

1 To see the final version of this paper please visit the publisher's website (Applied Animal behaviour
2 Science). Access to the published version may require a subscription.

3

4 Authors: Krista M. McLennan^{a*}, Carlos J. B. Rebelo^a, Murray J. Corke^a, Mark A. Holmes^a, Matthew
5 C. Leach^b, Fernando Constantino-Casas^a.

6

7 Article title: Development of a facial expression scale using footrot and mastitis as models of pain in
8 sheep.

9 Year of publication: 2016

10

11 **Highlights**

12 SPFES can accurately identify sheep with painful diseases from healthy sheep.

13 Trained observers reliably and accurately used the SPFES to detect pain in sheep.

14 Treatment of disease reduced the total facial pain score of adult sheep.

15 Total pain scores positively correlated with lesion and lameness scores.

16

17

18

19

20

21 **Development of a facial expression scale using footrot and mastitis as models of pain in sheep.**

22 **Krista M. McLennan^{a*}, Carlos J. B. Rebelo^a, Murray J. Corke^a, Mark A. Holmes^a, Matthew C.**

23 **Leach^b, Fernando Constantino-Casas^a.**

24 ^aDepartment of Veterinary Medicine, University of Cambridge, Madingley Road, Cambridge CB3
25 0ES, UK

26 ^bSchool of Agriculture, Food & Rural Development, Agriculture Building, Newcastle University,
27 Newcastle upon Tyne NE1 7RU, UK

28 * Contact for correspondence:

29 Krista McLennan, Department of Biological Sciences, University of Chester, Parkgate Rd, Chester
30 CH1 4BJ UK

31 Telephone: +44 (0) 1244 513686

32 Email: k.mclennan@chester.ac.uk

33

34 **Abstract**

35 Management of pain in sheep is limited by the challenges of recognising and accurately quantifying
36 pain in this species. The use of facial expression scoring to assess pain is a well-utilised, practical tool
37 in both humans and non-human animals. The objective of this study was to develop a standardised
38 facial expression pain scale for adult sheep, that could be used reliably and accurately to detect pain
39 associated with naturally occurring painful diseases, such as footrot and mastitis. We also investigated
40 whether the scale could be reliably and accurately utilised by observers after training, enabling the
41 development of an on-farm pain assessment tool. The Sheep Pain Facial Expression Scale (SPFES)
42 was able to correctly identify sheep suffering from disease with a high degree of accuracy (AUC;
43 Footrot: 0.81, Mastitis: 0.80). Diseased sheep scored higher on the scale than controls on the day of
44 disease identification ($P < 0.05$) and diseased sheep showed changes in their facial expression after
45 treatment ($P < 0.001$). The abnormal facial expressions of diseased sheep reduced over time, and at
46 recovery were in line with control sheep. Control sheep did not change their facial expression over
47 time. Five scorers who were trained to use the developed scale also assessed the facial expressions of
48 sheep. The scorers were blind to treatment and session. Scorers reliably (ICC: 0.86) and accurately (α
49 = 0.86) identified changes in the facial expression of sheep with footrot over time ($P < 0.05$), and
50 scored control sheep consistently low over time. The SPFES offers a reliable and effective method of
51 assessing pain in sheep after minimal training.

52

53 **Keywords:**

54 Footrot; Sheep; Pain; Facial expression.

55

56 **1. Introduction**

57 Pain is an aversive experience with both sensory and affective components, often associated with
58 actual or potential tissue damage (Broom, 2001; IASP, 1994; Sneddon et al., 2014). Pain can have

59 considerable effects on the well-being of an animal and its quality of life. The management of pain in
60 farm animals however, is often inadequate, resulting in poor welfare (Crook, 2014; Huxley and Whay,
61 2006). Reasons commonly cited by veterinarians for not administering analgesia to farm animals
62 include cost to the farmer, withdrawal periods for drug residues and a lack of licensed analgesic
63 products in some animals such as sheep (although they can be used under “The Cascade System”)
64 (Lizarraga and Chambers, 2012). One of the major reasons limiting the use of pain relieving drugs in
65 farm animals is difficulties in recognising and quantifying pain (Flecknell, 2008; Huxley and Whay,
66 2006; Ison and Rutherford, 2014; Lizarraga and Chambers, 2012). There is an evident need for an
67 objective, reliable scoring tool that can be effectively used to recognise and assess pain severity in
68 sheep.

69

70 Disease is a major source of pain in sheep, impacting negatively upon welfare and adversely effecting
71 productivity. Footrot in sheep causes severe lameness, a direct sensory response to the tissue damage
72 caused by the bacteria *Dichelobacter nodosus* (Kaler et al., 2010a). As lesion severity increases the
73 degree of lameness observed also increases, indicating the presence of pain (Dolan et al., 2003; Kaler
74 et al., 2010b). Mechanical threshold responses are also significantly reduced when severe footrot is
75 present, indicating the presence of chronic pain; the application of a local anaesthetic block raises the
76 threshold to be in line with that of healthy sheep (Ley et al., 1989). Resolution of the lesions does not
77 necessarily remove this pain, as hyperalgesia to a mechanical stimulus may still be present for up to
78 three months in sheep that had previously suffered from severe footrot (Dolan et al., 2003; Ley et al.,
79 1989).

80

81 Mastitis is also regarded as an extremely painful disease in sheep. Mastitis is the inflammation of the
82 mammary glands usually in response to pathogens such as *Staphylococcus aureus* and *Mannheimia*
83 *haemolytica* (Jones, 1991). These pathogens can also cause painful lesions within the teat canal
84 (Mavrogianni et al., 2004). The development of the disease can be rapid, and in severe cases can lead

85 to death of the sheep. Sheep with mastitis also show mechanical hyperalgesia (Dolan et al., 2000),
86 which supports the hypothesis that this is a painful condition.

87

88 Non-steroidal anti-inflammatory drugs (NSAIDs) have antipyretic, anti-inflammatory and analgesics
89 properties. This supports their use alongside antimicrobials in treating inflammatory conditions such
90 as footrot and mastitis, to aid recovery and reduce the associated pain. Within veterinary practice,
91 sheep suffering from mastitis are more likely to receive an NSAID as part of their treatment as it can
92 rapidly reduce clinical signs of mastitis (Fthenakis, 2000). There is some evidence to also support the
93 use of NSAIDs when treating sheep with footrot; Welsh and Nolan (1995) administered an NSAID,
94 flunixin meglumine, to sheep suffering from footrot. They found mechanical hyperalgesia to be
95 reduced in these sheep compared with sheep that did not receive an NSAID, demonstrating its
96 analgesic property. Kaler et al. (2010a) assessed the anti-inflammatory property of flunixin
97 meglumine as an aid to recovery in sheep with footrot. However, they did not find any effect of
98 NSAIDs on time to recovery when compared with sheep that only received an antibiotic. In sheep,
99 meloxicam has a longer elimination half-life than flunixin meglumine (10.85 ± 1.21 h, 2.48 ± 0.12 h
100 respectively) (Cheng et al., 1998; Shukla et al., 2007) and is detectable in blood plasma for up to 72
101 hours (Shukla et al., 2007) compared to 32 hours for flunixin meglumine (Cheng et al., 1998). These
102 studies provide evidence for meloxicam to be a better alternative to flunixin meglumine in reducing
103 inflammation and pain associated with diseases such as footrot and mastitis in sheep. The effect of
104 meloxicam as an NSAID has not yet been assessed for its ability to reduce pain associated with
105 disease in sheep.

106

107 Current pain assessment tools commonly use behavioural indicators as these provide sensitive and
108 non-invasive measures of pain (Mogil and Crager, 2004). Pain related behaviours such as lip curling,
109 trembling, abnormal postures and vocalisations have been well documented when assessing pain in
110 lambs undergoing tail docking and castration (Grant, 2004; Guesgen et al., 2014; Molony et al.,

111 2002). Observing behavioural changes can be time consuming, making it impractical for on-farm
112 settings. Furthermore, the fluctuating nature of spontaneous pain can mean that smaller, more subtle
113 changes are likely to be missed (Foss et al., 2006).

114

115 Facial expression scoring systems for pain assessment have been recently developed for use in
116 rodents, rabbits and horses (Dalla Costa et al., 2014; Langford et al., 2010; Leach et al., 2012). Facial
117 expression scoring has shown to be successful in identifying and assessing the severity of pain in
118 animals, with minimal time and training required for observers (Langford et al., 2010; Leach et al.,
119 2012; Sotocinal et al., 2011). Changes in facial expression are likely to be an involuntarily response
120 by an animal in response to the fluctuating level of pain experienced (Langford et al., 2010) leading to
121 higher sensitivity in the assessment. The evolutionary stability of facial expression across species
122 (Williams, 2002) and their use within social contexts (Defensor et al., 2012), suggest that adult sheep
123 would also be likely to exhibit changes within their facial expression when experiencing pain.

124

125 The objective of the present study therefore, was to develop a standardised facial expression pain
126 scale that can be used accurately to detect pain associated with naturally occurring painful diseases
127 such as footrot and mastitis. This objective was achieved by visiting eleven commercial farms across
128 East Anglia, UK when disease was reported, and evaluating the changes in facial expressions before
129 and after treatment with antibiotics and during the recovery time. Some of the sheep with footrot were
130 also treated with an NSAID to evaluate the effect of initial analgesia on the expression of pain in
131 sheep during recovery from the disease. We also tested whether the SPFES we developed could be
132 reliably and accurately utilised by observers after training, and thus be a useful and practical on-farm
133 pain assessment tool.

134

135 **2. Methods**

136 *2.1 Ethical statement*

137 Ethical approval was provided by the Department of Veterinary Medicine, University of Cambridge
138 Ethics and Welfare Committee. All disease incidents were naturally occurring and all animals were
139 under the supervision of a veterinarian. All sheep suffering from disease were treated appropriately
140 and revisited throughout the recovery period. No treatment was withheld during the study. Stress to
141 sheep was minimised when handling or approaching animals. Information was provided to each
142 farmer before they gave consent for the study to commence on their farm. Informed consent was
143 obtained from each observer prior to scoring images. All data was anonymised before analysis and no
144 personal details of the participants were recorded or stored.

145

146 *2.2 Footrot*

147 *2.2.1 Study population*

148 One hundred and eleven sheep of differing breeds, gender and coat colour were involved in the study.
149 All the sheep were over one year of age. A total of 73 sheep were diagnosed as having footrot by a
150 veterinarian, using lameness and lesion scoring. These sheep were matched with 38 control sheep
151 from the same farm that had no signs of footrot or other disease. Data were collected from October
152 2012 through to July 2014 across all seasons from eight farms.

153

154 *2.2.2 Study design and treatments*

155 All sheep were assessed for lameness using the five point gait scoring method devised by Ley et al.
156 (1992). All sheep were assessed for footrot lesions using the four point scale developed by Egerton
157 and Roberts (1971). Sheep were categorised into three treatment groups. Group FA (N=37) were
158 treated for the presence of footrot with antibiotics, tulathromycin by subcutaneous injection (2.5mg/kg
159 Draxxin®, Zoetis, Ltd) and topical chlortetracycline (Animedazon® Spray 2.4%, AniMedica). Group

160 FAN (N=36) were treated for the presence of footrot as before and also received a non-steroidal anti-
161 inflammatory drug, meloxicam by sub-cutaneous injection (0.5mg/kg, Metacam®, Boehringer
162 Ingelheim Ltd). Group FC (N=38) showed no signs of lameness and were clinically assessed as being
163 free from clinical disease by a veterinarian and were used as controls. Controls were matched
164 carefully on each farm for breed, gender and age.

165

166 Photographic images of sheep faces were taken on the day of disease identification (day 0) after
167 lameness and lesions were scored. All sheep received an initial treatment on the same day (day 0)
168 after images had been collected. All sheep were revisited during their recovery period and received
169 additional treatment as required by the veterinarian, if signs of active disease were still present.
170 Animals were reassessed for lesions and lameness to establish that they were fully recovered and
171 facial images were recorded again on day 90.

172

173 *2.3 Mastitis*

174 *2.3.1 Study population*

175 Twenty nine primiparous and multiparous recently parturient ewes of differing breeds, coat colour
176 and number of lambs were involved in the study. A total of 17 sheep were identified as having acute
177 clinical mastitis by a veterinarian. These sheep were matched as closely as possible for days since
178 parturition and for number of offspring, with a total of 12 control sheep from the same farm identified
179 as having no signs of clinical mastitis. Data were collected over two lambing seasons (January to July)
180 in 2013 and 2014 from four farms.

181

182 *2.3.2 Study design and treatments*

183 All sheep were assessed for signs of acute clinical mastitis through udder colour and udder palpation
184 by a veterinarian. A milk sample from diseased sheep was taken to identify the pathogen and ensure
185 correct treatment was applied. Sheep were categorised into two treatment groups. Group MAN
186 (N=17) were treated with an appropriate antibiotic, either tulathromycin by subcutaneous injection
187 (2.5mg/kg Draxxin®, Zoetis Ltd) or Oxytetracycline by intramuscular injection (10mg/kg Alamycin
188 LA®, Norbrook Laboratories, Ltd), and all animals received a non-steroidal anti-inflammatory drug,
189 meloxicam by subcutaneous injection (0.5mg/kg, Metacam®, Boehringer Ingelheim Ltd). Group MC
190 (N=12) were assessed as being free from clinical disease by the veterinarian and were used as
191 controls.

192

193 Photographic images of sheep faces were taken on the day of disease identification (day 0) after
194 udders were assessed. All sheep were treated on the same day (day 0) after images had been recorded.
195 All sheep were revisited during their recovery period and further images were collected on day 7 and
196 again on day 42. Animals were reassessed for signs of clinical mastitis to ensure full recovery had
197 occurred by day 42. If sheep had not responded to the initial treatment, further treatment was provided
198 by the veterinarian. Sheep were assessed in small groups with their lambs and stress was kept to a
199 minimum.

200

201 *2.4 Image capture*

202 Multiple photographs of sheep were taken from a distance of approximately 1 m using a high
203 definition camera (Casio®, Exilim HS EX-ZR100, Casio Electronics Co., Ltd., Japan). Photographs
204 were taken on day 0 after animals had been assessed for presence or absence of disease and had been
205 left for twenty minutes to settle. Further photographs were taken on day 7 and 42 for mastitis and on
206 day 90 for footrot before sheep were handled or reassessed for disease presence or absence.

207

208 Profile and frontal pictures were taken for each animal on each occasion whilst they remained within
209 the group. All photographs were cropped to include only the head and to remove body posture, to
210 prevent observers being influenced by the posture of the animal when scoring the facial expressions as
211 in Langford et al. (2010) and Leach et al. (2012). The highest quality pictures were used for scoring
212 where possible.

213

214 *2.5 Sheep pain facial scale development*

215 The sheep pain facial expression scale (SPFES) was developed using sheep from the footrot study
216 group. Footrot was used as our pain model following previous research showing the link between
217 lameness due to footrot and mechanical hyperalgesia (Ley et al., 1989). We followed methods by
218 Langford et al. (2010), Sotocinal et al. (2011), Leach et al. (2012), Keating et al. (2012) and Dalla
219 Costa et al. (2014) to develop our scale. Images of sheep on days 0 and days 90 were compared to
220 identify changes in facial expression associated with the presence of the disease and lameness. Based
221 on these comparisons an initial scale was established and trialled in a pilot study (McLennan et al.,
222 2014). Minor adjustment to the scale with the addition of higher quality photographs and more
223 detailed descriptors allowed the development of the SPFES (Fig. 1). The scale is used to assess
224 expression within five facial areas; orbital tightness, cheek tightness, ear position, lip and jaw profile,
225 and nostril and philtrum position. These areas are scored as having abnormal expression not present,
226 partially present, or present.

227

228 *2.6 Scoring facial expression.*

229 The facial expressions of sheep from both footrot and mastitis were scored separately by an observer
230 (KM) who was experienced in the use of the scoring system. To reduce possible bias, scoring took
231 place three months after the scale had been finalised. Photographs that were not in focus or were of
232 poor quality for angle and light were not scored. To maintain a balanced design, only sheep that had a

233 complete set of photographs across all time points were included. Sheep that required more than one
234 treatment were removed from further analysis. A total of 51 sheep from the footrot group (n, FA=16,
235 FAN=19, FC=16) and 22 for mastitis (n, MAN =12, MC=10) were scored. The scores from KM were
236 used to test the sensitivity and specificity of the scale at detecting disease status and thus pain for
237 mastitis and footrot. The scores were also analysed to determine the effect of time, treatment and a
238 time*treatment interaction.

239

240 Five treatment and session blind observers who had been given training on how to score the facial
241 expressions of sheep, scored a sample data set of 60 images from the footrot group consisting of 20
242 sheep with footrot (n, FOA=9, FOAN=11) and 10 control sheep (FOC). Training consisted of viewing
243 a pictorial guide with descriptors as well as multiple example images of each of the five facial areas.
244 This file also included training and testing sections as well as instructions on how to fill out the
245 scoring file. Training images were not used within the scoring file. The scores from these individuals
246 after training were used to test the reliability and accuracy of the scale across each treatment group for
247 the footrot population. The training tool can be found at www.animalwelfarehub.com.

248

249 Two photographs, one profile and one frontal, were assessed for each time point. Images were
250 presented in a random order generated using a random number generation on Microsoft Excel 7.0.
251 Both photographs were used to give one score to each of the facial areas using the three-point scale (0
252 = not present, 1 = partially present, 2 = present). If the two photographs differed in value, or one area
253 was obscured from view (e.g. nostril and philtrum position can only be seen from the frontal view) the
254 highest score of the two photographs was given. If an area was not clear on either of the photographs,
255 it was scored as 'not able to score'. If two or more areas were scored as 'not able to score', the total
256 score for this image was not included in the analysis. A total pain score (TPS) was determined by
257 adding the individual scores for each of the five areas for each set of photographs. The maximum
258 possible score was 10 (i.e. a score of 2 for each of the 5 facial areas). The five observers were also

259 asked to make a global assessment of whether they thought the sheep was in pain or not, based on
260 their own previous experience, as used by Dalla Costa et al. (2014) and Keating et al. (2012).

261

262 2.7 Statistical analysis

263 Statistical analysis was carried out using R386 3.1.1 (R Core Team, 2014) except for receiver
264 operator curve (ROC) analysis which was carried out using SPSS 22.0 (IBM Corp, 2013). Differences
265 were considered statistically significant at $P \leq 0.05$ and results are reported as mean \pm SEM unless
266 where otherwise stated. Spearman's rank correlations were calculated to investigate the relationships
267 between TPS, lameness and the total lesion scores, as this data was not normally distributed.
268 Spearman's rank correlations were also calculated between each of the facial areas and the TPS. The
269 sensitivity (ability of a test to correctly identify animals with the disease) and specificity (ability of a
270 test to correctly identify animals without the disease) of the scale were calculated. Sensitivity is the
271 ratio of true positives (TP) to true positives plus false negatives (FN): sensitivity = $TP / TP + FN$.
272 Specificity is the ratio of true negatives (TN) to true negatives plus false positives (FP): specificity =
273 $TN / TN + FP$. ROC analysis was carried out by plotting for all cut-off points, the rate of false positives
274 against the rate of true positives. A value of 1.0 indicates a perfect test, whilst a value of 0.5 indicates
275 an inadequate test (Lalkhen and McCluskey, 2008). The sensitivity and specificity for each of the TPS
276 levels was also determined. For footrot groups the outcome was lameness with the predictor as TPS.
277 For the mastitis group sensitivity and specificity was calculated for the first day only with the
278 outcome being disease status and the predictor as TPS. A repeated measures linear mixed-effects
279 model fit by maximum likelihood was used to analyse the TPS across time points (footrot: day 0 and
280 90; mastitis: day 0, 7 and 42). Day, as the repeated measure was nested within sheep as the random
281 effect, with treatment group, day, breed and farm as fixed effects. Any time*treatment interactions
282 were further investigated using analysis of variance with data from separate time periods forming the
283 dependent variables and treatment as the fixed effect. Post-hoc analysis of treatment group effects was
284 conducted using Tukey contrast tests. The Kruskal-Wallis test was used to investigate time*treatment

285 interactions for the footrot group on day 90 due to data being not normally distributed. Kruskal-Wallis
286 was also used to investigate time*treatment interactions for the mastitis group for days 7 and 42 due
287 to data not being normally distributed. In addition, changes in facial expression across days for each
288 treatment group was calculated and compared to zero using a 1-sample t-test, or a Wilcoxon signed
289 rank test where data were not normally distributed.

290

291 The global accuracy of the facial pain score was determined by comparing the global pain or no pain
292 judgement made by treatment and session blinded scorers with the actual disease state of the sheep in
293 each photograph based upon the lameness and lesion scores (e.g. control or diseased on day 0 and day
294 90). Reliability of the scale was assessed by comparing the participants' scores for each area and the
295 TPS, using the intra-class correlation coefficient (ICC), Cronbach's alpha.

296

297 **3. Results**

298 *3.1 Footrot*

299 The TPS scores over the two time periods showed a good accuracy with the area under curve (AUC)
300 reaching 0.81, compared to lameness. Table 1 shows the sensitivity and specificity of each total facial
301 expression score. Table 2 gives details on the correlation between facial areas and the TPS. There
302 were no significant main effects of sheep gender (P=0.47), breed (P=0.12) or farm (P=0.75) on TPS.
303 Time, treatment and time*treatment had significant effects on TPS (P=0.0001, P=0.0007, P=0.0436
304 respectively). On day 0 TPS were significantly different between the three treatment groups ($F_{(2,48)} =$
305 9.02, P=0.0005), with the TPS being higher in sheep with footrot (groups FA and FAN) compared to
306 the control group (group FC) (Tukey post-hoc, P<0.01 for both comparisons). No differences were
307 found between groups that received just antibiotics (FA) and those that received an additional non-
308 steroidal anti-inflammatory drug (FAN) (Tukey post-hoc, P=0.86). At day 90 there were no
309 significant differences between treatment groups ($\chi^2 = 4.59$, df=2, P=0.10) (Fig. 2). Sheep that were

310 treated for footrot with antibiotic only, had a decrease in their facial expression score from day 0 to
311 day 90 ($t=-3.29$, $df=15$, $P=0.005$), as did sheep that received an additional non-steroidal anti-
312 inflammatory drug ($V=7.5$, $P=0.003$). Control sheep did not have a change in their facial expression
313 from day 0 to day 90 ($V=18$, $P=0.18$).

314

315 Lameness was correlated positively with total lesion scores ($r_s = 0.89$, $P<0.0001$). TPS increased as
316 lameness scores increased ($r_s = 0.51$, $P<0.0001$) and as total lesion scores increased the TPS also
317 increased ($r_s = 0.50$, $P<0.0001$).

318

319 *3.2 Mastitis*

320 The facial expression scale showed good accuracy at correctly identifying diseased sheep from control
321 sheep with AUC of 0.80. Table 1 shows the sensitivity and specificity of each TPS for mastitis sheep
322 on day 0. Table 2 gives details on the correlation between facial areas and TPS. There were no main
323 effects of breed ($P=0.22$) or farm ($P=0.31$) on TPS. TPS was affected by a time*treatment interaction
324 ($P=0.02$). Sheep in group MAN had a higher TPS score (4 ± 0.54) than did sheep in group MC ($2 \pm$
325 0.47) on day 0 ($F_{(1, 20)} = 7.52$, $P=0.01$). There were no significant differences in TPS between
326 treatment groups for day 7 ($\chi^2=0.01$, $df=1$, $P=0.92$) and day 42 ($\chi^2 = 0.03$, $df=1$, $P=0.87$) (Fig. 3).
327 Sheep in group MAN had a significant decrease in their facial expression score between days 0 and
328 day 7 ($t=-2.15$, $df=11$, $P=0.05$) and between days 0 and days 42 ($t=-9$, $df=11$, $P<0.001$). The TPS did
329 not change between day 7 and day 42 for sheep in group MAN ($t=-1.61$, $df=11$, $P=0.14$) and did not
330 change for sheep in group MC between days 0 and day 7 ($t=1.03$, $df=9$, $P=0.33$), days 0 and day 42
331 ($t=0.133$, $df=9$, $P=0.90$) and days 7 and 42 ($t=-0.58$, $df=9$, $P=0.58$).

332

333 *3.3 Five trained observers*

334 The average accuracy of the global pain assessment was 67%, with individual accuracy ranging from
335 60% to 75%. Of the errors, false positives (26.3%) were more common than false negatives (6.3%).
336 The TPS had a high accuracy in relation to lameness with an AUC of 0.84. Table 1 shows the
337 sensitivity and specificity of each level of the TPS given by observers. Table 2 gives details on the
338 correlation between facial areas and TPS. There was a high inter-rater reliability with an overall intra-
339 class correlation (ICC) value of 0.86. Each of the facial areas assessed also showed high (orbital
340 tightening, 0.90; cheek tightening, 0.82; abnormal ear position, 0.85) to medium ICC values
341 (abnormal lip and jaw profile, 0.63; and abnormal nose position, 0.65). The five facial areas were
342 scored easily by all participants as demonstrated by the percentage of “not able to score” ranging from
343 0% for orbital tightening to 12% for cheek tightening.

344

345 There was a main effect of breed ($P=0.02$); however, when performing contrasts, there were no
346 significant differences identified between breeds ($P>0.05$). There were no significant effects of gender
347 ($P=0.46$) or farm ($P=0.71$) on TPS. Time, treatment and time*treatment interaction had an effect on
348 TPS ($P=0.001$, $P=0.02$, $P=0.003$, respectively). There were differences between treatment groups on
349 day 0 ($F_{(2,27)} = 11.33$, $P=0.0003$). Sheep in group FOA (4.78 ± 0.49) had a higher TPS than sheep in
350 group FOC (2.70 ± 0.30) (Tukey post-hoc, $P=0.007$). Sheep in group FOAN (5.45 ± 0.47) also had a
351 higher TPS than sheep in group FOC (Tukey post-Hoc, $P=0.0002$). Sheep in group FOA and group
352 FOAN did not differ in TPS on day 0. There were no differences in TPS on day 90 between treatment
353 groups ($\chi^2 = 1$, $df = 2$, $P=0.61$). Participants did not score sheep in group FOC differently on day 0
354 compared to day 90 ($t=0.33$, $df=9$, $P=0.75$). Sheep in group FOA had a lower TPS on day 90
355 compared to day 0 ($V=4$, $P=0.05$) as did sheep in group FOAN ($t = -5.49$, $df=10$, $P=0.0003$) (Fig. 4).

356

357 Lameness was correlated positively with total lesion scores ($r_s = 0.82$, $P<0.0001$). TPS increased as
358 lameness scores increased ($r_s = 0.56$, $P<0.0001$) and as total lesion scores increased the TPS also
359 increased ($r_s = 0.54$, $P<0.0001$).

360

361 **4. Discussion**

362

363 The SPFES developed for this study showed a high degree of accuracy, differentiating between lame
364 and non-lame sheep correctly, through identifying changes in the facial expressions according to their
365 level of lameness. These changes in facial expression are similar to those described in other species
366 with respect to pain (Dalla Costa et al., 2014; Keating et al., 2012; Leach et al., 2012; Sotocinal et al.,
367 2011). Importantly, there were no changes in the facial expression of non-lame sheep. Sheep that had
368 been suffering from footrot showed high total pain scores that decreased as they recovered. Total pain
369 scores were positively related to both the total lesion scores and the lameness scores, providing further
370 evidence for pain in sheep to be both a sensory and emotional experience. The positive correlation
371 between the level of lameness and severity of footrot lesion observed in our study confirms our choice
372 of model and is in agreement with other studies (Dolan et al., 2003; Kaler et al., 2011).

373

374 Although we could not differentiate between groups FA and FAN at either time point the provision of
375 analgesic treatment to sheep with footrot at the time of disease diagnosis appeared to reduce the total
376 pain score over the 90 day observation period further, compared with sheep that only received an
377 antibiotic. This result was also noted by the five trained observers whereby the FOAN group had a
378 larger decrease in their scores between day 0 and day 90 compared with the FOA group. This
379 decrease in total pain score supports the need to manage pain in sheep with this disease. It is possible
380 that the use of a non-steroidal anti-inflammatory drug may have reduced the effects of potential
381 “wind-up” from persistent excitation of the nociceptors involved with the footrot lesions (Stein, 2013;
382 Viking Höglund and Frendin, 2002); however, further investigation is required. The reduction in pain
383 could have allowed the sheep to recover more efficiently and resume normal activity before sheep that
384 had not received this additional treatment. Treatment was given to sheep on day 0 after the
385 photographs had been taken and so no effect of analgesic would have been occurring at the time of

386 photography. Any effect of analgesia would have been detectable up to 72 hours after the
387 administration of the analgesic as suggested by its elimination half-life (Shukla et al., 2007). In future
388 studies, it would be beneficial to monitor the changes of facial expression over this time period.

389

390 The high level of specificity for a total pain score above 5 for each of the diseases indicates that a
391 sheep given this score or above are unlikely to be a false positive. Sheep scoring a total pain score
392 above 5 are therefore likely to benefit from the administration of pain relief. Although the sensitivity
393 of the test is low, meaning that some of the diseased animals may go undetected below a TPS of 5, the
394 overall accuracy of the test is high. It is preferable for a test such as this to have a higher specificity
395 rate where sheep reaching a high pain score are unlikely to be negative for the painful disease.

396

397 Total facial expression scores at day 90 were not zero. It is possible that hyperalgesia remained a
398 contributing factor within our study population. Ley et al. (1995) also found sheep previously
399 diagnosed with footrot were still showing an increased response to mechanical stimulation compared
400 to control sheep three months after they had seemingly recovered. Control sheep were also not scored
401 as zero on day 0 or day 90, a finding observed in other studies using facial expression as a pain
402 scoring system (Dalla Costa et al., 2014; Keating et al., 2012). There are several possible explanations
403 for this. Control sheep may have had a previous episode of footrot that was not evident at the time of
404 clinical examination on day 0 and the associated hyperalgesia may have still been present.
405 Additionally, facial expression may change due to other affective states such as fear and stress, which
406 can both be related to pain. The development of facial expression scales to help identify other
407 affective states, both positive and negative, would be beneficial.

408

409 The SPFES also showed a high degree of accuracy in correctly differentiating between sheep with
410 mastitis and controls. The total pain scores of sheep with mastitis were higher than control sheep on

411 day 0 and decreased rapidly in response to treatment by day 7. Facial expressions in sheep with
412 mastitis did not change significantly from day 7 to day 42 suggesting that the provision of a non-
413 steroidal anti-inflammatory drug as well as antibiotic treatment reduced the associated pain
414 substantially by day 7. Importantly there were no changes in the facial expression of sheep that acted
415 as controls across time. These results provide further evidence that the SPFES is accurate at
416 identifying changes in facial expression that suggest pain in sheep associated with disease.

417

418 The results from the five observers are in line with those given by the more experienced scorer and
419 demonstrates that the provision of basic training allowed for the effective use of the SPFES to be used
420 accurately and reliably. In addition, the similarity in results from the main observer and the treatment
421 and session blinded observers, provides evidence that bias was unlikely to be present in the main
422 observer. The total pain scores given by the observers correctly identified lame and non-lame sheep,
423 giving higher scores to lame sheep compared to control sheep on day 0. The observers also scored
424 sheep at day 90 as low and similar between groups. Observers' scores also correlated positively with
425 both the lameness and the lesions scores, supporting the use of the SPFES in identifying pain. The
426 global pain assessment given by observers was lower (67%) than that of other "Grimace Scales" (97%
427 for the Mouse Grimace Scale, (Langford et al., 2010), 84% for the Rabbit Grimace Scale (Keating et
428 al., 2012), but similar to the Horse Grimace Score (73.3%) (Dalla Costa et al., 2014). Scorers were
429 readily able to identify pain when present, but were cautious in diagnosing absence of pain. In terms
430 on animal welfare, this is the preferable result. However, the accuracy of the scale improved (up to
431 84%) when scores given to each area were combined to give a total pain score. This provides support
432 for the use of the SPFES at identifying pain in sheep in relation to disease, rather than giving a global
433 assessment. The increase in objectivity through the use of the scale potential helps to remove any fear
434 of not identifying a sheep in pain correctly.

435

436 The SPFES scale is reliable between scorers with an overall inter-rater reliability score of 0.86, and
437 there was high consistency in scores given to the orbital area, the cheek area and ear positioning,
438 similar findings to others (Keating et al., 2012; Sotocinal et al., 2011). The lip and jaw profile along
439 with the nostril and philtrum positioning were less reliable between scorers, a result also noticed for
440 the Horse Grimace Scale (Dalla Costa et al., 2014). The nostril and philtrum position also did not
441 correlate well with the other areas of the face. This is likely due to the way in which images were
442 captured. Images for this study were taken as individual photographs rather than still images captured
443 from video footage. Low image quality and photographs taken at poor angles were avoided wherever
444 possible; however, there may still be possible negative impacts on effective scoring, a problem noted
445 within other validation studies of facial expression scales in animals (Dalla Costa et al., 2014; Keating
446 et al., 2012; Langford et al., 2010).

447

448 Farm and gender did not have any significant effect on the total pain score across treatment group and
449 time supporting its use as an on farm assessment tool. Breed was only noted to have a significant main
450 effect on total pain scores in the trained observer group; however, on further investigation there were
451 no significant differences between breeds found. The anatomical differences between some breeds of
452 sheep, as well as different colours of the face, may have made it difficult for some observer's to score
453 areas effectively. However, the muscle groups involved in facial expression will be the same in each
454 breed and so the changes in facial areas will be the same movement (see Fig. 1 abnormal ear position
455 for an example of this). Facial areas were well correlated with the total pain score across diseases.
456 There are some areas of the face that correlate with each other well; orbital tightening, abnormal ear
457 position and abnormal lip and jaw profile. Sheep suffering with mastitis had several areas of the face
458 that were not well correlated with each other. This could be due to the smaller sample size for the
459 mastitis group, or it could be a factor of the disease state. The systemic nature of mastitis is more
460 likely to leave sheep dehydrated and therefore some areas of the face may be affected by this, such as
461 the orbital and cheek area may appear sunken. It is important that a full assessment of any animal is
462 carried out if disease is suspected and taken into account when scoring the animals. Changes in the

463 facial expression occur during other activities such as blinking or chewing which can change the
464 appearance of the orbital area and cheek area respectively. Every effort was made to eliminate
465 photographs that may have been taken during these activities; however, using the SPFES to score
466 animals 'live' rather than using still images would resolve many of these problems. Fluctuation in
467 pain will also result in a fluctuating facial expression. Scoring animals live would identify these
468 fluctuations through the changes in facial expression and may lead to a better ability at assessing the
469 intensity of the pain experienced. Future trials for scoring animals live after initial training are
470 currently being planned to further investigate the use of the SPFES on farm.

471

472

473 **5. Conclusion**

474 The major challenge for pain research is being able to assess the emotional side of pain (Flecknell et
475 al., 2011). Facial expression as a pain scoring method offers the potential to start to understand this
476 side of animal pain (Kunz et al., 2012, 2009) and the results from the current study support this. At
477 present, the SPFES has been assessed using footrot as the clinical model and successfully applied to
478 mastitis, a disease causing acute pain in sheep. It is likely that the scale can be used for other
479 conditions that are suspected of being painful, such as pregnancy toxaemia where the administration
480 of a non-steroidal anti-inflammatory is known to aid recovery (Zamir et al., 2009). The current scale
481 provides an accurate and reliable method to recognise and assess pain in sheep. It is also doubles as a
482 training tool for veterinarians and farmers to learn more about changes in the facial expression of
483 sheep when they are likely to be suffering from pain. Such a tool is likely to improve an observer's
484 ability to quantify pain in animals and allow observers to discriminate between different pain states
485 independent of disease status, as well as detect the effectiveness of pain relief. Prompt recognition of
486 pain through the use of the scale will enable farmers and veterinarians to treat and manage their flocks
487 better, reducing the impact of pain on their sheep, thus improving welfare and production. It is
488 important to stress that the scale should be used as part of other measures of pain and not as a

489 standalone assessment. The provision of the sensitivity and specificity of the scales at each level of
490 pain will aid scorers in their decision of when to intervene with pain management; something that is
491 often missing from such scales. This will lead to better management of flocks, leading to better
492 production values and higher welfare for the sheep.

493

494 **Acknowledgements**

495 We would like to thank the EU VII Framework Program (FP7-KBBE-2010-4) for funding this study
496 as part of the AWIN project. Boehringer Ingelheim Ltd are thanked for providing the meloxicam
497 used in this study. The funders did not have any influence or involvement in the study design, data
498 collection, analysis, interpretation of data, writing of the report or the decision to publish. We would
499 also like to thank the participating farmers for allowing us to collect data from their farms and the
500 observers who took time to complete the scoring of sheep faces.

501

502 **Conflict of interests**

503 There are no conflicts of interests for any of the authors.

504

505 **Author Contributions**

506 Gathered data on farms: KMM, CJR, MC. Analysed data: KMM and MH. Interpreted results and
507 drafted manuscript: KMM. Read, edited and approved the manuscript: KMM, CJR, MC, ML, MH,
508 FCC.

509

510 **References**

511 Broom, D.M., 2001. The evolution of pain. *Vlaams Diergeneesk. Tijdschrift* 70, 17–21.

- 512 Cheng, Z., Mckeller, Q., Nolan, A., 1998. Pharmacokinetic studies of flunixin meglumine and
513 phenylbutazone in plasma, exudate and transudate in sheep. *J. Vet. Pharmacol. Ther.* 21, 315–
514 321. doi:10.1046/j.1365-2885.1998.00144.x
- 515 Crook, A., 2014. Introduction. Pain: An Issue of Animal Welfare, in: Egger, C.M., Love, L., Doherty,
516 T. (Eds.), *Pain Management in Veterinary Practice*. Wiley-Blackwell, Oxford, pp. 3–8.
- 517 Dalla Costa, E., Minero, M., Lebelt, D., Stucke, D., Canali, E., Leach, M.C., 2014. Development of
518 the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine
519 castration. *PLoS One* 9, e92281. doi:10.1371/journal.pone.0092281
- 520 Defensor, E.B., Corley, M.J., Blanchard, R.J., Blanchard, D.C., 2012. Facial expressions of mice in
521 aggressive and fearful contexts. *Physiol. Behav.* 107, 680–5. doi:10.1016/j.physbeh.2012.03.024
- 522 Dolan, S., Field, L.C., Nolan, A.M., 2000. The role of nitric oxide and prostaglandin signaling
523 pathways in spinal nociceptive processing in chronic inflammation. *Pain* 86, 311–320.
524 doi:10.1016/S0304-3959(00)00262-1
- 525 Dolan, S., Kelly, J.G., Monteiro, A.M., Nolan, A.M., 2003. Up-regulation of metabotropic glutamate
526 receptor subtypes 3 and 5 in spinal cord in a clinical model of persistent inflammation and
527 hyperalgesia. *Pain* 106, 501–512. doi:10.1016/j.pain.2003.09.017
- 528 Egerton, J.R., Roberts, D.S., 1971. Vaccination Against Ovine Foot-rot. *J. Clin. Psychol.* 81, 179–
529 185.
- 530 Flecknell, P., 2008. Analgesia from a veterinary perspective. *Br. J. Anaesth.* 101, 121–124.
531 doi:10.1093/bja/aen087
- 532 Flecknell, P., Leach, M., Bateson, M., 2011. Affective state and quality of life in mice. *Pain* 152, 963–
533 4. doi:10.1016/j.pain.2011.01.030
- 534 Foss, J.M., Vania Apkarian, A., Chialvo, D.R., 2006. Dynamics of pain: fractal dimension of temporal
535 variability of spontaneous pain differentiates between pain States. *J. Neurophysiol.* 95, 730–6.
536 doi:10.1152/jn.00768.2005
- 537 Fthenakis, G.C., 2000. Field evaluation of flunixin meglumine in the supportive treatment of ovine
538 mastitis. *J. Vet. Pharmacol. Ther.* 23, 405–407. doi:10.1111/j.1365-2885.2000.00284.x
- 539 Grant, C., 2004. Behavioural responses of lambs to common painful husbandry procedures. *Appl.*
540 *Anim. Behav. Sci.* 87, 255–273. doi:10.1016/j.applanim.2004.01.011
- 541 Guesgen, M.J., Beausoleil, N.J., Minot, E.O., Stewart, M., Stafford, K.J., 2014. Social context and
542 other factors influence the behavioural expression of pain by lambs. *Appl. Anim. Behav. Sci.*
543 159, 41–49. doi:10.1016/j.applanim.2014.07.008
- 544 Huxley, J.N., Whay, H.R., 2006. Current attitudes of cattle practitioners to pain and the use of
545 analgesics in cattle. *Vet. Rec.* 159, 662–668.
- 546 IASP, 1994. Part III: Pain Terms, A Current List with Definitions and Notes on Usage, in: Merskey,
547 H., Bogduk, N. (Eds.), *Classification of Chronic Pain*. IASP Press, Seattle, pp. 209–214.
- 548 IBM Corp, 2013. *IBM SPSS Statistics for Windows, Version 22.0*. Armonk, NY. IBM SPSS.
- 549 Ison, S.H., Rutherford, K.M.D., 2014. Attitudes of farmers and veterinarians towards pain and the use
550 of pain relief in pigs. *Vet. J.* 202, 622–7. doi:10.1016/j.tvjl.2014.10.003
- 551 Jones, J.E.T., 1991. Mastitis in Sheep, in: Owen, J.B., Axford, R.E. (Eds.), *Breeding for Resistance in*
552 *Farm Animals*. CABI Publishing, Wallingford, pp. 412–423.
- 553 Kaler, J., Daniels, S., Wright, J., Green, L., 2010a. A randomised factorial design clinical trial to
554 investigate the impact of parenteral long acting oxytetracycline, foot trimming and flunixin
555 meglumine on time to recovery from lameness and foot lesions in sheep lame with footrot. *J Vet*
556 *Intern Med* 24, 420–425.

- 557 Kaler, J., George, T.R.N., Green, L.E., 2011. Why are sheep lame? Temporal associations between
558 severity of foot lesions and severity of lameness in 60 sheep. *Anim. Welf.* 20, 433–438.
- 559 Kaler, J., Medley, G.F., Grogono-Thomas, R., Wellington, E.M.H., Calvo-Bado, L.A., Wassink, G.J.,
560 King, E.M., Moore, L.J., Russell, C., Green, L.E., 2010b. Factors associated with changes of
561 state of foot conformation and lameness in a flock of sheep. *Prev. Vet. Med.* 97, 237–44.
562 doi:10.1016/j.prevetmed.2010.09.019
- 563 Keating, S.C.J., Thomas, A., Flecknell, P., Leach, M.C., 2012. Evaluation of EMLA cream for
564 preventing pain during tattooing of rabbits: changes in physiological, behavioural and facial
565 expression responses. *PLoS One* 7, e44437. doi:10.1371/journal.pone.0044437
- 566 Kunz, M., Lautenbacher, S., LeBlanc, N., Rainville, P., 2012. Are both the sensory and the affective
567 dimensions of pain encoded in the face? *Pain* 153, 350–8. doi:10.1016/j.pain.2011.10.027
- 568 Kunz, M., Prkachin, K., Lautenbacher, S., 2009. The smile of pain. *Pain* 145, 273–5.
569 doi:10.1016/j.pain.2009.04.009
- 570 Lalkhen, A.G., McCluskey, A., 2008. Clinical tests: sensitivity and specificity. *Contin. Educ.*
571 *Anaesthesia, Crit. Care Pain* 8, 221–223. doi:10.1093/bjaceaccp/mkn041
- 572 Langford, D.J., Bailey, A.L., Chanda, M.L., Clarke, S.E., Drummond, T.E., Echols, S., Glick, S.,
573 Ingrao, J., Klassen-Ross, T., Lacroix-Fralish, M.L., Matsumiya, L., Sorge, R.E., Sotocinal, S.G.,
574 Tabaka, J.M., Wong, D., van den Maagdenberg, A.M.J.M., Ferrari, M.D., Craig, K.D., Mogil,
575 J.S., 2010. Coding of facial expressions of pain in the laboratory mouse. *Nat. Methods* 7, 447–9.
576 doi:10.1038/nmeth.1455
- 577 Leach, M.C., Klaus, K., Miller, A.L., Scotto di Perrotolo, M., Sotocinal, S.G., Flecknell, P.A., 2012.
578 The assessment of post-vasectomy pain in mice using behaviour and the Mouse Grimace Scale.
579 *PLoS One* 7, e35656. doi:10.1371/journal.pone.0035656
- 580 Ley, S.J., Livingston, A., Waterman, A.E., 1989. The effect of chronic clinical pain on thermal and
581 mechanical thresholds in sheep. *Pain* 39, 353–357. doi:10.1016/0304-3959(89)90049-3
- 582 Ley, S.J., Livingston, A., Waterman, A.E., 1992. Effects of clinically occurring chronic lameness in
583 sheep on the concentrations of plasma noradrenaline and adrenaline. *Res. Vet. Sci.* 53, 122–5.
- 584 Ley, S.J., Waterman, A.E., Livingston, A., 1995. A field study of the effect of lameness on
585 mechanical nociceptive thresholds in sheep. *Vet. Rec.* 137, 85–7.
- 586 Lizarraga, I., Chambers, J.P., 2012. Use of analgesic drugs for pain management in sheep. *N. Z. Vet.*
587 *J.* 60, 87–94. doi:10.1080/00480169.2011.642772
- 588 Mavrogianni, V.S., Fthenakis, G.C., Burriel, A.R., Gouletsou, P., Papaioannou, N., Taitzoglou, I.A.,
589 2004. Experimentally Induced Teat Stenosis in Dairy Ewes: Clinical, Pathological and
590 Ultrasonographic Features. *J. Comp. Pathol.* 130, 70–74. doi:10.1016/S0021-9975(03)00070-7
- 591 McLennan, K.M., Rebelo, C.B.J., Corke, M.J., Holmes, M.A., Constatino-Casas, F., 2014. The
592 development of a facial grimace score in adult sheep, in: *Proceedings of ISAE UK and Ireland*
593 *Regional Meeting.* p. 9.
- 594 Mogil, J.S., Crager, S.E., 2004. What should we be measuring in behavioral studies of chronic pain in
595 animals? *Pain* 112, 12–5. doi:10.1016/j.pain.2004.09.028
- 596 Molony, V., Kent, J.E., McKendrick, I.J., 2002. Validation of a method for assessment of an acute
597 pain in lambs. *Appl. Anim. Behav. Sci.* 76, 215–238. doi:10.1016/S0168-1591(02)00014-X
- 598 R Core Team, 2014. *R: A language and environment for statistical computing.* Vienna, Austria.
- 599 Shukla, M., Singh, G., Sindhura, B.G., Telang, A.G., Rao, G.S., Malik, J.K., 2007. Comparative
600 plasma pharmacokinetics of meloxicam in sheep and goats following intravenous administration
601 145, 528–532. doi:10.1016/j.cbpc.2007.01.020

- 602 Sneddon, L.U., Elwood, R.W., Adamo, S.A., Leach, M.C., 2014. Defining and assessing animal pain.
603 *Anim. Behav.* 97, 201–212. doi:10.1016/j.anbehav.2014.09.007
- 604 Sotocinal, S.G., Sorge, R.E., Zaloum, A., Tuttle, A.H., Martin, L.J., Wieskopf, J.S., Mapplebeck,
605 J.C.S., Wei, P., Zhan, S., Zhang, S., McDougall, J.J., King, O.D., Mogil, J.S., 2011. The Rat
606 Grimace Scale: a partially automated method for quantifying pain in the laboratory rat via facial
607 expressions. *Mol. Pain* 7, 55. doi:10.1186/1744-8069-7-55
- 608 Stein, C., 2013. Opioids, sensory systems and chronic pain. *Eur. J. Pharmacol.* 716, 179–87.
609 doi:10.1016/j.ejphar.2013.01.076
- 610 Viking Höglund, O., Frendin, J., 2002. Analgesic effect of meloxicam in canine acute dermatitis - a
611 pilot study. *Acta Vet. Scand.* 43, 247–252. doi:10.1186/1751-0147-43-247
- 612 Welsh, E.M., Nolan, A.M., 1995. Effect of flunixin meglumine on the thresholds to mechanical
613 stimulation in healthy and lame sheep. *Res. Vet. Sci.* 58, 61–66. doi:10.1016/0034-
614 5288(95)90090-X
- 615 Williams, A.C.D.C., 2002. Facial expression of pain: an evolutionary account. *Behav. Brain Sci.* 25,
616 439–55; discussion 455–88.
- 617 Zamir, S., Rozov, A., Gootwine, E., 2009. Treatment of pregnancy toxemia in sheep with flunixin
618 meglumine. *Vet. Rec.* 165, 265–266.
- 619

620 **Table 1.** The sensitivity and 1 - specificity of total facial expression scores for each disease for different positive cut-off points, as scored by the experienced
621 observer (footrot and mastitis) and by the five trained observers (five observers – footrot). A high sensitivity value indicates a high percentage of sheep
622 identified as being positive for the disease, if the pain score is greater than or equal to the total pain score value listed. A low 1-specificity value
623 indicates a high percentage of sheep correctly identified as not having the disease if the pain score is greater than or equal to the total pain score. *Note,
624 the 1-specificity value of 0.000 indicates all sheep that did not have the disease were not given a total pain score above this level, i.e. they were correctly
625 identified as not having the disease.

Total pain score: positive if greater than or equal to	Footrot		Mastitis		Five observers - footrot	
	Sensitivity	1 - Specificity	Sensitivity	1 - Specificity	Sensitivity	1 - Specificity
1.5	0.927	0.459	1.000	0.600	1.000	0.791
2.5	0.829	0.311	0.667	0.400	0.941	0.581
3.5	0.512	0.180	0.667	0.200	0.882	0.395
4.5	0.293	0.016	0.333	0.000*	0.647	0.163
5.5	0.171	0.000*	0.167	0.000	0.412	0.047
6.5	0.049	0.000	-	0.000	0.176	0.000*
7.5	0.024	0.000	0.083	0.000	0.118	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000

626

627 **Table 2.** Correlations between each areas of the face and the total pain score from sheep scored by
 628 KM with footrot represented in the top row, sheep scored by the five trained observers in the middle
 629 row and sheep with mastitis in the bottom row. *P<0.05, **P<0.01, ***P<0.001.

	Orbital tightening	Cheek (masseter) tightening	Abnormal ear position	Abnormal lip and jaw profile	Abnormal nostril and philtrum shape	Total Pain Score (TPS)
Orbital	-	0.32***	0.41***	0.26**	0.20*	0.61***
tightening		0.51***	0.52***	0.42***	0.37***	0.73***
		0.09	0.37**	0.29*	0.25	0.66***
Cheek	0.32***	-	0.18	0.36***	0.32***	0.60***
(masseter)	0.51***		0.45***	0.45***	0.47***	0.75***
tightening	0.09		0.18	0.26	0.01	0.45***
Abnormal ear	0.41***	0.18	-	0.34***	0.20*	0.64***
position	0.52***	0.45***		0.52***	0.50***	0.78***
	0.37**	0.18		0.37**	0.02	0.59***
Abnormal lip	0.26**	0.36***	0.34***	-	0.36***	0.47***
and jaw profile	0.42***	0.45***	0.52***		0.63***	0.79***
	0.29*	0.26	0.37**		0.24	0.73***
Abnormal	0.20*	0.32***	0.20*	0.36***	-	0.62***
nostril and	0.37***	0.47***	0.50***	0.63***		0.77***
philtrum shape	0.25	0.01	0.02	0.24		0.55***
Total Pain	0.61***	0.60***	0.64***	0.47***	0.62***	-
Score (TPS)	0.73***	0.75***	0.78***	0.79***	0.77***	
	0.66***	0.45***	0.59***	0.73***	0.55***	

Orbital tightening



Not present = 0



Partially present = 1



Present = 2

There is a closing of the palpebral fissure by the eyelids and a narrowing of the eye aperture. If the eye closes more than half way it should be scored as present (2).

Cheek (masseter muscle) tightening



Not present = 0



Partially present = 1



Present = 2

There is a more convex shaping to the cheek in the area of the masseter muscle and the zygomatic arch as tension increases.

Abnormal ear position (front)



Not present = 0



Partially present = 1



Present = 2

The ears become fully rotated ventrally and caudally and the inner pinna of the ear becomes less visible. Note: Baseline (not present) ear carriage varies between breeds; however, changes in ear position are the same.

Abnormal ear position (side)



Not present = 0

Partially present = 1

Present = 2

The ears become fully rotated ventrally and caudally and the inner pinna of the ear becomes less visible. Note: Baseline (not present) ear carriage varies between breeds; however, changes in ear position are the same.

Abnormal lip and jaw profile

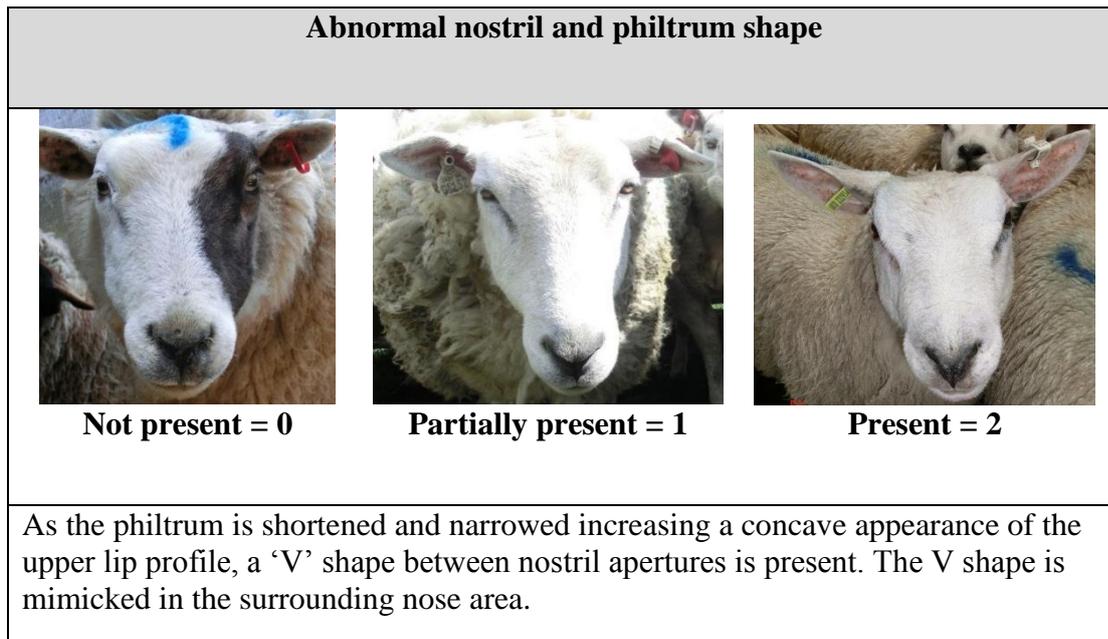


Not present = 0

Partially present = 1

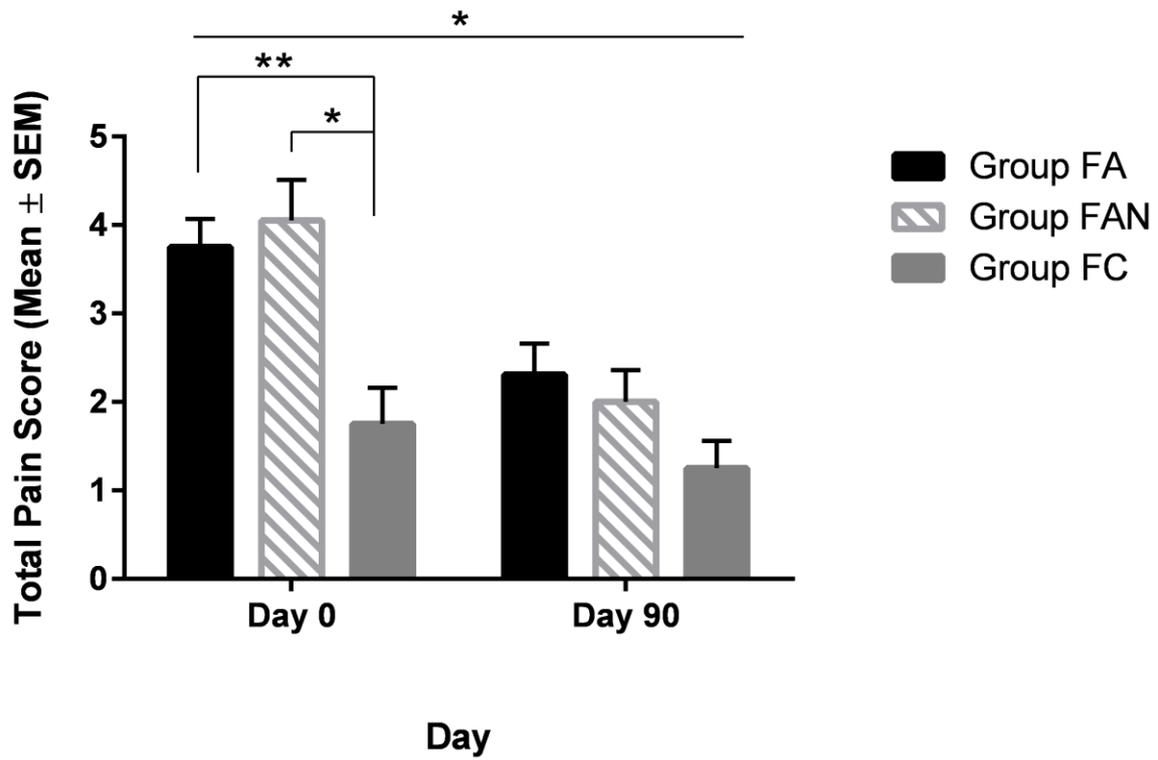
Present = 2

The lower lip is drawn back caudally and the jaw profile appears straight to concave. The chin and jaw line are straightened. The lip line to the commissure of the mouth is straight or even rotated ventrally.



631 **Fig. 1. The Sheep Pain Facial Expression Scale (SPFES).** The Sheep Pain Facial Expression Scale
 632 with images and descriptors of each facial area. Each facial area is scored according to whether it is
 633 not present (score of 0), partially present (score of 1) and present (score of 2). Note: not every facial
 634 area will be present when scoring the expression. Some areas may be expressed at the highest level,
 635 whilst others are not present, in the same sheep.

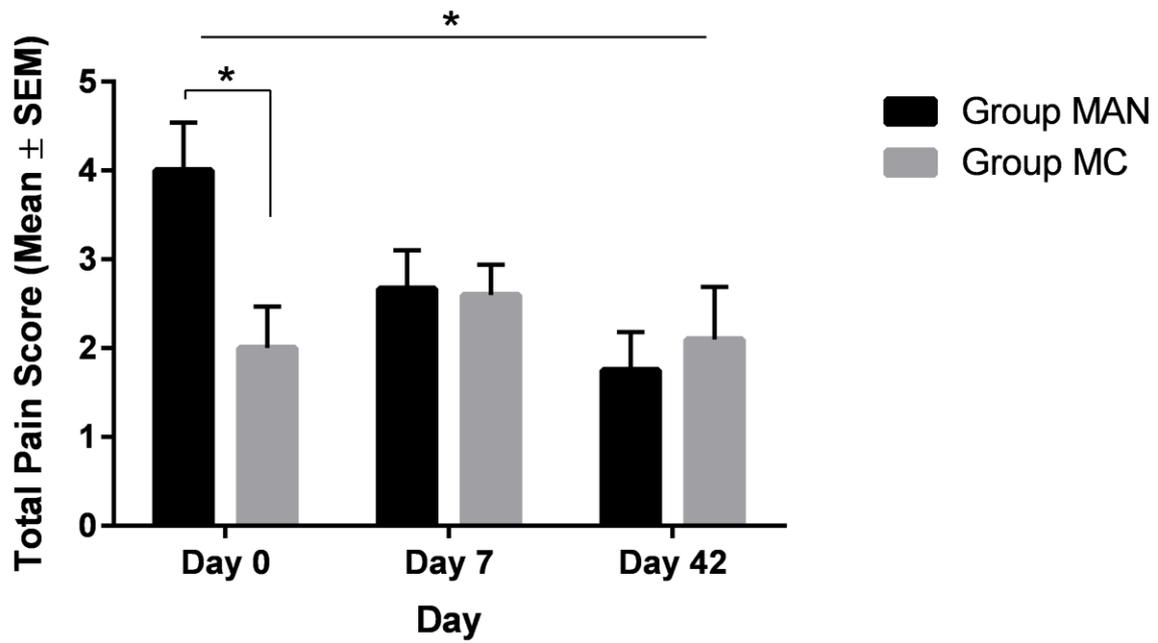
636



637

638 **Fig. 2.** Total facial expression pain score (mean ± SEM) of sheep treated for footrot with systemic
 639 antibiotics (FA), with antibiotics plus a non-steroidal anti-inflammatory drug (FAN) and control sheep
 640 (FC), as scored on day 0 and day 90 by an experienced observer. *P<0.05, ** P<0.01

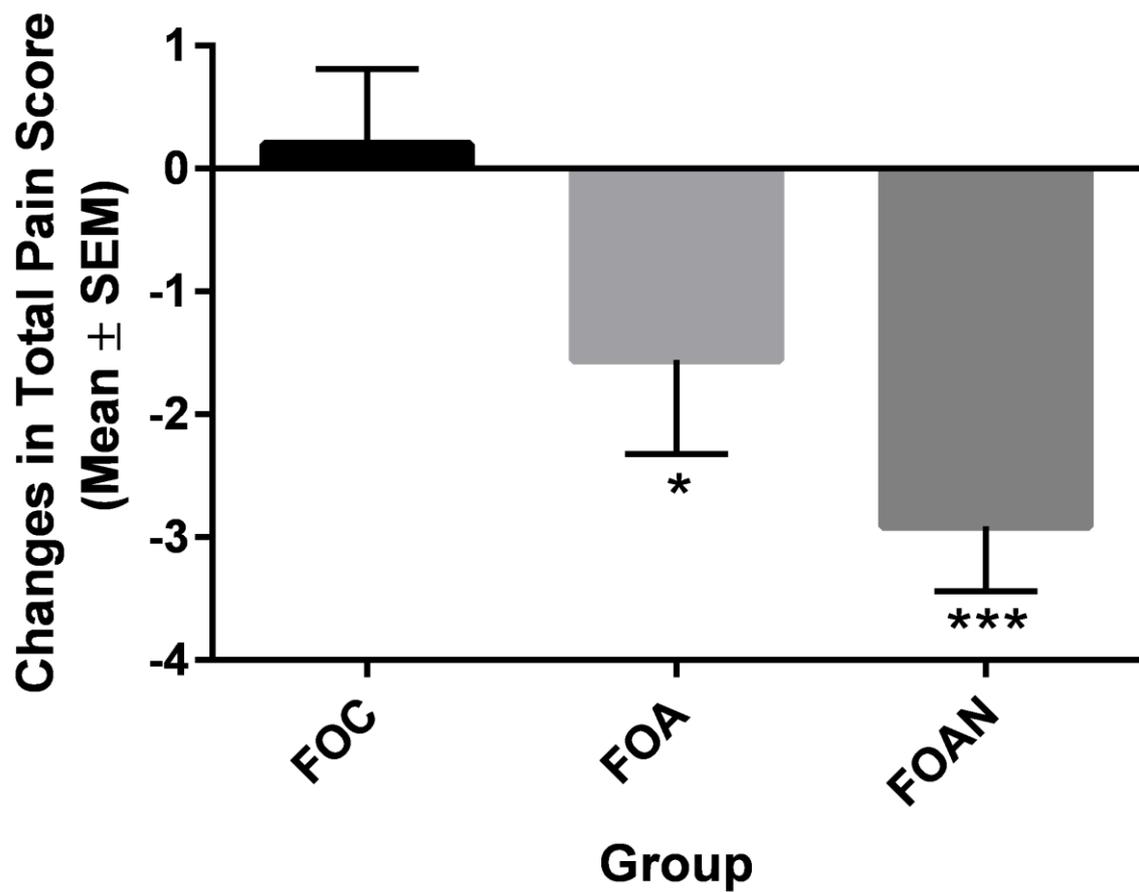
641



642

643 **Fig. 3.** Total facial expression pain score (mean ± SEM) of sheep treated for mastitis with systemic
 644 antibiotics and a non-steroidal anti-inflammatory drug (MAN) and control sheep (MC), as scored on
 645 day 0, day 7 and day 42 by an experienced observer. * P<0.05.

646



647

648 **Fig. 4.** Changes in facial expression total pain score (mean ± SEM) from day 0 to day 90 of sheep
 649 treated for footrot with systemic antibiotics (FOA), with antibiotics plus a non-steroidal anti-
 650 inflammatory drug (FOAN) and control sheep (FOC), as scored by five trained observers. * P<0.05,
 651 ***P<0.001.