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.[1] 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.[2] 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.[3–5] Investigators have evaluated owner-assessed subjective measurement scales of chronic pain in an effort to test them for validity and reliability.[6–13]

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).

Abbreviations

HCPI	Helsinki Chronic Pain Index
ITT	Intended to treat
QOL	Quality of life
VAS	Visual analogue scale

The VAS is a tool commonly used for owner assessment of an animal's pain in veterinary research.[6,7,9,13–20] 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).[21] 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.[9] 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.[22] Validity is the quality of the scale and is often divided into various types. Face validity[22–25]

of a pain VAS requires each evaluator to recognize pain. Criterion validity[22–25] is used when describing the correlation between a scale and another external measurement of the same phenomenon. Objective external evaluation tools, such as force plates,[26] pressure carpets,[27] or accelerometers,[28] are not always available; therefore, less objective measurement tools have been used.[11,13]

Reliability can be assessed via repeatability and sensitivity to change. Repeatability testing by the test-retest method requires 2 evaluations. Observers use the same scale twice under identical conditions, and the results are correlated via intraclass correlation.[29] Repeatability can also be evaluated according to the repeatability coefficient, which is equal to 2 SDs and to a 95% confidence limit of ± 1 SD.[29] In 1 study,[21] investigators suggested a confidence limit of ± 0.7 cm for a specific mark on a human self-rating pain VAS. Sensitivity to change reflects the capability of the instrument to measure changes in the amount of pain over time in response to clinical interventions.[23] For example, 1 group would receive an analgesic treatment (eg, an NSAID for osteoarthritis), whereas a second group would receive a placebo; it is presumed that the analgesic would affect the amount of pain more than would the placebo.

The purpose of the study reported here was to evaluate the validity and reliability of the original pain VAS used by owners. This was accomplished by use of scales and questionnaires completed at 5 time points by owners of dogs with chronic pain attributable to osteoarthritis. To test criterion validity of the original pain VAS used by owners, it was compared with 2 other external measures of the same phenomenon (pain): the HCPI and a QOL scale. To test reliability, a test-retest and reliability coefficient setting was chosen. Sensitivity to change was evaluated by 2 extreme groups: one using an NSAID and the other a placebo. To test face validity, the VAS was compared with another external measure, the HCPI, in 2 extreme groups (carprofen and placebo) and evaluated over time (ie, during no medication [2 baselines], medication phase [2 treatment-phase measurements], and again during no medication [follow-up]). We hypothesized that the VAS would have good validity, test-retest repeatability, and sensitivity to detect changes in pain. We also hypothesized that when combining the criterion validity tests with the sensitivity to change test, we would be able to determine the face validity of the VAS and the evolution of owners as evaluators of pain.

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 trial[17,18] 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 pretrial 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 carprofen[a] (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 5 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 HCPI[13] 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.[13] 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 5-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 second pain assessment questionnaire was completed 4 weeks later (second baseline evaluation [week 0]), before treatments or placebos were administered. The third and fourth pain assessment questionnaires were completed at weeks 4 and 8, which was 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 χ^2 test and cross tabulation for nonparametric 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.[25] 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 $\alpha = 0.05$ and $\beta = 0.2$, 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,[b] whereas all other tests and calculations were performed by use of 2 versions of another statistical program.[c,d]

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

Factor	All dogs	Carprofen	Placebo
No. of dogs	68	17	17
No. with dysplasia of a hip joint	58	14	14
No. with osteoarthritis of an elbow joint	10	3	3
Sex (male/female)	37/31	8/9	12/5
Age (y)			
Median	6	5[a]	6[b]
Range	1–11	1–9	1–11
Duration of signs (mo)			
Median	> 24	> 24[a]	12–24[b]
Range	1–> 24	12–> 24	1–> 24
Weight (kg)			
Median	34	35	34
Range	18–60	28–56	18–54
Mean ± SD HCPI score at week 0	15.96 ± 5.37	16.71 ± 5.91	15.29 ± 4.65
Mean ± SD pain VAS score at week 0 (cm)	3.90 ± 1.91	3.69 ± 2.06	3.69 ± 1.66

[a,b]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, 5 Newfoundlands, 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 hinds, 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 ($P \leq 0.001$ 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 ± 1.4).

Table 2—Analysis of correlations between pain VAS score and score for each HCPI item and total HCPI score assessed by owners of dogs with osteoarthritis at various time points.*

		All dogs†		Carprofen‡					Placebo‡				
HCPI item	Variable	Week –4	Week 0	Week –4	Week 0	Week 4	Week 8	Week 12	Week –4	Week 0	Week 4	Week 8	Week 12
Mood	*r*	0.07	0.13	–0.10	0.31	0.31	0.50	0.78	0.10	–0.16	0.27	0.41	0.49
	P value	0.556	0.310	0.720	0.237	0.330	0.040	< 0.001	0.733	0.531	0.351	0.103	0.053
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Play	*r*	0.04	0.19	0.01	0.48	0.33	0.35	0.42	0.05	0.09	–0.05	0.22	0.38
	P value	0.777	0.115	0.981	0.060	0.288	0.166	0.091	0.871	0.737	0.854	0.385	0.143
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Vocalization	*r*	0.31	0.19	0.25	–0.09	0.67§	–0.06	–0.21	0.53	0.33	0.37	0.42	0.03
	P value	0.011	0.121	0.357	0.729	0.018	0.827	0.413	0.041	0.190	0.195	0.096	0.919
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Walking	*r*	0.32	0.27	0.39	0.24	0.54	0.67§	0.77§	0.41	0.23	–0.12	0.16	0.51
	P value	0.010	0.027	0.138	0.364	0.071	0.004	< 0.001	0.126	0.367	0.686	0.544	0.046
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Trotting	*r*	0.34	0.42§	0.43	0.49	0.40	0.51	0.79§	0.31	0.12	0.20	0.44	0.62
	P value	0.006	< 0.001	0.097	0.056	0.204	0.035	< 0.001	0.256	0.659	0.483	0.078	0.011
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Galloping	*r*	0.05	0.12	–0.22	0.35	0.23	0.14	0.67§	–0.15	–0.25	–0.38	–0.07	0.26
	P value	0.712	0.322	0.413	0.190	0.472	0.580	0.004	0.582	0.343	0.186	0.800	0.329
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Jumping	*r*	0.39§	0.21	0.14	0.02	0.00	–0.04	0.49	0.29	0.09	0.37	0.00	–0.17
	P value	0.001	0.081	0.605	0.953	0.990	0.888	0.044	0.293	0.735	0.199	0.992	0.534
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Lying down	*r*	0.31	0.31	0.54	0.52	–0.13	0.34	0.66§	0.40	–0.15	0.42	0.25	0.09
	P value	0.013	0.011	0.030	0.041	0.696	0.188	0.004	0.135	0.564	0.133	0.343	0.738
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Getting up	*r*	0.32	0.32	0.56	0.57	0.18	0.16	0.72§	0.13	–0.12	–0.25	0.17	0.21
	P value	0.011	0.008	0.024	0.021	0.575	0.552	0.001	0.640	0.641	0.387	0.505	0.445
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Difficulty after rest	*r*	0.36§	0.33	0.18	0.66§	0.02	–0.03	0.64§	0.47	–0.06	–0.25	–0.12	0.26
	P value	0.004	0.007	0.516	0.006	0.942	0.899	0.006	0.079	0.808	0.392	0.642	0.333
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Difficulty after activity	*r*	0.18	0.17	0.02	–0.08	0.80§	0.66§	0.66§	0.20	0.00	–0.07	0.30	0.14
	P value	0.147	0.175	0.945	0.757	0.002	0.004	0.004	0.479	1.000	0.823	0.237	0.615
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16
Total HCPI	*r*	0.45§	0.40§	0.27	0.44	0.44	0.50	0.82§	0.38	–0.13	–0.01	0.37	0.48
	P value	< 0.001	0.001	0.304	0.085	0.149	0.040	< 0.001	0.167	0.631	0.982	0.139	0.059
	No. of dogs	64	67	16	16	12	17	17	15	17	14	17	16

*Owners completed a VAS and HCPI at 5 time points (2 baseline evaluations [week –4 and week 0], 2 evaluations during treatment [week 4 and week 8], and an evaluation 4 weeks after cessation of treatment [week 12]); onset of treatment was designated as week 0. †Represents 68 dogs in the ITT population; some dogs were excluded from some analyses. ‡Represents 17 dogs in each group (carprofen-treated dogs received carprofen [2 mg/kg, q 12 h for 8 weeks] and placebo-treated dogs received a placebo for 8 weeks); some dogs in each group were excluded from analyses because of a variety of reasons. §Significant ($P < 0.05$) correlations of $r > 0.35$ for a cohort of 68 and $r > 0.64$ for a cohort of 17 were considered meaningful.[32]

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 vocalization in the form of audible complaining 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 willingness to walk 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[e] and a homeopathic product[f]] that were administered in the clinical part of the study[17,18]).

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,[21] 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 reports[21] 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 meaningful[31]). 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' naïvete 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 certain behavioral and locomotion changes were caused by pain. The same phenomenon has been reported in the human literature. When chronic pain in children caused by juvenile rheumatoid arthritis was evaluated via the VAS by both physicians and parents, there was a correlation coefficient of only $r = 0.1$ between the 2 observer groups.[33] This indicates that one of these 2 groups of observers had trouble detecting chronic pain in children.[33] In a meta-analysis[34] of studies that used a VAS for evaluation of observed pediatric pain, the correlation coefficient (interobserver reliability) between parents and medical personnel was $r = 0.55$. Veterinary medicine is just beginning to develop tools for subjective chronic pain assessment that can be used by owners. Such questionnaires, indexes, or scales require extensive research through use of focus groups, item testing, and applications to specific clinical conditions before they will be useful tools.[6–13]

We concluded that the 3 first psychometric analyses conducted all indicated that the VAS could be a useful pain assessment instrument. However, the combined analysis indicated that the owner-assessed pain VAS lacked face validity, which is the very cornerstone of a valid instrument. This further emphasizes the need to adequately test existing instruments and also to develop new ones for use in veterinary clinical trials.[35–40] Poor correlation between scores for each HCPI item and the pain VAS score in the placebo-treated group indicated that the pain-naïve 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.

a. Rimadyl, 50 mg, Pfizer, Espoo, Finland.
b. StatXact-8, Cytel Software Corp, Cambridge, Mass.
c. SPSS 12.0, SPSS Inc, Chicago, Ill.
d. SPSS 15.0, SPSS Inc, Chicago, Ill.
e. Lyproflex, 500 mg, ICENI, OMNI Nutraceuticals, Cambridgeshire, Biofarm Oy, Finland.
f. Zeel ad us vet, 5 mL, Biologische Heilmittel Heel GmbH, Baden-Baden, Germany.

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;221:215–222.
2. Dobromylskyj P, Flecknell PA, Lascelles BD, et al. Pain assessment. In: Flecknell PA, Waterman-Pearson A, eds. *Pain management in animals*. London: WB Saunders Co, 2000;53–79.
3. American College of Veterinary Anesthesiologists. American College of Veterinary Anesthesiologists' position paper on the treatment of pain in animals. *J Am Vet Med Assoc* 1998;213:628–630.
4. Hardie EM. Recognition of pain behavior in animals. In: Hellebrekers LJ, ed. *Animal pain*. Utrecht, The Netherlands: Van der Wees, 2000;51–69.
5. Wiseman ML, Nolan AM, Reid J, et al. Preliminary study on owner-reported behaviour changes associated with chronic pain in dogs. *Vet Rec* 2001;149:423–424.
6. Innes JF, Barr ARS. Can owners assess outcome following treatment of canine cruciate ligament deficiency? *J Small Anim Pract* 1998;39:373–378.
7. Hielm-Björkman A, Kuusela E, Liman A, et al. Evaluation of methods for assessment of pain associated with chronic osteoarthritis in dogs. *J Am Vet Med Assoc* 2003;222:1552–1558.
8. Wiseman-Orr ML, Nolan AM, Reid J, et al. Development of a questionnaire to measure the effects of chronic pain on health-related quality of life in dogs. *Am J Vet Res* 2004;65:1077–1084.
9. Hudson JT, Slater MR, Taylor L, et al. Assessing repeatability and validity of a visual analogue scale questionnaire for use in assessing pain and lameness in dogs. *Am J Vet Res* 2004;65:1634–1643.
10. Wiseman-Orr ML, Scott EM, Reid J, et al. Validation of a structured questionnaire as an instrument to measure chronic pain in dogs on the basis of effects on health-related quality of life. *Am J Vet Res* 2006;67:1826–1836.
11. Brown DC, Boston RC, Coyne JC, et al. Development and psychometric testing of an instrument designed to measure chronic pain in dogs with osteoarthritis. *Am J Vet Res* 2007;68:631–637.
12. Brown DC, Boston RC, Coyne JC, et al. Ability of the canine brief pain inventory to detect response to treatment in dogs with osteoarthritis. J *Am Vet Med Assoc* 2008;233:1278–1283.
13. Hielm-Björkman A, Rita H, Tulamo R-M. Psychometric testing of the Helsinki chronic pain index by completion of a questionnaire in Finnish by owners of dogs with chronic signs of pain caused by osteoarthritis. *Am J Vet Res* 2009;70:727–734.
14. Welsh EM, Gettinby G, Nolan AM. Comparison of a visual analogue scale and a numerical rating scale for assessment of lameness, using sheep as the model. *Am J Vet Res* 1993;54:976–983.
15. Conzemius MG, Hill CM, Sammarco JL, et al. Correlation between subjective and objective measures used to determine severity of postoperative pain in dogs. *J Am Vet Med Assoc* 1997;210:1619–1622.
16. Holton L, Scott EM, Nolan AM, et al. Comparison of three methods used for assessment of pain in dogs. *J Am Vet Med Assoc* 1998;212:61–66.
17. Hielm-Björkman A, Tulamo R-M, Salonen H, et al. Evaluating complementary therapies for canine osteoarthritis. Part I: Green-lipped Mussel (*Perna canaliculus*). *Evid Based Complement Alternat Med* 2009;6:365–373.
18. Hielm-Björkman A, Tulamo R-M, Salonen H, et al. Evaluating complementary therapies for canine osteoarthritis. Part II: a homeopathic combination preparation (Zeel). *Evid Based Complement Alternat Med* 2009;6:465–471.
19. Quinn MM, Keuler NS, Lu Y, et al. Evaluation of agreement between numerical rating scales, visual analogue scoring scales, and force plate gait analysis in dogs. *Vet Surg* 2007;36:360–367.
20. Waxman AS, Robinson DA, Evans RB, et al. Relationship between objective and subjective assessment of limb function in normal dogs with an experimentally induced lameness. *Vet Surg* 2008;37:241–246.
21. Revill SI, Robinson JO, Rosen M, et al. The reliability of a linear analogue for evaluating pain. *Anaesthesia* 1976;31:1191–1198.
22. Carmines EG, Zeller RA. *Reliability and validity assessment*. Thousand Oaks, Calif: Sage Publications, 1979;9–70.
23. Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use*. 2nd ed. New York: Oxford University Press, 1995;4–161.
24. Nunnally JC. *Psychometric theory*. 2nd ed. New York: McGraw-Hill, 1978;86–323.
25. DeVellis RF. *Scale development—theory and applications. Applied social research methods series*. Vol 26. 2nd ed. Newbury Park, Calif: Sage Publications, 2003;1–137.
26. Budsberg SC. Long-term temporal evaluation of ground reaction forces during development of experimentally induced osteoarthritis in dogs. *Am J Vet Res* 2001;62:1207–1211.
27. Brach JS, Perera S, Studenski S, et al. The reliability and validity of measures of gait variability in community-dwelling older adults. *Arch Phys Med Rehabil* 2008;89:2293–2296.
28. Hansen BD, Lascelles BD, Keene BW, et al. Evaluation of an ac-

celerometer for at-home monitoring of spontaneous activity in dogs. *Am J Vet Res* 2007;68:468–475.
29. Cleophas TJ, Zwinderman AH, Cleophas TF. Testing reproducibility. In: *Statistics applied to clinical trials*. 3rd ed. Dordrecht, The Netherlands: Springer, 2006;307–310.
30. Vaillancourt JP, Martineau G, Morrow M, et al. Construction of questionnaires and their use in veterinary medicine, in *Proceedings*. Soc Vet Epidemiol Prevent Med 1991;94–106.
31. Norman GR, Streiner DL. Appendix: Table G. In: *Biostatistics—the bare essentials*. Hamilton, ON, Canada: BC Decker Inc, 2008;365.
32. Brown DC. Sources and handling of losses to follow-up in parallel-group randomized clinical trials in dogs and cats: 63 trials (2000–2005). *Am J Vet Res* 2007;68:694–698.
33. Huijer Abu-Saad HH, Uiterwijk M. Pain in children with juvenile rheumatoid arthritis: a descriptive study. *Pediatr Res* 1995;38:194–197.
34. van Dijk M, Koot HM, Huijer Abu-Saad HH, et al. Observational visual analog scale in pediatric pain assessment: useful tool or good riddance? *Clin J Pain* 2002;18:310–316.
35. Schulz KS, Cook JL, Kapatkin A, et al. Evidence-based surgery: time for change. *Vet Surg* 2006;35:697–699.
36. Cook JL. Outcomes-based patient care in veterinary surgery: what is an outcome measure? *Vet Surg* 2007;36:187–188.
37. Brown DC. Outcomes-based medicine in veterinary surgery: getting hard measures of subjective outcomes. *Vet Surg* 2007;36:289–292.
38. Kapatkin AS. Outcome-based medicine and its application in clinical surgical practice. *Vet Surg* 2007;36:515–518.
39. Innes JF. Outcomes-based medicine in veterinary surgery: levels of evidence. *Vet Surg* 2007;36:610–612.
40. Schulz KS. The outcomes measures program: what's in it for you? *Vet Surg* 2007;36:715–716.