

## Confiabilidade e pontuação mínima relacionada à intervenção analgésica de uma escala multidimensional para avaliação de dor pós-operatória em gatos

[Reliability and cut-off point related to the analgesic intervention of a multidimensional composite scale to assess postoperative pain in cats]

J.T. Brondani<sup>1</sup>, S.P.L. Luna<sup>1</sup>, B.W. Minto<sup>2</sup>, B.P.R. Santos<sup>1</sup>, S.L. Beier<sup>3</sup>,  
L.M. Matsubara<sup>1</sup>, C.R. Padovani<sup>4</sup>

<sup>1</sup>Faculdade de Medicina Veterinária e Zootecnia – UNESP – Universidade Estadual Paulista, Botucatu, SP

<sup>2</sup>Faculdade de Ciências Agrárias e Veterinárias – UNESP – Universidade Estadual Paulista, Jaboticabal, SP

<sup>3</sup>Escola de Veterinária – UFMG – Universidade Federal de Minas Gerais, Belo Horizonte, MG

<sup>4</sup>Instituto Biológico – UNESP – Universidade Estadual Paulista, Botucatu, SP

### RESUMO

Avaliou-se a confiabilidade e determinou-se a pontuação mínima relacionada à necessidade de intervenção analgésica de uma escala multidimensional para avaliação de dor pós-operatória em gatas submetidas à ovariectomia. Trinta gatas foram filmadas em quatro momentos predeterminados: pré-operatório, pós-operatório antes e depois da analgesia de resgate e às 24 horas após a cirurgia. Quatro observadores, não cientes dos momentos de avaliação, e um observador, considerado padrão-ouro – pesquisador que desenvolveu a escala –, avaliaram os vídeos em ordem aleatória, a partir da aplicação da escala, e determinaram a suficiência ou não de analgesia, baseando-se na experiência clínica. Os vídeos foram reavaliados dois meses após a primeira análise. O grau de concordância entre as avaliações dos diferentes observadores e das avaliações feitas pelo mesmo observador em diferentes ocasiões foi verificado pelo coeficiente de correlação intraclass, que variou entre bom e muito bom para todos os itens da escala, exceto atividade, atitude e miscelânea de comportamentos que apresentaram confiabilidade moderada. O ponto de corte ótimo (>7) foi selecionado pela análise descritiva e da curva *Receiver Operating Characteristic*. Os resultados suportam a confiabilidade em termos de reprodutibilidade e estabilidade e indicam que a terapia analgésica deve ser considerada em pontuações  $\geq 8$ .

Palavras-chave: gatos, dor, escalas, confiabilidade, intervenção analgésica

### ABSTRACT

*The aim with this study was to evaluate the reliability, and to establish the minimum score relating to the analgesic intervention of a multidimensional scale for use in assessing postoperative pain in cats undergoing ovariohysterectomy. Thirty cats were video recorded at 4 predetermined moments: prior to surgery, before and after rescue analgesia postoperatively, and 24 hours after surgery. Four observers blind to the times of assessment, and a gold standard observer (researcher who developed the scale) evaluated the video segments in random order using the scale, and determined the adequacy or not of analgesia based on clinical experience. The videos were reassessed two months after the first analysis. The agreement between the assessment of different observers, and the assessment made by the same observer on different occasions has been verified by the intra-class correlation coefficient that ranged from good to very good for all scale items, except activity, attitude and miscellaneous behaviors, which showed moderate reliability. The optimal cutoff point (>7) was selected based on descriptive analysis and Receiver Operating Characteristic curve. The findings support the reliability in terms of reproducibility and stability, and indicate that analgesic therapy should be considered in scores  $\geq 8$ .*

Keywords: cats, pain, scales, reliability, cut-off to rescue analgesia

---

Recebido em 28 de abril de 2011

Aceito em 6 de agosto de 2012

E-mail: jtbrondani@yahoo.com

## INTRODUÇÃO

A consciência da importância do controle adequado da dor nos animais já está bastante difundida entre os médicos veterinários. Entretanto, deve-se ressaltar que, para a instituição de uma terapia analgésica efetiva, que realmente alivie o sofrimento oriundo da dor, é necessária inicialmente uma avaliação criteriosa desta. Vista a complexidade e a subjetividade da experiência dolorosa, a mensuração da dor constitui um grande desafio, especialmente nos animais que são desprovidos da capacidade de autorrelato (Anil *et al.*, 2002). A necessidade de método observacional para o reconhecimento e a quantificação da dor nos animais tem estimulado o desenvolvimento de instrumentos de avaliação alicerçados em alterações comportamentais espécie-específicas (Holton *et al.*, 2001), haja vista a falência das variáveis fisiológicas em se correlacionarem com os escores de dor (Holton *et al.*, 1998a; Cambridge *et al.*, 2000), com ressalva à pressão arterial em gatos (Smith *et al.*, 1996; Brondani *et al.*, 2009).

Na medicina veterinária, a quantificação da dor ficou restrita, por um longo período, à utilização de instrumentos como a escala analógica visual, a escala numérica e a escala descritiva simples. Entretanto, a falta de um critério específico ou objetivo para a avaliação torna estas escalas sujeitas a grande viés e dependentes demasiadamente da experiência pessoal do observador (Baeyer e Spagrud, 2007). Já foi demonstrado em estudo de avaliação da dor em cães que tais instrumentos não produzem resultados consistentes quando utilizados por diferentes observadores (Holton *et al.*, 1998b).

Ao se considerar a natureza observacional da avaliação da dor nos animais, é imprescindível que os instrumentos de avaliação disponibilizados apresentem validade e confiabilidade estabelecidas (Morton *et al.*, 2005), ou seja, que produzam resultados acurados e consistentes quando utilizados por diferentes observadores. A confiabilidade avalia se o instrumento está medindo algo de uma forma reproduzível e pode ser determinada com base em uma única administração do teste (consistência interna), ou examinando-se a reprodutibilidade de uma medida quando avaliada por diferentes observadores ou

quando administrada em diferentes ocasiões (estabilidade) (Streiner e Norman, 2008).

A consistência interna avalia a inter-relação entre os diferentes itens da escala e assegura que os escores dos vários itens que compõem o instrumento possam ser somados, com o intuito de produzir um escore total relacionado à avaliação global da intensidade da dor. Entretanto, a avaliação da consistência interna isoladamente não apresenta uma base suficiente sobre a qual se pode fundamentar a confiabilidade de um instrumento, é necessário que esta também seja avaliada em termos de reprodutibilidade e estabilidade (Streiner e Norman, 2008). É essencial testar a confiabilidade relacionada à reprodutibilidade e à estabilidade quando a observação é o método de coleta dos dados (Beyer e Wells, 1989).

Adicionalmente, deve-se considerar que a utilização de instrumentos de avaliação de dor somente repercutirá em uma melhora na terapêutica analgésica quando os resultados obtidos pela aplicação destes forem interpretáveis, além de acurados e confiáveis. Em outras palavras, é imprescindível que exista um critério que direcione a instituição da terapia analgésica e avalie a eficácia ou não desta intervenção (Hodgins, 2002). A utilidade clínica de um instrumento de avaliação é realçada quando os escores de dor são informativos da necessidade ou não de intervenção analgésica (Reid *et al.*, 2007). Portanto, além da determinação da validade, da responsividade e da confiabilidade de um instrumento de avaliação de dor, é essencial definir as inferências que podem ser estabelecidas considerando-se os escores de dor obtidos com a aplicação da escala.

O objetivo deste estudo foi testar a confiabilidade de uma escala multidimensional para avaliação de dor aguda pós-operatória em gatos, mediante a verificação do grau de concordância entre diferentes observadores (confiabilidade interobservadores) e do grau de concordância das avaliações feitas pelo mesmo observador em diferentes ocasiões (confiabilidade intraobservadores). Adicionalmente, buscou-se estabelecer uma pontuação mínima relacionada à necessidade de intervenção analgésica.

## MATERIAL E MÉTODOS

Uma escala multidimensional para avaliação de dor aguda pós-operatória em gatas submetidas à ovariário-histerectomia foi desenvolvida e aperfeiçoada em estudos prévios (Brondani *et al.*, 2009; Brondani *et al.*, 2011), incluindo o teste de sua confiabilidade em termos de consistência interna (Brondani *et al.*, 2011). A escala apresenta 10 itens, distribuídos em quatro dimensões ou subescalas: 1) “alteração psicomotora”, composta pelos itens postura, conforto, atividade, atitude e miscelânea de comportamentos; 2) “proteção da área dolorosa”, que engloba os itens reação à palpação da ferida cirúrgica e reação à palpação do abdome/flanco; 3) “variáveis fisiológicas”, que compreendem os itens pressão arterial e apetite e 4) “expressão vocal da dor”, representada pelo item vocalização.

Cada item da escala é constituído por quatro níveis descritivos, que são escoreados categoricamente em 0, 1, 2 e 3, em que “0” representa normalidade ou não alteração e “3” a mais pronunciada alteração. O escore total da escala pode ser calculado com base no somatório dos escores dos itens e varia de 0 (ausência de dor arbitrária) até 30 pontos (dor máxima). Anteriormente ao início dos testes de confiabilidade deste estudo, a escala foi submetida a uma minuciosa análise do conteúdo, com a inclusão de explicações detalhadas sobre cada um dos níveis descritivos presentes, para limitar ao máximo a subjetividade nas avaliações.

Após a aprovação pela Comissão de Ética em Experimentação Animal da FVMZ-UNESP Campus de Botucatu (Protocolo nº 20/2008) e o consentimento dos proprietários, 30 gatas consideradas saudáveis baseando-se nos exames físico e laboratorial, com peso médio de  $2,8 \pm 0,5$  kg e idade de  $14 \pm 5$  meses, foram submetidas à ovariário-histectomia e filmadas durante o período perioperatório, para posterior avaliação dos escores de dor por diferentes observadores. Para a realização do procedimento cirúrgico, a anestesia foi induzida com propofol (8mg/kg) IV e mantida com isoflurano (1,5 - 2,0 CAM) em 100% de oxigênio, administrados por meio de um sistema sem reinalação. Logo após a indução anestésica, mas antes do início da cirurgia, foram aplicados, pela via IV,

cefazolina (30mg/kg), como antibiótico profilático, e fentanil (0,002mg/kg), como analgésico transoperatório.

Durante o período perioperatório, as gatas foram filmadas em quatro momentos predeterminados: M1) “pré-operatório” (de 18 a 24 horas antes da cirurgia); M2) “pós-operatório antes da analgesia de resgate” (de 30 minutos a 1 hora após o término da cirurgia); M3) “pós-operatório após o uso de analgesia de resgate” (4 horas após a administração de analgésico de resgate); e M4) “pós-operatório 24 horas após o término da cirurgia”. Após a realização das filmagens no M2, todos os animais receberam a associação de analgésicos: morfina (0,2mg/kg) IM, cetoprofeno (2mg/kg) SC e dipirona (25mg/kg) IV.

Para cada momento de avaliação especificado anteriormente foi gerado um videoclipe de cinco minutos de duração. Considerando-se as 30 gatas que participaram do estudo, o total de 10 horas de gravação foi distribuído em DVDs para posterior análise por observadores. Com o objetivo de tornar os observadores “cegos”, ou seja, não cientes quanto aos diferentes momentos de avaliação, foi utilizado um curativo micropore no local da ferida cirúrgica desde o período pré-operatório, e durante a organização dos DVDs a sequência cronológica dos vídeos foi alterada.

Para verificar a confiabilidade inter e intraobservadores do instrumento, quatro profissionais formados em medicina veterinária (três anestesiólogos e um cirurgião) foram contatados para a análise dos vídeos. Para a avaliação das filmagens, os observadores receberam uma versão modificada do instrumento, com o intuito de prevenir a tendenciosidade no preenchimento deste. Nesta escala modificada, a ordem dos níveis descritivos foi alterada e os escores de cada nível foram omitidos.

A confiabilidade interobservadores foi avaliada pelo grau de concordância dos escores de dor registrados pelos observadores “cegos”, para cada item da escala, para o escore total e para o escore parcial das subescala 1 “alteração psicomotora” e 2 “proteção da área dolorosa”. A subescala 3 “variáveis fisiológicas” não foi testada, devido à impossibilidade de avaliação da pressão arterial por meio de análise de vídeos. Em relação à subescala 4 “expressão vocal da

dor”, visto que ela é composta apenas pelo item vocalização, o resultado isolado deste item já expressa a resposta desta subescala.

A concordância interobservadores foi estimada pelo coeficiente de correlação intraclasse (CCI), com 95% de intervalo de confiança (IC) (Bartko, 1966). Foi escolhido o modelo de dois fatores, com o critério de concordância absoluta. Os resultados obtidos foram interpretados segundo a classificação de Altman (1991): 0,81 – 1,0 muito boa; 0,61 – 0,80 boa; 0,41 – 0,6 moderada; 0,21 – 0,4 razoável; <0,2 pobre. Para a realização das análises estatísticas, foram considerados os quatro momentos de avaliação agrupados e somente o M2 isolado.

Além dos observadores “cegos” mencionados anteriormente, o pesquisador que desenvolveu a escala e que apresenta experiência clínica na utilização desta também analisou os vídeos, sendo considerado o padrão-ouro. Desta forma, além da concordância entre os observadores “cegos”, avaliou-se também a concordância entre os observadores “cegos” adicionando-se o padrão-ouro, com o objetivo de observar se haveria uma melhora ou piora na confiabilidade com a inclusão das avaliações do padrão-ouro. Uma diminuição na concordância indicaria que os observadores “cegos” concordam entre si, mas discordam das avaliações consideradas ideais.

Para estabelecer a confiabilidade intraobservadores, os observadores “cegos” foram contatados novamente dois meses após a primeira avaliação e instruídos a examinar novamente os vídeos. Entretanto, para que as informações prévias não influenciassem na análise, foram entregues novos DVDs com uma reorganização na ordem dos animais e nos momentos de avaliação. A concordância intraobservadores foi avaliada para cada item da escala utilizando-se o CCI. Foram considerados os momentos de avaliação agrupados e o M2 isoladamente. Os resultados obtidos foram interpretados segundo a classificação de Altman (1991), já descrita anteriormente.

Com o intuito de se identificar uma pontuação mínima obtida com o uso da escala que estivesse relacionada à necessidade de administração de terapia analgésica, os observadores foram instruídos que, ao término da análise de cada vídeo, anteriormente ao preenchimento da escala,

identificassem a necessidade ou não de analgesia de resgate, baseando-se na experiência clínica para o julgamento. Para tal, foi incluída na versão modificada da escala apresentada aos observadores a seguinte pergunta: “De acordo com a sua experiência clínica você administraria analgesia de resgate?”

A pontuação relacionada à necessidade de intervenção analgésica foi determinada com base na análise descritiva dos dados e no estudo da curva ROC (*Receiver Operating Characteristic*). Para a análise descritiva, as pontuações de dor obtidas foram classificadas em dor leve (0 – 8), dor moderada (9 – 21) e dor intensa (22 – 30), considerando-se a informação difundida na literatura de que escores de 0 – 29% estariam relacionados com dor leve, de 30 – 70% relacionados com dor moderada e de 71 – 100% com dor intensa (Park et al., 2009).

A curva ROC é uma representação gráfica da relação entre “verdadeiros positivos” (sensibilidade) versus “falsos positivos” (1-especificidade) para todos os pontos de cortes possíveis (Martinez et al., 2003). O valor da área sob a curva (ASC) indica a capacidade discriminativa do teste, ou seja, avalia o quanto acurado é a habilidade do instrumento em classificar corretamente indivíduos com ou sem dor. Valores de ASC  $\geq 7$  são considerados adequados (Deyo et al., 1991). Tomando-se por base a análise da curva, é possível selecionar o ponto de corte ótimo para realizar tal discriminação, no qual mensurações menores ou iguais ao valor de corte são classificadas como indicativas de “ausência de dor” e, analogamente, escores acima do ponto de corte são classificadas como “presença de dor”.

## RESULTADOS E DISCUSSÃO

A confiabilidade interobservadores averigua a habilidade do instrumento em obter resultados similares quando utilizado por diferentes examinadores, sendo usualmente calculada pelo CCI, que varia de “0” (falta de concordância) até “1,0” (concordância perfeita) (Aissaoui et al., 2005). Entretanto, embora o CCI seja o método estatístico mais apropriado para a análise de concordância, ele apresenta limitações quando a amostra é homogênea (Deyo et al., 1991). Tal restrição foi observada neste estudo, o que exigiu a análise dos quatro momentos de avaliação

agrupados. Quando os momentos foram avaliados separadamente, somente o M2 apresentou variabilidade suficiente para permitir o cálculo do CCI. A avaliação deste momento isolado era primordial para adequada averiguação da confiabilidade do instrumento, visto que, no M2, os observadores devem concordar principalmente baseando-se na identificação de comportamentos relacionados com dor (categorias 1, 2 e 3 da escala) e não somente com ausência de dor (categoria 0).

Diante do exposto acima, ao se considerarem os momentos de avaliação agrupados, a

concordância dos observadores “cegos” foi boa a muito boa: 0,98 para o escore total, 0,97 para a subescala 1 (alteração psicomotora), 0,93 para a subescala 2 (proteção da área dolorosa), e para cada um dos itens da escala variou de 0,79 (vocalização) até 0,95 (miscelânea de comportamentos). Quando o M2 foi avaliado isoladamente, a concordância dos observadores “cegos” foi moderada a muito boa: 0,88 para o escore total, 0,67 para a subescala 1 (alteração psicomotora), 0,86 para a subescala 2 (proteção da área dolorosa), e para cada um dos itens da escala variou de 0,51 (atividade) até 0,92 (apetite) (Tab. 1).

Tabela 1. Concordância entre as avaliações dos observadores “cegos”, incluindo ou não o padrão-ouro, em todos os momentos de avaliação agrupados MT (pré-operatório e pós-operatório: antes e após analgesia de resgate e 24 horas após o término da cirurgia) e no M2 isolado (pós-operatório antes da analgesia de resgate).

Itens da escala	Observadores “cegos”		Observadores “cegos” e padrão-ouro	
	MT	M2	MT	M2
Postura	<b>0,93</b> (0,91–0,98)	<b>0,79</b> (0,68–0,88)	<b>0,94</b> (0,93–0,96)	<b>0,83</b> (0,73–0,90)
Conforto	<b>0,93</b> (0,90–0,97)	<b>0,61</b> (0,44–0,76)	<b>0,94</b> (0,92–0,95)	<b>0,68</b> (0,54–0,80)
Atividade	<b>0,89</b> (0,85–0,96)	<b>0,51</b> (0,33–0,69)	<b>0,90</b> (0,88–0,93)	<b>0,59</b> (0,43–0,74)
Atitude	<b>0,94</b> (0,92–0,95)	<b>0,58</b> (0,41–0,74)	<b>0,95</b> (0,93–0,96)	<b>0,66</b> (0,51–0,79)
Miscelânea de comportamentos	<b>0,95</b> (0,93–0,96)	<b>0,53</b> (0,35–0,70)	<b>0,96</b> (0,95–0,97)	<b>0,61</b> (0,46–0,76)
Palpação da ferida cirúrgica	<b>0,89</b> (0,86–0,92)	<b>0,82</b> (0,70–0,90)	<b>0,90</b> (0,88–0,93)	<b>0,84</b> (0,75–0,91)
Palpação do abdome	<b>0,89</b> (0,86–0,92)	<b>0,77</b> (0,64–0,87)	<b>0,91</b> (0,89–0,93)	<b>0,81</b> (0,70–0,89)
Apetite	<b>0,89</b> (0,85–0,92)	<b>0,92</b> (0,87–0,96)	<b>0,91</b> (0,88–0,93)	<b>0,93</b> (0,89–0,96)
Vocalização	<b>0,79</b> (0,73–0,84)	<b>0,74</b> (0,60–0,85)	<b>0,83</b> (0,79–0,87)	<b>0,79</b> (0,69–0,88)
<b>Escore Total</b>	<b>0,98</b> (0,97–0,99)	<b>0,88</b> (0,80–0,93)	<b>0,98</b> (0,97–0,99)	<b>0,90</b> (0,84–0,95)
<b>Subescala 1</b>	<b>0,97</b> (0,96–0,98)	<b>0,67</b> (0,52–0,81)	<b>0,97</b> (0,96–0,98)	<b>0,73</b> (0,60–0,84)
<b>Subescala 2</b>	<b>0,93</b> (0,91–0,95)	<b>0,86</b> (0,76–0,92)	<b>0,94</b> (0,93–0,96)	<b>0,88</b> (0,81–0,93)

Coefficiente de correlação intraclassa (95% de intervalo de confiança). Interpretação dos valores:

0,81 – 1,0 muito boa; 0,61 – 0,80 boa; 0,41 – 0,6 moderada; 0,21 – 0,4 razoável; < 0,2 pobre.

Subescala 1 “alteração psicomotora” engloba os itens postura, conforto, atividade, atitude e miscelânea de comportamentos. Subescala 2 “proteção da área dolorosa” engloba os itens reação à palpação da ferida cirúrgica e reação à palpação do abdome/flanco.

Com base nos dados da Tab. 1, observa-se que, quando as avaliações do padrão-ouro são adicionadas às análises dos observadores “cegos”, o grau de concordância aumenta, tanto na avaliação dos momentos agrupados quanto no M2 isolado. Tal resultado demonstra que, além de os avaliadores apresentarem satisfatória concordância entre si, as suas análises também estão de acordo com as avaliações acuradas do padrão-ouro.

Na determinação da confiabilidade intraobservadores, cada examinador precisa

concordar com a sua própria avaliação em tempos diferentes, sendo necessário estipular um intervalo adequado entre as avaliações, para que as respostas da segunda análise não sejam influenciadas pela memória da primeira. O intervalo de dois meses, instituído neste estudo, tem sido utilizado nas pesquisas de validação de escalas de dor em pediatria (Schultz *et al.*, 1999; Jong *et al.*, 2005). Adicionalmente, os observadores não foram informados sobre o procedimento de reanálise antecipadamente, para evitar qualquer tipo de tendenciosidade. Ao se considerarem os momentos de avaliação

agrupados, a confiabilidade intraobservadores foi muito boa para todos os itens da escala e variou de 0,84 (vocalização) até 0,99 (miscelânea de comportamentos) (Tab. 2). Quando o M2 foi avaliado isoladamente, a confiabilidade

intraobservadores variou de moderada a muito boa, com os itens apetite e atitude apresentando o melhor e o pior desempenho, respectivamente (Tab. 3).

Tabela 2. Confiabilidade intraobservadores para cada item da escala, abrangendo todos os momentos de avaliação agrupados (pré-operatório e pós-operatório: antes e após analgesia de resgate e 24 horas após o término da cirurgia).

Itens da escala	Observadores "cegos"			
	Anestesiologista 1	Anestesiologista 2	Anestesiologista 3	Cirurgião
Postura	<b>0,96</b> (0,94–0,97)	<b>0,98</b> (0,97–0,99)	<b>0,98</b> (0,98–0,99)	<b>0,96</b> (0,95–0,97)
Conforto	<b>0,96</b> (0,94–0,97)	<b>0,97</b> (0,95–0,98)	<b>0,98</b> (0,97–0,99)	<b>0,97</b> (0,96–0,98)
Atividade	<b>0,90</b> (0,87–0,93)	<b>0,96</b> (0,95–0,97)	<b>0,97</b> (0,96–0,98)	<b>0,93</b> (0,91–0,95)
Atitude	<b>0,97</b> (0,96–0,98)	<b>0,95</b> (0,93–0,96)	<b>0,98</b> (0,97–0,99)	<b>0,94</b> (0,92–0,96)
Miscelânea de comportamentos	<b>0,97</b> (0,96–0,98)	<b>0,99</b> (0,98–0,99)	<b>0,97</b> (0,95–0,98)	<b>0,98</b> (0,97–0,99)
Palpação da ferida cirúrgica	<b>0,92</b> (0,89–0,94)	<b>0,96</b> (0,94–0,97)	<b>0,95</b> (0,93–0,96)	<b>0,95</b> (0,93–0,97)
Palpação abdome	<b>0,96</b> (0,95–0,97)	<b>0,91</b> (0,81–0,95)	<b>0,95</b> (0,93–0,97)	<b>0,94</b> (0,91–0,95)
Apetite	<b>0,90</b> (0,86–0,93)	<b>0,95</b> (0,94–0,97)	<b>0,98</b> (0,97–0,98)	<b>0,92</b> (0,89–0,95)
Vocalização	<b>0,90</b> (0,86–0,93)	<b>0,97</b> (0,96–0,98)	<b>0,94</b> (0,91–0,95)	<b>0,84</b> (0,77–0,88)

Coefficiente de correlação intraclassa (95% de intervalo de confiança). Interpretação dos valores: 0,81 – 1,0 muito boa; 0,61 – 0,80 boa; 0,41 – 0,6 moderada; 0,21 – 0,4 razoável; < 0,2 pobre.

Tabela 3. Confiabilidade intraobservadores para cada item da escala, no M2 isolado (pós-operatório antes da analgesia de resgate).

Itens da Escala	Observadores "cegos"			
	Anestesiologista 1	Anestesiologista 2	Anestesiologista 3	Cirurgião
Postura	<b>0,74</b> (0,51–0,87)	<b>0,83</b> (0,68–0,92)	<b>0,86</b> (0,73–0,93)	<b>0,95</b> (0,89–0,97)
Conforto	<b>0,78</b> (0,59–0,89)	<b>0,81</b> (0,65–0,91)	<b>0,89</b> (0,79–0,95)	<b>0,87</b> (0,73–0,94)
Atividade	<b>0,61</b> (0,32–0,80)	<b>1,00</b>	<b>0,90</b> (0,81–0,95)	<b>0,78</b> (0,59–0,89)
Atitude	<b>0,81</b> (0,64–0,90)	<b>0,58</b> (0,27–0,77)	<b>0,81</b> (0,65–0,91)	<b>0,56</b> (0,23–0,77)
Miscelânea de comportamentos	<b>0,69</b> (0,44–0,84)	<b>0,80</b> (0,60–0,90)	<b>0,66</b> (0,40–0,82)	<b>0,81</b> (0,62–0,91)
Palpação da ferida cirúrgica	<b>0,98</b> (0,96–0,99)	<b>0,93</b> (0,86–0,97)	<b>0,92</b> (0,83–0,96)	<b>0,88</b> (0,75–0,94)
Palpação abdome	<b>0,92</b> (0,84–0,96)	<b>0,91</b> (0,81–0,95)	<b>0,91</b> (0,82–0,96)	<b>0,86</b> (0,77–0,93)
Apetite	<b>0,98</b> (0,96–0,99)	<b>0,94</b> (0,88–0,97)	<b>0,96</b> (0,92–0,98)	<b>0,93</b> (0,86–0,97)
Vocalização	<b>0,87</b> (0,75–0,94)	<b>0,97</b> (0,94–0,99)	<b>0,89</b> (0,79–0,95)	<b>0,80</b> (0,63–0,90)

Coefficiente de correlação intraclassa (95% de intervalo de confiança). Interpretação dos valores: 0,81 – 1,0 muito boa; 0,61 – 0,80 boa; 0,41 – 0,6 moderada; 0,21 – 0,4 razoável; < 0,2 pobre.

Considerando-se os resultados supracitados, pode-se observar que a escala apresentou adequada confiabilidade quando utilizada por veterinários formados e com experiência clínica. Atribuiu-se o satisfatório desempenho da escala em termos de reprodutibilidade ao minucioso detalhamento dos comportamentos a serem observados e, conseqüentemente, à redução na subjetividade da análise. Escalas consideradas

extremamente subjetivas, como a escala analógica visual, a escala numérica e a escala descritiva simples, têm apresentado resultados insatisfatórios de confiabilidade interobservadores na avaliação da dor aguda em cães (Holton *et al.*, 1998b).

No momento de maior desafio em termos de avaliação (M2), os itens atividade, miscelânea

de comportamentos e atitude apresentaram os valores mais baixos de confiabilidade inter e intraobservadores. Justifica-se a menor confiabilidade destes itens pela dependência da correta identificação de comportamentos bem específicos (atitude e miscelânea de comportamentos) e pela necessidade de avaliação dinâmica (atitude e atividade). Adicionalmente, a grande variabilidade observada no CCI intraobservadores destes itens sugere que as deficiências de confiabilidade possam ser supridas com um treinamento prévio para o reconhecimento dos comportamentos específicos descritos no instrumento. A criação de um *website* com a disponibilização de vídeos para o treinamento está em desenvolvimento.

A importância da interpretação dos escores de dor obtidos mediante aplicação de instrumentos de avaliação é uma preocupação que foi despertada recentemente na medicina humana (Hodgins *et al.*, 2002) e veterinária (Reid *et al.*, 2007). Com o intuito de estabelecer uma pontuação mínima que indicasse a necessidade de intervenção analgésica, o presente estudo utilizou, além da análise descritiva, uma técnica estatística inovadora na medicina veterinária na área de validação de escalas, a análise da curva ROC. Esta curva tem sido utilizada na medicina para determinar a habilidade de um teste em discriminar grupos, escolher um ponto de corte ótimo e comparar o desempenho de dois ou mais testes (Streiner e Cairney, 2007).

Em relação à análise descritiva, ao se considerarem as avaliações dos observadores “cegos”, no M1 e M3, os escores de dor variaram dentro da faixa de dor leve (M1: zero até oito pontos; M3: zero até seis pontos), e a analgesia de resgate não foi recomendada. No M2, os escores de dor de 17 a 19 gatas variaram dentro da faixa de dor moderada (de 12 até 21 pontos), e de 11 a 13 animais dentro da faixa de dor intensa (de 22 até 29 pontos), sendo recomendada a analgesia de resgate. No M4, os escores de dor de 27 a 28 gatas variaram dentro da faixa de dor leve (zero até oito pontos), e de duas a três gatas

dentro da faixa de dor moderada (nove até 15 pontos). Nos animais com pontuação de zero até sete, não foi recomendada analgesia de resgate. Por outro lado, nas gatas com escores de 10 até 15, o uso foi recomendado. Entre as gatas com pontuação oito, foi sugerida a administração de analgésicos em dois animais, e não recomendada em três. Nas gatas com pontuação nove, foi recomendado o uso de analgésico em um animal, e não recomendado em outro.

Com base na análise apresentada acima, observa-se que houve um consenso entre os observadores “cegos” quanto à recomendação de analgésicos em escores  $\geq 10$  pontos e quanto à não recomendação de analgesia de resgate em escores  $\leq 7$  pontos. No intervalo de oito a nove pontos (exatamente o valor de transição entre dor leve e moderada), não houve um consenso. Visto que a análise descritiva não foi capaz de identificar o ponto de corte ótimo, foi incorporada ao estudo a análise da curva ROC. Considerando-se a análise desta curva, diferentes pontos de corte foram sugeridos, sendo indicado o ponto ótimo representado pelo valor em que a sensibilidade e a especificidade são simultaneamente maiores.

Foi identificado um ponto de corte ótimo  $>7$ , com uma sensibilidade de 100% (95% IC: 97,8 – 100%) e especificidade de 98,85% (95% IC: 97,3 – 99,6 %) (Fig. 1 e 2). Embora o critério utilizado para definir o ponto ótimo de corte (maior otimização da sensibilidade em função da especificidade) seja amplamente recomendado pela literatura, o pesquisador deve analisar criticamente o resultado, porque, em algumas situações, pode ser interessante privilegiar uma ou outra (Martinez *et al.*, 2003). Adicionalmente, a alta área sob a curva observada ASC = 1,0 (95% IC: 0,993 – 1,000;  $p < 0,001$ ) indica que o instrumento apresentou uma habilidade discriminatória excelente. Desta forma, ao se considerar a análise descritiva e da curva ROC, sugere-se o uso de analgesia de resgate a partir de oito pontos (26,6% do escore total), sendo fortemente recomendada a partir de 10 pontos (33,3% do escore total).

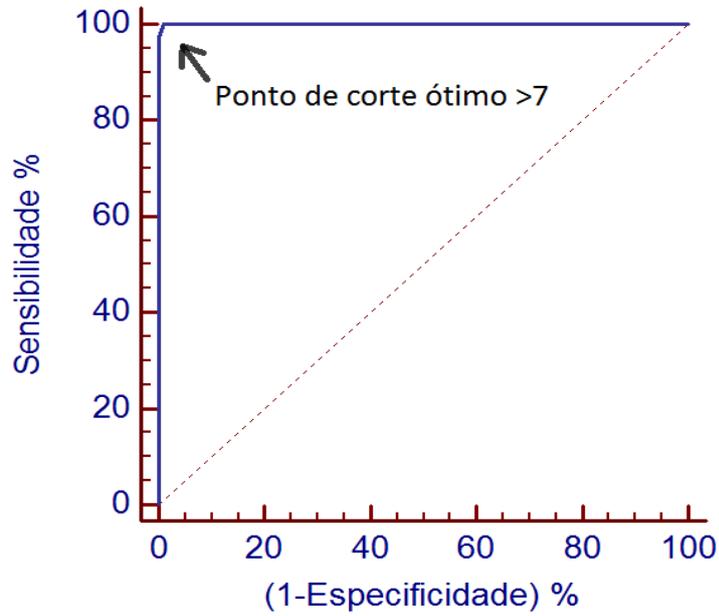


Figura 1. Curva ROC (*Receiver Operating Characteristic*) para a escala multidimensional de avaliação de dor pós-operatória em gatos: ponto de corte ótimo >7, com sensibilidade de 100% e especificidade 98,85% e área sob a curva de 1,0.

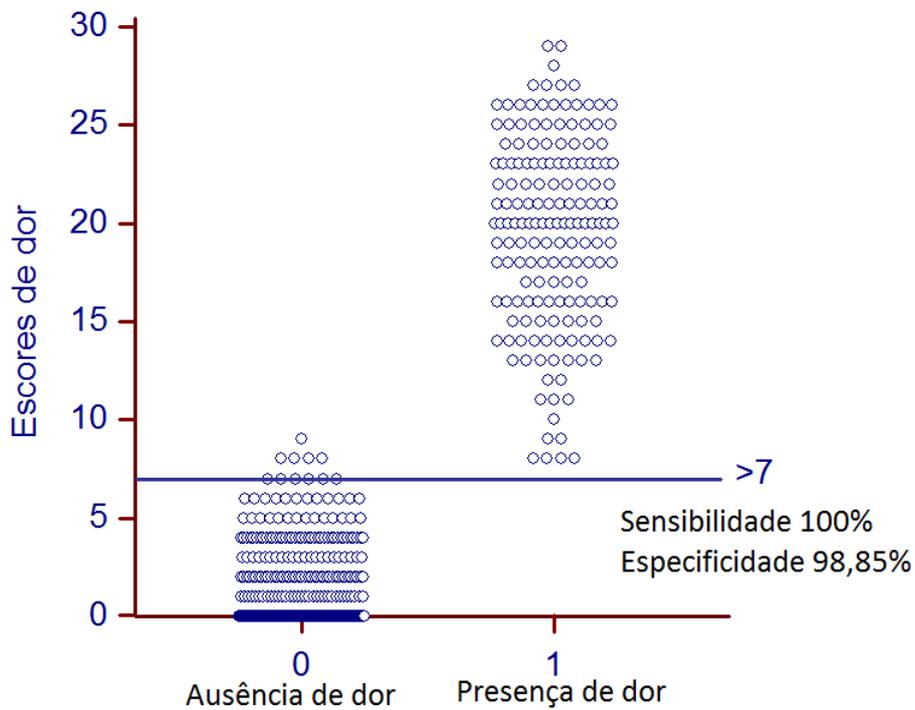


Figura 2. Diagrama ilustrando o ponto de corte ótimo identificado pela análise da curva ROC (*Receiver Operating Characteristic*), considerando as avaliações dos observadores “cegos” e padrão-ouro quanto à suficiência ou não de analgesia.

A disponibilidade de escalas de avaliação com um escore informativo sobre a necessidade de uso de analgésicos auxilia o profissional na decisão clínica sobre a terapêutica analgésica (Reid *et al.*, 2007). Adicionalmente, também apresenta um importante impacto nos estudos de pesquisa, visto que a analgesia de resgate tem sido utilizada juntamente com os escores de dor para a avaliação da eficácia analgésica de fármacos na espécie felina.

Na falta de critério para guiar o uso de analgesia de resgate, os pesquisadores têm adotado o valor empírico de  $\geq 50\%$  da pontuação total, seja por meio da escala analógica visual ou de contagem variável (Al-Gizawiy e Rudé, 2004; Gassel *et al.*, 2005; Tobias *et al.*, 2006). Entretanto, foi descrito que alguns animais com pontuações entre 38 e 48% da total apresentavam sinais clínicos de desconforto, mas não se beneficiaram do uso de analgesia de resgate (Al-Gizawiy e Rudé, 2004). Tal observação demonstra a importância de escalas de avaliação com critério de instituição de terapia analgésica estabelecido, para se evitar o sofrimento desnecessário dos animais nas pesquisas e conclusões imprecisas por parte dos pesquisadores.

Embora tenha sido definido um ponto ótimo de intervenção analgésica para a escala proposta como  $>7$ , não se deve negar o uso de analgésicos em animais com pontuação  $\leq 7$ , se o julgamento clínico do profissional indicar a necessidade de uso. A determinação dos escores de intervenção tem o objetivo de auxiliar na identificação da necessidade de instituição de terapia analgésica, mas não substitui a autonomia e a experiência clínica do profissional.

### CONCLUSÕES

Os resultados alcançados suportam a confiabilidade em termos de reprodutibilidade e estabilidade de uma escala multidimensional para avaliação de dor aguda pós-operatória em gatas submetidas à ovariectomia, quando utilizada por profissionais formados em medicina veterinária e com experiência clínica (anestesiologistas e cirurgiões). Pelo escore total obtido com o uso da escala, a instituição de terapia analgésica deve ser considerada em pontuações a partir de oito, sendo fortemente recomendada em escore igual ou maior que 10 pontos.

### AGRADECIMENTOS

FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo.

### REFERÊNCIAS

- AISSAOUI, Y.; ZEGGWAGH, A.A.; ZEKRAOUI, A. *et al.* Validation of a behavioral pain scale in critically ill, sedated, and mechanically ventilated patients. *Anesth. Analg.*, v.101, p.1470-1476, 2005.
- AL-GIZAWIY, M.M.; RUDÉ, E.P. Comparison of preoperative carprofen and postoperative butorphanol as postsurgical analgesics in cats undergoing ovariohysterectomy. *Vet. Anaesth. Analg.*, v.31, p.164-174, 2004.
- ALTMAN, D.G. (Ed). *Practical statistics for medical research*. London: Chapman & Hall, 1991. p.404-408.
- ANIL, S.S.; ANIL, L.; DEEN, J. Challenges of pain assessment in domestic animals. *J. Am. Vet. Med. Assoc.*, v.220, p.313-319, 2002.
- BAEYER, V.C.L.; SPAGRUD, L.J. Systematic review of observational (behavioral) measures of pain for children and adolescents aged 3 to 18 years. *Pain*, v.127, p.140-150, 2007.
- BARTKO, J.J. The intraclass correlation coefficient as a measure of reliability. *Psychol. Rep.*, v.19, p.3-11, 1966.
- BEYER, J.E.; WELLS, N. The assessment of pain in children. *Pediatr. Clin. North America*, v.36, p.837-854, 1989.
- BRONDANI, J.T.; LUNA, S.P.L.; BEIER, S.L. *et al.* Analgesic efficacy of perioperative use of vedaprofen, tramadol or their combination in cats undergoing ovariohysterectomy. *J. Feline Med. Surg.*, v.11, p.420-429, 2009.
- BRONDANI, J.T.; LUNA, S.P.L.; PADOVANI, C.R. Refinement and initial validation of a multidimensional composite scale for use in assessing acute postoperative pain in cats. *Am. J. Vet. Res.*, v.72, p.174-183, 2011.
- CAMBRIDGE, A.J.; TOBIAS, K.M.; NEWBERRY, R.C. *et al.* Subjective and objective measurements of postoperative pain in cats. *J. Am. Vet. Med. Assoc.*, v.217, p.685-690, 2000.

- DEYO, R.A.; DIEHR, P.; PATRICK, D.L. Reproducibility and responsiveness of a health status measures. *Control. Clin. Trials*, v.12, p.142-158, 1991.
- GASSEL, A.D.; TOBIAS, K.M.; EGGER, C.M. et al. Comparison of oral and subcutaneous administration of buprenorphine and meloxicam for preemptive analgesia in cats undergoing ovariohysterectomy. *J. Am. Vet. Med. Assoc.*, v.227, p.1937-1944, 2005.
- HODGINS, M.J. Interpreting the meaning of pain severity scores. *Pain Res. Manage.*, v.7, p.192-198, 2002.
- HOLTON, L.L.; SCOTT, E.M.; NOLAN, A.M. et al. Relationship between physiological factors and clinical pain in dogs scored using a numerical rating scale. *J. Small Anim. Pract.*, v.39, p.469-474, 1998a.
- HOLTON, L.L.; SCOTT, E.M.; NOLAN, A.M. et al. Comparison of three methods used for assessment of pain in dogs. *J. Am. Vet. Med. Assoc.*, v.212, p.61-66, 1998b.
- HOLTON, L.; REID, J.; SCOTT, M. et al. Development of a behaviour-based scale to measure acute pain in dogs. *Vet. Rec.*, v.28, p.525-531, 2001.
- JONG, A.E.E.; BREMER, M.; SCHOUTEN, M. et al. Reliability and validity of the pain observation scale for young children and the visual analogue scale in children with burns. *Burns*, v.31, p.198-204, 2005.
- MARTINEZ, E.Z.; LOUZADA-NETO, F.; PEREIRA, B.B. A curva ROC para testes diagnósticos. *Cad. Saúde Colet.*, v.11, p.7-31, 2003.
- MORTON, C.M.; REID, J.; SCOTT, M.E. et al. 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.*, v.66, p.2154-2166, 2005.
- PARK, J.; CHO, B.; PAEK, Y. et al. Development of a pain assessment tool for the older adults in Korea: the validity and reliability of a Korean version of the geriatric pain measure. *Arch. Gerontol. Geriatr.*, v.49, p.199-203, 2009.
- REID, J.; NOLAN, A.M.; HUGHES, J.M.L. et al. Development of the short-form Glasgow composite measure pain scale (CMPS-SF) and derivation of an analgesic intervention score. *Anim. Welf.*, v.16, p.97-104, 2007.
- SCHULTZ, A.A.; MURPHY, E.; MORTON, J. et al. Preverbal, early verbal pediatric pain scale (PEPPS): development and early psychometric testing. *J. Pediatr. Nurs.*, v.14, p.19-27, 1999.
- SMITH, J.D.; ALLEN, S.W.; QUANDT, J.E. et al. Indicators of postoperative pain in cats and correlation with clinical criteria. *Am. J. Vet. Res.*, v.57, p.1674-1678, 1996.
- STREINER, D.L.; CAIRNEY, J. What's under the ROC? An introduction to receiver operating characteristics curves. *Can. J. Psychiatry*, v.52, p.121-128, 2007.
- STREINER, D.L.; NORMAN, G.R. *Health measurement scales: a practical guide to their development and use*. 4nd ed. New York: Oxford University Press, 2008. 431p.
- TOBIAS, K.M.; HARVEY, R.C.; BYARLAY, J.M. A comparison of four methods of analgesia in cats following ovariohysterectomy. *Vet. Anaesth. Analg.*, v.33, p.390-398, 2006.