The Veterinary Journal 206 (2015) 356–364

Contents lists available at ScienceDirect

The Veterinary Journal

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

Monitoring acute equine visceral pain with the Equine Utrecht University Scale for Composite Pain Assessment (EQUUS-COMPASS) and the Equine Utrecht University Scale for Facial Assessment of Pain (EQUUS-FAP): A scale-construction study

Johannes P.A.M. van Loon [a,*], Machteld C. Van Dierendonck [a,b,c]

[a] *Department of Equine Sciences, Faculty of Veterinary Medicine, Utrecht University, Yalelaan 114, NL-3584 CM Utrecht, The Netherlands*
[b] *Veterinary Faculty Ghent University, Salisburylaan 133, 9820 Merelbeke, Ghent, Belgium*
[c] *Equus Research and Therapy, Tolnegenweg 39, 3776PT Stroe, The Netherlands*

A R T I C L E I N F O

Article history:
Accepted 20 August 2015

Keywords:
Colic
Equine
Pain scale
Facial expression
EQUUS-COMPASS
EQUUS-FAP

A B S T R A C T

Although recognition of equine pain has been studied extensively over the past decades there is still need for improvement in objective identification of pain in horses with acute colic. This study describes scale construction and clinical applicability of the Equine Utrecht University Scale for Composite Pain Assessment (EQUUS-COMPASS) and the Equine Utrecht University Scale for Facial Assessment of Pain (EQUUS-FAP) in horses with acute colic. A cohort follow-up study was performed using 50 adult horses ($n = 25$ with acute colic, $n = 25$ controls). Composite pain scores were assessed by direct observations, Visual Analog Scale (VAS) scores were assessed from video clips. Colic patients were assessed at arrival, and on the first and second mornings after arrival. Both the EQUUS-COMPASS and EQUUS-FAP scores showed high inter-observer reliability (ICC = 0.98 for EQUUS-COMPASS, ICC = 0.93 for EQUUS-FAP, $P < 0.001$), while a moderate inter-observer reliability for the VAS scores was found (ICC = 0.63, $P < 0.001$). The cut-off value for differentiation between healthy and colic horses for the EQUUS-COMPASS was 5, and for differentiation between conservatively treated and surgically treated or euthanased patients it was 11. For the EQUUS-FAP, cut-off values were 4 and 6, respectively.

Internal sensitivity and specificity were good for both EQUUS-COMPASS (sensitivity 95.8%, specificity 84.0%) and EQUUS-FAP (sensitivity 87.5%, specificity 88.0%). The use of the EQUUS-COMPASS and EQUUS-FAP enabled repeated and objective scoring of pain in horses with acute colic. A follow-up study with new patients and control animals will be performed to further validate the constructed scales that are described in this study.

Introduction

Animal pain and its recognition and management have received abundant attention over the past decades. Various studies describing tools for objective assessment of pain in farm animals (Prunier et al., 2013), companion animals (Hellyer, 2002; Hansen, 2003) and horses (Wagner, 2010) have laid the basis for the development of improved methods for objective pain assessment, which in turn have contributed positively to animal welfare (Valverde and Gunkel, 2005). Apart from species differences, pain expression is also dependent on the type and origin of pain. Somatic pain and visceral pain, for instance, are different phenomena that manifest differently and need to be treated differently (Robertson, 2002). Colic is one of the most important and often diagnosed diseases in the horse and the availability of purpose built, specific and validated pain assessment tools would help identify colic related pain and therefore add significantly to equine welfare and support quality of patient care.

Simple one-dimensional pain scales like the VAS (Visual Analog Scale) are very commonly used, but are deemed suboptimal instruments for pain evaluation in animals, partly because of poor inter-observer reliability (Lindegaard et al., 2010). Composite pain scales (CPS) offer the advantage that combining various indicators increases sensitivity and specificity of pain assessment (Abbott et al., 1995; Dobromylskyj et al., 2000; Prunier et al., 2013). In humans, multidimensional pain scales have been developed for recurrent abdominal pain in children (Malaty et al., 2005) and assessment of infant pain (Cong et al., 2013). Such pain scales have been described for different types of pain in horses like acute orthopaedic pain (Bussières et al., 2008) and pain after colic surgery (Pritchett

* Corresponding author. Tel.: +31 30 253 1111.
E-mail address: j.p.a.m.vanloon@uu.nl (J.P.A.M. van Loon).

http://dx.doi.org/10.1016/j.tvjl.2015.08.023

et al., 2003; Graubner et al., 2011; van Loon et al., 2014). Recently, a behaviour-based pain scale for horses with acute colic has been described and validated (Sutton et al., 2013a, 2013b).

Another strategy to assess objectively the amount of pain in an individual is to quantify facial expression. This technique has been studied in humans (Ahola Kohut et al., 2012) and rodents (Sotocinal et al., 2011) and has led to the development of grimace scales for the latter species. Recently, the Horse Grimace Scale (HGS) following surgical castration has been described (Dalla Costa et al., 2014) and the Equine Pain Face has been described after induced experimental pain (Gleerup et al., 2014).

In previous studies, the CPS as described by Bussières et al. (2008) was used to assess pain in a cross-section of equine patients in a referral centre (van Loon et al., 2010) and after emergency laparotomy and treatment in an intensive care unit (van Loon et al., 2014). This pain scale, although originally developed for orthopaedic pain, contains various elements that can also be applied to visceral pain (van Loon et al., 2014). In the current study, the CPS was used as the basis for the development of the Equine University Utrecht Scale for Composite Pain Assessment (EQUUS-COMPASS), a scale aiming at the optimal assessment of acute colic pain. Furthermore, a composite facial expression pain scale was constructed, the Equine University Utrecht Scale for Facial Assessment of Pain (EQUUS-FAP), based on several facial expression characteristics.

The aims of the current study were (1) to assess inter-observer variability of the EQUUS-COMPASS and the EQUUS-FAP; (2) to determine cut-off points for determination between healthy and painful animals; and (3) to assess the clinical applicability for the identification and follow-up of pain in horses with acute colic. The hypotheses were that EQUUS-COMPASS and EQUUS-FAP would have better inter-observer reliability than a VAS scale, would be easily clinically applicable and would be able to differentiate between control horses and colic patients.

Materials and methods

Animals

The study design was approved by the institutional Ethics Committee on the Care and Use of Experimental Animals in compliance with Dutch legislation on animal experimentation. Because the procedures used in this study only contain behavioural observations and physiological assessments (heart rate, breathing rate, borborygmi, rectal temperature) taken from clinical patients and therefore are not likely to cause pain, suffering or distress or lasting harm equivalent to, or higher than, that caused by the introduction of a needle (article 1.5f EU directive 2010/63/EU), ethical approval was obtained without an official approval number. Furthermore, owner's consent was obtained for all horses and ponies participating in this study.

Twenty-five horses that had been admitted to the equine referral centre with acute colic were included (Table 1). Twenty-five control horses (healthy mares that were used as recipients for embryo transfer and horses that came in for regular shoeing) that were admitted in the same period were included as well (Table 1). All control horses were free from lameness and/or teeth problems.

The total study population consisted of 30 mares and 20 geldings. Stallions, foals and mares with foals were excluded from the study because of possibly disturbing effects of sexual arousal or mare–foal interaction on the assessment of pain scores. Breeds included Warmbloods (37), Thoroughbreds (crossbreeds) (3), Friesians (5), Irish Cob (1), Fjorden horse (1), Haflinger (1), Icelandic (1) and Quarterhorse (1). Analgesic treatment and clinical decision-making were at the discretion of the attending veterinarian and independent of pain scores. The observers were not involved with day-to-day patient care and were unaware of any analgesic treatment. All colic patients were given non-steroidal anti-inflammatory drugs (NSAIDs) before referral to the university clinic. If α_2-agonists were required after arrival at the clinic, the horses were excluded from the study because of possible interference with pain scores.

Table 1
Data of horses that were included in the study ($n = 50$).

	Colic	Control
Number of horses	25	25
Conservative treatment (CT)	15	–
Surgical treatment/euthanasia (STE)	5/5	–
Warmblood/Thoroughbred	17	23
Other breeds	8	2
Mean (±SD) weight (kg)	540 (71.8)	593 (37.6)
Mean (±SD) age (years)	11 (6.4)	9 (4.3)

Equine Utrecht University Scale for Composite Pain Assessment (EQUUS-COMPASS)

The EQUUS-COMPASS is based on the CPS described by Bussières et al. (2008). For the development of EQUUS-COMPASS, the CPS was modified by deleting parameters that are not possible to assess in horses with acute abdominal pain (e.g. appetite) and by adding parameters that are thought to be more specific for visceral pain (such as tail flicking, laying down and sounds produced as an expression of pain like teeth grinding or moaning). The EQUUS-COMPASS is a multifactorial simple descriptive scale (SDS) based on 14 parameters. It includes physiological parameters, responses to stimuli, and spontaneous behavioural parameters (Table 2). Each of the 14 parameters can be scored from 0 to 3, leading to a total pain score range from 0 (no signs of pain) to 42 (maximal pain score).

Equine Utrecht University Scale for Facial Assessment of Pain (EQUUS-FAP)

The EQUUS-FAP is a multifactorial SDS based on nine parameters, describing different elements of facial expression, like appearance of eyelids, nostrils and muscle tone (Table 3). Each of the nine parameters can be scored from 0 to 2, leading to a total pain score ranging from 0 (no signs of pain) to 18 (maximal pain score).

Experimental design

Observations were performed by four observers (veterinary students) who performed their observations pairwise and simultaneously. The observers did not discuss their findings. Prior to commencement of the study, all observers were given the chance to familiarize and train themselves with the parameters in the EQUUS-COMPASS and EQUUS-FAP using pain-free horses (not included in the study). The observers were not blinded for the clinical diagnosis.

Patients were evaluated shortly after admission to the university hospital (T0) when horse owners were registered to the patient database, the first morning after admission (T12-24) and the second morning after admission (T36-48). Each observation period lasted 10 min. Scoring was performed with the animals in the colic box, where a video camera recorded the box during scoring. These videos were used to obtain Visual Analog Scores (VAS) on a continuous scale between 0 and 10 (Hawker et al., 2011), performed by two observers (equine veterinarians not involved in the treatment and blinded for time and treatment). The control horses were observed in the same colic box and EQUUS-COMPASS and EQUUS-FAP scores were obtained once, with simultaneous video recordings for VAS scores.

Data processing and statistical analysis

All data are expressed as medians and quartiles. Inter-observer reliability was assessed using Intraclass Correlation Coefficients (ICC). Bland–Altman plots were used to visually evaluate correlations and determine limits of agreement (average difference ± 1.96 standard deviation of the difference) (Bland and Altman, 1986; Myles, 2007). Differences in scores between control animals and colic patients and between conservatively treated animals (CT) and surgically treated or euthanized animals (STE) were analyzed using the Mann–Whitney U test. Cut-off values for EQUUS-COMPASS, EQUUS-FAP and VAS were determined to obtain maximal differentiation between colic patients and healthy animals and between CT and STE treatments. Internal sensitivity, specificity, and positive and negative predictive values were determined for EQUUS-COMPASS, EQUUS-FAP and VAS scores using these cut-off values. Sensitivity and specificity for individual parameters of both scores were also determined. Based on these values, weighting factors for future validation for the different individual parameters were determined retrospectively. When sensitivity or specificity was ≤25%, a weighting factor of 0 was applied; between 25% and 50% the weighting factor was 1; between 50% and 75% the weighting factor was 2; and when both sensitivity and specificity were ≥75% a weighting factor of 3 was applied. The effects over time for both EQUUS-COMPASS and EQUUS-FAP scores in colic patients were assessed by means of the Friedman test. Statistical analysis was performed using SPSS version 20.0 (IBM). Statistical significance was accepted at $P < 0.05$.

Results

Inter-observer reliability

Fig. 1 shows the results of correlation analysis between the different pain scores of two independent observers. Both the EQUUS-COMPASS and EQUUS-FAP scores showed strong and significant correlation (ICC = 0.98, $P < 0.001$ for EQUUS-COMPASS scores,

Table 2
Score sheet of the Equine Utrecht University Scale for Composite Pain Assessment (EQUUS-COMPASS).

Data	Categories	Score
Physiological data	24–44 beats/min	0
Heart rate	45–52 beats/min	1
	53–60 beats/min	2
	>60 beats/min	3
Respiratory rate	8–13 breaths/min	0
	14–16 breaths/min	1
	17–18 breaths/min	2
	>18 breaths/min	3
Rectal temperature	36.9 °C–38.5 °C	0
	36.4 °C–36.9 °C or 38.5 °C–39.0 °C	1
	35.9 °C–36.4 °C or 39.0 °C–39.5 °C	2
	35.4 °C–35.9 °C or 39.5 °C–40.0 °C	3
Digestive sounds	Normal motility	0
	Decreased motility	1
	No motility	2
	Hypermotility or steel band	3
Behaviour	Quietly standing and/or one hind leg resting, explores environment	0
Posture	Slightly tucked up abdomen, still explores environment (with possible unrest)	1
	Extremely tucked up abdomen, hunched back and/or stretching of body/limbs	2
	Does not stand or for short amounts of time (<1 min), sits on hindquarters	3
Laying down, rolling	Does not lie down or rests lying down	0
	Lies down in normal posture, rolls or tries to roll (once or twice/5 min)	1
	Alternates lying down and standing, rolls or tries to roll (more than twice/5 min)	2
	Constantly lies in an abnormal position: on its side with stretched limbs, on its back, or does not stop rolling	3
Sweating	No signs of sweating	0
	Warm or damp to touch, no sweat or wet spots visible	1
	Wet spots visible, no droplets or streams	2
	Excessive sweating, may include streams or droplets	3
Tail flicking (excluding flicking to chase off insects)	No tail flicking	0
	Occasional tail flicking (once or twice/5 min) and/or holds tail away from body	1
	Frequent tail flicking (three to four times/5 min), may hold tail away from body	2
	Excessive tail flicking (more than five times/5 min)	3
Kicking at abdomen	Quietly standing, no kicking	0
	Occasional kicking at abdomen (once or twice/5 min)	1
	Frequent kicking at abdomen (three to four times/5 min)	2
	Excessive kicking at abdomen (more than five times/5 min)	3
Pawing at floor (number of episodes)	Quietly standing, does not paw at floor	0
	Occasional pawing at floor (once or twice times/5 min)	1
	Frequent pawing at floor (three to four times/5 min)	2
	Excessive pawing at floor (more than five times/5 min)	3
Head movements	No fast movements, head mostly at same height/in same direction	0
	Occasional head movements laterally/vertically, looking at flank (once or twice/5 min)	1
	Frequent/fast head movements laterally/vertically, looking at flank (three to four times/5 min)	2
	Excessive head movements, excessive looking at flank (more than five times/5 min), biting at flank (more than once/5 min)	3
Pain sounds	No audible signs of pain	0
	Occasional teeth grinding or moaning (once or twice/5 min)	1
	Frequent teeth grinding or moaning (three to four times/5 min)	2
	Excessive teeth grinding or moaning (more than five times/5 min)	3
Overall appearance, reaction to observer(s)	Quiet but alert, approaches/turns to observer	0
	Alert, no reluctance to move, obvious reaction to sounds and/or movements	1
	Restless, constantly moving, exaggerated reaction to sounds and/or movements	2
	Stupor: the horse is not moving, head is lowered, reluctance to move	3
Reaction to palpation of painful area in the flank	No reaction to palpation	0
	Mild reaction to palpation	1
	Resistance to palpation	2
	Violent reaction to (attempt to) palpation	3
Total		. . ./42

ICC = 0.93, $P < 0.001$ for EQUUS-FAP scores), while a moderate but significant correlation for the VAS scores was found (ICC = 0.63, $P < 0.001$).

Differences between subgroups of control and colic patients

Both EQUUS-COMPASS and EQUUS-FAP scores showed significant differences between control and colic patients ($P < 0.001$). The EQUUS-COMPASS showed a statistically significant difference between CT and STE patients ($P < 0.01$), while the EQUUS-FAP scores between these two groups were not significantly different ($P = 0.84$). VAS scores were not different between control and colic patients ($P = 0.051$), nor between CT and STE patients ($P = 0.77$) (Fig. 2).

Relation of breed, age and sex with pain scores

There was no relationship between EQUUS-COMPASS and EQUUS-FAP scores and breed or sex (Figs. 3A–D). EQUUS-COMPASS scores in the group of young horses (0–5 years) were significantly lower than in elderly horses (>15 years) ($P < 0.05$, Fig. 3E). EQUUS-FAP scores between these two age groups were not significantly different. However, the distribution of CT and STE patients was also

Table 3
Score sheet of the Equine Utrecht University Scale for Facial Assessment of Pain (EQUUS-FAP).

Data	Categories	Score
Head	Normal head movement/interested in environment	0
	Less movement	1
	No movement	2
Eyelids	Opened, sclera can be seen in case of eye/head movement	0
	More opened eyes or tightening of eyelids. An edge of the sclera can be seen 50% of the time	1
	Obviously more opened eyes or obvious tightening of eyelids. Sclera can be seen >50% of the time	2
Focus	Focussed on environment	0
	Less focussed on environment	1
	Not focussed on environment	2
Nostrils	Relaxed	0
	A bit more opened	1
	Obviously more opened, nostril flaring and possibly audible breathing	2
Corners mouth/lips	Relaxed	0
	Lifted slightly	1
	Obviously lifted	2
Muscle tone head	No fasciculations	0
	Mild fasciculations	1
	Obvious fasciculations	2
Flehming and/or yawning	Not seen	0
	Seen	2
Teeth grinding and/or moaning	Not heard	0
	Heard	2
Ears	Position: Orientation towards sound/clear response with both ears or ear closest to source	0
	Delayed/reduced response to sounds	1
	Position: backwards/no response to sounds	2
Total		.../18

unequal in those two age groups (0–5 years: eight horses CT and one horse STE, >15 years: three horses CT and five horses STE).

Effects over time in horses admitted with acute colic

Fig. 4 shows both EQUUS-COMPASS and EQUUS-FAP scores over time of CT patients. For both pain scores, a significant decrease over time was found ($P < 0.05$).

Cut-off values between different categories of horses

Cut-off values for the EQUUS-COMPASS were 5 for differentiation between controls and colic patients and 11 as cut-off value for the distinction between CT and STE patients. Cut-off values for the EQUUS-FAP were 4 for differentiation between controls and colic patients and 6 for the distinction between CT and STE patients. Cut-off values for the VAS were 2 for differentiation between controls and colic patients and 5 as cut-off value for the distinction between CT and STE patients. Cut-off values for individual parameters of both EQUUS-COMPASS and EQUUS-FAP were 0 for controls and >0 for colic patients; for the differentiation between CT and STE patients the cut-off value was 0 (one respectively >1).

Internal sensitivity and specificity of composite pain scores and their individual parameters

Table 4 shows internal sensitivity and specificity of EQUUS-COMPASS, EQUUS-FAP and VAS scores using the cut-off values as described above. Sensitivity and specificity of the individual parameters and the determined weighting factors are shown in Table 5.

Discussion

This study describes the construction of both the EQUUS-COMPASS and EQUUS-FAP, a composite pain scale and a facial expression pain scale, respectively, for an objectified evaluation of the severity of pain symptoms in horses with acute colic. Scale development was supported by the high inter-observer reliability of both scores and the clinically acceptable limits of agreement that were found using Bland–Altman analysis. Good internal specificity and sensitivity for differentiating between pain-free healthy horses and colic patients and between colic patients that were treated conservatively and colic patients that were treated surgically or were euthanased show that the scoring systems were able to differentiate different pain levels (no pain, mild to moderate pain, and severe pain). In a follow-up study, a new cohort of horses will be used to further assess clinical applicability and to assess external sensitivity and specificity.

In the original CPS by Bussières et al. (2008) the key specific and most sensitive behavioural indices for acute orthopaedic pain were response to palpation of the painful area, posture and, to a lesser extent, pawing on the floor, kicking at the abdomen and head movements. For visceral postoperative pain (van Loon et al., 2014), pawing on the floor, overall appearance, head movements and interactive behaviour were the most important parameters of the CPS. The most sensitive parameters of the EQUUS-COMPASS for acute colic were the character of borborygmi, posture, sweating and reaction to observer and palpation of the painful flank. The importance of borborygmi for the assessment of severity of colic corresponds with the results of White et al. (2005), who found significantly increased odds ratios for need of surgery in animals with decreased or no intestinal sounds.

In the current study, sensitivity and specificity were determined for all individual parameters of both the EQUUS-COMPASS and EQUUS-FAP and these were used to determine weighting factors. By using these weighting factors for individual parameters, sensitivity, specificity and positive and negative predictive values can potentially be improved with a limited number of parameters. Cut-off points between healthy individuals and those suffering from colic pain and between different levels of pain were determined by optimizing sensitivity and specificity for differentiation between them. The cut-off values and weighting factors we determined in this scale construction study will be used in a follow-up study for determination of external sensitivity and specificity.

The EQUUS-FAP could potentially be used by horse owners and non-veterinarians after proper training, while the EQUUS-COMPASS contains several physiological parameters (borborygmi, heart and breathing frequency) that can be assessed only by veterinarians and

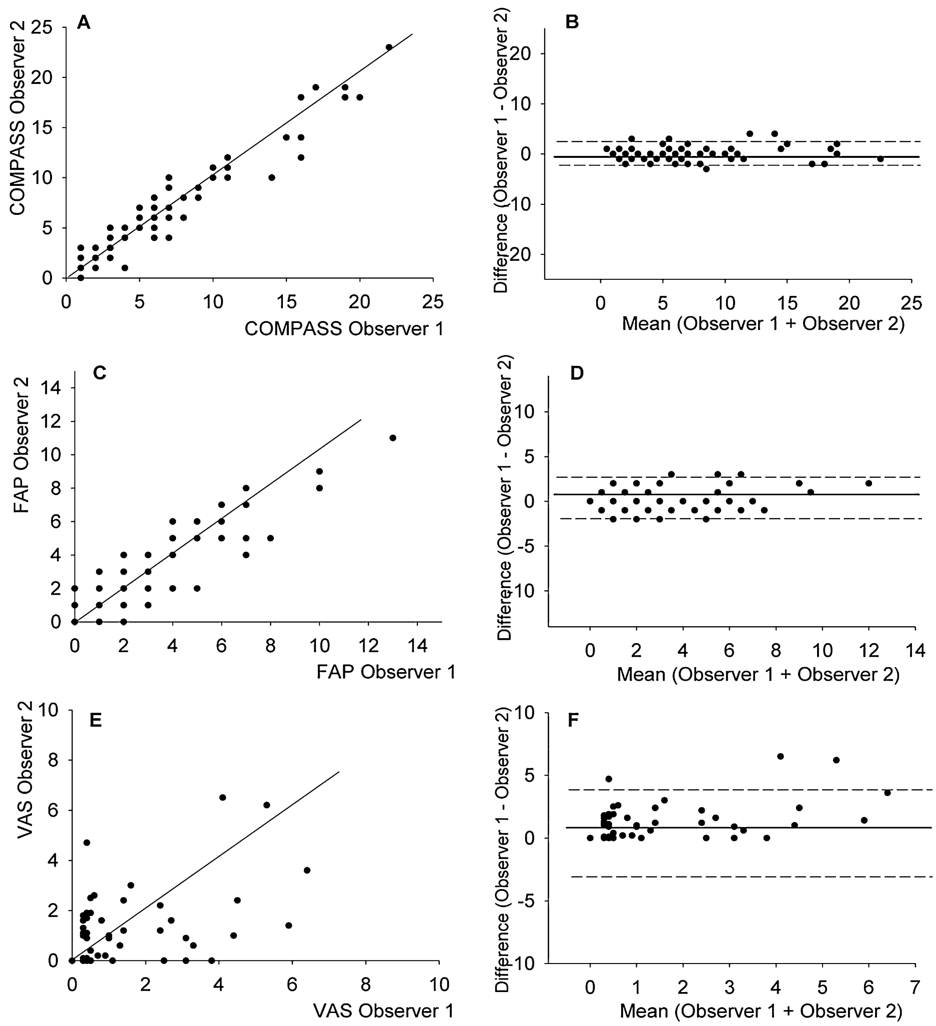

Fig. 1. Scatter plot of EQUUS-COMPASS scores (A) ($n = 72$, $r = 0.94$, ICC = 0.98, $P < 0.001$), EQUUS-FAP scores (C) ($n = 72$, $r = 0.84$, ICC = 0.93, $P < 0.001$) and VAS scores (E) ($n = 57$, $r = 0.34$, ICC = 0.63, $P < 0.05$), assessed by two different observers at the same moment. Bland–Altman plots of EQUUS-COMPASS (B), EQUUS-FAP (D) and VAS scores (F) show limits of agreement of –2.8 to +2.6 for EQUUS-COMPASS, –2.2 to +2.8 for EQUUS-FAP and –2.9 to +3.4 for VAS.

well-educated owners. A condensed form of the EQUUS-COMPASS without physiological parameters could also be of potential use for non-veterinarians. Objectified multi-dimensional cut-off values for differentiation between healthy and sick animals can help determine the necessity for veterinary help or for referral to a veterinary clinic or, after treatment, evaluation of disease/pain progression.

Pain scores will never replace clinical decision making (with physical examination, blood results, ultrasound etc.), but can aid follow-up of a patient and objectify responses to treatment. The importance of pain as a prognosticating factor for the severity of colic has also been described by others (Grulke et al., 2001; White et al., 2005), who found strongly increased odds ratios for the need for surgery in horses with constant pain or with pain that returned after treatment.

Recently, in the studies by Sutton et al. (2013a, 2013b), another concept has been applied to construct and validate pain scales for horses that experience acute colic (Equine Acute Abdominal Pain Scale: EAAPS-1 and EAAPS-2). These authors used a clinimetric approach to generate, select and appoint weights to items that were involved in pain expression in horses with acute colic. Expert opinion was used in a Delphi process to independently judge several items that best reflected expression of pain. Although this procedure is scientifically sound, item selection and the application of weighting factors still depend on expert opinion, which can be subjective. Furthermore, the readout parameters of these two pain scales were based on expression of only one behavioural parameter and increasing pain scores were attributed to different types of behaviour. In EAAPS-2, the frequency of expression of the specific behavioural items influenced the pain score. The disadvantage of the EAAPS is that it does not take into account that factors such as breed, sex or individual effects may influence expression of pain through differences in behavioural patterns. Composite pain scales offer the advantage of these differences in individual expression.

In the follow-up article by Sutton et al. (2013b), both behaviour-based pain scales (EAAPS-1 and -2) were validated by means of digital film clips showing horses with various degrees of acute colic and control horses. Inter-observer reliability was very good for both EAAPS scales (ICC > 0.75), while the Numerical Rating Scale (NRS) that was used in this study for comparison showed only fair inter-observer reliability (ICC = 0.67). This is in accordance with our results, but we found higher inter-observer reliability for the EQUUS-COMPASS and the EQUUS-FAP (ICC > 0.93) and only moderate inter-observer reliability for the Visual Analog Scale (VAS) (ICC = 0.63).

In contrast to our study, Sutton et al. (2013b) also assessed intra-observer reliability. However, only one video was shown twice to the observers for assessment of intra-observer reliability, which limits the value of this analysis. With respect to sensitivity and specificity for discrimination between colic cases and controls and between

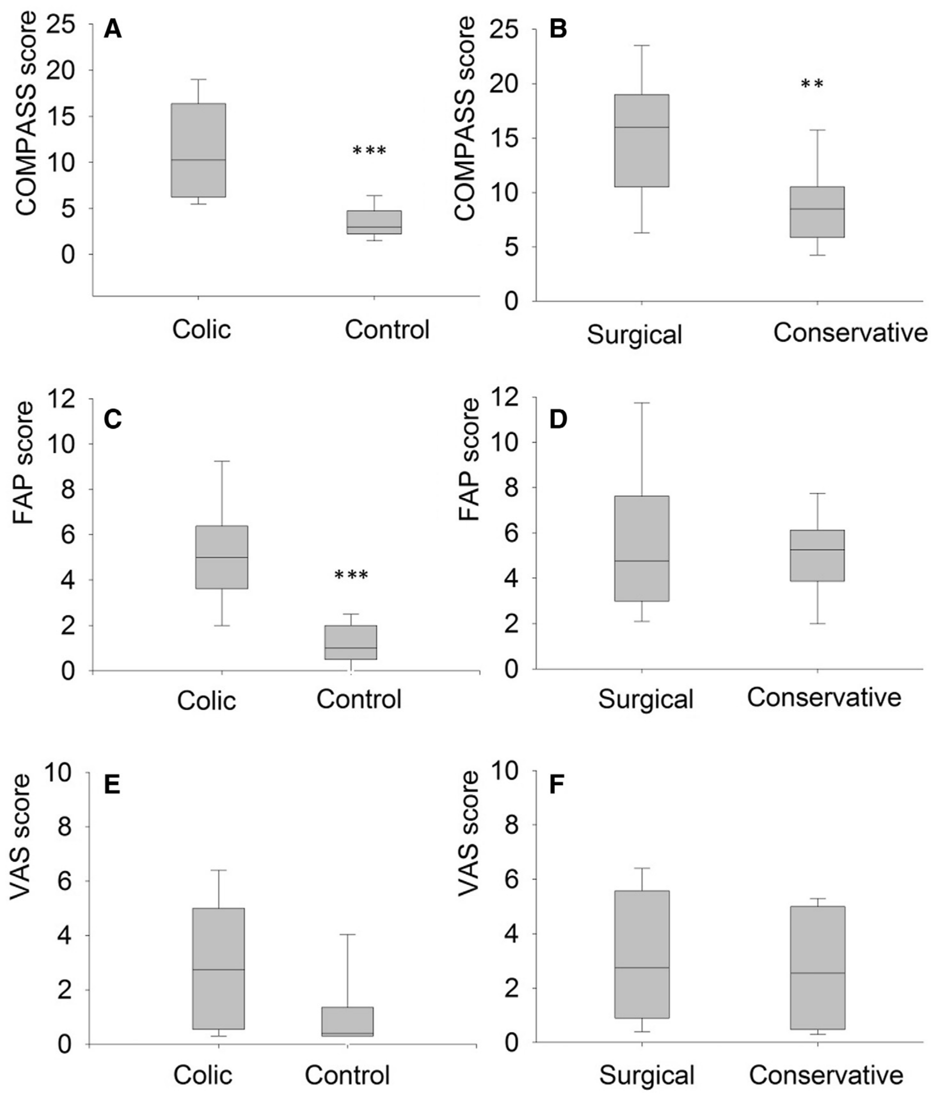

Fig. 2. Median EQUUS-COMPASS, EQUUS-FAP and VAS scores for colic ($n = 25$) versus control ($n = 25$) horses (A + C + E) and surgically treated/euthanased (STE) ($n = 10$) versus conservatively treated (CT) ($n = 15$) colic patients (B + D + F). Lines in boxes show median scores; boxes show 25–75th percentiles; error bars show 5–95th percentiles. ** = $P < 0.01$, *** = $P < 0.001$.

colic patients with different treatments (surgical versus conservative), the EAAPS-1 pain scale obtained best results with a sensitivity of 89% and a specificity of 75% for discrimination between colic cases and controls and of 70% and 57%, respectively, for discrimination between animals with different treatments.

By means of the EQUUS-FAP, the amount of pain expressed by horses with acute colic was assessed by means of facial expression only. Facial expression has been proven suitable for quantification of pain in rodents (Sotocinal et al., 2011) and infants (Ahola Kohut et al., 2012). Validation of grimace scales has been successful for rodents and such a scale has been proven to be suitable for horses after surgical castration in a recent study (Dalla Costa et al., 2014). We chose to build a multifactorial numerical rating scale, with descriptions of various aspects of facial expression such as ear and head movements and appearance of eyelids and nostrils, and pain-related sounds like teeth grinding, moaning and yawning. Our results showed that facial expression was very suitable for recognition and quantification of the severity of acute visceral pain in horses. Apart from acute visceral pain, pilot studies have shown that the EQUUS-FAP is also very suitable for recognition of acute dental and ocular pain in horses (unpublished data).

One of the limitations of our study is that observers were not blinded for clinical diagnosis. However, this is inevitable with live observations, and because observations were strictly structured and the observers were not involved with clinical decision-making and treatment, this was deemed unimportant. Another limitation is that a 'gold standard' for scoring acute visceral pain needed for validation of our pain scores is lacking. Therefore, we decided to obtain video clips from all horses, taken when composite pain scores were assessed. These videos were used to obtain VAS scores for comparison with the composite pain scores. However, in accordance with other studies (Lindegaard et al., 2010; Sutton et al., 2013a), inter-observer reliability of VAS scores was low with wide limits of agreement and 95% confidence intervals and VAS scores proved not useful for clinical validation of the composite pain scores.

Another possibility for better clinical validation of the pain scales could be to assess the effect of visceral analgesics like NSAIDs or opioids (more specifically κ-agonists) on development of pain scores in patients with acute colic. In our study, this was not possible because most horses with colic were referred to the university hospital after being treated with an NSAID. A follow-up study carried out in horses with acute colic at the first visit of the attending

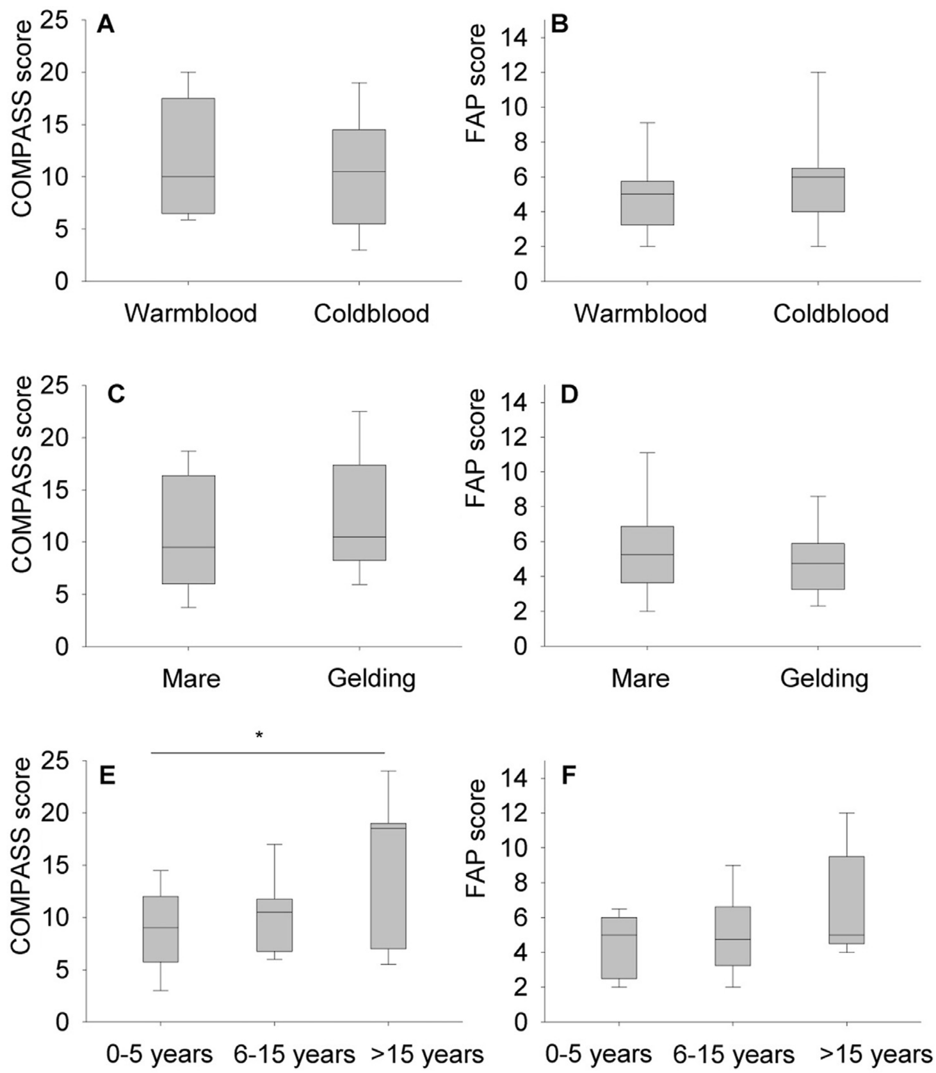

Fig. 3. Median EQUUS-COMPASS and EQUUS-FAP values for Warmbloods/Thoroughbreds ($n = 17$) and Cold bloods ($n = 8$: Friesians, $n = 3$; Irish Cob, $n = 1$; Quarterhorse, $n = 1$; Fjorden, $n = 1$; Haflinger, $n = 1$; Icelandic, $n = 1$) animals (A + B), horses with different sexes (mares, $n = 13$; geldings, $n = 12$) (C + D) and horses with different ages (0–5 years, $n = 9$; 6–15 years, $n = 8$; > 15 years, $n = 8$) (E + F). Lines in boxes show median scores; boxes show 25–75th percentiles; error bars show 5–95th percentiles. * = $P < 0.05$.

veterinarian would be suitable in order to test this. External validation with a new dataset of colic patients and healthy control horses will be performed to further determine the reliability and clinical usefulness of both pain scales and to complete their validation.

Assessment of the pain scoring systems by other veterinarians or veterinary technicians would elucidate more on the overall added value of the systems. Another step forward for practical implementation of the pain scoring systems will be to determine optimal

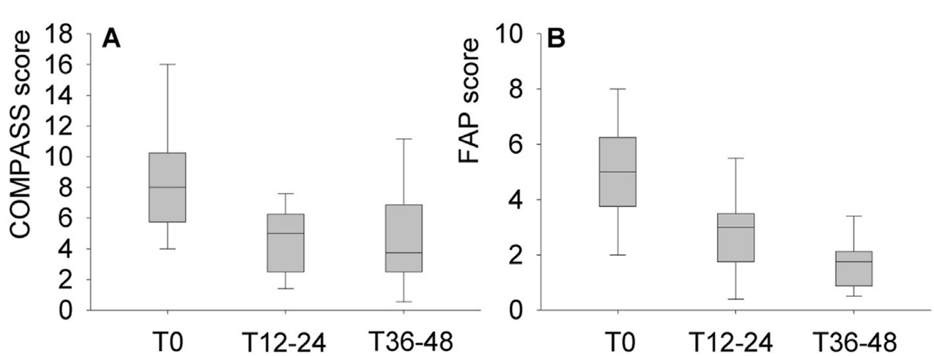

Fig. 4. Median EQUUS-COMPASS (A) and EQUUS-FAP (B) values at admission to clinic (T0), at the first morning after admission (T12-24) and at the second morning after admission (T36-48) ($n = 13$). Lines in boxes show median scores; boxes show 25–75th percentiles; error bars show 5–95th percentiles. Friedman tests show a significant decrease over time for both EQUUS-COMPASS ($P = 0.021$) and for EQUUS-FAP scores ($P = 0.012$).

Table 4
Internal sensitivity, specificity, positive and negative predictive value of the Equine Utrecht University Scale for Composite Pain Assessment (EQUUS-COMPASS), the Equine Utrecht University Scale for Facial Assessment of Pain (EQUUS-FAP) and the Visual Analog Scale (VAS).

	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Colic ($n = 25$) versus controls ($n = 25$)				
EQUUS-COMPASS	95.8	84.0	85.2	95.5
EQUUS-FAP	87.5	88.0	87.5	88.0
VAS	62.5	84.0	44.0	87.5
CT ($n = 15$) versus STE ($n = 10$)				
EQUUS-COMPASS	80.0	85.7	80.0	85.7
EQUUS-FAP	30.0	64.3	37.5	56.3
VAS	25.0	75.0	50.0	50.0

Weighting factors of individual parameters can be found in Table 5.
CT, conservatively treated animals; STE, surgically treated or euthanased animals.

Table 5
Sensitivity, specificity and weighting factors for individual parameters of the Equine Utrecht University Scale for Composite Pain Assessment (EQUUS-COMPASS) and the Equine Utrecht University Scale for Facial Assessment of Pain (EQUUS-FAP).

	Colic ↔ control		Weighting factor[a]	CT ↔ STE		Weighting factor[a]
	Sensitivity	Specificity		Sensitivity	Specificity	
	(%)			(%)		
EQUUS-COMPASS						
Heart frequency	46	54	1	100	40	1
Breathing frequency	42	54	1	86	67	2
Rectal temperature	08	100	0	–		
Borborygmi	92	96	3	80	36	1
Posture	88	72	2	44	75	1
Laying down, rolling	29	100	1	60	50	2
Sweating	54	100	2	38	100	1
Tail flicking	58	08	0	–		
Kicking at abdomen	29	100	1	40	50	1
Pawing at floor	63	48	1	86	38	1
Head movements	17	80	0	–		
Pain sounds	38	72	1	0	50	0
Reaction to observer	58	80	2	100	44	1
Reaction to palpation	54	80	2	50	83	2
EQUUS-FAP						
Head	75	96	3	14	73	0
Eyelids	58	52	2	14	100	0
Focus	58	100	2	40	60	1
Nostrils	58	40	1	50	83	2
Lips	71	64	2	50	91	2
Muscle tone	29	92	1	0	80	0
Flehmen	0	96	0	–		–
Teeth grinding	29	100	1	67	0	0
Ears	71	100	2	50	73	2

CT, conservatively treated animals; STE, surgically treated or euthanased animals.
[a] Where sensitivity or specificity is ≤25%, the weighting factor = 0; 25–50%, weighting factor = 1; 50%–75%, weighting factor = 2; ≥75%, weighting factor = 3.

reduction of the number of parameters in both the EQUUS-COMPASS and EQUUS-FAP while optimizing sensitivity and specificity with the most condensed composite pain scales.

Conclusions

The use of the EQUUS-COMPASS and the EQUUS-FAP improved objectivity in pain assessment of horses with acute colic. Because of the high inter-observer reliability over a range of severity of pain scores, these pain scales enabled comparisons between different observers and so offered the possibility of monitoring the pain status of a patient more objectively over a period of time. This is a major asset in larger veterinary hospitals where often several attending clinicians are involved in the treatment of a single patient and will also offer possibilities for non-veterinarians to participate in evaluations.

Conflict of interest statement

This study was partly funded by Boehringer Ingelheim BV, but this played no role in the study design, in the collection, analysis and interpretation of data, or in the manuscript writing or submission for publication. None of the authors has any other financial or personal relationships that could inappropriately influence or bias the content of the paper.

Acknowledgements

We would like to thank Marianne Beikes, Larissa Vermin, Anne-Marie Sas and Marlijn van Gent for their help with data collection and René van Weeren for critical reading of the manuscript.

References

Abbott, F.V., Franklin, K.B.J., Westbrook, R.F., 1995. The formalin test: Scoring properties of the first and second phases of the pain response in rats. Pain 60, 91–102.
Ahola Kohut, S., Pillai Riddell, R., Flora, D.B., Oster, H., 2012. A longitudinal analysis of the development of infant facial expressions in response to acute pain: Immediate and regulatory expressions. Pain 153, 2458–2465.
Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 8, 307–310.

Bussières, G., Jacques, C., Lainay, O., Beauchamp, G., Leblond, A., Cadoré, J.L., Desmaizières, L.M., Cuvelliez, S.G., Troncy, E., 2008. Development of a composite orthopaedic pain scale in horses. Research in Veterinary Science 85, 294–306.
Cong, X., McGrath, J.M., Cusson, R.M., Zhang, D., 2013. Pain assessment and measurement in neonates. Advances in Neonatal Care 13, 379–395.
Dalla Costa, E., Minero, M., Lebelt, D., Stucke, D., Canali, E., Leach, M.C., 2014. Development of the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine castration. PLoS ONE 9, e92281.
Dobromylskyj, P., Flecknell, P.A., Lascelles, B.D., Livingston, A., Taylor, P., Waterman-Pearson, A., 2000. Pain assessment. In: Flecknell, P.A., Waterman-Pearson, A. (Eds.), Pain Management in Animals. W.B. Saunders, Philadelphia, PA, USA, pp. 53–79.
Gleerup, K.B., Forkman, B., Lindegaard, C., Andersen, P.H., 2014. An equine pain face. Veterinary Anaesthesia and Analgesia 42, 103–114.
Graubner, C., Gerber, V., Doherr, M., Spadavecchia, C., 2011. Clinical application and reliability of a post abdominal surgery pain assessment scale (PASPAS) in horses. The Veterinary Journal 188, 178–183.
Grulke, S., Olle, E., Detilleux, J., Gangl, M., Caudron, I., Serteyn, D., 2001. Determination of a gravity and shock score for prognosis in equine surgical colic. Journal of the Veterinary Medical Association 48, 465–473.
Hansen, B.D., 2003. Assessment of pain in dogs: Veterinary clinical studies. ILAR Journal 44, 197–205.
Hawker, G.A., Mian, S., Kendzerska, T., French, M., 2011. Measures of pain. Arthritis Care & Research 63, S240–S252.
Hellyer, P.W., 2002. Treatment of pain in dogs and cats. Journal American Veterinary Medical Association 221, 212–215.
Lindegaard, C., Thomsen, M.H., Larsen, S., Andersen, P.H., 2010. Analgesic efficacy of intra-articular morphine in experimentally induced radiocarpal synovitis in horses. Veterinary Anaesthesia and Analgesia 37, 171–185.
Malaty, H.M., Abudayyeh, S., O'Malley, K.J., Wilsey, M.J., Fraley, K., Gilger, M.A., Hollier, D., Graham, D.Y., Rabeneck, L., 2005. Development of a multidimensional measure for recurrent abdominal pain in children: Population-based studies in three settings. Pediatrics 115, e209–e215.
Myles, P.S., 2007. Using the Bland-Altman method to measure agreement with repeated measures. British Journal of Anaesthesia 99, 309–311.
Pritchett, L.C., Ulibarri, C., Roberts, M.C., Schneider, R.K., Sellon, D.C., 2003. Identification of potential physiological and behavioural indicators of postoperative pain in horses after exploratory celiotomy for colic. Applied Animal Behaviour Science 80, 31–43.
Prunier, A., Mounier, L., Le Neidre, P., Leterrier, C., Mormède, P., Paulmier, V., Prunet, P., Terlouw, C., Guatteo, R., 2013. Identifying and monitoring pain in farm animals: A review. Animal: An International Journal of Animal Bioscience 7, 998–1010.
Robertson, S.A., 2002. What is pain? Journal American Veterinary Medical Association 221, 202–205.
Sotocinal, S.G., Sorge, R.E., Zaloum, A., Tuttle, A.H., Martin, L.J., Wieskopf, J.S., Mapplebeck, J.C., Wei, P., Zhan, S., Zhang, S., et al., 2011. The Rat Grimace Scale: A partially automated method for quantifying pain in the laboratory rat via facial expressions. Molecular Pain 7, 55.
Sutton, G.A., Dahan, R., Turner, D., Paltiel, O., 2013a. A behaviour-based pain scale for horses with acute colic: Scale construction. The Veterinary Journal 196, 394–401.
Sutton, G.A., Paltiel, O., Soffer, M., Turner, D., 2013b. Validation of two behaviour-based pain scales for horses with acute colic. The Veterinary Journal 197, 646–650.
van Loon, J.P.A.M., Back, W., Hellebrekers, L.J., van Weeren, P.R., 2010. Application of a composite pain scale to objectively monitor horses with somatic and visceral pain under hospital conditions. Journal of Equine Veterinary Science 30, 641–649.
van Loon, J.P.A.M., Jonckheer-Sheehy, V.S.M., Back, W., van Weeren, P.R., Hellebrekers, L.J., 2014. Monitoring equine visceral pain with a composite pain scale score and correlation with survival after emergency gastrointestinal surgery. The Veterinary Journal 200, 109–115.
Valverde, A., Gunkel, C.I., 2005. Pain management in horses and farm animals. Journal Veterinary Emergency and Critical Care 15, 295–307.
Wagner, A.E., 2010. Effects of stress on pain in horses and incorporating pain scales for equine practice. Veterinary Clinics of North America. Equine Practice 26, 481–492.
White, N.A., Elward, A., Moga, K.S., Ward, D.L., Sampson, D.M., 2005. Use of web-based data collection to evaluate analgesic administration and the decision for surgery in horses with colic. Equine Veterinary Journal 37, 347–350.