The Veterinary Journal 236 (2018) 72–79

Contents lists available at ScienceDirect

The Veterinary Journal

journal homepage: www.elsevier.com/locate/tvjl

Invited review

Measuring pain in dogs and cats using structured behavioural observation

J. Reid[a,*], A.M. Nolan[b], E.M. Scott[c]

[a] NewMetrica Ltd., 19 Woodside Place, Glasgow G3 7QL, UK
[b] Edinburgh Napier University, Sighthill Campus, Sighthill Court, EH 11 4BN, UK
[c] School of Mathematics and Statistics, 15 University Gardens, University of Glasgow, Gl2 8QW, UK

ARTICLE INFO

Article history:
Accepted 19 April 2018

Keywords:
Cat
Dog
Health-related quality of life
Measurement
Pain

ABSTRACT

The contemporary approach to pain measurement in people and animals seeks to measure the affective (emotional) component of the pain experience using structured questionnaires with formal scoring methodology. Chronic pain has wide-ranging impacts which affects the quality of life (QOL) of the individual, whether that is a person or an animal. Accordingly instruments to measure chronic pain are designed to measure its impact on QOL and are called health-related quality of life (HRQL) instruments. In veterinary science instruments to measure pain are based on behavioural observation by the veterinary surgeon/nurse in the case of acute pain and by the owner in the case of chronic pain. The development of HRQL instruments is an expanding field in veterinary science, not just for the measurement of pain, but for other chronic diseases, and it has a wide application in pharmaceutical research and clinical practice to improve patient care. This review highlights the challenges involved in creating such measures for dogs and cats, seeking to provide the reader with an understanding of their development process. It then provides an overview of the current status with regard to acute and chronic pain measurement.

Introduction

Our ability to measure pain in a valid and reliable way is a crucial component of effective pain management. Furthermore the current emphasis on evidence-based veterinary medicine requires that appropriate measures of clinical impact are developed and in that regard it is essential that instruments to monitor pain effectively in an individual, while providing data to enable the selection of treatments with known efficacy and impact are developed. In veterinary medicine, many pain scales have been constructed on an *ad hoc* basis, but the importance of applying rigorous methods to the development and testing of pain measures in order to ensure their validity and reliability is now recognised.

Pain

Pain is a multi-dimensional experience with sensory (discriminative), evaluative and affective (emotional) components; it is an abstract construct, like happiness, that is not directly measurable. Sensory and evaluative components tell us when and where the pain occurs, how intense it is and whether it is associated with a mechanical or thermal insult. Historically, intensity was the focus of pain measurement in human and veterinary medicine, but the contemporary approach to pain measurement focuses on the affective dimension which describes pain's unpleasantness, 'how it makes you feel'. The unpleasant feelings we experience cause the suffering we associate with pain. It has been suggested that a more comprehensive understanding of animal pain, and, in particular the affective component, may be of fundamental importance to the development of treatments for chronic and neuropathic pain (Flecknell et al., 2008).

If pain is considered to be an entirely subjective experience, valid measurement must attempt to access that subjective perception. Dawkins (2004) described animal behaviour as the expression of the emotions'. According to Griffin (1992) if we recognise that we can make useful and generally correct assumptions about the feelings of other people through observation of their behaviour, then it is clear that animal behaviour could equally be used to provide evidence for their mental experiences. In support of this, work has shown that inexperienced raters may be able to identify, with good agreement, subjective states in pigs (Wemelsfelder et al., 2001) and personality traits in dogs (Svartberg and Forkman, 2002)

* Corresponding author.
E-mail address: jacky.reid@newmetrica.com (J. Reid).

https://doi.org/10.1016/j.tvjl.2018.04.013
1090-0233/

Similarly, observer assessments of cats' behaviour (e.g. aggressive, playful, sociable) were found mainly to be valid and reliable (Mendl and Harcourt, 2000). Additionally, work done by Wiseman-Orr et al. in which they conducted semi-structured interviews with owners showed that owners are capable of reporting behavioural styles that signified the 'hidden' emotional or subjective states of their dogs, and could identify degrees of and changes in these subjective states (Wiseman-Orr et al., 2004). In addition researchers have suggested that in the course of their domestication dogs, which have evolved along with man over many thousands of years, have been selected for certain social-cognitive abilities that expedite their communication with humans, so that dogs can interpret human social cues and react accordingly with signals that humans are able to interpret (Soproni et al., 2001; Albuquerque et al., 2016; Nagasawa et al., 2009). This ability for man and dog to communicate means that the dog is a good candidate for the development of instruments to measure pain that depend upon subjective judgement. In contrast the cat has a more independent character and unlike the dog, has not been genetically selected as a companion for man or to work with man, making pain measurement potentially more challenging in this species.

Measuring pain

Given the longstanding focus on measuring only the intensity of pain it is hardly surprising that the use of objective measures has been extensively investigated over the years (Morton and Griffiths, 1985; Chapman, 1989; Bateson, 1991). Although these are useful in experimental situations, their use in the clinical situation has been disappointing. Apart from the need for equipment which may not be to hand in a veterinary practice, measures such as heart rate, respiratory rate and pupil dilation have been demonstrated to be unreliable in the dog in a clinical setting (Holton et al., 1998a, 1998b) and it is generally accepted that changes in hormonal markers such as cortisol are not specific for pain (Morton and Griffiths, 1985). However more recently attention has focused on force plate/gait analysis and activity monitors as objective measures of chronic orthopaedic pain with evidence of their usefulness in certain circumstances (Klinck et al., 2017).

The end product of the development of pain measures is often described an as instrument, typically in the form of a structured questionnaire. Instruments that measure how people feel are increasingly valued as outcome measures in human medicine (Emery et al., 2005) and are likely to become so in veterinary medicine as suitable instruments become available. The purpose of this review is to provide the reader with an appreciation of the significant challenges involved in the creation of such pain measures for animals and an understanding of their development with special reference to the dog and cat, so that when faced with a choice of instruments to choose from, veterinary surgeons can make a value judgement as to which to choose for their purpose, be that for clinical monitoring or for research.

The challenges of measuring pain

Species and breed differences

Behavioural disturbances have been recognised as potential indicators of pain in animals for many years (Morton and Griffiths, 1985; Mathews, 2000; Rutherford, 2002; Wiseman et al., 2001). According to Wiseman-Orr (2005) these include changes in demeanour, aggressiveness, submissiveness, fearfulness, restlessness, lethargy, activity, inquisitiveness, vocalisation, self-mutilation, appetite, drinking, urination, grooming and social behaviour. Nevertheless each species will manifest its own pain-related behaviours or behavioural disturbances which are unique, which means that they cannot be applied to another species. Breed differences will also affect pain behaviour. Accordingly instruments to measure pain must be developed for individual species (and sometimes these must be condition-specific) and, in some cases, must be sensitive to individual differences.

Use of a proxy

Self-report is the 'gold standard' in assessing human pain, but some people are unable to self-report effectively, for example, infants and those with cognitive impairment. In such cases carers and parents have been used as observers, in order to recognise and interpret behavioural indicators of pain (McGrath et al., 1998; van Dijk et al., 2000; Kappesser and Williams, 2002; Prkachin et al., 2002; Stallard et al., 2002). Similarly in veterinary practice reporting of acute pain-associated behavioural changes by a veterinary surgeon or nurse has been the focus of acute pain assessment research in dogs and cats (Firth and Haldane, 1999; Reid et al., 2007; Holopherne-Doran et al., 2010; Brondani et al., 2011, 2013; Reid et al., 2017a, 2017b). However in canine chronic pain, recent studies have highlighted that the owner is the preferred proxy rater because behavioural changes may be so subtle and gradual in onset that they are apparent only to someone very familiar with the individual animal (Flecknell, 1986). Similarly with chronic pain compared with acute pain, these subtle behavioural changes may not be obvious in a clinical setting where they may be masked by fear, excitement or anxiety associated with the unfamiliar environment. Inter-observer variability is not an issue when the owner is the only person consistently completing the assessment.

Respondent bias

Respondent bias is a threat to the valid measurement of pain by a proxy. For example, if scoring a dog with a cruciate repair shortly after surgery, a veterinary practitioner who has him/herself undergone similar surgery may be affected by that experience and score the pain more severely than him/herself would otherwise have done. This is compounded in chronic pain by the complexities of the human animal bond. Consciously or unconsciously an owner may bias their responses for several reasons, including fear of the veterinary surgeon suggesting euthanasia. Respondent bias can be reduced with comparative methods which use expert judgement to scale the value of each item response in advance. The result of this is that when the questionnaire instrument is used, the extent to which each response option represents the 'right' or 'wrong' answer is to some extent hidden from the respondent, thus making biased responding more difficult (McColl et al., 2001).

Developing an instrument that is scientifically robust and fit for purpose

This is without doubt the greatest challenge of all, regardless of whether the instrument measures pain in animals or man. Our medical colleagues have addressed this challenge over the last two decades by using psychometric methods which were first used in psychiatry to measure intangible constructs, such as anxiety and depression, using formally-assessed structured questionnaires. The methods used to create psychometric instruments are well established and comprise three phases (Streiner and Norman, 2008; Abell et al., 2009).

Phase 1 involves the identification of measurement objectives, classification of the target population, and the development of an initial collection of items for possible inclusion in the instrument. In phase 2, appropriate items are selected from the initial item pool and these are then subjected to expert validation. An instrument is constructed using the validated items with consideration given to layout, response options for the items, instructions for use and other details of administration. The resulting prototype is pre-tested with a group of target respondents to ensure that they can use the instrument correctly and without difficulty. Phase 3 comprises a field-test of the instrument to assess its key psychometric properties — validity, reliability and responsiveness to clinical change over time, if the instrument is designed to measure such change.

Validity (criterion, content and construct)

Validity provides evidence that the instrument is measuring what it is intended to measure and is the most fundamental property of an instrument (Streiner, 1993; Jensen, 2003).

Criterion validity. Criterion validity is the agreement of a new instrument with some existing 'gold standard', but when that does not exist, evidence can be gathered to support concurrent criterion validity (comparison with a validated measure of a related construct) or predictive criterion validity where performance of the new measure successfully predicts that of a later measure.

Content validity. Content validity relates to the appropriateness and completeness of the items comprising the instrument and if the items cover all the relevant aspects being measured with no extraneous features included then that is evidence to support content validity. Traditionally established using subjective expert judgement, the quantification of content validity has been introduced in human medicine and the social sciences (Polit and Beck, 2006). Relevant experts are asked to rate the relevance and clarity of items using a rating scale, and those ratings are used to calculate a content validity index (CVI) for each individual item on the scale, providing objective information to guide researchers in revising, deleting, or substituting items. This approach has recently been described in veterinary medicine (Noble et al., 2018).

Construct validity. Construct validity can be assessed in several ways, one of which is factorial validity. Factor analysis is a statistical technique that identifies the correlations between responses to the items of an instrument, clustering these into a smaller group of 'factors' to produce a factor model. Factorial validity is demonstrated if an interpretable factor model fits the construct that the instrument was designed to measure (Feinstein, 1987; Nunnally and Bernstein, 1994; Floyd and Widaman, 1995; Johnston, 1997). In a 'known-groups' approach to construct validation, predictions are made about how scores obtained with the instrument will differ between groups, such as healthy and sick animals, or will reflect disease burden, and these predictions are tested (Streiner, 1993; Johnston, 1997). For example, by showing that pain scores rise and fall predictably over time following surgery (Morton et al., 2005).

Reliability

A reliable instrument will produce the same score when an unchanging subject is measured at two time points by the same observer (repeatability/intra-rater reliability), or when two people measure the same subject at one time (reproducibility/inter-rater reliability) (Streiner and Norman, 2008). Inter and intra-rater reliability are good ways to estimate reliability when the measure is an observation, but frequently another form of objective reliability testing called internal consistency, measured by Cronbach's Alpha, is used to assess the consistency of results across items in the questionnaire (Tavakol and Dennick, 2011). If an instrument is valid then it is likely also to be reliable, but it may be highly reliable yet lack validity because it is measuring something other than that which it was intended to measure (Fallowfield, 1990).

Responsiveness

Responsiveness in a clinical instrument is that property which makes sure that the instrument can detect differences in health status that are important to the clinician or to the patient and these need not be statistically significant. In that context, the nature of the measurement provided by an instrument is important and is determined by the construction of the response options to any item. Response options are an important consideration in instrument development since, if their answers to items in the questionnaire are likely to lie on a continuum, it is important that respondents have the opportunity to answer in this way to ensure minimum loss of information and to minimise error (Streiner and Norman, 2008). Response options to an item may be binary such as yes/no, or may be more complex, e.g. ordinal or Likert, or continuous (Streiner and Norman, 2008). The resulting measurement may have nominal, ordinal, interval or ratio scale properties. A nominal scale which simply records into which category a response falls provides the least information. Ordinal and interval level measurements are both practicable and desirable for the assessment of pain. An ordinal scale, such as a 0–10 numerical rating scale is rather general but may offer the precision required, although it is important to be aware that, if the ordered categories are broad, the scale's sensitivity and responsiveness to change may be compromised. An interval level scale, an example of which is a Celsius temperature scale as measured by a thermometer, is more demanding to create, but provides more precise measurement (Morton et al., 2005; Scott et al., 2007).

Utility

Utility refers to the instrument's 'user friendliness' which means that for the owner it must be short, quick and easy to complete and not require lengthy training. For the clinician it must be easy to administer, score and interpret. An instrument that is valid and reliable, but lacking in utility is of little use in the clinical arena (Teasdale and Jennett, 1974).

Acute pain

Historically simple unidimensional tools, namely the simple descriptive scale (SDS), numerical rating scale (NRS) and the visual analogue scale (VAS) have been employed in the measurement of the intensity of acute pain. Although these tools are still in use to a certain extent, Holton et al. (1998a, 1998b) showed that inter-observer variability was unacceptable when three and four veterinarians simultaneously scored pain on the day of and the day following surgery in the dog, using the SDS, NRS and VAS. Furthermore these scales are often not standardized and because they have a limited number of response options they provide inadequate information, especially when used to measure a complex construct like pain. In contrast, multi-item or composite scales assess different components or aspects of a construct, making these much more suitable for the measurement of pain. Currently available composite scales for dogs and cats are shown in Table 1.

In veterinary practice, the practical worth of a pain assessment instrument is markedly improved if the score can be linked to an intervention level which guides the user as to whether or not an animal requires analgesic treatment. The CMPS-SF for the dog, both CMPS-Feline scales and the UNESP-Botucatu multidimensional composite pain scale for the cat have an intervention level defined.

The 'Pain Face'

Although changes in behaviour have been recognised for many years as the mainstay of acute pain measurement in animals, more recently, facial expression has been described as a possible means

Table 1
Currently available instruments to measure acute pain in dogs and cats.

Scale	Target species	Behavioural observations	Physiological measurements	Validated	Intervention level derived	Ref.
University of Melbourne Pain Scale	Dog	Yes	Yes	Yes	No	Firth and Haldane (1999)
Glasgow Composite Measure Pain Scale	Dog	Yes	No	Yes	Yes	Reid et al. (2007)
CMPS – SF						http://www.newmetrica.com/acute-pain-measurement/
4AVet	Dog & cat	Yes	Yes	Yes	No	Holopherne-Doran et al. (2010)
Colorado State acute pain scale	Dog	Yes	No	No	No	http://www.vasg.org/pdfs/CSU_Acute_Pain_Scale_Canine.pdf
Colorado State acute pain scale	Cat	Yes	No	No	No	http://www.vasg.org/pdfs/CSU_Acute_Pain_Scale_Kitten.pdf
UNESP-Botucatu multidimensional composite pain scale	Cat	Yes	Yes	Yes	Yes	Brondani et al. (2011, 2013) http://www.animalpain.com.br/assets/upload/escala-en-us.pdf
Glasgow CMPS-Feline	Cat	Yes	No	Yes	Yes	Calvo et al., 2014
Definitive Glasgow CMPS-Feline	Cat	Yes	No	Yes	Yes	Reid et al. (2017a, 2017b) http://newmetrica.com/acute-pain-measurement/

of pain assessment in non-human animals. The mouse and rat grimace scales (MGS and RGS) (Langford et al., 2010; Sotocinal et al., 2011) are standardised facial coding systems developed by recording changes in facial expression after application of a noxious stimulus. Other scales have been developed in rabbits (RbGS) (Keating et al., 2012) and in horses (Dalla Costa et al., 2014). Furthermore facial components have been included in multidimensional pain measures for children where they have been combined with behavioural and physiological parameters (Stevens et al., 1996; Hand et al., 2010). Although at the time of writing the authors are unaware of this technology being used in the dog, facial expression has been incorporated in the CMPS-Feline with the effect of increasing the sensitivity of the behavioural component (Calvo et al., 2014; Holden et al., 2014; Reid et al., 2017a, 2017b).

Chronic pain

The complexity of the pain experience is even greater when the pain becomes chronic, because chronic pain in people interacts in a complex way with a person's emotional (social and psychological) and physical well-being. Consequently, many human chronic pain instruments are concerned primarily with the patient's quality of life (QOL). Quality of life is a widely used term in which it is accepted that QOL is, like pain, a multi-dimensional construct that is a uniquely personal emotional experience. Health-related quality of life (HRQL) is the term given to those aspects of QOL that change with ill-health and medical treatment.

Measuring HRQL

Health-related quality of life instruments are designed to measure chronic pain's wide-ranging impacts, and in doing so offer a holistic approach to measurement of the pain experience as well as treatment effects and side-effects. However, HRQL instruments measure the impact of chronic disease whether or not that is associated with pain and as early as 1993 social and medical scientists recognised the need for valid HRQL measures for use in clinical trials (Berzon et al., 1993). Since then a plethora of disease specific instruments have been developed for this purpose in a large number of human diseases and generic instruments have also featured heavily in human clinical trials. Furthermore assessing HRQL has become important for assessing the standard of care in general medical practice (Tian-hui and Lu, 2005). Indeed, more recently 'patient-centred healthcare' has highlighted that incorporating patients' needs and views into healthcare delivery is important (Rozenblum and Bates, 2013). Essential to that process is an understanding of the patient's subjective experience through prospective clinical comparative effectiveness research (CER), which according to the Federal Coordinating Council for Comparative Effectiveness Research[1] is 'the conduct and synthesis of research comparing the benefits and harms of different interventions and strategies to prevent, diagnose, treat, and monitor health conditions in 'real world' settings'. The purpose of CER is to assist multiple stakeholders including consumers, clinicians, purchasers, and those concerned with policy to make educated decisions regarding health care, to benefit both individuals and the general population (Dreyer et al., 2010).

Pivotal to CER is the measurement of HRQL.

Properties of HRQL Instruments

Instruments to quantify pain and HRQL can be used to measure the difference between patients at a point in time (discriminative purposes) or differences within a patient over time (evaluative purposes). They can be specific, focusing on particular conditions (disease specific), or they can be generic, designed for use in a variety of contexts. Disease specific instruments may be more sensitive to clinical change, but generic instruments are useful to quantify a range of impacts related to disease and its treatment, and may be the only choice when a patient has more than one condition, a situation encountered commonly in veterinary medicine (Mattin et al., 2014; Noble et al., 2018). Instruments either generate a single index score which indicates that a patient is better or worse (Brazier et al., 2017), or a profile of scores which offers more information and may be more sensitive to group differences and to changes in health status over time (Streiner and Norman, 1989).

These instruments may be particularly useful in situations where carers have to make decisions regarding illness and pain in animals that are not approachable, for example zoo animals. Regular observation of these animals using a HRQL tool which targets species-specific behaviours should highlight deviations from the norm and indicate where further, more invasive investigation is warranted.

In veterinary medicine a number of disease specific HRQL instruments have been created for use in the dog to measure the

[1] Federal Coordinating Council for Comparative Effectiveness Research, Report to the President and the Congress, June 30 2009. http://www.reesfrance.com/en/IMG/pdf/2009_cerannualrpt_light_.pdf (accessed 7 September 2017).

Table 2
Health-related quality of life instruments for the dog and cat. References refer to initial development of the instrument.

Ref.	Target Species	Type Disease specific (DS) Generic (G)	Target condition	Number of items	Scoring
Freeman et al. (2005)	Dog	DS	Cardiac	17	Single index
Yazbek and Fantoni (2005) https://avmajournals.avma.org/doi/abs/10.2460/javma.2005.226.1354	Dog	DS	Cancer	12	Single index
Wojciechowska and Hewson (2005) https://avmajournals.avma.org/doi/abs/10.2460/ajvr.2005.66.1453	Dog	G	All	27	Single index
Budke et al. (2008)	Dog	DS	Spinal cord injuries	5	Single index
Favrot et al. (2010) https://www.ncbi.nlm.nih.gov/pubmed/20187912	Dog	DS	Atopic dermatitis	13	Single index
Niessen et al. (2010) https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1939-1676.2010.0579mellitus	Cat	DS	Diabetes mellitus	29	Single index
Zamprogno et al. (2010)	Cat	DS	Degenerative joint disease	15	Single index
Lynch et al. (2011)	Dog	DS	Cancer	21	Single index
Noli et al. (2011) https://www.ncbi.nlm.nih.gov/pubmed/21410569	Dog	DS	Skin disease	15	Single index
Niessen et al. (2012) https://onlinelibrary.wiley.com/doi/full/10.1111/j.1939-1676.2012.00947	Dog	DS	Diabetes mellitus	29	Single index
Freeman et al. (2012)	Cat	DS	Cardiac	17	Single index
Lavan (2013) https://www.sciencedirect.com/science/article/pii/S1090023313001391	Healthy dog	G	N/A	15	Single index
Reid et al. (2013)	Dog	G	All	46	Profile of scores in vitality, pain, distress, anxiety
Bijsmans et al. (2016) http://journals.sagepub.com/doi/abs/10.1177/1098612X166573865	Cat	DS	Chronic kidney disease	16	Average weighted score
Noli et al. (2016) https://www.ncbi.nlm.nih.gov/pubmed/2729213	Cat	DS	Skin disease	15	Single item score
Freeman et al. (2016) http://journals.sagepub.com/doi/full/10.1177/1098612X16657386	Healthy cat	G	N/A	33 items	Profile of scores in mobility, emotion, energy, engagement, eyes, coat, appetite, fitness
Reid et al. (2017a, 2017b) http://www.newmetrica.com/vetmetrica-hrql/	Dog	G	All	22 items	Profile of scores in energy, happiness, comfort and calmness
Noble et al. (2018) http://www.newmetrica.com/vetmetrica-hrql/	Cat	G	All	20	Profit of scores in vitality, comfort and emotional wellbeing

impact of chronic diseases which may or may not be associated with pain (Table 2).

Chronic pain in companion animals

Orthopaedic pain

The incidence of chronic pain in companion animals is known to be high with one in five dogs older than one year estimated to be suffering from osteoarthritis (Johnston, 1997). Much research has focused on measuring the pain associated with OA and because changes in mobility are a major feature in the dog, tools have been developed to measure functional changes, including locomotor activity monitoring, kinetic evaluation and owner-report of functional limitations. (Klinck et al., 2017). The latter are termed clinical metrology instruments and include a range of tools (Table 3).

Although it was not recognised until relatively recently the incidence of OA in the cat is also high, but in contrast to the dog, lameness is not a feature (Bennett and Morton, 2009). Despite that, the Feline Musculoskeletal Pain Index (FMPI — 17 items, single index score), a validated owner reported clinical metrology instrument has been developed which focuses on functional limitations associated with the disease (Zamprogno et al., 2010; Benito et al., 2013). The reader should note that the references supplied in Tables 2 and 3 refer to the initial development of these OA measures only and other published studies are available, the description of which is not included in this review. Many of these related publications relate to validity of the individual instruments and it is important to remember that validity is not determined by a single statistic, but by a body of research that upholds the claim that the instrument is valid for particular purposes, with defined populations and in specified contexts (Streiner and Norman, 2008).

Guidelines have been published for orthopaedic studies (Cook, 2014) which include a recommendation that at least one validated functional outcome such as kinematics, kinetics, activity monitors, and clinical metrology should be included, as well as at least one observer reported QOL instrument, thus recognising the important contribution of the latter.

Non-orthopaedic pain

Unfortunately, because of the high profile associated with orthopaedic disease, especially OA, there is a tendency to overlook the fact that other chronic painful conditions are also commonplace in cats and dogs, including, but not restricted to, dental disease, certain cancers and chronic inflammatory disease such as otitis and bowel disease in dogs and cystitis in cats. Apart from osteosarcoma, specific changes in mobility are often not a feature of these diseases so the use of clinical metrology instruments does not apply. In the absence of a disease specific instrument to measure the impact of these conditions and their treatment, a

Table 3
Clinical metrology instruments for osteoarthritis in the dog and cat. References refer to initial development of the instrument.

Name	Target species	Number of items	Score	Ref.
Canine Brief Pain Inventory (CBPI)	Dog	11	1. Severity of pain 2. Interference with function	Brown et al. (2007) http://www.vet.upenn.edu/research/clinical-trials/vcic/pennchart/cbpi-tool
Helsinki Chronic Pain Index (HCPI)	Dog	11	Single index	Hielm-Björkman et al. (2009) https://www.fourleg.com/media/Helsinki%20Chronic%20Pain%20Index.pdf
Liverpool Osteoarthritis in Dogs questionnaire (LOAD)	Dog	14	Single index	Hercock et al. (2009) https://dspace.uevora.pt/rdpc/bitstream/10174/19611/2/liverpool%20OA%20in%20dogs%20-%20load.pdf
Feline Musculoskeletal Pain Index (FMPI)	Cat	17	Single index	Zamprogno et al. (2010) http://journals.plos.org/plosone/article/file?type=supplementary&id=info:doi/10.1371/journal.pone.0131839.s001
Canine Orthopaedic Index (COI)	Dog	21	1. Stiffness 2. Gait 3. Function 4. Quality of Life Overall summary score	Brown (2014) http://www.vet.upenn.edu/research/clinical-trials/vcic/pennchart/canine-orthopedic-index

generic HRQL instrument is a useful option, as it is when co-morbidities exist. In a recent study to develop and validate a feline HRQL instrument, where it was shown that the presence of comorbidities were associated with decreased HRQL, 31/34 cats with OA had between and 6 co-morbidities (Noble et al., 2018).

Conclusions

The past two decades have seen a shift in focus from recording a 'global' score for the intensity of pain using simple unidimensional scores to measuring the affective component of the pain experience with composite instruments developed using the psychometric methods, well established in human medicine for that purpose. In particular, HRQL measurement is at the forefront of medical research because of its wide-ranging applications, not only to measure the impact of chronic pain, but also in clinical trials relating to non-painful chronic disease, and in clinical practice to enhance patient care. The last decade has seen a growing interest in the development of such instruments in veterinary science, and we are at the beginning of a fascinating journey as more instruments become available which will benefit animal care.

Instrument development is an iterative process, in which instruments are refined and re-tested with new populations in new contexts and for new purposes. Their development is a time-consuming, challenging and costly undertaking, but the importance of the psychometric approach to such instrument development is widely accepted (Cook et al., 2003). By adopting rigorous methods to construct pain and HRQL instruments which ensure their validity and reliability, veterinary practitioners can be more confident in the management and treatment of pain in animals under their care, whatever the cause of that pain might be.

Conflict of interest statement

Jacqueline Reid is a director in NewMetrica Ltd, a company that distributes pain and HRQL scales.

References

Abell, N., Springer, D.W., Kamata, A., 2009. Developing and Validating Rapid Assessment Instruments. Oxford University Press, Oxford.
Albuquerque, N., Guo, K., Wilkinson, A., Savalli, C., Otta, E., Mills, D., 2016. Dogs recognize dog and human emotions. Biology Letters 12, 20150883.
Bateson, P., 1991. Assessment of pain in animals. Animal Behaviour 42, 827–839.
Benito, J., DePuy, V., Hardie, E., Zamprogno, H., Thomson, A., Simpson, W., Roe, S., Hansen, B., Lascelles, B.D.X., 2013. Reliability and discriminatory testing of a client-based metrology instrument, feline musculoskeletal pain index (FMPI) for the valuation of degenerative joint disease-associated pain in cats. The Veterinary Journal 196, 368–373.
Bennett, D., Morton, C., 2009. A study of owner observed behavioural and lifestyle changes in cats with musculoskeletal disease before and after analgesic therapy. Journal of Feline Medicine and Surgery 11, 997–1004.
Berzon, R., Hays, R.D., Shumaker, S.A., 1993. International use, application and performance of health-related quality of life instruments. Quality Life Research 2, 367–368.
Bijsmans, E.S., Jepson, R.E., Syme, H.M., Elliott, J., Niessen, S.J.M., 2016. Psychometric validation of a general health quality of life tool for cats used to compare healthy cats and cats with chronic kidney disease. Journal of Veterinary Internal Medicine 30, 183–191.
Brazier, J., Ratcliffe, J., Saloman, J., Tsuchiya, A., 2017. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford University Press, Oxford, UK.
Brondani, J.T., Luna, S.P.L., Padovani, C.R., 2011. Refinement and initial validation of a multidimensional composite scale for use in assessing acute postoperative pain in cats. American Journal of Veterinary Research 72, 174–183.
Brondani, J.T., Mama, K.R., Luna, S.P., Wright, B.D., Niyom, S., Ambrosio, J., Vogel, P.R., Padovani, C.R., 2013. Validation of the English version of the UNESP-Botucatu multidimensional composite pain scale for assessing postoperative pain in cats. BMC Veterinary Research 9, 143.
Brown, D.C., Boston, R.C., Coyne, J.C., Farrar, J.T., 2007. Development and psychometric testing of an instrument designed to measure chronic pain in dogs with osteoarthritis. American Journal of Veterinary Research 68, 631–637.
Brown, D.C., 2014. The canine orthopedic index. Step 1: devising the items. Veterinary Surgery 43, 232–240.
Budke, C.M., Levine, J.M., Kerwin, S.C., Levine, G.J., Hettlich, B.F., Slater, M.R., 2008. Evaluation of a questionnaire for obtaining owner-perceived, weighted quality-of-life assessments for dogs with spinal cord injuries. Journal of the American Veterinary Medical Association 223, 925–930.
Calvo, G., Holden, E., Reid, J., Scott, E.M., 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, 622–629.
Chapman, C.B., 1989. Pain assessment and pain control. Proceedings of the 11th Bain-Fallon Memorial Lectures. Equine Pharmacology and Therapy 2–11.
Cook, K.F., Monahan, P.O., McHorney, C.A., 2003. Delicate balance between theory and practice: health status assessment and item response theory (Editorial). Medical Care 41, 571–574.
Cook, J.L., 2014. Canine orthopedic outcome measures program: where are we now? Veterinary Surgery 43, 229–231.
Dalla Costa, E., Minero, M., Lebelt, D., Stucke, D., Canali, E., Leach, M.C., 2014. Development of the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine castration. PLoS One 9, e92281.
Dawkins, M.S., 2004. Using animal behaviour to assess animal welfare. Animal Welfare 13, 3–7.
Dreyer, N.A., Schneeweiss, S., McNeil, B.J., Berger, M.L., Walker, A.M., Ollendorf, D.A., Gliklich, R.E., 2010. GRACE principles: recognizing high-quality observational studies of comparative effectiveness. American Journal of Managed Care 16, 467–471.
Emery, M.P., Perrier, L.L., Acquadro, C., 2005. Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID): frequently asked questions. Health and Quality of Life Outcomes 3, 1–2.

Fallowfield, L., 1990. The Quality of Life: the Missing Measurement in Health Care. Souvenir Press, London.
Favrot, C., Linek, M., Mueller, R., Zini, E., 2010. International Task Force on Canine Atopic Dermatitis. Development of a questionnaire to assess the impact of atopic dermatitis on health-related quality of life of affected dogs and their owners. Veterinary Dermatology 21, 64–70.
Feinstein, A.R., 1987. Clinimetric perspectives. Journal of Chronic Diseases 40, 635–640.
Firth, A.M., Haldane, S.L., 1999. Development of a scale to evaluate postoperative pain in dogs. Journal American Veterinary Medical Association 214, 651–659.
Flecknell, P.A., 1986. Recognition and alleviation of pain in animals. In: Fox, M.W., Mickley, L.D. (Eds.), Advances In Animal Welfare Science 1985. Martinus Nijhoff, Dordrecht, pp. 61–77.
Flecknell, P., Firth, A.M., Haldane, S.L., 2008. Analgesia from a veterinary perspective. British Journal of Anaesthesia 101, 121–124.
Floyd, F.J., Widaman, K.F., 1995. Factor analysis in the development and refinement of clinical assessment instruments. Psychological Assessment 7, 286–299.
Freeman, L.M., Rush, J.E., Farabaugh, A.E., Must, A., 2005. Development and evaluation of a questionnaire for assessing health-related quality of life in dogs with cardiac disease. Journal of the American Veterinary Medical Association 226, 1864–1868.
Freeman, L.M., Rush, J.E., Oyama, M.A., MacDonald, K.A., Cunningham, S.M., Bulmer, B., MacGregor, J.M., Laste, N.J., Malakoff, R.L., Hall, D.J., et al., 2012. Development and evaluation of a questionnaire for assessment of health-related quality of life in cats with cardiac disease. Journal of the American Veterinary Medical Association 240, 1188–1193.
Freeman, L.M., Rodenberg, C., Narayanan, A., Olding, J., Gooding, M.A., Koochaki, P.E., 2016. Development and initial validation of the Cat HEalth and Wellbeing (CHEW) Questionnaire: a generic health-related quality of life instrument for cats. Journal of Feline Medicine and Surgery 18, 689–701.
Griffin, D.R., 1992. Animal Minds. University of Chicago Press, Chicago.
Hand, I.L., Noble, L., Geiss, D., Wozniak, L., Hall, C., 2010. COVERS neonatal pain scale: development and validation. International Journal of Pediatrics 2010, 496–719.
Hercock, C.A., Pinchbeck, G., Giejda, A., Clegg, P.D., Innes, J.F., 2009. Validation of a client-based clinical metrology instrument for the evaluation of canine elbow osteoarthritis. Journal of Small Animal Practice 50, 266–271.
Hielm-Björkman, A.K., Rita, H., Tulamo, R.M., 2009. Psychometric testing of the Helsinki chronic pain index by completion of a questionnaire in Finnish by owners of dogs with chronic signs of pain caused by osteoarthritis. American Journal of Veterinary Research 70, 727–734.
Holden, E., Calvo, G., Collins, M., Bell, A., Reid, J., Scott, E.M., Nolan, A.M., 2014. Evaluation of facial expression in acute pain in cats. Journal of Small Animal Practice 55, 615–621.
Holopherne-Doran, D., Laboissière, B., Gogny, M., 2010. Validation of the 4Avet postoperative pain scale in dogs and cats. Veterinary Anaesthesia and Analgesia 37, 1–17.
Holton, L.L., Scott, E.M., Nolan, A.M., Reid, J., Welsh, E., 1998a. Relationship between physiological factors and clinical pain in dogs scored using a numerical rating scale. Journal of Small Animal Practice 39, 469–474.
Holton, L.L., Scott, E.M., Nolan, A.M., Reid, J., Welsh, E., Flaherty, D., 1998b. Comparison of three methods used for assessment of pain in dogs. Journal of the American Veterinary Medical Association 212, 61–66.
Jensen, M.P., 2003. Questionnaire validation: a brief guide for readers of the research literature. The Clinical Journal of Pain 19, 345–352.
Johnston, S.A., 1997. Osteoarthritis. Joint anatomy, physiology, and pathobiology. The Veterinary Clinics of North America. Small Animal Practice 27, 699–723.
Kappesser, J., Williams, A.C., 2002. Pain and negative emotions in the face: judgments by health care professionals. Pain 99, 197–206.
Keating, S.C.J., Thomas, A.A., Flecknell, P.A., Leach, M.C., 2012. Evaluation of EMLA cream for preventing pain during tattooing of rabbits: changes in physiological, behavioural and facial expression responses. PLoS One 7, e44437.
Klinck, M.P., Mogil, J.S., Moreau, M., Lascelles, B.D.X., Flecknell, P.A., Poitteg, T., Troncy, E., 2017. Translational pain assessment: could natural animal models be the missing link? Pain 158, 1633–1646.
Langford, D.J., Bailey, A.L., Chanda, M.L., Clarke, S.E., Drummond, T.E., Echols, S., Glick, S., Ingrao, J., Klassen-Ross, T., LaCroix-Fralish, M.L., et al., 2010. Coding of facial expressions of pain in the laboratory mouse. Nature Methods 7, 447–449.
Lavan, R.P., 2013. Development and validation of a survey for quality of life assessment by owners of healthy dogs. The Veterinary Journal 197, 578–582.
Lynch, S., Savary-Bataille, K., Leeuw, B., Argyle, D.J., 2011. Development of a questionnaire assessing health-related quality-of-life in dogs and cats with cancer. Veterinary and Comparative Oncology 9, 172–182.
Mathews, K.A., 2000. Pain assessment and general approach to management. Veterinary Clinics of North America: Small Animal Practice 30, 729–755.
Mattin, M., O'Neill, D., Church, D., McGreevy, P.D., Thomson, P.C., Brodbelt, D., 2014. An epidemiological study of diabetes mellitus in dogs attending first opinion practice in the UK. Veterinary Record 174, 349.
McColl, E., Jacoby, A., Thomas, L., Soutter, J., Bamford, C., Steen, N., Thomas, R., Harvey, E., Garratt, A., Bond, 2001. Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. Health Technology Assessment 5 Executive Summary.
McGrath, P.J., Rosmus, C., Canfield, C., Campbell, M.A., Hennigar, A., 1998. Behaviours caregivers use to determine pain in non-verbal, cognitively impaired individuals. Developmental Medicine and Child Neurology 40, 340–343.
Mendl, M., Harcourt, R., 2000. Individuality in the domestic cat: origins, development and stability. In: Turner, D.C., Bateson, P.P. (Eds.), The Domestic Cat: The Biology of Its Behaviour. Cambridge University Press, pp. 47–64.
Morton, D., Griffiths, P.H.M., 1985. Guidelines on the recognition of pain, distress and discomfort in experimental animals and a hypothesis for assessment. Veterinary Record 116, 431–436.
Morton, C.M., Reid, J., Scott, E.M., Holton, L.L., Nolan, A.M., 2005. Application of a scaling model to establish and validate an interval level pain scale for assessment of acute pain in dogs. American Journal Veterinary Research 66, 2154–2166.
Nagasawa, M., Mogi, K., Kikusui, T., 2009. Attachment between humans and dogs. Japanese Psychological Research 51, 209–221.
Niessen, S.J.M., Powney, S., Guitian, J., Niessen, A.P.M., Pion, P.D., Shaw, J.A.M., Church, D.B., 2010. Evaluation of a quality-of-life tool for cats with diabetes mellitus. Journal of veterinary internalmedicine 24, 1098–1105.
Niessen, S.J.M., Powney, S., Guitian, J., Niessen, A.P.M., Pion, P.D., Shaw, J.A.M., Church, D.B., 2012. Evaluation of a quality-of-life tool for dogs with diabetes mellitus. Journal of Veterina Internal Medicine 26, 953–961.
Noble, C.E., Wiseman-Orr, L.M., Scott, M.E., Nolan, A.M., Reid, J., 2018. Development, initial validation and reliability testing of a web-based, generic feline health-related quality-of-life instrument. Journal of Feline Medicine and Surgery doi: http://dx.doi.org/10.1177/1098612X18758883.
Noli, C., Minafò, G., Galzerano, M., 2011. Quality of life of dogs with skin diseases and their owners. Part 1: development and validation of a questionnaire. Veterinary Dermatology 22, 335–343.
Noli, C., Borio, S., Varina, A., Schievano, C., 2016. Development and validation of a questionnaire to evaluate the quality of life of cats with skin disease and their owners, and its use in 185 cats with skin disease. Veterinary Dermatology 27, 247–258.
Nunnally, J.C., Bernstein, I.H., 1994. Psychometric Theory, Third Ed. McGraw-Hill, New York.
Polit, D.F., Beck, C.T., 2006. The content validity index: are you sure you know what's being reported? Critique and recommendations. Research in Nursing and Health 29, 489–497.
Prkachin, K.M., Schultz, I., Berkowitz, J., Hughes, E., Hunt, D., 2002. Assessing pain behaviour of low-back pain patients in real time: concurrent validity and examinersensitivity. Behavioural Research and Therapy 40, 595–607.
Reid, J., Nolan, A.M., Hughes, J.M.L., Lascelles, D., Pawson, P., Scott, E.M., 2007. Development of the short-form Glasgow Composite Measure Pain Scale (CMPS-SF) and derivation of an analgesic intervention score. Animal Welfare 16 (Suppl), 97–104.
Reid, J., Wiseman-Orr, M.L., Scott, E.M., Nolan, A.M., 2013. Development, validation and reliability of a web-based questionnaire to measure health-related quality of life in dogs. Journal of Small Animal Practice 54, 227–233.
Reid, J., Wiseman-Orr, L., Scott, M., 2017a. Shortening of an existing generic online health-related quality of life instrument for dogs. Journal of Small Animal Practice doi:http://dx.doi.org/10.1111/jsap.12772.
Reid, J., Scott, E.M., Calvo, G., Nolan, A.M., 2017b. Definitive Glasgow acute pain scale for cats: validation and intervention level. Veterinary Record 108, 18.
Rozenblum, R., Bates, D.W., 2013. Patient-centred healthcare, social media and the internet: the perfect storm? BMJ Quality & Safety 22, 183–186.
Rutherford, K.M.D., 2002. Assessing pain in animals. Animal Welfare 11, 31–53.
Scott, E.M., Nolan, A.M., Reid, J., Wiseman-Orr, L.M., 2007. Can we really measure animal quality of life? Methodologies for measuring quality of life in people and other animals. Animal Welfare 16 (Suppl), 17–24.
Soproni, K., Miklósi, Á., Topál, J., Csányi, V., 2001. Comprehension of human communicative signs in pet dogs (Canis familiaris). Journal of Comparative Psychology 115, 122.
Sotocinal, S.G., Sorge, R.E., Zaloum, A., Tuttle, A.H., Martin, L.J., Wieskopf, J.S., Mapplebeck, J.C., Wei, P., Zhan, S., Zhang, S., et al., 2011. The Rat Grimace Scale: a partially automated method for quantifying pain in the laboratory rat via facial expressions. Molecular Pain 7, 55–126.
Stallard, P., Williams, L., Velleman, R., Lenton, S., McGrath, P.J., Taylor, G., 2002. The development and evaluation of the pain indicator for communicatively impaired children (PICIC). Pain 98, 145–149.
Stevens, B., Johnston, C., Petryshen, P., Taddio, A., 1996. Premature infant pain profile: development and initial validation. Clinical Journal of Pain 12, 13–22.
Streiner, D.L., 1993. Research methods in psychiatry: a checklist for evaluating the usefulness of rating scales. Canadian Journal of Psychiatry 38, 140–148.
Streiner, D.L., Norman, G.R., 2008. Health Measurement Scales: A Practical Guide to Their Development and Use, Fourth Ed. Oxford University Press, Oxford.
Svartberg, K., Forkman, B., 2002. Personality traits in the domestic dog (Canis familiaris). Applied Animal Behaviour Science 79, 133–155.
Tavakol, M., Dennick, R., 2011. Making sense of Cronbach's alpha. International Journal of Medical Education 2, 53–55.
Teasdale, G., Jennett, B., 1974. Assessment of coma and impaired consciousness. A practical scale. Lancet 2, 81–84.
Tian-hui, C., Lu, L., 2005. A systematic review: How to choose appropriate health-related quality of life (HRQOL) measures in routine general practice? Journal of Zhejiang University Science B 6, 936–940.
van Dijk, M., de Boer, J.B., Koot, H.M., Tibboel, D., Passchier, J., Duivenvoorden, H.J., 2000. The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. Pain 84, 367–377.
Wemelsfelder, F., Hunter, T.E., Mendl, M.T., Lawrence, A.B., 2001. Assessing the 'whole animal': a free choice profiling approach. Animal Behaviour 62, 209–220.

Wiseman, M.L., Nolan, A.M., Reid, J., Scott, E.M., 2001. Preliminary study on owner-reported behaviour changes associated with chronic pain in dogs. The Veterinary Record 149, 423–424.
Wiseman-Orr, M.L., Nolan, A.M., Reid, J., Scott, E.M., 2004. Development of a questionnaire to measure the effects of chronic pain on health-related quality of life in dogs. American Journal of Veterinary Research 65, 1077–1084.
Wiseman-Orr, M.L., 2005. The Development of an Instrument to Measure Chronic Pain in Dogs. Doctoral Dissertation. University of Glasgow.
Wojciechowska, J.I., Hewson, C.J., 2005. Quality-of-life assessment in pet dogs. Journal of the American Veterinary Medical Association 226, 722–728.
Zamprogno, H., Hansen, B.D., Bondell, H.D., Sumrell, A.T., Simpson, W., Robertson, I. D., Brown, J., Pease, A.P., Roe, S.C., Hardie, E.M., et al., 2010. Item generation and design testing of a questionnaire to assess degenerative joint disease-associated pain in cats. American Journal of Veterinary Research 71, 1417–1424.