Reliability and validity of a visual analogue scale used by owners to measure chronic pain attributable to osteoarthritis in their dogs

Anna K. Hielm-Björkman, DVM, PhD; Amy S. Kapatkin, DVM, MS; Hannu J. Rita, PhD

Objective—To assess validity and reliability for a visual analogue scale (VAS) used by owners to measure chronic pain in their osteoarthritic dogs.

Sample—68, 61, and 34 owners who completed a questionnaire.

Procedures—Owners answered questionnaires at 5 time points. Criterion validity of the VAS was evaluated for all dogs in the intended-to-treat population by correlating scores for the VAS with scores for the validated Helsinki Chronic Pain Index (HCPI) and a relative quality-of-life scale. Intraclass correlation was used to assess repeatability of the pain VAS at 2 baseline evaluations. To determine sensitivity to change and face validity of the VAS, 2 blinded, randomized control groups (17 dogs receiving carprofen and 17 receiving a placebo) were analyzed over time.

Results—Significant correlations existed between the VAS score and the quality-of-life scale and HCPI scores. Intraclass coefficient (r = 0.72; 95% confidence interval, 0.57 to 0.82) for the VAS indicated good repeatability. In the carprofen and placebo groups, there was poor correlation between the 2 pain evaluation methods (VAS and HCPI items) at the baseline evaluation, but the correlation improved in the carprofen group over time. No correlation was detected for the placebo group over time.

Conclusions and Clinical Relevance—Although valid and reliable, the pain VAS was a poor tool for untrained owners because of poor face validity (ie, owners could not recognize their dogs’ behavior as signs of pain). Only after owners had seen pain diminish and then return (after starting and discontinuing NSAID use) did the VAS have face validity. (Am J Vet Res 2011;72:601–607)

Osteoarthritis is a major cause of chronic pain in dogs. It is estimated that 20% of the dogs in the United Kingdom and United States have pain attributable to osteoarthritis. Canine patients with chronic pain can be apprehensive or excited when they are examined by a veterinarian and may mask signs of pain in a clinic environment. Therefore, owners’ assessments of their dogs’ pain in the home environment may be a better measurement and should be taken into consideration by practitioners and researchers. Investigators have evaluated owner-assessed subjective measurement scales of chronic pain in an effort to test them for validity and reliability.

The VAS is a tool commonly used for owner assessment of an animal’s pain in veterinary research. A VAS is a 100-mm line with 2 end points; observers place a mark on the line corresponding to their interpretation of the patient’s pain intensity. The original pain VAS, as applied to human patients able to respond to verbal commands and who could describe their own amount of pain, had the end points of no pain (left end of the line) and worst possible pain or pain could not be worse (right end of the line). Questionnaires on pain, mood, behavior, and lameness have been provided to dog owners in a VAS format, and at least 1 such VAS questionnaire has been tested for reliability and validity and found to be psychometrically sound. Also, the original pain VAS has been used in research in dogs, but to our knowledge, it has not been validated for use by dog owners.

To be a good measure, a scale must be both valid and reliable. Validity is the quality of the scale and is often divided into various types. Face validity.

Received November 24, 2009.
Accepted March 22, 2010.

From the Department of Equine and Small Animal Medicine, Faculty of Veterinary Medicine (Hielm-Björkman), and the Department of Forest Resource Management, Faculty of Agriculture and Forestry/Statistics (Rita), University of Helsinki, 00014 Helsinki, Finland; and the Department of Surgical and Radiological Sciences, School of Veterinary Medicine, University of California-Davis, Davis, CA 95616 (Kapatkin). The study was performed at the Department of Equine and Small Animal Medicine, University of Helsinki. Supported by the Helvi Knuuttila Foundation. Address correspondence to Dr. Hielm-Björkman (anna.hielm-bjorkman@helsinki.fi).
Materials and Methods

Animals—Sixty-eight dogs with chronic pain attributable to osteoarthritis were used in the study. Dogs were included if they had clinical lameness and radiographic changes as a result of moderate or severe osteoarthritis in a hip joint or elbow joint. All dogs were part of an unrelated clinical trial.17,18 Owners provided written informed consent, and the study was approved by the University of Helsinki Ethical Board.

Study design—The unrelated trial17,18 was designed as a 4-group, randomized, double-blind clinical trial in which 2 new pain treatments were tested and both a negative control treatment (ie, placebo) and a positive control treatment (ie, carprofen) were used. Data from the 2 baseline evaluations (before any treatment) were used in the criterion validity study (n = 68). For the analyses on sensitivity to change in this study, only the dogs of the placebo and carprofen groups were used (n = 34). The 2 other treatment groups were irrelevant for the study reported here.

Four weeks before the treatment phase started (first baseline [week –4]) and again when the treatment phase started (second baseline [week 0]), each owner completed a pain assessment questionnaire. All pretreatment analgesics were recorded in the questionnaires. At the second baseline, dogs were also assigned into the 4 groups (2 control and 2 treatment groups; n = 17 dogs/group) by use of a computer-generated random number list. From the start of the treatment phase (week 0), dogs in the 2 control groups received carprofen2 (2 mg/kg, q 12 h) or a placebo for 8 weeks. For ethical reasons, all owners were additionally provided with rescue analgesics in the form of carprofen tablets (50 mg/tablet) at the start of the trial. This rescue analgesic could be used as additional pain relief (1 tablet for a dog with a body weight of 20 to 30 kg, 2 tablets for a dog with a body weight of 31 to 40 kg, and 3 tablets for a dog with a body weight of 41 to 60 kg) if the owner believed that their dog had pain. Administration of rescue analgesia was also recorded. All dogs were reassessed at weeks 4 and 8 (during carprofen or placebo treatment) and at week 12 (follow-up).

Only 61 dogs were used in the repeatability study because some dogs were excluded on the basis of 2 factors. We wanted to evaluate the pain VAS by use of the test-retest method; thus, some dogs were excluded because they lacked data for both baseline evaluations (weeks –4 and 0). Amount of NSAIDs that had been given to the dogs by their owners 4 weeks prior to each of the 2 baseline evaluations was also determined.

Owners that reported an NSAID medication change that was >1 step on a 5-step medication scale (1 = no NSAIDs during past 4 weeks, 2 = NSAIDs 1 to 2 times during past 4 weeks, 3 = NSAIDs approx once per week during past 4 weeks, 4 = NSAIDs approx 3 to 3 times/wk during past 4 weeks, and 5 = NSAIDs daily or almost daily during past 4 weeks) were omitted from the repeatability analyses.

Owners were blinded to the treatments until the end of the study. Veterinarians and technical personnel were also blinded to treatment and therefore could not influence the responses owners provided.

Pain assessment questionnaire—All owner material and questionnaires were in the Finnish language. The basic pain assessment questionnaire was composed of 2 parts and the one used later in the study was composed of 3 parts. Part 1 was the HCPI14 and consisted of 11 questions about mood, behavior, and locomotion of the dog. The HCPI has been validated and tested for reliability, and it has been found that changes in the index correlate with changes in the chronic pain level of a dog.15 Part 2 was a pain VAS score, with the left end of the line (0 cm) indicative of no pain and the right end of the line (10 cm) indicative of the worst possible pain. Values for the VAS were reported to an
accuracy of 0.1 cm and were intended to correspond to the level of pain that owners believed their dogs had. At weeks 4, 8, and 12, the pain assessment questionnaire also contained part 3, which consisted of a question on change in QOL. Part 3 used a standard 3-point relative response for the following question: Compared to before the beginning of this trial, the dog’s QOL is now: 1 = much better, 2 = a bit better, 3 = the same, 4 = a bit worse, or 5 = much worse.

All evaluators were owners who lived in the same households as the dogs. None of the dog owners had received training as evaluators. Owners provided responses for the HCPI and completed a VAS with regard to their dogs 5 times and answered the QOL questions 3 times. Owners were instructed that the same person should complete the pain assessment questionnaire each time; owners were asked to sign each completed questionnaire to enable investigators to verify this.

The first pain assessment questionnaire was completed at the first baseline evaluation (week −4). Owners were provided the HCPI and VAS and given instructions on how to answer the questions and complete the VAS. Owners were not given guidance with regard to interpretation of the questions, and there was nothing in the materials to indicate that the HCPI questions would correlate with pain in any way. None of the headings on the questionnaire included words such as pain or assessment to avoid respondent bias.30 The headings on the questionnaire included words such as pain or assessment to avoid respondent bias.30 The VAS. Owners were not given guidance with regard to interpretation of the questions, and there was nothing in the materials to indicate that the HCPI questions would correlate with pain in any way. None of the headings on the questionnaire included words such as pain or assessment to avoid respondent bias.30 The second pain assessment questionnaire was completed 4 weeks later (second baseline evaluation [week 0]), before treatments or placebo were administered. The third and fourth pain assessment questionnaires were completed at weeks 4 and 8, which were 4 and 8 weeks after the start of carprofen or placebo administration. The fifth questionnaire was completed at week 12 and was a follow-up evaluation 4 weeks after discontinuation of all treatments.

Statistical analysis—Data for the VAS and HCPI were tested for normality by use of the White test. Baseline bias between the carprofen and placebo groups was assessed by use of a χ2 test and cross tabulation for non-parametric variables or a t test for parametric variables. We controlled for age and duration of signs in the analysis of treatment effect.

Criterion validity of the pain VAS was tested against 2 external pain measurement tools. First, scores for the pain VAS were correlated with scores for the HCPI at the same time points. Then, the change in the pain VAS from evaluation at weeks 0 to 8 was compared with the change in the QOL variable for the same time interval. All tests were performed with the Spearman rank correlation test.

Repeatability of the pain VAS was tested via a test-retest method by use of intraclass correlation.29 The repeatability coefficient was calculated as 2 SDs of the differences, with the assumption that 95% of the differences should be within the repeatability interval of ± 1 SD.31 An error of ± 0.7 cm was allowed.21

Sensitivity to change of the VAS was studied via 2 independent samples with the Mann-Whitney U test to compare VAS values between the carprofen-treated and placebo-treated groups at all 5 time points. Lower scores for the VAS indicated less pain, and higher scores for the VAS indicated more pain. Lower scores for the VAS in the carprofen-treated group during the period of medication administration were indicative of the index’s sensitivity to change. Differences in VAS scores between treatment and placebo groups were analyzed.

To test face validity, the mean VAS scores were also correlated (Spearman rank correlation test) with the summed values for the HCPI and each of the 11 HCPI items at all 5 time points for the 2 groups (carprofen and placebo). Because sample size is of importance for correlation calculations, a sample size table was used.31 By use of α = 0.05 and β = 0.2, a significant (P < 0.05) correlation of r > 0.35 is meaningful for a cohort of 68 and a significant correlation of r > 0.64 is meaningful for a cohort of approximately 17.

All analyses were 2-tailed tests. Significance was set at P < 0.05. To avoid postrandomization selection bias, analyses were made with all dogs in the ITT population and use of rescue analgesia was not taken into consideration in the criterion validity study.32 Some dogs had to be excluded to enable us to test repeatability. Controlling for variables with baseline bias was accomplished by use of a statistical program, whereas all other tests and calculations were performed by use of 2 versions of another statistical program.34

Table 1—Comparison of values among groups of dogs with osteoarthritis at week 0 (time of second baseline evaluation and initiation of product administration).

<table>
<thead>
<tr>
<th>Factor</th>
<th>All dogs</th>
<th>Carprofen</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of dogs</td>
<td>68</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>No. with dysplasia of a hip joint</td>
<td>58</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Sex</td>
<td>male/female</td>
<td>37/31</td>
<td>8/9</td>
</tr>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>6</td>
<td>5a</td>
<td>6b</td>
</tr>
<tr>
<td>Range</td>
<td>1–11</td>
<td>1–9</td>
<td>1–11</td>
</tr>
<tr>
<td>Duration of signs (mo)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>&gt; 24</td>
<td>&gt; 24a</td>
<td>12–24b</td>
</tr>
<tr>
<td>Range</td>
<td>1–&gt; 24</td>
<td>12–&gt; 24</td>
<td>1–&gt; 24</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>34</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>Range</td>
<td>18–60</td>
<td>28–56</td>
<td>18–54</td>
</tr>
<tr>
<td>Mean ± SD HCPI score at week 0</td>
<td>15.96 ± 5.27</td>
<td>16.71 ± 5.91</td>
<td>15.29 ± 5.68</td>
</tr>
<tr>
<td>Mean ± SD pain VAS score at week 0 (cm)</td>
<td>3.90 ± 1.91</td>
<td>3.69 ± 2.06</td>
<td>3.69 ± 1.88</td>
</tr>
</tbody>
</table>

*Within a row, values with different superscript letters differ significantly (P < 0.05).
Results

Baseline evaluation—Baseline data for the 68 dogs as well as the 2 groups (17 dogs in each of the carprofen and placebo groups) were summarized (Table 1). There were 25 breeds in the cohort of 68 dogs (17 German Shepherd Dogs, 5 Rottweilers, 5 Golden Retrievers, 3 Newfoundland, 4 Boxers, 4 Samoyeds, and 1 to 3 dogs for all other breeds). There was no significant difference at baseline between the carprofen and placebo groups for pain VAS score, HCPI total score, breed distribution, number of dogs with osteoarthritis in the forelimbs or hind limbs, sex, or body weight. However, there were significant differences in age and duration of signs of osteoarthritis between the carprofen and placebo groups. When we controlled for age and duration of signs, all results were similar and significant and nonsignificant findings remained. Normality evaluations revealed that both VAS and HCPI data at week 0 were normally distributed.

One page of the HCPI was missing for 13 of the 68 dogs at the evaluation conducted at week 4; thus, data for those questions were not used in all evaluations.

This was evident as a lower ITT number and in week 4 data (Table 2).

Criterion validity—A significant (P ≤ 0.001 for all analyses) Spearman correlation was detected between the VAS score and total HCPI score at each of the 5 time points (week –4, r = 0.45; week 0, r = 0.40; week 4, r = 0.49; week 8, r = 0.66; and week 12, r = 0.71 [n = 52 to 68 dogs in the ITT population]). The correlation for the changes in VAS and total HCPI scores from week 0 to week 8 (baseline to 8 weeks after start of treatment) was significant (r = 0.65; P < 0.001) for all dogs. There was also a significant correlation (r = 0.48; P < 0.001 [n = 67 dogs]) between the change in VAS score and the change in QOL score from week 0 to week 8.

Repeatability or intraobserver reliability—Analysis of the correlation between the pain VAS score at week –4 and that at week 0 (n = 61 dogs) revealed an intraclass correlation of 0.72 (95% confidence interval, 0.57 to 0.82). The mean change between the 2 baseline evaluations was 0.09, and the range was –3.5 to 3.7 cm. The repeatability coefficient for the VAS was 2.8 (SD was 0.4).

![Table 2](https://example.com/table2.png)
Sensitivity to change—We did not detect a significant difference between the carprofen and placebo groups in owner assessment of pain by use of the VAS at the 2 baseline evaluations at week –4 (P = 0.830) and week 0 (P = 0.817) and at the follow-up evaluation at week 12 (P = 0.387). However, there was a significant (P < 0.001) difference in owner assessment of pain determined by use of the VAS between the groups for both evaluations performed during treatment at weeks 4 and 8.

Face validity—Analysis of the correlation between pain VAS score and score for each of the 11 HCPI items for the carprofen and placebo groups separately revealed that correlations differed greatly depending on group and over time (Table 2). At the baseline evaluations at weeks –4 and 0, the pain VAS score for carprofen-treated dogs correlated with the score for only 1 HCPI item (the dog’s difficulty in movement after a long rest). At week 4, the pain VAS score for carprofen-treated dogs correlated with the scores for 2 HCPI items (dog’s willingness to walk and difficulty in movement after major activity or heavy exercise). At week 8, the pain VAS score for carprofen-treated dogs was again correlated with the scores for 2 HCPI items (dog’s vocalization in the form of audible complaining and difficulty in movement after major activity or heavy exercise). At week 12, the pain VAS score for carprofen-treated dogs was correlated with the scores of 8 HCPI items (dog’s mood, willingness to walk, willingness to trot, willingness to gallop, ease in lying down, ease in getting up from a lying position, difficulty in movement after a long rest, and difficulty in movement after major activity or heavy exercise).

No significant correlations were detected between pain VAS score and score for any of the HCPI items for the placebo group at any of the time points.

Discussion

In the study reported here, validity and reliability of a pain VAS used by untrained dog owners to measure chronic pain in their dogs were evaluated. Three of the analyses indicated that the VAS was valid and reliable, but the repeatability coefficient was not excellent, and sufficient face validity was not established.

Criterion validity correlation was one of the validity tests performed. The score for the pain VAS was increasingly correlated with the total score of the HCPI over time. This was evident in data for the ITT population (n = 68 dogs). As was found later in the study, owners learned to use the VAS only after providing their dogs with an effective pain treatment and then discontinuing that treatment. Thus, this increasing correlation can be explained by the fact that 51 of the dogs were given some type of effective pain treatment (carprofen or 2 other pain-relieving products [a green-lipped mussel and a homeopathic product] that were administered in the clinical part of the study17,18).

Repeatability analyses revealed that the mean difference between the 2 baseline evaluations conducted 4 weeks apart was 0.9 cm (± 10% of the VAS). This appeared to be good repeatability. However, the range of differences (–3.5 to 3.7 cm) and SD (1.39 cm) were large, and the mean ± SD error of an individual VAS score was 2.8 ± 1.39 cm. This imprecision of ± 1.4 cm is twice the acceptable amount of change of visual and motor error of an individual score reported in another study,31 which suggests that the repeatability is not as good as it first appeared. Potential reasons for this discrepancy for imprecision between the study reported here and other reports31 could have been the 4-week interval between the 2 baseline evaluations or the undulating nature of clinical signs of osteoarthritis in dogs. Investigators did their best to control test conditions (eg, pain medications, nutritional products, or diet) between the 2 baseline evaluations, but conditions such as weather could still have influenced results of this clinical trial.

Sensitivity to change was established. We did not detect a significant difference in the pain VAS scores between the carprofen- and placebo-treated dogs at the 2 baselines evaluations (weeks –4 and 0) or at the follow-up evaluation (week 12). However, there was a significant difference between the groups for evaluations during the treatment period (weeks 4 and 8), which indicated that the VAS was sensitive to change and could detect a clear treatment effect that was evident only in the carprofen-treated dogs.

Finally, when the criterion validity and the sensitivity to change data were combined (Table 2), it was possible to determine that there were no significant meaningful correlations between pain VAS score and scores for each of the HCPI items in the placebo group at any time (although results appeared to be significant, the sample size and the correlation coefficient made these results not meaningful31). However, there were an increasing number of HCPI item scores that correlated with the pain VAS scores in the carprofen-treated group.

When evaluating specific items, there was only 1 significant correlation between groups detected during the baseline evaluation: difficulty to move after rest for the carprofen-treated group at week 0. Because only 1 variable of 22 had a significant correlation, and because there was no baseline bias between groups, this could have been the result of chance. The next significant correlation was for difficulty to move after major activity or heavy exercise, which remained significantly correlated with the VAS score at both evaluations during treatment (weeks 4 and 8) and at the follow-up evaluation (week 12) and was therefore probably a recognized sign of pain. Vocalization was a sign recognized by owners after initial carprofen treatment, but it was not correlated with VAS score at the follow-up evaluation in this study. Therefore, the initial correlation was probably not meaningful. Dogs vocalize for a variety of reasons; therefore, it is usually a poor indicator of chronic pain. Most (8/11) correlations were significant and meaningful for the HCPI items only at the follow-up evaluation (week 12), which was the evaluation conducted after treatment had been discontinued for 4 weeks.

In the placebo-treated group, the owners’ naiveté toward behaviors indicative of pain remained similar throughout the study period (baseline, treatment, and follow-up evaluations), which was indicated by no correlation between the 2 scales. On the basis of these results, the face validity of the pain VAS used by
untrained owners was weak. The VAS relies on a dog owner to detect a behavioral response to chronic pain in their dog, and clearly, untrained owners did not realize that their dogs were displaying signs of chronic pain. In this study, the pain VAS proved to be a useful instrument only after owners had become self-trained by seeing obvious changes in their dog’s behavior and lameness attributable to administration of carprofen and then withdrawal of it. Therefore, we reject the hypothesis that the pain VAS is a valid and reliable tool for pain evaluation by untrained owners. To this end, we believe that veterinarians could teach owners to recognize their dogs’ pain by first administering pain medication and then withdrawing it.

References
1. Lascelles BDx, Main DCJ. Surgical trauma and chronically painful conditions—within our comfort level but beyond theirs? J Am Vet Med Assoc 2002;221:213–222.