Data Article

Gait quality scoring data of Franches-Montagnes stallions at walk and trot on a treadmill by experts of the breed and their reliability

Annik Imogen Gmel [1,2,*], Gerhard Gmel [3], Michael A. Weishaupt [2], Markus Neuditschko [1]

[1] *Animal GenoPhenomics, Agroscope, Route de la Tioleyre 4, 1725 Posieux, Switzerland*
[2] *Equine Department, Vetsuisse Faculty University of Zurich, Winterthurerstrasse 260, 8057 Zurich, Switzerland*
[3] *Alcohol Treatment Centre, Lausanne University Hospital and University of Lausanne, Beaumont 21bis, CHUV, 1011 Lausanne, Switzerland*

ARTICLE INFO

ABSTRACT

This article presents the data obtained from the scoring of 24 stallions of the Franches-Montagnes (FM) horse breed by six experts of this breed. The experts scored six traits at walk and eight at trot from the video recordings of these stallions walking and trotting on a treadmill during an incremental speed test. The scores were given on a scale of one to nine. All experts scored the same videos twice (two scoring tests) with a time interval of two years, and without feedback from the first scoring. Video sequences were presented in a different order between first and second scoring. The inter- and intra-rater reliability of the data was assessed

---

* Corresponding author.
  E-mail address: annik.gmel@agroscope.admin.ch (A.I. Gmel).

using intraclass correlation coefficients (ICC) to evaluate its quality.

Specifications Table

Subject	Veterinary science
Specific subject area	Inter- and intra-rater reliability of scores describing gait quality in the Franches-Montagnes (FM) horse breed.
Type of data	Tables and video recordings (in repository)
How data were acquired	Video recordings of FM stallions were obtained during an incremental speed test, observational scoring data was obtained using a pre-designed scoring sheet. Statistical analysis using Microsoft Excel and R.
Data format	Raw data: two ".csv" files, video files as ".mp4" files (in repository)
Parameters for data collection	Video recordings: stallions needed to move regularly for 20 seconds. Observational data: experts needed to be officially approved experts from the FM breeding association.
Description of data collection	Six experts of the FM horse breed observed four-minute video clips of 24 FM stallions walking and trotting on a treadmill during an incremental speed test. The experts then scored the stallions using a scoring sheet describing six traits at walk and eight traits at trot on a scale from one to nine.
Data source location	Institutions: Agroscope Country: Switzerland
Data accessibility	The raw scoring data are available as .csv files in the following repository: Repository name: Mendeley Data Data identification number: doi: 10.17632/n9g4v3w894.1 Direct URL to data: https://data.mendeley.com/datasets/n9g4v3w894/1 All scored videos are provided in the following repository: Repository name: Mendeley Data Data identification number: doi: 10.17632/bwfkyvy593.1 Direct URL to data: https://data.mendeley.com/datasets/bwfkyvy593/1

Value of the Data

- This dataset provides information on the scores attributed by experts of the FM horse breed on gait quality when observing the videos of stallions moving at the walk and trot on a treadmill during an incremental speed test.
- This dataset was used to estimate the inter- and intra-rater reliability of the scoring by the experts.
- The provided videos can be used by equine professionals for teaching and training.
- This dataset can be used to correlate expert scorings of gait quality with objective kinematic and kinetic measurements, to understand the external validity of the appraisal of horses.

1. Data Description

The data described here consists of two spreadsheets (Raw_scores_first_scoring.csv and Raw_scores_second_scoring.csv, https://data.mendeley.com/datasets/n9g4v3w894/1) of gait quality scores from two separate video scoring sessions of Franches-Montagnes (FM) stallions walking and trotting on a treadmill during an incremental speed test. The video clips of the stallions are accessible on the Mendeley Data repository (https://data.mendeley.com/datasets/bwfkyvy593/1). The video clips are in mp4 format with one video per stallion, at the walk and

Table 1
List of scored traits for the present dataset. Textual definitions from Gmel et al. [1]

Variable name	Trait	Definition	Score 1	Score 9
Walk				
W_Coverage	Ground coverage	How far and how wide the forelimb travels. Related to the length of the stride. The longer the stride, the more ground is covered (scope).	Undesirable	Ideal
W_Overtracking	Over-tracking	The distance between the hoof imprint of the hind limb in relation to the ipsilateral forelimb.	Undesirable	Ideal
W_Relaxation	Suppleness and relaxation	The horse, calm and concentrated, moves in a relaxed way, with high mobility of the head, neck, back and a free movement of the tail.	Undesirable	Ideal
W_Regularity	Regularity and harmony	The walk has a clear four-beat rhythm. The stride length and frequency remain constant over the length of the triangle, all body parts are used equally.	Undesirable	Ideal
W_Activity	Activity and impulsion	Represents the activity of the hindquarters and the transmission of stored energy from the hindquarters to the forehand.	Undesirable	Ideal
W_SL_LP	Walk (LP)	The length of the stride.	Short	Long
Trot				
T_Coverage	Ground coverage	How far and how wide the forelimb travels. Related to the length of the stride. The longer the stride, the more ground is covered (scope).	Undesirable	Ideal
T_Overtracking	Over-tracking	The position of the hind limb hoof in relation to the tuber coxae at impact	Undesirable	Ideal
T_Relaxation	Suppleness and relaxation	The horse, calm and concentrated, moves in a relaxed way, with high mobility of the head, neck, back and a free movement of the tail.	Undesirable	Ideal
T_Regularity	Regularity and harmony	The trot has a clear two-beat rhythm. The stride length and frequency remain constant over the length of the triangle, all body parts are used equally.	Undesirable	Ideal
T_Activity	Activity and impulsion	Represents the activity of the hindquarters and the transmission of stored energy from the hindquarters to the forehand.	Undesirable	Ideal
T_SL_LP	Trot (LP)	The length of the stride.	Short	Long
T_Impulsion_LP	Trot: Impulsion (LP)	Represents the activity of the hindquarters and the transmission of stored energy from the hindquarters to the forehand.	Little	Much
T_Elasticity_LP	Trot: Elasticity (LP)	The horse moves with high flexibility of the head, neck, back and a free movement of the tail.	Stiff	Elastic

trot at two different speeds, from three different perspectives. Within each gait and speed, the stallion was viewed 20 seconds from behind, then the left side and then from the front; first the two walk sessions followed by the two trot sessions, in order of speed (walk at 1.7 m/s, peak walk, trot at 4.5 m/s, peak trot), for a total of four minutes per stallion.

Six experts of the FM breed (coded A to F) scored these video clips of 24 FM stallions for six gait quality traits at the walk and eight traits at the trot. In total, each spreadsheet contains 144 entries (one score per horse and expert) for 14 variables, i.e. 13121 data points per scoring session. The traits are scored from 1 to 9, and the variable names in the spreadsheets Raw_scores_first_scoring.csv and Raw_scores_second_scoring.csv and the trait names are presented in Table 1. The textual definitions were first described in Gmel et al. [1], during which the same traits were scored for the same 24 stallions during over-ground locomotion by seven experts of the breed [1].

Table 2
Descriptive statistics (mean, standard deviation, median, mode and range) of gait quality traits of 24 FM stallions walking and trotting on the treadmill scored for the first time over video

| Trait	Mean	SD	Median	Mode	Minimum	Maximum
Walk						
Ground coverage	6.67	1.05	7.00	7	3	8
Over-tracking	6.83	1.21	7.00	8	3	9
Suppleness and relaxation	6.50	1.13	7.00	7	4	9
Regularity and harmony	6.31	1.15	6.00	7	3	9
Activity and impulsion	6.66	1.26	7.00	8	3	9
Walk (LP)	6.70	1.12	7.00	7	3	9
Trot						
Ground coverage	6.58	1.11	7.00	6	4	9
Over-tracking	6.33	1.22	6.00	6	3	9
Suppleness and relaxation	6.33	1.19	6.00	6	3	9
Regularity and harmony	6.48	1.15	7.00	7	3	9
Activity and impulsion	6.40	1.33	6.00	6	3	9
Trot (LP)	6.73	1.03	7.00	7	4	9
Trot: Impulsion (LP)	6.38	1.17	6.00	7	3	9
Trot: Elasticity (LP)	6.33	1.19	6.00	7	3	9

Table 3
Descriptive statistics (mean, standard deviation, median, mode and range) of scores of gait quality traits of 24 FM stallions walking and trotting on the treadmill scored for the second time over video

Trait	Mean	SD	Median	Mode	Minimum	Maximum
Walk						
Ground coverage	6.63	1.04	7.00	7	3	9
Over-tracking	6.61	1.27	7.00	7	3	9
Suppleness and relaxation	6.29	1.18	6.00	6	3	9
Regularity and harmony	6.43	1.08	6.50	7	3	9
Activity and impulsion	6.55	1.27	7.00	6	3	9
Walk (LP)	6.62	1.15	7.00	7	3	9
Trot						
Ground coverage	6.78	1.07	7.00	7	3	9
Over-tracking	6.32	1.17	6.00	6	2	9
Suppleness and relaxation	6.41	1.16	6.00	6 and 7	3	9
Regularity and harmony	6.53	1.18	7.00	6	2	9
Activity and impulsion	6.47	1.26	6.00	6	2	9
Trot (LP)	6.85	1.04	7.00	7	4	9
Trot: Impulsion (LP)	6.34	1.28	6.00	6	2	9
Trot: Elasticity (LP)	6.50	1.25	7.00	7	2	9

Descriptive statistics of the scores from the six experts of the breed are presented by scoring session (first: Table 2; second: Table 3). For the first video scoring, the mean scores ranged from 6.31 to 6.83 (Table 2). For the second video scoring, the means ranged from 6.29 to 6.85 (Table 3).

For each scoring, the scores for all traits were summarised for each expert in violin plots to visualise the scale anchoring (frequency of use for each score and effective range of scores, Fig. 1). Expert E had the widest use of the scale. Expert A had consistently the narrowest use of the scale.
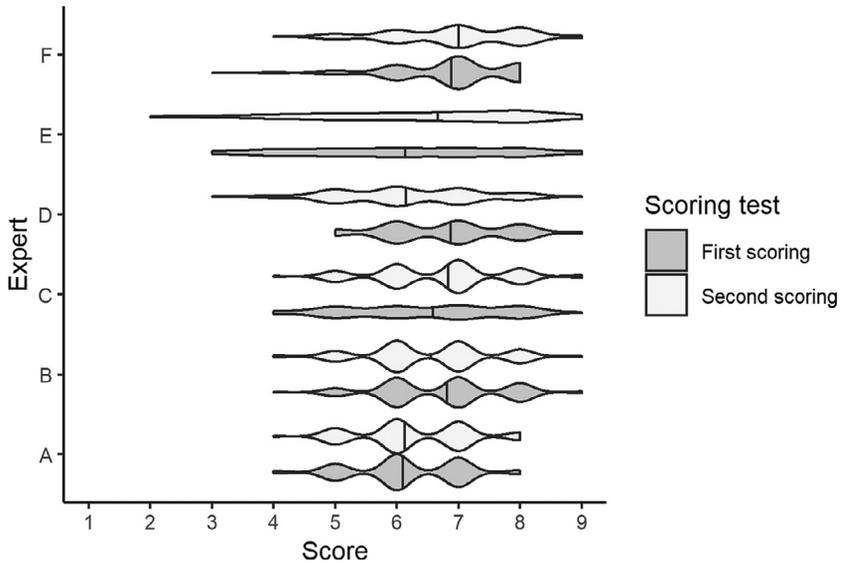
Fig. 1. Violin plots of the scores given by each expert for the scoring of 24 FM stallions at the walk and trot for each scoring. The middle line represents the median. The width of the violin plot represents the number of times a score was used.

The inter-rater reliability was estimated for each trait, gait and scoring (first and second video scoring, Table 4). The inter-rater reliabilities were poor for all traits during the first video scoring. Only one trait (suppleness and relaxation), had moderate reliability during the second video scoring (all other reliabilities were poor). The second video scoring had higher inter-rater reliability estimates than the first video scoring.

Table 4

Intraclass correlation coefficients (ICC) and their confidence intervals of gait quality traits for the first and second video scoring of 24 FM stallions at the walk and trot by six FM breeding experts

Trait	First scoring			Second scoring		
	ICC	Low CI	High CI	ICC	Low CI	High CI
Walk						
Ground coverage	0.26	0.12	0.47	0.34	0.18	0.55
Over-tracking	0.22	0.08	0.42	0.21	0.06	0.43
Suppleness and relaxation	0.21	0.07	0.42	0.22	0.09	0.43
Regularity and harmony	0.13	0.03	0.31	0.31	0.16	0.52
Activity and impulsion	0.26	0.12	0.47	0.32	0.16	0.53
Walk (LP)	0.32	0.16	0.53	0.38	0.22	0.59
Trot						
Ground coverage	0.26	0.12	0.47	0.42	0.24	0.63
Over-tracking	0.25	0.11	0.45	0.29	0.14	0.51
Suppleness and relaxation	0.15	0.03	0.34	0.50	0.33	0.69
Regularity and harmony	0.10	-0.01	0.27	0.38	0.22	0.58
Activity and impulsion	0.29	0.14	0.49	0.36	0.20	0.57
Trot (LP)	0.26	0.12	0.46	0.49	0.32	0.69
Trot: Impulsion (LP)	0.30	0.15	0.51	0.43	0.26	0.64
Trot: Elasticity (LP)	0.21	0.08	0.40	0.49	0.32	0.68

The pairwise agreement between experts was estimated with correlation coefficients (Fig. 2). In the first video scoring, the highest pairwise correlation over all scores was between experts E and F, and the lowest between experts A and C. In the second video scoring, the highest correlation was between experts B and C and the lowest between experts C and E.
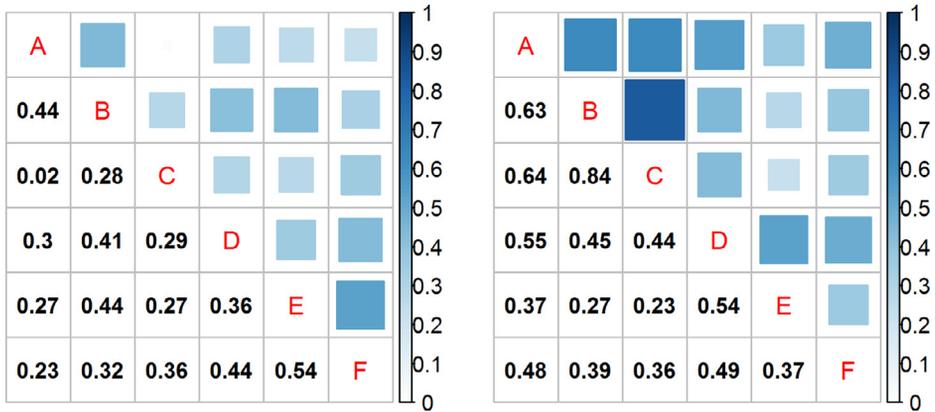

Fig. 2. cross-correlation matrix of the experts across all traits (at the walk and trot) for the first (a) and second (b) video scoring, with the corresponding numerical values below the diagonal.

Expert B had the highest intra-rater reliability at the walk for the trait "suppleness and re-laxation" (Table 5), and Expert E had the highest intra-rater reliability at the trot for the trait "ground coverage" (Table 6). Expert A had the lowest intra-rater reliability for the trait "regular-ity and harmony" at both the walk and trot.

Table 5

Intra-rater reliability of gait quality traits of 24 FM stallions at the walk of video scoring from six experts of the breed using the intraclass-correlation coefficient (ICC) for absolute agreement, with their confidence intervals (low and high CI).

Trait	A ICC	A Low CI	A High CI	B ICC	B Low CI	B High CI	C ICC	C Low CI	C High CI	D ICC	D Low CI	D High CI	E ICC	E Low CI	E High CI	F ICC	F Low CI	F High CI
Ground coverage	0.35	-0.02	0.65	0.38	-0.02	0.68	0.41	0.01	0.70	0.18	-0.11	0.49	0.49	0.11	0.74	0.26	-0.11	0.58
Over-tracking	0.31	-0.07	0.62	0.26	-0.17	0.60	0.17	-0.23	0.54	0.23	-0.10	0.55	0.43	0.03	0.71	0.29	-0.12	0.62
Suppleness and relaxation	-0.09	-0.48	0.33	0.56	0.21	0.78	0.46	0.07	0.73	0.30	-0.06	0.61	0.10	-0.33	0.48	0.34	-0.06	0.65
Regularity and harmony	-0.10	-0.49	0.31	0.19	-0.25	0.55	0.25	-0.18	0.59	0.17	-0.19	0.52	0.41	0.00	0.70	0.46	0.08	0.72
Activity and impulsion	0.18	-0.24	0.54	0.26	-0.17	0.60	0.44	0.05	0.71	0.32	-0.05	0.62	0.21	-0.21	0.56	0.40	0.00	0.69
Walk (LP)	0.22	-0.20	0.57	0.26	-0.17	0.60	0.44	0.06	0.71	0.34	-0.02	0.64	0.39	-0.01	0.68	0.43	0.04	0.71
Mean	0.15	-	-	0.32	-	-	0.36	-	-	0.26	-	-	0.34	-	-	0.36	-	-
Minimum	-0.10	-	-	0.19	-	-	0.17	-	-	0.17	-	-	0.10	-	-	0.26	-	-
Maximum	0.35	-	-	0.56	-	-	0.46	-	-	0.34	-	-	0.49	-	-	0.46	-	-

Table 6

Intra-rater reliability of gait quality traits of 24 FM stallions at the trot of video scoring from six experts of the breed using the intraclass-correlation coefficient (ICC) for absolute agreement, with their confidence intervals (low and high CI).

Trait	A			B			C			D			E			F		
	ICC	Low CI	High CI	ICC	Low CI	High CI	ICC	Low CI	High CI	ICC	Low CI	High CI	ICC	Low CI	High CI	ICC	Low CI	High CI
Ground coverage	0.48	0.12	0.73	0.57	0.23	0.78	0.20	-0.11	0.51	0.34	-0.02	0.64	0.69	0.40	0.85	0.16	-0.13	0.47
Over-tracking	0.21	-0.19	0.56	0.41	-0.03	0.71	0.46	0.07	0.73	0.36	-0.04	0.66	0.62	0.28	0.82	0.52	0.16	0.76
Suppleness and relaxation	0.06	-0.35	0.45	0.33	-0.07	0.64	0.26	-0.10	0.58	0.39	0.03	0.68	0.59	0.24	0.80	0.33	-0.10	0.64
Regularity and harmony	-0.04	-0.44	0.36	0.28	-0.15	0.61	0.18	-0.15	0.51	0.10	-0.24	0.45	0.56	0.22	0.78	0.15	-0.28	0.52
Activity and impulsion	0.25	-0.17	0.59	0.48	0.11	0.73	0.25	-0.13	0.58	0.54	0.20	0.77	0.52	0.17	0.76	0.15	-0.28	0.52
Trot (LP)	0.49	0.13	0.74	0.65	0.36	0.83	0.32	-0.04	0.62	0.30	-0.06	0.60	0.57	0.22	0.79	0.42	0.02	0.70
Trot: Impulsion (LP)	0.24	-0.19	0.58	0.47	0.06	0.74	0.38	-0.03	0.68	0.48	0.12	0.73	0.58	0.25	0.79	0.42	0.01	0.71
Trot: Elasticity (LP)	0.12	-0.26	0.48	0.30	-0.09	0.62	0.41	0.05	0.69	0.41	0.03	0.69	0.39	0.01	0.68	0.20	-0.22	0.55
Mean	0.23	-	-	0.44	-	-	0.31	-	-	0.37	-	-	0.57	-	-	0.29	-	-
Minimum	-0.04	-	-	0.28	-	-	0.18	-	-	0.10	-	-	0.39	-	-	0.15	-	-
Maximum	0.49	-	-	0.65	-	-	0.46	-	-	0.54	-	-	0.69	-	-	0.52	-	-

2. Experimental Design, Materials and Methods

24 Franches-Montagnes (FM) stallions were kinematically and kinetically measured and recorded on video in walk and trot during an incremental speed test (walking speed range 1.4 to 2.0 m/s at 0.1 m/s increments; trotting speed range 3.3 to 6.5 m/s at 0.5 m/s increments above 4.0 m/s). During the incremental speed test, the stallions were video-recorded from the front, from behind and from the left side with three Sony Camcorders (HDR-CX760). The videos were cut into 20 second clips for each walking and trotting speed using the software iMovie (Apple Inc.). The video sequences were prepared as follows: within each gait and speed, the horse was viewed from behind, then the left side and then from the front; first the two walk sessions followed by the two trot sessions, in order of speed (walk at 1.7 m/s, peak walk, trot at 4.5 m/s, peak trot). A 10 second black screen with the code name (FMXXXX) of the stallion to be appraised was added at the beginning of the video to ensure that the experts filled out the scoring sheet for the corresponding stallion. All videos are provided in the repository.

In the FM breed, there are nine official experts of the breed, selected from a pool of regional experts, and elected for a period of four years by a delegation of breeders, for a maximum of three consecutive mandates. Six out of these nine official experts agreed to score the videos based on a previously published scoring scheme [1], and were further designated with a letter from A to F. The experts were separated into three groups (three groups of two: A-B, C-D, E-F), each group presented with a randomised sequence of the 24 stallion videos (Table S1). To anticipate the challenge of scoring treadmill exercise videos, which none of the experts were familiar with, video sequences of three randomly selected stallions were put in the same order for all the experts before the official video scoring, during which the traits to evaluate were briefly discussed within each expert pair and the first author. The experts did not consult each other during the official video scoring. The video for each stallion was cut to four minutes, whereby the entire video scoring lasted approximately two hours.

The same six experts were invited a second time, two years after the first video scoring test and without any feedback concerning the former, to re-score the same videos to estimate both inter- and intra-rater reliabilities. This time, the experts were separated into three groups (two groups of three) divided by geographic regions and language (A-B-C, D-E-F). The same three video sequences were presented again before the official scoring, and the experts used these first three sequences to discuss the traits to be scored with the first author. Subsequently, the videos of the 24 stallions were presented in another randomised order than the first scoring (Table S2), which the experts assessed independently from one another.

The scores from each expert were transcribed from paper onto two electronic files (files Raw_scores_first_scoring.csv and Raw_scores_second_scoring.csv), and analysed for each scoring test (first and second video scoring). Besides the descriptive statistics (mean, sd, median, mode, range), the inter-rater reliability of each trait was evaluated with an intra-class correlation coefficient (ICC) for absolute agreement [2]. The intra-rater reliability was estimated for each expert using an ICC for absolute agreement [2].

3. Implication of the Data

The second video scoring showed higher inter-rater reliabilities than the first video scoring. The scores from the second video scoring should be used when comparing scores to kinematics measurements. The intra-rater reliabilities were expert-specific and depended on the gait. Considering the low ICC for most of the traits, the current scoring system is not reliable enough to accurately assess the gait quality of Franches-Montagnes horses on the treadmill.

Ethics Statements

The videos used for the scorings were recorded during a study conducted under animal experiment permit number VD 3164 complying with Swiss Federal Legislation. No animal was harmed during the experiment. The experts participated in the appraisal voluntarily and data was anonymised to ensure privacy.

CRediT Author Statements

Annik Imogen Gmel: Conceptualisation, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualisation, Funding acquisition; **Gerhard Gmel:** Validation; **Michael Andreas Weishaupt:** Conceptualisation, Investigation, Data curation, Writing – review & editing, Resources, Supervision; **Markus Neuditschko:** Writing – review & editing, Funding acquisition, Resources, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this manuscript.

Data Availability

Video clips of Franches-Montagnes stallions at walk and trot on a treadmill during an incremental speed test (Original data) (Mendeley Data).
Gait quality scoring data of Franches-Montagnes stallions at walk and trot on a treadmill by experts of the breed (Original data) (Mendeley Data).

Acknowledgments

Supplementary Materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.dib.2022.108123.

References

[1] A.I. Gmel, G. Gmel, R. von Niederhäusern, M.A. Weishaupt, M. Neuditschko, Should we agree to disagree? An evaluation of the inter-rater reliability of gait quality traits in Franches-Montagnes stallions, Journal of Equine Veterinary Science (2020) 102932.

[2] T.K. Koo, M.Y. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, Journal of chiropractic medicine 15 (2) (2016) 155–163.