Facial Expression Pain Scale (FEPS) for donkeys:

Improving the FEPS for donkeys with acute pain at 'The Donkey Sanctuary' Sidmouth, England

A.I. Smalbroek
Studentnr. 4250850
Master Thesis, University of Utrecht, Faculty of Veterinary Medicine
October 2018

Supervisors:
J.P.A.M. van Loon
M.C. van Dierendonck

Contents

Abstract	3
Introduction	4
Material and methods	
Animals	6
The facial expression pain scale	6
Data collection	8
Data processing and statistical analysis	8
Results	10
Discussion	12
Conclusion	19
Literature	20
Appendices	23

Abstract

Assessment of pain in horses is difficult and can only be evaluated by changes in behaviour and physiology as they cannot verbally express their pain. Enough data is collected about facial expressions of pain in horses but there is a hardly any information available about obvious pain expressions shown by the donkey. This study describes the inter-observer reliability, differentiation between patients and control donkeys and condition related behaviour to further develop and improve the Facial Expression Pain Scale (FEPS) for donkeys in acute pain at 'The Donkey Sanctuary'.

This study was performed using 159 donkeys (n=44 patients, n=115 controls). The facial expression pain scores were assessed by 2 observers and followed over time. The patients were categorised by specific types of pain: post-operation pain (n=6), facial pain (n=7), colic (n=7) and lameness (n=24). Cut-off value for the FEPS was determined at 1 to obtain maximal differentiation between patients and healthy control donkeys.

The FEPS showed good inter-observer reliability (ICC=0.77 $P<0.001$) and the FEPS could differentiate between patients and control donkeys ($P<0.001$). Age, gender, size and social relations were taken into account in choosing the control donkeys, matching the patient's characteristics but no significant differences were found between the specific control groups and the total control group.

A statistical analysis between a specific group of patients and their specific control donkeys showed different results for every subgroup of patients: colic $P<0.001$, facial pain $P<0.01$, lameness $P<0.001$. No significant differences were found between donkeys with post-operative pain and control donkeys ($P=0.248$). Internal sensitivity and specificity were good for the FEPS for the total patient group (sensitivity 68% and specificity 75%).

The FEPS offers an effective and reliable method of assessing various types of acute pain in donkeys. Further studies are necessary to increase the patient data, specifically patients after more invasive surgery, to improve the FEPS even more.

Introduction

Assessment of pain in horses is difficult and can only be evaluated by changes in behaviour and physiology as they cannot verbally express their pain. With donkeys we face an even bigger challenge as they are more stoic and show only subtle changes in their behaviour when in pain (1).

Developing methods for valid and reliable pain recognition have been subject of intensive study, like the visual analogue scales, numerical rating scales, composite pain scales and facial expression based pain scales (2). These pain scales must be reliable, consistent and applicable in various clinical settings.

Facial expressions have been investigated in all kinds of animals. Different facial expression-based pains scales have been developed for mice, rats and piglets and the results show that these scales could prove to be a sufficient means of assessing pain (3–5). Human studies also created facial pain scales for infants or humans who are not able to communicate (6–8). Data collected from these studies may provide important evidence to use these facial expression scales as assessment tools that are valid and practical.

In the last several years, especially facial features were studied in relation to the intensity of pain. Development of 'grimace' or 'facial expression' pain scales can show a more detailed recognition of subtle or early pain. But there are numerous other advantages these facial expression pain scales can bring. For example the fact that humans instinctively tend to focus on the face and head when observing pain in animals and people (9,10).

The Horse Grimace Scale (HGS) was described by Dalla Costa et al. (11) for horses undergoing castration. The HGS consists of 6 facial action units (FAUs): Stiffly backward ears, orbital tightening, tension above the eye area, prominent strained chewing muscles, mouth strained and pronounced chin, strained nostrils and flattening of the profile. The last two expressions combined are considered to be one facial action unit. Each FAU could be scored by using a 3-point scale with zero meaning that the FAU is not present, one meaning that the FAU is moderately present and two meaning that the FAU is obviously present. The HGS score was significantly higher in horses undergoing the routine castration (including analgesics treatment) compared to the control group. The HGS showed to be a reliable and effective method for pain assessment in this type of pain. They then applied their HGS to horses admitted to a clinic with acute laminitis, as the method currently used (the Obel grading system) requires the observation of the moving and therefore is it likely to cause further pain (12). The HGS again showed to be a potentially effective method for this specific painful condition.

Another equine facial expression pain scale, named the Equine Pain Face, was developed by Gleerup et al. (13). These authors describe detailed facial cues during application of noxious stimuli. During the pain stimuli, all horses in this study displayed specific changes in their facial expression such as asymmetrical/low ears, orbital tightening, square-like dilated nostrils, a tense stare, tension of the muzzle and mimic muscles.

Recently, the Equine Utrecht University Scale for Facial Assessment of Pain (EQUUS-FAP) was developed based on facial expression characteristics (14). This EQUUS-FAP consists of 9 parameters: movement of the head, tightening of the eyelids, focus on the environment, positioning corners mouth/lips, muscle tone of the head, flehming/yawning, teeth grinding and positioning of the ears. With each parameter to be scored, will be scored from 0 to 2, this results in a total pain score ranging from 0 to 18 with 0 meaning no signs of pain and 18 meaning maximal achievable pain. Studies show that the EQUUS-FAP proved to be an objective tool for the assessment of horses with acute visceral and head-related pain (14–16). The EQUUS-FAP effectively distinguished between controls and horses with these types of pain. It allowed reliable and reproducible assessment of visceral and head-related pain over time and demonstrated a high inter-observer reliability over a range of low and high pain scores.

Enough data is collected about facial expressions of pain in horses, however there is a hardly any information available about obvious pain expressions shown by the donkey. It is long known that donkeys and horses differ greatly. Not only in physiology and behaviour but also in showing pain (1). Donkeys tend to show more subtle behavioural changes and maybe even have a higher tolerance level even though there is no difference in cortical processing (17). The donkey is a prey species and will mostly run away from predators as horses do but will also sometimes stay and engage their fight response. Donkeys often live on their own which makes that individual animal the only target for the predator and fleeing is not an ideal choice when the donkey has a foal. This “fight” instinct is different from the horse that always tends to “flight”. Donkeys are less expressive in their behaviour when terrified so they tend not to panic as easily as horses (1,18). Therefore the donkey has evolved to minimise signs of pain to reduce a predator’s advantage while fighting instead of flighting (19).

Previously described pain scales were only applied in horses, but it is now clear that it is not valid to compare a donkey with a small horse. Therefore, there is a need to adapt these pain scales to donkeys and after that validate them for the use in donkeys. This could contribute to an improved and more effective recognition and treatment of pain in donkeys and therefore, improve the welfare of donkeys.

The aim of this study is 1) to modify and optimize the previously developed facial expression-based pain scales from preceding students for assessment of acute pain in donkeys. 2) To investigate if the facial expression pain scale can significantly differentiate between control donkeys and patients. 3) To assess the contribution of individual parameters to total pain score to see which parameters are most relevant for donkeys. 4) To assess the inter-observer variability of the facial expression pain scale. 5) To assess the monitoring of pain status over a period of time. The overall goal is to develop a facial expression-based pain scale for donkeys to improve recognition of pain and therefore improve donkey welfare.

The hypothesis for this study is that the Facial Expression Pain Scale (FEPS) would be able to differentiate between patients and controls and have a strong inter-observer reliability and would show a good development of scores over time during hospitalisation and analgesic treatment.

Materials & methods

Animals
Data was collected from October till November 2017 at the Veterinary Department of The Donkey Sanctuary, Sidmouth, for six weeks. A total of 29 donkeys with acute pain from various origins have been observed at The Donkey Sanctuary. The patients were categorised by specific equine pain types: post-operation pain, facial pain, colic and lameness. The obtained pain scores had no influence on decisions about the treatment protocol and were used for documentary purposes only. We aimed to have two or three control donkeys per patient. One control donkey is the bonded companion of the patient, one control donkey is from the same barn as the patient and the third donkey is from another barn. Control donkeys had to be free of any clinical problems and not on pain medication. Age, gender, size and social relations were taken into account in choosing the control donkeys, matching the patient's characteristics. A total of 78 control donkeys have been observed.
All the data collected by previous students was added to obtain a larger dataset. This previously collected dataset consisted of another 19 patients and 49 control donkeys. After reviewing the data, 4 patients and their control donkeys were excluded from the dataset, because they were on pain medication before the start or cause of their pain was found. Tables 1 and 2 show the data of the donkeys included in this study.

Table 1 The animals used in this study

	Patients (n=44)	**Controls (n=115)**
Age (years)		
Mean	16,75	16,47
Min-Max	1-37	0-35
Gender		
Stallions	3	1
Geldings	27	72
Mares	14	42
Condition		
Lameness	24	-
Colic	7	-
Facial pain	7	-
Postop pain	6	-
Total number	**44**	**115**

The facial expression pain scale.
The FEPS was developed by Daja van Nunen for her master thesis. She investigated a lot of data about normal donkey behaviour, interviewed specialists with experience in donkeys and researched studies of other facial pain scales. Last year, two veterinary master students, tested the facial expression pain scale for donkeys in patients with acute pain at the 'The Donkey Sanctuary' in Sidmouth, England. The FEPS for donkeys consists of 12 categories, named: head, eyelids, focus, nostrils, corners mouth/lips, facial muscles, flehming/yawning/smacking, teeth grinding and/or moaning, ear response, ear position, startle/headshaking and sweating behind the ears. Table 3 shows the FEPS for donkeys. The highest score in each category is two and the lowest score is zero. Most categories can be scored with a 0, 1 or 2. Some categories can only be scored in a binary fashion, with a 0 or 2, as in those cases it can only be decided whether the behaviour is normal or abnormal with no in between. For example: the ear position is either normal or abnormal. When all the scores are added together, the total score can range from 0, being no signs of pain, to a maximum of 24.
With the outcomes and remarks from the previous students, the FEPS had not been changed itself but needed more patients and data to improve the inter-observer reliability and validity. One of their adjustments would be to exclude some of the elements. It was decided to not exclude these elements but test them again with an increased number of patients.

Table 2 Patient data

Patient	*Name*	*Gender*	*Age*	*Medical condition*	*Patient group*
1	Sparkle Walker	Jenny / mare	27	Solar abcess	Lameness
2	Dolly KS	Jenny / mare	23	Laminitis	Lameness
3	Jack Tidball	Gelding	21	Mouth ulcerations	Facial pain
4	Zebedee Montague	Gelding	23	Solar abcess	Lameness
5	Curry Walsh	Jack / stallion	1	Castration	Postoperation pain
6	Ralph Walsh	Jack / stallion	1	Castration	Postoperation pain
7	Beauty Walks	Jenny / mare	14	Solar abcess	Lameness
8	William Woodland	Gelding	22	Solar abcess	Lameness
9	Rosschap Eire	Gelding	8	Ostheoarthritis	Lameness
10	Archie Keevans	Gelding	17	Keratitis	Facial pain
11	Peter C	Gelding	28	Corneal ulceration	Facial pain
12	Patsey Eire	Gelding	25	Solar abcess	Lameness
13	Coco Culling	Gelding	31	Solar abcess	Lameness
14	Crackers Hall	Gelding	37	Laminitis and white line abcess	Lameness
15	Rosa Lewis	Jenny / mare	18	Ostheoarthritis	Lameness
16	Edward Hancock	Gelding	6	Solar abcess	Lameness
17	Tayto	Gelding	14	Conjunctivitis and uveitis	Facial pain
18	Mr McGregor Eire	Gelding	28	Impaction colic	Colic
19	Penny Starsmore	Jenny / mare	17	Solar abcess	Lameness
20	Paddy Stevens	Gelding	9	Impaction colic	Colic
21	Smokey Stiles	Gelding	18	Impaction colic	Colic
22	Jack F Eire	Gelding	23	Solar abcess	Lameness
23	Malty Eire	Gelding	26	Laminitis	Lameness
24	Ciara Eire	Jenny / mare	9	Solar abcess	Lameness
25	Cocoa Clews	Gelding	22	Impaction colic	Colic
26	Ganty Eire	Gelding	7	Laminitis	Lameness
27	Fourmay Ake	Jenny / mare	16	Solar abcess	Lameness
28	Sile M Eire	Jenny / mare	8	Conjunctivitis	Facial pain
29	Willie Drennan Eire	Gelding	9	Sarcoid removal	Postoperation pain
30	Oreo	Jack / stallion	1	Castration	Postoperation pain
31	Buzz MI	Jenny / mare	29	Other colic	Colic
32	Sandon Proud Playboy	Gelding	23	Molar removal	Postoperation pain
33	Violet Davies	Jenny / mare	25	Painful frog	Lameness
34	Charlie Chuck Devlin	Gelding	9	Solar abcess	Lameness
35	Abbie Eire	Jenny / mare	26	Other colic	Colic
36	Brodaha Eire	Jenny / mare	8	Solar abcess	Lameness
37	Jacko Buttle	Gelding	17	Hoofwall bruise	Lameness
38	Cobweb PSNI	Jenny / mare	5	Solar abcess	Lameness
39	Henry Gribben	Gelding	13	Corneal ulceration	Facial pain
40	Eey-ore Eire	Gelding	13	Unknown cause	Lameness
41	Rodney Barrett	Gelding	10	Other colic	Colic
42	Camalan Jester	Gelding	15	Fractured jaw	Facial pain
43	Bruno D Eire	Gelding	9	Sarcoid removal	Postoperation pain
44	Phoebe Wing	Jenny / mare	26	Solar abcess	Lameness

Data collection.
The donkeys were observed as soon as possible in the acute pain phase and preferably before the veterinarian had started the clinical examination or had administrated any analgesics. A few times it was not possible to obtain the pain scores before the veterinarian already started the clinical examination or before analgesic treatment. The observers would always let the donkey get adjusted to having people in its surroundings, by standing quietly and a few meters away from the donkey. The observers would then obtain the facial expression pain scores and composite pain scores. While collecting data for the FEPS, the Composite Pain Scale (CPS) was tested as well by another student. This CPS was also developed and improved by other previously students and now tested with similar aims by Julie Smolenaers. Because both pain scales were assessed simultaneously but independently by two students, the inter-observer reliability can be tested.
After assessing both the FEPS and the CPS the observers would catch the donkey, sometimes with the help of a groom, and put on his/her head collar.
At last, the observers obtained the physical parameters like heart rate, respiratory rate, rectal temperature and gut sounds. These parameters were included in the CPS.
After the observations, the donkey was filmed to obtain video footage which can be used for blind scoring, as education material or as a reference to verify the outcome of the observation. In this video, first the whole body of the donkey was filmed, then the head of the donkey was filmed and at last the donkey was filmed while walking. After the observations and filming were done, the observers would receive the VAS score from the treating veterinarian, whenever there was one present. The veterinarian would then do her clinical exam and give a diagnosis.
This whole protocol of collecting the data was replicated for the patients and for the control donkeys. For every patient, observations were performed for several days (if possible) to follow-up on the patients and to be able to assess possible effects over time. Control donkeys were observed only once. All observations were performed simultaneously with another student (Julie Smolenaers) to be able to assess the inter-observer reliability and therefore findings could not be discussed before, during and after the scoring. The observers knew the diagnosis and exact initiation of the treatment of the patient during the follow-up observations. Data was obtained before and the day after the start of the treatment of the patient. Accordingly, it can be known which observations are done with and without pain treatment.

Data processing and statistical analysis.
After observing the donkey and assessing the CPS and FAP scores, the Visual Analogue Score (VAS) was performed by the treating veterinarian if possible. All of the data collected by the previous students was added to obtain a larger dataset. When observing the donkey and assessing the pain scores, the scores were registered on paper sheets. An Excel spreadsheet was used to organize the collected scores on paper. In this spreadsheet, all pain scores were collected together with other characteristics of the donkey such as age, weight, location and information on the diagnosis, additional testing and treatment information. To collect the clinical data of the patients and control donkeys, the Donkey Sanctuary patient system has been accessed. Any important information such as medical status and medication was also registered in the Excel spreadsheet. The data of the patients consists of the mean score of the two observers at the first observation of the acute pain
The inter-observer reliability was assessed using the Spearman correlation analysis. The Mann-Whitney U test was used to analyse the differences between patients

and the control donkeys. Within the different pain groups (lameness, colic, facial pain and postoperative pain) the patients were compared to their specific control donkeys and the other control donkeys. Therefore, the Mann-Whitney U test was used as well.
Cut-off value for the FEPS was determined at 1 to obtain maximal differentiation between patients and healthy control donkeys. Both sensitivity and specificity were determined for every type of pain and for every individual parameter. Based on these values, weighting factors for the various parameters were determined. Statistical analysis was performed with SPSS Statistics version 24.0. Statistical significance was accepted at $p < 0.05$.

Table 3 The Facial Expression Pain Scale for Donkeys

Head	Normal movement Less/no or more/ exaggerated movement	0 2
Eyelids	Opened More opened eyes or tightening of eyelids Obviously more opened eyes or obvious tightening of eyelids	0 1 2
Focus	Focused on environment Less focused on environment Not focused on environment	0 1 2
Nostrils	Relaxed A bit more opened, nostrils lifted, wrinkles seen Obviously more opened, nostril flaring, possibly audible breathing	0 1 2
Corners mouth/lips	Relaxed Lifted	0 2
Muscle tone head	No fasciculation’s Mild fasciculation’s Obvious fasciculation’s	0 1 2
Flehming/yawning/smacking	Not seen Seen	0 2
Teeth grinding and/or moaning	Not been heard Heard	0 2
Ear response	Clear response with both ears or ear closest to source Delayed/reduced response to sounds No response to sounds	0 1 2
Ear position	Normal position Abnormal position (hang down/backwards)	0 2
Startle/headshaking	No startle/headshaking At least one startle (a sudden abrupt movement with the head as if suddenly aware of danger)/period of head shaking	0 2
Sweating behind the ears	No signs of sweating Signs of sweating	0 2
Total		/24

Results

Inter-observer reliability.
Figure 1 shows the correlation analysis between the different pain scores of two independent observers. The FEPS showed a good and significant correlation. The R-square for the FEPS was 0,77 ($p < 0.001$)
Patients vs control donkeys.
Figure 2 shows that when all patients are compared to all control donkeys, the FEPS scores of the patients are significantly higher compared to control donkeys ($P < 0.001$).

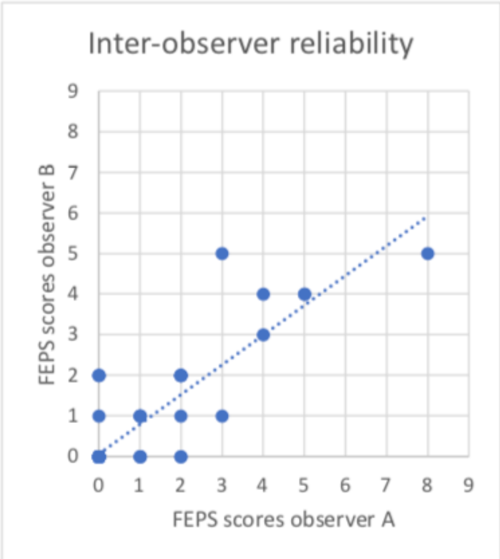

Figure 1 Inter-observer reliability of the FEPS scores. (n=130, R2 =0.77, P<0.001)

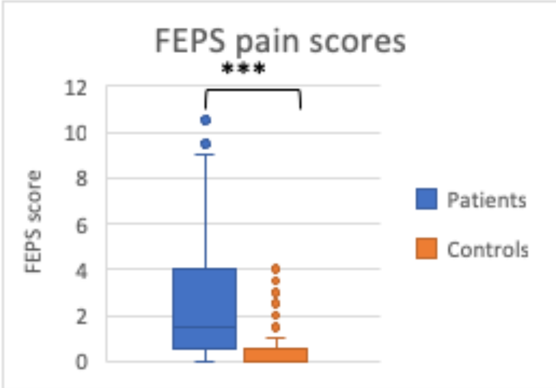

*Figure 2 Mean FEPS scores of all patients (n=44) versus all control donkeys (n=115). Lines in boxes show median scores;boxes show 25-75th percentiles;error bars show 5-95th percentiles; ***=P<0.001*

Patient subgroups vs specific control donkeys.
The same comparison was made for every subgroup of patients. No significant difference was found between the specific control donkeys and the other control donkeys. All subgroups of patients were compared to their specific control donkeys. A statistical analysis between a specific group of patients and their specific control donkeys showed good results for 3 out of the 4 subgroups of patients: colic $P<0.001$, facial pain $P<0.01$ and lameness $P<0.001$. No significant differences were found between donkeys with post-operative pain and control donkeys: $P=0.248$ (NS). Figure 3 shows the boxplots with the results for each subgroup of patients with their specific and total control groups.

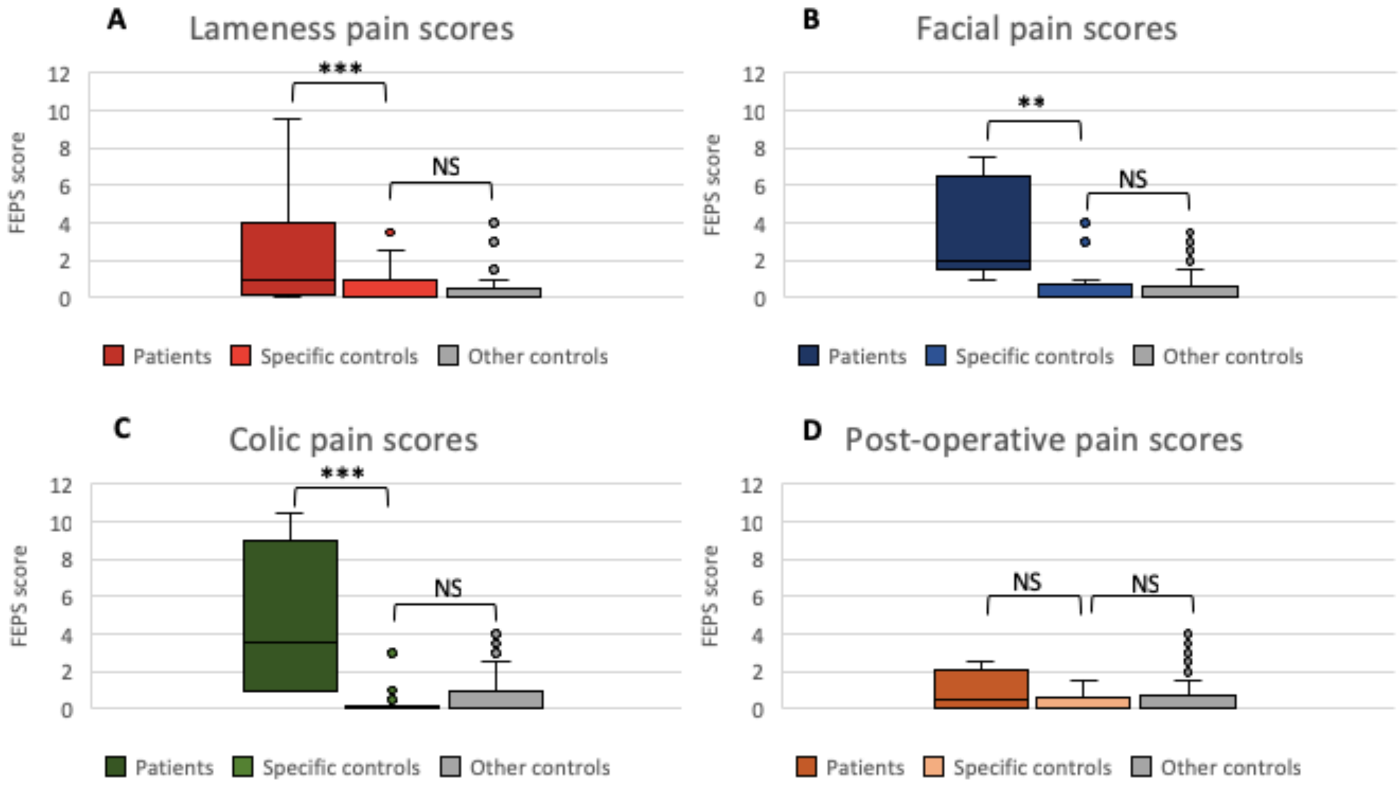

*Figure 3 Mean FEPS scores for lameness (n=24) versus specific control donkeys (n=66) and the other control donkeys (n=49) (A), facial pain (n=7) versus specific control donkeys (n=17) and the other control donkeys (n=98) (B), colic (n=7) versus specific control donkeys (n=18) and the other control donkeys (n=97) (C) and postoperative pain (n=6) versus specific control donkeys (n=14) and the other control donkeys (n=101) (D). Lines in boxes show median scores; boxes show 25-75th percentiles; error bars show 5-95th percentile; ***=P<0.001, **=P<0.01, NS=not significant*

Effects over time.
Figure 4 shows the FEPS scores over time of the patients for every type of pain.

Internal sensitivity and specificity of FEPS and their individual parameters.
A cut-off value of 1 was used to obtain all the sensitivity and specificity. Appendix 1 shows the internal sensitivity, specificity, negative predictive value and positive predictive value for the data collected in 2016 and 2017 and these two datasets combined. It also shows these results for the data without post-operative pain patients because no significant difference was found between these patients and the control donkeys. Appendix 2 shows the sensitivity, specificity, negative predictive value and positive predictive value for every subgroup of patients. Appendix 3 shows the sensitivity and specificity for every element of the FEPS in the overall dataset. The sensitivity and specificity for the individual parameters of the FEPS for every subgroup of pain is shown in Appendices 3-7.

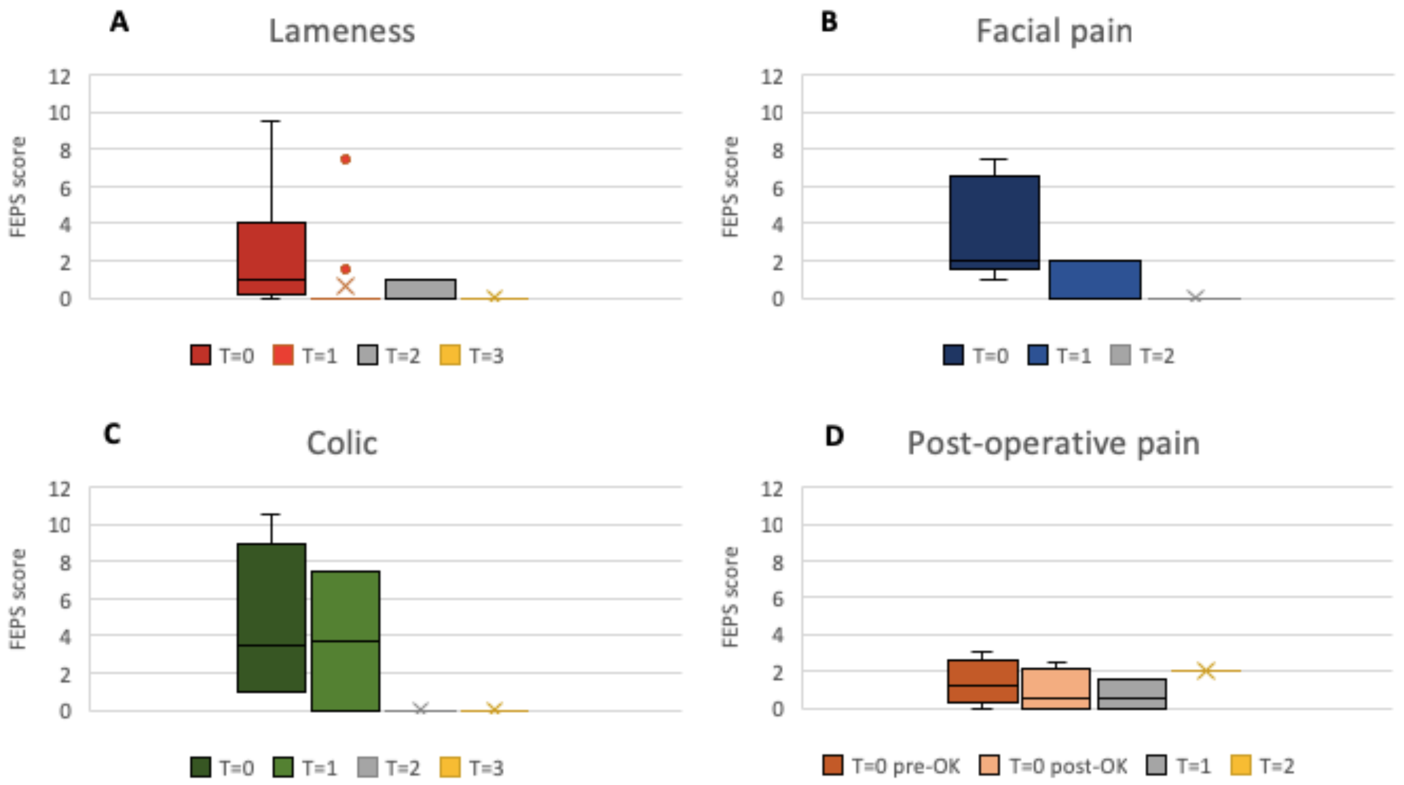

Figure 4 Mean FEPS scores for lameness at the first observation (T=0) (n=24) versus the first day after the first observation (T=1) (n=14) versus the second day after the first observation (T=2) (n=6) versus the third day after the first observation (T=3) (n=5) (A), facial pain at the first observation (T=0) (n=7) versus the first day after the first observation (T=1) (n=5) versus the second day after the first observation (T=2) (n=1) (B), colic at the first observation (T=0) (n=7) versus the first day after the first observation (T=1) (n=2) versus the second day after the first observation (T=2) (n=2) versus the third day after the first observation (T=3) (n=1) (C) and post-operative pain before the operation (T=0 pre-OK) (n=4) versus after the operation (T=0 post-OK) (n=6) versus the first day after the operation (T=1) (n=3) versus the second day after the operation (T=2) (n=1) (D). Lines in boxes show median scores; boxes show 25-75th percentiles; error bars show 5-95th percentiles.

Discussion

The FEPS for donkeys was tested in this study in patients with acute pain and control donkeys. An inter-observer reliability correlation coefficient of 0.77 was found. The FEPS for donkeys showed significant difference between patients and control donkeys. This shows that the FEPS is able to assess pain in donkeys and differentiate between donkeys with pain and their healthy control donkeys.
Dividing the patients in different subgroups of pain, showed significant difference between those specific patients and their specific control donkeys for colic pain, facial pain and lameness pain. There was no significant difference between patients with post-operative pain and their control donkeys. Good sensitivity and specificity in the overall dataset show that the FEPS is able to differentiate between patients and control donkeys.
Van Loon and van Dierendonck (14–16) also concluded that the EQUUS-FAP is a reliable assessment of pain in horses with acute visceral pain and head-related pain. These findings matched with our findings for donkeys with colic and facial pain. Another pain scale was developed for geldings after castration (11). This Horse Grimace Scale (HGS) showed to be a reliable and effective method for pain assessment in this type of pain. However, in the current study there were no significant results for patients undergoing operations like castrations. A reason could be that this group of patients was very

small (n=6), with only 3 patients undergoing castrations even though they showed higher FEPS scores after the surgery. More patient data could improve the effectiveness and reliability of this subgroup of patients. Secondly, a reason could be that the post-operative pain in these cases was suppressed by the adequate pain relief or that the operations weren't invasive to an extent for which the pain was measurable with the FEPS for donkeys. This could then lead to the conclusion that the patients received a very effective analgesic treatment. The patients in the HGS study did not receive an intra-testicular block with lidocaine but all patients in the current study did receive an intra-testicular block with lidocaine. A study by Sanz et al. (20) showed that it is possible that this intra-testicular administration of lidocaine could provide a very potent perioperative analgesia. This could be another reason why no significant results were found for patients undergoing operations like castrations. Further studies should ideally include patients that undergo more invasive operations. Even though this study did not show significant increases in pain scores after castration, more data could be collected to see if the FEPS for donkeys is an effective method to assess pain in donkeys undergoing castration (or other more invasive surgeries).
Another remarkable finding is that the observation before the surgery showed high FEPS scores. A possible explanation could be that most of these operations were castrations of young stallions. The pre-operation scores could therefore be due to the more stressed behaviour of this type of donkey. These patients mostly showed open nostrils and headshaking. This likely correlated to the excitation rather than to pain. The control donkeys used for these patients undergoing castrations were mostly Jenny's as there were no other stallions housed at the Donkey Sanctuary. Open nostrils and headshaking were not seen in the control donkeys for this type of pain. This supports the thought that the stallions' more stressed behaviour influenced the FEPS scores. This should be tested in further studies.
Another important finding of this research is that it is not necessary to use a specific control group since no significant differences were found between specific control donkeys and the other control donkeys. This information could be used to adjust the protocol in further studies.
This study showed that the FEPS for donkeys is able to assess acute pain in donkeys. However, some elements of the FEPS were never or only once scored. Firstly, the parameter 'pain sounds' was not seen or heard for our patients nor for the control donkeys. The grooms did address this 'teeth grinding and/or moaning' as heard before the observations started but because it was never heard by the observers themselves, it was never scored in this study. Other studies show that 'pain sounds' are a very important parameter in horses (21) but in our study it was never scored in any donkey patient. It is important to see more patients in order to determine if donkeys in pain show this type of behaviour. Secondly, the element 'muscle tone of the head' was never scored in the FEPS for donkeys. Again this is a very important parameter shown by horses experiencing pain or fear (11–13) but is not described in donkeys (19). Ashley et al. (22) even discarded tension of facial muscles as a behaviour element because it did not meet the criteria of exclusivity in a study focussing on working donkeys. Maybe it is not possible to see this muscle tone of the head because of the thick fur donkeys have on their heads. Again, it is important to see more patients in order to determine if donkeys show this type of behaviour.
The element 'Flehmen, yawning, smacking' was scored almost as many times in control donkeys (n=8) as in patient donkeys (n=11). This gave one of the lowest specificities in the overall dataset compared to the other elements.

(appendix 3) Contradictory, 'flehmen' had one of the highest sensitivities of all elements. This element is highly influenced by the environment. Many donkeys showed flehming when a mare was walking by. Therefore, it is difficult to conclude that this element is related to pain in donkeys. More donkeys with colic pain should be assessed as flehming is an important parameter in horses with abdominal pain (23). A reason for this could be that horses are housed alone and therefore are less influenced by their environment. Flehming is then mostly shown as a sign of pain instead of sexual behaviour.

The element 'sweating behind the ears' was only seen once in a control donkey. The study by Ashley at al. (19) suggested that because of the stoical behaviour of the donkey, they will rarely sweat or roll like a horse would in an acute stage of colic. This sweating might only be seen in the end-stage of diseases like colic. A few donkeys in this study were diagnosed with a terminal stage of colic and were put to sleep after the first observation. Still, none of these patients showed 'sweating behind the ears'. A next study should include more severe colic patients to determine if sweating behind the ears is an important parameter in donkeys.

In this study, an inter-observer reliability correlation coefficient of 0.77 was found. This is a good inter-observer reliability that indicates that the FEPS for donkeys is mainly independent of the observer. However, this obtained value is not as good as the excellent inter-observer reliabilities found in other equine facial pain scales like the HGS (0.92 and 0.85) (11,12) and the EQUUS-FAP (0.93) (14,16) but it seems promising that the FEPS for donkeys can reach these levels of reliability when more data is collected and when improvements for the FEPS for donkeys are applied. Because humans instinctively tend to focus on the face and head when observing pain in other animals and people (9,10), using facial expression

Figure 5 Patient 17, Rosa Lewis, showing lifted corners of the mouth. FEPS score was 4, including lifted corners of the mouth, less focus and an abnormal ear position.

pain scales could be a reliable method to assess acute pain.

Unfortunately, the VAS score that should have been obtained for every patient, was not analysed in this study. Most of the time the observers had to assess the donkeys when no veterinarian was present, or the VAS score was obtained after the veterinarian had already done the clinical examination on the donkey which has an influence on the objectivity of the VAS score. When the veterinarian has diagnosed the patient, they would have too much information compared to a veterinarian that has not yet examined the patient. This could give very different VAS scores. Literature describes that the VAS is not an ideal scoring technique for the assessment of pain in equines due to moderate (ICC = 0.63) and fair (ICC=0.34) found reliabilities (14,24,25).

Donkeys seem to only start showing facial expressions of pain when the pain is quite severe. When the obtained Composite Pain Score was high, only then the donkey also showed elements of the FEPS. But this does not apply to facial pain. All donkeys in the facial pain group showed one or more elements of the FEPS. Oppositely the CPS scores were rather low. It seems plausible that when a donkey has facial pain, it shows more in its facial expression pain scale. In conclusion, the results of this study show that the FEPS for donkeys is better applicable to facial pain and colic pain then for lameness and post-operative pain. Van Loon and van Dierendonck (2015, 2016, 2017) also concluded that their EQUUS-FAP proved useful for the assessment of horses with head-related and visceral pain (14–16).

In future studies, it could be possible to adapt the general FEPS to specific FEPS for different types of pain or to use weighing factors. If there is condition related behaviour, it should be seen in the scoring frequency in the FEPS elements. The lameness patient group mostly showed an abnormal ear position, flehming/yawning/smacking and opened nostrils. Patients that were suffering from lameness, were the only patient group that showed lifted corners of the mouth. An example of lifted corners of the mouth is shown in figure 5. The two patients that showed lifted corners of the mouth, had high overall FEPS scores which could mean that this element is only shown when the pain is quite severe. This is also supported by the fact that this element was never shown by any control donkey from every subgroup of patients. Because there were only two patients that showed lifted corners of the mouth, it is too early to conclude that this element could be specific for lameness. The Horse Grimace Scale by Dalla Costa et al. (12) showed good reliability for this element in horses with acute laminitis. Also, the elements opened nostrils' and 'abnormal ear position' proved good reliability in this type of pain, which was also frequently seen in our lameness patients. Especially the 'abnormal ear position' proved a very good reliability. This element had the highest sensitivity of all elements in this type of pain in our patients (appendix 4). This element is also very important in horses with laminitis (12) and donkey literature (1) also concluded that a lower/backward ear carriage is a sign of pain. One study investigated the behaviour of donkeys in Kenya, India and Pakistan (22). These donkeys were living under poor conditions and held their ears in a very low/backward position compared to donkeys living in the UK, who held their ears in a more forward position. Figure 6 shows a patient with an abnormal ear position.

Figure 6 Patient 25, Jack F, showing an abnormal position of the ears. FEPS score was 4, including abnormal position of the ears and flehming.

Literature shows that flaring nostrils might be a useful element in assessment of pain in donkeys and other species (3,7,11,12,14,19,26). Even though the element 'opened nostrils' was scored many times in lameness patients (n=7), it was also frequently scored in their specific control donkeys (n=9). But this element was only scored in control donkeys by the previous students that conducted this study as well. This could have a big influence on the validity of the FEPS for donkeys. An explanation could be that they scored the element 'open nostrils' because of the wrinkles that they might have seen. They also conducted this study in a warmer period of the year, which could also have contributed to scoring this element more frequently. In a future study it is important to make a better differentiation between 'open nostrils' and 'wrinkles seen'. Seeing wrinkles does not explicitly mean more opened nostrils. Excluding this part of 'wrinkles seen' could therefore help improve the validity of the FEPS for donkeys.

The facial pain patient group mostly showed tightening of the eyelids, less movement of the head and less focus on the environment, whereas tightening of the eyelids was most frequently seen in patients with eye problems. Orbital tightening is an important parameter in horses undergoing castration and in horses with acute laminitis or colic pain (11,12,14). But no research is available about orbital tightening in horses with facial pain. A study by Burden and Thiemann (1) showed that excessive lacrimation, rubbing of eyes, and blinking are indicative of eye issues in donkeys but are influenced by weather conditions and fly irritation. When the sun is shining bright or when donkeys are sleeping, they could show tightening of the eyelids. Therefore, it could sometimes be scored in control donkeys as well as in patient donkeys. It is difficult to exclude these false positive observations even though this element does differentiate between patient and control donkeys. Also, it is difficult to determine the degree of orbital tightening. Many times, the two observers scored the degree of orbital tightening (1 or 2 points) differently. Even though it is a difficult element to be asses, the element tightening of the eyelids seems to be an important element of the FEPS for donkeys. A suggestion could be to downsize the scoring options into 2 instead of 3. The eyelids are then scored opened (0 points) or more opened/tightened (2 points) instead of opened (0 points), more opened/tightened (1 point) or obviously more opened/tightened (2 points). The degree of orbital opening/tightening is then no longer a point of discussion.

When scoring the element 'head movement', no observation was made of more/exaggerated head movement. Only normal or less/no head movement was observed in the donkey patients. More head movement was never scored. This behaviour can be seen when a donkey suffers from trigeminal neuralgia, but this was never diagnosed in the current study.

The colic patient group mostly showed tightening of the eyelids, less focus, opened nostrils and an abnormal ear position. The element 'less focus' could be a very good parameter for pain in donkeys. Dullness, including this decrease in focus, is frequently described in literature as an indicator of pain in donkeys (1,22). They show less interest in the environment. In general, literature says that dullness and a depressed appetite are most frequently observed signs of abdominal pain (27,28). But it is very important to know an individual donkey's normal behaviour and attitude in order to see this change in its behaviour.

The post-operative patient group mostly showed flehming/yawning/smacking. This was observed in two donkey patients and they were the only two patients that showed a sign of pain behaviour at all. One of these patients had undergone a castration and the other patient had undergone a sarcoid removal. In this study, most operations were castrations. This contributes to the outcome of the study as all of those patients were very young stallions. Their excited behaviour, caused by their young age and getting adjusted to a new home, had an influence on the elements shown by the patients and control donkeys. In future studies, it is very important to assess more invasive operations.

The element 'startle/headshaking' was not specifically seen in one of the subgroup of patients. It can be seen in certain types of head related pain (like trigeminal neuralgia) but in our study, it was not obviously seen in this type of patients. The donkeys that showed a startle/headshake

had very high overall FEPS scores. Therefore, it could be concluded that this element is only shown by donkeys when they suffer from severe pain. When a startle was visible, it was very subtle and difficult to see. When donkeys were sleeping, they also could have shown a startle. It is very important to differentiate between snoozing donkeys that show a startle and donkeys that are awake and show a startle from pain. Also, equines can show headshaking as a reaction to seasonal and climate changes. Horses show headshaking more frequently in spring and summer (29). Even though this element could be influenced by the environment and by the state of consciousness of the donkey, the element showed to be a relevant indicator of acute pain in donkeys.
The element 'ear response' also was not specifically seen in just one group of patients. Literature (1, 22) shows that the movement of the ears is a very important indicator of pain in donkeys. When donkeys show no or less movement of the ears in response to noise changes, they are feeling very uncomfortable (1). A study by Ashley et al. (22) showed that healthy donkeys, even whilst sleeping, hardly kept their ears still, constantly moving between forward and sideways positions. No or less movement of the ears was never shown by any of our control donkeys but was shown several times in patients. This implies that this element can differentiate well between control donkeys and patients. Therefore, this is a very important element in the FEPS for donkeys

In this study, it was sometimes seen that when a donkey is sleeping, this had an influence on the observation. Elements such as 'eyelids', 'focus' and 'ear position/movement' can be assessed differently in donkeys that are sleeping than in donkeys that are awake. Therefore, it is possible to observe these elements in healthy control donkeys as well. Consequently, it is very important to wake the patient before starting the observation. This can be achieved by making some noise like clapping hands or to get their attention by showing that you have candy. Waking up the donkey is very important to ensure that the surroundings and the activity state of the donkey are the same for every observation. Also, donkeys tended to freeze when the observers entered the stable or came close to the donkey. This changed their behaviour, especially when the camera was set in place to film the donkey. For several minutes, the donkeys would show less behavioural elements. Therefore, it is necessary to wait a few minutes before starting the observation and to place the camera at an appropriate distance from the donkey such as 2 meters.

The environment could have an influence on the behaviour elements shown by the donkey. When donkeys are standing in mixed groups, males and females will react on each other and could therefore show an abnormal ear position or flehming. Temperature also has an influence. When it is a warm day, donkeys could show opened nostrils. It is very important to be aware of these influences when doing the FEPS assessment.

Because every donkey shows different normal behaviour when he/she is healthy, there are very big individual variations in signs of pain showed by the animal. Donkey owners should frequently measure the donkey's normal behaviour with the FEPS to know exactly when the donkey's behaviour changes.

Anxious or excitable animals can show more signs of pain compared to tranquil animals. Therefore, it is important that when these pain scales are used, the observations are made by its own veterinarian/caretaker/owner that is familiar with this specific animal and knows its normal behaviour. When this is not possible, it should always be kept in mind that every animal has a different personality and could show a complete different range of signs of pain (26,30). To reduce the amount of stress a donkey is experiencing, it is also very important to assess pain scores when a donkey is in his/her own environment, as a change in their environment can provoke stress and therefore influence the donkey's behaviour (18). Furthermore, it is important to observe a donkey only when his friend is present. Most donkeys have a very strong bond with another donkey or animal. When this friend is taken away from them, they can become stressed and/or dull. Again, this could influence their behaviour (1,31).

Conclusion

This study proved to be a successful study to test the Facial Expression Pain Scale for donkeys suffering from various types of acute pain. It showed a significant difference between donkey patients and control donkeys and had a very good inter-observer reliability. This is indicates that the FEPS for donkeys is a reliable tool to assess the acute pain in donkeys. The FEPS for donkeys was especially effective in patients suffering from colic, lameness and facial pain. For patients with post-operative pain, the FEPS for donkeys did not differentiate significantly between patients and control donkeys. A possible explanation could be that the FEPS is not effective enough in this type of pain or that this type of pain was not severe enough in these patients due to the protocol of pain relief or the invasiveness of the operation. In a next study, it is important to see more patients that undergo more invasive operations or that have a different protocol of pain relief. This study also showed that it was not necessary to have a specific control group for every type of pain. This could be used in the protocol of a follow-up study. At this point, there seems to be no indication to adjust the FEPS for donkeys. Collecting more data for the four types of pain would be beneficial and could improve the FEPS for donkeys even more. Only the element "corners mouth lifted" was specifically shown by one type of pain, lameness. But this is not indicative for pronounced condition related behaviour and therefore there is no indication to create specific FEPS's for every type of pain. It is important to always keep in mind the influence of the observers and the environment on the facial expression pain signs shown by the patient donkey.
Future research will improve the validity and reliability of the FEPS for donkeys
The FEPS for donkeys could ultimately be helpful for veterinarians, caretakers and owners to assess acute pain in donkeys more sufficiently.

Literature

1. Burden F, Thiemann A. Donkeys are different. J Equine Vet Sci. 2015;35(5):376–82.

2. de Grauw JC, van Loon JPAM. Systematic pain assessment in horses. Vet J. 2016;209:14–22.

3. Sotocina SG, Sorge RE, Zaloum A, Tuttle AH, Martin LJ, Wieskopf JS, et al. The Rat Grimace Scale: A Partially Automated Method for Quantifying Pain in the Laboratory Rat via Facial Expressions. Mol Pain. 2011;7:7–55.

4. Leach MC, Klaus K, Miller AL, Scotto di Perrotolo M, Sotocinal SG, Flecknell PA. The Assessment of Post-Vasectomy Pain in Mice Using Behaviour and the Mouse Grimace Scale. PLoS One. 2012;7(4):e35656.

5. Viscardi A V, Hunniford M, Lawlis P, Leach M, Turner P V. Development of a Piglet Grimace Scale to Evaluate Piglet Pain Using Facial Expressions Following Castration and Tail Docking: A Pilot Study. Front Vet Sci. 2017;4.

6. Hicks CL, Von Baeyer CL, Spafford PA, Van Korlaar I, Goodenough B. The Faces Pain Scale - Revised: Toward a common metric in pediatric pain measurement. 2001 ;93:173–83.

7. Arif-Rahu M, Grap MJ. Facial expression and pain in the critically ill non-communicative patient: State of science review. Intensive and Critical Care Nursing. 2010;26:343–52.

8. Ahola Kohut S, Pillai Riddell R, Flora DB, Oster H. A longitudinal analysis of the development of infant facial expressions in response to acute pain: Immediate and regulatory expressions. Pain. 2012;153(12):2458–65.

9. Leach MC, Coulter CA, Richardson CA, Flecknell PA. Are we looking in the wrong place? implications for behavioural-based pain assessment in rabbits (Oryctolagus cuniculi) and beyond? PLoS One. 2011;6(3).

10. Williams A. Facial expression of pain: An evolutionary account. Behav Brain Sci. 2002;25(4):439–88.

11. Dalla Costa E, Minero M, Lebelt D, Stucke D, Canali E, Leach MC. Development of the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine castration. PLoS One. 2014;9(3).

12. Costa ED, Stucke D, Dai F, Minero M, Leach MC, Lebelt D. Using the horse grimace scale (HGS) to assess pain associated with acute laminitis in horses (Equus caballus). Animals. 2016;6(8).

13. Gleerup KB, Forkman B, Lindegaard C, Andersen PH. An equine pain face. Vet Anaesth Analg. 2015;42(1):103–14.

14. van Loon JPAM, Van Dierendonck MC. 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. Vet J. 2015;206:356–64.

15. Van Dierendonck MC, van Loon JPAM. 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 validation study. Vet J. 2016;216(3):175–7.

16. van Loon JPAM, Van Dierendonck MC. Monitoring equine head-related pain with the Equine Utrecht University scale for facial assessment of pain (EQUUS-FAP). Vet J. 2017;220:88–90.

17. Grint NJ, Johnson CB, Clutton RE, Whay HR, Murrell JC. Spontaneous electroencephalographic changes in a castration model as an indicator of nociception: A comparison between donkeys and ponies. Equine Vet J. 2015;47(1):36–42.

18. Taylor TS, Matthews NS. Mammoth asses—selected behavioural considerations for the veterinarian. Appl Anim Behav Sci. 1998;60:283–9.

19. Ashley F, Waterman-Pearson a E, Whay HR. Behavioural assessment of pain in horses and donkeys: application to clinical practice and future studies. Equine Vet J. 2005;37(6):565–75.

20. Sanz MG, Sellon DC, Cary JA, Hines MT, Farnsworth KD. Analgesic effects of butorphanol tartrate and phenylbutazone administered alone and in combination in young horses undergoing routine castration. J Am Vet Med Assoc. 2009;235(10):1194–203.

21. Casey RA. Clinical problems associated with the intensive management of performance horses. Waran N, editor. Welf Horses. 2004;19–44.

22. Ashley FH, Waterman-Pearson AE, Whay HR. Development of an ethogram to record potential behavioural indicators of pain in working donkeys. Fifth International Colloquium on Working Equines. 2007;15–23.

23. Pritchett LC, Ulibarri C, Roberts MC, Schneider RK, Sellon DC. Identification of potential physiological and behavioral indicators of postoperative pain in horses after exploratory celiotomy for colic. Appl Anim Behav Sci. 2003;80(1):31–43.

24. Sutton GA, Dahan R, Turner D, Paltiel O. A behaviour-based pain scale for horses with acute colic: scale construction. Vet J. 2013;196(3):394–401.

25. Lindegaard C, Thomsen MH, Larsen S, Andersen PH. Analgesic efficacy of intra-articular morphine in experimentally induced radiocarpal synovitis in horses. Vet Anaesth Analg. 2010;37(2):171–85.

26. Wagner AE. Effects of Stress on Pain in Horses and Incorporating Pain Scales for Equine Practice. Vol. 26, Veterinary Clinics of North America - Equine Practice. 2010;481–92.

27. Matthews N. Anaesthesia and analgesia for donkeys. Equine Vet J Ltd, Newmarketd Newmarket. 2002;(Proceedings of the 41st British Equine Veterinary Association Congress):140.

28. Crane M. Colic in the donkey. Equine Vet J Ltd, Newmarket. 2002;(Proceedings of the 41st British Equine Veterinary Association Congress):33–4.

29. Ross SE, Murray JK, Roberts VLH. Prevalence of headshaking within the equine population in the UK. Equine Vet J. 2018;50(1):73–8.

30. Bufalari A, Adami C, Angeli G, Short CE. Pain Assessment in Animals. Vet Res Commun. 2007;31(S1):55–8.

31. Svendsen, E.D. Duncan, J. Hadrill D. Chapter 2: An approach to the dull donkey. In: The professional handbook of the donkey. 4th edition. 2008;28–36.

Appendices

Appendix 1 Sensitivity, specificity, positive and negative predictive value for the FEPS

	Sensitivity	*Specificity*	*Positive predictive value*	*Negative predictive value*
Data 2016 (n=19)	79%	49%	38%	86%
Data 2017 (n=25)	60%	94%	79%	86%
All data (n=44)	68%	75%	51%	86%
All data without post-operative patients (n=38)	74%	74%	52%	88%

Appendix 2 Sensitivity, specificity, positive and negative predictive value for the different subgroups of patients

	Sensitivity	*Specificity*	*Positive predictive value*	*Negative predictive value*
Colic (n=7)	100%	79%	64%	100%
Facial pain (n=7)	100%	75%	64%	100%
Lameness (n=24)	58%	74%	45%	83%
Post-operative pain (n=6)	33%	79%	40%	73%

Appendix 3 Sensitivity and specificity for individual parameters of the FEPS

	Sensitivity	*Specificity*
Head movement	13%	99%
Eyelids	20%	97%
Focus	26%	100%
Nostrils	22%	87%
Corners mouth lip	4%	100%
Muscle tone head	0%	100%
Flehmen	24%	93%
Teeth grinding	0%	99%
Ear response	11%	100%
Ear position	28%	97%
Startle/headshaking	9%	97%
Sweating	0%	99%

Appendix 4 Sensitivity and specificity for individual parameters of the FEPS in the lameness patient group

	Sensitivity	Specificity
Head movement	4%	100%
Eyelids	4%	100%
Focus	21%	100%
Nostrils	29%	86%
Corners mouth lip	8%	100%
Muscle tone head	0%	100%
Flehmen	25%	91%
Teeth grinding	0%	98%
Ear response	8%	100%
Ear position	33%	95%
Startle/headshaking	4%	100%
Sweating	0%	100%

Appendix 5 Sensitivity and specificity for individual parameters of the FEPS in the facial pain patient group

	Sensitivity	Specificity
Head movement	43%	94%
Eyelids	71%	88%
Focus	43%	100%
Nostrils	0%	82%
Corners mouth lip	0%	100%
Muscle tone head	0%	100%
Flehmen	14%	94%
Teeth grinding	0%	100%
Ear response	14%	100%
Ear position	29%	100%
Startle/headshaking	14%	82%
Sweating	0%	100%

Appendix 6 Sensitivity and specificity for individual parameters of the FEPS in the colic patient group

	Sensitivity	Specificity
Head movement	29%	100%
Eyelids	43%	100%
Focus	57%	100%
Nostrils	43%	89%
Corners mouth lip	0%	100%
Muscle tone head	0%	100%
Flehmen	29%	94%
Teeth grinding	0%	100%
Ear response	29%	100%
Ear position	43%	100%
Startle/headshaking	29%	100%
Sweating	0%	94%

Appendix 7 Sensitivity and specificity for individual parameters of the FEPS in the post-operative pain group

	Sensitivity	*Specificity*
Head movement	0%	100%
Eyelids	0%	86%
Focus	0%	100%
Nostrils	0%	93%
Corners mouth lip	0%	100%
Muscle tone head	0%	100%
Flehmen	33%	100%
Teeth grinding	0%	100%
Ear response	0%	100%
Ear position	0%	100%
Startle/headshaking	0%	100%
Sweating	0%	100%