Université de Montréal

Construct validity, responsiveness and reliability of the Feline Grimace Scale© in kittens

Par

Alice J. Cheng

Département de sciences cliniques, Faculté de médecine vétérinaire

Mémoire présenté à la Faculté de médecine vétérinaire en vue de l'obtention du grade de Maîtrise es sciences (M. Sc.) en sciences vétérinaires, option sciences cliniques

Décembre 2023

Université de Montréal

Département de sciences cliniques, Faculté de médecine vétérinaire

Ce mémoire intitulé

Construct validity, responsiveness and reliability of the Feline Grimace Scale© in kittens

Présenté par

Alice J Cheng

A été évalué par un jury composé des personnes suivantes

Dr Mouhamadou Diaw
Président-rapporteur

Dr Paulo V Steagall
Directeur de recherche

Dre Beatriz P Monteiro
Codirectrice

Dre Hélène LM Ruel
Membre du jury

Résumé

Cette étude prospective, randomisée et à l'aveugle a évalué la validité, la réactivité et fiabilité de l'échelle de grimaces félines (Feline Grimace Scale; FGS) chez les chatons.

Trente-six jeunes chattes en santé (âgées de 10 semaines à 6 mois) étaient filmées avant puis 1 et 2 h après ovariohystérectomie. La procédure a été effectuée avec un protocole d'anesthésie injectable (sans opioïde), avec ou sans analgésie multimodale. Les chatons en douleur étaient également filmés avant et 1h après avoir reçu une analgésie de secours (buprénorphine 0.02 mg/kg IM). Quatre évaluateurs, aveugles aux conditions expérimentales, ont évalué deux fois à cinq semaines d'intervalle les expressions faciales sur 111 images extraites des vidéos. Les cinq unités d'action (action unit; AU) de la FGS ont été évaluées (position des oreilles, serrage orbital, tension du museau, position des moustaches, position de la tête; avec un score possible de 0 à 2 pour chacune). La validité de construit, la réactivité et la fiabilité inter- et intra-évaluateur de la FGS ont été analysées en utilisant un modèle linéaire avec correction Benjamini–Hochberg, un test Wilcoxon signed-rank et un coefficient de corrélation intra-classes unique (ICC_{single}), respectivement (*P* <0.05).

Les ratios des scores FGS totaux (médiane [étendue interquartile, EI]) étaient augmentés 1 et 2 h après l'ovariohystérectomie (médiane [EI] : 0.30 [0.20–0.40] et 0.30 [0.20–0.40], respectivement) comparativement à la mesure de base (médiane [EI] : 0.10 [0.00–0.30]) (*P* <0.001), et inférieurs après l'analgésie (médiane [EI] : 0.40 [0.20–0.50]) qu'avant son administration (médiane [EI] : 0.20 [0.10–0.38]) (*P* <0.001). Pour la fiabilité inter-évaluateur, les ICC_{single} des ratios des scores FGS totaux étaient 0.68 et compris entre 0.35 et 0.70 pour chaque AU, individuellement. Pour la fiabilité intra-évaluateur, les ICC_{single} des ratios des scores FGS totaux étaient compris entre 0.77–0.91 et 0.55–1.00 pour chaque AU.

La FGS est un outil d'évaluation de la douleur aiguë valide et réactif chez les chatons avec une fiabilité inter-évaluateur modérée et intra-évaluateur bonne à excellente.

Mots-clés : félin, chaton, analgésie, douleur, évaluation de la douleur, expressions faciales, Feline Grimace Scale.

Abstract

This prospective, randomized, blinded study investigated the construct validity, responsiveness and reliability of the Feline Grimace Scale (FGS) in kittens.

Thirty-six healthy female kittens (aged 10 weeks to 6 months) were video recorded before, 1 and 2 h after ovariohysterectomy using an opioid-free injectable anesthetic protocol with or without multimodal analgesia. Painful kittens were additionally filmed before and 1 h after administration of rescue analgesia (buprenorphine 0.02 mg/kg IM). One hundred eleven facial images collected from video recordings were randomly scored by 4 observers, blinded to treatment groups and time points, twice with a 5 weeks interval using the FGS. The five action units (AU) of the FGS were scored (ear position, orbital tightening, muzzle tension, whiskers position and head position; 0–2 each). Construct validity, responsiveness, inter- and intra-rater reliability of the FGS were evaluated using linear models with Benjamini–Hochberg correction, Wilcoxon signed-rank test and single intra-class correlation coefficients (ICC_{single}), respectively (*P* <0.05).

The FGS total ratio scores were higher 1 and 2 h after ovariohysterectomy (median [interquartile range, IQR]: 0.3 [0.20–0.40], and 0.30 [0.20–0.40], respectively) than at baseline (median [QR]: 0.10 [0.00–0.30]) (*P* <0.001), and lower after the administration of rescue analgesia (median [QR]: 0.40 [0.20–0.50]) than before (median [QR]: 0.20 [0.10–0.38]) (*P* <0.001). The inter-rater ICC_{single} was 0.68 for the FGS total ratio scores and 0.35–0.70 for AUs considered individually. The intra-rater ICC_{single} was 0.77–0.91 for the FGS total ratio scores and 0.55–1.00 for AUs considered individually.

The FGS is a valid and responsive acute pain scoring instrument with moderate inter-rater reliability and good to excellent intra-rater reliability in kittens.

Keywords : feline, kitten, analgesia, pain, pain assessment, facial expressions, Feline Grimace Scale.

Table des matières

Liste des tableaux

Liste des figures

Liste des sigles et abréviations

AU: action unit; unité d’action

CG: Control Group

CI: confidence interval

CMPS-F: Glasgow Composite Measure Pain Scale – Feline

FGS: Feline Grimace Scale

$ICC_{average}$: average intraclass correlation coefficients

ICC_{single}: single intraclass correlation coefficients

IM: intramuscular

IQR: interquartile range

MMG: Multimodal Group

n: number of kittens

OVH: ovariohysterectomy

rCMPS-F: revised Glasgow Composite Measure Pain Scale – Feline

SC: subcutaneous

TPR: Temperature, Pulse, Respiration

UFEPS: UNESP-Botucatu Multidimensional Feline Pain Assessment Scale

UFEPS-SF: Unesp-Botucatu Feline Pain Scale – Short Form

To Kitty, Fido, Kiwi, and Matteo

Remerciements

To the veterinarians of the Steagall Laboratory who accompanied me on this project, thank you for this beautiful opportunity. Your love and care for feline well-being will always be reflected in my future work.

To Mom and Dad, thank you for your relentless love and support. Your warm hugs and delicious meals were a huge relief for me. 谢谢你们的爱和支持。

To Jiaqi, thank you for the countless hours talking about the past to our future. You were always there to guide me and support me. No matter what project I start, I know you will always be there to cheer me on, thank you.

To Kitty, my first who taught me everything.

To Fido, my forever big baby.

To Kiwi, my strong and resilient fluffy boy.

To Matteo, my crazy little rat.

Each of you was a motivation to pursue this career and make a difference in the lives of pets.

To Sean, thank you for encouraging me at every step of this project and helping me balance everything in life. I would not have made it this far without you. You are always a loving source of comfort and escape. Time to eat now!

Chapter 1 – Introduction and Literature Review

> As a member of the veterinary medical profession, I solemnly swear that I will use my scientific knowledge and skills for the benefit of society. I will strive to: promote animal health and welfare, prevent and relieve animal suffering, protect the health of the public and the environment, and advance comparative medical knowledge. I will perform my professional duties conscientiously, with dignity, and in keeping with the principles of veterinary medical ethics. I will strive continuously to improve my professional knowledge and competence and to maintain the highest professional and ethical standards for myself and the profession (CVMA, 2004).

This oath made by every Canadian veterinarian since 2004 is the core foundation of veterinary medicine (CVMA, 2004). The primacy of animal welfare and pain alleviation is evident. For veterinarians to effectively manage pain as part of their ethical and medical duty, they must be able to identify, understand, and quantify pain in animals.

Acute Pain

Definition and Physiology

The International Association for the Study of Pain defines pain as "an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage" (Raja et al., 2020). In its most recent 2020 revision, a new note states that "verbal description is only one of several behaviours to express pain; inability to communicate does not negate the possibility that a human or a non-human animal experiences pain" (Raja et al., 2020). This addition is an important advancement for the field of veterinary medicine as it recognizes the possibility of feeling pain despite our patients' inability to verbally express their experience of pain.

Acute pain, also known as adaptive pain, is generally described as appearing subsequent to an injury or disease, and ends when healing has occurred after a relatively short period of time. One of the purposes of acute pain is to change the animal's behaviour to optimize healing (Mathews et al., 2014). For example, some common acute pain presentations in veterinary medicine include sensory and affective experiences resulting from tissue damage due to a wound or elective surgery. On the other end of the spectrum, pain lasting beyond the point of healing, with no clear end or biological function, would be qualified as chronic, maladaptive, or pathological pain (B. Monteiro et al., 2023; Monteiro & Steagall, 2019b).

The conscious perception of pain is the manifestation of a complex, multifactorial, and subjective experience. Broadly, pain begins with a thermal, mechanical and/or chemical stimulus activating nociceptors distributed across the skin, periosteum, joints, muscles, and viscera. These peripheral receptors are free nerve endings from, primarily, two afferents nervous fibers: myelinated Aδ-type fibers and unmyelinated C-type fibers. Aδ-type fibers and C-type fibers trigger, respectively, a rapid "stabbing" and a slow "burning" sensation. Both of these are part of the acute pain response. Other fibers typically not associated with nociception, like myelinated Aβ-type fibers, can be solicited following changes caused by tissue inflammation, which can lead to allodynia (Jensen & Finnerup, 2014). Nociceptive stimuli are transduced into electric inputs that are then transmitted to the dorsal horn of the spinal cord. From there, modulation of nociceptor information occurs through different up- and down- regulations, whereby noxious stimulation can be increased or decreased via positive or negative feedback (Yam et al., 2018). The pain perceived is the final sum of all nociceptor influxes from activator and inhibitor signals reaching the cerebral cortex of the central nervous system. Other types of pain, like neuropathic and inflammatory pain, exist and have their own mechanisms. In fact, the experience of pain is not exclusive to only one pain type at a time (Mathews et al., 2014; B. Monteiro et al., 2023; Yam et al., 2018).

Acute Pain in Veterinary Medicine

Today, pain is considered the fourth vital sign. It ought to be evaluated systematically alongside the TPR (temperature, pulse, respiration) as part of every routine veterinary physical exam, regardless of the reason for consultation (B. Monteiro et al., 2023). However, pain assessment has not always been incorporated into physical examination. Indeed, pain relief practices in Canada using analgesic treatments in companion animals have undergone rapid changes over the last few decades.

Nearly thirty years ago, a national survey published in 1996 showed that only about 50% of veterinarians reported administering analgesics following non-ovariohysterectomy (OVH) abdominal surgeries (Dohoo & Dohoo, 1996a). When considering the type of surgeries, animals received analgesics in a majority of, but not all, orthopedic surgeries. The same could only be said for around 10% of castrations. When comparing between species, Canadian veterinarians in 1996 assigned similar or higher pain perception scores to dogs versus cats for the same type of surgery. Likewise, a similar or higher prevalence of analgesics was administered postoperatively to dogs versus cats (Dohoo & Dohoo, 1996b). Five years later in 2001, there was a marked increase in analgesic usage with 62% of Canadian veterinarians choosing between two classes of analgesics perioperatively. However, up to 12% of veterinarians still reported not using any analgesic for the castrations and ovariohysterectomies of dogs and cats (Hewson et al., 2006b). Another decade later in 2012, over 60% of veterinarians in the Canadian province of Ontario administered preemptive analgesia. Despite the increased usage, 9% of OVHs and castrations in dogs versus 16-22% of the same in cats were still performed without provision of perioperative analgesia (Reimann et al., 2017). This latest statistic is concerning because it suggests that, while analgesic administration is on the rise overall, cats continue to be perceived as less painful than dogs for the same procedures. In turn, cats receive less analgesia than dogs (Reimann et al., 2017).

Different factors can explain the significant progress in pain management by Canadian veterinarians. First, veterinarians are more aware of pain in animals being caused by different painful procedures. More precisely, veterinarians were asked in questionnaires if they believed certain procedures were painful or not to the animal. The survey results revealed an increase over time in the belief of pain (Beswick et al., 2016; Hugonnard et al., 2004; Menéndez et al., 2023; Williams et al., 2005). Furthermore, new and improved school curricula in veterinary programs offer greater emphasis on pain management. Recent graduates show higher pain perception scores than graduates from farther ago. This finding holds regardless of the school of graduation (Beswick et al., 2016; Hewson et al., 2006a). Intriguingly, some minor differences also persist due to the sex of the veterinarian, with female veterinarians scoring higher on pain perception than their male counterparts (Beswick et al., 2016; Dohoo & Dohoo, 1996a; Lorena et al., 2014; Raekallio et al., 2003; Williams et al., 2005). In all cases, continuing education has been shown to be important for improving pain perception, particularly to counter the potential difference in education of older graduates (Dohoo & Dohoo, 1996a).

Next, there has been an increase in the accessibility and usage analgesic drugs. In the 1996 survey, Canadian veterinarians were concerned by the use of potent opioids and the associated risk of adverse reactions. They were consequently motivated to minimize the use of postoperative analgesics (Dohoo & Dohoo, 1996a). Since that time, new molecules and techniques, such as preventive analgesia and multimodal analgesia, have been evaluated and promoted as safe analgesic protocols for cats and dogs (Bortolami & Love, 2015; Monteiro & Steagall, 2019a; Simon & Steagall, 2017; Steagall et al., 2014; Valverde & Skelding, 2019). New guidelines are likewise published regularly to help orient veterinarians in the best practices for optimal assessment, prevention, and management of pain in clinical contexts (Epstein et al., 2015; Gruen et al., 2022; Mathews et al., 2014; B. Monteiro et al., 2023; Ryan et al., 2019; Steagall et al., 2022).

Lastly, there have been improvements in the methods for objective pain recognition. While veterinarians may know that certain procedures are painful for the animal, their ability to recognize that pain may be compromised, causing deficient pain management. In the subsequent section, the various obstacles to effective pain recognition are outlined.

Differences in Acute Pain Among Domesticated Dogs and Cats

Cats are the most popular household animal in Canada, counting approximately 8.5 million individuals nationwide in 2022. Cats have overtaken the number of dogs by about 0.6 million individuals (CAHI, 2022). Despite their greater popularity, the primary demand for companion animal veterinary services in Canada is driven by dog consultations. While 86% of dogs in Canada saw a veterinarian at least once in 2022, only 61% of cats consulted within the same time frame. A small but significant 3% increase in the administration of medications to cats from 2020 to 2022 seems to have helped narrow this historical gap between dogs and cats (CAHI, 2022).

In addition to fewer absolute veterinary visits, cats have long been under-diagnosed and under-treated for similar perioperative and traumatic pain in comparison to dogs. This trend has only begun to reverse in the early 2000s (Taylor & Robertson, 2004). Before the prevalence of feline-specific studies, pain management principles in cats had been extrapolated from so-called canine equivalents (Steagall, 2020). However, this shortcut is a false equivalence as major differences exist and must be considered between these two most popular domesticated species.

Physiological Differences

Cats have physiological peculiarities unique to their species and pharmacological studies cannot be cross-applied. For example, cats are deficient in several metabolic conjugation pathways. The result is less effective drug metabolism and elimination with many drugs having a longer half-life than in dogs (Court, 2013). These specific differences must be considered in the choice, posology, and route of administration of certain drugs as cats suffer from a consequent higher risk of toxicity. Conversely, other drugs use metabolic pathways that are similar or even more efficient in cats than in dogs (Steagall, 2020; Taylor & Robertson, 2004). For example, piroxicam, a nonsteroidal anti-inflammatory drug, has a much faster elimination half-life in cats than in dogs and humans (Court, 2013). As can be seen, there is no broad equivalence. Continued studies in feline-specific metabolic pathways, pharmacokinetics, and pharmacodynamics are necessary to optimize pain management in cats.

Behavioural Differences

Behavioural changes are one of the most useful signs of pain in animals (B. Monteiro et al., 2023; Reid et al., 2018). In addition to their physiological differences, cats tend to be more subtle with their outward expression of pain in comparison to dogs (Gruen et al., 2022; Steagall, 2020). This difference in behaviour can be potentially attributed to the unique history of cat domestication. Until fairly recently, the main role of a cat in human society was to be a solitary rodent hunter (Young, 1985). The preservation of their status as mid-level predators can serve to explain their display of both predator and prey behaviours (Gruen et al., 2022).

When in pain, cats tend not to overtly show pain-specific behaviours. While the evolutionary rationale may be to avoid resembling prey, it results in a more challenging pain assessment for the veterinarian (Gruen et al., 2022; Steagall, 2020). Other factors like fear, stress, as well as human and other animal interaction may exacerbate the drive to conceal these pain-specific behaviours (Hernandez-Avalos et al., 2019; Steagall, 2020).

In contrast, dogs have evolved alongside humans for between 32,000 and 16,0000 years (Tancredi & Cardinali, 2023). As humanity's first domesticated animal, the dog has been artificially selected and bred to enhance desirable behavioural traits such as friendliness and companionship (Perri et al., 2021; Young, 1985). In turn, dogs are rewarded for their ability to communicate and exhibit social behaviour. Indeed, animals are broadly known to be capable of recognizing and expressing emotions among individuals of their own species to share intentions and motivations. Dogs in particular, however, have set themselves apart by progressing to perceiving the emotions of humans (Albuquerque et al., 2016).

The veterinarian's expert understanding of differences between species and ability to identify pain related behaviours is essential to efficiently treat pain. At the same time, the importance of owner-animal-veterinarian collaboration helps tremendously in the identification of anomalous individual-specific behaviour across the large variety of existing cats. Educating and engaging owners in the detection of early signs of pain in their animal remains the key to optimal pain management and animal welfare (Gruen et al., 2022; B. Monteiro et al., 2023; Väisänen et al., 2007).

Differences in Acute Pain According to Age

There is varied pain presentation within a species when observing subpopulations of different ages. In veterinary medicine, the question of pain is salient in younger animals, since they are the patient group routinely recommended to undergo one of the most common surgical procedures: neutering. Around the world and across domesticated species, the threat of overpopulation is a major concern due to potential ecological, public health, and financial taxpayer repercussions. Spay-neuter programs are essential to curb the exponential growth that would otherwise occur without intervention (Deak et al., 2019; Griffin et al., 2016; Sparkes et al., 2013).

Specifically in cats, the International Society of Feline Medicine currently supports neutering of kittens aged four months or below (Sparkes et al., 2013). This timing is significantly earlier than the traditional six months of age milestone, but prepubertal gonadectomy assures a higher efficiency of population control if done before animal sexual maturity. The recommendation is supported by numerous international veterinary associations such as the American Veterinary Medical Association, the Canadian Veterinary Medical Association, the American Animal Hospital Association, and the American Society for the Prevention of Cruelty to Animals (AAHA, 2018; ASPCA, n.d.; AVMA, n.d.; CVMA, 2019).

Lending support to the practice of early neutering, the safety of neutering procedures has been demonstrated as early as at six weeks of age (Howe, 2015). Despite concerns raised by some animal welfare and private clinical practice groups, studies looking at short- and long-term effects found no significant physical or behavioural problems associated with early neutering (Joyce & Yates, 2011; Moons et al., 2018). To further promote safety, increasingly more studies are investigating analgesic-anesthetic protocols that would provide improved postoperative pain relief in kittens (Diep et al., 2020; Joyce & Yates, 2011; Malo et al., 2023; Olson et al., 2001; Porters et al., 2015).

When observing pain levels during a routine OVH, kittens were reported as being less painful and as recovering faster than older cats, based on subjective descriptive pain scales. Wound sensitivity, as measured by mechanical nociceptive threshold from pressure applied to the surgical wound site, was similar. This finding confirms that kittens do not have a sensory difference that would better explain the reduced pain response (Polson et al., 2014). In a subsequent study, a lower prevalence of rescue analgesic was administered to kittens aged four to six months compared to adult cats (Diep et al., 2020).

There is hesitation in conclusively stating that kittens are less painful than their adult counterparts. The possibility of shortfalls in kitten pain evaluation is ever-present. Similarly, there is a conspicuous lack of evidence in the literature to describe pain in kitten populations. For optimal pain management in kittens, further investigation in the field is required to better determine the differences in painful behaviours compared to adults and to guide the development of objective instruments that evaluate this type of pain.

Acute Pain Recognition

Since the presence of pain suggests the potential for an underlying medical problem, accurately identifying it is important for the diagnosis and treatment of injuries or disease processes. In terms of more symptomatic relief, assessing pain is also simply the first step to efficiently managing pain. In veterinary medicine, pain recognition is complicated by the incapacity of the patient population to verbalize how they feel (Sneddon et al., 2014). Not only does patient discomfort become difficult to express, but the inadequacy of treatment and the need for additional treatment is likewise difficult to convey. Similar studies in human medicine have shown that non-verbal patients are at increased risk of having their pain both under-recognized and under-treated (Bellieni, 2012; Witt et al., 2016).

Identifying pain-specific behaviours and adapting pain assessment tools to the specific patient population are two important skills that healthcare providers must master to identify and treat pain appropriately (Booker & Haedtke, 2016; McGuire et al., 2016). In addition, veterinary medicine adds the complexity of having to treat multiple species with knowledge of their specific anatomy, physiology, and behavioural differences. As a result, there is a glaring need for objective measures adapted to each species in order to guide pain evaluation with the appropriate subsequent treatments.

Acute Pain Assessment Instruments in Cats

Following the recent interest in feline pain management, multiple objective instruments have been developed to offer a consistent and systematic method of evaluating pain behaviours specific to adult cats. The three major acute feline pain assessment instruments with reported validity and analgesic intervention score will be presented in this section.

Glasgow Composite Measure Pain Scale - Feline (CMPS-F)

Inspired by its equivalent in dogs, the CMPS-F evaluates eight components of acute pain: vocalisation, posture and activity, attention to wound, ear position, shape of muzzle, response to interaction with the observer, palpation of painful area, and qualitative behaviour (NewMetrica, 2019; Steagall et al., 2022). A definitive version (rCMPS-F) was published in 2017 after revision of two questions (Calvo et al., 2014). The rCMPS-F has reported evidence of construct validity as well as responsiveness. An analgesic intervention score was determined to be ≥ 5/20 (Reid et al., 2017).

UNESP-Botucatu Multidimensional Feline Pain Assessment Scale (UFEPS)

The UFEPS is an extensive and complex tool that takes into consideration several variables of acute pain in cats including behavioural, psychomotor, and physiological changes (Brondani et al., 2013). A short form (UFEPS-SF) was recently published making the evaluation less time-consuming (Belli et al., 2021). Initially developed in Brazilian Portuguese, it has been translated into multiple languages and has undergone extensive validity, reliability, responsiveness, sensitivity, and specificity testing in several painful conditions. Both full and short forms of the UFEPS have analgesic intervention scores of ≥7/30 and 4/12, respectively (Brondani et al., 2013; Luna et al., 2022).

Feline Grimace Scale (FGS)

The FGS is a facial expression-based scoring system comprising five action units (AUs): ear position, orbital tightening, muzzle tension, whiskers position, and head position. Each AU is scored from 0 to 2: 0 = AU is absent; 1 = moderate appearance of the AU or uncertainty over its presence or absence; and 2 = obvious appearance of the AU. If the presence or absence of an AU is uncertain, a score of 1 can be given. The sum of all five scores gives the total FGS score, with a maximum of 10. Total ratio scores are obtained by considering the sum of scores for each visible AU divided by the maximum possible score of all visible AUs. It is important for evaluators to not interact or stimulate the cat while scoring the FGS. An FGS total ratio score of ≥0.39/1 suggests that analgesia should be administered (Evangelista et al., 2019).

The FGS has demonstrated validity and reliability in various acute painful conditions (Evangelista et al., 2019). Its reliability was shown by evaluating cats before and after dental extractions (Watanabe et al., 2020). While FGS evaluation can be applied using real-time or image assessment, the former does lead to a slight overestimation when compared to the latter.

However, this is probably of minimal clinical relevance (Evangelista et al., 2020). Good agreement and reliability were observed when the FGS was used by individuals with differing levels of expertise, including cat caregivers (Evangelista & Steagall, 2021; B. Monteiro et al., 2023). Furthermore, the FGS showed strong validity, reliability, and discriminative ability in a recent systematic review assessing the measurement properties of grimace scales among different species (Evangelista et al., 2022). Since December 2023, a fully automated deep learning model was developed to predict FGS pain scores based on automated facial landmark position and FGS scoring (Steagall et al., 2023).

No pain assessment tool currently exists for kittens to the best of the authors’ knowledge. As has been discussed, this specific subpopulation has unique differences that prevent the automatic application of existing tools with no further supporting analysis. Kittens clearly stand to benefit from a high-quality pain assessment instrument.

Instrument Validation

As discussed in the previous sections, pain assessment tools play a critical role in veterinary medicine. In a clinical setting, veterinarians can use the results provided by these instruments to base their medical decisions and analgesic treatment plans. Ideally, pain assessment should be rigorous to maximize clinical utility. At the same time, the instrument chosen should be simple and easily integrated into a routine examination. The most important characteristics of a pain assessment tool remain, however, its psychometric properties such as its validity, reliability, and responsiveness. The goal is to have an objective guarantee as to the integrity of the various results provided by the instrument (Roach, 2006). In other words, the quality of the tool is most dictated by its psychometric test performance (Cypress, 2017).

It is commonly accepted that pain is universally present in animals. However, the measured outcome from a pain assessment instrument may not be universally applicable to all animals or even within subpopulations of the same species. The validity, reliability, and responsiveness of a tool are therefore characteristics pertaining to a defined population with specific clinical and demographic attributes (Roach, 2006). While the FGS has been shown to be valid and reliable in adult cats, it is evident that these same factors need to be evaluated in kittens.

Validity

Test validity is defined as “the degree to which an [...] instrument measures the construct(s) it purports to measure” (Mokkink et al., 2010). In this thesis, the precise construct is pain in kittens. Validity is a complex measurement property that regroups various subtypes. In a multidisciplinary international taxonomy consensus study, experts decided on three key subtypes based on study designs and methods: content validity, construct validity, and criterion validity. Some validity subtypes even englobe their own sub-subtypes (Mokkink et al., 2010).

Content Validity

Content validity is defined as “the degree to which the content of an [...] instrument is an adequate reflection of the construct to be measured” (Mokkink et al., 2010). This validity subtype inspects each item of an instrument and measures whether they appropriately capture the whole range of attributes of the construct (Ahmed & Ishtiaq, 2021; DeVon et al., 2007; Souza et al., 2017). Recall that pain is a complex phenomenon and that veterinary medicine suffers from additional complexities in pain evaluation. As a result, pain is a difficult construct to evaluate in our kitten population. It was decided the content validity of the FGS would not be measured due to its limitation in assessing only one domain of pain: facial expression. Indeed, other known dimensions of pain such as behavioural and psychological are omitted.

Construct Validity

Construct validity is defined as "the degree to which the scores of an [...] instrument are consistent with hypotheses (for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the [...] instrument validly measures the construct to be measured" (Mokkink et al., 2010). In this thesis, construct validity is measured via hypothesis testing with known-group discrimination. Since the evaluated construct, pain, cannot be directly observed, the actual analysis is based on a theoretical framework (Souza et al., 2017). The chosen test hypothesis provides the direction expected of the score for the measure. Results that align with the hypothesis are associated with greater construct validity of the instrument (DeVon et al., 2007; Souza et al., 2017). Beyond hypothesis testing, construct validity may also include structural validity and cross-cultural validity (Mokkink et al., 2010). These sub-subtypes are outside the scope of this work, however.

Criterion Validity

Criterion validity is defined as "the degree to which the scores of an [...] instrument is an adequate reflection of a 'gold standard'" (Mokkink et al., 2010). The chosen gold standard must be a criterion known to be valid for the same underlying construct. An agreement between the two instruments implies an accurate measure of that construct, here pain. The extent of the validity of the new instrument is limited by the validity of the criterion selected as the gold standard (DePoy & Gitlin, 2016; DeVon et al., 2007). In the seminal FGS study, criterion validity was measured via comparison with rCMPS-F, a validated pain scale for cats. A very strong correlation was observed between the two pain assessment tools (Evangelista et al., 2019). Due to the global absence of pain measures for kittens at the time of this experiment, criterion validity is not evaluated in this work.

Reliability

Test reliability is defined as “the degree to which the measurement is free from measurement error.” More specifically, we evaluate “the extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: for example, [...] over time (test–retest) by different persons on the same occasion (inter-rater) or by the same persons (i.e., raters or responders) on different occasions (intra-rater)” (Mokkink et al., 2010). In this thesis, we ask if the pain assessment tool is well-designed to allow veterinarians, the raters, to give the same score between each of them, for the same kitten, at the same time (inter-rater reliability), as well as again with themselves at a later time (intra-rater reliability).

Reliability is an important factor to verify in pain assessment, and can be said to be necessary for a high-quality instrument. However, it is not sufficient on its own due to the possibility of repeated and consistent wrongly measured outcomes (Cypress, 2017; DeVon et al., 2007; Souza et al., 2017). In other words, high reliability suggests that there is strong consensus among veterinarians for each pain score, but says nothing about the accuracy of these scores. Since pain scores direct the treatment plan in clinical contexts, high test validity is equally essential.

As a summary of test validity versus reliability, the metaphor of a shooting target is useful. The accuracy of the shots, whether they are on or off the center of the target, denote validity. The precision or distribution of the shots, on the other hand, represent reliability. Shots that consistently hit the same location on the target are surely reliable, but are invalid if the majority land off-centered. Shots that land near the center of the target are certainly valid, but are unreliable if distributed like a random mass around the center (Souza et al., 2017).

With regard to the FGS, perfect reliability would entail all veterinarians reporting identical scores for the same kittens, at the same time as well as later at a second moment. Perfect validity would involve all painful kittens having a score at least 4/10 and all non-painful kittens a score strictly under this predetermined threshold. To be clear, strong validity with weak reliability would imply that painful versus non-painful kittens are well-discriminated, but for the same kitten, there is large variance in assigned scores within their range above or below the threshold. In contrast, strong reliability with weak validity would arise when all veterinarians agree on the same precise score for each kitten, but the majority of painful kittens are given scores strictly below 4/10 while most non-painful kittens receive scores above this threshold.

Responsiveness

Test responsiveness is defined as "the ability of an [...] instrument to detect change over time in the construct to be measured" (Mokkink et al., 2010). In this thesis, we ask if the FGS is capable of detecting the evolution of pain levels in kittens. More precisely, do pain scores respond as expected by increasing significantly after a painful event and decreasing significantly after administration of an analgesic with known efficacy? Responsiveness is highly clinically relevant to identify changes in a cat's pain level. Furthermore, a responsive pain assessment tool allows veterinarians to individualize their treatment plan to address the cat's specific needs.

Objectives and Hypothesis

The objective of this thesis is to evaluate the construct validity, responsiveness, as well as inter- and intra-rater reliability of the FGS in kittens undergoing an elective OVH.

The hypothesis is that the FGS will be a valid, responsive, and reliable pain assessment instrument for kittens. It will offer, respectively, significant discrimination between pain scores before and after surgery, before and after the administration of analgesics to painful kittens, and consistent scores between raters at one time and within raters over time.

Chapter 2 – Scientific Article

Original Article

Construct validity, responsiveness and reliability of the Feline Grimace Scale© in kittens

Alice J Cheng[1], Annie Malo[1], Marta Garbin[1], Beatriz P Monteiro[1], Paulo V Steagall[1,2]

[1]Department of Clinical Sciences, Faculty of Veterinary Medicine, Université de Montréal, Saint-Hyacinthe, QC, Canada

[2]Department of Veterinary Clinical Sciences, Centre for Companion Animal Health and Welfare, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Hong Kong S.A.R., China

Keywords: analgesia; kitten; pain; pain assessment; facial expressions; Feline Grimace Scale

Corresponding author: Paulo V Steagall DVM, MSc, PhD, Diplomate ACVAA

This article has been published by the *Journal of Feline Medicine and Surgery*, Volume 25, Issue 12, December 2023

https://doi.org/10.1177/1098612X231211765

Abstract

Objectives

This study aimed to investigate construct validity, responsiveness and reliability of the Feline Grimace Scale (FGS) in kittens.

Methods

A total of 36 healthy female kittens (aged 10 weeks to 6 months) were included in a prospective, randomized, blinded study. Video recordings of all kittens were made before and 1 and 2 h after ovariohysterectomy using an opioid-free injectable anesthetic protocol with or without multimodal analgesia. Additional recordings were taken before and 1 h after administration of rescue analgesia (buprenorphine 0.02 mg/kg IM) to painful kittens. Screenshots of facial images were collected from the video recordings for FGS scoring. Four observers blinded to treatment groups and time points scored 111 randomized images twice with a 5-week interval using the FGS. Five action units (AUs) were scored (ear position, orbital tightening, muzzle tension, whiskers position and head position; 0–2 each). Construct validity, responsiveness, and inter- and intra-rater reliability were evaluated using linear models with Benjamini–Hochberg correction, Wilcoxon signed-rank test and single intraclass correlation coefficients (ICC_{single}), respectively ($P < 0.05$).

Results

FGS total ratio scores were higher at 1 and 2 h after ovariohysterectomy (median [interquartile range, IQR]: 0.30 [0.20–0.40] and 0.30 [0.20–0.40], respectively) than at baseline (median [IQR]: 0.10 [0.00–0.30]) ($P < 0.001$). FGS total ratio scores were lower after the administration of rescue analgesia (median [IQR] before and after rescue analgesia) 0.40 [0.20–0.50] and 0.20 [0.10–0.38],

respectively ($P < 0.001$). Inter-rater ICC_{single} was 0.68 for the FGS total ratio scores and 0.35–0.70 for all AUs considered individually. Intra-rater ICC_{single} was 0.77–0.91 for the FGS total ratio scores and 0.55–1.00 for all AUs considered individually.

Conclusion and Relevance

The FGS is a valid and responsive acute pain-scoring instrument with moderate inter-rater reliability and good to excellent intra-rater reliability in kittens.

Introduction

There have been significant advances in feline perioperative pain management with the publication of three instruments for acute pain assessment :[1] the Glasgow Composite Measure Pain Scale-Feline (CMPS-F); [2] the UNESP-Botucatu multidimensional feline pain assessment scale (UFEPS) and its short form (UFEPS-SF), [3, 4] and the Feline Grimace Scale (FGS). [5-7] The FGS is a facial expression-based scoring system comprising five action units (AUs): ear position, orbital tightening, muzzle tension, whiskers position, and head position. Each AU is scored from 0 to 2, for a maximum FGS score of 10. Total ratio scores are obtained by considering the sum of scores for each visible AU divided by the maximum possible score. An FGS total ratio score ≥0.39/1 suggests that analgesia should be administered.[5] The FGS has demonstrated validity and reliability in various acute painful conditions, [5, 8] when evaluated using real-time or image assessment, [6] and when used by individuals with differing levels of expertise, including cat caregivers. [9, 10]

Although some of these assessment tools have gone through extensive validation in adult cats, none have been uniquely validated in kittens. Nonetheless, early-age neutering (≤4 months of age) is now common practice in feline medicine and is supported by organizations such as the International Society of Feline Medicine for optimal health, social development and population control. [11,12] Considering that thousands of kittens undergo early-age neutering every year, it is concerning that pain assessment tools have yet not been validated in this population to ensure that pain is recognized and proper analgesic treatment is administered based on objective pain scores.

The objective of the present study was to investigate the construct validity, responsiveness, and inter- and intra-rater reliability of the FGS in kittens undergoing ovariohysterectomy (OVH). The hypothesis was that the FGS total ratio scores would be significantly different before and after surgery, and before and after the administration of analgesics, and would present good reliability between raters and within raters over time.

Materials and methods

Ethical Statement

The present study was conducted in parallel with a prospective, randomized, blinded clinical trial comparing an opioid-free injectable anesthetic protocol with or without multimodal analgesia in kittens undergoing OVH. [13] The study protocol was approved by the *Comité d'éthique de l'utilisation des animaux* (Animal Use Ethics Committee) of the Faculty of Veterinary Medicine, Université de Montréal (protocol 21-Rech-2132), performed in accordance with the Canadian Council on Animal Care and reported according to the ARRIVE guidelines. [14]

Animals

A total of 40 domestic healthy female kittens of any breed, aged between 10 weeks and 6 months, weighting ≥1 kg and with a body condition score of 4–6/9 were recruited by one of the investigators (AJC) from two local shelters between July and August 2021. Kittens were excluded if they were beyond the age limits, male, or showing shy or feral behaviours or signs of disease (pain, hyperthermia, etc.). Any other individual was included in the study. Kittens were admitted to the veterinary teaching hospital (*Centre hospitalier universitaire vétérinaire*) of the Faculty of Veterinary Medicine, Université de Montréal approximately 16 h before elective OVH. Written consent was obtained for each kitten.

After admission, kittens were individually and randomly housed in adjacent stainless-steel cages in a cat ward with temperature control. Each kitten had access to water *ad libitum*, a litter box, a blanket, a cardboard box and toys, except during filming periods (see 'Pain assessment and video recording'). Kittens were left undisturbed for 3 h in their cages to acclimatize to the study site and personnel. Thereafter, a detailed physical examination was performed by two veterinarians (AM, BPM) using low-stress, feline-friendly interactive techniques.[15,16] Soft food (Royal Canin Gastrointestinal Kitten; Royal Canin) was offered at specific time points in proportion to each kitten's individual daily calorie requirements. Kittens were discharged 24 h after surgery and returned to their respective shelters for adoption.

Anesthesia and Surgery

Briefly, food was removed 6–10 h before the induction of general anesthesia, but water remained accessible. Each kitten received 0.5 ml of corn syrup applied directly to the gingival membrane on the morning of surgery to prevent hypoglycemia. All kittens received an intramuscular (IM) injection of ketamine (4 mg/kg; Ketaset; Zoetis), dexmedetomidine (40 µg/kg; Dexdomitor; Zoetis) and midazolam (0.25 mg/kg; Midazolam; Sandoz) administered together into the lumbar epaxial muscles using a 1 ml syringe. A randomization plan assigned each kitten to either the multimodal group (MMG) or control group (CG). In the MMG, animals received subcutaneous (SC) meloxicam (0.1 mg/kg; Metacam 0.5%; Boehringer Ingelheim) before surgery and bupivacaine hydrochloride 0.25% (2 mg/kg; Bupivacaine injection BP; Sterimax) administered intraperitoneally before OVH. In the CG, kittens received saline (0.9% sodium chloride solution) at equal volumes and routes of administration as the meloxicam and bupivacaine in the MMG. The same veterinarian with experience in surgery (BPM) performed an elective OVH using the pedicle tie technique. During anesthetic recovery, kittens received 5 ml/kg of lactated Ringer's solution (Lactated Ringer's Inj. Bag/500 ml; McCarthy & Sons Service) SC between the shoulder blades. An IM injection of atipamezole (0.4 mg/kg; Antisedan; Zoetis) was administered 15 mins after the end of surgery and each kitten was returned to its respective cage for pain assessment and postoperative care.

Pain Assessment and Video Recording

Pain was assessed in real time using the UFEPS-SF before surgery (baseline) and 1, 2, 4, 6, 8, 12 and 24 h postoperatively. Pain assessment was performed by a female veterinarian (AM) who was blinded to the treatment groups. This individual had 7 years of experience in clinical practice. Before to the study, she completed training using the UFEPS with videos available on the website www.animalpain.org . Rescue analgesia was administered to kittens with UFEPS-SF scores ≥4/12 using IM buprenorphine (0.02 mg/kg; Vetergesic; Champion Alstoe) in the MMG and CG, and SC meloxicam (0.1 mg/kg) in the CG. The administration of the rescue analgesia was performed by veterinarians who were not involved with pain assessment. Regardless of the treatment group, kittens weighing over 2 kg or older than 16 weeks received a second dose of meloxicam by the oral route of administration (0.05 mg/kg; Metacam 0.5 mg/ml oral suspension; Boehringer Ingelheim) 24 h after the first dose.

Video recordings of the kittens were made by an observer (AJC) using high-definition wide-angle cameras (GoPro Hero 5 and GoPro Hero 9; GoPro Inc.) at 60 frames per second. Cameras were attached to the cage bars at the kitten’s eye level and adjusted when needed. The cage walls were covered with a light-colored cardboard to ensure background uniformity during filming and image assessment. Diffused external lights were used for appropriate lighting. Cages were emptied apart from a blanket and the kitten. During FGS video recording, kittens were left undisturbed. Videos with a duration of 3 mins videos were recorded at baseline, 1 and 2 h after OVH and before and 1 h after the administration of rescue analgesia. In some cases, extended videos (total of 9 mins) were collected for the purpose of a future study involving kitten behaviour, but only the first 3 mins were used for the purpose of this study.

Image Capturing, Selection and Cropping

A conversion software (Free Video to JPG Converter; DVDVideoSoft) was used to collect still images from video recordings with the output set at two images per second. An observer (AJC) who was blinded to the treatment group and UFEPS-SF pain scores selected one image for each third of the video, overall collecting three images per video in total. For an image to be included in the study, the kitten had to be facing the camera with visible AUs. If it was leaning on to a surface (eg, the cage wall or floor), the image could still be included if the AUs were visible on the contralateral side. Images were excluded if they were of poor quality or if kittens were vocalizing, grooming or sleeping. Of the three images selected per video, the best image (in terms of quality and definition) was chosen to characterize the kitten at that time point for later scoring. Subsequently, images were cropped to include the entire face of the kitten using software (Adobe Photoshop CS6 V13.0; Adobe). Brightness was increased for dark images. A short video is provided, illustrating the procedure for video recording, image capturing, selection and cropping (see video in supplementary material).

Image Scoring

Four raters (AM, BPM, MG and PVS) received the FGS training manual and participated in a training session provided by the principal investigator (PVS) before to image scoring.[5] In this session, raters independently scored 10 images from the 'Practice your Skills' menu of the FGS website (www.felinegrimacescale.com) and discussed their scores. Thereafter, an online survey (LimeSurvey; LimeSurvey GmbH) was built to allow scoring of the final selected images. Images were randomly ordered using a random sequence generator (www.random.org). For each image, the rater, blinded to treatment groups and time points, scored the five AUs of the FGS from 0 to 2: 0 = AU is absent; 1 = moderate appearance of the AU or uncertainty over its presence or absence; and 2 = obvious appearance of the AU. If the AU was not clearly visible, raters had the option of checking 'not possible to score'. An image was excluded from statistical analysis when two or more AUs were not possible to be scored by at least one rater. A FGS total ratio score for each image scored by each rater was calculated by summing the scores from each AU and dividing by the maximum possible score, excluding the AUs marked as 'not possible to score'. Image scoring was divided into two parts; raters were instructed to complete evaluations at least 24 h apart to avoid fatigue. Five weeks later, the same procedure was repeated with raters scoring the same images using a different randomized sequence. Image scoring was performed between November 2021 and January 2022.

Statistical Analysis

Statistical analyses were performed using R (R Core Team V4.0.3; R Foundation for Statistical Computing). Construct validity was assessed by comparing FGS total ratio scores at baseline vs 1 and 2 h. A repeated measures linear mixed model with the Benjamini–Hochberg correction was used. The kitten was treated as a random effect and time as fixed effect. Responsiveness was assessed by comparing FGS total ratio scores before and 1 h after the administration of rescue analgesia. Normality was verified with the Kolmogorov–Smirnov test and through visual inspection. The FGS total ratio scores were not normally distributed and the Wilcoxon signed-rank test was used for group comparisons. Inter- and intra-rater reliability were assessed for each AU and for the FGS total ratio scores using the intraclass correlation coefficient (ICC) with the 95%

confidence interval (CI). Inter-rater reliability was assessed using a two-way random effect model for absolute agreement, whereas intra-rater reliability was assessed using a two-way mixed effect model. The single and average ICCs (ICC_{single} and $ICC_{average}$, respectively) were interpreted as follows: <0.5 = poor reliability, between 0.5 and 0.75 = moderate reliability, between 0.75 and 0.9 = good reliability and >0.90 = excellent reliability.[17] *P* values of $P < 0.05$ were considered statistically significant.

Results

A total of 111 images of 36 female kittens (mean age 15 ± 4.7 weeks of age; mean body weight 1.5 ± 0.4 kg) were included in the statistical analysis (Figure 1). The mean ± standard deviation for the duration of FGS image scoring were 73 ± 33 mins and 68 ± 23 mins for the first and second round of assessments, respectively.

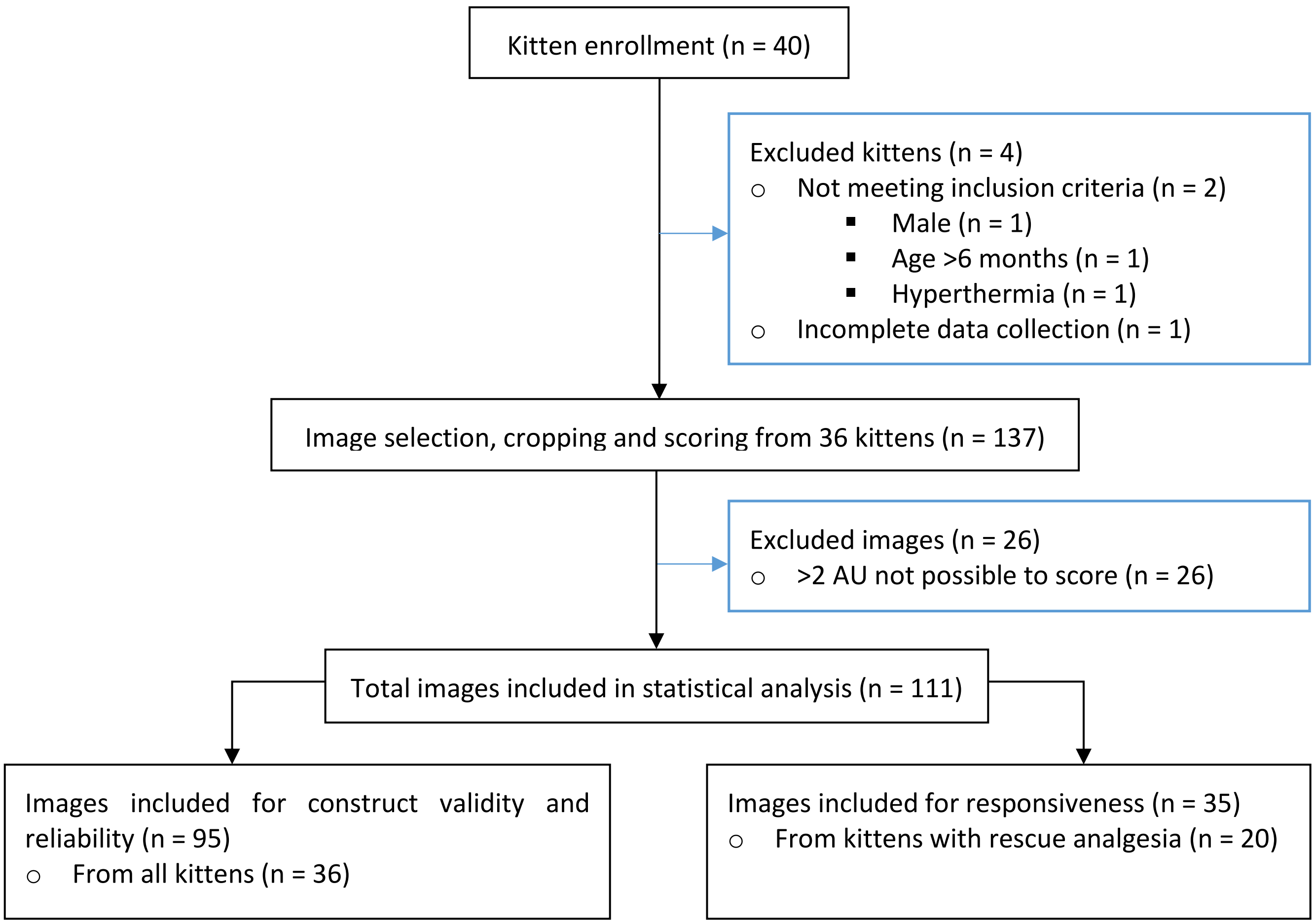

Figure 1. – Flowchart of a prospective, randomized, and blinded study investigating the construct validity, responsiveness, inter-rater and intra-rater reliability of the Feline Grimace Scale using facial images of kittens undergoing ovariohysterectomy

A total of 95 images of 36 kittens were used to assess construct validity by known-group discrimination. The FGS total ratio scores were increased at 1 and 2 h (median [interquartile range, IQR range]: 0.30 [0.20–0.40] and 0.30 [0.20–0.40], respectively) when compared with baseline (median [IQR]: 0.10 [0.00–0.30]) (*P* <0.001) (Figures 2 and 3).

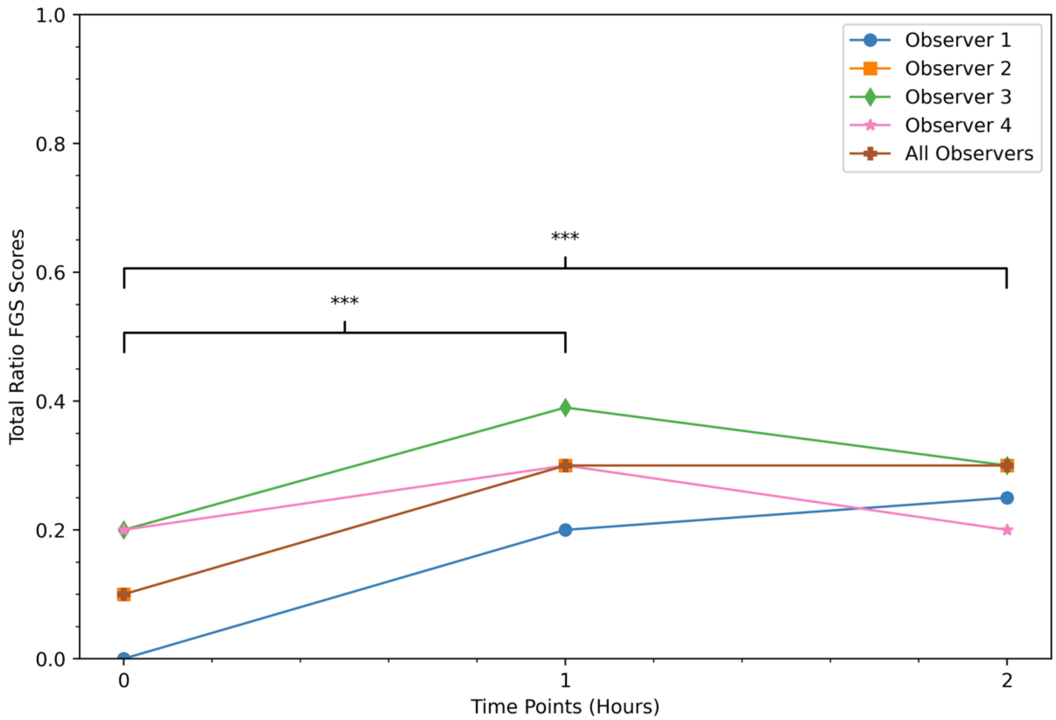

Figure 2. – Median total ratio Feline Grimace Scale (FGS) scores in kittens assessed using known-group discrimination. There were significant increases of the FGS total ratio scores at 1 and 2 h after compared with baseline (*P* < 0.001), which demonstrates construct validity

Figure 3. – An example of a kitten with changes in the Feline Grimace Scale total ratio scores (a) before and (b) after surgery, demonstrating construct validity of the tool

A total of 35 images of 20 kittens were used to assess responsiveness. The FGS total ratio scores were significantly decreased 1 h after administration of rescue analgesia (median [IQR] before and after rescue: 0.40 [0.20–0.50] and 0.20 [0.10–0.38], respectively) (*P* <0.001) (Figures 4 and 5).

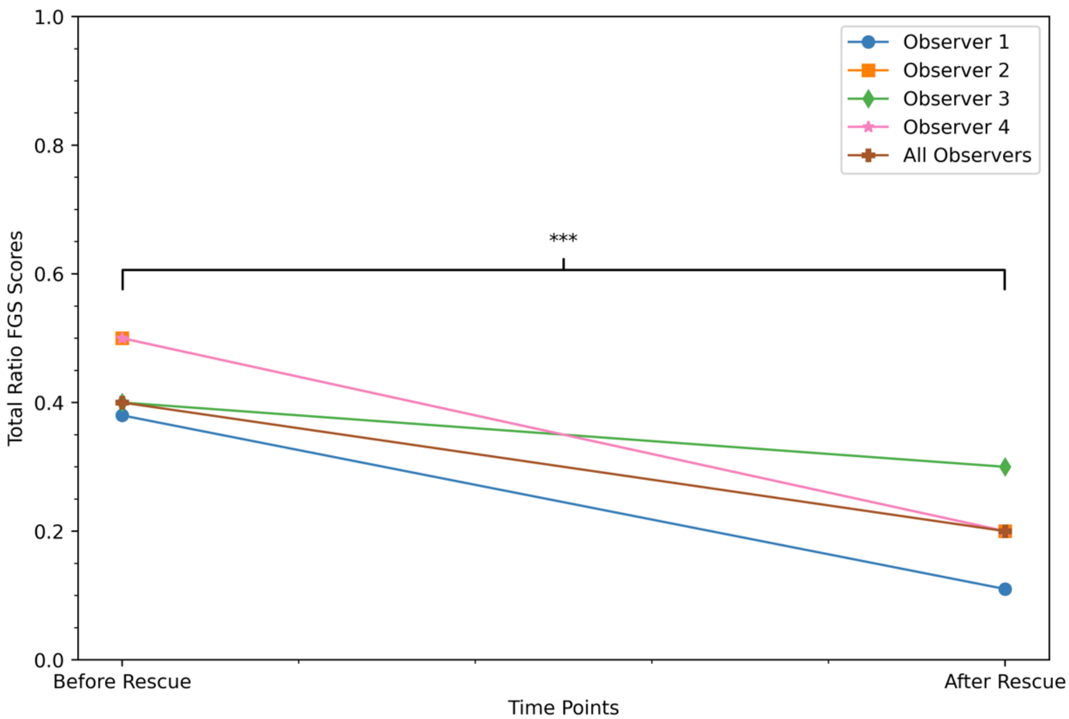

Figure 4. – Median total ratio Feline Grimace Scale (FGS) scores after an intramuscular injection of buprenorphine with or without subcutaneous meloxicam in kittens showing signs of pain after ovariohysterectomy. There were significant decreases of the FGS total ratio scores 1 h after rescue analgesia (*P* <0.001), demonstrating responsiveness of the FGS in kittens

Figure 5. – An example of a kitten with changes in the Feline Grimace Scale total ratio scores (a) before and (b) after the administration of buprenorphine and meloxicam by the intramuscular and subcutaneous routes, respectively, demonstrating responsiveness of the tool

A total of 95 images of 36 kittens were used to assess reliability. The inter-rater reliability of the FGS total ratio scores was moderate and excellent according to the ICC_{single} and $ICC_{average}$, respectively. The ICC_{single} was moderate for ear position, orbital tightening and head position, and poor for muzzle tension and whiskers position. The $ICC_{average}$ was excellent for orbital tightening, good for ear position, muzzle tension and head position, and moderate for whiskers position (Table 1). The intra-rater reliability for the FGS total ratio scores of the four raters was good to excellent. The ICC_{single} ranged from moderate to excellent for ear position and orbital tightening, moderate to good for muzzle tension, moderate for whiskers position and good to excellent for head position (Table 2).

Action unit	ICC_{single} (95% CI)	$ICC_{average}$ (95% CI)
Ear position	0.62 (0.53 – 0.71)	0.77 (0.61 – 0.85)
Orbital tightening	0.70 (0.61 – 0.77)	0.92 (0.89 – 0.95)
Muzzle tension	0.47 (0.36 – 0.59)	0.83 (0.68 – 0.90)
Whiskers position	0.35 (0.22 – 0.48)	0.63 (0.42 – 0.77)
Head position	0.66 (0.53 – 0.76)	0.89 (0.83 – 0.93)
FGS total ratio scores	0.68 (0.57 – 0.77)	0.91 (0.87 – 0.94)

Tableau 1. – Inter-rater reliability of the Feline Grimace Scale in kittens undergoing ovariohysterectomy. Data were obtained from round 1 assessments and are reported as single and average intraclass correlation coefficient (ICC_{single} and $ICC_{average}$, respectively) and 95% confidence interval (CI). The ICC_{single} and $ICC_{average}$ were interpreted as follows: <0.5 = poor reliability; 0.5–0.75 = moderate reliability; 0.75–0.9 = good reliability; and >0.90 = excellent reliability[21]

FGS = Feline Grimace Scale

Action unit	Rater 1	Rater 2	Rater 3	Rater 4
Ear position	0.71 (0.59 – 0.80)	0.75 (0.64 – 0.82)	0.65 (0.52 – 0.76)	0.93 (0.90 – 0.95)
Orbital tightening	0.69 (0.57 – 0.78)	1.00 (1.00 – 1.00)	1.00 (1.00 – 1.00)	0.92 (0.88 – 0.95)
Muzzle tension	0.70 (0.58 – 0.79)	0.75 (0.64 – 0.82)	0.68 (0.55 – 0.77)	0.86 (0.79 – 0.90)
Whiskers position	0.55 (0.39 – 0.68)	0.57 (0.42 – 0.69)	0.66 (0.53 – 0.76)	0.75 (0.65 – 0.83)
Head position	0.89 (0.85 – 0.93)	0.79 (0.69 – 0.85)	0.78 (0.68 – 0.85)	0.95 (0.92 – 0.97)
FGS total ratio scores	0.82 (0.75 – 0.88)	0.78 (0.69 – 0.85)	0.77 (0.67 – 0.84)	0.91 (0.86 – 0.94)

Tableau 2. – Intra-rater reliability of the Feline Grimace Scale scores in kittens undergoing ovariohysterectomy. Raters scored each image twice in a different randomized order 5 weeks apart. Data are reported as single intraclass correlation coefficient (ICC_{single}) and 95% confidence interval (CI). The ICCsingle was interpreted as follows: <0.5 = poor reliability; 0.5–0.75 = moderate reliability; 0.75–0.9 = good reliability; and >0.90 = excellent reliability[21]
FGS = Feline Grimace Scale

Discussion

In the present study, the validation of the FGS in kittens is reported by evaluating its construct validity, responsiveness and reliability using image assessment. Overall, the results indicated that the inter-rater reliability was moderate using ICC_{single} and excellent using $ICC_{average}$, whereas intra-rater reliability ranged from good to excellent for the FGS total ratio scores. Similar to what is observed in adult cats, the FGS total ratio scores increased after surgery and decreased after the administration of rescue analgesia.[5] These results are important for feline health and welfare, as the FGS is now the first acute pain assessment tool with reported validity specifically in kittens. With the widespread adoption of early-age spaying in cats, the tool can now be used in high-volume, high-quality spay/neuter programs to ensure pain recognition and treatment based on a more objective assessment.

Construct validity was assessed using known-groups discrimination by comparing the FGS total ratio scores before surgery (non-painful) vs after surgery (painful). The hypothesis was that kittens before surgery would have lower FGS total ratio scores than after surgery. Comparisons were limited to baseline vs 1 and 2 h, when acute pain is suspected to be most significant after OVH and when rescue analgesia is commonly required based on our previous studies. [5, 6, 8, 18] The FGS discriminated non-painful and painful kittens, as scores were significantly increased at 1 and 2 h when compared with baseline. Similar results were observed in adult cats in a previous study: client-owned cats with painful conditions had higher FGS scores than non-painful control cats. [5]

Responsiveness consists of the ability of an instrument to detect clinically important changes over time in the construct measured. [7, 19, 20] In this study, the discrimination between the FGS total ratio scores before and 1h after analgesic intervention was assessed. The intervention in this study consisted of administration of buprenorphine with or without meloxicam for rescue analgesia when kittens received UFEPS-SF pain scores ≥4/12. Significant decreases in FGS total ratio scores were observed 1 h after the administration of rescue analgesia to kittens in pain. Thus, the FGS is responsive to analgesic treatment in kittens undergoing OVH. Similar results were observed in adult cats with medical or surgical pain after the administration of analgesia. [5, 6]

Reliability describes the capacity of measurements to be inherently reproducible. [21] The measurement can be tested between raters (inter-rater reliability) and over time for the same rater (intra-rater reliability) using the same outcomes (ie, images, videos, etc). The inter-rater reliability (ICC_{single}) of the FGS total ratio scores was moderate in this study. Previous studies showed good inter-rater reliability using the FGS in adult cats.[5, 8, 10] The reliability of whiskers position was relatively lower compared with the other AUs, which is also similar to previous findings.[5,8] The interpretation of this AU could be difficult when assessment is performed using an image, as it can be affected by camera position, lighting and image quality. It is also important to note that the 95% CI (ie, 95% chance that the true ICC value will be at any point within the CI for that AU) includes values for which the ICC interpretation could change. In addition, the $ICC_{average}$, an index for the reliability of mean of k raters, revealed moderate inter-rater reliability for whiskers position, good inter-rater reliability for ear position, muzzle tension and head position, and excellent inter-rater reliability for orbital tightening and FGS total ratio scores, according to the classification used. Comparisons among studies are not easy as different interpretations and calculations can be used for ICC, and it is not always known what ICC type (ie, single or average) is actually reported in studies.[7] In our study, ICC calculations were performed using two-way random effect models because the results can be generalized to any raters who possess the same characteristics as the selected raters in our study (ie, veterinarians). Absolute agreement, and not consistency, was selected for the ICC definition, as it is not expected that the raters' scores would be correlated in an additive manner.[17,22] The $ICC_{average}$ is usually higher than the ICC_{single} and it can often be used to 'inflate the results'. However, the ICC_{single} is best selected when the basis of the actual clinical measurement (ie, FGS scoring) involves a single rater during its application.[17] Finally, our classification for ICC values is more conservative than the one used by Altman (<0.2 = poor; 0.21–0.4 = reasonable; 0.41–0.60 = moderate; 0.61–0.80 = good; and

0.81–1.0 = very good),[23] which may contribute to different interpretations of our results and the use of the FGS in the clinical setting.[9]

The intra-rater reliability of the FGS total ratio scores in kittens was good to excellent, which was similar to previous studies with images from adult cats, when ICCs from the FGS total ratio scores from veterinarians were excellent.[5,10] Orbital tightening and ear position ranged from moderate to excellent, muzzle tension ranged from moderate to good, whiskers position was moderate and head position ranged from good to excellent. The intra-rater reliability of veterinarians for individual AUs was evaluated in adult cats using the same method for classification.[10,17] The ICC classification between this previous and the present study was similar for all AUs, except for head position, which ranged from good to excellent in the present study and was moderate in the previous study. While knowing the inter- and intra-rater reliability of each AU is important in research, the FGS score is considered exclusively in the clinical setting as it informs the need for analgesic treatment based on the interventional score. Similar results were reported in adult cats for the FGS total ratio scores and for each AU, but with altogether higher reliability than in kittens.[5, 8, 10] This was surprising as raters participated in a training session before the study began. As much as the effects of training on FGS scoring are not known, we expected that our results would be similar to previous studies using the FGS in adult cats. This difference could be because raters may be less familiar with kittens than with adult cats.

The present study has some limitations. Image quality could still be improved. Despite efforts to add external lighting, a similar background and to use a better camera than previous FGS studies,[5,8] the process of image selection and cropping still led to the exclusion of many images due to poor quality. Suboptimal images may have had greater impact when scoring fine

AUs, such as whiskers position and muzzle tension, than when scoring head position, ear position and orbital tightening; the latter may be easier to discern than the former ones. It is important to consider that the construct validity, responsiveness and reliability of the FGS were studied only in female kittens undergoing OVH. As much as this is certainly the target population considering early-age spaying practices, it is not known if similar results would have been obtained with medical pain, other types of surgeries or trauma, and with kittens presenting feral or shy behaviors. The exclusion of behavior-specific populations represents a potential bias to the validation and application of the FGS in kittens, as spay/neuter programs may include individuals with all types of behavioral traits. Responsiveness of the FGS was not tested for control, non-painful kittens, to demonstrate that scores would not change over time for these individuals in a hospital setting. On the other hand, data were robust, including a heterogenous set of images of both painful and pain-free kittens for comparisons.

Conclusion

The FGS demonstrated high discriminative ability and responsiveness to an analgesic treatment with overall moderate inter-rater reliability, and good to excellent intra-rater reliability, in kittens. The FGS is a valid and reliable instrument for acute pain assessment in kittens, representing a substantial improvement for feline pain management in the context of spay/neuter programs.

Acknowledgements

We are thankful to all co-authors of the clinical trial that involved video recordings and Sean MacRae for manuscript revision and plotting.

Authors' note

An abstract of this study was part of an oral presentation at the 2022 Nafplio Association of Veterinary Anaesthetists Spring Meeting. This manuscript represents a portion of a MSc degree thesis of the first author to the Université de Montreal.

Supplementary material

The following files are available as supplementary material:

Supplementary material video

ARRIVE Checklist.

Conflict of interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Paulo V Steagall's laboratory is funded by a Discovery Grant of the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-03831).

Ethical Approval

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

Informed Consent

Informed consent (verbal or written) was obtained from the owner or legal custodian of all animal(s) described in this work (experimental or non-experimental animals, including cadavers) for all procedure(s) undertaken (prospective or retrospective studies). For any animals or people individually identifiable within this publication, informed consent (verbal or written) for their use in the publication was obtained from the people involved.

References

1. Steagall PV and Monteiro BP. Acute pain in cats: Recent advances in clinical assessment. *Journal of Feline Medicine and Surgery* 2019; 21: 25-34. DOI: 10.1177/1098612x18808103.
2. Reid J, Scott EM, Calvo G, et al. Definitive Glasgow acute pain scale for cats: validation and intervention level. *Vet Rec* 2017; 180: 449. 20170127. DOI: 10.1136/vr.104208.
3. Belli M, de Oliveira AR, de Lima MT, et al. Clinical validation of the short and long UNESP-Botucatu scales for feline pain assessment. *PeerJ* 2021; 9: e11225. 20210412. DOI: 10.7717/peerj.11225.
4. Luna SPL, Trindade PHE, Monteiro BP, et al. Multilingual validation of the short form of the Unesp-Botucatu Feline Pain Scale (UFEPS-SF). *PeerJ* 2022; 10: e13134. 20220323. DOI: 10.7717/peerj.13134.
5. Evangelista MC, Watanabe R, Leung VSY, et al. Facial expressions of pain in cats: the development and validation of a Feline Grimace Scale. *Sci Rep* 2019; 9: 19128. 20191213. DOI: 10.1038/s41598-019-55693-8.
6. Evangelista MC, Benito J, Monteiro BP, et al. Clinical applicability of the Feline Grimace Scale: real-time versus image scoring and the influence of sedation and surgery. *PeerJ* 2020; 8: e8967. 20200414. DOI: 10.7717/peerj.8967.
7. Evangelista MC, Monteiro BP and Steagall PV. Measurement properties of grimace scales for pain assessment in nonhuman mammals: a systematic review. *Pain* 2022; 163: e697-e714. 20210909. DOI: 10.1097/j.pain.0000000000002474.
8. Watanabe R, Doodnaught GM, Evangelista MC, et al. Inter-Rater Reliability of the Feline Grimace Scale in Cats Undergoing Dental Extractions. *Frontiers in Veterinary Science* 2020; 7. Brief Research Report. DOI: 10.3389/fvets.2020.00302.
9. Monteiro BP, Lee NH and Steagall PV. Can cat caregivers reliably assess acute pain in cats using the Feline Grimace Scale? A large bilingual global survey. *J Feline Med Surg* 2023; 25: 1098612x221145499. DOI: 10.1177/1098612x221145499.
10. Evangelista MC and Steagall PV. Agreement and reliability of the Feline Grimace Scale among cat owners, veterinarians, veterinary students and nurses. *Scientific Reports* 2021; 11: 5262. DOI: 10.1038/s41598-021-84696-7.

11. Sparkes AH, Bessant C, Cope K, et al. ISFM guidelines on population management and welfare of unowned domestic cats (Felis catus). *J Feline Med Surg* 2013; 15: 811-817. DOI: 10.1177/1098612x13500431.

12. Joyce A and Yates D. Help stop teenage pregnancy! Early-age neutering in cats. *J Feline Med Surg* 2011; 13: 3-10. DOI: 10.1016/j.jfms.2010.11.005.

13. Malo A, Cheng AJ, Ruel HL, et al. Randomised, prospective, blinded, clinical trial of opioid-free injectable anaesthesia with or without multimodal analgesia in kittens undergoing ovariohysterectomy. *Journal of Feline Medicine and Surgery* 2023; 25: 1098612X231158582. DOI: 10.1177/1098612x231158582.

14. Percie du Sert N, Hurst V, Ahluwalia A, et al. The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *PLoS Biol* 2020; 18: e3000410. 20200714. DOI: 10.1371/journal.pbio.3000410.

15. Rodan I, Dowgray N, Carney HC, et al. 2022 AAFP/ISFM Cat Friendly Veterinary Interaction Guidelines: Approach and Handling Techniques. *J Feline Med Surg* 2022; 24: 1093-1132. DOI: 10.1177/1098612x221128760.

16. Rodan I, Sundahl E, Carney H, et al. AAFP and ISFM feline-friendly handling guidelines. *J Feline Med Surg* 2011; 13: 364-375. DOI: 10.1016/j.jfms.2011.03.012.

17. Koo TK and Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* 2016; 15: 155-163. 20160331. DOI: 10.1016/j.jcm.2016.02.012.

18. Grint NJ, Murison PJ, Coe RJ, et al. Assessment of the influence of surgical technique on postoperative pain and wound tenderness in cats following ovariohysterectomy. *Journal of Feline Medicine & Surgery* 2006; 8: 15-21. DOI: https://doi.org/10.1016/j.jfms.2005.06.002.

19. Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: a critical review and recommendations. *Journal of Clinical Epidemiology* 2000; 53: 459-468. DOI: https://doi.org/10.1016/S0895-4356(99)00206-1.

20. Guyatt GH, Deyo RA, Charlson M, et al. Responsiveness and validity in health status measurement: A clarification. *Journal of Clinical Epidemiology* 1989; 42: 403-408. DOI: https://doi.org/10.1016/0895-4356(89)90128-5.

21. Lachin JM. The role of measurement reliability in clinical trials. *Clin Trials* 2004; 1: 553-566. DOI: 10.1191/1740774504cn057oa.

22. McGraw KO and Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1996; 1: 30-46. DOI: 10.1037/1082-989X.1.1.30.

23. Altman DG. *Practical statistics for medical research*. 1st ed. London: Chapman and Hall London, 1991.

Chapter 3 – General Discussion

Summary and Impact

The construct validity, reliability, and responsiveness of the FGS were successfully evaluated in kittens undergoing an OVH, in line with the objectives and hypothesis of this work. Construct validity was assessed by comparing the FGS total ratio scores of kittens at baseline to after OVH, a pain-inducing event. A significant increase in pain scores was observed at 1 and 2 h after OVH, thus supporting the construct validity of the FGS in kittens. Reliability was evaluated by comparing the FGS total ratio scores between raters at one time (inter-rater reliability) and of the same rater over time (intra-rater reliability). Moderate inter-rater reliability and good to excellent intra-rater reliability of the FGS total ratio scores were observed. Inter-rater reliability specific to each AU ranged from poor to moderate. Intra-rater reliability specific to each AU ranged from moderate to excellent. Responsiveness was evaluated by comparing the FGS total ratio scores of painful kittens postoperatively to 1 h after the administration of rescue analgesia. A significant decrease was observed after analgesic treatment was given, thus supporting the responsiveness of the FGS in kittens.

The consequences of a successful FGS for pain in kittens are several. First, better pain detection improves client trust in the veterinary care institution. More adherent patients who regularly attend their routine appointments improve the chances of early diagnosis of painful conditions. The early detection of pain then increases the odds of catching a disease process in its acute phase where treatment has a greater chance of preventing progression, pain chronification, and delayed healing. At the same time, a positive and collaborative owner-animal-veterinarian environment is promoted. From a clinic bottom-line perspective, happy and regular clients also equate to more compliant follow-up care.

Second, the FGS is especially useful for large-scale spay-neuter programs where minimal patient follow-up can be done due to the impractical logistics related to their setting and size. The FGS is a rapid, simple, valid, and effective pain assessment tool that can be used prior to release to ensure kitten welfare.

Third and on a more philosophical level, two core principles at the heart of medical ethics are to maximize beneficence and minimize maleficence. In other words, the goal of care should always be to provide the highest patient benefit with the least suffering. It is our ethical duty as clinicians to offer the best care to the extent possible in order to maximize quality of life.

“A veterinarian should be dedicated to providing competent veterinary medical care, with compassion and respect for animal welfare and human health. [...]” (Bourque & Horney, 2016).

Limitations

The FGS has some important limitations, both intrinsic to its use as a pain scale and extrinsically due to factors outside reasonable control. First, the application of the FGS to individual extremes in appearance in the kitten population is an issue. In the seminal FGS study, black cats were excluded due to the generally poor image quality, including low contrast and brightness, associated with them. The consequence was increased difficulty in identifying facial landmarks for the AUs comprising the FGS measurements. A similar limitation was noted in dark-coated horses when evaluating the Horse Grimace Scale (Dalla Costa et al., 2014). No such exclusion was done in this experiment, which could explain the slightly lower reliability encountered.

In addition, brachycephalic feline breeds, such as Persian and Himalayan, were also excluded from the original FGS study’s final analysis. In this present study, no brachycephalic or breed-specific kittens were included due to their low yield in shelters at that age. Though breed-specific facial characteristics have been documented in prior studies (Künzel et al., 2003), their impact on the FGS has not been investigated. Morphological differences like round-shaped skulls and flatter nasal bones could directly affect the evaluation of AUs and be a source of bias in the application of the FGS. Brachycephalic cat breeds are not uncommon in clinical settings and are just as at risk of pain as domestic short-haired and long-haired cats.

Second, kittens do differ anatomically from adult cats. The slight differences could explain the lower reliability for this population. Facial characteristics are known to change as individuals grow. A kitten at six months old, denoting the maximum inclusion age for this experiment, is often closer to its adult appearance than it is at 10 weeks old, representing the minimum inclusion age. The mean age for the study's population was 15 ± 4.7 weeks, which is significantly closer to the latter situation following this logic. We wonder if there could have been further differences among kittens of intermediate age within the 10 weeks to 6 months range that affected the psychometric evaluation of the AUs.

Furthermore, beyond the direct impact on AUs, anatomical differences can influence the rater's degree of familiarity. The role of experience and training programs in pain assessment have been documented in prior studies (Benito et al., 2017; Doodnaught et al., 2017; Mich et al., 2010), but not investigated for the FGS. The kitten life stage is a rapidly changing and ephemeral period when compared to the whole longevity of a cat. Although many routine visits are recommended during this period, veterinarians still examine kittens less commonly overall than adult cats, as the former are a small proportion of total clinic cat load. Lesser familiarity with pain evaluation in kittens could explain why the training session among the four raters had less than the desired effect.

Future Work

Since the publication of the FGS, multiple studies have evaluated the application of this pain assessment instrument. Despite its validity and reliability in different painful medical conditions, and its simplicity of use adapted to clinical settings, further studies remain necessary as some of its limitations are concerning.

First, it is important to further investigate validity and reliability in different cat subpopulations. As previously explained with black coloured cats and brachycephalic feline breeds, the application of the FGS to these cats might be problematic. In a clinical setting, the potential for encountering bias could be detrimental to the pain management and welfare of these subpopulations.

Second, with the new development of automated FGS pain assessment using a deep learning model, future work will be necessary to establish its accuracy in comparison to a human evaluator. If such differences exist, an ethical debate about who to trust between human and machine will be salient.

Chapter 4 – Conclusion

In conclusion, the objectives of the present work were achieved by demonstrating the validity, responsiveness, and inter- and intra-rater reliability of the FGS in kittens undergoing an elective OVH.

The results of this study are an important advancement in pain recognition and management for kittens. The FGS is the first acute pain assessment instrument validated as well as shown to be reliable and responsive in kittens undergoing an OVH. Indeed, despite commonly undergoing painful surgeries, no instrument specifically measured pain in this population until today. With the help of the FGS, veterinarians will be able to ensure appropriate pain management in kittens, from early on to all throughout their nine lives.

References

AAHA. (2018). *Pediatric neutering*. https://www.aaha.org/about-aaha/aaha-position-statements/pediatric-neutering/

Ahmed, I., & Ishtiaq, S. (2021). Reliability and validity: Importance in medical research. *Journal of the Pakistan Medical Association*, *71*(10), 2401-2406. https://doi.org/10.47391/jpma.06-861

Albuquerque, N., Guo, K., Wilkinson, A., Savalli, C., Otta, E., & Mills, D. (2016). Dogs recognize dog and human emotions. *Bioloy Letters*, *12*(1), 20150883. https://doi.org/10.1098/rsbl.2015.0883

ASPCA. (n.d.). *Companion animal birth control*. https://www.aspca.org/about-us/aspca-policy-and-position-statements/companion-animal-birth-control

AVMA. (n.d.). *AVMA policy: Pediatric spay/neuter of dogs and cats.* . https://www.avma.org/resources-tools/avma-policies/pediatric-spayneuter-dogs-and-cats

Belli, M., de Oliveira, A., de Lima, M., Trindade, P., Steagall, P., & Luna, S. (2021). Clinical validation of the short and long UNESP-Botucatu scales for feline pain assessment. *PeerJ*, *9*, e11225. https://doi.org/10.7717/peerj.11225

Bellieni, C. (2012). Pain assessment in human fetus and infants. *The AAPS Journal*, *14*(3), 456-461. https://doi.org/10.1208/s12248-012-9354-5

Benito, J., Monteiro, B., Beauchamp, G., Lascelles, B., & Steagall, P. (2017). Evaluation of interobserver agreement for postoperative pain and sedation assessment in cats. *Journal of American Veterinary Medical Association*, *251*(5), 544-551. https://doi.org/10.2460/javma.251.5.544

Beswick, A., Dewey, C., Johnson, R., Dowsett-Cooper, J., & Niel, L. (2016). Survey of Ontario veterinarians' knowledge and attitudes on pain in dogs and cats in 2012. *The Canadian Veterinary Journal*, *57*(12), 1274-1280. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5109631/

Booker, S., & Haedtke, C. (2016). Assessing pain in nonverbal older adults. *Nursing*, *46*(5), 66-69. https://doi.org/10.1097/01.Nurse.0000480619.08039.50

Bortolami, E., & Love, E. (2015). Practical use of opioids in cats: a State-of-the-art, evidence-based review. *Journal of Feline Medicine and Surgery*, *17*(4), 283-311. https://doi.org/10.1177/1098612x15572970

Bourque, T., & Horney, B. (2016). *Principles of veterinary medical ethics of the CVMA*. Canadian Veterinary Medical Association. https://www.canadianveterinarians.net/media/o5qjghc0/principles-of-veterinary-medical-ethics-of-the-cvma.pdf

Brondani, J., Mama, K., Luna, S., Wright, B., Niyom, S., Ambrosio, J., Vogel, P., & Padovani, C. (2013). Validation of the English version of the UNESP-Botucatu multidimensional composite pain scale for assessing postoperative pain in cats. *BMC Veterinary Research*, *9*(1), 143. https://doi.org/10.1186/1746-6148-9-143

CAHI. (2022). *Latest Canadian pet population figures released*. https://www.globenewswire.com/news-release/2022/09/22/2521210/0/en/Latest-Canadian-Pet-Population-Figures-Released.html

Calvo, G., Holden, E., Reid, J., Scott, E., Firth, A., Bell, A., Robertson, S., & Nolan, A. M. (2014). Development of a behaviour-based measurement tool with defined intervention level for assessing acute pain in cats. *Journal of Small Animal Practice*, *55*(12), 622-629. https://doi.org/10.1111/jsap.12280

Court, M. (2013). Feline drug metabolism and disposition: Pharmacokinetic evidence for species differences and molecular mechanisms. *Veterinary Clinics of North America: Small Animal Practice*, *43*(5), 1039-1054. https://doi.org/10.1016/j.cvsm.2013.05.002

CVMA. (2004). *The Canadian veterinary oath*. https://www.canadianveterinarians.net/about-cvma/the-canadian-veterinary-oath/

CVMA. (2019). *Neutering of dogs and cats*. https://www.canadianveterinarians.net/policy-and-outreach/position-statements/statements/neutering-of-dogs-and-cats/

Cypress, B. (2017). Rigor or reliability and validity in qualitative research: Perspectives, strategies, reconceptualization, and recommendations. *Dimensions of Critical Care Nursing*, *36*(4), 253-263. https://doi.org/10.1097/dcc.0000000000000253

Dalla Costa, E., Minero, M., Lebelt, D., Stucke, D., Canali, E., & Leach, M. (2014). Development of the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine castration. *PLoS One*, *9*(3), e92281. https://doi.org/10.1371/journal.pone.0092281

Deak, B., Ostendorf, B., Taggart, D., Peacock, D., & Bardsley, D. (2019). The significance of social perceptions in implementing successful feral cat management strategies: A global review. *Animals*, *9*(9). https://doi.org/10.3390/ani9090617

DePoy, E., & Gitlin, L. (2016). Collecting data through measurement in experimental-type research. In *Introduction to Research* (5th ed., pp. 227-247). Mosby. https://doi.org/10.1016/B978-0-323-26171-5.00017-3

DeVon, H., Block, M., Moyle-Wright, P., Ernst, D., Hayden, S., Lazzara, D., Savoy, S., & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*, *39*(2), 155-164. https://doi.org/10.1111/j.1547-5069.2007.00161.x

Diep, T., Monteiro, B., Evangelista, M., Balleydier, A., Watanabe, R., Ruel, H., Doodnaught, G., Le Quang, T., & Steagall, P. (2020). Anesthetic and analgesic effects of an opioid-free, injectable protocol in cats undergoing ovariohysterectomy: A prospective, blinded, randomized clinical trial. *The Canadian Veterinary Journal*, *61*(6), 621-628. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7238485/

Dohoo, S., & Dohoo, I. (1996a). Factors influencing the postoperative use of analgesics in dogs and cats by Canadian veterinarians. *The Canadian Veterinary Journal*, *37*(9), 552-556. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1576376/

Dohoo, S., & Dohoo, I. (1996b). Postoperative use of analgesics in dogs and cats by Canadian veterinarians. *The Canadian Veterinary Journal*, *37*(9), 546-551. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1576375/

Doodnaught, G., Benito, J., Monteiro, B., Beauchamp, G., Grasso, S., & Steagall, P. (2017). Agreement among undergraduate and graduate veterinary students and veterinary anesthesiologists on pain assessment in cats and dogs: A preliminary study. *The Canadian*

Veterinary Journal, *58*(8), 805-808. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5508929/

Epstein, M., Rodan, I., Griffenhagen, G., Kadrlik, J., Petty, M., Robertson, S., & Simpson, W. (2015). 2015 AAHA/AAFP pain management guidelines for dogs and cats. *Journal of American Animal Hospital Association*, *51*(2), 67-84. https://doi.org/10.5326/jaaha-ms-7331

Evangelista, M., Benito, J., Monteiro, B., Watanabe, R., Doodnaught, G., Pang, D., & Steagall, P. (2020). Clinical applicability of the Feline Grimace Scale: real-time versus image scoring and the influence of sedation and surgery. *PeerJ*, *8*, e8967. https://doi.org/10.7717/peerj.8967

Evangelista, M., Monteiro, B., & Steagall, P. (2022). Measurement properties of grimace scales for pain assessment in nonhuman mammals: a systematic review. *Pain*, *163*(6), e697-e714. https://doi.org/10.1097/j.pain.0000000000002474

Evangelista, M., & Steagall, P. (2021). Agreement and reliability of the Feline Grimace Scale among cat owners, veterinarians, veterinary students and nurses. *Scientific Reports*, *11*(1), 5262. https://doi.org/10.1038/s41598-021-84696-7

Evangelista, M., Watanabe, R., Leung, V., Monteiro, B., O'Toole, E., Pang, D., & Steagall, P. (2019). Facial expressions of pain in cats: The development and validation of a Feline Grimace Scale. *Scientific Reports*, *9*(1), 19128. https://doi.org/10.1038/s41598-019-55693-8

Griffin, B., Bushby, P., McCobb, E., White, S., Rigdon-Brestle, Y., Appel, L., Makolinski, K., Wilford, C., Bohling, M., Eddlestone, S., Farrell, K., Ferguson, N., Harrison, K., Howe, L., Isaza, N., Levy, J., Looney, A., Moyer, M., Robertson, S., & Tyson, K. (2016). The Association of Shelter Veterinarians' 2016 veterinary medical care guidelines for spay-neuter programs. *Journal of American Veterinary Medical Association*, *249*(2), 165-188. https://doi.org/10.2460/javma.249.2.165

Gruen, M., Lascelles, B., Colleran, E., Gottlieb, A., Johnson, J., Lotsikas, P., Marcellin-Little, D., & Wright, B. (2022). 2022 AAHA pain management guidelines for dogs and cats. *Journal of American Animal Hospital Association*, *58*(2), 55-76. https://doi.org/10.5326/jaaha-ms-7292

Hernandez-Avalos, I., Mota-Rojas, D., Mora-Medina, P., Martínez-Burnes, J., Casas Alvarado, A., Verduzco-Mendoza, A., Lezama-García, K., & Olmos-Hernandez, A. (2019). Review of different methods used for clinical recognition and assessment of pain in dogs and cats. *International Journal of Veterinary Science and Medicine*, *7*(1), 43-54. https://doi.org/10.1080/23144599.2019.1680044

Hewson, C., Dohoo, I., & Lemke, K. (2006a). Factors affecting the use of postincisional analgesics in dogs and cats by Canadian veterinarians in 2001. *The Canadian Veterinary Journal*, *47*(5), 453-459. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1444907/

Hewson, C., Dohoo, I., & Lemke, K. (2006b). Perioperative use of analgesics in dogs and cats by Canadian veterinarians in 2001. *The Canadian Veterinary Journal*, *47*(4), 352-359. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1405826/

Howe, L. (2015). Current perspectives on the optimal age to spay/castrate dogs and cats. *Veterinary Medicine (Auckland, N.Z.)*, *6*, 171-180. https://doi.org/10.2147/VMRR.S53264

Hugonnard, M., Leblond, A., Keroack, S., Cadoré, J., & Troncy, E. (2004). Attitudes and concerns of French veterinarians towards pain and analgesia in dogs and cats. *Veterinary*

Anaesthesia and Analgesia, *31*(3), 154-163. https://doi.org/10.1111/j.1467-2987.2004.00175.x

Jensen, T., & Finnerup, N. (2014). Allodynia and hyperalgesia in neuropathic pain: Clinical manifestations and mechanisms. *Lancet Neurology*, *13*(9), 924-935. https://doi.org/10.1016/s1474-4422(14)70102-4

Joyce, A., & Yates, D. (2011). Help stop teenage pregnancy! Early-age neutering in cats. *Journal of Feline Medicine and Surgery*, *13*(1), 3-10. https://doi.org/10.1016/j.jfms.2010.11.005

Künzel, W., Breit, S., & Oppel, M. (2003). Morphometric investigations of breed-specific features in feline skulls and considerations on their functional implications. *Anatomia Histologia Embryologia*, *32*(4), 218-223. https://doi.org/10.1046/j.1439-0264.2003.00448.x

Lorena, S., Luna, S., Lascelles, B., & Corrente, J. (2014). Current attitudes regarding the use of perioperative analgesics in dogs and cats by Brazilian veterinarians. *Veterinary Anaesthesia and Analgesia*, *41*(1), 82-89. https://doi.org/10.1111/vaa.12104

Luna, S., Trindade, P., Monteiro, B., Crosignani, N., Della Rocca, G., Ruel, H., Yamashita, K., Kronen, P., Tseng, C., Teixeira, L., & Steagall, P. (2022). Multilingual validation of the short form of the Unesp-Botucatu Feline Pain Scale (UFEPS-SF). *PeerJ*, *10*, e13134. https://doi.org/10.7717/peerj.13134

Malo, A., Cheng, A., Ruel, H., Monteiro, B., Lutevele, N., Marangoni, S., Garbin, M., Watanabe, R., & Steagall, P. (2023). Randomised, prospective, blinded, clinical trial of opioid-free injectable anaesthesia with or without multimodal analgesia in kittens undergoing ovariohysterectomy. *Journal of Feline Medicine and Surgery*, *25*(3), 1098612X231158582. https://doi.org/10.1177/1098612x231158582

Mathews, K., Kronen, P., Lascelles, D., Nolan, A., Robertson, S., Steagall, P., Wright, B., & Yamashita, K. (2014). Guidelines for recognition, assessment and treatment of pain. *Journal of Small Animal Practice*, *55*(6), E10-68. https://doi.org/10.1111/jsap.12200

McGuire, D., Kaiser, K., Haisfield-Wolfe, M., & Iyamu, F. (2016). Pain assessment in noncommunicative adult palliative care patients. *Nursing Clinics of North America*, *51*(3), 397-431. https://doi.org/10.1016/j.cnur.2016.05.009

Menéndez, S., Cabezas, M., & Gomez de Segura, I. (2023). Attitudes to acute pain and the use of pain assessment scales among Spanish small animal veterinarians. *Frontiers in Veterinary Science*, *10*. https://doi.org/10.3389/fvets.2023.1302528

Mich, P., Hellyer, P., Kogan, L., & Schoenfeld-Tacher, R. (2010). Effects of a pilot training program on veterinary students' pain knowledge, attitude, and assessment skills. *The Journal of Veterinary Medical Education*, *37*(4), 358-368. https://doi.org/10.3138/jvme.37.4.358

Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P., Knol, D., Bouter, L., & de Vet, H. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, *63*(7), 737-745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Monteiro, B., Lascelles, B., Murrell, J., Robertson, S., Steagall, P., & Wright, B. (2023). 2022 WSAVA guidelines for the recognition, assessment and treatment of pain. *Journal of Small Animal Practice*, *64*(4), 177-254. https://doi.org/10.1111/jsap.13566

Monteiro, B., Lee, N., & Steagall, P. (2023). Can cat caregivers reliably assess acute pain in cats using the Feline Grimace Scale? A large bilingual global survey. *Journal of Feline Medicine and Surgery*, *25*(1), 1098612x221145499. https://doi.org/10.1177/1098612x221145499

Monteiro, B., & Steagall, P. (2019a). Antiinflammatory drugs. *Veterinary Clinics of North America: Small Animal Practice*, *49*(6), 993-1011. https://doi.org/10.1016/j.cvsm.2019.07.009

Monteiro, B., & Steagall, P. (2019b). Chronic pain in cats: Recent advances in clinical assessment. *Journal of Feline Medicine and Surgery*, *21*(7), 601-614. https://doi.org/10.1177/1098612x19856179

Moons, C., Valcke, A., Verschueren, K., Porters, N., Polis, I., & de Rooster, H. (2018). Effect of early-age gonadectomy on behavior in adopted shelter kittens—The sequel. *Journal of Veterinary Behavior*, *26*, 43-47. https://doi.org/10.1016/j.jveb.2018.04.001

NewMetrica. (2019). *Acute pain measurement*. https://www.newmetrica.com/acute-pain-measurement/

Olson, P., Kustritz, M., & Johnston, S. (2001). Early-age neutering of dogs and cats in the United States (a review). *Journal of Reproductive and Fertility*, *57*, 223-232. https://www.researchgate.net/publication/368246778_Early_age_neutering_in_dogs_and_cats

Perri, A., Feuerborn, T., Frantz, L., Larson, G., Malhi, R., Meltzer, D., & Witt, K. (2021). Dog domestication and the dual dispersal of people and dogs into the Americas. *The Proceedings of the National Academy of Sciences*, *118*(6). https://doi.org/10.1073/pnas.2010083118

Polson, S., Taylor, P., & Yates, D. (2014). Effects of age and reproductive status on postoperative pain after routine ovariohysterectomy in cats. *Journal of Feline Medicine and Surgery*, *16*(2), 170-176. https://doi.org/10.1177/1098612x13503651

Porters, N., de Rooster, H., Moons, C., Duchateau, L., Goethals, K., Bosmans, T., & Polis, I. (2015). Prepubertal gonadectomy in cats: Different injectable anaesthetic combinations and comparison with gonadectomy at traditional age. *Journal of Feline Medicine and Surgery*, *17*(6), 458-467. https://doi.org/10.1177/1098612x14546919

Raekallio, M., Heinonen, K., Kuussaari, J., & Vainio, O. (2003). Pain alleviation in animals: Attitudes and practices of Finnish veterinarians. *The Veterinary Journal*, *165*(2), 131-135. https://doi.org/10.1016/S1090-0233(02)00186-7

Raja, S., Carr, D., Cohen, M., Finnerup, N., Flor, H., Gibson, S., Keefe, F., Mogil, J., Ringkamp, M., Sluka, K., Song, X., Stevens, B., Sullivan, M., Tutelman, P., Ushida, T., & Vader, K. (2020). The revised International Association for the Study of Pain definition of pain: Concepts, challenges, and compromises. *Pain*, *161*(9), 1976-1982. https://doi.org/10.1097/j.pain.0000000000001939

Reid, J., Nolan, A., & Scott, E. (2018). Measuring pain in dogs and cats using structured behavioural observation. *The Veterinary Journal*, *236*, 72-79. https://doi.org/10.1016/j.tvjl.2018.04.013

Reid, J., Scott, E., Calvo, G., & Nolan, A. (2017). Definitive Glasgow acute pain scale for cats: Validation and intervention level. *Veterinary Record*, *180*(18), 449. https://doi.org/10.1136/vr.104208

Reimann, J., Dewey, C., Bateman, S., Kerr, C., & Johnson, R. (2017). Perioperative analgesic use by Ontario veterinarians, 2012. *The Canadian Veterinary Journal*, *58*(2), 149-156. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5234314/

Roach, K. (2006). Measurement of health outcomes: Reliability, validity and responsiveness. *Journal of Prosthetics and Orthotics*, *18*, P8-P12. https://doi.org/10.1097/00008526-200601001-00003

Ryan, S., Bacon, H., Endenburg, N., Hazel, S., Jouppi, R., Lee, N., Seksel, K., & Takashima, G. (2019). WSAVA animal welfare guidelines. *Journal of Small Animal Practice*, *60*(5), E1-e46. https://doi.org/10.1111/jsap.12998

Simon, B., & Steagall, P. (2017). The present and future of opioid analgesics in small animal practice. *Journal of Veterinary Pharmacology and Therapeutics*, *40*(4), 315-326. https://doi.org/10.1111/jvp.12377

Sneddon, L., Elwood, R., Adamo, S., & Leach, M. (2014). Defining and assessing animal pain. *Animal Behaviour*, *97*, 201-212. https://doi.org/10.1016/j.anbehav.2014.09.007

Souza, A., Alexandre, N., & Guirardello, E. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiology and Health Services*, *26*(3), 649-659. https://doi.org/10.5123/s1679-49742017000300022 (Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade.)

Sparkes, A., Bessant, C., Cope, K., Ellis, S., Finka, L., Halls, V., Hiestand, K., Horsford, K., Laurence, C., MacFarlaine, I., Neville, P., Stavisky, J., & Yeates, J. (2013). ISFM guidelines on population management and welfare of unowned domestic cats (Felis catus). *Journal of Feline Medicine and Surgery*, *15*(9), 811-817. https://doi.org/10.1177/1098612x13500431

Steagall, P. (2020). Analgesia: What makes cats different/challenging and what is critical for cats? *Veterinary Clinics: Small Animal Practice*, *50*(4), 749-767. https://doi.org/10.1016/j.cvsm.2020.02.002

Steagall, P., Monteiro, B., Marangoni, S., Moussa, M., & Sautié, M. (2023). Fully automated deep learning models with smartphone applicability for prediction of pain using the Feline Grimace Scale. *Scientific Reports*, *13*(1), 21584. https://doi.org/10.1038/s41598-023-49031-2

Steagall, P., Monteiro, B., & Taylor, P. (2014). A review of the studies using buprenorphine in cats. *Journal of Veterinary Internal Medicine*, *28*(3), 762-770. https://doi.org/10.1111/jvim.12346

Steagall, P., Robertson, S., Simon, B., Warne, L., Shilo-Benjamini, Y., & Taylor, S. (2022). 2022 ISFM consensus guidelines on the management of acute pain in cats. *Journal of Feline Medicine and Surgery*, *24*(1), 4-30. https://doi.org/10.1177/1098612x211066268

Tancredi, D., & Cardinali, I. (2023). Being a dog: A review of the domestication process. *Genes*, *14*(5). https://doi.org/10.3390/genes14050992

Taylor, P., & Robertson, S. (2004). Pain management in cats--past, present and future. Part 1. The cat is unique. *Journal of Feline Medicine and Surgery*, *6*(5), 313-320. https://doi.org/10.1016/j.jfms.2003.10.003

Väisänen, M., Tuomikoski, S., & Vainio, O. (2007). Behavioral alterations and severity of pain in cats recovering at home following elective ovariohysterectomy or castration. *Journal of American Veterinary Medical Association*, *231*(2), 236-242. https://doi.org/10.2460/javma.231.2.236

Valverde, A., & Skelding, A. (2019). Alternatives to opioid analgesia in small animal anesthesia: Alpha-2 agonists. *Veterinary Clinics North America: Small Animal Practice*, *49*(6), 1013-1027. https://doi.org/10.1016/j.cvsm.2019.07.010

Watanabe, R., Doodnaught, G., Evangelista, M., Monteiro, B., Ruel, H., & Steagall, P. (2020). Inter-rater reliability of the Feline Grimace Scale in cats undergoing dental extractions. *Frontiers in Veterinary Science*, *7*. https://doi.org/10.3389/fvets.2020.00302

Williams, V., Lascelles, B., & Robson, M. (2005). Current attitudes to, and use of, peri-operative analgesia in dogs and cats by veterinarians in New Zealand. *New Zealand Veterinary Journal*, *53*(3), 193-202. https://doi.org/10.1080/00480169.2005.36504

Witt, N., Coynor, S., Edwards, C., & Bradshaw, H. (2016). A guide to pain assessment and management in the neonate. *Current Emergency Hospital Medicine Reports*, *4*, 1-10. https://doi.org/10.1007/s40138-016-0089-y

Yam, M., Loh, Y., Tan, C., Adam, S., Manan, N., & Basir, R. (2018). General pathways of pain sensation and the major neurotransmitters involved in pain regulation. *International Journal of Molecular Sciences*, *19*(8). https://doi.org/10.3390/ijms19082164

Young, M. (1985). The evolution of domestic pets and companion animals. *Veterinary Clinics of North America: Small Animal Practice*, *15*(2), 297-309. https://doi.org/https://doi.org/10.1016/S0195-5616(85)50302-2