Validation of the Italian version of the UNESP-Botucatu multidimensional composite pain scale for the assessment of postoperative pain in cats

Giorgia Della Rocca[1*], Alice Catanzaro[1], Maria Beatrice Conti[1], Antonello Bufalari[1], Valentina De Monte[1], Alessandra Di Salvo[1], Juliana Tabarelli Brondani[2] and Stelio Pacca Loureiro Luna[2]

[1] *Departement of Veterinary Medicine, University of Perugia, Perugia, Italy.*
[2] *School of Veterinary Medicine and Animal Science, Sao Paulo State University (UNESP), Botucatu, Brasil.*

[*] *Corresponding author at: Departement of Veterinary Medicine, University of Perugia, Via S.Costanzo 4, 06126 Perugia, Italy. Tel.: +39 075 5854612, e-mail: giorgia.dellarocca@unipg.it.*

Veterinaria Italiana 2018, **54** (1), 49-61. doi: 10.12834/Vetlt.567.2704.2
Accepted: 26.06.2016 | Available on line: 18.01.2018

Keywords
Cat,
Italian validation,
Pain scale,
UNESP-Botucatu MCPS.

Summary
The study described in this paper had the goal to validate the Italian version of the UNESP-Botucatu multidimensional composite pain scale (UNESP-Botucatu MCPS) to assess postoperative pain in cats using video analysis and psycometric testing. The English version of the scale was translated into Italian. Thirty videos of the perioperative period of ovariohysterectomy surgery were analysed by 5 Italian observers with the aim to determine the pain score using the Italian version of the scale and to verify the need for analgesic treatment for each cat. Obtained scores were submitted to psycometric validity, responsiveness, and reliability tests. Of the 3 domains identified by factor analysis, the internal consistency was excellent for 'Psychomotor changes' and 'Protection of the painful area and vocal expressions of pain', while 'Physiological variables' showed moderate internal consistency. Significant changes in pain scores in response to surgery and analgesics confirmed content and construct validity. The agreement between the 'gold standard' and the blinded observers supported the criterion validity. Inter- and intra-rater reliability ranged from good to very good for all scale items. The cut-off point for rescue analgesia was > 7. The study concluded that the Italian version of the UNESP-Botucatu MCPS is a valid and reliable instrument for assessing postoperative pain in cats. The cut-off point for rescue analgesia provides an additional tool for guiding analgesic therapy.

Validazione della versione italiana della UNESP-Botucatu multidimensional composite pain scale *per la misurazione del dolore post-operatorio nel gatto*

Parole chiave
Gatto,
Scala del dolore,
Validazione,
UNESP-Botucatu MCPS.

Riassunto
Il presente articolo riporta i risultati di uno studio finalizzato alla validazione, mediante l'impiego di video-analisi e di test psicometrici, della versione italiana della *UNESP-Botucatu multidimensional composite pain scale* (Unesp-Botucatu MCPS) per misurare il dolore post-operatorio nel gatto. Cinque osservatori italiani hanno analizzato i video di 30 gatti ripresi nel periodo perioperatorio dopo un intervento di ovarioisterectomia, attribuendo i punteggi con la traduzione italiana della scala inglese, per verificare la necessità di somministrare un trattamento analgesico a ciascun gatto. I risultati ottenuti sono stati sottoposti a test psicometrici per la valutazione della validità, responsività e affidabilità della scala. Dei 3 domini identificati mediante analisi fattoriale, la consistenza interna è risultata eccellente per le categorie "Modificazioni psicomotorie" e "Protezione dell'area dolente e espressioni vocali del dolore", mentre la categoria "Variabili fisiologiche" ha ottenuto una consistenza interna moderata. Cambiamenti significativi nei punteggi del dolore in risposta alla chirurgia e alla somministrazione di analgesici hanno confermato la validità di contenuto e costrutto. Il grado di concordanza tra il gold standard e gli osservatori italiani è risultato essere tra buono e molto buono per tutti gli elementi della scala, così come il grado di concordanza inter- e intra-osservatori. Il limite per evitare l'analgesia è risultato essere > 7. I risultati del presente studio evidenziano come la versione italiana della UNESP-Botucatu MCPS sia uno

strumento valido, sensibile ed affidabile per la valutazione del dolore postoperatorio nel gatto. La definizione di un limite per la somministrazione dell'analgesia fornisce un ulteriore strumento per guidare la terapia.

Introduction

Pain in domestic animal may be due to a variety of reasons (traumatic, surgical, and patological). It is important to treat pain in order to reduce sufferance and to facilitate recovery.

Recognising pain and assessing its intensity are essential for its correct management and treatment. Moreover, if the intensity of pain is not correctly assessed, the selection of the right analgesic will be hampered, resulting in the lack of pain relief.

The use of pain scales represents a valuable diagnostic aid, as they provide the veterinarian doctors with an objective and ready-to-use tool.

Unidimensional instruments such as analogical, visual, numerical or simply descriptive scales are universal and widely used for the assessment of pain in small animals (Anil *et al.* 2002). However, the extremely subjective nature of these scales cause inconsistency, due to variations in the observations when used by different observers (Holton *et al.* 1998). In order to limit subjectivity and increase the accuracy of the assessment, specific pain scoring systems have been developed for the evaluation of acute pain in dogs or cats. The 'University of Melbourne Pain Scale', the '4AVet scales', the 'Glasgow Composite Pain Tool - Short Form' and, the 'UNESP-Botucatu Multidimensional Composit Pain Scale' are some examples (Firth and Haldane 1999, Laboissière 2006, Reid *et al.* 2007, Brondani *et al.* 2011, Brondani *et al.* 2012, Brondani *et al.* 2013a). These scales are multiparametric tools consisting of numerical gradation (scoring) systems that include various categories (i.e. comfort, movement, behaviour, vocalizations, and physiological parameters). The sum of the scores obtained for each category gives a measure of the pain experienced by the animal, thus capturing the 'multidimensional' aspect of pain.

In order to produce more consistent and accurate results and to enable the comparison of outcomes from different studies, pain assessment tools need a validation process (Hellyer *et al.* 2007) to evaluate its validity, responsiveness, and reliability.

Validity is defined as the ability of the scale to measure what it intends to measure (Morton *et al.* 2005). It has traditionally been separated into 3 aspects, namely: content, criterion, and construct. The primary focus should be the hypothesis test methodology (Streiner and Norman 2008, Cook and Beckman 2006).

Content validity refers to the degree and adequacy with which the instrument actually measures the phenomenon of interest, in this case the pain. It determines whether the scale assesses the pain effectively and not other variables such as, for example, fear or anxiety (Bullock and Tenebein 2002). In other words, it is the degree to which the measurement represents the measured concept (Streiner and Norman 2008).

Construct validity examines whether the instrument detects changes in the construct theoretically conjectured, which provides the strongest evidence for validation (Crellin *et al.* 2007). Construct validity can also be checked using factor analysis, which distinguishes the underlying dimensions ('dimensionality') that establish the relationship between the instrument items.

Criterion validity establishes the validity of a measuring instrument by comparing it with some external criterion (Souza and Silva 2005).

Responsiveness or sensitivity to change assesses the ability of the scale to detect a significant change in pain scores in an expected direction, in response to events that produce pain, as surgical procedures, or reduce pain, as analgesics administration (Baeyer and Spagrud 2007).

Reliability assesses whether the instrument is measuring something in a reproducible way. Reliability of a scale is initially assessed by testing its 'internal consistency', which verifies the interrelations among the different items of the scale. It is then also necessary to assess the ability of the instrument to produce similar results when used by different individuals ('inter-observer reliability or reproducibility') or when used at different times by the same individual ('intra-observer reliability or stability') (Streiner and Norman 2008).

A validated instrument in one language is not automatically valid when translated in another language and culture (Guillemin *et al.* 1993, Souza and Rojjanasrirat 2011) and so a rigorous and thorough process of translation, cultural adaptation, and evaluation of the psychometric properties are necessary (Guillemin *et al.* 1993, Beaton *et al.* 2000). This ensures that the meaning and intent of the original items are maintained and that the scale

remains relevant (Sperber, 2004). As part of this process, it is suggested that the validation of the tool or scale should be performed using psychometric tests assessing the validity, responsiveness, and reliability of the instrument.

In 2011, an Italian version (ICMPS-SF) of the short form of the Glasgow Composite Measure Pain Scale (CMPS-SF) to measure acute pain in dogs was created and validated (Reid *et al.* 2011).

In the same year, an instrument for the assessment of postoperative pain in cats (the UNESP Botucatu-MCPS) was also developed and validated, first in Brazilian Portuguese (Brondani *et al.* 2011, Brondani *et al.* 2012, Brondani *et al.* 2013a) and, later, in English and Spanish (Brondani *et al.* 2013b, Brondani *et al.* 2014).

Due to the absence of validated tools to assess acute pain in cats and in view of the positive results obtained with the English and Spanish validation of Brazilian Portuguese scale, the aim of this study was to validate the Italian version of the UNESP Botucatu-MCPS following international guidelines proposed in the literature for cross-cultural validation (Beaton *et al.* 2000, Streiner and Norman 2008, Souza and Rojjanasrirat 2011). The hypothesis of this study was that if the translation and cultural adaptation were adequate, the Italian version would demonstrate validity, responsiveness, and reliability similar to the original Brazilian Portuguese scale as well as to the English and Spanish versions.

Materials and methods

Translation, back-translation and semantic equivalence

Initially, the scale was translated from English (using the already validated English version of the scale - Brondani *et al.* 2013b) to the Italian language by 2 independent bilingual (Italian-English) translators, whose native language was Italian. Both translated versions were synthesized into 1 version by a third person, also an Italian native speaker.

Subsequently, the scale in Italian (synthesized version) was translated back to the English language (reverse or back-translation) by a different bilingual, English native speaker (blinded to the original scale). Then the scale in English originated by reverse translation was compared with the original English scale and minor adjustments were made to the Italian scale in order to maintain maximal semantic equivalence.

Content validity - analysis by a committee of experts

Two Italian anesthesiologists (who were not involved in the previously mentioned translations) with expertise in the area of pain performed a detailed critical analysis of the content and comprehensibility of the scale (focused on the nuances of cultural adaptation) and judged the appropriateness of each item of the instrument. For this purpose, the following classification was applied: 1 = relevant, 0 = not sure, -1 = not relevant.

After modifying the contents following the suggestions of the experts, the final version of the Italian scale was generated.

Validity, responsiveness, and reliability testing based on video analysis

The psychometric properties of the Italian version of the scale were studied by analysing video previously recorded during the perioperative period of 30 cats subjected to ovariohysterectomy (Brondani *et al.* 2009)[1]. Cats were anesthetized with propofol IV (8 mg/kg), fentanyl (0.002 mg/kg) IV, and isoflurane in 100% of oxygen using a non-rebreathing system.

Cats were filmed for 5 minutes at 4 predetermined points in time within the perioperative period: T_1 (preoperative: 18 to 24 hours before surgery), T_2 (postoperative, before use of rescue analgesia: 30 minutes to 1 hour after surgery), T_3 (postoperative, after administration of rescue analgesia: approximately 4 hours after administering together the following analgesics: morphine 0.2 mg/kg IM, ketoprofen 2 mg/kg SC and dipyrone 25 mg/kg IV), T_4 (postoperative, approximately 24 hours after the end of surgery). Cats received the aforementioned association of analgesics at the conclusion of the T_2 video, approximately 1 hour after the end of the surgery.

The order of videos taken from each cat at each time point were randomized so that knowledge of the time point would not influence the results given by the observers, who would later evaluate these recordings. Moreover, before preoperative assessments a small piece of micropore™ medical tape was placed over the surgical area to avoid visualization of the presence or absence of the surgical wound.

Five Italian observers (a surgeon, an anesthesiologist, 2 critical care doctors, and a pharmacologist) watched the videos and recorded pain scores using the Italian version of the UNESP-Botucatu-MCPS. These blinded observers were provided with directions but not trained in the use of the UNESP-Botucatu-MCPS.

With the aim to determine the intraobserver reliability (see later), new DVDs with a rearrangement of the order of animals and time points were

[1] Study approved by the Ethics Committee for animal experimentation of the FMVZ-UNESP Botucatu with Protocol No 20/2008.

provided (in order to avoid the influence of the previous assessment). The second analysis occurred 1 month after the first one.

Dimensionality (construct validity)

In order to confirm the multidimensional structure of the original scale, the Italian version was submitted to an explorative factorial analysis based on the principal components analysis with varimax rotation in the correlation matrix. The varimax rotation method seeks to maximize the variance and the factors, which allow for a better representation of the variables (Kaiser 1958). The identification of factors was based on the Kaiser's criterion, which suggests retaining all components with an eigenvalue > 1 (Kaiser 1958). The factorial structure was determined by considering items with factor loading and communality > 0.5.

Content and construct validity by hypothesis testing

The methodology used to establish content and construct validity was based on hypothesis testing.

Content validity was based on the premise that if the scale measures pain, scores obtained preoperatively (before performing an elective surgery such as ovariohysterectomy) should be significantly lower when compared with those obtained in the postoperative period (after undergoing a painful situation). Thus, pain scores should increase significantly in T_2 compared with T_1.

Construct validity was determined by considering that if analgesics reduce pain, therefore scores obtained after administration of analgesics should decrease significantly with respect to those obtained after surgery, but before analgesic therapy. Thus, pain scores in T_3 should be significantly lower than in T_2.

As acute pain is of limited duration and tends to decrease, construct validity was also based on alterations in pain scores during the postoperative period. Thus, pain scores should diminish over time (T_2 vs. T_4).

In order to confirm content and construct validity, pain scores were summarized as median and semi-range and the Wilcoxon signed rank test was used for statistical comparisons ($p < 0.05$).

Criterion validity by comparison with a gold standard

Criterion validity was established from the correlation between pain scores recorded by blinded observers and those recorded by the Brazilian researcher, who developed the scale and who had advanced training and significant experience in feline pain assessment (considered as the gold standard).

The agreement between each blinded observer and the 'gold standard' was determined by the weighted Kappa coefficient, with 95% confidence interval (CI) (Cohen 1968).

The Altman's classification (0.81-1.00 very good; 0.61-0.80 good; 0.41-0.6 moderate; 0.21-0.4 fair and < 0.2 poor) (Altman 1991) was used to interpret the weighted Kappa coefficient and the CI, calculated for each item of the scale. This was done for cumulative results from all time points (preoperative, postoperative before and after analgesia, and 24 hours after the end of surgery).

Responsiveness (sensitivity to change) by hypothesis testing

The methodology used to establish content and construct validity was also used to assess the sensitivity to change of the scale.

Internal consistency

The Cronbach's alpha coefficient (Cronbach 1951) was used to assess the interrelationship among the different items of the scale (internal consistency) of the Italian version of the UNESP-Botucatu-MCPS. Values > 0.7 were considered acceptable (Jensen 2003). The coefficient was calculated for both the overall scale and each subscale identified by factor analysis.

Inter- and intra rater reliability

The inter- and intra-observer reliability was assessed for each scale items. The degree of agreement among different Italian observers and the degree of concordance among the assessments made by the same observer at different times, respectively, were determined.

The inter- and intra-rater reliability was evaluated using the intra-class correlation coefficient (ICC) (Bartko, 1966), consisting of a two-way random effect model and absolute agreement method with 95% CI. The results were interpreted using the Altman's classification as previously described. The ICC was calculated for each scale item considering all time points (preoperative, postoperative before and after analgesia and 24 hours after the end of surgery).

Cut-off point for rescue analgesia

To identify the minimum score related to the need for analgesic intervention (cut-off point), at the end

of each video analysis, blinded observers were asked to indicate the need for analgesics, based on their clinical experience.To this end the evaluation form included the following question: 'According to your clinical experience, do you think that it is necessary to provide rescue analgesia?'.

The rescue analgesic score was identified by analysing the Receiver Operating Characteristic (ROC) curve (Streiner and Cairney 2007). The ROC curve plots true positive rates (sensitivity) against false positive rates (1 - specificity) for a series of cut-off values. The optimal point is represented by the value at which the sensitivity and specificity are higher simultaneously.

The area under the curve (AUC), which indicates the discriminative ability of the method (i.e., evaluates the accuracy of the instrument's ability to classify correctly individuals with and without pain) (Deyo *et al*. 1991), was also calculated. This area ranges from 0.5 (no accuracy) to 1.0 (perfect accuracy). Values between 0.50 and 0.70, 0.70 and 0.90, and over 0.90 represent low, moderate, and high accuracy, respectively (Streiner and Cairney 2007).

Statistical analysis

Statistical analyses were performed with the programs SPSS® version 12.0.1 (SPSS, Chicago, Illinois, USA) and MedCalc® version 12.4.0 (MedCalc Software bvba, Ostend, Belgium).

Results

Content validity - analysis by a committee of experts

The items posture, comfort, miscellaneous behaviours, reaction to the palpation of the surgical wound, and vocalization were considered relevant (score = 1) by the experts. With regard to activity – i.e. attitude, reaction to the palpation of abdomen/flank, and appetite – there was no agreement among experts regarding the relevance of these items that were scored 1 or 0 by one or the other expert, respectively. Both experts were 'not sure' about the content validity of the item arterial pressure. However, because of this lack of agreement and because none of the items was considered irrelevant, the researchers decided to keep all items in the scale. Following their review, experts suggested minor semantic changes. The final scale included 10 items: posture, comfort, activity, attitude, miscellaneous behaviours, reaction to palpation of the surgical wound, reaction to palpation of the abdomen/flank, vocalization, arterial blood pressure, and appetite. Each item was assigned a score of 0-3, with 0 indicating normal or no change and 3 the most marked change for the item. The total score, calculated from the sum of the item scores, ranged from 0 (arbitrary absence of pain) to 30 (maximum pain).

Table I. *Results of factor analysis of the Italian version of the scale.* Factor analysis distinguishes the underlying dimensions ('dimensionality') that establish the relationship between the instrument items. The identification of factors is based on the Kaiser's criterion, which suggests retaining all components with an eigenvalue > 1. The factorial structure is determined by considering items with factor loading and communality > 0.5.

Item	Factor Loading*			Communality†
	Factor 1	Factor 2	Factor 3	
Posture	0.857‡	0.278	0.223	0.862
Comfort	0.859‡	0.261	0.236	0.862
Activity	0.836‡	0.152	0.281	0.801
Attitude	0.805‡	0.236	0.342	0.820
Miscellaneours behaviours	0.796‡	0.291	0.291	0.803
Reaction to palpation of surgical wound	0.247	0.903‡	0.157	0.900
Reaction to palpation of abdomen and flank	0.193	0.924‡	0.152	0.914
Appetite	0.291	0.033	0.833‡	0.779
Vocalization	0.482	0.547‡	0.055	0.534
Arterial blood pressure	0.296	0.290	0.681‡	0.634
Eigenvalue	3.95	2.36	1.59	NA
Variance (%)	39.57	23.61	15.92	NA
Accumulated variance (%)	39.57	63.18	79.11	NA

*Factor loading represents correlations between the variables and factors; †Communality represents the proportion of the variance for each item that can be explained by the factor; ‡Item was substantially loaded for the factor; NA = Not applicable.

Dimensionality (construct validity) by factor analysis

Factor analysis of the Italian version by use of the principal component extraction method and varimax rotation with the Kaiser criterion resulted in a 3-factor solution with eigenvalue magnitudes of 3.95, 2.36 and 1.59. Factor 1 explained 39.57% of the variance, and was denominated 'psychomotor changes' including the items posture, comfort, activity, attitude, and miscellaneous behaviours. Factor 2 accounted 23.61% of the variance, and was represented by the dimension 'protection of the painful area and vocal expressions of pain' including the items reaction to palpation of surgical wound, reaction to palpation of abdomen/flank, and vocalization. The third factor composed by the item arterial blood pressure and appetite was named 'physiological variables' and contributed with 15.92% of the total variance (Table I).

Content and construct validity and sensitivity to change by hypothesis testing

Since factor analysis confirmed the multidimensionality of the Italian version of the scale, the content and construct validity as well as the sensivity to change were determined for both total and partial or subscale scores. These increased significantly at T_2 (after surgery but before postoperative analgesics), when compared to T_1 (preoperative), this supporting content validity; and decreased significantly after cats received postoperative analgesics (T_2 vs. T_3) and over time (from T_2 to T_4), hence suporting construct validity (Table II).

Criterion validity by comparison with a gold standard

At all time points the agreement among blind observers and the 'gold-standard' observer, as evaluated by weighted kappa coefficient (95% CI), ranged from good to very good for all the scale items but activity, where 1 blind observed (Critical Care 2) showed a moderate correspondence. The same observer showed a lower agreement in quite all items. The item activity, attitude, and appetite showed the lowest agreement (Table III).

Internal consistency

Cronbach's alpha coefficient of the scale total score was 0.949, which indicated excellent internal consistency. The internal consistency of the subscale 1 (Psychomotor changes) and subscale 2

Table II. *Medians and semi-range of the pain scores determined by blinded observers and gold-standard by assessing video recordings from perioperative period of cats submitted to ovariohysterectomy.* The content and construct validity as well as the sensivity to change are determined for both total and partial or subscale scores.

Evaluation times	Pain Scores	Gold-standard	Blinded observers				
			Surgeon	Anesthesiologist	Critical Care 1	Critical Care 2	Pharmacologist
T_1 Preoperative	Total (0-30)	0.0 ± 2.5	1.0 ± 4.5	0.0 ± 1.5	1.5 ± 3.5	4.0 ± 5.0	0.0 ± 3.0
	Subscale 1 (0-15)	0.0 ± 2.5	0.0 ± 3.5	0.0 ± 1.5	1.5 ± 3.5	2.0 ± 3.5	0.0 ± 3.0
	Subscale 2 (0-9)	0.0 ± 0.0	0.0 ± 0.5	0.0 ± 0.5	0.0 ± 0.5	0.0 ± 2.0	0.0 ± 0.0
	Subscale 3 (0-6)	0.0 ± 0.0	0.0 ± 1.0	0.0 ± 1.0	0.0 ± 0.5	0.0 ± 1.0	0.0 ± 0.0
T_2 Postoperative: before rescue analgesia	Total (0-30)	20.5 ± 7.0*	21.5 ± 8.0*	18.5 ± 10.0*	19.0 ± 8.0*	22.0 ± 12.5*	19.5 ± 9.0*
	Subscale 1 (0-15)	12.0 ± 3.0*	13.5 ± 3.0*	11.0 ± 5.5*	11.0 ± 5.0*	13.0 ± 6.0*	12.0 ± 3.0*
	Subscale 2 (0-9)	6.0 ± 4.5*	5.5 ± 4.5*	6.0 ± 4.5*	4.0 ± 4.5*	6.0 ± 4.5*	4.0 ± 4.5*
	Subscale 3 (0-6)	3.0 ± 3.0*	3.0 ± 3.0*	2.5 ± 2.5*	3.0 ± 3.0*	3.5 ± 3.0*	3.0 ± 3.0*
T_3 Postoperative: after rescue analgesia	Total (0-30)	0.0 ± 3.0†	0.5 ± 3.5†	0.0 ± 2.0†	2.0 ± 3.5†	3.0 ± 4.5†	0.0 ± 2.5†
	Subscale 1 (0-15)	0.0 ± 2.5†	0.0 ± 2.5†	0.0 ± 2.0†	1.0 ± 3.0†	3.0 ± 3.5†	0.0 ± 2.0†
	Subscale 2 (0-9)	0.0 ± 0.5†	0.0 ± 0.5†	0.0 ± 0.0†	0.0 ± 0.5†	0.0 ± 0.5†	0.0 ± 1.0†
	Subscale 3 (0-6)	0.0 ± 1.0†	0.0 ± 1.0†	0.0 ± 1.0†	0.0 ± 1.5†	0.0 ± 1.5†	0.0 ± 1.0†
T_4 Postoperative: 24 hours after end of surgery	Total (0-30)	4.0 ± 7.0†	5.5 ± 8.5†	2.0 ± 7.0†	5.5 ± 6.5†	7.5 ± 7.5†	3.0 ± 6.5†
	Subscale 1 (0-15)	0.0 ± 3.5†	3.0 ± 5.0†	0.0 ± 1.5†	3.0 ± 3.5†	4.0 ± 4.5†	0.0 ± 3.0†
	Subscale 2 (0-9)	2.0 ± 3.0†	2.0 ± 3.5†	0.5 ± 3.0†	2.0 ± 2.5†	2.0 ± 3.0†	1.0 ± 3.0†
	Subscale 3 (0-6)	0.0 ± 1.0†	0.0 ± 1.5†	0.0 ± 1.0†	0.0 ± 1.5†	0.0 ± 1.5†	0.0 ± 1.5†

*Indicate content validity: pain scores in T_2 significantly higher than T_1 ($p < 0.001$); †Indicate construct validity: pain scores in T_3 and T_4 significantly lower than T_2 ($p < 0.001$); Subscale 1 = Posture, comfort, activity, attitude and miscellaneous behaviours; Subscale 2 = Reaction to palpation of surgical wound, reaction to palpation of abdomen/flank and vocalization; Subscale 3 = Arterial blood pressure and appetite.

(Protection of the painful area and vocal expressions of pain) was also excellent (0.949 and 0.836, respectively). The subscale 3 (Physiological variables) showed moderate internal consistency (0.563).

Table IV shows the Italian version of the UNESP-Botucatu-MCPS after content analysis and rearrangement of domains.

Inter-rater reliability (riproducibillity)

At all time points, the degree of agreement among the different Italian observers determined by ICC (95% CI) was good for all the scale items, with the exception of items 'reaction to the palpation of the surgical wound' and 'reaction to the palpation of abdomen/flank', where the reliability was very good. The items 'reaction to the palpation of the surgical wound' and 'appetite'/'vocalization' showed the highest and the lowest agreement, respectively (Table V).

Intra-rater reliability (stability)

At all time points, the level of agreement between the 2 assessments made 1 month apart by each

Table III. *Agreement between blinded observers and 'gold standard' for each scale item – video analysis, assessed by the weighted Kappa coefficient (95% CI), for each item of the scale considering all time points (preoperative, postoperative before and after rescue analgesia, and 24 hours after the end of surgery).* The Altman's classification was used to interpret the weighted Kappa coefficient.

	Surgeon	Pharmacologist	Critical care 1	Critical Care 2	Anesthesiologist
Posture	0.80 (0.70-0.90)	0.92 (0.88-0.97)	0.85 (0.80-0.90)	0.78 (0.70-0.86)	0.90 (0.81-0.97)
Comfort	0.83 (0.77-0.90)	0.91 (0.86-0.96)	0.74 (0.64-0.85)	0.89 (0.83-0.96)	0.88 (0.82-0.95)
Activity	0.78 (0.69-0.88)	0.84 (0.76-0.92)	0.69 (0.59-0.79)	0.51 (0.40-0.63)	0.87 (0.80-0.94)
Attitude	0.78 (0.70-0.86)	0.88 (0.82-0.93)	0.78 (0.70-0.86)	0.71 (0.60-0.82)	0.82 (0.73-0.91)
Miscellaneous behaviors	0.93 (0.88-0.97)	0.95 (0.92-0.98)	0.91 (0.85-0.97)	0.66 (0.54-0.77)	0.87 (0.79-0.95)
Reaction to palpation of surgical wound	0.89 (0.83-0.94)	0.90 (0.83-0.96)	0.96 (0.94-0.98)	0.89 (0.83-0.94)	0.87 (0.80-0.95)
Reaction to the palpation of abdomen/flank	0.90 (0.85-0.94)	0.86 (0.79-0.94)	0.87 (0.82-0.93)	0.86 (0.80-0.93)	0.83 (0.74-0.92)
Appetite	0.88 (0.80-0.97)	0.88 (0.77-0.99)	0.72 (0.56-0.88)	0.74 (0.62-0.86)	0.76 (0.59-0.93)
Vocalization	0.67 (0.47-0.87)	0.89 (0.81-0.97)	0.81 (0.67-0.95)	0.64 (0.46-0.83)	0.89 (0.77-0.99)

Interpretation: 0.81-1.0 = very good; 0.61-0.80 = good; 0.41-0.6 = moderate; 0.21-0.4 = fair; < 0.2 = poor.

Table IV. *UNESP-Botucatu MCPS – final Italian version. —cont'd*

Parte 1. Modificazioni psicomotorie (0-15)		
Postura	Il gatto assume una posizione naturale con i muscoli rilassati (si muove normalmente)	0
	Il gatto ha una postura normale ma è contratto (si muove poco o è riluttante al movimento)	1
	Il gatto è seduto o in decubito sternale con il dorso inarcato e la testa bassa; oppure è in decubito dorso-laterale con gli arti posteriori estesi o contratti	2
	Il gatto cambia di continuo posizione nel tentativo di trovare una postura confortevole	3
Comfort	Il gatto è a suo agio, sta sveglio o dorme ed interagisce se stimolato (interagisce con l'osservatore e/o è interessato all'ambiente circostante)	0
	Il gatto è quieto e scarsamente recettivo quando stimolato (interagisce poco con l'osservatore e/o non appare molto interessato all'ambiente)	1
	Il gatto è quieto e "dissociato dall'ambiente" (anche quando stimolato non interagisce con l'osservatore e/o non è interessato all'ambiente). Può stare con il muso rivolto verso la parte posteriore della gabbia	2
	Il gatto non è a suo agio, è agitato (cambia continuamente posizione) ed è poco recettivo agli stimoli o "dissociato dall'ambiente". Può stare con il muso rivolto verso la parte posteriore della gabbia	3
Attività	Il gatto si muove normalmente (si muove immediatamente non appena la gabbia viene aperta; fuori dalla gabbia si muove spontaneamente quando stimolato o manipolato)	0
	Il gatto si muove più del normale (spostandosi di continuo da una parte all'altra della gabbia)	1
	Il gatto è più tranquillo del normale (all'apertura della gabbia esita ad uscire e se portato fuori tende a ritornare all'interno, al di fuori della gabbia si muove poco se stimolato o manipolato).	2
	Il gatto è riluttante a muoversi (all'apertura della gabbia esita ad uscire e se portato fuori tende a ritornare all'interno, al di fuori della gabbia non si muove anche se stimolato o manipolato).	3

continued

Table IV. *UNESP-Botucatu MCPS – final Italian version. —cont'd*

Parte 1. Modificazioni psicomotorie (0-15)		
Attitudine	Osservare e rilevare la presenza delle seguenti condizioni mentali:	
	A – soddisfatto: il gatto è vigile ed interessato all'ambiente circostante (esplora l'ambiente), è socievole e interagisce con l'osservatore (gioca e/o risponde agli stimoli) * il gatto può inizialmente interagire con l'osservatore per distrarsi dal dolore. Va quindi osservato con attenzione per distinguere tra il gioco da distrazione e il gioco da soddisfazione	A
	B – non interessato: il gatto non interagisce con l'osservatore (non è interessato ai giochi o gioca poco; non risponde se l'osservatore lo chiama o lo accarezza) * nel caso di gatti che non amano giocare, valutare l'interazione con l'osservatore (risposte a richiami vocali e/o tattili)	B
	C – indifferente: il gatto non è interessato all'ambiente circostante (non è curioso; non esplora l'ambiente) * inizialmente il gatto può avere paura ad esplorare l'ambiente. L'osservatore deve accarezzare il gatto ed incoraggiarlo a muoversi da solo (tirandolo fuori dalla gabbia e/o facendogli cambiare posizione)	C
	D – ansioso: il gatto è spaventato (cerca di nascondersi e scappare) o nervoso (dimostra impazienza, ringhia, geme e soffia quando lo si accarezza o manipola)	D
	E – aggressivo: il gatto è aggressivo (tenta di mordere o graffiare se accarezzato o manipolato)	E
	Presenza dello stato mentale A	0
	Presenza di uno degli stati mentali B, C, D o E	1
	Presenza di due degli stati mentali B, C, D o E	2
	Presenza di tre o tutti gli stati mentali B, C, D o E	3
Comportamenti vari	Osservare l'animale e rilevare la presenza dei seguenti comportamenti:	
	A – il gatto è in decubito ed è tranquillo, ma muove la coda	A
	B – il gatto contrae ed estende gli arti posteriori e/o contrae i muscoli addominali (fianco)	B
	C – gli occhi del gatto sono parzialmente chiusi (occhi socchiusi)	C
	D - Il gatto lecca e/o morde la ferita chirurgica	D
	Nessuno dei comportamenti indicati è presente	0
	Presenza di uno dei comportamenti sopraelencati	1
	Presenza di due dei comportamenti sopraelencati	2
	Presenza di tre o tutti i comportamenti sopraelencati	3
Parte 2. Protezione dell'area dolente ed espressioni vocali di dolore (0–9)		
Reazione alla palpazione della ferita chirurgica	Il gatto non reagisce quando la ferita chirurgica viene toccata o compressa; oppure, non si osservano variazioni rispetto alle risposte evocate nella fase pre-operatoria (nel caso in cui sia stata effettuata una valutazione basale)	0
	Il gatto non reagisce quando la ferita chirurgica viene toccata, ma risponde alla compressione. In tal caso può vocalizzare e/o tentare di mordere	1
	Il gatto reagisce sia quando la ferita chirurgica viene toccata sia quando viene compressa. Può vocalizzare e/o tentare di mordere	2
	Il gatto reagisce già quando l'osservatore si avvicina alla ferita chirurgica. Può vocalizzare e/o tentare di mordere. Il gatto non permette all'osservatore di palpare la ferita	3
Reazioni alla palpazione dell'addome/fianco	Il gatto non reagisce quando si tocca o si comprime l'addome/fianco; oppure, non si osservano variazioni rispetto alle risposte evocate nella fase pre-operatoria nel caso in cui sia stata effettuata una valutazione basale). I muscoli dell'addome/fianco non sono tesi	0
	Il gatto non reagisce se l'addome/fianco viene toccato, ma reagisce alla compressione. I muscoli dell'addome/fianco sono tesi	1
	Il gatto reagisce sia al tocco che alla compressione dell'addome/fianco. I muscoli dell'addome/fianco sono tesi	2
	Il gatto reagisce al solo avvicinarsi dell'osservatore all'addome/fianco. Può vocalizzare e/o tentare di mordere. Il gatto non permette all'osservatore di palpare l'addome/fianco	3
Vocalizzazioni	Il gatto è tranquillo, fa le fusa se stimolato, o miagola quando interagisce con l'osservatore, ma non ringhia, non si lamenta e non soffia	0
	Il gatto fa le fusa spontaneamente (senza essere stimolato o manipolato dall'osservatore)	1
	Il gatto ringhia, si lamenta o soffia quando manipolato dall'osservatore (specie quando gli venga cambiata posizione)	2
	Il gatto ringhia, si lamenta o soffia spontaneamente (senza essere stimolato o manipolato dall'osservatore)	3
Parte 3. Variabili fisiologiche (0 - 6)		
Pressione arteriosa	Aumentata dello 0-15% rispetto al valore preoperatorio	0
	Aumentata dal 16% al 29% rispetto al valore preoperatorio	1
	Aumentata dal 30% al 45% rispetto al valore preoperatorio	2
	Aumentata di > 45% rispetto al valore preoperatorio	3

continued

Table IV. *UNESP-Botucatu MCPS – final Italian version. —cont'd*

Parte 3. Variabili fisiologiche (0 - 6)		
Appetito	il gatto mangia normalmente	0
	il gatto mangia più del normale	1
	il gatto mangia meno del normale	2
	il gatto non è interessato al cibo	3
		Punteggio totale (0-30)

Linee guida per l'impiego della scala
Inizialmente si deve osservare il comportamento del gatto senza aprire la gabbia. Osservare se il gatto sta riposando o se è attivo, se è interessato o disinteressato all'ambiente, se è silenzioso o se vocalizza. Verificare la presenza di specifici comportamenti (vedi sezione "comportamenti vari").
Aprire la gabbia ed osservare se il gatto esce rapidamente o se esita ad uscire. Avvicinarsi al gatto e valutare la sua reazione: se è socievole, aggressivo, spaventato, indifferente o se vocalizza. Toccare il gatto ed interagire con lui, verificando se è reattivo (se gradisce essere accarezzato e/o è interessato al gioco). Se il gatto esita ad uscire dalla gabbia, stimolarlo a muoversi (chiamandolo per nome e accarezzandolo) e manipolarlo (fargli cambiare posizione e/o tirarlo fuori dalla gabbia). Osservarlo quando è fuori dalla gabbia, verificando se si muove spontaneamente, con circospezione o se è riluttante a muoversi. Offrirgli cibo appetibile e valutare la sua risposta.*
Infine, porre il gatto in decubito laterale o sternale e misurare la pressione arteriosa. Valutare la sua reazione prima toccando la parete addominale/fianco (far scorrere le dita sulla zona) e di seguito esercitando una leggera compressione sull'area medesima. Aspettare un momento, e ripetere la stessa procedura per valutare la reazione del soggetto alla palpazione della ferita chirurgica.
Per valutare l'appetito nell'immediato post-operatorio, offrire al gatto una piccola quantità di alimento appetibile subito dopo il risveglio dall'anestesia. In questo momento la maggior parte dei gatti mangia normalmente, a prescindere dalla presenza o meno di dolore. Aspettare un po', offrire di nuovo il cibo e valutare la sua reazione.

blinded observer and assessed using ICC (95% CI) varied from good to very good for all scale items. The 2 Critical Care observers showed the lowest agreement (Table VI).

Cut-off point for rescue analgesia

From the analysis of the ROC curve, different cut-off points were suggested, highlighting the point represented by the greatest value of the sensitivity and specificity, simultaneously. The identified optimal cut-off point was > 7 (scale range 0-30 points), with a sensitivity of 94.4% (95% CI: 90.1-97.2%) and specificity of 97.0% (95% CI: 94.8-98.9%). The high AUC = 0.992 (95% CI: 0.981-0.997; $p < 0.001$) indicated that the instrument has excellent discriminatory ability (Figures 1 and 2). This way, the use of additional analgesia is recommended in scores ≥ 8 (0-30 points). This represents the 26.6% in relation to the maximum total score of the scale.

Table V. *Inter-rater reliability for each scale item – video analysis.* The inter-rater reliability (degree of agreement among the different Italian observers) was assessed by the Intra-class Correlation Coefficient (ICC) (95% CI) for each scale item considering all time points (preoperative, postoperative before and after rescue analgesia, and 24 hours after the end of surgery). The Altman's classification was used to interpret the ICC.

Items	Inter-rater reliability
Posture	0.76 (0.69 – 0.83)
Comfort	0.80 (0.74 – 0.85)
Activity	0.62 (0.50 – 0.72)
Attitude	0.71 (0.63 – 0.78)
Miscellaneous behaviors	0.78 (0.70 – 0.85)
Reaction to palpation of surgical wound	0.87 (0.84 – 0.91)
Reaction to the palpation of abdomen/flank	0.84 (0.79 – 0.88)
Appetite	0.69 (0.61 – 0.76)
Vocalization	0.69 (0.62 – 0.75)

Interpretation: 0.81-1.0 = very good; 0.61-0.80 = good; 0.41-0.6 = moderate; 0.21-0.4 = fair; < 0.2 = poor.

Discussion

This study was performed in order to validate the Italian version of the UNESP-Botucatu multidimensional composite pain scale (UNESP-Botucatu MCPS) to evaluate postoperative pain in cats.

Translation and cultural adaptation and psychometric tests were performed according to the rules commonly reported in the literature.

The results obtained from this analysis support the multidimensional nature of the Italian version of the scale, confirming its validity, reliability and interpretability (definition of the score for an analgesic action) for the evaluation of pain in cats undergoing ovariohysterectomy, as well as the possibility of its use by veterinarians with different cultural backgrounds.

The multidimensionality observed in the original Brazilian Portuguese scale and then in its English and Spanish version was confirmed also in the Italian version.

Little difference was found regarding the factor structure: the original Brazilian Portuguese scale

Table VI. *Intra-rater reliability for each scale item – video analysis.* The intra-rater reliability (degree of concordance among the assessments made by the same observer at different times) was assessed by the Intra-class Correlation Coefficient (ICC) (95% CI) for each scale item considering all time points (preoperative, postoperative before and after rescue analgesia, and 24 hours after the end of surgery). The Altman's classification was used to interpret the ICC.

	Surgeon	Pharmacologist	Critical care 1	Critical Care 2	Anesthesiologist
Posture	0.84 (0.70-0.89)	0.90 (0.85-0.93)	0.88 (0.83-0.92)	0.82 (0.75-0.88)	0.96 (0.94-0.97)
Comfort	0.92 (0.88-0.94)	0.89 (0.85-0.92)	0.72 (0.63-0.80)	0.88 (0.83-0.91)	0.95 (0.93-0.97)
Activity	0.85 (0.79-0.90)	0.83 (0.76-0.88)	0.74 (0.63-0.82)	0.63 (0.51-0.73)	0.96 (0.94-0.97)
Attitude	0.81 (0.74-0.87)	0.86 (0.81-0.90)	0.78 (0.70-0.85)	0.70 (0.59-0.78)	0.86 (0.80-0.90)
Miscellaneous behaviors	0.93 (0.90-0.95)	0.95 (0.93-0.96)	0.95 (0.93-0.96)	0.74 (0.65-0.82)	0.92 (0.88-0.95)
Reaction to palpation of surgical wound	0.91 (0.87-0.94)	0.92 (0.89-0.94)	0.98 (0.97-0.98)	0.88 (0.84-0.92)	0.89 (0.83-0.92)
Reaction to the palpation of abdomen/flank	0.90 (0.86-0.93)	0.89 (0.84-0.92)	0.93 (0.90-0.95)	0.84 (0.78-0.89)	0.87 (0.81-0.91)
Appetite	0.85 (0.79-0.89)	0.86 (0.80-0.90)	0.79 (0.71-0.85)	0.64 (0.50-0.74)	0.85 (0.78-0.90)
Vocalization	0.73 (0.63-0.81)	0.91 (0.87-0.94)	0.87 (0.82-0.91)	0.68 (0.57-0.77)	0.96 (0.95-0.98)

Interpretation: 0.81-1.0 = very good; 0.61-0.80 = good; 0.41-0.6 = moderate; 0.21-0.4 = fair; < 0.2 = poor.

consists of 4 dimensions, 1 of which includes the single item 'vocalizations' and was then called 'vocal expressions of pain' (Brondani *et al.* 2011). Factor analysis of the Italian version, in analogy with the English and the Spanish versions led to the identification of 3 dimensions only. Indeed, in the Italian version the item 'vocalizations' was included in the same dimension together with 'reaction to palpation of the surgical wound' and 'reaction to palpation of the abdomen/flank'. Therefore, this dimension has been called 'protection of the painful area and vocal expressions of pain'. Unlike the English version, but in analogy with results obtained for the Spanish scale, in the Italian version the category 'miscellaneous behaviours' has been included in the dimension 'psychomotor changes'. Thus, the Italian version of the scale shows a better multidimensional structure with respect to the original scale, as the presence of factors containing more than 2 categories is considered as desirable (Wiseman-Orr *et al.* 2006).

The proposed hypotheses to assess the validity of content and construct of the scale, i.e. the increase in pain scores after the surgical procedure and their decrease after the administration of analgesic therapy, respectively, have been widely used in human medicine for the validation of pain scales addressed to pediatric patients (Bullock and Tenenbein 2002, Manworren and Hynan 2003). This method has already been described in other studies of validation of pain scales addressed to the dog (Morton *et al.* 2005, Murrell *et al.* 2008), and allows for

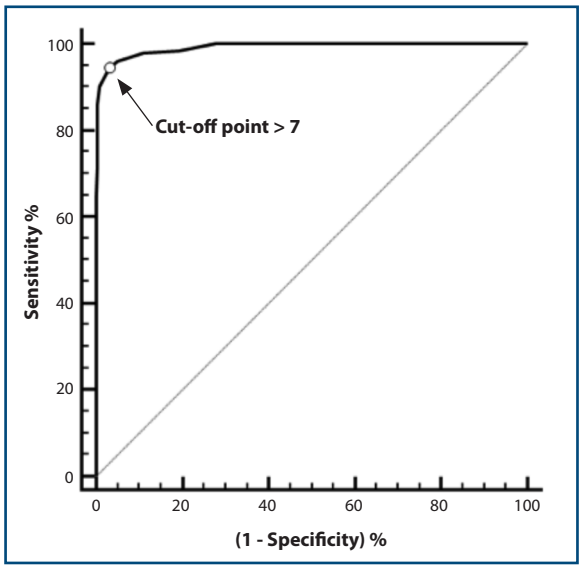

Figure 1. *ROC curve and the optimal cut-off point > 7 for rescue analgesia.*

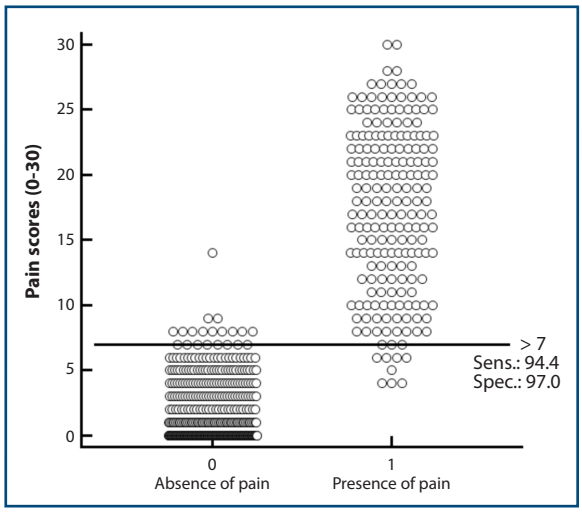

Figure 2. *Diagram illustrating the optimal cut-off point for rescue analgesia, represented by the greatest value of the sensitivity and specificity, simultaneously.*

evaluating both the content and construct validity as well as the responsiveness of the instrument (Baeyer and Spagrud, 2007). Indeed, the sensitivity to change is confirmed by variation in pain scores obtained during the postoperative period.

In the present study, the total score and the partial scores of the 3 subscales were significantly different in response to surgery, analgesic administration, and over time during the postoperative period. This confirms the content and construct validity and responsiveness of the Italian version of the scale. This result coincides with the one obtained during the validation process of the original scale in Brazilian Portuguese, as well as with those obtained as a result of the validation process of the English and Spanish version (Brondani *et al.* 2012, Brondani *et al.* 2013a, Brondani *et al.* 2013b, Brondani *et al.* 2014). The validation of the partial score of the various subscales permits to assess separately each aspect of the construct. This means that, it is possible to omit the assessement of variables in a given subscale without compromising the global assessment of pain if there are some technical difficulties for the evaluation of the elements in a given subscale (e.g. the lack of equipment to measure the blood pressure).

Criterion validity is usually assessed by the correlation of the results obtained using the scale to be validated with another tool, regarded as the gold-standard (Morton *et al.* 2005). Due to the lack of a validated scale for the assessment of pain in cats to be used as a reference model, the criterion validity was determined by comparing the scores assigned by the Italian blinded evaluators with those obtained by the observer who developed the scale, which is considered the gold-standard. This alternative method is similar to the one used by Gauvain-Piquard and colleagues (Gauvain-Piquard *et al.* 1999) to validate a scale for the assessment of pain in children. The high degree of correlation obtained in this study emphasizes the criterion validity of the Italian version of the scale.

As for the internal consistency of the Italian version of the scale, excellent results have been observed both in relation to the total score (for which a value of 0.949 was obtained) and in relation to subscales 1 and 2, namely 'psychomotor changes' and 'protection of the painful area and vocal expressions of pain'. Both scales were analysed individually (obtained values: 0.949 and 0.836, respectively). These findings, similar to the ones obtained for the Brasilian Portuguese, the English, and the Spanish versions (Brondani *et al.* 2011, Brondani *et al.* 2013b, Brondani *et al.* 2014) show that the results obtained by assessing the pain experienced by an animal with the aid of the scale can be interpreted using both the total score (overall assessment of the severity of the pain) or the partial score of each of the above mentioned subscales.

In contrast, the internal consistency of the subscale called 'Physiological variables' was moderate (obtained score: 0.563). This finding differed from that obtained both in the original Brasilian Portuguese version (0.80) (Brondani *et al.* 2011) and the English version (0.28) (Brondani *et al.* 2013b), but was similar to that obtained in the Spanish version of the scale (0.55) (Brondani *et al.* 2014). This result is likely due to the close similarity of the 2 languages (Italian and Spanish). The variability observed in different languages with regard to the internal consistency of the third subscale clearly reflects the limits of its categories. Therefore, the dimension 'physiological changes' should always be assessed in combination with the other 2 dimensions.

The reliability of a rating scale, in terms of reproducibility and stability, is a fundamental requirement in all those cases in which the collected data are derived from observational assessments (Beyer and Wells 1989). Due to the observational nature of the evaluation of pain in animals (Anil *et al.* 2002), the reliability both inter- and intra-observer was evaluated with regard to the scores given by the various blinded observers during the 2 subsequent videoanalises. In the Italian version, all the elements of the scale showed an appropriate degree of reliability, since the level of correlation between the assessments made both by different observers and by the same observer at different times ranged between good and very good. The calculation of the ICC is considered as the most appropriate statistical method for the analysis of reliability, even if it shows some limitations with homogeneous samples (Deyo *et al.* 1991). To ensure the heterogeneity of data, in the present study analyses were performed by grouping the 4 evaluation times (T_1, T_2, T_3 and T_4). The reliability of the Italian version of the scale was satisfactory and coincides with results obtained with the original scale in Brazilian Portuguese and with the English and Spanish versions (Brondani *et al.* 2013a, Brondani *et al.* 2013b, Brondani *et al.* 2014). The reproducibility and stability of the scale can be attributed to the detailed description of the categories included in the tool; these are described in such a way as to minimize the subjectivity of the analysis.

The availability of a minimum score (cut-off) beyond which it is necessary to apply a rescue anagesia is an essential requirement in a rating pain scale, as it helps the physician in choosing the analgesic treatment (Reid *et al.* 2007). In the present study, the analysis of the ROC curve was used for the determination of this score. The same method was also used in the validation process of the original scale, as well as in the English and Spanish versions (Brondani *et al.* 2013a, Brondani *et al.* 2013b, Brondani *et al.* 2014).

Although this statistical technique is innovative in veterinary medicine, it has already been used in studies focusing on human medicine (Hünseler *et al.* 2011). The ROC curve analysis allows for determining the ability of a test to discriminate groups, establish an optimal cut-off, and compare the performance of 2 or more tests (Streiner and Cairney, 2007). As for the original Brazilian Portuguese scale, the English, and Spanish versions (Brondani *et al.* 2013a, Brondani *et al.* 2013b, Brondani *et al.* 2014), the cut-off identified for the Italian version is > 7, which means that there is the need for an analgesic therapy once a score ≥ 8 (26.6% of the total score) is obtained. However, the use of painkillers in animals with score ≤ 7 should not be denied if the clinician believes it as necessary.

The results of this study show that the Italian version of the UNESP-Botucatu MCPS is a viable, sensitive, and reliable tool for the assessment of postoperative pain in cats undergoing ovariohysterectomy.

Using this scale, the veterinarian is eased in take the appropriate clinical decisions related to analgesic therapy within the postoperative period. Standardized tools for pain assessment, validated in different languages and cultures, may also be used to compare information that can derive from scientific studies.

References

Altman D.G. 1991. Some common problems in medical research. *In* Practical statistics for medical research. Chapman and Hall, London, 404-408.

Anil S.S., Anil I. & Deen J. 2002. Challenges of pain assessment in domestic animals. *J Am Vet Med Assoc*, **220**, 313-319.

Baeyer V.C. & Spagrud L.J. 2007. Systematic review of observational (behavioral) measures of pain for children and adolescents aged 3 to 18 years. *Pain*, **127**, 140-150.

Bartko J.J. 1966. The intraclass correlation coefficient as a measure of reliability. *Psycol Rep*, **19**, 3-11.

Beaton D.E., Bombardier C., Guillemin F. & Ferraz M.B. 2000. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, **25**, 3186-3191.

Beyer J.E. & Wells N. 1989. The assessment of pain in children. *Pediatr Clin North America*, **36**, 837-854.

Brondani J.T., Luna S.P.L., Beier S.L. & Minto B.W. 2009. Analgesic efficacy of perioperative se of vedaprofen, tramadol or their combination in cats undergoing ovariohysterectomy. *J Feline Med Surg*, **11**, 420-429.

Brondani J.T., Luna S.P.L. & Padovani C.R. 2011. Refinement and initial validation of a multidimensional composite scale for use in assessing acute postoperative pain in cats. *Am J Vet Res*, **72**, 174-183.

Brondani J.T., Luna S.P.L., Minto B.W., Santos B.P.R., Beier S.L., Matsubara L.M. & Padovani C.R. 2012. Validade e responsividade de uma escala multidimensional para avaliação de dor pós-operatória em gatos. *Arq Bras Med Vet Zootec*, **64**, 1529-1538.

Brondani J.T., Luna S.P.L., Minto B.W., Santos B.P.R., Beier S.L., Matsubara L.M. & Padovani C.R. 2013a. Confiabilidade e pontuação minima relacionada à intervenção analgésica de uma escala multidimensional para avaliação de dor pós-operatória em gatos. *Arq Bras Med Vet Zootec*, **65**, 153-162.

Brondani J.T., Mama K.R., Luna S.P.L., Wright B.D., Niyom S., Ambrosio J. & Vogel P. 2013b. Validation of the English version of the UNESP-Botucatu multidimensional composite pain scale for assessing postoperative pain in cats. *BMC Veterinary Research*, **9**, 143.

Brondani J.T., Luna S.P.L., Crosignani N., Redondo J.I., Granados M.M., Bustamante H., Palacios C. & Otero P. 2014. Validez y confiabilidad de la versión en español de la escala multidimensional de la UNESP-Botucatu para evaluar el dolor postoperatorio en gatos. *Arch Med Vet*, **46**, 477-486.

Bullock B. & Tenenbein M. 2002. Validation of 2 pain scales for use in the pediatric emergency department. *Pediatrics*, **110**, 1-6.

Cohen J. 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psych Bull*, *70*, 213-220.

Cook D.A. & Beckman T.J. 2006. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med*, **119**, 166.e-166.e16.

Cronbach L.J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297-333.

Crellin D., Sullivan T.P., Babl F.E., O'Sullivan R., Hutchinson A. 2007. Analysis of the validation of existing behavioral pain and distress scales for use in the procedural setting. *Paediatr Anaesth*, **17**, 720-733.

Deyo R.A., Diehr P. & Patrick D.L. 1991. Reproducibility and responsiveness of a health status measures. *Control Clin Trials*, **12**, 142-158.

Firth A.M. & Haldane S.L. 1999. Development of a scale to evaluate postoperative pain in dogs. *J Am Vet Med Assoc*, **214**, 651-659.

Gauvain-Piquard A., Rodary C., Rezvani A. & Serbouti S. 1999. The development of the DEGRR: a scale to assess pain in young children with cancer. *Eur J Pain*, **3**, 165-176.

Guillemin F., Bombardier C. & Beaton D. 1993. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*, **46**, 1417-1432.

Hellyer P., Rodan I., Brunt J., Downing R., Hagedorn J.E. & Robertson S.A. 2007. AAHA/ AAFP Pain management guidelines for dogs & cats. *J Am Anim Hosp Assoc*, **43**, 235-248.

Holton L.L., Scott E.M., Nolan A.M., Reid J., Welsh E. & Flaherty D. 1998. Comparison of three methods used for assessment o pain in dogs. *J Am Vet Med Assoc*, **212**, 61-66.

Hünseler C., Merkt V., Gerloff M., Eifinger F., Kribs A. & Roth B. 2011. Assessing pain in ventilated newborns and infants: validation of the Hartwig score. *Eur J Pediatr*, **170**, 837-853.

Jensen M.P. 2003. Questionnaire validation: a brief guide for readers of the research literature. *Clin J Pain*, **19**, 345-352.

Kaiser H.F. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187-200.

Laboissière B.L.A. 2006. Validation statistique des grilles 4A-Vet d'évaluation de la douleur post-opératoire chez le chien et le chat. Thesis, Ecole Nationale Vétérinaire de Nantes.

Manworren R.C. & Hynan L. 2003. Clinical validation of FLACC: preverbal patient pain scale. *Pediatr Nurs*, **29**, 140-146.

Morton C.M., Reid J., Scott M.E., Holton L.L. & Nolan A.M. 2005. Application of a scaling model to establish and validate an interval level pain scale for assessment of acute pain in dogs. *Am J Vet Res*, **66**, 2154-2166.

Murrell J.C., Psatha E.P., Scott E.M., Reid J. & Hellebrekers L.J. 2008. Application of a modified form of the Glasgow pain scale in a veterinary teaching centre in the Netherlands. *Vet Rec*, **162**, 403-408.

Reid J., Nolan A.M., Hughes J.M., Lascelles D., Pawson P. & Scott E.M. 2007. Development of the short- form Glasgow composite measure pain scale (CMPS-SF) and derivation of an analgesic intervention score. *Anim Welf*, **16**, 97-104.

Reid J., Colpo R., Scott E.M. & Nolan A.M. 2011. Creation of an Italian version (ICMPS-SF) of the short form of the Glasgow Composite Measure Pain Scale (CMPS-SF) to measure acute pain in dogs and its validation. British Small Animal Veterinary Association (BSAVA) Congress 2011, Birmingham, 31 March - 3 April 2011. http://www.vin.com/doc/?id=4823443.

Souza F.F. & Silva J.A. 2005 A métrica da dor (dormetria): problemas teóricos e metodológicos. *Rev Dor*, **6**, 469-513.

Souza V.D. & Rojjanasrirat W. 2011. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract*, **17**, 268-274.

Sperber A. 2004. Translation and validity of study instruments for cross-cultural research. *Gastroenterology*, **126** (Suppl 1), s124-128.

Streiner D.L. & Cairney J. 2007. What's under the ROC? An introduction to receiver operating characteristics curves. *Can J Psychiatry*, **52**, 121-128.

Streiner D.L. & Norman G.R. 2008. Health measurement scales: a practical guide to their development and use. 4th ed. Oxford University Press, New York, USA,.

Wiseman-Orr M.L., Scott E.M., Reid J. & Nolan A.M. 2006. Validation of a structured questionnaire as an instrument of measure chronic pain in dogs on the basis of effects on health-related quality of life. *Am J Vet Res*, **67**, 1826-1836.