Evaluation of a welfare assessment tool to examine practices for preventing, recognizing, and managing pain at companion-animal veterinary clinics

Lauren C. Dawson, Cate E. Dewey, Elizabeth A. Stone, Cornelia I. Mosley, Michele T. Guerin, Lee Niel

Abstract

Successful prevention, recognition, and treatment of pain are integral to ensuring veterinary patient welfare. A canine and feline welfare assessment tool, incorporating verbal interviews with veterinarians using open-ended questions, was developed to assess pain management practices that safeguard and improve patient welfare. The tool was evaluated in 30 companion- and mixed-animal veterinary clinics in Ontario in order to assess its reliability, feasibility, and validity, while also benchmarking current practices. Responses were analyzed according to a scoring scheme developed based on published literature and expert opinion. Based on weighted *kappa* statistics, interview scoring had substantial inter-observer (K_w = 0.83, 0.73) and near-perfect intra-observer (K_w = 0.92) agreement, which suggests that the tool reliably collects information about pain management practices. Interviews were completed at all recruited clinics, which indicates high feasibility for the methods. Validity could not be assessed, as participants were reluctant to share information about analgesic administration from their clinical records. Descriptive results indicated areas for which many veterinarians are acting in accordance with best practices for pain management, such as pre-emptive and post-surgical analgesia for ovariohysterectomy patients, and post-surgical care instructions. Areas that offer opportunity for enhancement were also highlighted, e.g., training veterinary staff to recognize signs of pain and duration of analgesia in ovariohysterectomy patients after discharge. Overall, based on this limited sample, most veterinarians appear to be effectively managing their patients' pain, although areas with opportunity for enhancement were also identified. Further research is needed to assess trends in a broader sample of participants.

Résumé

Être en mesure de prévenir, reconnaitre, et traiter la douleur avec succès est essentiel pour assurer le bien-être des patients vétérinaires. Un outil d'évaluation du bien-être des chiens et des chats, incorporant une entrevue orale avec des vétérinaires avec des questions ouvertes, a été développé pour évaluer les pratiques de gestion de la douleur qui sauvegarde et améliore le bien-être des patients. L'outil a été évalué dans 30 cliniques vétérinaires pour animaux de compagnie et cliniques mixtes en Ontario afin de vérifier la fiabilité, la faisabilité, et la validité, tout en réalisant un étalonnage des pratiques actuelles. Les réponses ont été analysées selon un schéma de pointage basé sur la littérature publiée et l'opinion d'expert. Sur la base des statistiques kappa *pondérées, les pointages des entrevues avaient un accord inter-observateur marqué (K_w = 0,83, 0,73) et un accord intra-observateur presque parfait (K_w = 0,92), ce qui suggère que l'outil a permis d'obtenir des informations fiables sur les pratiques de gestion de la douleur. Les entrevues ont été complétées dans toutes les cliniques recrutées, ce qui indiquait une excellente faisabilité pour les méthodes utilisées. La validité n'a pu être vérifiée car les participants étaient réfractaires à partager de l'information sur l'administration d'analgésique à partir de leurs dossiers médicaux. Les résultats indiquent que plusieurs vétérinaires agissent en concordance avec les bonnes pratiques de gestion de la douleur pour l'analgésie préventive et post-chirurgicale des patients subissant une ovariohystérectomie et les instructions pour les soins post-chirurgie. D'autres domaines ont été identifiés comme nécessitant des améliorations, e.g. former le personnel de la clinique à reconnaitre les signes de douleurs et la durée de l'analgésie chez les patients ayant eu une ovariohystérectomie après leur congé. De manière générale, sur la base de cet échantillonnage limité, la plupart des vétérinaires semble gérer la douleur de leurs patients de manière efficace, bien que des améliorations à faire aient été identifiées. De la recherche supplémentaire est requise pour évaluer les tendances dans un échantillonnage plus grand de participants.*

(Traduit par Docteur Serge Messier)

Department of Population Medicine (Dawson, Dewey, Guerin, Niel) and Department of Clinical Studies (Stone, Mosley), Ontario Veterinary College, University of Guelph, Guelph, Ontario N1G 2W1.

Address all correspondence to Dr. Lee Niel; telephone: (519) 824-4120, ext. 53030; fax: (519) 763-8621; e-mail: niell@uoguelph.ca

Dr. Mosley's current address is 404 Veterinary Emergency and Referral Hospital, Newmarket, Ontario L3Y 0B3.

Received October 27, 2016. Accepted April 3, 2017.

Introduction

Many aspects of veterinary care have the potential to affect patient welfare, both within the veterinary clinic and in the animal's home. While on-farm animal welfare assessment schemes, e.g., Welfare Quality, have been developed for most agricultural species, similar assessment tools are not available for companion animals in the veterinary clinic setting. To begin the process of developing assessment tools, a multi-stage survey was administered to companion-animal veterinarians and animal welfare experts to identify factors related to veterinary care that affect canine and feline welfare (1). The optimization of analgesic regimens associated with surgical procedures, as well as post-surgical and chronic pain recognition and control, were among the factors identified as having the highest impact on patient welfare (1).

A welfare assessment tool was developed for companion-animal veterinary hospitals, with the long-term goal of encouraging veterinary clinics to adopt practices that optimize patient welfare. It was based on expert opinion (1), current scientific literature, and veterinary pain management guidelines (2). Given its impact on companion animal welfare, exploring veterinary approaches to managing canine and feline pain is a major subsection of this tool.

Pain has the potential to directly impair animal welfare (3) and can also have detrimental effects on animal health. For example, pain can impair post-surgical recovery, resulting in extended veterinary hospital stays due to inappetence, self-injury, and immune suppression, which increase risk of infections (4,5).

It is challenging to identify the number of cats and dogs that are affected by pain. With approximately 7.9 million cats and 5.9 million dogs owned in Canada alone (6), however, a significant number of animals will undergo procedures, such as ovariohysterectomy and castration, which can cause post-surgical pain if not properly managed. In addition, chronic pain often goes undiagnosed, which makes its prevalence difficult to determine, however, estimates have suggested that 14% of cats and 20% of dogs are affected by chronic pain (7).

Effective pain management requires accurate identification and appropriate treatment of affected animals. Identification depends on an observer's ability to recognize and evaluate non-verbal signs of pain. This can be complicated by the fact that many pain-related behaviors, such as withdrawal, are also observed in response to fear (5) and fear is common in veterinary contexts (8,9). It is thus not surprising that veterinarians consider the recognition of pain as one of the most significant barriers to effective pain management, especially for the control of chronic pain (10). Moreover, veterinarians' knowledge and attitudes can also play an important role in diagnosing and treating pain. For example, gender and year of graduation from veterinary school influence veterinarians' attitudes toward pain and use of analgesics (10–15). An effective welfare assessment tool as it relates to pain management must therefore assess a veterinarian's knowledge about pain assessment and current veterinary practices related to treatment.

An ideal welfare assessment tool uses measures that are reliable, feasible, and valid. It should provide data that lead to the same conclusion within and across observers (reliability), allow for ease of measurement without excessive effort, time, or cost (feasibility), and accurately measure what it is attempting to measure (validity) (16). The primary objective of this study was to assess the reliability, feasibility, and validity of the trial tool to assess animal welfare, specifically with regard to pain recognition and management, in companion-animal veterinary hospitals. The secondary objective was to collect information about current pain management practices in the selected veterinary clinics to determine which aspects of the tool best differentiate between welfare practices across clinics.

Materials and methods

This study was approved by the Research Ethics Board (REB #13JN017) and the Animal Care Committee (AUP #2272) at the University of Guelph. Further details about methodology can be found in the thesis from which this work is derived (17). Veterinary clinics were selected from a publicly available, online listing through the College of Veterinarians of Ontario. A sample size of 30 veterinary clinics was necessary to have enough power to assess reliability (18). In order to recruit 30 veterinary hospitals, all hospitals that fit the selection criteria were invited to participate in the survey (N = 474). The owner of every companion- and mixed-animal veterinary clinic located within 100 km of the University of Guelph received an invitation by mail, with specialty, emergency, teaching, and humane society hospitals excluded. Invitations were also sent to 6 veterinary clinics outside the 100-km radius because staff members had previously participated in research and had expressed interest in participating in future studies. To encourage participation, a follow-up phone call was made within 14 d and up to 2 follow-up phone calls and 1 follow-up e-mail were sent to each clinic, until 30 veterinary clinics had been recruited and scheduled visits.

As part of a larger study, each participating veterinary clinic was visited in-person for a minimum of 1 full day. During each veterinary visit, a welfare assessment tool was evaluated that consisted of appointment observations, a verbal interview, a written questionnaire, and a tour of the facility. Pain assessment and management were evaluated using verbal interviews only. These interviews used open-ended questions and followed a script (Table I) and were conducted with either the clinic owner or another senior veterinarian. When the clinic owner was unavailable, he or she suggested a suitable alternate veterinarian from the clinic staff, based on seniority to ensure familiarity with the clinic's practices. Questions were developed based on results obtained from a previous study in which experts identified veterinary care-related factors that affect patient welfare (1).

One section focused on aspects of pain management, including the recognition of patient pain, staff and owner training, and perioperative analgesic regimens. Veterinarians were further asked about perioperative analgesic regimens for ovariohysterectomy, for which they were also asked to retrieve veterinary records. Interviews covered a number of other topics that are discussed in detail in other manuscripts, i.e., veterinary-client communication about aspects of animal care and welfare, behavioral health, and clinic management.

Before data collection, interview questions were pilot-tested with a veterinarian for comprehension and no modifications were required. Although most interviews were conducted in person, 3 veterinarians were interviewed by telephone at a later date as

Table I. Pain assessment and management interview script.

Category	Questions
Surgical pain	1. Post-operatively, how do you recognize pain in cats?
	2. Post-operatively, how do you recognize pain in dogs?
	3. How do you train staff (including volunteers) to recognize signs of pain?
	4. When planning for and about to perform surgical procedures:
	a) How do you develop a pain management plan for your patients?
	b) How do you decide about providing analgesics for surgical patients?
	c) Please describe the common points when you routinely provide analgesics.
	5. At the time of post-surgical discharge:
	a) How do you communicate post-surgical instructions to owners?
	b) How do you train owners to recognize signs of pain?
	6. For the last 6 ovariohysterectomies performed, please pull your records:
	a) Did you provide pre-surgical analgesia?
	b) Did you provide post-surgical analgesia?
	c) Were analgesics provided to owner?
	d) Did you follow up?
Chronic pain	7. How would you know if a cat had chronic pain? What does a cat in chronic pain look like?
	8. How would you know if a dog had chronic pain? What does a dog in chronic pain look like?

they were unavailable due to pre-scheduled clinical duties. Each interview was recorded using a digital audio recorder (Zoom H2n Handy Recorder; Zoom Corporation, Tokyo, Japan) and audio files were manually transcribed into text for analysis.

Responses were evaluated according to a scoring scheme that was developed based on published literature [World Small Animal Veterinary Association recommendations and consultation with a board-certified veterinary anesthesiologist (CIM)]. During development of the scoring scheme, a single, experienced observer (LCD) scored 5 interview transcripts for comprehensiveness and clarity using a draft scoring scheme. The draft scoring scheme was then modified accordingly to create a final scoring scheme that included sufficient detail to enable consistent scoring (see Table II for a simplified version). The responses to most questions were scored on a 3-point scale, representing 'insufficient' (score of 1), 'acceptable' (score of 2), and 'excellent' (score of 3) practices, respectively. For 4 questions, an adjustment to the score (+0.5) was possible (adjustment column, Table II). Adjustment scores were included to fulfill the long-term goal of welfare improvement by encouraging veterinary clinics to consider practices that might be above and beyond best practices. This would not lead to improvement among this sample of veterinary clinics, but rather in the future once the developed tool is formally implemented. For 2 questions, i.e., analgesia critical control points and provision of analgesics to surgical patients, it was only possible to distinguish between insufficient and acceptable practices; responses were therefore scored on a 2-point scale. Interviews from all 30 veterinary clinics were completed before beginning any scoring.

To investigate inter-observer reliability, 3 observers independently scored all transcripts in random order. The experienced observer was a doctoral student in animal welfare (observer 1, LCD) and the 2 inexperienced observers were a master's student in animal behavior and welfare (observer 2) and an upper-year undergraduate Bachelor of Science student with an interest in animal welfare (observer 3). Observers had no prior training in veterinary medicine, thus eliminating any potential scoring biases due to knowledge or experience in clinical veterinary medicine. Before scoring, the 2 inexperienced observers underwent a training session led by the experienced observer. Training involved reviewing the scoring scheme, independently scoring 3 transcripts, and reviewing scores verbally as a group. After training, the 2 inexperienced observers were only permitted to ask the experienced observer for clarification until the end of the fifth interview, so that reliability could be properly assessed between independent observers. To assess intra-observer reliability, the experienced observer re-scored all interview transcripts; this was done in a random order with a minimum of 6 wk between the first and second scoring to reduce the likelihood of recalling original scores when re-scoring responses.

Weighted *kappa* statistics, using exact Monte Carlo estimates and quadratic weighting, were computed using statistical software (SAS, Version 9.4; SAS Institute, Cary, North Carolina, USA) to statistically assess inter- and intra-observer scoring reliability. Quadratic weighting penalizes larger scoring disagreements more harshly than smaller disagreements by weighting the magnitude of the disagreement quadratically, e.g., scores that differ by 2 points on the scoring scale are given a weight of 4 (2^2) and scores that differ by 3 points are given a weight of 9 (3^2). Weighted *kappa* statistics were calculated at both the question and overall interview level (all questions pooled), based on pair-wise comparisons to the experienced observer (observer 1). Landis and Koch's guidelines (19) were used for interpretation: a weighted *kappa* (K_w) greater than 0.80 suggests almost perfect agreement, a K_w of 0.61 to 0.80 suggests substantial agreement, a K_w of 0.41 to 0.60 suggests moderate agreement, and a K_w of 0.21 to 0.40 suggests fair agreement. Descriptive statistics were calculated at the question level, based only on the experienced observer's scores, in order to derive frequencies, medians, and ranges in scores for each question.

Table II. Simplified scoring scheme for veterinarians' responses to interview questions about pain assessment and management.

Question	Excellent (3)	Acceptable (2)	Insufficient (1)	Adjustment (0.5)
Analgesia critical control points*	NA	Before and/or during surgery	After surgery or none	NA
Development of pain management plan	Individualized	Standard protocol	No protocol	NA
Form of information provided to owners	Written and oral	Oral only	Written only	Demonstration
OHE: analgesics provided to owner	3 or more days for cats, 5 or more days for dogs	2 or less days for cats, 4 or less days for dogs	None	NA
OHE: follow-up with owner	Phone call or check	Encourage owner to contact if questions or issues	No follow-up	NA
OHE: perioperative analgesia	Pre-surgery (opioid) and post-surgery (opioid and/or NSAID)	Pre-surgery (opioid)	During or after surgery, or none	Local analgesics
Provision of analgesics to surgical patients*	NA	All surgical patients	Not all surgical patients	NA
Recognition of chronic pain in cats	Objective scoring scheme	Subjective: pain-specific behavior	Subjective: non-specific behavior	NA
Recognition of chronic pain in dogs	Objective scoring scheme	Subjective: pain-specific behavior	Subjective: non-specific behavior	NA
Recognition of postoperative pain in cats	Behavior through interaction	Behavior at distance	Non-specific behavior	Use of pain scale
Recognition of postoperative pain in dogs	Behavior through interaction	Behavior at distance	Non-specific behavior	Use of pain scale
Training owners to recognize pain	Written and oral	Written only	Do not train owners	NA
Training staff to recognize pain	Formal training	Informal training	No training	NA

* Aspects of pain management were scored on a 2-point scale (insufficient/acceptable).
OHE — Ovariohysterectomy; NA — not applicable.

Results

Of the 474 veterinary clinics invited to participate, 41 (9%) agreed to be involved, 105 (22%) declined to participate, 209 (44%) did not respond despite follow-up, and 119 (25%) were not individually contacted to follow up because our target sample size of 30 veterinary clinics had been reached. Of the 41 clinics that agreed to be involved, the first 30 to schedule visits were included. Companion-animal veterinary practices accounted for 80% (24/30) and mixed-animal practices accounted for 20% (6/30) of all participating clinics. Participation rates were 6% (24/417) and 11% (6/57) for companion- and mixed-animal practices. Of the 6 veterinary clinics located outside the 100-km radius, 5 agreed to participate, 4 of which were mixed-animal practices. All participating clinics treated cats and all except 1 clinic treated dogs; 5 clinics (17%) were accredited by the American Animal Hospital Association and 5 clinics (17%) were participants in the American Association of Feline Practitioners Cat Friendly Practice program.

Interviews were conducted with practice owners at 22 clinics (73%) and with senior veterinarians at 8 clinics (27%). Most of the

Table III. Inter- and intra-observer reliability of scoring responses given by veterinarians from 30 veterinary clinics to interview questions about pain assessment and management, across 3 observers in pair-wise comparisons.

	Observer 1 vs. 2		Observer 1 vs. 3		Observer 1 vs. 1	
Question	K_w [95% CI]	% Agree	K_w [95% CI]	% Agree	K_w [95% CI]	% Agree
Analgesia critical control points	0.47 [-0.13 to 1.00]	93	0.36 [-0.07 to 0.79]	83	0.66 [0.24 to 1.00]	93
Development of pain management plan	0.65 [0.44 to 0.85]	73	0.33 [-0.05 to 0.71]	77	0.92 [0.82 to 1.00]	90
Form of information provided to owners	0.77 [0.66 to 0.88]	90	0.41 [-0.14 to 0.97]	93	0.83 [0.76 to 0.91]	97
OHE: analgesics provided to owner	0.92 [0.81 to 1.00]	93	0.73 [0.44 to 1.00]	85	0.93 [0.81 to 1.00]	96
OHE: follow-up with owner	1.00 [1.00 to 1.00]	100	0.79 [0.42 to 1.00]	93	0.81 [0.45 to 1.00]	96
OHE: perioperative analgesia	0.97 [0.92 to 1.00]	93	0.86 [0.64 to 1.00]	85	0.99 [0.98 to 1.00]	96
Provision of analgesics to surgical patients	U	87	U	87	U	100
Recognition of chronic pain in cats	0.79 [0.50 to 1.00]	93	0.66 [0.29 to 1.00]	90	0.90 [0.71 to 1.00]	97
Recognition of chronic pain in dogs	U	97	U	93	U	100
Recognition of postoperative pain in cats	0.85 [0.67 to 1.00]	87	0.84 [0.68 to 1.00]	87	0.99 [0.97 to 1.00]	97
Recognition of postoperative pain in dogs	0.57 [0.25 to 0.88]	72	0.60 [0.34 to 0.87]	72	0.85 [0.64 to 1.00]	93
Training owners to recognize pain	0.63 [0.36 to 0.90]	67	0.45 [0.20 to 0.70]	63	0.90 [0.78 to 1.00]	87
Training staff to recognize pain	0.80 [0.63 to 0.98]	80	0.67 [0.44 to 0.90]	67	0.96 [0.90 to 1.00]	93
Overall	0.83 [0.78 to 0.89]	85	0.73 [0.65 to 0.82]	82	0.92 [0.87 to 0.97]	94

Observer 1 — Experienced observer; Observers 2 and 3 — Trained, inexperienced observers; OHE — Ovariohysterectomy; K_w — Weighted *kappa*; % Agree — Raw percent agreement; U — Uninformative: not enough variation in responses to calculate a *kappa* statistic.

* As aspects of pain management were scored on a 2-point scale (insufficient-acceptable), it was not possible to receive a score of excellent.

interviews (77%, 23/30) were conducted with a female veterinarian. In 3 veterinary clinics, more than 1 veterinarian answered questions. Although each question was only answered once per participating clinic, different individuals answered different sections of the interview, either because veterinarians noted that another staff member would be better suited to provide an answer or for logistical reasons, such as limited time due to scheduled clinical duties.

Scoring reliability

Inter-observer reliability was relatively high overall [see Table III for weighted *kappa* statistics and 95% confidence intervals (CI)]. When pooling values for all questions together, for each of the 2 inexperienced, trained observers compared to the experienced observer, raw percent agreement was 85% and 82% and weighted *kappa* was 0.83 and 0.73, which reflected almost perfect and substantial agreement. Inter-observer reliability was among the lowest for responses related to analgesic control points and among the highest for responses related to perioperative analgesia for ovariohysterectomy. The graduate student observer (observer 2) had at least moderate agreement ($K_w > 0.40$, range: 0.47 to 1.00) with the experienced observer for all questions, with 4 of 13 questions having almost perfect agreement ($K_w > 0.80$). The undergraduate student observer's scoring (observer 3) was generally less consistent with the experienced observer's, although it still showed at least fair agreement for each question ($K_w > 0.20$, range: 0.33 to 0.86).

Intra-observer reliability was higher than inter-observer reliability, with an overall raw percent agreement of 94% (range: 87% to 100%, Table III) and an overall weighted *kappa* statistic of 0.92, equating to almost perfect agreement ($K_w > 0.80$). At the individual question level, weighted *kappa* statistics ranged from 0.66 to 0.99; similar to results for inter-observer reliability, scoring for analgesic control points was the least reliable, whereas scoring for perioperative analgesia for ovariohysterectomy showed the highest reliability. Nevertheless, only 1 question showed below almost perfect agreement ($K_w < 0.80$), yet still remained above the threshold for substantial agreement ($K_w > 0.60$).

Descriptive statistics

No single aspect of pain assessment and management scored excellent (score = 3) at every veterinary clinic (Figure 1). However, pain management practices relating to the provision of analgesics to surgical patients, the development of a pain management plan, the provision of accessible information to owners after their pets had undergone surgery, and recognition of chronic pain in dogs scored acceptable (score = 2) or higher at every veterinary clinic (Figure 1, Table IV). Most veterinarians received a score of excellent

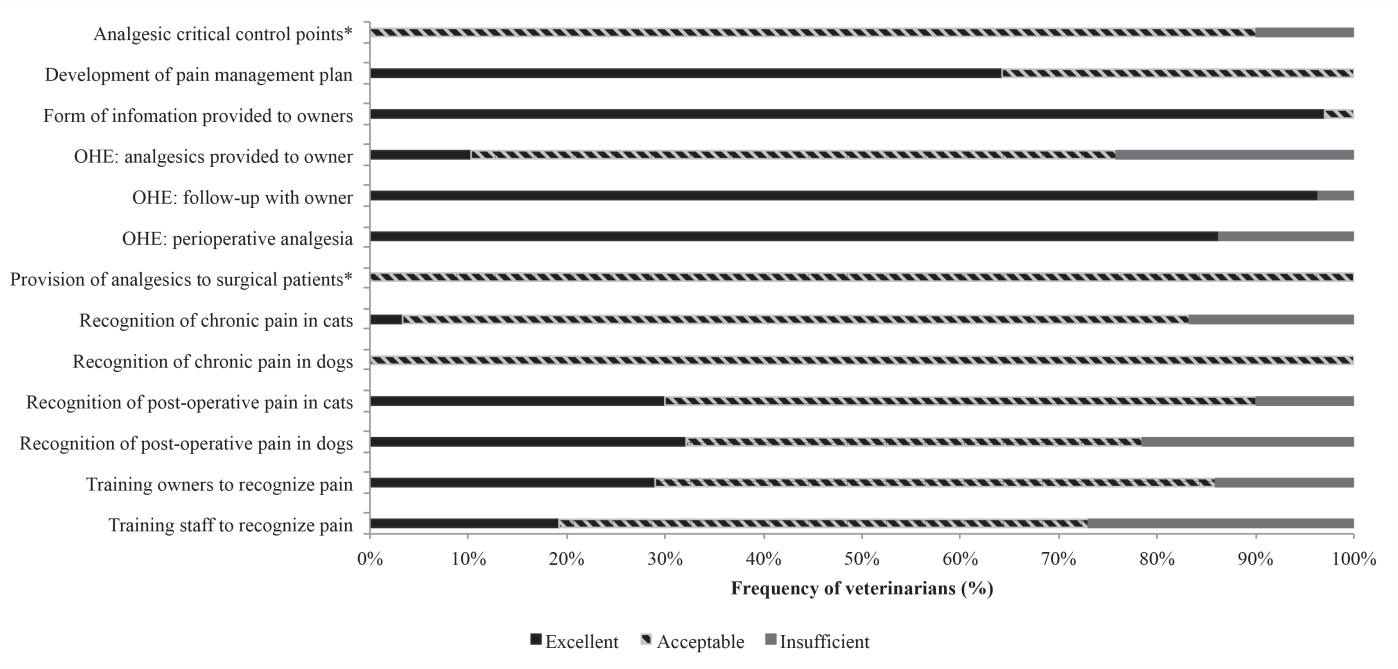

Figure 1. Frequency distribution of veterinarians from 30 veterinary clinics according to their interview response scores to questions about pain assessment and management
*** As aspects of pain management were scored on a 2-point scale (insufficient-acceptable), it was not possible to receive a score of excellent.**
OHE — Ovariohysterectomy.

for providing owners with both oral and written post-surgical care instructions (97% of participants) and following up with pet owners, either by phone or by an in-clinic check, after routine elective ovariohysterectomy (96% of participants) (Figure 1). Most veterinarians (86%) indicated that they provided appropriate pre-emptive analgesia, with an opioid and post-surgical analgesia or with an opioid and/or a nonsteroidal anti-inflammatory drug (NSAID), to patients undergoing ovariohysterectomy, which earned an excellent score in this area. Some veterinarians (14%) received insufficient scores, however, for reporting that they do not provide pre-emptive analgesia. Conversely, very few veterinary clinics attained excellent scores for the methods used to recognize chronic pain in cats and dogs or for providing analgesics for the ideal recommended duration after ovariohysterectomy.

No single aspect of pain assessment and management scored insufficient (score = 1) across every clinic. The greatest proportion of clinics was categorized as insufficient in the following areas: training staff to recognize pain in veterinary patients; providing owners with analgesics for their pets at discharge after ovariohysterectomy; and recognizing post-operative pain in dogs (Figure 1). Across all participating clinics, 27% reported that they do not train their staff members to identify patient pain, 24% do not provide any analgesics to owners to administer to their pets for pain management at home after an ovariohysterectomy, and 21% rely on non-specific behavioral signs, e.g., heart rate or vocalization, to identify post-surgical pain in dogs. Other veterinary clinics received a score of excellent in these 3 areas. With the exception of 4 items, i.e., development of a pain management plan, provision of analgesics to surgical patients, form of information provided to owners, and recognition of chronic pain in dogs, veterinary clinics scored across the full range of possible scores for each item (Table IV).

For 4 questions, a 'bonus' adjustment score was possible (Table II). While these assessment and management practices were used at a very low frequency, each was still employed by at least 1 participating veterinary clinic. A standardized, objective pain scale was routinely used to identify postoperative pain in cats and dogs by only 10% and 7% of veterinary clinics (3 and 2 clinics), respectively. During surgical patient discharge appointments, only 7% of veterinary clinics (2 clinics) demonstrated post-surgical instructions to owners, whereas 3% (1 clinic) used any form of local analgesic, e.g., line blocks at incision site, during routine ovariohysterectomy.

Discussion

Overall, the welfare assessment scheme used to assess pain management practices in companion-animal veterinary clinics in the current study showed high levels of reliability. Both inter- and intra-observer reliability were generally high, showing at least substantial agreement overall according to weighted *kappa* statistics (19). This suggests that scoring interview responses was dependable both between and within observers. Scoring by individuals with various backgrounds and levels of formal welfare training and repeat scoring by 1 individual over time should therefore lead to similar conclusions about animal welfare in relation to pain management. In addition, reliability was relatively high for most questions, which suggests that the scoring scheme adequately outlined scoring criteria for the type of responses routinely given by veterinarians. While reliability assessment is important during tool development, high reliability would also be important to maintain consistency between observers as part of a formal assessment program, assuming a team of observers score interviews from different veterinary clinics.

Table IV. Measures of central tendency for the experienced observer's scores for responses to interviews with veterinarians from 30 veterinary clinics about pain assessment and management.

Question	Median	Range
Analgesic critical control points*	2	1 to 2
Development of pain management plan	3	2 to 3
Form of information provided to owners	3	2 to 3.5
OHE: analgesics provided to owner	2	1 to 3
OHE: follow-up with owner	3	1 to 3
OHE: perioperative analgesia	3	1 to 3.5
Provision of analgesics to surgical patients*	2	2 to 2
Recognition of chronic pain in cats	2	1 to 3
Recognition of chronic pain in dogs	2	2 to 2
Recognition of postoperative pain in cats	2	1 to 3.5
Recognition of postoperative pain in dogs	2	1 to 3
Training owners to recognize pain	2	1 to 3
Training staff to recognize pain	2	1 to 3

Score of 1 — insufficient; 2 — acceptable; 3 — excellent; OHE — Ovariohysterectomy.

* As apects of pain management were scored on a 2-point scale (insufficient – acceptable), it was not possible to receive a score of excellent.

Criteria for acceptable levels of reliability do not exist. If a weighted *kappa* value of 0.40 is arbitrarily set as a cut-off for inclusion, as has been suggested by Fleiss and colleagues (20), all interview items would be retained in the welfare assessment tool. In other words, a minimum weighted *kappa* statistic of 0.40 was obtained for at least 1 pair of observers for every interview question. Only a single question, asking about critical control points for the provision of analgesia for surgical procedures, showed relatively low reliability across all pairs of observers (K_w = 0.47 and 0.36 for 2 inexperienced *versus* 1 experienced observer), including the experienced observer with repeat testing (K_w = 0.66). Despite this, raw percent agreement remained relatively high, at 93% and 83% agreement with the experienced observer and 93% agreement within the experienced observer. This discrepancy is likely due to nuances of the *kappa* statistic itself, rather than to issues with the assessment tool or scoring scheme; weighted *kappa* statistics become unreliable when variability is low or observations are rare or very common (18,21). In fact, responses to this question were quite uniform, yielding identical scores across 27 out of 30 veterinary clinics, according to the experienced observer's scores. For the purpose of practicality, and because this was an initial evaluation of a draft assessment scheme, inter- and intra-observer reliability was evaluated through the use of 3 observers. The use of additional observers could have strengthened the findings by providing evidence that similar levels of reliability could be achieved across a larger number of observers. Although there is no evidence to suggest otherwise, this could be further investigated in future research. A third type of reliability, test-retest reliability, has not been assessed here, but warrants further investigation before the tool is used in a formal welfare assessment scheme.

Although reliability tends to be given the highest importance when evaluating animal welfare assessment programs, it is also important to optimize feasibility and validity (16,22). In terms of feasibility, we were able to complete a full interview with at least 1 appropriate representative from every veterinary clinic. Since full interviews, including other sections not discussed here, can be completed in under an hour, they do not require an unreasonable amount of time for participants. Interviews can also be completed by telephone without introducing any systematic bias related to accessibility. This can be advantageous for scheduling according to the veterinarian's preferences and availabilities, while also reducing travel-related costs, further increasing the feasibility of the tool. Moreover, for interview scoring, only a single 2-h training session and 1 follow-up session was necessary to reach an acceptable degree of reliability, even with an observer who had minimal formal training in animal welfare. This suggests that observers can be adequately trained in a feasible amount of time and that advanced welfare training is not necessary. As such, suitable observers could include those with an undergraduate science degree, veterinarians, or veterinary technicians.

Although validity, or the extent to which a measure correctly assesses what it intends to assess (16), was not explicitly explored in this study, it still warrants discussion. The assessment tool was developed based on results of a survey of animal welfare researchers and companion-animal veterinarians (1). Contents of the tool are therefore expert validated and the tool has content validity. Moreover, the scoring scheme was based on published scientific evidence and guidelines in combination with expert opinion, which adds to the tool's content validity. In some cases, e.g., post-ovariohysterectomy follow-up procedures, there is a lack of scientific evidence to demonstrate that following acceptable *versus* excellent practices results in a measurable, positive effect on companion animal welfare. For the most part, however, extensive research is available to inform the areas of pain assessment and management that are addressed in this tool. This is particularly true for examining the efficacy of perioperative analgesic drugs, e.g., Slingsby and Waterman-Pearson (23), and the timing of their administration, e.g., Ingwersen et al (24), as well as identifying patient pain, e.g., Cambridge et al (25). Previous research has therefore facilitated the development of a tool with content validity.

Another type of validity relates to the ability of the tool to produce interview responses that correspond to true pain management practices. During an interview, the desire to provide responses that are deemed to be more generally acceptable (26) might bias responses. In this study, veterinarians might have provided what they perceived to be the "correct" answers regardless of whether they reflected their actual practices. This would be especially likely if veterinarians were aware of their own deficiencies with respect to pain management. It is difficult to quantify the extent of this bias and to evaluate the validity of interviewing as a method to assess pain management. Interviews yielded a large range of responses, including many inadequate scores, which suggests that participants were not simply providing what they perceived to be the "correct" response.

Interview responses should ideally be validated by comparing with actual practices, either through direct observation or by video or by checking against veterinary records. In this study, examination

of veterinary records was considered the best way to validate interview responses compared to direct observation of delivered pain medications or animal pain scoring. The use of records does not depend on surgical schedules, which would make it difficult to obtain and analyze an adequate sample size. At the time of enrollment, all veterinary clinics agreed that they would provide details from their veterinary records for the 6 most recent ovariohysterectomy procedures. At the time of the interview, however, only 3 veterinarians were willing to do so. Instead, most veterinarians provided general information from memory. As client records are confidential documents, interviewers were unable to retrieve these records themselves and participants were free to decline participation in any aspects of the study, based on ethics regulations for human research. It might be possible to investigate this aspect of tool validity in future studies, however, as veterinarians might be less reluctant to reveal actual surgical records as part of an already established welfare assessment program.

Since within- and between-observer scoring agreement suggest good reliability and responses did not appear to be systematically biased towards "correct" answers, the descriptive statistics from this study appear to be accurate about current practices related to pain management in this sample of companion-animal veterinary clinics. Thus, the descriptive statistics can be examined to determine which aspects of the tool best differentiate welfare standards among clinics. The overall goal is to use the final developed tool in a welfare assessment program to improve companion animal welfare in veterinary clinics through assessment and education. Therefore, it is important to set criteria high enough to encourage progress and improvement, yet moderate enough to be attainable by at least some proportion of veterinary hospitals. Given the wide ranges in scores for the majority of questions, with no median scores below 2, criteria seem to be attainable by most participating veterinary clinics, and are perhaps even too relaxed. Although many volunteer veterinary clinics were able to attain excellent scores for some areas, these items should still tentatively remain as part of the tool until it has been evaluated in a broader population of veterinary clinics. The pain management practices used by participating veterinary clinics might not represent those used by all veterinary clinics. Moreover, measures with little variation in responses might not be suitable for a welfare assessment program, as uniformity does not allow discrimination between 'good' and 'great' clinics.

In this study, 2 areas showed little variation in scoring: the provision of analgesics to surgical patients and the recognition of chronic pain in dogs. Every veterinary clinic received identical scores for their responses to these 2 questions. Regardless, assessing these aspects of pain management could still be informative if converted to a different type of measure. For example, reviewing records for the provision of analgesia might yield different, more discriminatory results than those provided through interviews. Moreover, scores for recognizing chronic pain in dogs showed an opportunity for enhancement, which suggests that it should be retained, regardless of low variability between veterinary clinics.

Conversely, if a large proportion of veterinary clinics score insufficient in any one area, or if very few veterinary clinics receive excellent scores, this would highlight aspects of pain management that should be prioritized in order to encourage enhancement. Training staff to recognize patient pain was one such area among this sample of veterinary clinics. Many participants noted that they expect most veterinarians and technicians to perfect these skills through their formal education, although relying on prior training might not be sufficient. For example, French veterinarians indicated in a survey that their veterinary education did not provide them with adequate skills to assess or manage patient pain (14) and Canadian veterinarians noted that veterinary education was the least important source of information on pain recognition and treatment (27). Since this second study, veterinary curriculums have included more information about pain and its treatment, so there is less reliance on continuing education. Regardless, veterinary curriculums differ by institution, which further complicates the idea of relying on prior training. Hewson and colleagues (13) discovered that Canadian veterinarians who graduated from certain veterinary schools, i.e., Atlantic Veterinary College and Ontario Veterinary College, used more analgesics than those trained at other institutions. A lack of in-clinic staff training might be mitigated by encouraging regular attendance at continuing education sessions, which allow veterinarians to stay current with ever-improving analgesic practices in a manner not possible through formal education alone.

The duration of analgesia provided after an ovariohysterectomy is another area that may benefit from enhancement based on this limited sample of veterinary clinics. Guidelines written by the World Small Animal Veterinary Association outline that post-ovariohysterectomy pain relief should be provided to cats and dogs for 3 and 5 d, respectively (2). As Dohoo and Dohoo have noted (27), it is important to not only provide analgesia, but also to do so for suitable durations. Given the frequency of ovariohysterectomies, providing a shortened analgesic regimen can potentially leave a large number of animals in pain. Furthermore, inadequate provision of analgesics can also leave pet owners with unmet expectations. A British study of analgesic use found that 61% of pet owners expect analgesics to be sent home with their pet following surgery, yet only 18% of veterinary clinics did so routinely (28). Similarly, only 16% of veterinarians surveyed in the United Kingdom, Australia, and New Zealand provided analgesia to ovariohysterectomy and castration patients post-discharge for 24 h or more (29). The results of this study are thus in line with other studies and show that analgesic practices might leave a number of cats and dogs in preventable pain upon discharge after an ovariohysterectomy, at least among this sample.

Furthermore, many veterinary clinics in this sample used insufficient methods to recognize pain, e.g., using non-specific behavioral signs to identify postoperative pain in dogs, and few used objective pain scales to identify both post-surgical and chronic pain in cats and dogs, e.g., the Glasgow Short Form Composite Measure Pain Scale (30) and the Helsinki Chronic Pain Index (31). Identifying pain is the cornerstone of effective pain management. Standard analgesic dosages might be insufficient to fully mitigate pain as animals vary in their response to procedures. Again, pain must first be identified in order to be successfully treated. Although no single pain scale can perfectly predict or quantify patient pain, validated pain scales can help to ensure an objective decision-making process when prescribing analgesics (2) and can mitigate the potentially detrimental effects of individual attitudes (27). In humans, when practitioners use a paper pain assessment form, the use of analgesics increases and

pain control improves (32). The finding that objective pain assessment methods are underused is consistent with previous research: 73% of French veterinarians did not use pain scales and 58% admitted to restricting the prescription of analgesics due to difficulty in recognizing pain (14).

It is worth noting that in every weak area that was identified in the current sample, some veterinary clinics attained excellent scores. This suggests that improvement is possible, likely through enhanced training and continuing education opportunities for all staff involved in pain management. Furthermore, if closed-ended questions had been used, it is possible that fewer individuals would have scored insufficiently or more individuals would have scored excellent. For example, directly asking about the use of an objective scoring scheme for recognizing chronic pain might have elicited more excellent responses when veterinarians did not think to discuss pain scales. Conversely, responses to closed-ended questions would be less valid, as veterinarians might be likely to respond positively when provided with the "correct" answer.

As with any study, the current results and conclusions are constrained by certain limitations. Participant selection was limited to clinics within a reasonable distance of the Ontario Veterinary College. Clinics that are geographically closer to a major teaching hospital might have different standards of practice than the general veterinary population. Furthermore, in order to recruit a sufficient number of veterinary clinics, all animal hospitals that fit the selection criteria were invited, resulting in a non-random sample. Since participation was voluntary, non-response and self-selection bias might have influenced the results. Overall, those who were willing to participate might not be representative of the larger population of veterinary clinics in Ontario or Canada, which limits our ability to generally apply results concerning specific pain assessment practices to all veterinary clinics in Ontario and across the country. Participants might be different than non-participants in terms of demographics, practices, or attitudes towards pain management. Multiple studies have demonstrated that both year of graduation and the veterinarian's gender influence attitudes towards pain management, with more recent graduates and females more likely to provide analgesics to their patients post-surgically (11–15). As most participating veterinarians were female, they may therefore have been more likely to use analgesics than male veterinarians. Conversely, many were also practice owners, so were less likely to be recent graduates. As pain assessment and management receive increased attention in veterinary curricula, the effect of year of graduation is likely to decrease over time.

Additionally, the sample encompassed veterinarians working in various settings, i.e., small- and mixed-animal practices and urban and rural settings. A limited number of clinics were enrolled in voluntary accreditation programs, i.e., American Animal Hospital Association (AAHA) accreditation and Cat Friendly Practice program. Although this indicates an increased interest in providing a higher standard of care, most clinics were not participants in these programs and the proportion of participants enrolled in these programs was similar to national enrollment statistics, e.g., 5/30 clinics (17%) AAHA Accredited *versus* an estimated 12% to 15% of hospitals across Canada and the United States (33). While senior veterinarians, particularly clinic owners, should be aware of their clinic's pain management practices, in reality different veterinarians can take different approaches or have individual attitudes or preferences that may not accurately reflect all aspects of the clinic's practices. The use of a single individual representative might therefore be biased. Overall, although these limitations might affect descriptive results, they should have minimal detrimental effect on the evaluation of the tool itself, as scoring reliability does not depend on the specific content of the response, but rather on the ability of different observers to score a response in a similar way.

This animal welfare assessment tool, including the scoring scheme for interview responses, was developed based on a combination of expert opinion, scientific evidence, and published guidelines for best practices. Despite this, there is some variation in expert opinion concerning certain aspects of pain management, e.g., use of nonsteroidal anti-inflammatory drugs (NSAIDs) *versus* opioids pre-surgically. Results might differ slightly if the scoring scheme was modified to reflect other opinions. Moreover, one of the main documents used was published in 2014 (2), and some interviews were completed before it was released. Even though this document (2) is based primarily on past scientific studies, some veterinary staff might not have been aware of the recommendations deemed best practice in our scoring scheme.

In summary, the pain assessment tool developed and evaluated in this study appears to have a good level of inter- and intra-observer reliability and is also feasible for use in the veterinary clinic setting. Its validity requires further investigation, however, and future research should focus on validating that interview responses are a good proxy for what actually happens in veterinary clinics. Overall, veterinary clinics scored across the full range of scores for most questions, which suggests that the questions included in the current tool will be useful for differentiating among welfare standards across clinics. The results of the current study identify aspects of pain management that could be improved and highlight a number of areas in which veterinary clinics are acting in accordance with best practices. Further research should include a larger and more geographically diverse sample of veterinary clinics in order to draw conclusions about current pain management practices across Canada.

Acknowledgments

This work was supported by the Ontario Veterinary College Pet Trust Fund. We wish to thank Brittany Lostracco, Melissa Speirs, and Bailey Kleefstra for their assistance with veterinary clinic visits, interview scoring, and data entry.

References

1. Dawson LC, Dewey CE, Stone EA, Guerin MT, Niel L. A survey of animal welfare experts and practicing veterinarians to identify and explore key factors thought to influence canine and feline welfare in relation to veterinary care. Anim Welf 2016;25:125–134.
2. Mathews K, Kronen P, Lascelles D, et al. Guidelines for recognition, assessment and treatment of pain. J Small Anim Pract 2014;55:E10–E68.
3. Weary DM, Niel L, Flower FC, Fraser D. Identifying and preventing pain in animals. Appl Anim Behav Sci 2006;100:64–76.

4. Hewson CJ, Dohoo IR, Lemke KA. Perioperative use of analgesics in dogs and cats by Canadian veterinarians in 2001. Can Vet J 2006;47:352–359.
5. Mathews KA. Pain assessment and general approach to management. Vet Clin North Am Small Anim Pract 2000;30:729–755.
6. Canada's Pet Wellness Report 2011. Canadian Veterinary Medical Association. Available from: https://www.canadianveterinarians.net/documents/canada-s-pet-wellness-report2011 Last accessed July 15, 2017.
7. Muir WW 3rd, Wiese AJ, Wittum, TE. Prevalence and characteristics of pain in dogs and cats examined as outpatients at a veterinary teaching hospital. J Am Vet Med Assoc 2004;224:1459–1463.
8. Döring D, Roscher A, Scheipl F, Küchenhoff H, Erhard MH. Fear-related behaviour of dogs in veterinary practice. Vet J 2009;182:38–43.
9. Glardon OJ, Hartnack S, Horisberger L. Analyse du comportement des chiens et des chats pendant l'examen physical en cabinet vétérinaire. Schweiz Arch Tierheilkd 2010;152:69–75.
10. Bell A, Helm J, Reid J. Veterinarians' attitudes to chronic pain in dogs. Vet Rec 2014;175:428–434.
11. Capner CA, Lascelles BDX, Waterman-Pearson AE. Current British veterinary attitudes to perioperative analgesia for dogs. Vet Rec 1999;145:95–99.
12. Dohoo SE, Dohoo IR. Factors influencing the postoperative use of analgesics in dogs and cats by Canadian veterinarians. Can Vet J 1996;37:553–556.
13. Hewson CJ, Dohoo IR, Lemke KA. Factors affecting the use of postincisional analgesics in dogs and cats by Canadian veterinarians in 2001. Can Vet J 2006;47:453–459.
14. Huggonnard M, Leblond A, Keroack S, Cadoré JL, Troncy E. Attitudes and concerns of French veterinarians towards pain and analgesia in dogs and cats. Vet Anaesth Analg 2004;31:154–163.
15. Williams VM, Lascelles BDX, Robson MC. Current attitudes to, and use of, peri-operative analgesia in dogs and cats by veterinarians in New Zealand. N Z Vet J 2005;53:193–202.
16. Martin P, Bateson P. Measuring Behaviour: An Introductory Guide. 2nd ed. Cambridge: Cambridge University Press, 1993.
17. Dawson LC. Development and evaluation of a canine and feline welfare assessment tool for use in companion animal veterinary clinics [PhD dissertation]. Guelph, Ontario: University of Guelph, 2016.
18. Sim J, Wright CC. The kappa statistics in reliability studies: Use, interpretation, and sample size requirements. Phys Ther 2005;85:257–268.
19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.
20. Fleiss JL, Levin B, Paik MC. Statistical Methods for Rates and Proportions. 3rd ed. Hoboken: John Wiley & Sons, 2003.
21. Viera AJ, Garrett JM. Understanding interobserver agreement: The kappa statistic. Fam Med 2005;37:360–363.
22. Meagher RK. Observer ratings: Validity and value as a tool for animal welfare research. Appl Anim Behav Sci 2009;119:1–14.
23. Slingsby LS, Waterman-Pearson AE. Postoperative analgesia in the cat after ovariohysterectomy by use of carprofen, ketoprofen, meloxicam or tolfenamic acid. J Small Anim Pract 2000;41:447–450.
24. Ingwersen W, Fox R, Cunningham G, Winhall M. Efficacy and safety of 3 versus 5 days of meloxicam as an analgesic for feline onychectomy and sterilization. Can Vet J 2012;53:257–264.
25. Cambridge AJ, Tobias KM, Newberry RC. Subjective and objective measurements of postoperative pain in cats. J Am Vet Med Assoc 2000;217:685–690.
26. Tourangeau R, Yan T. Sensitive questions in surveys. Psychol Bull 2007;133:859–883.
27. Dohoo SE, Dohoo IR. Postoperative use of analgesics in dogs and cats by Canadian veterinarians. Can Vet J 1996;37:546–551.
28. Demetrious JL, Geddes RF, Jeffery ND. Survey of pet owners' expectations of surgical practice within first opinion veterinary clinics in Great Britain. J Small Anim Pract 2009;50:478–487.
29. Farnworth M, Adams NJ, Keown A, Waran N, Stafford K. Veterinary provision of analgesia for domestic cats *(Felis catus)* undergoing gonadectomy: A comparison of samples from New Zealand, Australia and the United Kingdom. N Z Vet J 2014;62:117–122.
30. Reid J, Nolan AM, Hughes JML, Lascelles D, Pawson P, Scott EM. Development of the short-form Glasgow Composite Measure Pain Scale (CMPS-SF) and derivation of an analgesic intervention score. Anim Welf 2007;16:97–104.
31. Hielm-Björkman AK, Rita H, Tulamo RM. 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.
32. Baillie L. Clinical pain management: A review of pain assessment tools. Nurs Stand 1993;7:25–29.
33. About AAHA. American Animal Hospital Association. Available from: https://www.aaha.org/professional/about_aaha/default.aspx. Last accessed July 15, 2017.