Original Article

Journal of Feline Medicine and Surgery
2022, Vol. 24(2) 142–151

DOI: 10.1177/1098612X211011984
journals.sagepub.com/home/jfm

This paper was handled and processed by the American Editorial Office (AAFP) for publication in *JFMS*

Refinement of the Feline Musculoskeletal Pain Index (FMPI) and development of the short-form FMPI

Masataka Enomoto[1], B Duncan X Lascelles[1–5], James B Robertson[6] and Margaret E Gruen[2,3,7]

Abstract

Objectives The aim of this study was to investigate the reliability and responsiveness of the Feline Musculoskeletal Pain Index (FMPI) using the collective results of multiple clinical studies and iteratively refine the FMPI for future use.
Methods Data were compiled from previously conducted studies involving client-owned cats with degenerative joint disease (DJD) and which used the FMPI. The reliability of the FMPI was assessed using the data from the initial visits of those studies. For the assessment of responsiveness of the FMPI, only placebo-controlled studies that used analgesic treatments were included. Treatment groups from each study were combined and categorized as 'placebo' group and 'analgesic' group. Then, the mean change from baseline in score of each FMPI item and across all items within and between these groups was assessed. Based on the results of the reliability and responsiveness of the FMPI, stepwise elimination was used to remove the items that were least able to distinguish between the placebo and analgesic groups. Finally, after the stepwise elimination, a proposed new FMPI-short form (FMPI-sf) was constructed and its reliability was reassessed using the data sets described above. Individual and combined data sets of the studies were also used to compare the responsiveness of the original FMPI and the FMPI-sf.
Results The data from 180 cats from four studies were included. The original FMPI had a reasonable reliability, but low/no responsiveness. The elimination process of FMPI items refined the responsiveness of the instrument while maintaining its reliability. When the responsiveness was compared between the original FMPI (17 items) and the FMPI-sf (nine items), the treatment effect between groups was always greater when the FMPI-sf was used.
Conclusions and relevance The proposed FMPI-sf may be able to better distinguish between placebo and analgesic effects in cats with DJD.

Keywords: Osteoarthritis; pain; musculoskeletal disease; degenerative joint disease; behavioral change; outcome measure

Accepted: 22 March 2021

[1]Translational Research in Pain (TRiP) Program, Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, USA
[2]Comparative Pain Research and Education Centre, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, USA
[3]Comparative Medicine Institute, Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, USA
[4]Thurston Arthritis Centre, UNC School of Medicine, Chapel Hill, NC, USA
[5]Center for Translational Pain Research, Department of Anaesthesiology, Duke University, Durham, NC, USA
[6]Biostatistician, Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, USA
[7]Behavioral Medicine, Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, USA

Corresponding author:
Margaret E Gruen DVM, PhD, DACVB, North Carolina State University College of Veterinary Medicine, 1060 William Moore Drive, Raleigh, NC 27607, USA
Email: megruen@ncsu.edu

Introduction

As veterinary healthcare has advanced, many companion animals are living longer and the prevalence of chronic conditions has increased. In cats, one of the most common chronic conditions is degenerative joint disease (DJD); this condition is associated with negative consequences including pain, mobility impairment and decreased quality of life.[1] Research studies investigating DJD in cats have demonstrated that this condition is common, with an estimated 30–90% of cats affected with radiographic signs, and up to 45% of all cats affected with at least moderate clinical signs associated with chronic pain.[2–6] As such, assessment of chronic pain should be considered fundamental to feline medicine; however, the behavioral changes present with DJD-associated pain develop gradually, are often subtle and may be most evident in the home environment. Thus, in spite of the high prevalence of disease in cats, DJD remains underdiagnosed and undertreated.[7]

Several recent studies have described the development of tools for assessing chronic pain in cats, including objective measures such as activity monitoring and gait analysis.[8–11] These tools show promise, but have practical limitations for clinical use. Thus, the detection of behavioral changes still relies heavily on owner report. Clinical metrology instruments (CMIs) have been designed to standardize the collection of owner reports and facilitate clinical research. There are currently four owner-completed CMIs for use in cats with DJD;[12–16] these are the Feline Musculoskeletal Pain Index (FMPI), Client Specific Outcome Measure (CSOM), Montreal Instrument for Cat Arthritis Testing (MI-CAT), and the Feline Physical Function Formula (FPFF).

The FMPI is the most widely studied of the CMIs; it has been evaluated for construct validity, internal consistency, reliability and discriminatory ability.[16] In its most recent version,[17] it consists of a series of 17 fixed items regarding individual behaviors likely to be affected by DJD-associated pain.[14] Owners are asked to rate their cat's ability to perform each of these behaviors on a Likert scale, and summary scores are generated based on their responses. To date, one weakness of the FMPI is a relative lack of responsiveness (ability to detect the effect of an analgesic treatment) in the form of clinically relevant changes, compared with placebo treatment, after administration of analgesics, particularly when other outcome assessments such as accelerometry have shown an effect.[14]

As the FMPI has been used as an outcome assessment in several recent studies, this is an appropriate time for refinement of the FMPI.[14,18,19] All 17 items included in the current version of the FMPI were selected based on significant differences in answers obtained from owners of cats with DJD-associated pain compared with owners of healthy, non-painful cats; however, not all items may be responsive to treatment effects. In addition, growing awareness of DJD and associated pain in cats is likely to drive interest in owner assessments and development of novel therapeutics. Thus, the objective of this study is to investigate the reliability and responsiveness of the FMPI using the collective results of multiple clinical studies and iteratively refine the FMPI for future use.

Materials and methods

This study used data collected previously at the Translational Research in Pain (TRiP) Program (formerly the Comparative Pain Research Laboratory) at North Carolina State University (NCSU). All data were collected in accordance with the Institutional Care and Use Committee guidelines (IACUC #s 06-056; 08-124; 08-125; 11-102; 14-009; 14-043), and procedures were carried out under informed, written owner consent. The work in the original studies involved the use of only client-owned cats, and followed established internationally recognized best practice standards of individual veterinary clinical patient care.

Study overview

To investigate the reliability and responsiveness of the FMPI, we used data from three published studies[14,18,19] and one unpublished study that included client-owned cats with DJD and which used the FMPI; details are shown in Table 1. These included studies focused on CMI development[14] and studies evaluating the efficacy of potential therapeutics (supplement, low-dose non-steroidal anti-inflammatory drug [NSAID; low-dose][18] and anti-nerve growth factor antibody [antibody][19]). In these studies, all cats had a history of DJD-related mobility impairment and joint pain detected during veterinary orthopedic examination. The presence of DJD was confirmed by radiographs and radiographic features considered indicative of the presence of DJD including osteophytes, enthesophytes, sclerosis, subluxation and mineralization. In each study, cats did not receive any analgesics for at least 2 weeks prior to study enrollment. All owners of cats that participated in the studies signed a study consent form before enrollment and agreed that patient data could be used for research. All cats enrolled in the previous studies were included in the present study; however, cats were excluded from analysis if FMPI data were not available.

Internal consistency, test-retest reliability and repeatability were assessed to investigate the reliability of the FMPI (see Table 2 for brief definitions of terms). For internal consistency, data from the initial visit were evaluated. In the supplement data, the FMPI was not repeated before cats started receiving treatment; therefore, this study was included only for the evaluation of internal consistency. In the other studies, the FMPI was repeated 2 weeks after the initial visit at the end of the baseline period (baseline visit), and these data were used for test-retest reliability and repeatability (Bland–Altman analysis).

Table 1 Studies used to evaluate the reliability of Feline Musculoskeletal Pain Index (FMPI) items and which analyses each study was included in (internal consistency, test-retest reliability and Bland–Altman analysis). The time interval between visits and number of cats in each analysis are also shown

	Time between initial and baseline visits	Total number of cats		
		Internal consistency	Test-retest reliability	Bland–Altman analysis
Supplement (unpublished study)	–	59		
Clinical metrology instrument (CMI)[14]	2 weeks	29	25*	25*
Low-dose[18]	2 weeks	58	58	58
Antibody[19]	2 weeks	34	34	34
Total number of cats		180	117	117

*Four cats were excluded at the baseline visit in the CMI study because they did not meet study inclusion criteria of that study

Table 2 Definition of validity testing terms

Internal consistency	Measure of how closely related a set of items are as a group based on the correlations between different items
Responsiveness	The ability of a questionnaire to detect clinically important changes over time
Discriminatory ability	Ability to differentiate one from the other
Reliability	The extent to which the same test or procedure will yield the same result either over time or with different observers
Repeatability	How well a questionnaire can reproduce an outcome in unchanged conditions
Construct validity	The extent to which a test measures the concept or construct that it is intended to measure
Sensitivity	The ability of a test to correctly identify patients with the disease
Specificity	The ability of a test to correctly identify patients without the disease

For the assessment of responsiveness of the FMPI, only studies that used analgesic treatments for musculoskeletal pain were included;[14,18,19] therefore, the unpublished supplement trial was excluded from the analysis. The three included studies were all placebo-controlled, double-blind, randomized studies, and all cats were reassessed 2–3 weeks after treatment (Table 3). The cats that received meloxicam [CMI and low-dose] or anti-nerve growth factor (NGF) monoclonal antibody (mAb) (antibody) were combined and categorized as having received an analgesic ('analgesic' group). The cats that received a placebo from each study were combined and categorized as the 'placebo' group. The mean change from baseline in score of each FMPI item and across all items within and between these groups were assessed to investigate the responsiveness of the FMPI. Based on the results of the reliability and responsiveness of the FMPI, stepwise elimination was used to remove the items that were least able to distinguish between placebo and analgesic groups. Finally, after the stepwise elimination, a proposed new FMPI-short form (FMPI-sf) was constructed and its reliability was reassessed using the three data sets described above. Individual and combined data sets of three studies were also used to compare the responsiveness of the original FMPI and the new FMPI-sf.

To assess the discriminatory ability of the FMPI-sf, FMPI data from musculoskeletally normal cats were also gathered from another published study.[20] The average FMPI score across 17 items and FMPI-sf items that were obtained at the initial visit from normal cats and DJD cats (from the above four studies) were calculated, and then they were compared.

FMPI score conversion

In the included studies, scoring of the FMPI data had been performed in various ways (for statistical reasons) as previously described.[20] To align each study, the scoring for each was converted to a 0–4 scale (0 = activity performed normally; 4 = activity was impossible to perform). For example, if the original study employed a 5 (normal) to 1 (impossible) scale for each item, the data were converted as shown in Table 4.

Statistical analysis

Statistical analysis was performed using JMP software (JMP Pro 14; SAS), except for the calculation of intraclass correlation coefficient (R package version 0.84.1: available at https://cran.r-project.org/web/packages/irr/index.html). Cronbach's α was used to investigate the internal consistency of each FMPI item and was calculated for

Table 3 Studies used to evaluate the responsiveness of Feline Musculoskeletal Pain Index (FMPI) items. The number of cats in each study, the analgesic and route of administration and the treatment period between assessments are shown

Study	Total number of cats (analgesic; placebo)	Analgesic	Dosage, route	Treatment period
Clinical metrology instrument (CMI)[14]	25 (10; 15)	Meloxicam	0.1 mg/kg, then 0.05 mg/kg q24h PO	2 weeks
Low-dose[18]	58 (29; 29)	Meloxicam	0.035 mg/kg q24h PO	3 weeks
Antibody[19]	34 (23; 11)	Anti-NGF mAb	0.4 or 0.8 mg/kg, single dose SC	3 weeks

NGF = nerve growth factor; mAb = monoclonal antibody

Table 4 Example of score conversion of Feline Musculoskeletal Pain Index (FMPI) from the original study to the present study

Original FMPI score	Converted FMPI score	Description in FMPI
5	0	Normal
4	1	Not quite normal
3	2	Somewhat worse than normal
2	3	Barely, or with great effort
1	4	Not at all

FMPI data collected at the initial visit. Interpretation of the data was based on the following criteria: unacceptable ($0.5 > \alpha$); poor ($0.6 > \alpha \geq 0.5$); questionable ($0.7 > \alpha \geq 0.6$); acceptable ($0.8 > \alpha \geq 0.7$); good ($0.9 > \alpha \geq 0.8$); and excellent ($\alpha \geq 0.9$).[21] A score above 0.70 was taken as an indication of sufficient homogeneity of the items in the FMPI.[22] Test-retest reliability for all FMPI items was assessed using intraclass correlation coefficient (ICC) and repeatability of each FMPI item was assessed using Bland–Altman analysis. These were calculated between the initial and baseline visits. Strength of agreement was based on: poor ($0.0 < \text{ICC} < 0.2$); fair ($0.2 \leq \text{ICC} < 0.4$); moderate ($0.4 \leq \text{ICC} < 0.6$); strong ($0.6 \leq \text{ICC} < 0.8$); and near complete ($0.8 \leq \text{ICC}$).[23] A paired *t*-test was performed to assess the difference in each FMPI score between the initial and baseline visits. For the assessment of responsiveness of the FMPI, unpaired *t*-tests or Mann–Whitney U tests were used to evaluate the difference from baseline in score between groups (placebo vs analgesic). Paired *t*-tests or Wilcoxon rank-sum tests were used to evaluate the treatment effect within subject (baseline vs revisit after treatment). Mann–Whitney U tests were used for the evaluation of the discriminatory ability of FMPI; $P < 0.05$ was considered significant. The Shapiro–Wilk test was used to evaluate the normality of FMPI scores. Parametric tests were used when data were normally distributed, and non-parametric tests were used when data were not normally distributed. No adjustment for multiple comparisons was made in order to maximize the chances of selecting items that appeared responsive. The decision to not correct for multiple comparisons when examining the responsiveness of items was made a priori, as in this case we accepted an increased type 1 error rate in favor of reducing the type 2 error rate. That is, we sought to avoid spurious elimination of questions and so allowed for greater spurious inclusion. This decision was made because we foresaw much greater reduction in utility in the exclusion of potentially relevant questions than in the inclusion of potentially useless ones.

Based on the results of the reliability and responsiveness analyses, the FMPI items were eliminated in a stepwise manner. Selection of items for removal was as follows: first, an item was removed if responsiveness data showed that the placebo effect was larger than the active drug effect. If multiple items had higher placebo effects than active drug effects, we started by removing the item with a higher *P* value. Next, an item was removed if prevalence of impairment (of that item) in the population of cats with DJD was <50% based on our prevalence data.[20] After each item was removed, statistical analysis was performed on the remaining items to calculate the group difference across the set of remaining items. After the stepwise elimination, a proposed new FMPI-sf was constructed. The internal consistency, test-retest reliability and Bland–Altman analysis were used to reassess reliability of FMPI-sf using the combined data sets of three studies. The individual and combined data sets of three studies were then used to compare the responsiveness of the original FMPI and the new FMPI-sf.

Results

Animals

Demographic data are reported in the relevant publications and available from the authors on request. However, there was no difference in age, sex, body weight and scores at the initial visit between placebo and analgesic groups. In this report, we simply report on the number of cats in each evaluation.

Internal consistency of the FMPI

Across studies, 180 cats were included for the evaluation of internal consistency. Table 5 shows the α of each FMPI

Table 5 Internal consistency of each Feline Musculoskeletal Pain Index (FMPI) item (n = 180) using Cronbach's α. Numbers of answers for each item reflect differences in the home environments for some cats (ie, absence of stairs or other pets in a home)

Item	α	Number of cats
Walk/move	0.896	180
Run	0.893	179
Jump up	0.897	180
Kitchen counter	0.902	173
Jump down	0.894	180
Stairs up	0.891	133
Stairs down	0.894	133
Toys	0.899	176
Other pets	0.896	161
Get up	0.893	180
Sit down	0.896	180
Stretch	0.894	170
Groom	0.899	180
Interaction	0.899	180
Touch	0.903	180
Eat	0.905	180
Litter box	0.902	174
Overall	0.903	133

item and the overall FMPI. As some homes did not have stairs or other pets, the response number varies and the number of responses for each item is shown. Overall, the original FMPI had an α of 0.903, ranging from 0.891 to 0.905 for each item. Four items had an α of more than 0.90 and the others had an α of just below 0.90.

Test-retest reliability and repeatability of the FMPI

Across the studies, 117 cats were included for the evaluation of internal consistency. The ICC for all 17 items was 0.814 ($P<0.001$) with a 95% confidence interval of 0.757 to 0.858. Table 6 shows the Bland–Altman analysis of each FMPI item. Based on Bland–Altman results, the score at the baseline visit was significantly higher than that at the initial visit in two items (walk/move and stretch); however, the difference was relatively small with an average of 0.09 and 0.16 points for walk/move and stretch, respectively (the plots are available in Figure 1 in the supplementary material).

Responsiveness of the FMPI

In total, there were 55 and 62 cats in the placebo group and analgesic group, respectively. There were no significant differences in scores for each item and across 17 items between groups at baseline visit (Table 7).

When compared with the baseline value, a significant treatment effect was seen across many items in both the placebo and analgesic groups (Table 8). Two types of items were of interest from this analysis: (1) items where a significant difference was found between the treatment groups; and (2) items where a significant effect was found within the analgesic group and not the placebo group.

Table 6 Bland–Altman analysis of initial to baseline agreement for each Feline Musculoskeletal Pain Index (FMPI) item (n = 117 cats)

Item	Bland–Altman (bias: 95% CI)	*t*-test (*P* value)	Number of answers
Walk/move	0.09 (0.005–0.18)	0.04* (baseline >initial)	117
Run	−0.03 (−0.14 to 0.08)	0.62	115
Jump up	0.01 (−0.1 to 0.12)	0.84	117
Kitchen counter	−0.04 (−0.13 to 0.04)	0.33	110
Jump down	−0.006 (−0.13 to 0.12)	0.93	116
Stairs up	0.09 (−0.02 to 0.20)	0.13	92
Stairs down	0.04 (−0.08 to 0.16)	0.52	81
Toys	0.05 (−0.09 to 0.20)	0.46	113
Other pets	0.16 (−0.0009 to 0.32)	0.051	104
Get up	−0.01 (−0.12 to 0.10)	0.83	116
Sit down	0.09 (−0.02 to 0.20)	0.11	116
Stretch	0.16 (0.02 to 0.29)	0.02* (baseline >initial)	106
Groom	0.06 (−0.04 to 0.16)	0.21	117
Interaction	−0.03 (−0.13 to 0.07)	0.50	117
Touch	−0.05 (−0.17 to 0.07)	0.39	117
Eat	−0.03 (−0.10 to 0.04)	0.33	117
Litter box	−0.04 (−0.12 to 0.05)	0.35	112
Average of 17 items	0.03 (−0.03 to 0.09)	0.37	117

Initial and baseline refer to the first and second (respectively) times the FMPI was administered. The difference between the initial and baseline visits were assessed using a paired *t*-test: a significant difference (*$P<0.05$) was found for two items (walk/move and stretch; for both items the baseline scores were higher (more impaired) than the initial visit scores
CI = confidence interval

Table 7 Comparison of the mean score in each Feline Musculoskeletal Pain Index (FMPI) item between groups at baseline

Item	Placebo (average ± SD)	Analgesic (average ± SD)	*P* value
Walk/move	1.6 ± 0.8	1.5 ± 0.9	0.45
Run	2.2 ± 1.1	1.9 ± 1.2	0.21
Jump up	2.4 ± 0.9	2.4 ± 1.0	0.90
Kitchen counter	3.5 ± 0.9	3.4 ± 1.1	0.53
Jump down	2.0 ± 0.9	2.3 ± 0.9	0.07
Stairs up	1.9 ± 1.0	1.7 ± 0.9	0.51
Stairs down	1.7 ± 1.0	1.8 ± 0.9	0.72
Toys	2.2 ± 1.2	2.2 ± 1.3	0.97
Other pets	2.0 ± 1.4	1.9 ± 1.4	0.59
Get up	1.4 ± 0.9	1.5 ± 1.0	0.41
Sit down	1.1 ± 1.0	1.2 ± 1.0	0.63
Stretch	1.4 ± 1.2	1.7 ± 1.3	0.35
Groom	1.2 ± 1.2	1.3 ± 1.2	0.80
Interaction	1.0 ± 1.0	0.7 ± 0.9	0.17
Touch	1.0 ± 1.0	1.0 ± 1.0	0.66
Eat	0.1 ± 0.4	0.2 ± 0.5	0.64
Litter box	0.7 ± 0.8	0.8 ± 1.0	0.58
Average of 17 items	1.6 ± 0.64	1.6 ± 0.69	0.97

Higher value = more impairment. Mann–Whitney U tests were used for each item

Between groups there were significant differences for 'jump down' and 'get up', favoring the analgesic group, while both 'stretch' and 'grooming' showed a significant change in the analgesic group and not the placebo group.

Elimination of FMPI items

No significant difference was seen between the groups across the scores for 17 items.

The change in the placebo group was larger than the change in the analgesic group in six items, and these were removed one by one (Table 9). Additionally, two items were removed due to a relatively small change from baseline or a low prevalence of impairment in our population of cats. After stepwise elimination, nine FMPI items were retained. The nine items that ultimately remained were: jumping up and down; jumping up to kitchen counter; playing with toys; interaction with other pets; getting up from a resting position; sitting down; stretching; and grooming.

Comparing the 17-item FMPI and nine-item FMPI-sf

The mean FMPI score across the 17-item FMPI for the DJD cats was 1.51 ± 0.65 and for the normal cats (n = 16) was 0.004 ± 0.015. The mean FMPI score across the nine-item FMPI-sf for the DJD cats was 1.78 ± 0.76 and for the normal cats was zero. There was a significant difference between the DJD cats and the normal cats in

Table 8 Comparison of the mean change (± SD) in score from baseline for each Feline Musculoskeletal Pain Index (FMPI) item within and between groups

Item	Placebo (average ± SD)	Placebo (*P* value: within group)	Analgesic (average ± SD)	Analgesic (*P* value: within group)	*P* value (between groups)
Walk/move	−0.53 ± 0.72	0.0003*	−0.52 ± 0.78	0.0002*	0.94
Run	−0.70 ± 1.0	0.0005*	−0.50 ± 1.0	0.008*	0.30
Jump up	−0.52 ± 0.79	0.0009*	−0.85 ± 1.0	<0.0001*	0.11
Kitchen counter	−0.25 ± 0.62	0.34	−0.48 ± 1.2	0.21	0.21
Jump down	−0.33 ± 1.0	0.045*	−0.95 ± 0.85	<0.0001*	0.0004†
Stairs up	−0.76 ± 0.72	0.0009*	−0.67 ± 0.85	0.0039*	0.61
Stairs down	−0.59 ± 0.80	0.004*	−0.60 ± 0.78	0.0061*	0.98
Toys	−0.60 ± 1.0	0.002*	−0.84 ± 1.1	0.0005*	0.23
Other pets	−0.71 ± 1.2	0.007*	−0.95 ± 1.2	<0.0001*	0.32
Get up	−0.39 ± 0.76	0.003*	−0.74 ± 0.85	<0.0001*	0.02†
Sit down	−0.31 ± 0.81	0.04*	−0.55 ± 0.88	0.0005*	0.13
Stretch	−0.35 ± 0.79	0.26	−0.64 ± 1.1	0.0205*	0.11
Groom	−0.25 ± 0.62	0.26	−0.55 ± 0.95	0.0186*	0.053
Interaction	−0.45 ± 0.74	0.02*	−0.35 ± 0.73	0.0027*	0.46
Touch	−0.42 ± 0.83	0.02*	−0.42 ± 0.82	0.0016*	0.99
Eat	−0.05 ± 0.45	0.62	−0.08 ± 0.46	0.1659	0.75
Litter box	−0.30 ± 0.74	0.005*	−0.44 ± 0.79	0.0012*	0.32
Average of 17 items	−0.43 ± 0.42	<0.0001*	−0.59 ± 0.49	<0.0001*	0.057

Mann–Whitney U tests were used for each item when comparing within group. Wilcoxon rank-sum tests were used for each item when comparing between groups, except for the average of 17 items
Negative values = improvement
*Significant difference from baseline ($P<0.05$)
†Significant difference between groups ($P<0.05$)

Table 9 Elimination process of Feline Musculoskeletal Pain Index (FMPI) items, and the comparison of the mean change from baseline across the remaining items and the group difference

Number of items retained	Items removed	Placebo (average ± SD)	Analgesic (average ± SD)	*P* value
17	NA	−0.43 ± 0.42	−0.59 ± 0.49	0.057
16	Run	−0.41 ± 0.41	−0.60 ± 0.48	0.027*
15	Interaction	−0.41 ± 0.42	−0.62 ± 0.49	0.016*
14	Stairs up	−0.39 ± 0.42	−0.61 ± 0.48	0.011*
13	Walk/move	−0.38 ± 0.43	−0.62 ± 0.48	0.0057*
12	Touch	−0.38 ± 0.44	−0.64 ± 0.50	0.0041*
11	Stairs down	−0.37 ± 0.44	−0.64 ± 0.50	0.0026*
10	Eat	−0.40 ± 0.46	−0.70 ± 0.54	0.0022*
9	Litter box	−0.41 ± 0.48	−0.72 ± 0.58	0.0022*

Wilcoxon rank-sum tests were used for each item. Negative value = improvement; NA = not applicable
*Significant difference ($P < 0.05$)

Table 10 Comparison of the mean change from baseline in Feline Musculoskeletal Pain Index (FMPI) score between groups, when using the 17-item FMPI and nine-item FMPI

Study	Number of items	Placebo (average ± SD)	Analgesic (average ± SD)	*P* value
Clinical metrology instrument (CMI)[14]	17	−0.46 ± 0.50	−0.52 ± 0.72	0.79
	9	−0.52 ± 0.51	−0.69 ± 0.85	0.53
Low-dose[18]	17	−0.39 ± 0.39	−0.43 ± 0.42	0.69
	9	−0.33 ± 0.45	−0.56 ± 0.52	0.079
Antibody[19]	17	−0.49 ± 0.42	−0.82 ± 0.39	0.027*
	9	−0.49 ± 0.49	−0.94 ± 0.46	0.013*

Wilcoxon rank-sum tests were used for the CMI and antibody studies, and paired *t*-tests were used for the low-dose study
*Significant difference ($P < 0.05$)

Table 11 Internal consistency of each Feline Musculoskeletal Pain Index-short form (FMPI-sf) item (n = 117) using Cronbach's α

Item	Cronbach's α
Jump up	0.823
Kitchen counter	0.839
Jump down	0.831
Toys	0.834
Other pets	0.829
Get up	0.814
Sit down	0.824
Stretch	0.817
Groom	0.833
Overall	0.844

both the 17-item FMPI ($P < 0.001$) and nine-item FMPI-sf ($P < 0.001$).

When the original 17-item FMPI and the nine-item FMPI-sf (see the 'Feline Musculoskeletal Pain Index – short form' in the supplementary material) were compared in three studies, the difference in the mean change from baseline score between groups was always greater when the FMPI-sf was used (Table 10). In the antibody study (anti-NGF mAb), the significance was seen for both the original FMPI and the FMPI-sf; however, the difference between groups was greater with the FMPI-sf. In contrast, in the two studies using an NSAID, while the difference between groups was greater using the FMPI-sf, significance was not reached in either study. After the reduction of items, the ICC was 0.837 ($P < 0.001$) with a 95% confidence interval of 0.781 to 0.879 and no significant difference was observed in the mean of FMPI-sf scores between initial and baseline visits. Additionally, the internal consistency of all items in FMPI-sf was more than 0.8, which was higher than the minimum standard of reliability of 0.7 (Table 11).

Discussion

In this study, we used a data-driven approach to assess the reliability and responsiveness of the original FMPI, and subsequently refined the instrument to improve the responsiveness. The original 17-item FMPI seemed to have a reasonable reliability, but low/no responsiveness. The FMPI-sf was constructed by reducing the number of items from 17 to nine, which increased the responsiveness of the instrument while maintaining its reliability. Thus, the proposed FMPI-sf may be able to better distinguish

between placebo and analgesic effects in cats with DJD. However, a future prospective study is needed to further assess responsiveness.

Consistent with previous work, our results showed that the original 17-item FMPI seems to have an excellent internal consistency with an overall α of more than 0.90.[16,24] Generally, better internal consistency is reflected by a higher α. However, in empirical studies, a maximum α value of 0.90 has been recommended as optimal.[22] If items have an α of more than 0.90, it suggests there may be redundancies in the questionnaire, and the number of items could be reduced.[22] The α values for each item in the original FMPI were all near or above 0.90, while those of the FMPI-sf are below 0.90.

Overall, the original 17-item FMPI seems to have a reasonable reliability. In this study, the ICC between the initial and baseline visits across the 17 FMPI items was in near-complete agreement (0.814), and this result was similar to the previous study.[14] Based on paired *t*-tests, there were significant differences in two items between the scores at the initial and baseline visits (ability to 'walk/move' and 'stretch'). In both cases, scores for the baseline visits were higher (indicating more impairment); this may be explained by increased owner attention to the performance of these behaviors after filling out the FMPI at the initial visit. These results were corroborated by the Bland–Altman bias scores, where the bias was close to zero across each of the individual items and the total score, and the limits of agreement encompassed the majority of the points. The least agreement was seen with scores between one and two; ratings of cats at the extremes of the scale were in close agreement. However, overall differences were relatively small, suggesting that the FMPI results from the initial visit could serve as the baseline value.

Previous studies have investigated which behaviors are responsive to analgesic treatment in cats with DJD and associated pain.[25,26] Although no placebo group was included for comparison, these studies showed that owner-rated positive changes were seen in some behaviors including jumping, gait stiffness, grooming habits and temperament. Indeed, in the present study, a significant treatment effect was observed in the majority of items in the 17-item FMPI in cats receiving analgesics. However, similar changes were also detected in cats receiving placebo. As a consequence, there was no significant difference in FMPI scores across 17 items between these groups. The lack of responsiveness may be explained by the caregiver placebo effects, known to be high in studies of client-owned cats with DJD.[17] This effect appears to be higher than in similar studies with dogs. Our results, showing improvement in the placebo group for most FMPI items, highlights the difficulty in developing outcome assessments for use with cats. The FMPI-sf should be tested in new clinical trials using known analgesics.

After the stepwise elimination of FMPI items, nine items were retained to construct the FMPI-sf. As shown, the FMPI-sf had an acceptable reliability, similar to the original FMPI. Therefore, the reliability of the instrument did not suffer in any way from the reduction of items and the procedures described above improved responsiveness of the instrument while maintaining the reliability.

The original FMPI was clearly able to discriminate between musculoskeletally normal cats and mildly to severely impaired DJD cats.[24] The FMPI-sf also showed good discriminatory ability when applied to our data. However, this should be confirmed in a study using a new cohort of affected animals and healthy controls. Furthermore, placebo-controlled studies using NSAIDs failed to demonstrate responsiveness of the original FMPI in cats with DJD-associated pain. When the original FMPI and the FMPI-sf were compared for the individual studies (CMI, low-dose, antibody), the difference in the mean change from baseline score between groups of cats with DJD pain that were treated with placebo or active treatment was always greater when the FMPI-sf was used. This difference (treatment effect) was significant when cats were given an anti-NGF mAb, but not when given an NSAID. We do not know if this difference in detected efficacy is due to a true difference between treatments or to features of the study design, or a combination of both. There was a difference in patient populations between the CMI study and the other studies; the CMI study included mildly to severely impaired cats, and more than half of the population were mildly impaired. In contrast, the other studies enrolled moderate to severely impaired cats (resulting in recruitment of more highly impaired cats than the CMI study). Therefore, the improvement after treatment might be less easily detected in the CMI study. Overall, the responsiveness appears to be improved when the FMPI-sf is used.

Recently, our research group proposed a clinical checklist that could be used to identify cats likely to have DJD-associated pain.[20] The checklist was derived from the FMPI and 17 items were reduced to six items based on the sensitivity and specificity of each item. All six items involved movement and included: running; jumping up; jumping down; stairs up; stairs down; and playing with toys. On the other hand, the FMPI-sf contains items that involve not only movement but also the owner's rating of their cat's ability to perform the activities of daily living and social relationships. While the checklist serves as an effective screening tool, this highlights that chronic pain impacts multiple dimensions in cats, and behavioral changes in those dimensions need to be assessed in outcome measures.[27]

As an important limitation, this study is a retrospective design and includes data from studies that were performed in only a single site, which could result in a case selection bias. Prospective studies using the FMPI-sf will evaluate the performance of the instrument in other populations of cats. Further, future work will investigate the FMPI-sf's relationship with objective outcome measures including activity monitoring and gait analysis.

Conclusions

The present retrospective study showed that the original FMPI seemed to have a reasonable reliability, but low/no responsiveness. The stepwise elimination of items was performed to refine the responsiveness of FMPI, and the FMPI-sf was proposed. The proposed FMP-sf seemed to have better responsiveness while maintaining its reliability. Although a future prospective study is needed, the proposed FMPI-sf may be able to better detect analgesic effects in cats with DJD.

Acknowledgements The authors acknowledge all the cat owners and referring veterinarians who participated in these studies. They especially thank Andrea Thomson for her assistance in conducting these studies.

Supplementary material The following files are available online:

Supplementary Figure 1: Bland–Altman plots for FMPI items (showing the mean difference [±1.96 SD] between the initial and baseline visits).

Feline Musculoskeletal Pain Index – short form with instructions.

Conflict of interest The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: MEG and BDXL are both paid consultants for Zoetis. This study (analysis of previously collected data) was funded by Zoetis. MEG has current funding from Zoetis for other (unrelated) research. BDXL has a $0 collaborative research agreement with Zoetis for other (unrelated) work. Zoetis did not participate in developing the concept nor design of the study, nor in the analysis of the data nor writing of the manuscript.

Funding The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this study was supported by Zoetis in the form of salary support for ME.

Ethical approval The work described in this manuscript involved the use of non-experimental (owned or unowned) animals. Established internationally recognized high standards ('best practice') of veterinary clinical care for the individual patient were always followed and/or this work involved the use of cadavers. Ethical approval from a committee was therefore not specifically required for publication in *JFMS*. Although not required, where ethical approval was still obtained, it is stated in the manuscript.

Informed consent Informed consent (verbal or written) was obtained from the owner or legal custodian of all animal(s) described in this work (experimental or non-experimental animals, including cadavers) for all procedure(s) undertaken (prospective or retrospective studies). No animals or people are identifiable within this publication, and therefore additional informed consent for publication was not required.

ORCID iD Masataka Enomoto https://orcid.org/0000-0003-2516-635X
B Duncan X Lascelles https://orcid.org/0000-0002-2950-9009
Margaret E Gruen https://orcid.org/0000-0002-6036-8849

References

1 Monteiro BP and Steagall PV. **Chronic pain in cats: recent advances in clinical assessment.** *J Feline Med Surg* 2019; 21: 601–614.
2 Lascelles BD, Henry JB, Brown J, et al. **Cross-sectional study of the prevalence of radiographic degenerative joint disease in domesticated cats.** *Vet Surg* 2010; 39: 535–544.
3 Hardie EM, Roe SC and Martin FR. **Radiographic evidence of degenerative joint disease in geriatric cats: 100 cases (1994–1997).** *J Am Vet Med Assoc* 2002; 220: 628–632.
4 Slingerland LI, Hazewinkel HA, Meij BP, et al. **Cross-sectional study of the prevalence and clinical features of osteoarthritis in 100 cats.** *Vet J* 2011; 187: 304–309.
5 Clarke SP, Mellor D, Clements DN, et al. **Prevalence of radiographic signs of degenerative joint disease in a hospital population of cats.** *Vet Rec* 2005; 157: 793–799.
6 Lascelles BD. **Feline degenerative joint disease.** *Vet Surg* 2010; 39: 2–13.
7 Taylor PM and Robertson SA. **Pain management in cats – past, present and future. Part 1. The cat is unique.** *J Feline Med Surg* 2004; 6: 313–320.
8 Gruen ME, Alfaro-Cordoba M, Thomson AE, et al. **The use of functional data analysis to evaluate activity in a spontaneous model of degenerative joint disease associated pain in cats.** *PloS One* 2017; 12. DOI: 10.1371/journal.pone.0169576.
9 Lascelles BD, Hansen BD, Thomson A, et al. **Evaluation of a digitally integrated accelerometer-based activity monitor for the measurement of activity in cats.** *Vet Anaesth Analg* 2008; 35: 173–183.
10 Lascelles BD, Findley K, Correa M, et al. **Kinetic evaluation of normal walking and jumping in cats, using a pressure-sensitive walkway.** *Vet Rec* 2007; 160: 512–516.
11 Corbee RJ, Maas H, Doornenbal A, et al. **Forelimb and hindlimb ground reaction forces of walking cats: assessment and comparison with walking dogs.** *Vet J* 2014; 202: 116–127.
12 Klinck MP, Monteiro BP, Lussier B, et al. **Refinement of the Montreal Instrument for Cat Arthritis Testing, for Use by Veterinarians: detection of naturally occurring osteoarthritis in laboratory cats.** *J Feline Med Surg* 2018; 20: 728–740.
13 Klinck MP, Rialland P, Guillot M, et al. **Preliminary validation and reliability testing of the Montreal Instrument for Cat Arthritis Testing, for Use by Veterinarians, in a colony of laboratory cats.** *Animals (Basel)* 2015; 5: 1252–1267.
14 Benito J, Hansen B, Depuy V, et al. **Feline musculoskeletal pain index: responsiveness and testing of criterion validity.** *J Vet Intern Med* 2013; 27: 474–482.
15 Lascelles BDX, Hansen BD, Roe S, et al. **Evaluation of client-specific outcome measures and activity monitoring to measure pain relief in cats with osteoarthritis.** *J Vet Intern Med* 2007; 21: 410–416.

16 Stadig S, Lascelles BDX, Nyman G, et al. **Evaluation and comparison of pain questionnaires for clinical screening of osteoarthritis in cats.** *Vet Rec* 2019; 185: 757. DOI: 10.1136/vr.105115.
17 Gruen ME, Dorman DC and Lascelles BDX. **Caregiver placebo effect in analgesic clinical trials for cats with naturally occurring degenerative joint disease-associated pain.** *Vet Rec* 2017; 180: 473. DOI: 10.1136/vr.104168.
18 Gruen ME, Griffith EH, Thomson AE, et al. **Criterion validation testing of clinical metrology instruments for measuring degenerative joint disease associated mobility impairment in cats.** *PloS One* 2015; 10: e0131839. DOI: 10.1371/journal.pone.0131839.
19 Gruen ME, Thomson AE, Griffith EH, et al. **A feline-specific anti-nerve growth factor antibody improves mobility in cats with degenerative joint disease-associated pain: a pilot proof of concept study.** *J Vet Intern Med* 2016; 30: 1138–1148.
20 Enomoto M, Lascelles BDX and Gruen ME. **Development of a checklist for the detection of degenerative joint disease-associated pain in cats.** *J Feline Med Surg* 2020; 22: 1137–1147.
21 Streiner DL. **Starting at the beginning: an introduction to coefficient alpha and internal consistency.** *J Pers Assess* 2003; 80: 99–103.
22 Tavakol M and Dennick R. **Making sense of Cronbach's alpha.** *Int J Med Educ* 2011; 2: 53–55.
23 Fledelius J, Khalil A, Hjorthaug K, et al. **Inter-observer agreement improves with PERCIST 1.0 as opposed to qualitative evaluation in non-small cell lung cancer patients evaluated with F-18-FDG PET/CT early in the course of chemo-radiotherapy.** *EJNMMI Res* 2016; 6. DOI: 10.1186/s13550-016-0223-6.
24 Benito J, Depuy V, Hardie E, et al. **Reliability and discriminatory testing of a client-based metrology instrument, feline musculoskeletal pain index (FMPI) for the evaluation of degenerative joint disease-associated pain in cats.** *Vet J* 2013; 196: 368–373.
25 Clarke SP and Bennett D. **Feline osteoarthritis: a prospective study of 28 cases.** *J Small Anim Pract* 2006; 47: 439–445.
26 Bennett D and Morton C. **A study of owner observed behavioural and lifestyle changes in cats with musculoskeletal disease before and after analgesic therapy.** *J Feline Med Surg* 2009; 11: 997–1004.
27 Lascelles BDX, Brown DC, Conzemius MG, et al. **Measurement of chronic pain in companion animals: discussions from the Pain in Animals Workshop (PAW) 2017.** *Vet J* 2019; 250: 71–78.