Key Points

- Although an owner’s detailed assessment is routinely relied upon in clinical decision making for the management of many musculoskeletal conditions, there is a lack of valid and reliable tools available to quantify the owner’s assessment of pain, function, and quality of life in their pet.

- Established principles of questionnaire development can be used to develop valid and reliable behavior-based assessment instruments for owner appraisal of chronic pain, function, and quality of life in companion animals.

- Such instrument will allow surgeons to quickly collect valid and reliable data on the owner’s assessment of pre- and post-intervention outcome in their patients, which is invaluable to investigators designing studies, as well as to practitioners that use the results of such studies to guide their clinical decision making.

In clinical practice, the veterinarian relies heavily on the owner’s assessment of the efficacy of an intervention (surgical procedure, drug etc.) through the description of the dog’s behavior in its normal environment during its activities of daily living (walking, running, jumping etc). For example, an owner may describe their dog ‘slowing down’, ‘refusing to climb stairs’, or ‘stopping during walks’. These complaints could prompt an evaluation of the animal that leads to a diagnosis of coxofemoral osteoarthritis and initiation of treatment with a nonsteroidal anti-inflammatory drug (NSAID). In monitoring the progression of the dog’s condition over time, the veterinarian relies heavily on the owner’s report of improvement in the dog’s behaviors. When the owner reports improvement in behaviors without substantial adverse effects, the veterinarian continues to offer the treatment. However, when an owner reports no improvement in behaviors, the veterinarian may increase the dose of the NSAID, or switch to another medication. If the owner continues to report negative behaviors, the veterinarian may begin a conversation about surgical management of the condition. In the end, if the dog does have surgery, it will likely only be considered a success if the owner considers the procedure a success, reporting improved function and quality of life for their pet. Although an owner’s detailed assessment of pain, function, and quality of life is routinely relied upon in clinical decision making for the management of many musculoskeletal conditions, there is a lack of valid and reliable tools available to quantify the owner’s assessment of pain, function, and quality of life in their pet.

This lack of validated owner assessments can be attributable to the subjective nature of an owner’s assessment and the complex, multi-step, resource intensive methods required for the development and validation of assessment instruments that can reliably quantify peoples’ perception of subjective states. Although behaviors are the result of a complex set of inputs and qualities unique to each animal, appropriate, established principles of questionnaire development can be used to develop valid and reliable behavior-based assessment instruments for owner appraisal of chronic pain, function, and quality of life in companion animals.
The Stepwise Process of Developing a Health Measurement Instrument

Step 1—Devising the Items: The first step in developing an instrument (questionnaire) involves devising the items (questions) themselves. Whereas this seems obvious, it is far from trivial, because no amount of statistical manipulation after the fact can compensate for questions that are poorly worded, ambiguous, or irrelevant. Items are generated through multiple focus groups or interviews with owners and/or veterinarians of animals that have the disease or condition for which we are hoping to determine the efficacy of interventions. While devising the items of a questionnaire appears on the surface to be straightforward, it is a very involved and time consuming process that can greatly impact the validity, reliability, and ultimate utility of the questionnaire.

Step 2—Selecting the Items: Typically, not all of the items that are developed are ultimately included in the new instrument. Some may be confusing, interpreted differently by different respondents, or not deliver the desired information. Various criteria can be used to determine which of the developed items should be retained for the preliminary instrument. Items that are difficult to interpret are removed or rephrased. Once a preliminary set of items is agreed upon, the instrument can be pretested. Pretesting involves the administration of the preliminary instrument to a small group of respondents to determine whether some items may not perform well. In most situations, when measuring a trait, the instrument should be homogeneous. That is, all of the items should be tapping into different aspects of the same trait, not different aspects of different traits. Therefore, the items should be moderately correlated with each other. An interitem correlation matrix can be analyzed from the data collected from this group of respondents to identify items that have consistently low correlations with other items in the instrument. These items can either be revised and the preliminary instrument pretested again in a different group of respondents, or the items can be removed from the instrument before moving onto large scale testing for reliability and validity.

Step 3—Reliability and Validity Testing: Once a data-gathering instrument is developed, it must be established that it will target what it is supposed to measure. This is defined as the validity of the instrument. In addition, the instrument must measure what it is supposed to measure in a consistent manner. The tendency toward consistency is referred to as reliability. There are a number of ways in which an instrument can be tested for reliability. An assessment of internal consistency can be based on the data collected from a single administration of the instrument to a large group of respondents. Whereas there are a number of methods of consistency calculation, all represent the average of the correlations among all of the items. An assessment of internal consistency alone; however, is not sufficient to declare an instrument reliable, because it relies on only a single administration. To account for the day-to-day variability in responses, an assessment of the stability (i.e. reproducibility) of the instrument must also be made. By administering the instrument to the same population of respondents on 2 different occasions, the test–retest reliability of the instrument can be assessed. It is only when we can demonstrate that the instrument is measuring something reliably, that we can begin the process of determining what that something is. To determine that the instrument is measuring what is intended requires more than peer judgments (face validity). Validating an instrument is a process by which we determine the degree of confidence we can place on conclusions we draw about an animal based on their score from that instrument. If other validated instruments designed to measure the same attribute exist, then an obvious approach is to administer the
experimental instrument along with the existing one and see whether there is a strong correlation between the 2. More likely, however, no other measure exists and developers must test ‘‘construct’’ validity. Construct validity is evaluated when the attribute being measured cannot be directly observed. For example, chronic pain cannot be ‘‘seen,’’ but behaviors can be observed which, according to our theories about chronic pain in companion dogs, result from it. There is no one single experiment or statistic, which can unequivocally ‘‘prove’’ a construct. It is through multiple analyses and assessments that a construct appears to be valid. While there is no one single experiment to ‘‘prove’’ the validity of the instrument, multiple, well-designed, hypothesis driven studies can build the body of evidence that the instrument is measuring what we have intended. It is necessary to conduct validation studies for each new instrument that is developed and the task is an on-going one.

Applicability to Veterinary Surgery

Very few attempts have been made in veterinary medicine to develop valid and reliable instruments to measure companion animal health. It is not uncommon to find published efficacy studies where one of the outcomes is the summative score of a set of items that were made up for the study, but underwent no tests for reliability or validity. It is impossible to know for sure, therefore, whether the results (positive or negative) of such studies are because of the intervention rather than inconsistencies or inaccuracies associated with the outcome score. One reason for the prevalent use of unvalidated outcomes may be a lack of awareness of the necessity for the process and the fact that sound methodology for development and validation exists. Another reason may be that, as outlined above, the process of instrument development and validation is an onerous one. It takes a great amount of time (years), expertise, and financial support. The ACVS Foundation, with support from a variety of industry sponsors, has supported the development and validation of such an instrument. This instrument will allow surgeons to quickly collect valid and reliable data on the owner’s assessment of pre- and post-intervention outcome in their patients, which is invaluable to investigators designing studies, as well as to practitioners that use the results of such studies to guide their clinical decision making.