.; Egenvall, A. Head and Pelvic Movement Asymmetry during Lunging in Horses with Symmetrical Movement on the Straight. *Equine Vet. J.* **2015**, *48*, 315–320. [[CrossRef](#)]
24. Rhodin, M.; Egenvall, A.; Andersen, P.H.; Pfau, T. Head and Pelvic Movement Asymmetries at Trot in Riding Horses in Training and Perceived as Free from Lameness by the Owner. *PLoS ONE* **2017**, *12*, e0176253. [[CrossRef](#)]
25. Hawkins, A.; O’Leary, L.; Bolt, D.; Fiske-Jackson, A.; Berner, D.; Smith, R. Retrospective Analysis of Oblique and Straight Distal Sesamoidean Ligament Desmitis in 52 Horses. *Equine Vet. J.* **2021**, *54*, 312–322. [[CrossRef](#)] [[PubMed](#)]
26. Pfau, T.; Robilliard, J.J.; Weller, R.; Jespers, K.; Eliashar, E.; Wilson, A.M. Assessment of Mild Hindlimb Lameness during over Ground Locomotion Using Linear Discriminant Analysis of Inertial Sensor Data. *Equine Vet. J.* **2007**, *39*, 407–413. [[CrossRef](#)] [[PubMed](#)]
27. Keegan, K.G.; Wilson, D.A.; Wilson, D.J.; Smith, B.; Gaughan, E.M.; Pleasant, R.S.; Lillich, J.D.; Kramer, J.; Howard, R.D.; Bacon-Miller, C. 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*, 1370–1377.
28. Keegan, K.G.; Dent, E.V.; Wilson, D.A.; Janicek, J.; Kramer, J.; Lacarrubba, A.; Walsh, D.M.; Cassells, M.W.; Esther, T.M.; Schiltz, P.; et al. Repeatability of Subjective Evaluation of Lameness in Horses. *Equine Vet. J.* **2010**, *42*, 92–97. [[CrossRef](#)]
29. Pfau, T.; Reilly, P. How Low Can We Go? Influence of Sample Rate on Equine Pelvic Displacement Calculated from Inertial Sensor Data. *Equine Vet. J.* **2020**, *53*, 1075–1081. [[CrossRef](#)] [[PubMed](#)]
30. Pfau, T.; Witte, T.H.; Wilson, A.M. A Method for Deriving Displacement Data during Cyclical Movement Using an Inertial Sensor. *J. Exp. Biol.* **2005**, *208*, 2503–2514. [[CrossRef](#)] [[PubMed](#)]
31. Warner, S.M.; Koch, T.O.; Pfau, T. Inertial Sensors for Assessment of Back Movement in Horses during Locomotion over Ground. *Equine Vet. J.* **2010**, *42*, 417–424. [[CrossRef](#)]
32. Starke, S.D.; Witte, T.H.; May, S.A.; Pfau, T. Accuracy and Precision of Hind Limb Foot Contact Timings of Horses Determined Using a Pelvis-Mounted Inertial Measurement Unit. *J. Biomech.* **2012**, *45*, 1522–1528. [[CrossRef](#)]
33. Starke, S.D.; Willems, E.; May, S.A.; Pfau, T. Vertical Head and Trunk Movement Adaptations of Sound Horses Trotting in a Circle on a Hard Surface. *Vet. J.* **2012**, *193*, 73–80. [[CrossRef](#)]
34. Haeb-Umbach, R.; Ney, H. Linear Discriminant Analysis for Improved Large Vocabulary Continuous Speech Recognition. In Proceedings of the ICASSP 1992, San Francisco, CA, USA, 23–26 March 1992; IEEE Computer Society: Washington, DC, USA, 1992.
35. Pytkkonen, J. LDA Based Feature Estimation Methods for LVCSR. In Proceedings of the INTERSPEECH 2006-ICSLP, Pittsburgh, PA, USA, 17 September 2006.
36. Pfau, T.; Persson-Sjodin, E.; Gardner, H.; Orssten, O.; Herlund, E.; Rhodin, M. Effect of Speed and Surface Type on Individual Rein and Combined Left–Right Circle Movement Asymmetry in Horses on the Lunge. *Front. Vet. Sci.* **2021**, *8*, 692031. [[CrossRef](#)]
37. Hammarberg, M.; Egenvall, A.; Pfau, T.; Rhodin, M. Rater Agreement of Visual Lameness Assessment in Horses during Lunging. *Equine Vet. J.* **2016**, *48*, 78–82. [[CrossRef](#)] [[PubMed](#)]
38. Sepulveda Caviedes, M.F.; Forbes, B.S.; Pfau, T. Repeatability of Gait Analysis Measurements in Thoroughbreds in Training. *Equine Vet. J.* **2018**, *50*, 513–518. [[CrossRef](#)] [[PubMed](#)]
39. Starke, S.; Raistrick, K.; May, S.; Pfau, T. The Effect of Trotting Speed on the Evaluation of Subtle Lameness in Horses. *Vet. J.* **2013**, *197*, 245–252. [[CrossRef](#)] [[PubMed](#)]
40. Persson-Sjodin, E.; Serra Braganca, F.; Pfau, T.; Egenvall, A.; Weishaupt, M.; Rhodin, M. Movement Symmetry of the Withers Can Be Used to Discriminate Primary Forelimb Lameness from Compensatory Forelimb Asymmetry in Horses with Induced Lameness. *Equine Vet. J.* **2016**, *48*, 32.



41. Rhodin, M.; Persson-Sjodin, E.; Egenvall, A.; Serra Braganca, F.; Pfau, T.; Roepstorff, L.; Weishaupt, M.; Thomsen, M.H.; van Weeren, P.R.; Hernlund, E.; et al. Vertical Movement Symmetry of the Withers in Horses with Induced Forelimb and Hindlimb Lameness at Trot. *Equine Vet. J.* **2018**, *50*, 818–824. [[CrossRef](#)]
42. Pfau, T.; Noordwijk, K.; Caviedes, M.F.S. Head, Withers and Pelvic Movement Asymmetry and Their Relative Timing in Trot in Racing Thoroughbreds in Training. *Equine Vet. J.* **2018**, *50*, 117–124. [[CrossRef](#)]
43. Buchner, H.H.; Savelberg, H.H.; Schamhardt, H.C.; Barneveld, A. Head and Trunk Movement Adaptations in Horses with Experimentally Induced Fore- or Hindlimb Lameness. *Equine Vet. J.* **1996**, *28*, 71–76. [[CrossRef](#)]
44. May, S.A.; Wyn-Jones, G. Identification of Hindleg Lameness. *Equine Vet. J.* **1987**, *19*, 185–188. [[CrossRef](#)]
45. Maliye, S.; Marshall, J.F. Objective Assessment of the Compensatory Effect of Clinical Hind Limb Lameness in Horses: 37 Cases (2011–2014). *Am. J. Vet. Res.* **2016**, *249*, 940–944. [[CrossRef](#)]
46. Maliye, S.; Voute, L.C.; Marshall, J.F. Naturally-Occurring Forelimb Lameness in the Horse Results in Significant Compensatory Load Redistribution during Trotting. *Vet. J.* **2015**, *204*, 208–213. [[CrossRef](#)]
47. Uhler, C.; Licka, T.; Kübber, P.; Peham, C.; Scheidl, M.; Girtler, D. Compensatory Movements of Horses with a Stance Phase Lameness. *Equine Vet. J.* **1997**, *23*, 102–105. [[CrossRef](#)] [[PubMed](#)]
48. Weishaupt, M.A.; Wiestner, T.; Hogg, H.P.; Jordan, P.; Auer, J.A. Compensatory Load Redistribution of Horses with Induced Weightbearing Hindlimb Lameness Trotting on a Treadmill. *Equine Vet. J.* **2004**, *36*, 727–733. [[CrossRef](#)] [[PubMed](#)]
49. Weishaupt, M.A.; Wiestner, T.; Hogg, H.P.; Jordan, P.; Auer, J.A. Compensatory Load Redistribution of Horses with Induced Weight-Bearing Forelimb Lameness Trotting on a Treadmill. *Vet. J.* **2006**, *171*, 135–146. [[CrossRef](#)] [[PubMed](#)]
50. Pfau, T.; Boulthbee, H.; Davis, H.; Walker, A.; Rhodin, M. Agreement between Two Inertial Sensor Gait Analysis Systems for Lameness Examinations in Horses. *Equine Vet. Educ.* **2016**, *28*, 203–208. [[CrossRef](#)]
51. Pfau, T.; Caviedes, M.F.S.; McCarthy, R.; Cheetham, L.; Forbes, B.; Rhodin, M. Comparison of Visual Lameness Scores to Gait Asymmetry in Racing Thoroughbreds during Trot In-Hand. *Equine Vet. Educ.* **2018**, *32*, 191–198. [[CrossRef](#)]
52. Hewetson, M.; Christley, R.M.; Hunt, I.D.; Voute, L.C. Investigations of the Reliability of Observational Gait Analysis for the Assessment of Lameness in Horses. *Vet. Rec.* **2006**, *158*, 852–857. [[CrossRef](#)]
53. Rey, A.; Goldstein, R.M.; Perruchet, P. Does Unconscious Thought Improve Complex Decision Making? *Psychol. Res.* **2009**, *73*, 372–379. [[CrossRef](#)]