The Veterinary Journal xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

The Veterinary Journal

journal homepage: www.elsevier.com/locate/tvjl

Comparison of subjective scoring systems used to evaluate equine laminitis

I. Viñuela-Fernández [a,*], E. Jones [b], M.E. Chase-Topping [c], J. Price [d]

[a] Royal (Dick) School of Veterinary Studies, Division of Veterinary Biomedical Sciences, University of Edinburgh, Summerhall, Edinburgh EH9 1QH, UK
[b] Royal (Dick) School of Veterinary Studies, Division of Veterinary Clinical Studies, University of Edinburgh, Easter Bush Veterinary Centre, Easter Bush, Roslin EH25 9R, UK
[c] Centre for Infectious Diseases, Ashworth Laboratories, Kings Buildings, University of Edinburgh, Edinburgh EH9 3JY, UK
[d] Saint George's University Veterinary School, Saint George's, Grenada, West Indies

ARTICLE INFO

Article history:
Accepted 7 May 2010
Available online xxxx

Keywords:
Clinical grading system
Horses
Obel score
Visual analogue scale
Generalizability theory
Lameness

ABSTRACT

This study compared three subjective scoring systems used to assess lameness associated with equine laminitis: (1) visual analogue scale, (2) Obel score and (3) clinical grading system (CGS). Two groups of 12 observers, consisting of equine veterinarians and final-year veterinary students, scored lameness severity after watching video footage of 14 horses on two occasions. Generalizability theory was used to investigate the reliability of the three systems and the effects of observer experience.

Overall reliability across all times and observers was high. Intra-observer reliability was higher than inter-observer reliability for all scoring systems, with student reliability being consistently lower than veterinarians, especially for Obel and CGS. All three methods were reasonably reliable tools for assessing lameness, but they were more limited in the hands of inexperienced observers.

Introduction

Accurate evaluation of pain is of fundamental importance to equine welfare and veterinary management of the equine patient. The horse is particularly prone to musculoskeletal injury and disease due to its conformation, anatomy and different uses, with pain typically manifested as lameness (Schatzmann, 2000).

Equine laminitis is a common, potentially devastating disease, which is a major welfare concern. Hinckley and Henderson (1996) identified a prevalence of 7.1% in the UK. The condition is characterised by an acute onset of lameness of variable severity. Pain management in horses with laminitis is crucial since, although the majority will recover, some horses have chronic severe foot pain that may lead to recumbency and eventually require euthanasia on welfare grounds (Swanson, 1999; Pollitt, 2003). Equines suffering from laminitis are often refractory to standard (anti-inflammatory) analgesic therapy (Herthel and Hood, 1999).

Accurate assessment of the efficacy of analgesia requires reliable tools with which to measure pain. In clinical practice, subjective methods of assessment are commonly used to evaluate equine laminitis, often supplemented by use of hoof testers to evaluate solar sensitivity. The Obel score (Obel, 1948) and the clinical grading system (CGS) (Taylor et al., 2002) are simple descriptive scales (SDS) developed to evaluate equine lameness. The Obel was specifically developed to assess laminitis-associated lameness. SDS consists of a number of expressions used to describe pain intensity, such as 'no pain' 'mild', 'moderate' and 'severe'. Each descriptor is assigned an index value (generally 0–5), which becomes the pain score for the animal. Although Obel and CGS are extensively used, their performance has not been thoroughly validated and they are associated with significant inter-observer variability and poor agreement with objective assessments (Silver et al., 1983; Peloso et al., 1993; Keegan et al., 1998; Peham et al., 1999; Hood et al., 2001; Taylor et al., 2002).

A visual analogue scale (VAS) is a line, the extremes of which represent the limits of the pain experience, defined as 'no pain' and 'worst possible pain'. A vertical line is placed by the observer at a point corresponding to the considered severity of the pain assessed. VAS has been used to assess lameness in dogs (Hielm-Björkman et al., 2003) and sheep (Welsh et al., 1993), but it has not been validated in equines.

This study aimed to evaluate the performance (with regard to intra- and inter-observer reliability) of the three scoring systems for the assessment of laminitis-associated lameness. A second aim was to compare the performance of each scoring system when used by equine clinicians and final-year veterinary students, in order to evaluate the role of experience in lameness assessment.

Materials and methods

Video clips were edited from footage of 12 horses admitted to the Royal (Dick) School of Veterinary Studies (R(D)SVS) for management of clinical laminitis (diagnosed using gait evaluation, clinical examination, sensitivity to hoof testers and radiography). Video clips were also made of two control horses deemed sound on clinical exam by an experienced equine orthopaedic surgeon. The duration of each video clip was 60 s, consisting of 10 s where the horse was stationary and 50 s

* Corresponding author. Tel.: +44 131 242 94 94; fax: +44 131 242 79 87.
E-mail address: I.Vinuela-Fernandez@sms.ed.ac.uk (I. Viñuela-Fernández).

1090-0233/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tvjl.2010.05.011

Table 1
Modified Obel score (adapted from Owens et al. (1995)).

Grade	Definition
0	No gait abnormalities at a walk or trot
1	At rest, the horse exhibited foot lifting. The horse exhibited a normal gait at a walk. The trot showed a shortened stride and showed even head and neck lifting for each foot
2	The walk was stilted, but showed no abnormal head or neck lifting. The trot showed obvious lameness with uneven head and neck lifting. A forefoot could be lifted off the ground easily
3	The lameness was obvious at a walk and trot. The horse resisted attempts to have a forefoot lifted and was reluctant to move
4	The horse experienced difficulty bearing weight at rest or was very reluctant to move

Table 2
Clinical grading system (Taylor et al., 2002).

Grade	Definition
0	Horse is capable of full athletic function
1	Horse is capable of a minimum amount of pleasure riding but not full athletic function
2	Horse cannot be ridden but is suitable for breeding; horse can be maintained without lameness on pasture with minimal use of systemically administered analgesics
3	Horse must be maintained on systemically administered analgesics to function
4	Horse must be euthanased because of severe pain that does not respond to systemically administered analgesics

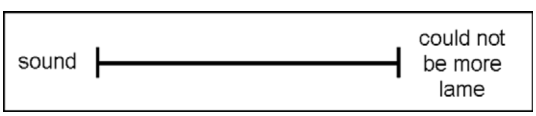

Fig. 1. Visual analogue scale adapted for lameness assessment (Welsh et al., 1993). VAS consisted of a 10 cm long horizontal line, where the extremes are defined as 'sound' and 'could not be more lame'.

where the horse was led at the walk, turned and returned to the observer. The video clips were combined randomly into a single observation session of 28 video clips. Each video clip was presented twice for assessment. A 30 s interval between video clips allowed time for scoring.

A group of 12 final-year veterinary science students and a group of 12 experienced equine practitioners (senior clinicians and postgraduates with at least 3 years experience) from the R(D)SVS watched the video compilation in separate sessions. Observers were placed two seats apart to ensure independence of scoring and were provided with written descriptions of the Obel and CGS (Tables 1 and 2). After watching each video clip, observers scored lameness severity on an individual scoring sheet that had the three scoring systems printed in random order. Obel and CGS data were discrete scores (0–4) and VAS scores were expressed as the distance (mm) from the left extreme to the vertical mark (Fig. 1).

Data analysis

Descriptive statistics (mean, standard error, median) were used to characterise the scores from each scoring system.

The heterogeneity of the scores from each scoring system were further analysed by the Generalizability theory (*G*-theory) (Brennan, 2001; Streiner and Norman, 2008). *G*-theory is a statistical theory for evaluating the dependability ('reliability') of behavioural measurements, which acknowledges multiple sources of error variance (Streiner and Norman, 2008). Based on the analysis of variance, *G*-theory is divided into two parts: (1) the Generalizability study (*G*-study), and (2) the Decision study (*D*-study). The *G*-study estimates the magnitude of variation associated with all effects. These variance estimates are then used to determine reliability or Generalizability coefficients for the various facets (or factors). The *D*-study uses the information of the *G*-study to determine the reliability of a particular protocol, by means of investigating the effects of altering specified aspects of the study.

G-study

A fully crossed, two-facet mixed effects analysis of variance (ANOVA) was calculated for each of the three scoring systems. From the results of the ANOVAs, variance components were estimated for each main effect (horse (*H*), observer (*O*), and time (*T*)), their interactions and error terms (Streiner and Norman, 2008). In this study, seven sources of error variance, *H*, *O*, *T*, *HT*, *HO*, *OT* and *HOT* (which includes the pure error) were identified. Variance components estimates were expressed as percentages of total variance. The variance components were used to calculate coefficients corresponding to inter-observer reliability, intra-observer (test–re-test) reliability and overall reliability (Streiner and Norman, 2008), using the following equations:

Inter-observer reliability:

$$R_1 = \frac{\sigma_H^2 + \sigma_{HT}^2}{\sigma_H^2 + \sigma_O^2 + \sigma_{HT}^2 + \sigma_{HO}^2 + \sigma_{OT}^2 + \sigma_{HOT,e}^2}$$

where σ^2 = variance.

Intra-observer or test–re-test reliability:

$$R_2 = \frac{\sigma_H^2 + \sigma_O^2 + \sigma_{HO}^2}{\sigma_H^2 + \sigma_O^2 + \sigma_{HT}^2 + \sigma_{HO}^2 + \sigma_{OT}^2 + \sigma_{HOT,e}^2}$$

Overall reliability:

$$R_3 = \frac{\sigma_H^2}{\sigma_H^2 + \sigma_O^2/n_O + \sigma_{HT}^2/n_T + \sigma_{HO}^2/n_O + \sigma_{OT}^2/n_{O\times T} + \sigma_{HOT,e}^2/n_{O\times T}}$$

D-study

Follow-up *D*-studies were conducted to explore the effect of altering the number of observers and observation times, for students and veterinarians separately. Data were presented as the relative Generalizability (*G*) coefficient, which represents the relative amount of variation associated with a given facet and its interactions. *G*-coefficients range from 0 (null reliability) to 1 (perfect reliability), with values above 0.75 considered as good reliability, those between 0.50 and 0.75 as moderate and those under 0.5 as poor (Portney and Watkins, 2000).

Descriptive statistics and variance component analysis were conducted using SAS version 9.3.1 (SAS Institute). For the inter- and intra-reliability estimates, distributions were generated using bootstrap sampling with 3000 iterations. In each iteration, data were sampled from the overall distribution with replacement. The mean, upper (97.5 percentile) and lower (2.5 percentile) confidence limits were derived from the respective bootstrap distributions. The *D*-study was done with the GENOVA program.[1]

Results

Overall performance of the scoring systems: G-study

Table 3 contains the raw scores generated from the study for the three scoring systems: VAS, CGS and Obel. Fig. 2a shows the proportion of variance attributable to the various sources derived from the ANOVA. The main source of variance in all methods is due to systematic differences between horses (*H*) (65–66% of total variance). Less variance is due to the main effects of observer (*O*) (3–6%) and time (*T*) (<1%). The interaction of horse and observer (*HO*) and the interaction of horse, observer and time and the error term (*HOT*, *e*) were relatively large (9–10% and 12–19%, respectively). The *HOT*, *e* variance, however, was lower for the VAS (12%) in comparison to the CGS (19.0%) and Obel (18.7%). As a result, the variance of other components including observer (*O*) and observer-by-time interaction (*OT*) was higher for the VAS.

Performance of the scoring systems (students versus veterinarians): G-study

Fig. 2b shows the proportion of variance attributable to the various sources derived from the ANOVA for the students and veterinarians separately. As above, the main sources of variance were attributable to systematic differences between horses (*H*) (60–73% of total variance) and the interaction of horse, observer and time and the error (*HOT*, *e*) variance (12–25%). The proportion of variance attributable to all sources was similar for veterinarians and students for the VAS scoring system. However, the proportion of variance attributable to *H* and *HOT*, *e* differed between veterinar-

[1] See: http://www.education.uiowa.edu/casma/computer_programs.htm#genova.

Please cite this article in press as: Viñuela-Fernández, I., et al. Comparison of subjective scoring systems used to evaluate equine laminitis. The Veterinary Journal (2010), doi:10.1016/j.tvjl.2010.05.011

Table 3
Summary statistics for raw values for the different measurement scores. VAS, visual analogue scale; CGS, clinical grading system; OBEL, Obel score.

Observer	Time	VAS		CGS		Obel	
		Mean (SE)	Median	Mean (SE)	Median	Mean (SE)	Median
All	1	39.7 (1.51)	35.5	2.0 (0.06)	2.0	1.8 (0.06)	2.0
	2	42.9 (1.55)	40.5	2.1 (0.06)	2.0	1.9 (0.06)	2.0
Student	1	42.1 (2.14)	37.5	2.1 (0.08)	2.0	1.9 (0.09)	2.0
	2	45.6 (2.18)	46.0	2.1 (0.08)	2.0	2.0 (0.09)	2.0
Veterinarians	1	37.2 (2.11)	33.0	1.9 (0.09)	2.0	1.8 (0.09)	2.0
	2	40.2 (2.20)	38.0	2.0 (0.08)	2.0	1.9 (0.09)	2.0

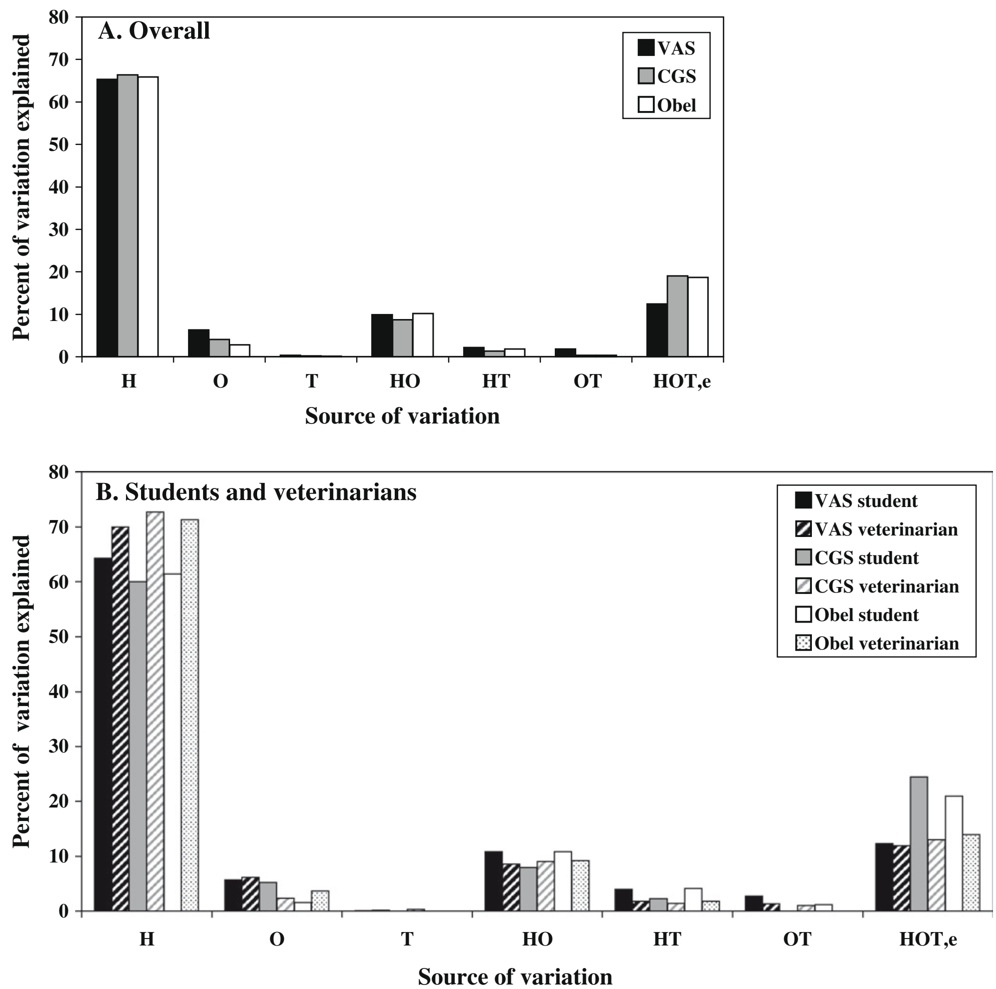

Fig. 2. Estimated variance components for the three different measurement scores (visual analogue scale (VAS), clinical grading system (CGS) and Obel score); (A) overall and (B) separately for students and veterinarians. *H*, horse; *O*, observer; *T*, time; *HT*, horse-by-time interaction; *HO*, horse-by-observer interaction; *OT*, observer-by-time interaction; *HOT*, horse-by-observer-by-time interaction plus the pure error component.

ians and students for the CGS and Obel. The student error variance was higher (24.5 and 21% for CGS and Obel, respectively) than the veterinarians (13.0 and 14.0% for CGS and Obel, respectively), whereas the veterinarian error attributed to horses (*H*) (72.7% and 71.3% for CGS and Obel, respectively) was higher than the students (60.0% and 61.4% for CGS and Obel, respectively).

Overall reliability for both veterinarians and students was high (>0.75) for the three scoring systems (Fig. 3). Inter-observer reliability was lower than intra-observer reliability for all scoring systems and experience levels. Intra- and inter-observer reliability was higher for veterinarians than students, particularly when using the SDS. Overall, confidence intervals show overlapping reliability estimates for student and veterinarians for all scoring systems and reliability coefficients (Fig. 3).

Performance of the scoring systems (students versus veterinarians) – D study

Fig. 4 shows the results of the *D*-studies. With the number of horses held constant, the change in Generalizability (*G*) coefficient for the three scoring systems is shown for students (Fig. 4A) and veterinarians (Fig. 4B) for different numbers of observers (3, 6, 9, 12, 15, 18 and 21) and observing times (1, 2 and 3). *G*-coefficients increased as the number of observers increased, although the change was much greater for the students. Increasing the number of observing times above one provided only a small increase in the *G*-coefficients for the veterinarians. However, the increase from one to two observing times for the students, greatly increased the Generalizability.

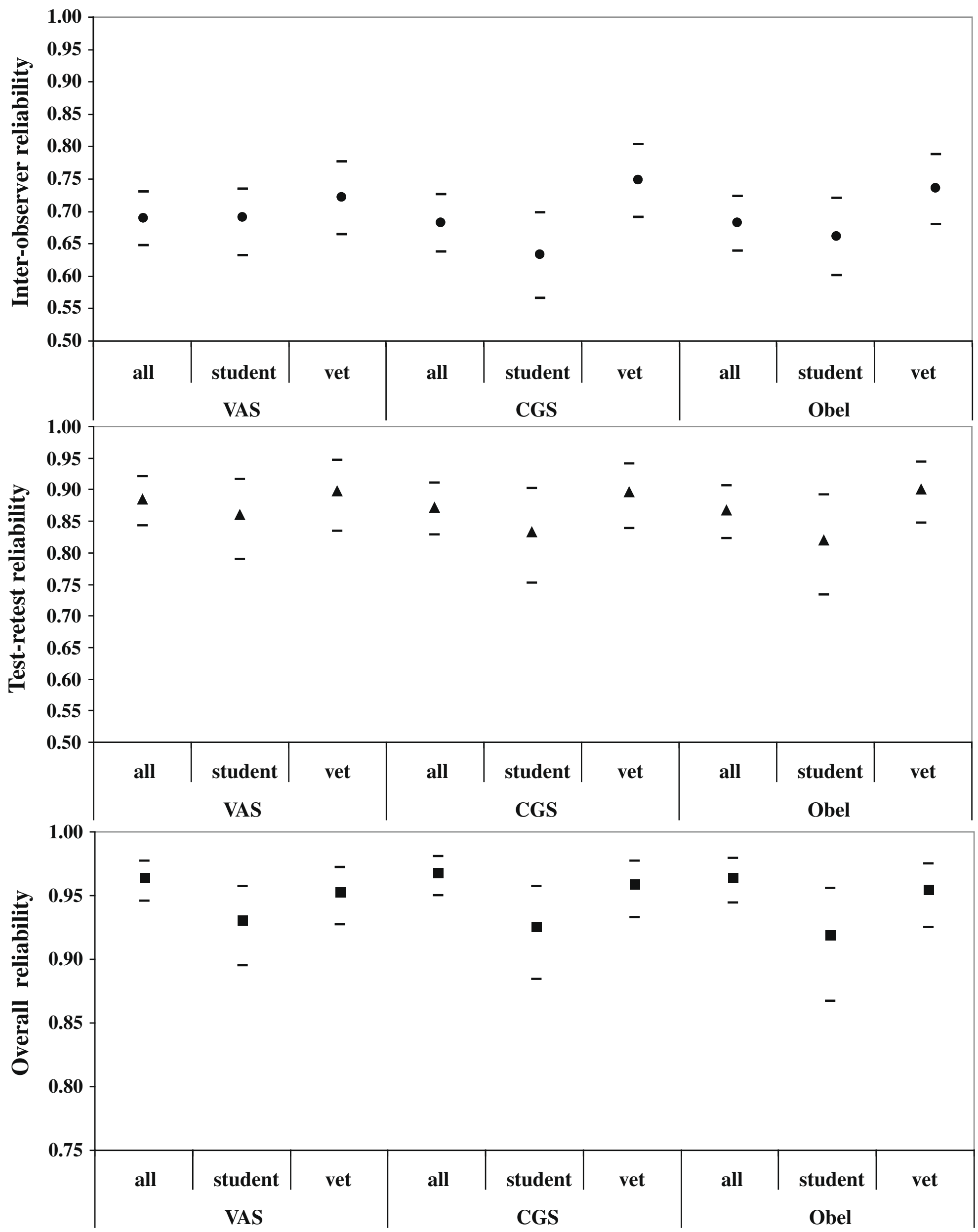

Fig. 3. Mean, upper and lower confidence limits for the inter-observer (R1), Intra-observer (test–re-test) (R2) and overall (R3) reliability coefficient for the different scoring systems (VAS, visual analogue scale; CGS, clinical grading system; Obel: Obel score) for all observers, students only and veterinarians only.

Differences in scoring tendencies of veterinarians and students

Using the VAS, students assigned higher scores (means (±SE)) than veterinarians, at 43.85 (8.01) and 38.71 (8.08), respectively. Similarly, veterinarians gave more scores of zero and fewer scores of 4 compared to students when using SDS. For all three scoring systems, veterinarians were more consistent in identifying the control horses as 'sound', compared with students.

In general, Obel scores of zero corresponded to VAS values <15 mm, and this range was slightly wider for scores of zero using the CGS. Scores of 1 using the CGS corresponded to VAS scores <40 mm, and this range was wider for an Obel of 1. Scores of 2 and 3 in both SDS methods had the widest range of VAS values, going from 10 to 70 mm (Obel) and from 20 to 30 mm (CGS) to the end of the scale. Scores of 4 in both methods corresponded to a narrower range of VAS values grouped at the right end of the scale (Fig. 5). The range of VAS scores assigned by veterinarians to each category of SDS score was smaller than the range of scores assigned by students (Fig. 5). Graphical comparison of Obel and CGS, showed that each Obel score corresponded to more than one category of CGS score and vice versa, with wider ranges in the student group (Fig. 5).

Discussion

Effective evaluation of pain management relies on sensitive and reproducible assessment tools. In order to assess the reliability of three subjective scoring systems used to assess laminitis pain and lameness (VAS, CGS and Obel), we first performed an analysis of variance. Estimated variance components showed that the largest source of variance was attributed to the horses, which is ex-

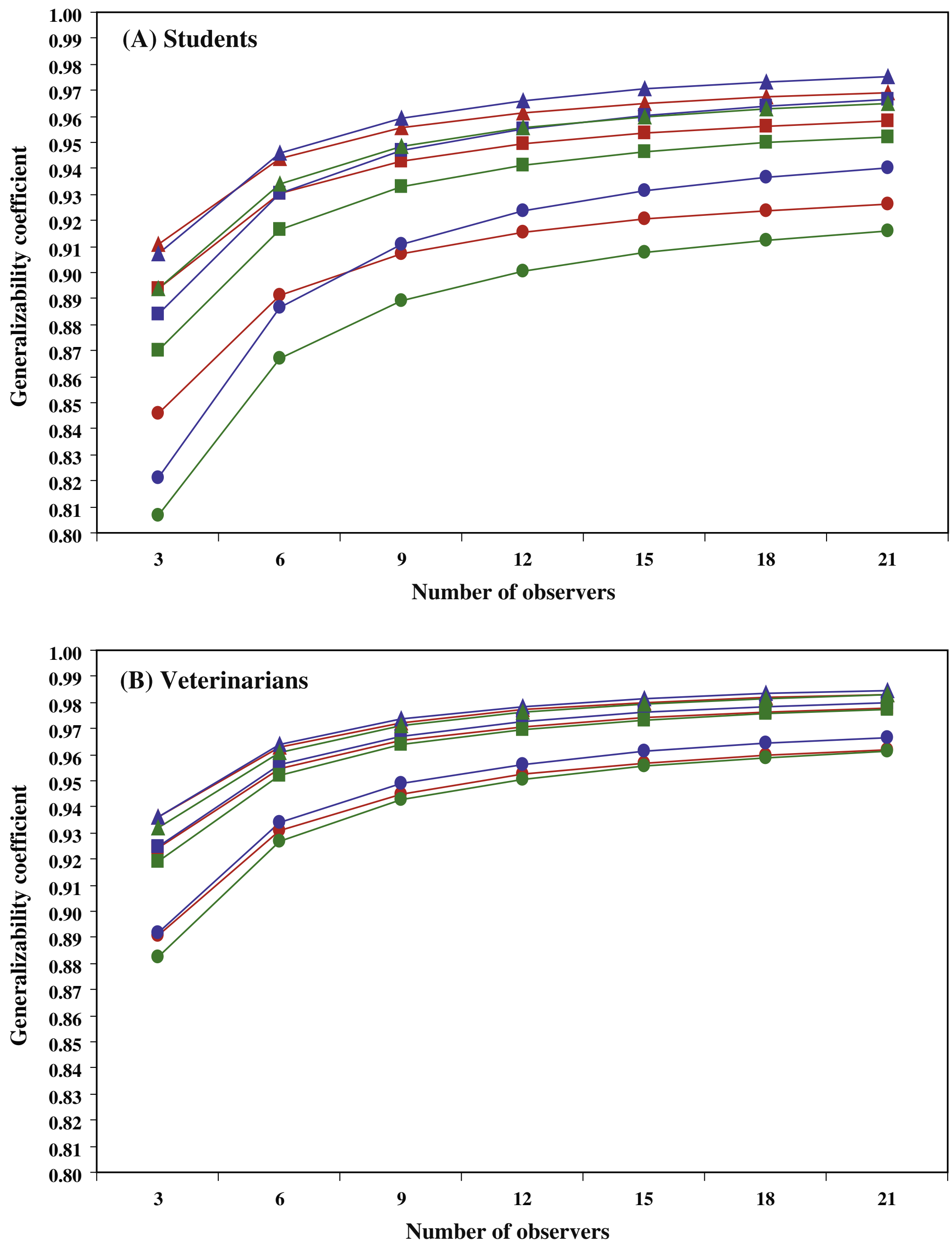

Fig. 4. Results of alternative *D*-studies for students (A) and veterinarians (B) summarised by the Generalizability coefficient for the visual analogue scale (VAS) (red), clinical grading system (CGS) (blue), and Obel score (green). Three separate observing times were considered: time 1 (circle), time 2 (square) and time 3 (triangle).

pected since horses with varying degrees of lameness (plus two sound horses) were used for this study.

The second biggest variance component was attributed to the error term (12–19% of total variance), which suggested that there might be unaccounted sources of variance. A relatively large variance component was the interaction of horse and observer (9–10% of total variance), which suggested that observer scores might be influenced by particular horses' gait features or particular degrees of lameness. The difficulty of assessing mild lameness using subjective scoring methods has been reported previously (Keegan et al., 1998). Relatively less variance was due to the main effects of observer (3–6% of total variance) and time (<1% of total variance), although the significance of these percentages might be confounded by the expected high variability of the horse component.

Mean overall, and intra-observer (test re-test) reliability coefficients were high (>0.75) for both students and veterinarians across all three scales, but students' reliability was slightly lower when using the CGS and Obel, as opposed to the VAS. There are conflicting opinions over the sensitivity inherent in the VAS pain scoring. It has been suggested that VAS is more sensitive than SDS since it measures pain on a continuous scale and observers are not forced to choose between predefined categories (Scott and Huskisson, 1976; Reid and Nolan, 1991; Lascelles et al., 1994; Robertson, 2003). Other authors have suggested that the continuous nature of VAS provides a 'false' impression of increased sensitivity (Holton et al., 1998). The consistency in VAS scores from students and veterinarians in this study suggested that VAS might be more appropriate for mixed experience groups.

Please cite this article in press as: Viñuela-Fernández, I., et al. Comparison of subjective scoring systems used to evaluate equine laminitis. The Veterinary Journal (2010), doi:10.1016/j.tvjl.2010.05.011

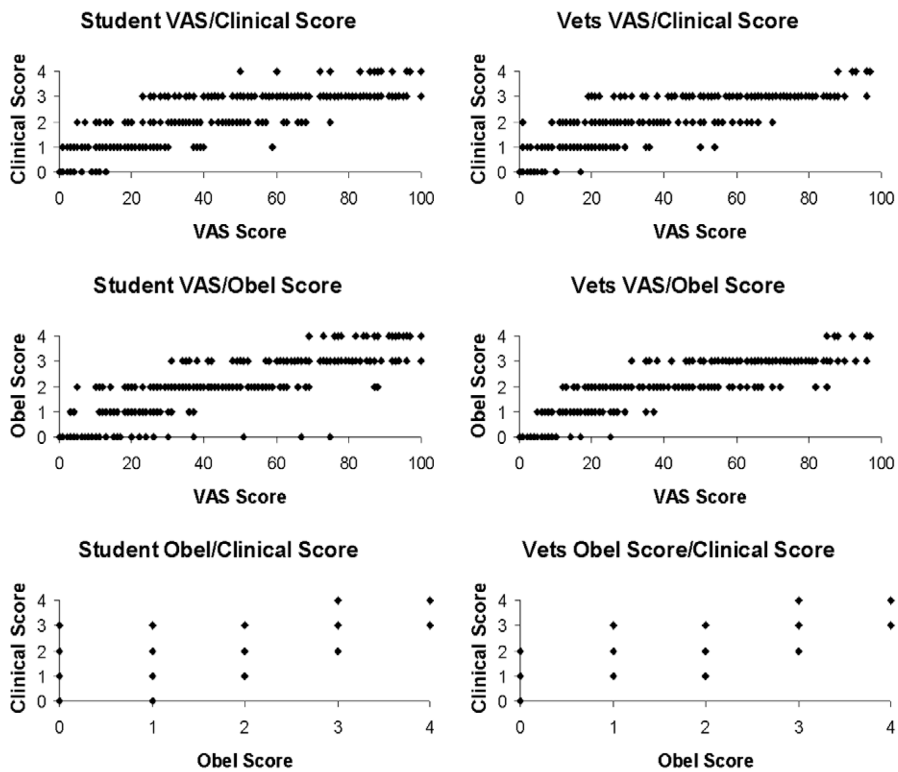

Fig. 5. Graphical representation of pairs of subjective scores plotted against one another for both final-year students (left column) and veterinarians (right column).

Close agreement between different observers using a scale is a pre-requisite of reliability (Rutherford, 2002). Inter-observer agreement has enormous implications in pain management where an animal may be re-examined by different clinicians. In our study, for all scoring systems student inter-observer reliability (mean and confidence intervals) was slightly under the 0.75 level, which was assumed to be the minimum cut-off point for good reliability, whereas veterinarians had 'good' inter-observer agreement using all scales. For the students, inter-observer agreement was lower when using CGS and Obel, in comparison to VAS.

It was surprising that SDS systems had lower intra- and inter-observer agreement than VAS when used by the students, since the provision of descriptors in both SDS systems and the fact that these scales are specifically designed to assess lameness should result in higher reliability. However, the results suggested that experience might be required to interpret the guidelines provided by the CGS and Obel. For both veterinarians and students, inter-observer coefficients were consistently smaller than intra-observer coefficients, which demonstrated that all three methods have a poorer performance when used by different observers. These results were consistent with previous studies investigating the reliability of SDS for assessment of postoperative pain in dogs (Holton et al., 1998) and lameness in horses (Keegan et al., 1998).

Differences between students and veterinarians were also identified in the categories (scores) they predominantly agreed or disagreed on. Students tended to agree predominantly on what they considered the most severe degrees of lameness, which was probably a reflection of their limited experience and of their tendency to assign higher scores than veterinarians. The lowest agreement was achieved for scores of 1 and 2, which reflected the difficulty in identifying the subtleties of mild lameness previously reported (Keegan et al., 1998, 2003). In contrast, veterinarians predominantly agreed on what they considered a 'sound horse' and disagreed mostly in scores assigned to the horses that were most severely lame. It was encouraging to observe the consensus between members of the profession with regard to identification of 'soundness'. The poorer consistency among veterinarians at the upper end of the scale could reflect perceptual differences of opinion relating to pain severity.

Our study had certain limitations. Observers made their assessments based on 60 s video clips, whereas in a clinical setting, lameness scoring would result from a more detailed evaluation. Also, certain aspects of the Obel and CGS scales could not be assessed, as the horses were not observed at the trot, since it was deemed unethical to force horses in severe pain to trot. Despite these restrictions, we considered that this study identified important limitations of these lameness scoring systems.

The use of subjective assessments alone is questionable. There is scope for incorporation of objective assessments into clinical evaluation of laminitis, including the use of calibrated hoof testers, force plates and motion capture and analysis (Owens et al., 1995; Taylor et al., 2002; McGuigan et al., 2005). Additionally, recent studies have shown that quantification of defined behaviours (i.e., weight shifting) could be used as a reliable indicator to measure pain associated with laminitis (Reitmann et al., 2004; Jones et al., 2007).

Conclusions

Although the overall reliability of the subjective scoring systems studied was high, intra-observer reliability was higher than inter-observer reliability, indicating that there are differences in the perceived degree of lameness between observers. Additionally, student reliability was consistently lower than veterinarians, especially for Obel and CGS, suggesting that these methods are more limited in the hands of inexperienced observers. Clinical evaluation of laminitis could be improved by the incorporation of objective pain assessment methods.

Conflict of interest statement

None of the authors of this paper has a financial or personal relationship with other people or organisations that could inappropriately influence or bias the content of the paper.

Acknowledgments

The authors would like to thank the veterinarians and students at the R(D)SVS who participated in this study, Joyce Kent for providing the video recordings, Dr. E. Welsh for critical appraisal of the manuscript, and World Horse Welfare for their sponsorship of the clinical position held by one of the authors.

References

Brennan, R.L., 2001. Generalizability Theory. Springer-Verlag, New York.

Herthel, D., Hood, D.M., 1999. Clinical presentation, diagnosis, and prognosis of chronic laminitis. Veterinary Clinics of North America: Equine Practice 15, 375–394.

Hielm-Björkman, A.K., Kuusela, E., Liman, A., Markkola, A., Saarto, E., Huttunen, P., Leppäluoto, J., Tulamo, R.-M., Raekallio, M., 2003. Evaluation of methods for assessment of pain associated with chronic osteoarthritis in dogs. Journal of the American Veterinary Medicine Association 222, 1552–1558.

Hinckley, K.A., Henderson, I.W., 1996. The epidemiology of equine laminitis in the UK. In: Proceedings of the 35th Congress of the British Equine Veterinary Association. Warwick, UK, p. 62.

Holton, L.L., Scott, E.M., Nolan, A.M., Reid, J., Welsh, E., Flaherty, D., 1998. Comparison of three methods used for assessment of pain in dogs. Journal of the American Veterinary Medical Association 212, 61–66.

Hood, D.M., Wagner, I.P., Taylor, D.D., Brumbaugh, G.W., Chaffin, M.K., 2001. Voluntary limb-load distribution in horses with acute and chronic laminitis. American Journal of Veterinary Research 62, 1393–1398.

Jones, E., Viñuela-Fernandez, I., Eager, R.A., Delaney, A., Anderson, H., Patel, A., Robertson, D.C., Allchorne, A., Sirinathsinghji, E.C., Milne, E.M., MacIntyre, N., Shaw, D.J., Waran, N.K., Mayhew, J., Fleetwood-Walker, S.M., 2007. Neuropathic changes in equine laminitis pain. Pain 132, 321–331.

Keegan, K.G., Wilson, D.A., Wilson, D.J., Smith, B., Gaugham, E.M., Pleasant, R.S., Lillich, J.D., Kramer, J., Howard, R.D., Bacon-Miller, C., Davis, E.G., May, K.A., Cheramie, H.S., Valentino, W.L., van Harreveld, P.D., 1998. Evaluation of mild lameness in horses trotting on a treadmill by clinicians and interns or residents and correlation of their assessment with kinematic gait analysis. American Journal of Veterinary Research 59, 1370–1377.

Keegan, K.G., Arafat, S., Skubic, M., Wilson, D.A., Kramer, J., 2003. Detection of lameness and determination of the affected forelimb in horses by use of a continuous wavelet transformation and neural network classification of kinematic data. American Journal of Veterinary Research 64, 1376–1381.

Lascelles, B.D.X., Butterworth, S.J., Waterman, A.E., 1994. Postoperative analgesic and sedative effects of carprofen and pethidine in dogs. Veterinary Record 134, 187–191.

McGuigan, M.P., Walsh, T.C., Pardoe, C.H., Day, P.S., Wilson, A.M., 2005. Deep digital flexor tendon force and digital mechanics in normal ponies and ponies with rotation of the distal phalanx as a sequel to laminitis. Equine Veterinary Journal 37, 161–165.

Obel, N., 1948. Studies on the Histopathology of Acute Laminitis. PhD Dissertation, Almqvist and Wiksells, Botryckeri, Uppsala, Sweden.

Owens, J.G., Kamerling, S.G., Stanton, S.R., Keowen, M.L., 1995. Effects of ketoprofen and phenylbutazone on chronic hoof pain and lameness in the horse. Equine Veterinary Journal 27, 296–300.

Peham, C., Licka, T., Girtler, D., Scheidl, M., 1999. Supporting forelimb lameness: clinical judgment vs. computerised symmetry measurement. Equine Veterinary Journal 31, 417–421.

Peloso, J.G., Stick, J.A., Soutas-Little, R.W., Caron, J.C., Decamp, C.E., Leach, D.H., 1993. Computer-assisted three-dimensional gait analysis of amphotericin-induced carpal lameness in horses. American Journal of Veterinary Research 54, 1535–1543.

Pollitt, C.C., 2003. Medical management of laminitis. In: Diagnosis and Management of Lameness in the Horse. W.B. Saunders, Missouri, USA, pp. 329–331.

Portney, L.G., Watkins, M.P., 2000. Foundations of Clinical Research; Application to Practice. Julie Alexander, Upper Saddle River, USA.

Reitmann, T.R., Stauffacher, M., Bernasconi, P., Auer, J.A., Weishaupt, M.A., 2004. The association between heart rate, heart rate variability, endocrine and behavioural pain measures in horses suffering from laminitis. Journal of Veterinary Medicine A 51, 218–225.

Reid, J., Nolan, A.M., 1991. A comparison of the postoperative analgesic and sedative effects of flunixin and papaveretum in the dog. Journal of Small Animal Practice 32, 603–608.

Robertson, S.A., 2003. How do we know if they hurt? Pain assessment in small animals. Veterinary Medicine 98, 70–79.

Rutherford, K.M.D., 2002. Assessing pain in animals. Animal Welfare 11, 31–53.

Schatzmann, U., 2000. Practical analgesic treatment in horses. In: Hellebrekers, L.J. (Ed.), Animal Pain. A Practice-Oriented Approach to an Effective Pain Control in Animals. Van Der Wees, Utrecht, The Netherlands, pp. 161–179.

Scott, J., Huskisson, E.C., 1976. Graphic representation of pain. Pain 2, 175–184.

Silver, I.A., Brown, P.W., Goodship, A.E., Lanyon, L.E., McCullough, K.G., Perry, G.C., Williams, I.F., 1983. A clinical and experimental study of tendon injury, healing and treatment in the horse. Equine Veterinary Journal Supplement 1, 1–43.

Streiner, D.L., Norman, G.R., 2008. Health Measurement Scales: A Practical Guide to Their Development and Use, fourth ed. Oxford University Press, New York, USA.

Swanson, T.D., 1999. Clinical presentation, diagnosis, and prognosis of acute laminitis. Veterinary Clinics of North America: Equine Practice 15, 311–319.

Taylor, D., Hood, D.M., Wagner, I.P., 2002. Short-term effect of therapeutic shoeing on severity of lameness in horses with chronic laminitis. American Journal of Veterinary Research 63, 1629–1633.

Welsh, E.M., Gettinby, G., Nolan, A.M., 1993. Comparison of a visual analogue scale and a numerical rating scale for assessment of lameness, using sheep as a model. American Journal of Veterinary Research 54, 976–983.

Please cite this article in press as: Viñuela-Fernández, I., et al. Comparison of subjective scoring systems used to evaluate equine laminitis. The Veterinary Journal (2010), doi:10.1016/j.tvjl.2010.05.011