

Variability of resting endoscopic grading for assessment of recurrent laryngeal neuropathy in horses

J. D. PERKINS*, R. O. SALZ, J. SCHUMACHER‡, L. LIVESEY‡, R. J. PIERCY and S. Z. BARAKZAI†

Department of Veterinary Clinical Sciences, Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield, Hertfordshire AL9 7TA, UK;

‡Department of Veterinary Clinical Sciences, Auburn University, Auburn, Alabama 36849, USA; and †Department of Veterinary Clinical Studies, Faculty of Veterinary Medicine, Royal (Dick) School of Veterinary Studies, Easter Bush Veterinary Centre, Easter Bush, Roslin, Midlothian EH25 9RG, UK.

Keywords: horse; equine; larynx; recurrent laryngeal neuropathy; endoscopy

Summary

Reasons for performing study: The extent to which variability affects endoscopic grading of arytenoid cartilage movement is uncertain.

Objective: To determine the observer and within horse variability of grading arytenoid cartilage movement in horses during resting endoscopic examination, using a 7-grade system.

Methods: Endoscopic recordings of the upper respiratory tract made at rest in 270 draught horses were reviewed independently by 2 veterinarians to assess interobserver variability when scoring horses' laryngeal function with a 7-grade system. Grading was repeated by both examiners in 80 randomly selected recordings in order to assess intraobserver variability. In 120 horses, endoscopy was repeated after 24–48 h, with videos graded by both veterinarians to assess intrahorse variability.

Results: The mean weighted κ statistic for concordance within examiners was 0.867, with a mean intraobserver agreement of 76.3%. The weighted κ statistic for concordance between the 2 examiners was 0.765, with an interobserver agreement of 63.1%. Of the horses receiving 2 endoscopic examinations, the same grade was assigned to 57.1% of horses at the second examination, when effects resulting from interobserver variability were removed. The mean weighted κ statistic for concordance between the grade assigned at first vs. second examinations was 0.588, indicating only moderate agreement.

Conclusions and potential relevance: Intra- and interobserver reliability of resting endoscopic grading of arytenoid cartilage movement using a 7-grade system was high when examinations were conducted by experienced veterinarians. However, there was moderate daily intrahorse variability, suggesting that results of resting endoscopic examinations performed on a single day should be interpreted with caution.

Introduction

Recurrent laryngeal neuropathy (RLN) has been recognised as a common cause of respiratory noise and reduced performance in

horses for many years (Cole 1946; Cook 1965; Beard 1996). RLN is a peripheral neuropathy caused by axonal degeneration of the recurrent laryngeal nerves (Cahill and Goulden 1986; Duncan *et al.* 1991). Corresponding neurogenic muscle atrophy of the *cricoarytenoid dorsalis* muscle, the primary abductor of the arytenoid cartilages, results in impaired arytenoid abduction. Large breed horses are more often affected, with signs being most severe on the left side (Marks 1970; Tetens *et al.* 2001).

Endoscopic examination of the larynx is the primary method used to assess arytenoid cartilage movement (Cook 1974; Lane 1987; Archer *et al.* 1991) and accurate and consistent grading between and within individuals is important for the appropriate classification of RLN severity. Inaccurate or inconsistent classification may lead to errors with regard to clinical decision making in, for example, sales examinations, vetting procedures, disease investigation and surgical management. Similarly, consistency of the grade assigned to an individual horse at different examinations (intraobserver variability) is important. Significant inter- and intraobserver, and intrahorse variability would probably introduce errors in studies designed to assess the extent to which severity of RLN progresses with time, perhaps explaining why disease progression is unclear (Anderson *et al.* 1997; Dixon *et al.* 2002).

Numerous grading systems of laryngeal function have been devised in attempts to create a practical method that allows for repeatable and standardised interpretation, and that correlates with laryngeal function at exercise (Hackett *et al.* 1991; Lane 1993; Dixon *et al.* 2001). The 4-grade system, first described by Ducharme *et al.* (1989) and subsequently modified by Hackett *et al.* (1991) is used by most clinicians and researchers in North America. With a 4-grade system, *grade 3* includes any horse that demonstrates asynchronous movement of the left arytenoid cartilage during any phase of respiration, in which full abduction of the left arytenoid cartilage cannot be induced or maintained by nasal occlusion or swallowing. Therefore, a horse with mild but constant laryngeal asymmetry would be allocated the same grade as a horse with marked laryngeal asymmetry. Concern over the wide range of anomalous motion grouped into *grade 3* has led to the evolution of grading systems that subdivide this grade. Lane (1993) and Dixon *et al.* (2001) have described 5- and 6-grade

*Author to whom correspondence should be addressed.

[Paper received for publication 17.11.08; Accepted 29.12.08]

systems, respectively. A 7-grade, ‘Havemeyer’ system has since been developed by a panel of experts in the hope that clearly defined criteria would help minimise variability and hence result in a reliable grading system that was consistent between observers (Dixon *et al.* 2003).

Intra- and interobserver variability of a 4-grade system have been investigated (Hackett *et al.* 1991), but there is still inadequate knowledge about the variability of endoscopic grading of the resting horse using other systems. Similarly, only one study has addressed intrahorse variability, but the study was limited to relatively few horses and a 4-grade system (Ducharme *et al.* 1991). In the current study, intra- and interobserver, and intrahorse variability of endoscopic grading of arytenoid cartilage movement was examined in a large number of Clydesdale and Belgian/Percheron horses, using the Havemeyer system.

Materials and methods

Animals

Laryngeal function was assessed in 120 Belgian/Percheron and 150 Clydesdale horses using digital videoendoscopy. The Belgian/Percheron horses were selected randomly from a herd in the USA and the Clydesdale horses in Scotland. The breed, age and sex of each horse were recorded.

Procedure

Horses were restrained without sedation using a nose twitch. A flexible 10 mm diameter, 1.4 m digital videoendoscope¹ was introduced through the right nostril, advanced to the pharynx and positioned centrally so that the appearance and movements of the larynx could be observed. Laryngeal activity was examined during respiration at rest and following induction of swallowing, (induced a minimum of 3 times in each horse). Each examination lasted 1–4 min and was recorded and stored for subsequent review. Endoscopic examinations were repeated in the 120 Belgian/Percheron horses 24–48 h later.

Recordings for horses with upper respiratory tract abnormalities other than RLN, such as severe pharyngeal lymphoid hyperplasia or arytenoid chondritis, or recordings where full abduction of the right arytenoid cartilage was not achieved were discarded from the study to prevent possible identification on

repeat blinded viewings. Remaining recordings were coded, randomised and then graded independently by 2 experienced, blinded clinicians who subjectively assessed arytenoid cartilage movements using the Havemeyer 4-grade system with subgrades (Dixon *et al.* 2003, Table 1). Assessment was repeated by both clinicians at a separate occasion, on 80 randomly chosen and renumbered recordings.

Statistical analysis

Agreement of grading laryngeal function within and between observers and within horses examined on different days was summarised as frequency of agreement and range of disagreement. To assess intraobserver variability, mean percentage agreements for first vs. second grading by each observer were calculated within each grade, by calculating the mean of the percentage agreement within each grade when first grading was compared to second grading, and when second grading was compared to first grading. Mean percentage agreements within each grade for interobserver variability were calculated similarly, by calculating the mean of the percentage agreement when *observer 1* was compared to *observer 2*, and when *observer 2* was compared to *observer 1*.

For the purpose of statistical analysis, assigned grades were renumbered 1–7 and a linear increase between grades was assumed. Thereafter, linear weighted Kappa (κ) statistics, were calculated for all pairs of observations (within and between observers and between first and second examination for each horse) to account for the magnitude of relative differences between observations (Cohen 1968). Agreement was considered poor ($\kappa < 0.20$), fair ($0.21 \leq \kappa \leq 0.40$), moderate ($0.41 \leq \kappa \leq 0.60$), good ($0.61 \leq \kappa \leq 0.80$) and excellent ($\kappa > 0.80$) (Brennan and Silman 1992).

In order to determine whether there was greater variability within certain grades, the relative differences between pairs of observations within grades were compared. The distribution of variability within assigned intra- and interobserver, and intrahorse grades were not distributed normally (Kolmogorov-Smirnov test), so differences in variability within grades were compared using Kruskal-Wallis tests.

Statistical significance was accepted when $P \leq 0.05$. *Post hoc* comparisons between individual pairs of observations were examined using Mann-Whitney U tests with a Bonferroni correction ($P \leq 0.0024$). Data were analysed using a statistical computer package (MedCalc)².

TABLE 1: Grading system of laryngeal function performed in the standing unsedated horse, outlined at the Havemeyer foundation (Dixon *et al.* 2003)

Grade	Description	Subgrade	Description
I	All arytenoid cartilage movements are synchronous and symmetrical and full arytenoid cartilage abduction can be achieved and maintained.		
II	Arytenoid cartilage movements are asynchronous and/or larynx is asymmetric at times but full arytenoid cartilage abduction can be achieved and maintained.	1	Transient asynchrony, flutter or delayed movements are seen.
		2	There is asymmetry of the <i>rima glottidis</i> much of the time due to reduced mobility of the affected arytenoid and vocal fold but there are occasions, typically after swallowing or nasal occlusion, when full symmetrical abduction is achieved and maintained.
III	Arytenoid cartilage movements are asynchronous and/or asymmetric. Full arytenoid cartilage abduction cannot be achieved and maintained.	1	There is asymmetry of the <i>rima glottidis</i> much of the time due to reduced mobility of the arytenoid and vocal fold but there are occasions, typically after swallowing or nasal occlusion, when full symmetrical abduction is achieved but not maintained.
		2	Obvious arytenoid abductor deficit and arytenoid asymmetry. Full abduction is never achieved.
		3	Marked but not total arytenoid abductor deficit and asymmetry with little arytenoid movement. Full abduction is never achieved.
IV	Complete immobility of the arytenoid cartilage and vocal fold.		

TABLE 2: Percentage of endoscopic recordings in which there was perfect agreement within observers (intraobserver agreement) and between observers (interobserver agreement), for each individual grade. For intraobserver agreement, the number of recordings receiving each grade at first and second viewings of the same recordings combined are shown for *observer 1* and *observer 2*. For interobserver agreement, the numbers of recordings receiving each grade by *observer 1* and *observer 2* combined are shown

Laryngeal function grade	160 observations					466 observations	
	No. horses (<i>observer 1</i>)	No. horses (<i>observer 2</i>)	Intraobserver agreement (<i>observer 1</i>) %	Intraobserver agreement (<i>observer 2</i>) %	Mean intraobserver agreement %	No. horses	Mean interobserver agreement %
I	20	31	84.7	80.8	82.8	67	90.1
II.1	43	34	77.3	74.8	76	80	53.4
II.2	31	29	71.1	64.6	67.8	116	56.6
III.1	17	27	76.7	70.9	73.8	101	61.6
III.2	27	19	73.9	89	81.5	56	64.6
III.3	8	8	80	80	80	26	69.2
IV	14	12	100	87.5	93.8	20	70.7

Results

Sixteen recordings of laryngeal movement of the Belgian/Percheron horses and 21 recordings of the Clydesdale horses were discarded due to other upper respiratory tract abnormalities or incomplete recordings. This left a population of 233 (82 Belgian, 22 Percheron and 129 Clydesdale) horses for assessment of interobserver variability and 104 horses for assessment of intrahorse variability. Overall, there were 160 mares, 58 geldings and 15 stallions, mean age 7.5 years (range 1–24 years).

Intraobserver variability

Both observers assigned the same grade on second viewing of the same endoscopic recording (intraobserver agreement) in 76.3% of recordings. The intraobserver agreements within each of the 7 grades are shown in Table 2. In order to determine whether errors were more consistently made within certain grades by each observer (i.e. for some grades it may be more difficult to be consistent), we compared the intraobserver disagreement of videos in both observers. Of the 80 recordings graded twice by each observer, there were 50 recordings (62.5%) in which there was perfect agreement within both observers and 8 recordings (10%) in which there was disagreement by both observers. All of the 8 horses in which both observers assigned different grades at their 2 viewings were graded II.1, II.2 or III.1 at one of the viewings, by each observer. The weighted κ statistics for concordance within observers was 0.87 (s.e. = 0.07) for *observer 1* and 0.86 (s.e. = 0.07) for *observer 2*.

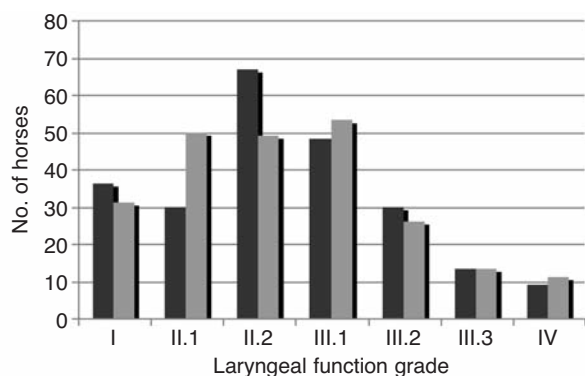


Fig 1: Bar chart showing the number of horses assigned each laryngeal function grade by each observer ($n = 233$ for each observer). ■ = Observer 1; ▒ = Observer 2.

Interobserver variability

The number of horses assigned each laryngeal grade by *observers 1* and *2* is shown in Figure 1. Both observers assigned the same grade to 63.1% of horses. The observers differed by 1 grade in 32.2% of horses and by 2 grades in 4.7% of horses. The weighted κ statistic for concordance between observers was 0.76 (s.e. = 0.04). The interobserver agreements within each of the 7 grades are shown in Table 2. The 2 observers were found to disagree significantly more often over a grade II.1, when compared to any other grade ($P < 0.0024$).

Intrahorse variability

Of the 104 horses that received 2 endoscopic examinations, fewer than half (41.7%) were graded the same at the second examination (mean percentage agreement of the 2 observers); 42.3% of horses were given a grade that differed by 1, with 16.6% receiving an improved grade and 25.7% receiving a worse grade; 13.2% of horses were given a grade that differed by 2, with 4.9% receiving an improved grade and 8.4% receiving a worse grade; 2.3% of horses were given a grade that differed by 3, with 1.3% receiving an improved grade and 1.1% receiving a worse grade. One horse improved by 4 grades. The weighted κ statistics for concordance between grade assigned at examinations 1 and 2 were 0.59 (s.e. = 0.06) for *observer 1* and 0.58 (s.e. = 0.07) for *observer 2*.

To eliminate the effect of interobserver variability, subsequent analysis was performed in 35 horses where there was perfect grading agreement between observers at each of the first and second examinations. Of these horses, 57.1% were assigned the same grade at the second examination, 34.3% were assigned a grade that differed by 1 and 8.6% were assigned a grade that differed by 2. The weighted κ statistic for concordance between grades assigned at examinations 1 and 2 was 0.74 (s.e. = 0.11). The percentages of horses that were assigned the same grade, a grade that differed by 1 and by 2, within each grade are shown in Figure 2. No particular grade was found to change significantly more often than any other grade ($P > 0.0024$).

Discussion

At yearling sales and some prepurchase examinations, endoscopic examinations are performed to confirm 'respiratory soundness' (Haynes 1992); indeed, evidence supports an association between yearlings' resting endoscopic grades and their subsequent racing

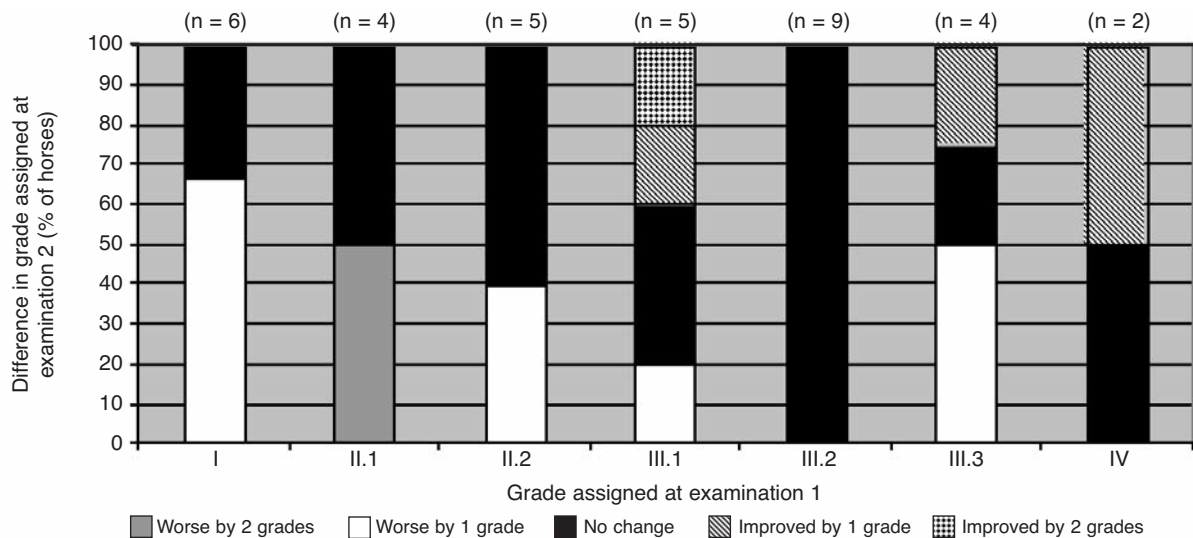


Fig 2: Bar chart showing how the laryngeal function grade changed between examinations 1 and 2 in horses that received 2 endoscopic examinations within 24–48 h. Only the 35 horses in which there was perfect agreement between observers are included.

performance (Stick *et al.* 2001). The extent of variability of endoscopic grading of RLN therefore has significant clinical implications. In the current study intraobserver agreement was excellent and interobserver agreement was good. Interestingly, however, only moderate agreement was found in the scores assigned to horses that were examined on 2 occasions within 48 h.

Intraobserver results compare favourably with published data based on the 4-grade system, in which mean κ statistics within each grade were 0.65 (for *grade 2*) to 0.98 (for *grade 4*) and intraobserver agreement was 83.3% (Hackett *et al.* 1991). Although the percentage agreement using the 7-grade system was lower than the agreement in Hackett *et al.* (1991), with a 4-grade system, there is an increased likelihood that observers agree by chance when compared to a 7-grade system (Cohen 1968). While κ statistics take account of chance agreement, percentage agreements do not, so percentage agreement must be interpreted with care when comparing systems with different numbers of grades.

Similarly, interobserver results in the present study compare favourably with those of a 4-grade system, where κ statistics for concordance between 3 observers within each grade ranged from 0.51 (for *grade 2*) to 0.90 (for *grade 4*) (Hackett *et al.* 1991). Interobserver variability has also been assessed for the 6-grade system in 35 Clydesdale horses, where grades assigned by 2 independent observers were not found to differ significantly (Hawe *et al.* 2001). In the present study, both observers assigned the same grade in 63.1% of horses, which is similar to findings with 4-level grading, where all observers assigned the same grade in 64% of horses (Hackett *et al.* 1991). Hackett *et al.* (1991) found 2 of the 3 observers assigned the same grade in an additional 36% of horses. With 3 observers, perfect agreement between observers is less likely, but with fewer grades, there is increased likelihood that the observers will agree by chance. Further studies are required that directly compare different grading systems.

Intraobserver agreement was lowest within grades II.1, II.2 and III.1 and interobserver agreement was lowest for grade II.1. This suggests that the most difficult decision faced by observers was whether full arytenoid cartilage abduction was maintained once it was achieved. Hackett *et al.* (1991) found most disagreements over grading were due to fluctuations between a

grade 1 and *2* and therefore concluded that most disagreement existed over the degree of synchrony rather than the ability to reach full abduction. Deciding whether full arytenoid cartilage abduction is maintained once it has been achieved is a purely subjective assessment and may be a limitation of the 7-grade system.

Baker (1982) proposed that in most horses, endoscopic examination of laryngeal function at rest remains unchanged throughout life. However, Dixon *et al.* (2002) suggested that RLN becomes more severe in 15% of severely affected horses over a median of 12 months. In contrast, Anderson *et al.* (1997) reported that laryngeal function grade improved in 29% of horses and worsened in 28% of horses when endoscopy was performed at a 16 month interval. Proposed explanations for the change in endoscopic grades within horses included observer variability and variations in the cycle of nerve degeneration and subsequent reinnervation of muscle fibres from surviving motor neurones (Anderson *et al.* 1997). However, a plausible alternative explanation is that the endoscopic appearance of laryngeal function in any particular horse may not be consistent on a day-to-day basis. In the present study, when horses received 2 endoscopic examinations within 24–48 h, after removing effects resulting from interobserver variability, only 57.1% of horses were assigned the same grade at both examinations. Changes in grades occurred in both directions with similar frequency, suggesting these changes were not simply due to the horses having adapted to the endoscopic procedure. It is conceivable that this intrahorse variability may be influenced by environmental effects, such as stress or anxiety and physiological effects, such as relative muscle fatigue. It is therefore suggested that single resting endoscopic examinations performed at the time of sale should be interpreted with some caution. Further studies are required to determine whether endoscopic examinations performed during exercise are similarly subject to inter- and intraobserver, and intrahorse variability.

To our knowledge there is only one previous study that has investigated intrahorse variation of endoscopic grading over a short period, (20 horses were examined twice within one week: Ducharme *et al.* 1991). In this previous study, using the 4-grade system, 21% of horses were assigned different grades at each examination, when endoscopy was performed via the right nostril (Ducharme *et al.* 1991). It is unsurprising that intrahorse

variability is lower for the 4-grade than the 7-grade system, since each grade is considerably broader in the former method.

Moderate intrahorse variability probably has significant implications for the reliability of progression studies: changes observed over a 12–16 month period is hard to distinguish from the changes that could be observed just because of day-to-day variation. This conclusion therefore has significant implications both clinically and for RLN research.

In conclusion, the results of this study show that the 7-grade system proposed at the Havemeyer foundation has excellent intra- and good interobserver reliability, probably because the grades and subgrades are clearly defined. However, the marked variability found to exist between grades assigned to horses examined twice in a 24–48 h period suggest that caution should be exercised when reaching important decisions on the basis of a single resting endoscopic examination. Work in the future should be directed at determining whether there is similar day-to-day variation in laryngeal function in horses at exercise and investigating more objective ways to assess severity of RLN.

Acknowledgement

Funding for endoscopic examination of Clydesdale horses was provided by the RCVS Trust.

Manufacturers' addresses

¹EV Veterinary Products Ltd., Brockhurst, Shropshire, UK.

²MedCalc Software, Mariakerke, Belgium.

References

- Anderson, B.H., Kannegieter, N.J. and Goulden, B.E. (1997) Endoscopic observations on laryngeal symmetry and movements in young racing horses. *N. Z. vet. J.* **45**, 188-192.
- Archer, R.M., Lindsay, W.A. and Duncan, I.D. (1991) A comparison of techniques to enhance the evaluation of equine laryngeal function. *Equine vet. J.* **23**, 104-107.
- Baker, G.J. (1982) Laryngeal asynchrony in the horse: Definition and significance. In: *Proceedings of the First International Conference on Equine Exercise Physiology*, Eds: D.H. Snow, S.G.B. Persson and R.J. Rose, Granta Publications, Cambridge. pp 46-50.
- Beard, W. (1996) Upper respiratory causes of exercise intolerance. *Vet. Clin. N. Am.: Equine Pract.* **12**, 435-455.
- Brennan, P. and Silman, A. (1992) Statistical methods for assessing observer variability in clinical measures. *Br. med. J.* **304**, 1491-1494.
- Cahill, J.I. and Goulden, B.E. (1986) Equine laryngeal hemiplegia. I. A light microscopic study of peripheral nerves. *N. Z. vet. J.* **34**, 161-169.
- Cohen, J. (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol. Bull.* **70**, 213-220.
- Cole, C.R. (1946) Changes in the equine larynx associated with laryngeal hemiplegia. *Am. J. vet. Res.* **7**, 69-77.
- Cook, W.R. (1965) The diagnosis of respiratory unsoundness in the horse. *Vet. Rec.* **77**, 516-521.
- Cook, W.R. (1974) Some observations on diseases of the ear, nose and throat in the horse, and endoscopy using a flexible fiberoptic endoscope. *Vet. Rec.* **94**, 533-541.
- Dixon, P.M., Robinson, E. and Wade, J.F. (2003) Havemeyer Foundation - Workshop Summary. In: *Proceedings of a Workshop on Equine Recurrent Laryngeal Neuropathy*, R & W Publications, Newmarket. pp 93-97.
- Dixon, P.M., McGorum, B.C., Raitlon, D.I., Hawe, C., Tremaine, W.H., Pickles, K. and McCann, J. (2001) Laryngeal paralysis: a study of 375 cases in a mixed-breed population of horses. *Equine vet. J.* **33**, 452-458.
- Dixon, P.M., McGorum, B.C., Raitlon, D.I., Hawe, C., Tremaine, W.H., Pickles, K. and McCann, J. (2002) Clinical and endoscopic evidence of progression in 152 cases of equine recurrent laryngeal neuropathy (RLN). *Equine vet. J.* **34**, 29-34.
- Ducharme, N.G., Hackett, R.P., Fubini, S.L. and Erb, H.N. (1989) Influence of sedation, side of endoscopy, and time of evaluation on subjective and objective evaluation of endoscopic examination in assessment of laryngeal function in racehorses. *Vet. Surg.* **18**, 63-64.
- Ducharme, N.G., Hackett, R.P., Fubini, S.L. and Erb, H.N. (1991) The reliability of endoscopic examination in assessment of arytenoid cartilage movement in horses. II. Influence of side of examination, reexamination, and sedation. *Vet. Surg.* **20**, 180-184.
- Duncan, I.D., Reifenrath, P., Jackson, K.F. and Clayton, M. (1991) Preferential denervation of the adductor muscles of the equine larynx. II: nerve pathology. *Equine vet. J.* **23**, 99-103.
- Hackett, R.P., Ducharme, N.G., Fubini, S.L. and Erb, H.N. (1991) The reliability of endoscopic examination in assessment of arytenoid cartilage movement in horses. I. Subjective and objective laryngeal evaluation. *Vet. Surg.* **20**, 174-179.
- Hawe, C., Dixon, P.M. and Mayhew, I.G. (2001) A study of an electrodiagnostic technique for the evaluation of equine recurrent laryngeal neuropathy. *Equine vet. J.* **33**, 459-465.
- Haynes, P.F. (1992) Examination of the upper and lower respiratory tract relevant to purchase. *Vet. Clin. N. Am.: Equine Pract.* **8**, 347-364.
- Lane, J.G. (1987) Fiberoptic endoscopy of the equine upper respiratory tract: A commentary on progress. *Equine vet. J.* **19**, 495-499.
- Lane, J.G. (1993) Recurrent laryngeal neuropathy (RLN): current attitudes to aetiology, diagnosis and treatment. In: *Proceedings of the 15th Bain-Fallon Memorial Lectures*, Artamon, Australia. pp 173-192.
- Marks, D., Mackay-Smith, M.P., Cushing, L.S. and Leslie, J.A. (1970) Observations on laryngeal hemiplegia in the horse and treatment by abductor muscle prosthesis. *Equine vet. J.* **2**, 159-167.
- Stick, J.A., Peloso, J.G., Morehead, J.P., Lloyd, J., Eberhart, S., Padungtod, P. and Derksen, F.J. (2001) Endoscopic assessment of airway function as a predictor of racing performance in Thoroughbred yearlings: 427 cases (1997-2000). *J. Am. vet. med. Ass.* **219**, 962-967.
- Tetens, J., Hillmann, D.J. and Derksen, F.J. (2001) Idiopathic laryngeal hemiplegia. *Comp. cont. Educ. pract. Vet.* **23**, 85-94.

Author contributions The initiation, conception and planning for this study were by J.P., R.J.P. and S.B. Its execution was by J.P., R.O.S., J.S., L.L. and S.B., with statistics by R.O.S. and R.J.P. The paper was written by J.P., R.O.S., R.J.P. and S.B.