

Standard Article

J Vet Intern Med 2017;31:1871–1876

Interobserver Variation in the Diagnosis of Neurologic Abnormalities in the Horse

W.J.A. Saville , S.M. Reed, J.P. Dubey, D.E. Granstrom, P.S. Morley, K.W. Hinchcliff , C.W. Kohn, T.E. Wittum, and J.D. Workman

Background: The diagnosis of equine protozoal myeloencephalitis (EPM) relies heavily on the clinical examination. The accurate identification of neurologic signs during a clinical examination is critical to the interpretation of laboratory results.

Objective: To investigate the level of agreement between board-certified veterinary internists when performing neurologic examinations in horses.

Animals: Ninety-seven horses admitted to the Veterinary Teaching Hospital at The Ohio State University from December 1997 to June 1998.

Methods: A prospective epidemiologic research design was used. Horses enrolled in the study were examined by the internist responsible for care of the horse, and later by an internist who was not aware of the presenting complaint or other patient history. Data were analyzed by descriptive statistics, and kappa (*K*) statistics were calculated to assess interobserver agreement.

Results: Ninety-seven horses were enrolled in the study. Overall, examiners, also referred to as observers, agreed that 60/97 (61.9%) were clinically abnormal, 21/97 (21.6%) were clinically normal, and the status of 16/97 (16.5%) of horses was contested. There was complete agreement among the examiners with regard to cranial nerve signs and involuntary movements. Disagreement involving severity of clinical signs occurred in 31 horses, and 25 of those horses (80.6%) were considered either normal or mildly affected by the primary observer. When examining the results of all paired clinical examinations for 11 different categories, there was wide variability in the results. When examiners rated the presence or absence of any neurologic abnormalities, lameness, or ataxia, the agreement among observers was either good or excellent for 80% of horses. When assessing truncal sway, the agreement among observers was good or excellent for 60% of the horses. When examining the horses for asymmetry of deficits, agreement was either good or excellent for 40% of the horses. Agreement among observers was excellent or good for only 20% of the horses when assessing muscle atrophy, spasticity (hypermetria), and overall assessment of the severity of neurologic abnormalities.

Conclusions and Clinical Importance: This study underscores the subjectivity of the neurologic examination and demonstrates a reasonable level of agreement that may be achieved when different clinicians examine the same horse.

Key words: Diagnosis; Horse; Neurologic.

Equine protozoal myeloencephalitis (EPM) is a common neurologic disease of the horse characterized by abnormalities of gait and other neurologic deficits that may be referable to any part of the central nervous

From the Department of Veterinary Preventive Medicine, College of Veterinary Medicine, The Ohio State University, Columbus, OH (Saville, Wittum, Workman); Rood and Riddle Equine Hospital, Lexington, KY (Reed); United States Department of Agriculture, Agricultural Research Service, Beltsville Agricultural Research Center, Animal Parasitic Diseases Laboratory, Beltsville, MD (Dubey); American Veterinary Medical Association, Schaumburg, IL (Granstrom); Colorado School of Public Health, James L. Voss Veterinary Teaching Hospital, Colorado State University, Fort Collins, CO (Morley); Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville, Vic Australia (Hinchcliff); Department of Veterinary Clinical Sciences, College of Veterinary Medicine, The Ohio State University, Columbus, OH (Kohn).

Corresponding author: William J. A. Saville, Department of Veterinary Preventive Medicine, Ohio State University, A186 Sisson Hall, 1920 Coffey Road, Columbus, Ohio 43210; e-mail: saville.4@osu.edu.

Submitted May 16, 2017; Revised July 21, 2017; Accepted August 3, 2017.

Copyright © 2017 The Authors. *Journal of Veterinary Internal Medicine* published by Wiley Periodicals, Inc. on behalf of the American College of Veterinary Internal Medicine.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

DOI: 10.1111/jvim.14822

Abbreviations:

EPM	equine protozoal myeloencephalitis
<i>K</i>	kappa statistic
<i>P_e</i>	proportion of expected agreement
<i>P_o</i>	proportion of observer agreement

system. Clinical signs most frequently are attributed to spinal cord disease. The Veterinary Teaching Hospital (VTH) at The Ohio State University had a very high caseload of horses with EPM because of referrals to the medicine service. In addition, the majority of all horses with neurologic disease admitted to the VTH had EPM. Thus, our study was primarily focused on EPM.

The diagnosis of EPM relies heavily on the clinical examination. The sensitivity and specificity of the Western blot analysis for antibodies to *S. neurona* in cerebrospinal fluid (CSF) have been reported to be 89 and 89%, respectively, based on comparison to postmortem evaluation in severely affected horses.¹ The prevalence of EPM must be considered when interpreting the results of the assay for *S. neurona* antibody to establish a diagnosis of EPM.² The positive predictive value of a positive test is markedly decreased when the prevalence of disease is low, as is the case for the prevalence of EPM among neurologically normal horses.² Thus, the accurate identification of neurologic signs during a

clinical examination is critical to the interpretation of laboratory results.

Since the original Western blot was described in 1997, several other antibody tests have been developed using other methods of testing.³ In California, an immunofluorescent antibody test was developed and then likelihood ratios of infection were based on antibody concentration in serum and CSF.³ A novel test was developed at Equine Diagnostic Solutions LLC^a in Lexington, Kentucky, using SnSAG2 and SnSAG4/3 ELISA and the ratio of the serum titer to CSF titer was determined.⁴ Likelihood ratios and diagnostic sensitivity and specificity were calculated based on serum titers, CSF titers, and the serum-to-CSF titer ratio. This combination resulted in excellent sensitivity (86.4%) and specificity (95.9%) when the cutoff was extremely rigorous (cutoff ratio of ≤ 50).⁴

Methods are available to assess agreement among clinicians or the repeatability of a clinician's assessment of the same patient, but these have not been used to rate interobserver agreement or repeatability of neurologic examination of horses.⁵ Other methods also are used to rate interobserver variation in neurologic examinations. One example is a standard 2×2 table to describe prevalence, sensitivity, specificity, and validity of the clinical examinations.⁶ In another study, a method using intraclass correlation coefficients (ICC) was applied.⁷ Other studies^{8,9} have used kappa to assess interexaminer agreement in assessment of lameness in horses. However, in yet another study,¹⁰ the Kendall coefficient of concordance was used. All of these methods achieve the same end result using different mathematical models.

There is good agreement among observers when performing neurologic examination on human patients regarding assessment of the speed of limb motion, fluency of movement, variability in patterns of arm movement, and global judgment of movement in human infants, but poor agreement in the assessment of the amplitude of movement, variability of movement patterns, and variability of leg movement patterns.¹¹ Similar information is not available regarding the clinical evaluation of signs of neurologic disease (especially EPM) in the horse. The purpose of our study was to determine the agreement among observers performing a standardized neurologic examination on horses.

Materials and Methods

Study Design

The study was conducted in accordance with the guidelines of the Institutional Animal Care and Use Committee (IACUC) of The Ohio State University. The study population consisted of horses admitted to the Veterinary Teaching Hospital at The Ohio State University from December 1997 to June 1998. Upon admission, the attending medicine clinician performed a complete neurologic examination. A second clinician who was masked to the horse's history then conducted an independent neurologic examination soon after the initial examination on the same day. There were no ancillary tests performed, and horses were not sedated for examination. Owners and trainers were advised not to talk to the

masked examiner, also referred to as observers. Four clinicians who were board-certified in internal medicine participated in the study.

All clinicians performed neurologic examinations according to a standardized protocol (Fig 1). Briefly, the horse was observed in the stall for evidence of behavior abnormalities and cranial nerve deficits that might be related to brainstem or peripheral cranial nerve lesions. The clinician also evaluated the horse for abnormalities in posture or coordination and proprioceptive deficits. Gait was evaluated outside the stall as the horse moved in a straight line at the walk and trot, while walking in wide and tight circles, and backing. Horses also were walked up and down a gradual (15° – 20°) as well as a steep (35° – 45°) incline, and made to walk over objects on the ground. Results of the neurologic examination were categorized as binary outcomes: abnormal neurologic signs (yes/no), asymmetry of signs (yes/no), abnormal cranial nerve signs (yes/no), weakness (yes/no), spasticity or hypermetria (yes/no), and ataxia (yes/no). The overall severity of neurologic signs was evaluated using a grading system: grade "0" for normal horses, grade "1" for abnormalities difficult to see by experienced clinicians, grade "2" for neurologic deficits readily detected by a thorough clinical examination, grade "3" for horses in which neurologic signs were obvious from a distance, grade "4" for horses that had neurologic deficits characterized by falling if turned in circles, and grade "5" for horses that were recumbent and unable to rise. Half scores were used when considered appropriate by clinicians (half scores denoted as "+"). Severity of neurologic deficits was further categorized to facilitate analysis (mild, moderate, and severe). Horses were considered to be mildly affected if abnormalities were grades 1 or 2, moderately affected if abnormalities were grades 2+ to 3+, and severely affected if abnormalities were grades 4 or 5.

Statistical Analysis

Clinical findings of both the attending clinician and the blinded observer were recorded and entered into a Corel Paradox 8^b database. Data regarding the clinical examinations were exported to Microsoft Excel^c for calculation of 2×2 tables for statistical analysis. Agreement among observers was assessed by the kappa statistic (K) which evaluates the proportion of agreement that occurred beyond that expected by chance.⁵ The agreement beyond chance was calculated as the proportion of observer agreement (P_o) minus the proportion of expected agreement (P_e). The maximum possible excess is $1 - P_e$.⁵ The kappa statistic is a ratio of these 2 differences: $K = (P_o - P_e)/(1 - P_e)$

Generally, kappa values of 0 to 0.4 are interpreted as indicating poor agreement, values of 0.40–0.74 are interpreted as indicating good agreement, and values of 0.75–1.0 are interpreted as indicating excellent agreement.¹² Variance was used to estimate the standard deviation of the K statistic as previously described.¹² A 95% confidence interval (CI) then was calculated using the following formula: $K \pm 1.96 \times \text{standard deviation (SD)}$.

Results

Ninety-seven horses were enrolled in the study. Examiners 1 and 2 performed neurologic examinations on 54 horses; 14 horses were evaluated by examiners 1 and 3, 11 horses were evaluated by examiners 1 and 4, 8 horses were evaluated by examiners 2 and 3, and 10 horses by examiners 2 and 4. Overall, examiners agreed that 60/97 (61.9%) were clinically abnormal, 21/97 (21.6%) were clinically normal, and examiners disagreed on results of 16/97 (16.5%) horses (Table 1).

Equine Neurologic Examination

The Ohio State University Veterinary Medical Center

Date: _____

Faculty/Resident: _____

History: _____

General Observations: _____

Cranial Nerve Exam:

Menace (II, VII)	_____	_____	Facial Symmetry:	_____
Pupil Size (II, III, sym)	_____	_____	Temporal/Masseter (V)	_____
Pupil Symmetry (II, sym)	_____	_____	Expressive Muscles (VII)	_____
PLR (I, III)	_____	_____	Palpebral Reflex (V, VII)	_____
Doll's Eye (VIII, III, IV, VI)	_____	_____	Retractor Oculi (V, VI)	_____
Ocular Position (VIII, III, IV, VI)	_____	_____	Gag Reflex (IX, X)	_____
Pathologic Nystagmus	_____	_____	Tongue (XII)	_____

Symmetry of Neck/Body (muscle mass, scoliosis, etc): _____

Manipulation of the Neck: L/R: _____
Up/Down: _____

Spontaneous Involuntary Movements (tremor, myoclonus, myotonia, etc.): _____

Description of Gait (at a walk and trot): _____

Circling: Large: L: _____ Small: L: _____
R: _____ R: _____

Backing: _____

Up/Down an Incline: _____

Elevation of Head: _____

Proprioception: LF: _____ RF: _____
LR: _____ RR: _____

Sway Reaction: Fore: _____
Rear: _____
Tail Pull: _____

Grading System: Write in grading according to deLahunta

0= No Gait Deficits	1= Deficits barely perceptible – worsened with head elevation or provocative testing
2= Deficits noted at a walk	3= Deficits noted at rest, walking, nearly falls with head elevation or provocative testing
4= Falls or nearly falls at normal gaits	5= Recumbent patient

Gait and Posture (graded 0 to +4)

	Motor		Sensory
	Weakness	Spasticity	Ataxia
LF	_____	_____	_____
RF	_____	_____	_____
LR	_____	_____	_____
RR	_____	_____	_____

Reflexes: Anal: _____
Patellar: _____
Triceps: _____
Other: _____

Nociceptive (withdrawal): _____

Tail Tone: _____

Autonomic: Urinary: _____
Rectal: _____
Sweating: _____

Cutaneous Sensation: _____

Assessment/Comments: _____

Lesion Localization: _____

Tentative Dx: _____

CSF

	Cells	Protein	Culture	Cytology
L/S	_____	_____	_____	_____
A/O	_____	_____	_____	_____

Radiographs: _____

Myelogram: _____

Comments: _____

Equine Neurologic Exam

Fig 1. Standardized protocol used to examine horses.

Table 1. Summary of agreement for all Examiners in assessing overall status (neurologic/not neurologic), each aspect of gait (truncal sway, asymmetry, lameness, ataxia, weakness, spasticity [hypermetria]), the presence or absence of muscle atrophy, and overall severity of the neurologic signs. Overall kappa statistics ($K \pm SD$) and 95% confidence intervals for interobserver agreement between the 4 internists for each category of clinical sign observed during a neurologic workup.

Clinical Sign	Agree (+/+)	Agree (-/-)	Disagree (+/-) or (-/+)	Kappa	Standard Deviation	95% Confidence Interval
Neurologic	60	21	16	0.61	0.09	0.44–0.78
Atrophy	4	79	14	0.29	0.14	0.004–0.57
Truncal Sway	8	70	19	0.36	0.13	0.12–0.61
Lameness	10	74	13	0.53	0.12	0.30–0.75
Asymmetry	43	28	26	0.46	0.09	0.28–0.63
Ataxia	58	21	18	0.57	0.09	0.39–0.75
Weakness	56	22	19	0.56	0.09	0.38–0.73
Spasticity (hypermetria)	49	18	30	0.34	0.10	0.13–0.54
Severity	48	18	31	0.30	0.10	0.10–0.49

Cranial nerve deficits were identified in 7.2% (7/97) of horses and involuntary movements such as muscle tremors or muscle fasciculations were identified in 6.2% of horses (6/97). There was complete agreement among the examiners with regard to these signs.

Examiner 1 was considered the primary examiner, and examiner 2 was considered the primary examiner

when examiner 1 was not involved in the case. Of the 97 horses enrolled in the study, the primary examiners considered 25 (25.8%) horses to have no neurologic deficits, 49 (50.5%) were classified as mildly affected, 17 (17.5%) were moderately affected, and 6 (6.2%) were classified as severely affected. Disagreement on assessment of severity of clinical signs occurred in 31 horses,

and 25 of these horses (80.6%) were considered to be normal or mildly affected by the primary examiner.

There was wide variability in the kappa statistics for all paired clinical examinations in the 11 categories assessed. For both cranial nerve signs and involuntary movements, there was excellent agreement 100% of the time ($K = 1$). When the examiners rated the presence or absence of neurologic signs, and the presence or absence of lameness or ataxia, the agreement was either good or excellent for 80% of the horses (Table 2). Agreement for truncal sway was good or excellent for 60% of the horses (Table 2). When rating the asymmetry of abnormalities or weakness, agreement was either good or excellent for 40% of horses (Table 2). Agreement was good or excellent for only 20% of horses when assessing muscle atrophy, spasticity, and the overall severity of clinical signs (Table 2).

Between examiners 1 and 2, after examination of 54 horses, there was excellent agreement ($K \geq 0.75$) with regard to ataxia and weakness. There was good agreement ($K = 0.4-0.74$) between these examiners with regard to the presence or absence of neurologic signs, lameness, symmetry of neurologic signs, spasticity, and severity of the neurologic signs. With regard to muscle atrophy and the presence or absence of truncal sway, there was poor agreement ($K < 0.4$).

Ten horses were examined by clinicians 1 and 3. There was good agreement ($K = 0.4-0.74$) regarding the presence or absence of neurologic signs, truncal sway, lameness, as well as ataxia. There was poor agreement ($K < 0.4$) between the clinicians for muscle atrophy, asymmetric neurologic signs, weakness, spasticity, and severity of the neurologic signs.

Eleven horses were evaluated by examiners 1 and 4. Good agreement ($K = 0.4-0.74$) was present for the presence or absence of neurologic signs, lameness, ataxia, and weakness. There was poor agreement ($K < 0.4$) regarding asymmetry of neurologic signs, spasticity, and severity of the neurologic signs, and no agreement for the presence or absence of muscle atrophy.

Table 2. Agreement (K) among 5 pairs of boarded internists when observing clinical signs of horses. $K =$ kappa.

Clinical Sign	Excellent	Good	Poor
	Agreement $K \geq 0.75$	Agreement $K = 0.74-0.40$	Agreement $K < 0.40$
Neurologic Signs	0/5 (0%)	4/5 (80%)	1/5 (20%)
Cranial Nerve Signs	5/5 (100%)	0/5 (0%)	0/5 (0%)
Involuntary Movements	5/5 (100%)	0/5 (0%)	0/5 (0%)
Muscle Atrophy	1/5 (20%)	0/5 (0%)	4/5 (80%)
Truncal Sway	1/5 (20%)	2/5 (40%)	2/5 (40%)
Lameness	1/5 (20%)	3/5 (60%)	1/5 (20%)
Asymmetry	1/5 (20%)	1/5 (20%)	3/5 (60%)
Ataxia	1/5 (20%)	3/5 (60%)	1/5 (20%)
Weakness	1/5 (20%)	1/5 (20%)	3/5 (60%)
Spasticity	0/5 (0%)	1/5 (20%)	4/5 (80%)
(Hypermetria)			
Severity of Signs	0/5 (0%)	1/5 (20%)	4/5 (80%)

Eight horses were evaluated by examiners 2 and 3. There was excellent agreement ($K \geq 0.75$) between these examiners for the presence or absence of muscle atrophy, truncal sway, and asymmetry of neurologic signs. There was no agreement for spasticity. There was good agreement ($K \geq 0.4$ and $K < 0.75$) for the presence or absence of neurologic signs and ataxia, but poor agreement ($K < 0.4$) for assessment of lameness, weakness, and severity of the neurologic signs.

Ten horses were evaluated by examiners 2 and 4. There was no agreement between these clinicians for muscle atrophy, truncal sway, and severity of neurologic signs. There was excellent agreement ($K \geq 0.75$) for the presence or absence of lameness. Poor agreement was found for the presence or absence of neurologic signs, asymmetry of neurologic signs, ataxia, weakness, and spasticity.

Examinations sometimes occurred several hours apart, and in some cases on different days. Reports by technical staff and students who were present during both examinations suggested that some horses were noticeably different from 1 examination to the next.

Discussion

The agreement among observers was either good or excellent for 80% of horses when evaluating for the presence or absence of neurologic abnormalities. This information demonstrates that, despite the subjectivity of the neurologic examination, a reasonable level of agreement among observers may be achieved. Agreement was best among observers when examiners were assessing the presence or absence of neurologic signs, lameness, ataxia, and truncal sway, and agreement was worst when assessing muscle atrophy, spasticity, and severity of neurologic deficits.

Despite the reasonable level of agreement seen in this study, these results also show that there can be considerable interobserver variability in the recognition of clinical signs of neurologic disease in the horse. This result may have been affected by the population of horses used in our study, of which, over 75% of the horses were considered normal or mildly affected. Not surprisingly, most of the disagreement among clinicians (>80%) occurred regarding horses in those 2 categories. There was little disagreement when examining severely or moderately affected horses. According to some authors, a kappa value of 0.5–0.6 would be expected regarding interobserver agreement among experienced clinicians when attempting to diagnose conditions that are moderately difficult to identify.⁵ Diagnosing neurologic disease in mildly affected horses would meet the definition of a moderately difficult diagnostic task. There were 55 possible opportunities for measuring agreement with 11 categories regarding clinical signs and 5 pairs of observers. In this investigation, agreement in 26 of 55 measurements had kappa values ≥ 0.5 . The highest kappa statistics were observed between examiners 1 and 2, based on examination of 54 horses. Good or excellent agreement was observed in all but 2 of the 11 categories. The other comparisons between

pairs of observers were only based on examination of 8 to 14 horses. If these examiners had examined more horses, the overall kappa statistics may have improved. Clinical examinations are inherently subjective and historical information can bias the examiner. We attempted to remove this bias by use of a blinded observer. However, we may have introduced another source of disagreement because both examiners did not have the same information regarding the horse involved. In this context, the estimates of agreement may have been conservative.

Perhaps the most important category evaluated by clinicians was the general assessment of the presence or absence of neurologic deficits. In our study, assessment of agreement of examiners in this category was good for 4 of 5 pairs of examiners. Considering that >75% of the horses were normal or mildly affected, this is an acceptable agreement level. Lameness and ataxia also had good agreement with 3 of 5 pairs of examiners.

Little information is available with regard to interobserver variation in clinical diagnoses of other conditions in veterinary medicine.⁵ One of the few published studies of interobserver variation in veterinary medicine involved abdominal auscultation in the horse.⁶ There was good intraobserver agreement ($K = 0.57$), but poor interobserver agreement ($K = 0.37$) in the assessment of abdominal sounds in these horses.¹³ However, even with poor agreement between observers, the authors concluded that the level of agreement documented was significant.¹³ Similar studies have been performed when examining horses for lameness. In 1 study, the level of agreement was only 66%, leaving observer error of 34% using the verbal rating scale (VRS) score whereas using the numerical rating scale (NRS) it was 72%, leaving observer error of 28%.¹³ In another study, the clinicians agreed 61.9% ($K = 0.23$) of the time when lameness score was less noticeable (score <1.5/5), whereas, the clinicians agreed 93.1% ($K = 0.86$) of the time when the lameness was more apparent (score >1.5/5).⁸ Such a finding is not unusual, because the same issues arise when horses with mildly neurologic signs are examined. For example, in another study, the intra-assessor scores were highly repeatable, but when comparing ≥ 2 examiners the results were disappointing ($K = 0.41$), when utilizing the scores of 1 assessor, however, the results were much better ($K = 0.58$ and $K = 0.78$).⁹

Neurologic signs may fluctuate between observation periods, particularly when signs are caused by infectious diseases.¹⁴ Neurologic signs may have varied from examination to examination, which may have contributed to disagreement among examiners. There are other reasons for poor agreement such as lack of a standardized diagnostic evaluation or a difference in the knowledge base among observers based on experience with different breeds of horses.⁵ However, the clinicians participating in the investigation had 18 to 26 years of experience in equine internal medicine and had worked together for a minimum of 5 years. A standardized clinical examination was used by all clinicians and a standardized form was used for recording clinical observations.

It has been suggested that some of the errors that frequently occur in neurologic examinations include erroneous interpretation of motor function, overlooking muscle fasciculations and muscle atrophy, missing subtle changes in gait, and frequent errors in sensory testing as a result of the examiner's technique.¹⁵ This suggestion is consistent with the findings of our study. Although there was 100% agreement in the diagnosis of involuntary movements, there was poor agreement in assessment of muscle atrophy and spasticity.

Another important reason for errors in the emergency room is a lack of attention to detail rather than lack of knowledge,¹⁶ and it is possible that this lack of attention to detail contributed to our observed lack of agreement because our study was conducted during a period of a heavy caseload.

Our study suggests that agreement among observers was good, considering the level of difficulty in examining this population of horses. This information is important to the veterinary profession and may provide a baseline for future research. These results attest to the reproducibility and value of clinical examinations when performed by competent, experienced equine internists.

Footnotes

^a SnSAG2 and SnSAG4/3 ELISAs were performed at Equine Diagnostic Solutions, LLC, Lexington, KY

^b Corel Paradox 8; Corel Corp., Ottawa, Ontario, Canada

^c Microsoft Excel 97, Seattle, WA

Acknowledgments

All work for this study was conducted at the Veterinary Teaching Hospital at The Ohio State University in Columbus, Ohio. The authors thank Duncan Alexander of Geneva, Illinois, for his donation to support this work.

Conflict of Interest Declaration: Kenneth W. Hinchcliff serves as Co-Editor in Chief for the Journal of Veterinary Internal Medicine. He was not involved in review of this manuscript.

Off-label Antimicrobial Declaration: Authors declare no off-label use of antimicrobials.

References

1. Granstrom DE. Equine Protozoal Myeloencephalitis: Parasite biology, experimental disease, and laboratory diagnosis. International Equine Neurology Conference 1997;4-6.
2. Morley PS, Saville WJA. Equine Protozoal Myeloencephalitis: What does a positive test mean? Proc Am Assoc Equine Pract 1997;1-5.
3. Duarte PC, Daft BM, Conrad PA, et al. Evaluation and comparison of an indirect fluorescent antibody test for detection of antibodies to Sarcocystis neurona, using serum and cerebrospinal fluid of naturally and experimentally infected, and vaccinated horses. J Parasitol 2004;90:379-386.

4. Reed SM, Howe DK, Morrow JK, et al. Accurate ante-mortem diagnosis of equine protozoal myeloencephalitis (EPM) based on detecting intrathecal antibodies against *Sarcocystis neurona* using the SnSAG2 and SnSAG4/3 ELISAs. *J Vet Intern Med* 2013;27:1193–1200.
5. Martin S, Meek A, Willeberg P. *Epidemiologic Concepts. Veterinary Epidemiology*. Ames, IA: Iowa State University Press, 1987. p. 3–21.
6. Mayhew IG. Measurements of the accuracy of clinical diagnoses of equine neurologic disease. *J Vet Intern Med* 1991;5:332–334.
7. Olsen E, Dunkel B, Barker WH, et al. Rater agreement on gait assessment during neurologic examination of horses. *J Vet Intern Med* 2014;28:630–638.
8. Keegan KG, Dent EV, Wilson DA, et al. Repeatability of subjective evaluation of lameness in horses. *Equine Vet J* 2010;42:92–97.
9. Fuller CJ, Bladon BM, Driver AJ, Barr ARS. The intra- and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses. *Vet J* 2006;171:281–286.
10. Hewetson M, Christley RM, Hunt ID, Voute LC. Investigations of the reliability of observational gait analysis for the assessment of lameness in horses. *Vet Rec* 2006;158:852–858.
11. van Kranen-Mastenbroek V, van Oostenbrugge R, Palmans L, et al. Inter- and intra-observer agreement in the assessment of the quality of spontaneous movement in the newborn. *Brain Develop* 1992;14:289–293.
12. Fleiss JL. *Statistical methods for rates and proportions*, 2nd ed. New York: John Wiley and Sons; 1981.
13. Ehrhardt E. Observer variation in equine abdominal auscultation. *Equine Vet J* 1990;22:182–185.
14. Mayhew I. *Large Animal Neurology*, 1st ed. Philadelphia, PA: Lea & Febiger; 1989.
15. van Den Berge J, Schouten H, Boomstra S, et al. Interobserver agreement assessment of ocular signs in coma. *J Neurol Neurosurg Psychiatry* 1979;42:1163–1168.
16. Wiener S, Nathanson M. Frequent errors made in doing a neurologic exam. *Medical Times* 1978;106:109–111.