ASSESSING THE REPEATABILITY AND VALIDITY OF A QUESTIONNAIRE ON PAIN AND LAMENESS IN THE CANINE

A Thesis

by

JONATHAN THOMAS HUDSON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2003

Major Subject: Epidemiology

ASSESSING THE REPEATABILITY AND VALIDITY OF A QUESTIONNAIRE ON PAIN AND LAMENESS IN THE CANINE

A Thesis

by

JONATHAN THOMAS HUDSON

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

Margaret Slater (Chair of Committee)	Lathrop Taylor (Member)
H. Morgan Scott (Member)	
Sharon Kerwin (Member)	Evelyn Tiffany-Castiglioni (Head of Department)

May 2003

Major Subject: Epidemiology

ABSTRACT

Assessing the Repeatability and Validity of a Questionnaire on Pain and Lameness in the Canine. (May 2003)

Jonathan Thomas Hudson, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Margaret Slater

The measurement of pain has had a growing importance in animals for both privately owned animals and those animals involved in clinical research. Lameness is considered to be 1 aspect of the pain experience. The ability of a veterinarian to assess lameness during a routine orthopedic examination can be difficult given the short amount of time in which the clinician can observe the animal, and the fact that the animal is in a stressful environment. Thus, the input of the owner concerning the animal's well-being over an extended time period may be extremely useful to the clinician in assessing the degree of lameness of the animal. It was the purpose of this study to establish an instrument that was both repeatable and valid in assessing the degree of lameness. The instrument used was a questionnaire containing 39 questions in a visual analog scale format. A force platform was used as the gold-standard for detecting mechanical lameness. Peak vertical, cranial-caudal, and their associated impulses were forces used to determine lameness, along with maximum slope in some cases. A test-retest measure of repeatability was conducted on a subset of 19 dogs that were confirmed to have less than a 10% change in vertical peak force. Nineteen of the

39 questions were found to be repeatable based on a Spearman rank correlation. These 19 questions were then used as predictor variables in several multiple regression models which predicted force plate measurements. The result was 3 different models each containing 7 independent variables that were thought to be valid representations of the forces measured (vertical peak, vertical impulse, and propulsion peak forces). Each reduced model was found to fit the data as well as the full model containing all 19 of the repeatable questions. The composite of 11 questions from the 3 different models was used to calculate a total score. This total score was found to be significantly correlated with force plate measurements. These 11 questions should be useful to a clinician in detecting the degree of lameness in the dog.

DEDICATION

I would like to dedicate this thesis to my wife, Denise, who has been supportive and encouraging through the entire process. She has inspired me and provided me with much of the motivation necessary to finish this project, and in a timely manner. I would also like to extend this dedication to the rest of my family who has been very supportive of my desire to further my education. Without the support of these people that I hold near and dear to me, I would not be nearly as successful in my endeavors.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Margaret Slater, who agreed to take me under her wing nearly 2 years ago. She allowed me the opportunity to receive a more advanced education which will undoubtedly help me accomplish future goals. When I found myself against an obstacle during this study, Dr. Slater found a way to direct me toward the solution without simply giving me the answer. I appreciate being challenged intellectually and in a constructive manner. Her numerous revisions of my work have given me a better understanding of the shape and format that scientific literature should take.

I would also like to thank Dr. Lathrop Taylor for allowing me to analyze her data for this thesis project. Working on data already collected always sounds appealing, but there is the necessity of learning and understanding all the work that went into the data collection. There is also the lack of first-hand knowledge which can be useful when other questions arise. Dr. Taylor was very helpful and patient in relaying some of this knowledge. She also helped me in understanding some of the relationships among the forces measured, and how forces may be redistributed.

Dr. Sharon Kerwin was also instrumental in this project by explaining the requirements of an orthopedic exam. She also helped me in understanding the differences in lameness scales and how they are used among different orthopedic specialists. Her cheerful attitude and disposition were also very much appreciated.

Some of the statistical analysis required in this project was very challenging. Dr. H. Morgan Scott was very helpful in my understanding of some of the issues of multicollinearity and their effects in linear regression. His input on model selection and the criteria involved was also very helpful in this project.

I would also like to extend my gratitude to those graduate students who have gone before me and paved the road. Dr. Kathy Hughes, Jacque Harbison, and Allison Bernstein all gave me a better perspective on the graduate experience and what I should expect and learn. They were always very supportive and provided guidance when it was needed. I would also like to express thanks to the current graduate students in the Epidemiology program. Ashley Loven, Summer Strickland, and Linda Campbell have all been great people to work with, and I am glad to have had the opportunity to share this experience with them.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Pain has been referred to as the fifth vital sign in people and measuring it has been an evolving process.[1] Likewise, in animals, the level of pain is difficult to determine and can become an even greater task when the patient cannot speak with the doctor. While pain can be a difficult phenomenon to define in both humans and animals, it has been described as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage or described in terms of such damage."[2] Pain in animals has also been defined as "an aversive sensory experience that elicits protective motor actions, results in learned avoidance and may modify species specific traits of behavior, including social behavior."[2] This "learned avoidance" is referred to as lameness, which is thought to be a part of the pain experience. The avoidance of limb-use prevents the stimulation of nociceptors in the leg which leads to the sensation of pain.[3] Lameness in the context of veterinary medicine has been most commonly described as a deviation in the normal gait of an animal.[4] Thus, lameness is an outcome of a painful experience. However, the pain experience is comprised of a multitude of contributing factors. Thus, most animals in which lameness is present are thought to be experiencing pain. However, it is possible for an animal to have a mechanical lameness not associated with pain.

The measurement of this lameness is thought to be associated with a level of pain that an animal may be experiencing. Thus, an accurate measurement of this lameness

---

This thesis follows the style and format of the *Journal of the American Veterinary Medical Association*.

may help in assessing the level of pain the animal is experiencing. However, the animal may not present its lameness in a clinical setting, which makes the task of diagnosing lameness difficult even for veterinary orthopedic surgeons who are specially trained. For this reason, information gathered from the owner may be especially useful in diagnosing lameness. This information would aid the clinician in diagnosing lameness in the same respect that taking a history from a patient aids in the diagnosis of a disease. The current method of measuring lameness is also not uniform across the profession. I believe that a more uniform assessment of lameness may aid the clinician in assessing both pain and lameness. This uniform assessment could be especially beneficial for a veterinarian who is not specially trained in orthopedics. A more objective and standardized evaluation would also aid veterinary orthopedic research and facilitate multi-investigator projects.

LITERATURE REVIEW

Orthopedic Exam

Veterinarians have traditionally assessed lameness in the dog by conducting a thorough orthopedic exam. This method has been refined over the years by orthopedic surgeons, but is still a somewhat subjective physical exam. A typical orthopedic evaluation consists of 2 main parts: localizing the problem or source of lameness (isolating the problem area) and diagnosing the cause of the lameness.[4] This can be accomplished by a routine which generally consists of the following steps.[4] The dog should be observed sitting and standing. The dog should then be observed while walking and/or trotting. Certain movements or tendencies while sitting, standing, or walking, can be indicative of certain causes of lameness. Any of these indicative characteristics should be noted. The clinician should also be familiar with certain breed characteristics, which may mask or imitate lameness. At this point the clinician should have some idea as to the affected limb, and can now gently palpate the affected limb while the dog is standing noting differences in symmetry and any proprioceptive deficits. The clinician should then conduct a thorough exam of each limb with the dog in lateral recumbency, beginning with limbs believed to be functioning properly and ending with the affected limb and beginning at the toes/pads and working proximally. The lameness would then be scored based on the veterinarian's scoring system of choice, which may or may not be consistent with other veterinarians.

Useful Owner Information

Information concerning the dog's lameness prior to its visit to the veterinary clinic may be very useful to the clinician. An animal that goes to the veterinary clinic could experience a certain amount of stress with the visit, based on previous unpleasant visits, and could mask some signs of lameness which could aid in diagnosis. Also, working or hunting dogs will often only display their lameness when actively worked.[4] These situations can make a clinical exam much less accurate in the diagnosis and detection of lameness. Thus, gaining detailed information from the owner about the dog's activity and behavior may aid the clinician in diagnosing lameness.

The level of anxiety exhibited by the animal can also be used to judge the level of pain experienced by the animal. The owner of the animal may be the best person to judge this level of anxiety since they are more familiar with the normal level of anxiety displayed by the animal. A preliminary study by Wiseman et al.[5] shows that owners provide useful and valuable information pertaining to the dog's behavior when chronic pain is considered. The study consisted of 13 animals diagnosed with degenerative joint disease. This study was an exploratory study of owner's assessment of chronic pain; it lacked a structured interview and any statistical analysis. The owners of these animals were asked questions about the dogs' behavior prior to the onset of the disease. From these interviews, the occurrences of different behavior traits were enumerated and some traits, such as mobility and activity, seem to be common to most animals in the study. Wiseman also suggested certain traits that are expected to either increase or decrease with chronic pain, such as mobility and dependence.

Behavior

Pain assessment in animals has been a difficult task for those involved in developing more specific, objective techniques. The preferred method to validate these techniques is by comparison to a "gold-standard." A gold-standard is defined as the most accurate method, procedure, or measurement that is known to represent the true value of what is being tested. The gold-standard for assessing pain in people has been verbal communication; this obviously is not possible for animals.[1] Thus, we rely on certain behaviors that seem to indicate pain in the animal. However, certain behaviors associated with pain can often be misinterpreted as the animal seeking attention due to a decrease in the behavior when the animal receives attention. Actually, the attention given by the owner may only be distracting the animal from its own pain. Some studies have tried to identify behaviors of dogs that appear to be associated with pain.[2,5,6] Most of these studies depend on suggestions by qualified personnel such as veterinary surgeons.

The McGill pain questionnaire by Melzack in 1975 was 1 of the first human studies that established a valid tool for assessing pain.[7] The questionnaire was designed to capture 3 main aspects of the pain experience: sensory, affective, and evaluative. The questionnaire consisted of 4 parts, 2 of these were analyzed statistically. One part was a word association where the patients chose words from a list that described their pain. The other part consisted of 6 questions based on a numeric rating scale of 1-5, with descriptors ranging from "mild" to "excruciating." Three different measures were possible with the McGill questionnaire. One was a "pain rating index" that was calculated based on word associations in which different words had a different ranked value. Another measurement obtained from the questionnaire was the "number of words chosen." A value for the "present pain intensity" was also calculated in the McGill questionnaire based on the numeric rating scale. The "pain rating index" was found to be the measurement that was most sensitive to change in pain, whereas "the number of words chosen" was found to be the least sensitive to change. The McGill questionnaire was found to be a valid assessment of the pain experience. It has since been used as a comparison tool for others trying to establish valid assessments of pain for both humans and animals. For example, Holton et al. developed a pain scale for dogs that was based on the same techniques as the McGill study.[8] However, they used more refined statistical techniques to establish the different groups of questions which will be discussed later.

Scales and Scores of Pain

There have been various pain scales that have been developed for use in veterinary medicine over the years. These scales are generally of 4 basic formats, a numeric rating scale (NRS), a simple descriptive scale (SDS), a visual analogue scale (VAS), or a multifactorial scale which incorporates scores being given for different NRS, SDS, or VAS type questions. One of the more common pain scales that has been used in veterinary orthopedics is a 0-4 NRS.[9] The Ontario Veterinary College developed a 0-10 VAS[10] where descriptions were available to aid the rater. This does have similarities to a NRS in which descriptions are set for each integer rating; however, the rater is free to mark anywhere along the 10cm line. Simple descriptive scales have also been developed which more or less correlate to the 0-4 NRS where the descriptors are no pain, mild, moderate, severe, and very severe. Multifactorial pain scales have also been developed and shown to be reliable in detecting postoperative pain in animals.[8,11]

Various studies have been conducted to determine which scale is the most reliable and most accurate. VAS scales have been reported to be more likely to detect subtle changes than NRS.[12-15] A study by Holton et al.[13] suggests that while VAS scales are more likely to detect small changes, they show more variability and are less reliable. SDS scales are said to be less sensitive to detecting subtle changes than both VAS and NRS.

A multifactorial or composite scale has also been developed by Holton et al.[8] (same questionnaire discussed in *Behavior* section) that was modeled after the McGill pain questionnaire established in humans. The validity of Holton et al.'s questionnaire was based on an analysis of construct validity (see different types of validity in *repeatability and validity*) which is used extensively in the human literature. The construct validity of the groups in Holton's questionnaire was shown using hierarchical agglomerative cluster analysis, Cronbach's alpha coefficient, and an analysis of variance using multiple comparisons. These tests seemed appropriate in developing categories of questions. However, the validity of the whole questionnaire should be evaluated against a known gold-standard which has yet to be identified in assessing pain in dogs. The similarities between their questionnaire and the McGill questionnaire provides Holton with construct validity since the McGill questionnaire has been widely used and accepted in the human literature since its creation in 1975. However, Holton acknowledged that this was only the development phase of the questionnaire, and that it still needed to be tested on dogs. The questionnaire needs to be shown to be sensitive to change using a more objective measurement or gold-standard to validate it.

The study by Conzemius et al.[12] discusses how certain objective pain measurements (heart rate, respiratory rate, blood pressure, and pain threshold) do not correlate with subjective VAS and NRS scores. One possible problem with Conzemius's study is that the correlation between the objective and subjective measures was done with Pearson's product moment correlation coefficient, which carries with it an assumption of normality. The results may have followed a normal distribution, but were not shown or described. If the results were not truly normal, this would account for some discrepancies in the correlations. The authors showed that there were some significant correlations between ratings on the VAS and NRS when compared to respiratory rate and vocalization when using the Pearson's product moment correlation coefficient. However, there were not significant correlations between the subjective pain scales and the heart rate, blood pressure, or pain threshold test. Thus, the authors concluded that the more objective physiological measures were not adequate in determining the level of pain, and that the correlations with respiratory rate and vocalization could be due to things other than pain (i.e. anxiety, fear, or anesthesia-induced delirium). However, it is also possible that the VAS and NRS scales that were used were not a valid and reliable assessment of pain. These VAS and NRS scales may have been measuring 1 aspect of pain, whereas the physiologic measurements were measuring another aspect of the pain experience.

There is an inherent element of subjectivity in many pain scales, and with the advancement of technology and diagnostic tools there are now more objective ways in which biological measurements may help us in diagnosing a degree of pain or lameness.

Physiologic Response

Measurements of physiologic parameters such as pulse, blood pressure, pupil dilation, and/or respiration rate are also thought to aid in pain assessment. There have been studies which showed that many of these physiologic parameters were unreliable indicators of pain.[1,12,16] However, while these physiologic parameters were not able to accurately predict pain independently, they could be used in conjunction with each other to give the clinician a better assessment of the animal's pain.

The study by Holton, et al.,[16] concluded that these parameters (heart rate, respiratory rate, and pupil dilation) were not useful indicators of pain in hospitalized dogs. However, I believe there to be certain fallacies in the study. For example, the validity of their gold-standard, a numeric rating scale (NRS), was based on its correlation with visual analogue scales (VAS) and simple descriptive scales (SDS) which were not known gold-standards. Also, it appeared that the dogs used in the study had multiple measurements taken and these observations were pooled, thus violating an assumption of independence in the statistical analysis. There were also different observers used for assessing pain, and there was no reference to inter-rater variability, specifically whether it was negligible or not.

The study by Conzemius et al.[12] also shows that certain objective pain measurements (heart rate, respiratory rate, blood pressure, and pain threshold test) do not correlate with subjective pain scores (VAS and NRS questions on pain). However, the validity of the VAS and NRS used in this study is questionable. The author states that the validity of using these scales was based on other VAS and NRS scales being reproducible in both human and animal studies. In my opinion, the reliability and validity of the comparison questionnaire should be determined if it is to be used as the gold-standard, otherwise it is not a true gold-standard.

Another important physiologic aspect of pain assessment is monitoring the patient's response to analgesia.[1] As analgesic agents are administered, changes in the patient should correlate with different levels of pain. This concept should also aid the clinician in assessing the animal's pain.

Gait Analysis

Pain can also be assessed to a certain degree by the amount of lameness exhibited by the dog. The lameness is assessed by analyzing the gait of the dog. The gait of the dog is a description of the motion of its limbs and has historically been somewhat subjectively evaluated. A visual gait analysis of the animal is usually done by the veterinarian as part of the orthopedic exam; however, there are now more objective ways in which gait can be described. There are currently 2 types of gait analysis: kinematic and kinetic; they can be used in conjunction with each other or used independently. Kinematic gait analysis employs the use of markers placed on the body and limbs and video cameras which relay information to a computer which then develops a three-dimensional graph that can be used for further analysis.

Kinetic gait analysis is accomplished by examining the forces exerted by a limb on a force plate (see Figure 1). For this reason, kinetic gait analysis is often referred to as force plate analysis. This kind of kinetic gait analysis provides the clinician with a quantitative assessment of the forces a dog exerts on a limb and a more objective interpretation of the gait. The forces recorded by each limb on the force plate are standardized so that they can be compared to other dogs. The force exerted by a limb is recorded in Newtons, divided by the dog's weight in Newtons, and multiplied by 100 to give a percentage of the animal's body weight which can then be compared to other dogs or other evaluations of the same dog.

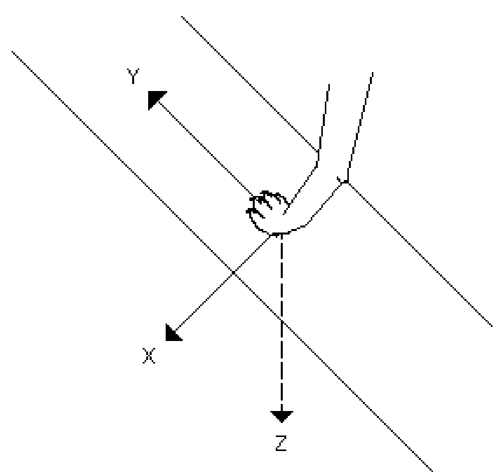

Figure 1 – Forces exerted by the dog's limb which can be recorded by the force plate.

Force plate analysis has had a growing acceptance in quantifying the severity of lameness in orthopedic patients. Force plate data has been collected on healthy animals and shown to correlate with morphometric measurements as expected.[17] Forces from healthy dogs have further been analyzed to see the amount of variation inherent in the different forces that can be measured.[18] This study along with others from clinically lame dogs (described below) show that the vertical force (Fz) and the craniocaudal forces (Fy) seem to be repeatable measures whereas mediolateral force (Fx) appears to be less reliable. Force plate data also shows that forelimbs appear to be more involved in decelerating or braking, whereas hindlimbs appear to be primarily used for propulsion.[19]

Other studies have clinically induced lameness in dogs to show that the force plate data was sensitive to this change and that a redistribution of forces occurs in other limbs when 1 is lame.[20-23] When 1 limb is lame, a decrease in the forces exerted by that limb may be present in addition to an increase in the forces placed on the contralateral limbs and/or the ipsilateral limb. Thus, the use of another limb as a control when estimating the amount of lameness in a limb is controversial and generally not thought to be good practice. Ideally, the amount of lameness should be calculated by comparing the forces of the suspected lame animal to the normal values for the animal with respect to breed, age, etc. However, since this information is not currently available, other means to describe the amount of lameness have to be implemented. Since different proportions of forces are exerted on forelimbs compared to hindlimbs and there is a redistribution of forces on limbs when lameness occurs,[20-23] a global outcome variable can be difficult to generate. However, when estimating the total amount of lameness exhibited by the dog, it may be possible to look at the absolute difference in hindlimbs and add the absolute difference seen in the forelimbs. This technique possesses certain sources of variation as well. For example, if there is a bilateral lameness in the hindlimbs and a compensatory effort made by the forelimbs, the amount of lameness would be negated. Likewise, if there was a bilateral lameness in the forelimbs, a compensatory effort by the hindlimbs would not be measured by this technique. Thus, this method would be ideal in describing the total lameness of the animal when the lameness is confined to 1 forelimb or 1 hindlimb.

Force plate analyses have also been used to determine the success of different types of surgical procedures in the canine. For example, Dupuis et al.[24] conducted a study which used force plate data along with other physiologic data to show that a fibular head transposition does not seem beneficial to the dog. Likewise, Budsberg et al.[25] used force plate data to quantify the success of an extracapsular repair of the ruptured cranial cruciate ligament. Jevens et al.[26] compared 2 techniques for cranial cruciate repair in dogs using force plate data and also found a correlation between the force plate data and a lameness scale that was used. McLaughlin et al.[27] used force plate analysis along with a lameness scoring system to show the efficacy of a triple pelvic osteotomy in treating hip dysplasia in the dog. Although there was not a formal comparison of the lameness score and the force plate data in McLaughlin's study, they did appear to be inversely related based on graphical representations. As the lameness score increased, the apparent force exerted by the limb decreased.

Another study by Budsberg et al.[28] evaluated ground reaction forces before and after a total hip replacement in the dog. They found that forces in the craniocaudal direction (Fy) appear to be unaffected by degenerative joint disease secondary to hip dysplasia. They also found that peak vertical force (Fz) may not be sensitive enough to detect changes in force in this type of population, suggesting the use of multiple forces including impulses when evaluating lameness via force plate. The impulse is simply the total force exerted by a limb and is depicted by a (Force x Time) curve. Also notable in this study by Budsberg was a correlation between the ground reaction forces and the lameness score. Although the coefficient of determination was moderate (the lameness score could only account for 60% of the force plate data), there was still a statistically significant correlation between the lameness score and the vertical force (Fz). This shows that while the subjective evaluation by the clinician is not without flaws, it does correlate with force plate data. It also suggests that additional information may be necessary when assessing lameness.

Another study by McLaughlin et al.[29] compared 2 different surgical approaches to the shoulder joint and used a lameness scale and ground reaction forces to compare the techniques. There was not a formal test of relationship between the force plate and the lameness scale. The relationship of the measurements was described as "poor" by the authors. The authors claim that differences in lameness scores between pre- and post-surgical intervention were not consistent with differences in force plate analysis. However, in my opinion, the interpretation of graphs presented by McLaughlin et al. seem to suggest that the lack of correlation may be due to the treatment group that was used in comparing the force plate data to the lameness score. The authors claim that the mean lameness scores indicated the dogs to be most lame 7-10 days after surgery, however, the graph of mean lameness scores tells us that the dogs appear to be most lame at 3-7 days after surgery. The largest decrease in mean peak vertical force in the operated treatment group also appeared to occur at 3 days after surgery. However, no statistical test of correlation appears to have been done. The power of this study may limit its usefulness also since there were only 10 dogs in the study (5 receiving new treatment, 5 regular treatment). Also, only the peak vertical force (Fz) was compared to the lameness scale and the authors suggest that other forces may have been better correlated with the lameness scale. This further supports the theory that lameness scales are not entirely accurate and that more information may aid in diagnosis of lameness. This also shows us that the choice of force may be important, and that all available forces should be analyzed.

Force plate data has also been used to show the efficacy of different drugs in controlling pain associated with lameness. In a study by Vasseur et al.[30] a nonsteroidal anti-inflammatory drug (NSAID) was compared to a placebo in a double-blinded clinical trial. Owners and veterinarians were asked if they thought the animal's condition was improved, worse, or unchanged. Force plate data was also used to verify a change in lameness. All forms of evaluation: owner, veterinarian, and force plate showed that the drug appeared to improve the dog's condition. Interestingly, the odds ratios for dogs having a positive response to the drug compared to the placebo were 4.2 for owner evaluation, 3.5 for veterinarian evaluation, and 3.3 for force plate evaluation (all odds ratios were also significant). These odds ratios are probably not significantly different from each other, even though there was no formal statistical comparison. Thus, the information provided by the owners appeared to be as valuable as that of the veterinarian and force plate in the diagnosis of lameness.

Force plate data has also been used in conjunction with different lameness scales.[24,26-30] However, there is a definite lack of agreement between some of these lameness scales and the force plate data. This could be due in part to the differences in lameness scales that are used. Another possibility is that there is a difference between what the lameness scales were actually measuring and what they were proven to be reliable in measuring. This also shows the necessity for a more valid and reliable instrument to measure pain and lameness.

Sources of Variation in Kinetic Gait Analysis

It has been shown that different sources of variation can influence force plate readings. A study by Jevens et al.[31] calculated the coefficient of variation (CV) for the major forces measured by the force plate. The peak vertical (Fz) force and its associated impulse appeared to have the least amount of variation at 5.8% and 8.5% respectively for 5 dogs who ran across the force plate 5 times. The craniocaudal forces (Fy) of the forelimbs had a CV of 26.4% and 30.5% for their respective peak force and impulse. The craniocaudal forces (Fy) of the hindlimbs had a CV of 63.0% and 25.9% for their respective peak force and impulse. The amount of variation associated with the dogs ranged from 14-69%, and the amount of variation attributed to repetition ranged from

29-85%. The authors suggested using approximately 5 repetitions per dog to increase the power of the experiment. This was estimated by assuming that the cost of including another dog was 50 times greater than using another repetition. The amount of variation attributed to the different handlers was between 0-7%. This study shows that vertical force (Fz) and its associated impulse are probably the most accurate followed by craniocaudal force (Fy) and its associated impulse.

Other sources of variation include the familiarity of the dog with the force plate setup and routine.[32] The day in which the force plate data is acquired is important also. Rumph et al.[33] shows that significant differences in forces can be found from 1 day to the next. This suggests that force plate data and questionnaire data should be collected on the same day. The velocity of the dog and stance time (the amount of time the limbs are in contact with the force plate) also need to be controlled for the force plate to give accurate results. As the velocity increases, the peak vertical force (Fz) increases, the stance time decreases, and the vertical impulse decreases.[34-36] Thus, either a constant velocity or constant stance time[37] should be maintained to reduce variation in the force plate data.

Repeatability and Validity

Repeatability or reliability is often defined as the extent to which a measurement can be repeated under identical conditions.[38,39] One method of testing the repeatability or reliability of a measurement is using the test-retest method. This test-retest method uses the same individual for 2 measurements.[39] This seems to be 1 of the most logical choices for a method to assess reliability because it reduces the amount of variability that may be introduced by other variables when using multiple individuals. In essence, the individuals are matched upon themselves. In this way, additional variables which may be unknown to the investigator are removed and additional sources of variation will be avoided. Hopefully there is no change in any factors other than those being analyzed between 1 measurement and the next. However, a lack of agreement in a test-retest assessment could have implications other than a lack of reliability.[39] The lack of agreement could signify that a change has occurred in the factor being measured rather a decrease in reliability. Time also plays an important role in test-retest assessment. Too much time between tests may allow the factor to change which was hoped to remain constant. On the other hand, too little time may introduce a recall bias, such that the subject remembers their answers to the first assessment. Another possible source of error in test-retest methodology is that a subject may have a heightened sensitivity to certain factors after the first assessment thus influencing the second assessment. Thus, when a lack of reliability is found using the test-retest method, other possibilities should be considered before ruling that the test being analyzed is unreliable.

The validity of a measurement is described as the degree to which an instrument measures that which it was intended to evaluate.[38,39] There are many different forms of validity which can be used when describing a measurement. *Construct validity* is simply the extent by which a measurement corresponds to a theoretical concept or construct.[38,39] *Content validity* is referred to when discussing the extent by which a measurement encompasses all possible variables associated with a health event. *Criterion validity* is used when discussing the correlation between the measurement in question and another external measurement of the same phenomenon. Criterion validity can be further broken down into concurrent validity and predictive validity. *Concurrent validity* is used when the measurement and the external criterion are done at the same time. On the other hand, *predictive validity* refers to the extent by which a measurement is able to predict some external criterion which will occur in the future.[38,39]

Human Literature

The test-retest method of assessing reliability has been used extensively in the human literature. While there is additional work required in the data collection of a repeated assessment, it is still regarded as the easiest way to evaluate reliability. Usually, only a subset of the sample population is required in the test-retest assessment. For example, in the validation of the West Haven-Yale Multidimensional Pain Inventory,[40] a subset of 60 individuals out of the original sample population of 120 was used in the test-retest assessment which required a second assessment 2 weeks after the original assessment. Likewise, the Juvenile Arthritis Functional Status Index (JASI) was determined to be reliable based on 2 different test-retest assessments of smaller subsets of individuals.[41,42] An interval of 1 week between tests was used for the first study[41] whereas the second[42] used an interval of 2-3 weeks and a 3 month interval to show it was reliable.

The time period between the initial assessment and the repeated assessment should be chosen carefully due to the possibility of introducing different kinds of bias (see *Repeatability and Validity*). For example, the reliability of the Pediatric Pain Questionnaire (PPQ) was done using the test-retest method over a 6 month period.[43] One of the stipulations of the test-retest method is that there is not a significant amount of change in the factors being analyzed. With an increase in the time interval, this problem usually grows. In the test-retest of the PPQ, it was possible that the level of pain changed in a 6 month interval. However, other measurements of the patient's disease and functionality had a significant correlation indicating that the reliability testing was unlikely to be adversely affected.

The theme of proving the construct validity of a questionnaire based on a correlation with the theoretical categories the questionnaire was designed to answer is quite common in the human literature.[39-41,44] This is primarily done by use of either factor analysis or a principal component analysis which shows that the common factors extracted from the questionnaire correlate with the predetermined categories.

The concurrent criterion validity of a questionnaire can also be demonstrated by the use of correlations with other questionnaires that have been previously validated.[45-48] However, problems can exist if the validity of the questionnaire which is used as a "gold-standard" is in question. For example, the Pediatric Pain Questionnaire (PPQ) was based on the McGill pain questionnaire which has a long-standing reputation of being a valid and reliable measure of pain.[43] The PPQ was then validated concurrently with parent, patient, and doctor assessments along with a correlation of activities of daily living and a disease activity index. These measurements of activities provide an objective measurement, but are definitely not a "gold-standard" of pediatric pain measurement. The Total Quality Pain Management (TQPM) was later developed to assess a child's postoperative pain and was validated against the PPQ.[47]

It is my belief that with the advent of new technology and new techniques, the scientific community should be able to more objectively validate these questionnaires using a combination of available inputs. More objective and quantitative measures have been used in the human literature to show the validity of some questionnaires. For example, the JASI was validated based on grip strength, joint counts, and a timed walk/run.[42] The validity of the Hughston clinic subjective knee questionnaire was shown with a correlation between the questionnaire and several kinematic variables.[49] The Gillette Functional Assessment Questionnaire has also been validated by correlating the questionnaire to kinematic data.[50]

Questionnaire

The questionnaire used in this experiment was developed prior to my involvement. It was developed based on some of the same criteria of human pain scales which may also apply to canine behavior. Specifically questions were asked in the following categories: 1) an overall assessment, 2) interaction with the owner, 3) general physical activity and exercise level, and 4) a listing of specific activities. The questionnaire was somewhat lengthy as it tried to find any and all possible associations with pain. In a preliminary test-retest assessment of the questionnaire, it was shown to have an excellent agreement in more than half of the questions, and poor agreement in less than 15% of the questions.[51]

Questionnaire Reduction

In statistics there often many roads which will lead to the same conclusion, and as responsible scientists we must try to use the most applicable and relevant procedure. In statistical modeling, the fewest number of variables which adequately represents the situation is desired. The smaller model allows for the easiest interpretation and most practical use in the real world. There are many different statistical methods that can be used to decrease the number of variables in a model. However, there should be some information about the model which is known to the investigator *a priori*. This information should be used in conjunction with other statistical methods in creating a valid model.[52] If models are only analyzed using statistical procedures, then the result may be a model that lacks information that is known to be important.

Principal component analysis (PCA) is a method of decreasing the number of variables by reducing the dimensionality of many variables that may be correlated with each other.[38,39,53] This procedure takes the entire set of independent variables, which are thought to be correlated, and creates a new set of principal components which are uncorrelated while retaining as much of the variation explained by the original variables. Ideally, PCA is done as an exploratory technique to describe themes in a data set. Thus, it can aid in showing construct validity if the principal components extracted correlate to theoretical constructs. For each of these principal components, a loading factor corresponding to independent variables is calculated which represents the degree by which this principal component is represented by each independent variable. This matrix of principal components and independent variables is usually rotated such that independent variables load higher in 1 component than others and allows easier interpretation. It is conceivable that each of the principal components could be represented by those variables that had high loading factors. The result would be a reduced model that accounts for nearly as much of the variation as the original model.

Another method of reducing the independent variables is by using selection procedures in the statistical packages known as stepwise, forward, and backward selection.[54] The stepwise procedure goes through a list of possible variables and adds the variable with the most significant coefficient and removes the least statistically significant variable, and then looks at the remaining variables and the procedure is repeated. The forward selection method simply starts with no variables in the model and adds variables to the equation that are significant. The backward selection method starts with all possible variables in the model and removes the least significant variables.

Multicollinearity can cause problems in interpreting a model.[54] Multicollinearity occurs when a supposedly independent variable is highly correlated with another independent variable. The problem is that the predictor variables are thought to be independent of each other and their coefficients are thought to measure the amount of change in the dependent variable related to a change in only that predictor variable. Thus, variables can also be removed from a model if they are highly correlated with others.

Conclusion

It was the purpose of this project to establish an instrument that could be used by a veterinarian to aid in their diagnosis of an orthopedic disorder accompanied by lameness and to quantify the degree of lameness. The instrument used was a questionnaire given to clients who could attest to the animal's abilities in the home environment. Ground reaction forces obtained from a force plate were used as a "gold standard" in this experiment as a measurement of the amount of mechanical lameness exhibited by the dog.

The aims of this study were to 1) test the questionnaire on a large number of dogs and further evaluate its reliability, 2) test the concurrent criterion validity of the questionnaire using the force plate measurements, and 3) to develop and evaluate the validity of a shortened questionnaire.

MATERIALS AND METHODS

Selection of Subjects

Animals used for this study were client-owned animals seen at the Texas A&M Veterinary Medical Teaching Hospital. Dogs of any breed or sex were eligible to participate in the study provided they showed clinical signs of lameness. The dogs chosen were determined to be skeletally mature based on the opinion of an orthopedic research scientist. Dogs were generally ≥ 13.6 kg [30 lb] so that 2 individual limb strikes could be recorded for force plate analysis.

All data for this study were previously collected as part of other projects (to be described). The dogs used came from 3 different cohorts of dogs, 2 of which were also used for other studies. The first cohort was a group of dogs used as a pilot to test the reliability of the questionnaire. The other 2 cohorts were then used to increase the power of the reliability evaluation and to test the questionnaire's validity. These cohorts consisted of dogs enrolled in a nutraceutical clinical trial and a weight loss study (to examine the effects of weight loss on lameness). Consequently, the period of time between evaluations, the inclusion criteria, the criteria for being reevaluated, and the examination procedures were different for each cohort. Dogs were recruited by the orthopedic specialist who saw the client if the dog fit into the inclusion criteria of 1 of the different cohorts.

Study 1

The first group of pilot dogs was used to test reliability of the questionnaire; that is, these dogs were chosen because their lameness was thought to be relatively constant. This was quantified with force plate analysis. If the dog had less than a 10% change in its peak vertical force its lameness was considered to be the same. This study was conducted between June and November, 1997. Dogs were included in this cohort if they had degenerative joint disease in any limb based on radiographic and clinical documentation. The dogs used in this pilot could have experienced lameness due to a variety of reasons, and could be any age, breed, or gender. The initial assessment of the pilot dogs included an orthopedic examination, questionnaire completion by the owner, and force plate evaluation. The owners of the pilot dogs were then instructed to bring their dogs back in 1 or 2 weeks to repeat the assessment. Fifteen dogs met the inclusion criteria and were used in our study. Of those 15 dogs, 11 showed a change of less than 10% in their peak vertical force and were used in the test-retest assessment of reliability.

Study 2

This study was conducted between April 1999 and May 2001. There were several criteria by which dogs from the nutraceutical study were included. A dog of any breed or gender, between 2 and 10 years of age, with a weight between 23 kg and 41 kg [50 - 90 lbs] were included if the following criteria was also met. The dog had a body condition score between 2 and 4 (out of 5). The dog exhibited chronic, recurring lameness for more than 4 months of 1 limb due to either degenerative joint disease secondary to hip dysplasia or repair of a unilateral cruciate-deficient stifle by fibular head transposition within the last 1 to 6 years. A lameness score between 2 and 4 (out of 5) was obtained. Radiographic evidence of osteoarthritis in 1 or more joints of the affected limb was also present. To be included in the study, ground reaction forces of the affected limb obtained from the force plate were less than the contralateral limb. Owners of the dogs must have signed an informed consent, evaluated the dog at home, and brought the dog back for scheduled reevaluation.

Dogs were excluded from the nutraceutical study if any of the following occurred: the dog received any nonsteroidal ant-inflammatory drug (NSAID) within the last 14 days, any injectable steroid within the last 8 weeks, any oral steroid within the last 8 weeks, any chondroprotective agent or nutraceutical product in the last 12 weeks, or any sulfonamide or tetracycline antibiotics within the last 7 days. The dog was also excluded if it was pregnant or medically ill as determined by clinical or laboratory evaluation. If the dog was lame or had gait abnormalities due to an immunologic, neurologic, infectious, or neoplastic condition, it was not included in the study. If the dog had received surgery within the last 6 months (however, elective neutering was acceptable) it was excluded. An initial lameness score of either 1 or 5 (out of 5) also excluded the animal.

Dogs included in Study 2 were initially assessed by a physical examination, CBC, serum biochemical profile, urinalysis, subjective orthopedic examination, radiographic evaluation, questionnaire completion, and force plate analysis. The dogs were then randomly prescribed either the nutraceutical or a placebo. The owners were instructed to bring back their dogs in 4 weeks and again 8 weeks from the original assessment. Upon return, the dogs were assessed in a similar fashion. Twelve dogs met all the inclusion criteria and were also used for our study of the lameness questionnaire. None of these dogs were used in the reliability evaluation because either a treatment or placebo was given between initial assessment and the next assessment. Thus, similar conditions necessary for a test-retest evaluation were not available.

Study 3

The third cohort of dogs came from a weight loss study which took place from June 1998 until November 2000. Dogs were included based on lameness due to hip osteoarthritis and the dog being overweight. Dogs were classified as overweight if they were ≥15% above their optimal body weight, had a body condition score of ≥4 out of a possible 5, and an estimated body fat of ≥25% as assessed by a board certified nutritionist. The diagnosis of hip osteoarthritis was based upon orthopedic examination and radiographs. Dogs were any gender and were required to be over 2 years old, between 18 kg and 55 kg [40 and 120 lbs], and leash trained. The dogs' history of osteoarthritis was determined as stable by the orthopedist. The owners provided informed consent and agreed to comply with a weight loss regimen which consisted of both diet and exercise. Dogs with osteoarthritis in any other joint, pain in another joint, or with neurologic diseases were excluded from the study. Initial assessment included a subjective orthopedic examination, radiographic evaluation, questionnaire completion, and force plate analysis. The dogs were then reevaluated when they lost 1/3 of their excess body weight, again when they lost 2/3 of the excess body weight, and when they reached their final optimal weight. Many dogs were also reevaluated 2 weeks after their initial examination and before the weight loss regimen began so that these dogs could be included in testing the reliability of the questionnaire before any known change occurred. Seventeen dogs met the inclusion criteria and were used in our study of the lameness questionnaire. Of these 17 dogs, 9 came back for a second evaluation prior to any diet change; 8 of these also showed a change of less than 10% in their peak vertical force and used in the test-retest assessment of reliability.

See Figure 2 below for a graphical presentation of the source of the animals for our study of the lameness questionnaire.

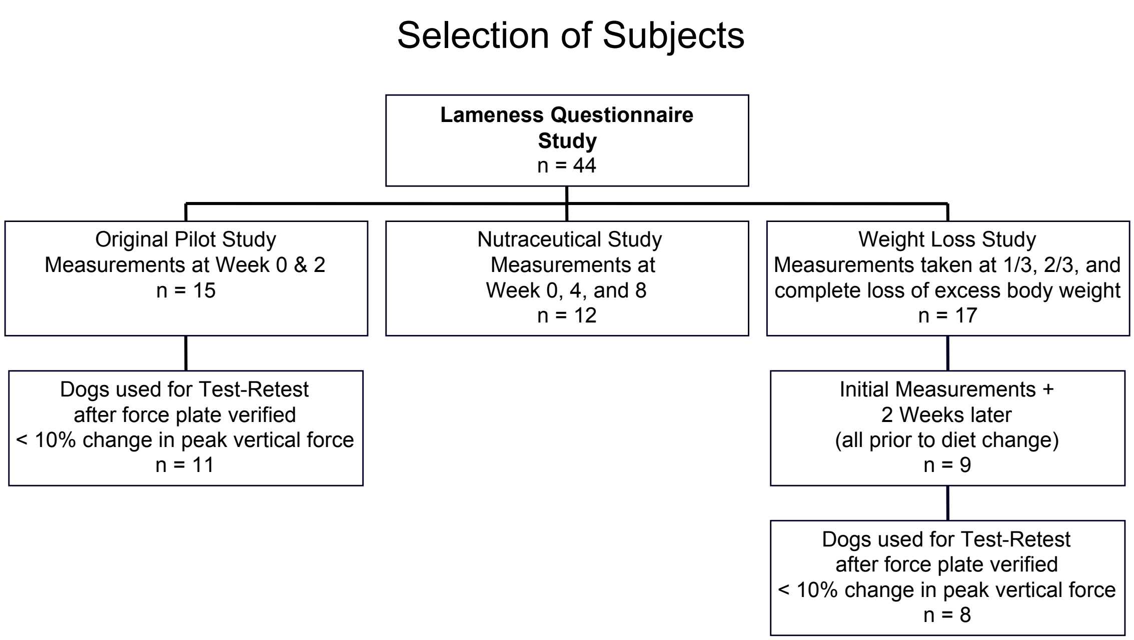

Figure 2 – Source of subjects for this study were enrolled in other studies. The source of subjects for the test-retest assessment of reliability is shown.

Questionnaire

The questionnaire consisted of 39 questions which followed the format of a 10cm Visual Analog Scale (VAS). Some of the questions followed the examples: “How willing is your dog to play voluntarily?”, “How difficult is it for your dog to lie down?” There was a corresponding VAS line for each question with descriptors at each end such as “not at all” vs. “very willingly” or “difficult” vs. “easy.” See *Appendix A* for more examples and the entire questionnaire. One open-ended question was included to allow the owner to list activities that may not have been captured by the questionnaire. There was also a question that asked the client which questions they perceived as most important. The response to this question was not evaluated further in this project.

Lameness Assessment Using Force Plate

The force plate data was collected in the Canine Lameness Assessment Laboratory at the Texas A&M Veterinary Medical Teaching Hospital. The force plate was an AMTI OR-6-5 force plate and was installed centrally and flush along an 11m runway and covered with a thin carpet that also served to camouflage the actual force plate. Signals from the force plate and its integrated photoelectric cells proceed through a Vishay 2200 amplifier and then into a designated IBM compatible 486 computer. The data from each foot strike was recorded for 650 milliseconds at 2 millisecond intervals. Forces were analyzed using the Acquire package from Sharon Software. Peak vertical, cranial/caudal, and medial/lateral forces with their associated impulses were recorded.

The dogs were familiarized with the Canine Lameness Assessment Laboratory for a few minutes and walked across the runway a few times before the actual trials were recorded. A valid trial consisted of the dog being led at a trot across the runway in which both the forelimb and ipsilateral hind limb contacted the surface of the force plate. The forces used for analysis were peak vertical, cranial-caudal, and their associated impulses. The maximum slope values of the vertical force curve during the rising and falling of the curve were also recorded on those dogs seen following the software upgrade. Vertical forces are those exerted in the Z direction (see Figure 1) and were recorded by either the peak/maximum force detected or the impulse, which is the force measured over time. Cranial-caudal forces are measured in the Y direction (see Fig. 1) and likewise were either a peak/maximum value or an impulse value. The cranial-caudal forces were also further dichotomized into propulsion and braking forces that each limb exerted. The difference between propulsion and braking depends on the direction in which the force was exerted in the Y direction. These forces were standardized by dividing the force by the dog's weight. A minimum of 5 trials was recorded for both the left and right sides of the animal during each assessment, and the average of these trials was used to assess lameness. A constant velocity [of 1.6-2.1m/s] and minimum acceleration [$0 \pm 0.5 m/s^2$] was required for all valid trials and was determined by the use of photoelectric cells. The same handler was used throughout the study. All trials were videotaped to record the dog's gait and verify acceptable limb strikes.

Analysis of Reliability

The questionnaire was tested for reliability using a test-retest method. The force plate data was used as the gold-standard in predicting changes in and the extent of lameness for this study. Those dogs whose degree of lameness remained unchanged from 1 assessment to the next based on force plate data were used to test the reliability of the questionnaire. A change of less than 10% in the vertical peak force (Fz) was the criteria by which a dog was considered unchanged with respect to lameness. Only dogs from the pilot cohort and from the weight loss study (who had a repeated evaluation before diet began) were used for examining the reliability because there may have been a placebo effect for those dogs undergoing a treatment which could skew the results of the questionnaire (n=19). The questionnaire data, which consist of continuous variables for each question, were compared using the Spearman rank correlation[55] since the data were not normally distributed. Questions were considered repeatable if there was a response from at least 18 owners; these questions had a correlation of at least 0.6, and a significant p-value (less than 0.05).

Analysis of Validity

The concurrent criterion validity of the questionnaire was tested by comparing the assessments of dogs from all cohorts which had some degree of lameness determined by force plate analysis. The first questionnaire from all cohorts of dogs was used for this portion of the analysis since it was before any treatment was given, and all dogs had some degree of lameness perceived by the owner. Although multiple measurements were taken on the dogs at different points in time, these measurements were not used in further analysis in this study because the measurements were not truly independent. Also, the criteria by which the repeated measures were taken were not uniform across all dogs.

A multiple linear regression model was created to determine if the questionnaire could accurately predict the degree of lameness depicted by the force plate data. Each question represented a continuous variable with a corresponding score of 0.00-10.00. Multiple dependent variables and their respective regression models were possible since multiple forces were recorded by the force plate. In this study, all animals had hindlimb lameness which was confirmed with the force plate; although, it was possible that some animals also had forelimb lameness based on the inclusion criteria of the pilot cohort. As it turned out, 1 dog was diagnosed as having forelimb lameness, and 1 other dog was diagnosed has having lameness in both the fore and hindlimbs which was disproportionate among limbs and not consistent between assessments. The amount of lameness recorded by each force measurement (peak vertical, vertical impulse, etc.) was tabulated by taking the absolute difference in the hindlimb forces and adding it to the absolute difference in forelimb forces. This calculation was thought to estimate a degree of lameness exhibited by the dog by quantifying the redistribution of forces (see Table 1 for a sample calculation). Thus, for the data presented in Table 1, a value of 13.36 would be used as the value for the dependent variable "vertical (Z) peak" which will subsequently be called "Z Peak Diff" and represents the absolute percent change in body weight based on the vertical force.

Table 1 – An example of how forces were used in calculating a dependent variable for each type of force measured.

LIMB	VERTICAL (Z) PEAK	Absolute Difference in % Body Weight	Total Difference in % Body Weight
LF	99.89	Forelimbs	13.36
RF	97.23	2.66	
LH	58.47	Hindlimbs	
RH*	47.77	10.70	

* Lame Limb

Different linear models were developed using the total amount of lameness detected by each of the forces as dependent variables and each of the remaining repeatable questions as an independent variable. These were named in a similar fashion; for example, "Z Impulse Diff" was used for the absolute differences in percent of body weight based on the vertical impulse. "Prop Peak Diff" was used to denote the absolute differences in percent of body weight of the peak propulsion force in the Y direction, and "Prop Impulse Diff" was used for the absolute difference in percent body weight of the impulse propulsion force. Likewise, "Brake Peak Diff," and "Brake Impulse Diff" are the dependent variables used to represent the absolute differences in percent body weight of forces in the Y direction (opposite of propulsion). The dependent variables "Rise Slope Diff" and "Fall Slope Diff" were also created in the same manner to represent the maximum values of the slope of the vertical force curve when it is rising and falling.

For each of these full models there were 19 independent variables which each represented a question. There were also observations from 44 dogs used, except for those models that used "Rise Slope Diff" and "Fall Slope Diff" as the dependent variable. For these models, only 29 dogs were used because the software to detect this measurement was not acquired until later in the data collection.

The assumptions of multiple linear regression were evaluated. These assumptions were based on the value of the residuals, which is the vertical distance between an observed case and the regression line. The assumptions were that the mean of the residuals equal to 0, and that the residuals were both independent and normally distributed. Some transformations were necessary to maintain these assumptions and were obtained using the Box-Cox method.

Questionnaire Reduction

Principal Component Analysis (PCA) of all repeatable questions was the original goal to see if the components extracted were the same as those categories by which the questionnaire was developed. However, due to sample size restrictions this could not be done. Approximately 95 dogs would be needed to use PCA on 19 variables. Instead, questions were grouped into those areas representing "owner assessment," "mobility," and "behavior." Each of these groupings were analyzed using PCA.[55] As with the linear regression models, all first questionnaires of dogs were used for the PCA. However, it was possible to include 4 more dogs in this procedure because there were 4 dogs from the pilot study in which questionnaire data was obtained, but no corresponding force plate measurements were obtained. All other inclusion/exclusion criteria were adhered to for these dogs. Thus, 48 dogs were used for the PCA. The principal components extracted had to have an eigenvalue of greater than 1 to be used. The resulting factor loading matrix was then rotated using a varimax rotation such that the independent variables would appear to load higher in 1 component than the others and allow for easier interpretation. A factor loading of greater than 0.6 was considered to load onto the corresponding factor.

Spearman correlations were also determined in order to compare all questions against each other. As with the reliability test, Spearman correlations were used since the data were not normally distributed. The reduction of the questionnaire was accomplished by examining the Spearman correlations among questions, the principal component analysis of groups of questions, and the results of different variable reduction procedures in linear regression.[55]

The value of the coefficient of determination, R^2, was used along with the adjusted R^2 value and the p-value for the regression model to determine the fit of the model. Models were also built on some trial and error techniques to see the change in the model fit and the significance of the coefficients. A priori knowledge of the areas in which the questionnaire was designed to measure was also considered in the reduction of variables.

Multiple linear regression models were created from the reduced questionnaire. These results were then compared to the model from the full set of repeatable questions to determine if the shortened questionnaire accurately predicted the degree of lameness in the dog.

A total score based on the VAS responses from the shortened questionnaire was also calculated to determine if this would adequately represent the lameness detected by the force plate. Spearman rank correlations between this total VAS score and force plate data were analyzed.

RESULTS

Repeatability

From the 2 cohorts of dogs used in assessing repeatability, 24 dogs (15 from the pilot study and 9 from the weight loss study) were considered eligible for the test-retest assessment. However, only 19 of these dogs (11 from the pilot study and 8 from the weight loss study) were verified by the force plate as having less than a 10% change in peak vertical force and used in the repeatability analysis.

From the original 39 VAS questions in the questionnaire, 19 (70%) were found to be repeatable. These 19 questions had a Spearman rank correlation coefficient ranging from 0.68 – 0.90. The questions also had a $p < 0.001$. There was a 100% response rate to all questions considered repeatable except for 1 in which 18 out of 19 (95%) clients responded. The predetermined criteria for 18 cases being required for the question to remain in the model resulted in the removal of only 6 questions in which no other criteria were violated (correlation > 0.6 and $p < 0.05$). These 6 questions had a response of no more than 14 clients from the pool of 19. See Table 2 regarding the removal of questions that were not repeatable.

The categories of questions which were used to develop the questionnaire (an overall assessment by the owner, interaction with the owner, general physical activity and exercise level, and various specific activities) were all represented in the remaining repeatable questions. Further scrutiny of those questions remaining in the predetermined categories led to a better definition of the categories of remaining questions. The category of "overall assessment by the owner" was unchanged. The category of "interaction with the owner" seemed to be better defined as "the dog's attitude/behavior." The categories of "physical activity and exercise level" and "various specific activities" were lumped together into a more representative category termed "the dog's mobility."

Table 2 – Repeatable questions based on correlation coefficient, p-value, and n.

Question	Spearman rank correlation coefficient	N	P-value	Remaining questions	Group
1a-Overall(month)	0.741	19	< 0.001	1a-Overall(month)	Owner
1b-Overall(week)	0.688	19	0.001	1b-Overall(week)	Owner
2a-Mood(month)	0.688	19	0.001	2a-Mood(month)	Behavior
2b-Mood(week)	0.681	19	0.001	2b-Mood(week)	Behavior
3a-Attitude(month)	0.736	19	< 0.001	3a-Attitude(month)	Behavior
3b-Atttitude(week)	0.808	19	< 0.001	3b-Atttitude(week)	Behavior
4-Happy dog post	0.742	18	< 0.001	4-Happy dog post	Behavior
5-Frustrated Owner	0.605	13	0.028		
6-Offer Spec Treats	0.361	17	0.154		
7-Massaged Dog	0.893	12	< 0.001		
8-Amt inter w/family	0.346	17	0.173		
9-Type inter w/family	0.213	18	0.397		
11-Chg Freq activities	0.495	19	0.031		
12-Chg Amt activities	0.687	19	0.001	12-Chg Amt activities	Behavior
13-Freq other activities	0.398	19	0.092		
14a-Play Voluntarily	0.886	19	< 0.001	14a-Play Voluntarily	Behavior
14b-Chg Vol Play	0.510	19	0.026		
15a-Amt Vol Exer	0.468	19	0.044		
15b-Chg Vol Exer	0.086	18	0.734		
16-Often get Exercise	0.800	19	< 0.001	16-Often get Exercise	Mobility
17a-Difficult jumping	0.577	19	0.010		
17b-Chg amt assist	0.265	16	0.321		
18a-Diff lie down	0.393	19	0.096		
18b-Diff sitting	0.716	19	0.001	18b-Diff sitting	Owner
19a-Diff rise from lying	0.903	19	< 0.001	19a-Diff rise from lying	Owner
19b-Diff rise from sitting	0.482	19	0.037		
20a-Diff squat urine/def	0.689	19	0.001	20a-Diff squat urine/def	Owner
20b-Diff lift leg urinate	0.939	10	< 0.001		
21a-Walk up stairs	0.819	14	< 0.001		
21b-Prev walk up stairs	0.632	13	0.021		

Table 2 – Continued.

Question	Spearman Rank correlation coefficient	N	P-value	Remaining questions	Category
22a-Walk down stairs	0.758	14	0.002		
22b-Prev walk dn stairs	0.650	13	0.016		
23-Vocally indicate	0.455	19	0.050		
24a-Stiff rising for day	0.782	19	< 0.001	24a-Stiff rising for day	Behavior
24b-Stiff at end of day	0.744	19	< 0.001	24b-Stiff at end of day	Behavior
25-Indic Lame (walk)	0.819	19	< 0.001	25-Indic Lame (walk)	Mobility
26-Indic Lame (trot)	0.799	19	< 0.001	26-Indic Lame (trot)	Mobility
27-Indic Lame (run)	0.870	19	< 0.001	27-Indic Lame (run)	Mobility
28-Pain turning (walk)	0.706	19	0.001	28-Pain turning (walk)	Mobility

Validity

The validity of the 19 repeatable questions was then ascertained. This was done by setting up multiple linear regression models with the 19 questions used as independent variables, and the different forces measured as the possible dependent variables. Upon initial analysis the regression model using "Z Impulse Diff" as the dependent variable had an R^2 value of 0.926, an adjusted R^2 of 0.859, and a $p < 0.001$. However, further analysis found that this was due to the inclusion of 2 independent variables (questions 3a and 3b) which were highly collinear. The graphical representation also did not seem to agree with these results. When 1 of these variables was omitted, the model lost a large amount of predictive power, and the variable included (3a or 3b) became insignificant. Further analysis showed that the inclusion of these 2 variables was due to a highly influential case in which there was a large difference between these 2 values. Since these 2 questions were very similar (one asking for an assessment of attitude in the last month, and the other for an assessment of attitude in the last week), it was thought that these questions should be highly correlated and that there should not be a huge discrepancy. Since this discrepancy was causing results that were counterintuitive, the responses for these 2 questions in the data were omitted from further analysis for that 1 case.

After adjusting for this problem, 8 multiple linear regression models were run using the 19 repeatable questions. These 8 models were based on the 8 different forces that were tabulated for each limb strike and were subsequently used as the dependent variables. Some transformations of the dependent variables were necessary to accommodate model assumptions. The results can be seen in Table 3. Note that for the forces of "Rise Slope Diff" and "Fall Slope Diff" there were only 29 observations due to the inability of the software to determine this at the beginning of the study. This could account for the high value of R^2 of "Rise Slope Diff" and "Fall Slope Diff," and the insignificance of the regression models ($p>0.05$). All assumptions associated with multiple linear regression were met for all models except for "Brake Peak Diff" as the dependent variable. For this model, no transformation could be found that would result in the validation of all assumptions, and the summary statistics for the fit of this model did not seem to warrant further investigation (the R^2 value was rather low, the adjusted R^2 value was very low, and the model was not statistically significant).

Table 3 – The summary of models using the 8 possible dependent variables and the 19 repeatable independent variables using 44 dogs.

Full Model Summaries			
Dependent Variable	**R-squared**	**Adjusted R-Square**	**P-value for Regression**
Z Peak Diff	0.731	0.488	0.008
Z Impulse Diff [a]	0.680	0.390	0.031
Prop Peak Diff [b]	0.663	0.359	0.043
Prop Impulse Diff	0.575	0.191	0.185
Brake Peak Diff	0.516	0.079	0.354
Brake Impulse Diff	0.427	-0.092	0.664
Rise Slope Diff [a] *	0.851	0.497	0.103
Fall Slope Diff [c] *	0.862	0.536	0.081
* n = 29			
Transformations on Y using: [a]Sqrt(y), [b]1/Sqrt(y), [c]1/y			

The scatterplots of the dependent variable vs. the predicted value of the dependent variable are shown in Figures 3-5 for the significant regressions using "Z Peak Diff," "Z Impulse Diff," and "Prop Peak Diff." These graphs represent the model with all 19 independent variables and 44 observations.

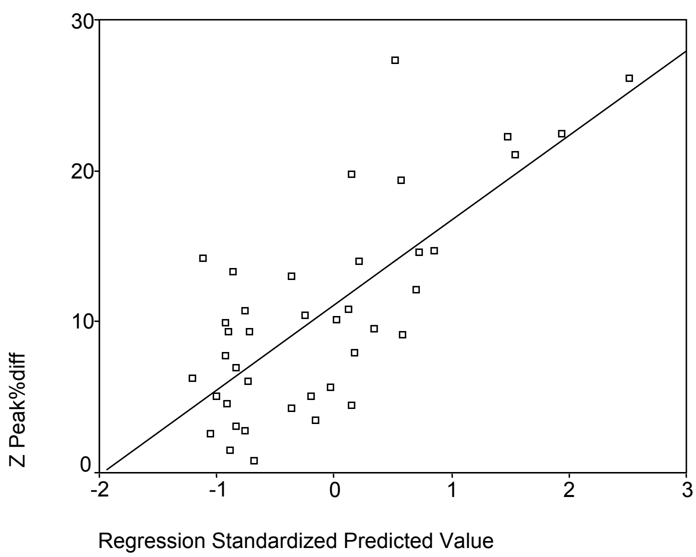

Figure 3 – Scatterplot of the absolute difference in the peak vertical force among limbs against its predicted value using the regression model with the 19 repeatable questions and 44 observations.

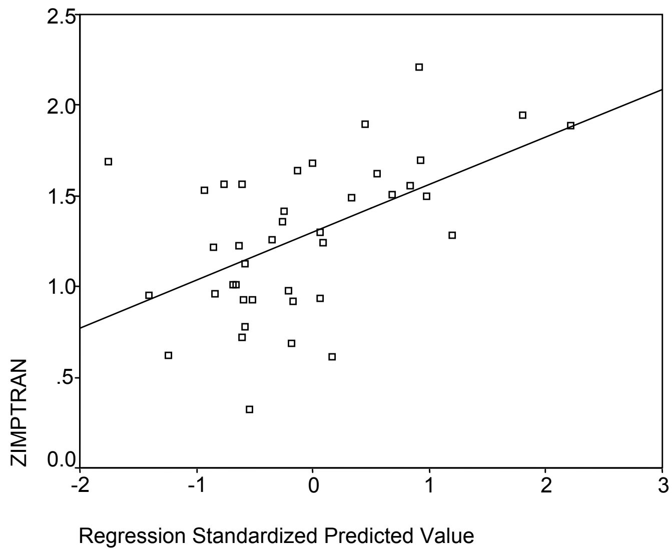

Figure 4 – Scatterplot of the absolute difference in the peak impulse force among limbs against its predicted value using the regression model with the 19 repeatable questions and 44 observations.

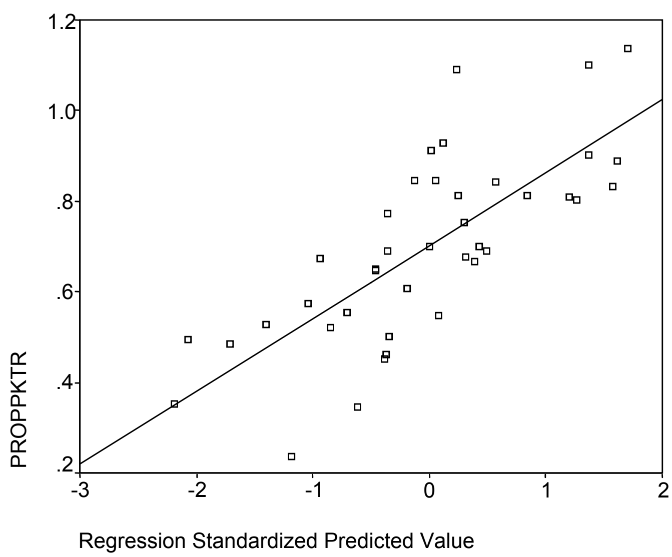

Figure 5 – Scatterplot of the absolute difference in the peak propulsion force among limbs against its predicted value using the regression model with the 19 repeatable questions and 44 observations.

Variable Reduction

Step 1

Regression models were run using each group of questions to get an idea of how well each group was represented by each of the best dependent variables (from Table 3: Z Peak, Z Impulse, and Prop Peak). These models were run with questions coming from each group as identified in Table 2. See Table 4 below for a further description.

Table 4 – Regression models based on the groupings of questions using the best dependent variables previously identified by using full models.

Group of Questions	Dependent Variable	R-squared	Adjusted R-Square	P-value for Regression
Owner Assessment	Z Peak Diff	0.196	0.087	0.136
Owner Assessment	Z Impulse Diff	0.257	0.157	0.043
Owner Assessment	Prop Peak Diff	0.143	0.027	0.314
Behavior	Z Peak Diff	0.465	0.315	0.009
Behavior	Z Impulse Diff	0.461	0.310	0.010
Behavior	Prop Peak Diff	0.277	0.073	0.247
Mobility	Z Peak Diff	0.279	0.179	0.032
Mobility	Z Impulse Diff	0.286	0.187	0.027
Mobility	Prop Peak Diff	0.337	0.244	0.009

Step 2

A principal component analysis was conducted on each group of questions to help identify the relationship among the variables contained in each grouping. The "Overall Assessment by Owner" category resulted in only 1 principal component (see Table 5). The category of "Dog's Attitude/Behavior" resulted in 3 principal components being extracted (see Table 6), and the category of "Dog's Mobility" resulted in 2 principal components being extracted (see Table 7). The factor loadings of the independent variables (questions) can be seen in the tables below.

Table 5 – Principal component analysis on the overall owner assessment group of questions.

Variables	Factor loadings for Component A
1a-Overall (month)	**0.849**
1b-Overall (week)	**0.877**
18b-Diff sitting	**0.743**
19a-Diff rise from lying	**0.783**
20a-Diff squat urine/def	**0.836**
Factor loadings >0.6 in bold.	

Table 6 – Principal component analysis on the behavior group of questions.

	Factor loadings for Components		
Variables	B	C	D
2a-Mood (month)	**0.917**	0.172	0.111
2b-Mood (week)	**0.861**	0.185	0.070
3a-Attitude (month)	**0.915**	-0.056	0.195
3b-Attitude (week)	**0.918**	-0.059	0.193
12-Chg Amt activities	0.134	**0.614**	-0.097
24a-Stiff rising for day	0.029	**0.864**	0.099
24b-Stiff at end of day	-0.022	**0.847**	-0.067
4-Happy dog post	0.185	-0.074	**0.857**
14a-Play Voluntarily	0.156	-0.068	**0.859**
Factor loadings >0.6 in bold.			

Table 7 – Principal component analysis on the mobility group of questions.

Variables	Factor loadings for Components	
	E	**F**
16-Often get Exercise	0.017	**0.971**
25-Indic lame (walk)	**0.886**	-0.128
26-Indic lame (trot)	**0.959**	0.045
27-Indic Lame (run)	**0.856**	0.110
28-Pain turning (walk)	**0.713**	0.408

Factor loadings >0.6 in bold.

Step 3

A model was created by taking 1 question from each component that had a high loading factor. It was thought that since the variables in a component were highly correlated that using 1 with a high loading factor may account for a large amount of variability among questions in the same component. Spearman correlations between questions were also calculated and were taken into account when choosing the representative question for each component (See *Appendix B*). Some comparisons were also made in a trial and error method of substituting questions from the same component for each other. Thus, the model was picked based on loading factors, Spearman correlations, and some trial and error. The result was a model with 6 independent variables. The dependent variables used were those that appeared to be the best when using the full model: "Z Peak Diff," "Z Impulse Diff," and "Prop Peak Diff." See Table 8 below for the results of this modeling scheme.

Table 8 – Multivariate model results in which the independent questions were based on the principal component analysis of groups of questions, Spearman correlations, and trial and error.

Independent Variable	Component Represented	Dependent Variable: Z Peak Diff[g] β	Z Peak Diff[g] p	Z Impulse Diff[h] β	Z Impulse Diff[h] p	Prop Peak Diff[i] β	Prop Peak Diff[i] p
(Constant)		6.055	0.372	1.488	0.004*	0.87	0.000*
1b-Overall (week)	A[a]	-2.885	0.001*	-0.136	0.022*	0.009	0.707
2a-Mood (month)	B[b]	1.959	0.019*	0.054	0.358	0	0.997
24b-Stiff at end of day	C[c]	-0.453	0.331	-0.038	0.258	0.011	0.45
14a-Play Voluntarily	D[d]	0.506	0.204	0.023	0.425	-0.003	0.779
28-Pain turning (walk)	E[e]	1.026	0.048*	0.098	0.011*	-0.047	0.006*
16-Often get Exercise	F[f]	0.61	0.149	0.003	0.933	-0.024	0.069

[a]A is the only component extracted from the "Owner" group of questions
[b]B is 1 of 3 components extracted from the "Behavior" group of questions.
[c]C is 1 of 3 components extracted from the "Behavior" group of questions.
[d]D is 1 of 3 components extracted from the "Behavior" group of questions.
[e]E is 1 of 2 components extracted from the "Mobility" group of questions.
[f]F is 1 of 2 components extracted from the "Mobility" group of questions.
[g]For this model: R^2 = 0.407, adjusted R^2 = 0.308, and the p-value = 0.003.
[h]For this model: R^2 = 0.375, adjusted R^2 = 0.271, p-value = 0.007, and Dep. Var. = Sqrt (y).
[i]For this model: R^2 = 0.249, adjusted R^2 = 0.123, p-value = 0.094, and Dep. Var. = 1/Sqrt(y).
[*]Coefficient was statistically significant at α=0.05.

A comparison of the different models resulted in many coefficients that were not statistically significant. Likewise, the model fit based on R^2 and adjusted R^2 was not very high for any of the models. However, the p-values for the regression models seem to indicate that that at least 1 coefficient in each model was not equal to 0, and that the models did have some predictive ability based on the necessity of 1 or more of these questions in the model.

A comparison of these 3 models presented in Table 8 to their respective full models of all 19 of the independent variables was done. The comparison test was an F-test of whether or not the coefficients of all the independent variables not included in the reduced model could equal 0 and the reduced model still be as good as the full model. The comparison of reduced models to their full counterparts resulted in the 3 models based on principal components and 6 independent variables to be as good as the original full model with all 19 variables ($p>0.05$ for both "Z Peak Diff" and "Prop Peak Diff," $p>0.1$ for "Z Impulse Diff"). This was based on a null hypothesis that the coefficients not included equal 0. However, as could be seen in Table 8, the R^2 values for these models were relatively low. Also, the regression model using "Prop Peak Diff" was not significant, that is, there is not enough information to argue that the coefficients included in this model were not all equal to 0, which was also evident in Table 8.

Step 4

Other variable reduction procedures were employed from the statistics software (forward, backward, and stepwise elimination) to see what results occur based on significance testing of the coefficients. There was some representation of the principal components we wanted to include. Backward elimination was considered to be the most appropriate in this matter, since the goal was to start with a full model and reduce it to a smaller model. The results varied depending on the dependent variable used ("Z Peak Diff," "Z Impulse Diff," or "Prop Peak Diff"). The "Adjusted R-Square" value was used to determine if the additional variables were contributing to the model. The best models from the backward elimination were based on the highest value of the adjusted R^2. The use of the adjusted R^2 value is thought to be more important than the significance of the coefficients due to assumptions inherent in the significance testing on the coefficients and the issue of multicollinearity.[54] The results of using "Z Peak Diff" as the dependent variable resulted in almost all of the principal components being represented (5 of 6), and some components are represented more than others. See Table 9 below for the results of the backward elimination using "Z Peak Diff."

Table 9 – Results of backward elimination of variables and the remaining variables using the absolute difference in peak vertical force as the dependent variable.

Predictor Variables	**Question Group (Component)**
1a-Overall (month)	Owner (A)
2a-Mood (month)	Behavior (B)
2b-Mood (week)	Behavior (B)
3a-Attitude (month)	Behavior (B)
12-Chg Amt activities	Behavior (C)
4-Happy dog	Behavior (D)
25-Indic Lame (walk)	Mobility (E)
27-Indic Lame (run)	Mobility (E)
The fit for this model using "Z Peak Diff" as the dependent variable was: R^2= 0.699, adjusted R^2=0.612, and p<0.001.	

Using the backward elimination procedure for the use of "Z Impulse Diff" as the dependent variable resulted in a different set of predictor variables. Again, the model obtained was based on the best "Adjusted R-Square" value. Also, almost every principal component was represented by the predictor variables (5 of 6) that remained in this model. See Table 10 below for the results of the backward elimination using "Z Impulse Diff."

Table 10 – Results of backward elimination of variables and the remaining variables using the absolute difference in the vertical impulse force as the dependent variable.

Predictor Variables	Question Group (Component)
1a-Overall (month)	Owner (A)
1b-Overall (week)	Owner (A)
2a-Mood (month)	Behavior (B)
2b-Mood (week)	Behavior (B)
3a-Attitude (month)	Behavior (B)
24b-Stiff at end of day	Behavior (C)
4-Happy dog	Behavior (D)
14a-Play Voluntarily	Behavior (D)
28-Pain turning (walk)	Mobility (E)
The fit for this model using Sqrt(Z Impulse Diff) as the dependent variable was: R^2= 0.623, adjusted R^2=0.513, and p<0.001.	

The backward elimination procedure of the full model using "Prop Peak Diff" was representative of all principal components (See Table 11 below). Again, maximizing the "Adjusted R-Square" value was used as the criterion for the best model.

Table 11 – Results of backward elimination of variables and the remaining variables using the absolute difference in the peak propulsion force as the dependent variable.

Predictor Variables	Question Group (Component)
1a-Overall (month)	Owner (A)
1b-Overall (week)	Owner (A)
20a-Diff squat urine/def	Owner (A)
2a-Mood (month)	Behavior (B)
2b-Mood (week)	Behavior (B)
3a-Attitude (month)	Behavior (B)
3b-Attitude (month)	Behavior (B)
12-Chg Amt activities	Behavior (C)
24a-Stiff rising for day	Behavior (C)
24b-Stiff at end of day	Behavior (C)
4-Happy dog	Behavior (D)
25-Indic Lame (walk)	Mobility (E)
27-Indic Lame (run)	Mobility (E)
28-Pain turning (walk)	Mobility (E)
16-Often get Exercise	Mobility (F)
The fit for this model using 1/Sqrt(Prop Peak Diff) as the dependent variable was: R^2= 0.644, adjusted R^2=0.413, and p=0.007.	

<u>Step 5</u>

Using the results from the previous sets of backward elimination procedures with *a priori* knowledge, new models were constructed that predicted the force plate data reasonably well. Based on the knowledge of principal components in each group of questions, inclusion of a minimum number of questions from each of the principal component was the goal. Thus, if there was not already a component represented in the model, it was added. We also removed some questions that were thought to be represented by other questions in the same component. Some trial and error was also used in adding questions for each component and removing extra questions from some components. This resulted in 3 models with some common predictor variables, and some differences based on which dependent variable was used. The results of these 3 models can be seen in Table 12 below.

Table 12 – Variables remaining in the model by incorporating principal components with backward elimination. Components that each variable came from are given. Summary statistics for each model are also specified.

Predictors For Z Peak Diff[a] - (Component)	**Predictors For Z Impulse Diff[b] - (Component)**	**Predictors For Prop Peak Diff[c] - (Component)**
1a-Overall (month) - (A)	1a-Overall (month) - (A)	1a-Overall (month) - (A)
2a-Mood (month) - (B)	2a-Mood (month) - (B)	2a-Mood (month) - (B)
16-Often get Exercise - (F)	16-Often get Exercise - (F)	16-Often get Exercise - (F)
3a-Attitude (month) - (B)	3a-Attitude (month) - (B)	24a-Stiff rising for day - (C)
12-Chg Amt activities - (C)	24b-Stiff at end of day (C)	24b-Stiff at end of day - (C)
4-Happy dog - (D)	14a-Play Voluntarily - (D)	4-Happy dog - (D)
25-Indic Lame (walk) - (E)	28-Pain turning (walk) - (E)	28-Pain turning (walk) - (E)

[a]For this model: R^2 = 0.631, adjusted R^2 = 0.555, and the p-value < 0.001.
[b]For this model: R^2 = 0.496, adjusted R^2 = 0.396, p-value = 0.001, and Dep. Var. = Sqrt (y).
[c]For this model: R^2 = 0.444, adjusted R^2 = 0.332, p-value = 0.003, and Dep. Var. = 1/Sqrt(y).

A formal test of these reduced models compared to their full counterparts resulted in a null hypothesis for each comparison in which the coefficients not included in the model are equal to 0. The comparison of the reduced model with "Z Peak Diff" as the dependent variable compared to the full model tells us that the reduced model is as good as the full model ($p>0.25$). The test of comparison between the reduced model using "Z Impulse Diff" as the dependent variable compared to the full model tells us that the reduced model is as good as the full model ($p>0.25$). The comparison of the reduced model using "Prop Peak Diff" as the dependent variable compared to the full model tells us that the reduced model is as good as the full model ($p>0.05$). Thus, there is not sufficient evidence to show that the coefficients not included in the models are required for their respective models.

An analysis of the coefficients for these 3 models resulted in some coefficients not appearing significant. See Tables 13 below. Some of the coefficients forced into the model were not significant, however, it was thought inclusion of these variables was essential since all principal components should be represented. Some of these principal components may be measuring an aspect of pain other than mechanical lameness.

Table 13 – Coefficients of the reduced model based on backward elimination and principal component analysis and using peak vertical, vertical impulse, and peak propulsion forces as the dependent variables.

Independent Variable	Component Represented	Dependent Variable: Z Peak Diff[g] β	p	Z Impulse Diff[h] β	p	Prop Peak Diff[i] β	P
(Constant)		5.064	0.344	1.563	0.001*	1.140	<0.001*
1a-Overall (month)	A[a]	-2.819	<0.001*	-0.117	0.021*	-0.030	0.165
2a-Mood (month)	B[b]	6.266	<0.001*	0.295	0.003*	0.011	0.640
3a-Attitude (month)	B[b]	-4.773	<0.001*	-0.266	0.002*		
12-Chg Amt activities	C[c]	0.974	0.097				
24a-Stiff rising for day	C[c]					-0.052	0.002*
24b-Stiff at end of day	C[c]			-0.035	0.258	0.032	0.033*
4-Happy dog	D[d]	0.696	0.045*			0.008	0.503
14a-Play Voluntarily	D[d]			0.024	0.356		
25-Indic Lame (walk)	E[e]	0.722	0.285				
28-Pain turning (walk)	E[e]			0.101	0.004*	-0.045	0.002*
16-Often get Exercise	F[f]	-0.012	0.971	-0.047	0.863	-0.024	0.041*

[a]A is the only component extracted from the "Owner" group of questions
[b]B is 1 of 3 components extracted from the "Behavior" group of questions.
[c]C is 1 of 3 components extracted from the "Behavior" group of questions.
[d]D is 1 of 3 components extracted from the "Behavior" group of questions.
[e]E is 1 of 2 components extracted from the "Mobility" group of questions.
[f]F is 1 of 2 components extracted from the "Mobility" group of questions.
[g]For this model: R^2 = 0.631, adjusted R^2 = 0.555, and the p-value < 0.001.
[h]For this model: R^2 = 0.496, adjusted R^2 = 0.396, p-value < 0.001, and Dep. Var. = Sqrt(y).
[i]For this model: R^2 = 0.444, adjusted R^2 = 0.332, p-value = 0.003, and Dep. Var. = 1/Sqrt(y).
[*]Coefficient was statistically significant at α=0.05.

A comparison of each of these 11 questions to the 3 dependent variables used was done using linear regression models with only one independent variable entered at a time. This allowed for the relationship between each predictor variable and dependent variable to be analyzed without other predictor variables in the model. See Table 14 for these comparisons.

Table 14 – Bivariate relationships between the remaining 11 questions and the dependent variables based on linear regression models.

		Dependent Variables		
		Z Peak Diff	Z Impulse Diff [c]	Prop Peak Diff [d]
β Sig. (2-tailed)	**1a-Overall (month)**	-1.485 0.013	-0.105 0.013	0.010 0.555
β Sig. (2-tailed)	**2a-Mood (month)**	0.323 0.688	-0.030 0.595	0.009 0.701
β Sig. (2-tailed)	**3a-Attitude (month)**	-0.850[a] 0.255	-0.088[a] 0.092	0.006[a] 0.763
β Sig. (2-tailed)	**4-Happy dog**	-0.014 0.975	-0.050 0.102	0.009 0.472
β Sig. (2-tailed)	**12-Chg Amt activities**	-0.254[b] 0.727	-0.063[b] 0.218	0.011[b] 0.590
β Sig. (2-tailed)	**14a-Play Voluntarily**	0.104 0.803	-0.006 0.837	-0.004 0.761
β Sig. (2-tailed)	**16-Often get Exercise**	-0.207[a] 0.641	-0.049[a] 0.110	-0.011[a] 0.360
β Sig. (2-tailed)	**24a-Stiff rising for day**	1.342 0.004	0.077 0.020	-0.038 0.003
β Sig. (2-tailed)	**24b-Stiff at end of day**	0.475 0.309	0.025 0.451	-0.013 0.321
β Sig. (2-tailed)	**25-Indic Lame (walk)**	1.136 0.002	0.064 0.016	-0.037 <0.001
β Sig. (2-tailed)	**28-Pain turning (walk)**	1.330 0.006	0.115 0.001	-0.037 0.005

44 pairwise comparisons were made for every combination except where indicated.
[a]Only 43 pairwise comparisons possible, [b]Only 42 pairwise comparisons possible
[c] Sqrt (y) transformation used
[d] 1/Sqrt (y) transformation used

Graphical representations of the 3 models depicted in Table 13 can be seen by plotting the dependent variable against the standardized predicted value for the dependent variables and can be seen below in Figures 6-8.

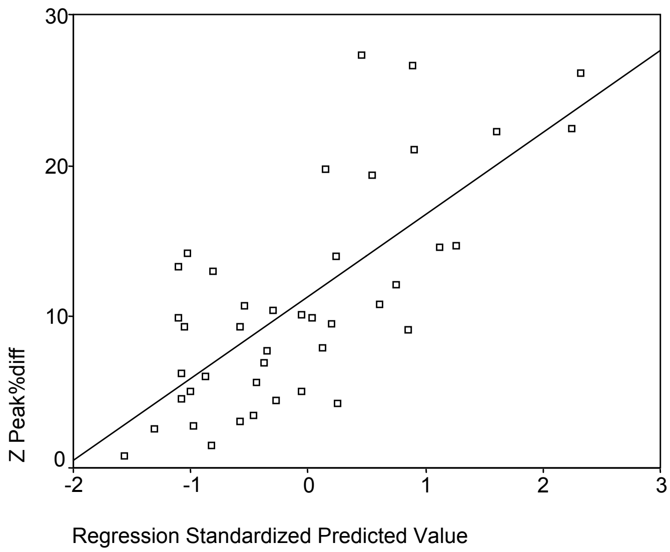

Figure 6 – Scatterplot of the absolute difference in the peak vertical force among limbs against its predicted value using the reduced regression model based on 7 questions and 44 observations.

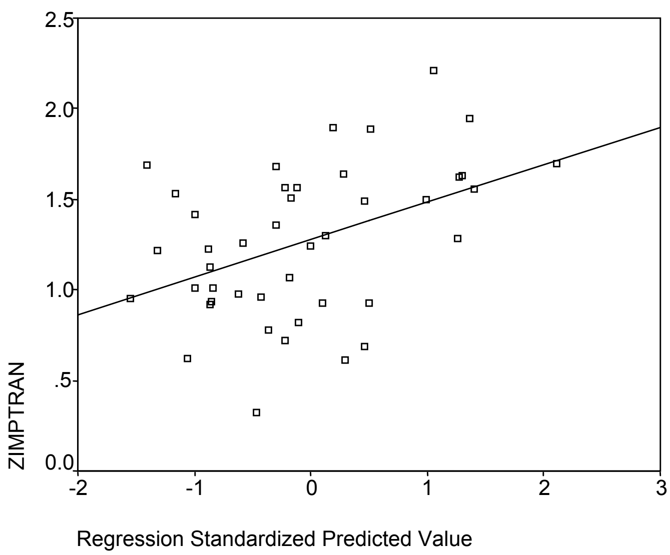

Figure 7 – Scatterplot of the absolute difference in the vertical impulse force among limbs against its predicted value using the reduced regression model based on 7 questions and 44 observations.

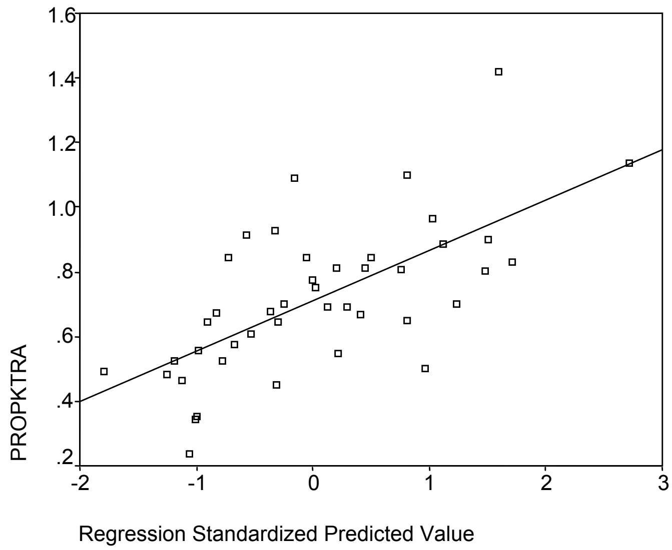

Figure 8 – Scatterplot of the absolute difference in the peak propulsion force among limbs against its predicted value using the reduced regression model based on 7 questions and 44 observations.

The graphical representations of the models further show that there is a definite relationship between the force plate data and the predictions based on the questions included. The scatter about the regression line is also indicative of the variability in the models. The models based on "Z Peak Diff" as the dependent variable seemed to have the best predictive power.

Composite Questionnaire

The composite of the 11 questions from the 3 different models based on backward elimination and principal components was then used to find an outcome that would be useful in a clinical setting. A total score for these 11 questions was tabulated by simply adding up the VAS scores for these questions. However, the responses for questions 24a, 24b, 25, and 28 had to be converted such that a maximum value of the VAS corresponded with a sound dog. Likewise, a high total VAS score represented a sound dog. The relationship between this total VAS score of the composite questionnaire and the 3 force plate measurements that were found to be the most useful in the regression analysis was then analyzed. Spearman rank correlations between the total VAS score and the force plate measurements were all significant. See Table 15 below. The correlation coefficients were also all negative which was expected based on a high VAS score being related to a sound dog, and a high % difference in the force plate measurements related to a very lame dog. Only 42 dogs were available for this comparison because only 42 of the 44 in our study completed all 11 questions.

Table 15 – Spearman rank correlations between the total VAS scores of the composite questionnaire and force plate measurements of 42 dogs.

		Z Peak % Diff	**Z Impulse % Diff**	**Propulsion Peak % Diff**
Corr. Coeff.	**Total of Composite Questionnaire**	-0.381	-0.515	-0.363
P-value (2-tailed)		0.013	<0.001	0.018

CONCLUSIONS

Quantifying the amount of pain that an animal may be experiencing is a difficult task. It has been generally accepted that whatever causes pain in humans should be thought to also cause pain in animals.[2,6] However, since animals lack the ability tell us exactly what they are feeling, we must rely on other indications of pain. The lameness aspect of pain was the focus of this study. It was hypothesized that the questionnaire used in this study would capture information about both the entire concept of pain, and in particular, the lameness aspect of pain. Although, only the lameness aspect of pain was analyzed in this study. The degree of lameness was then quantified by force plate analysis of the animal and compared to the questionnaire. The original questionnaire of 39 visual analogue scale (VAS) questions was reduced into 3 different combinations of questions that were slightly different depending on the force plate measurement used in the model. These combinations were based on those force plate measurements that seemed to be best predicted by multivariate models developed from the questionnaire data.

Determining the Final Models

In our study, we began with a lengthy questionnaire comprised of 39 VAS questions and tried to reduce that to a few questions that could accurately predict the degree of lameness based on force plate data. The questionnaire was reduced to 19 very repeatable questions (see Table 2) using a Spearman rank correlation. Regression models were then analyzed using each of the 8 possible dependent variables and all 19 repeatable questions as the independent variables. Three dependent variables, “Z Peak Diff,” “Z Impulse Diff,” and “Prop Peak Diff,” were found to be the best variables based on the R^2 value, the adjusted R^2 value, and the significance of the regression model (see Table 3).

There were 3 main groups of questions among those remaining in the questionnaire: “overall assessment by the owner,” “the dog’s attitude/behavior,” and “the dog’s mobility.” Regression models were run in which each group of questions were set as the predictors of the 3 best dependent variables (“Z Peak Diff,” “Z Impulse Diff,” and “Prop Peak Diff”). This analysis showed that the behavior assessment seemed to be the most helpful indicator in determining the degree of lameness (see Table 4). Assessment of mobility by the owner also seemed to have information that aided in determining the degree of lameness. The group of overall owner assessment questions did not appear to be quite as good at predicting the degree of lameness, based on force plate assessment, as the other 2 groups of questions. However, this did not mean that the group of questions pertaining to owner assessment was not useful. It only meant that this group of questions was not measuring mechanical lameness. The group of owner assessment questions may have been measuring another aspect of pain that was not quantified in this study.

A principal component analysis (PCA) on each of these groups of questions revealed distinct components that had questions loading onto them. This was performed to help identify the multicollinearity that was thought to exist among the independent variables. The PCA on each category revealed a total of 6 components: 1 from the "overall owner assessment" category, 3 from the "attitude/behavior" category, and 2 from the "mobility" category (See Tables 5-7).

It was thought that using some questions from each of these components that loaded highly would yield a model that would be nearly as good as the original full model. However, the models resulting from this strategy resulted in 3 models (based on "Z Peak Diff," "Z Impulse Diff," and "Prop Peak Diff" used as dependent variables) that did not seem to do nearly as good a job at prediction as the original full models (See Tables 8 and 3).

A backward elimination procedure was performed on each of the dependent variables to see what kind of results came from using an algorithm method of variable reduction. One problem with these selection methods is that it is based on analyzing the significance of a coefficient given that the other variables are included and truly independent. This is a problem when there is a known multicollinearity issue as was the case in our study. Thus, it was assumed that the values of R^2 and adjusted R^2 would more adequately predict the best model since these values should not be affected by multicollinearity.[56] Surprisingly, there was a fairly good representation of the components in the models produced by backward elimination (See Tables 9-11). The resultant models had multiple questions from different components. Interestingly, the proportion of questions coming from each of the categories related to the analysis of how well the categories predicted the different dependent variables (Table 4).

The results of the backward elimination procedures were used in conjunction with the results of the principal component analysis on the different categories of questions to try to obtain adequate models. Some of the models from the backward elimination procedures contained variables that were highly correlated with each other. If these variables were included in the model they would have coefficients that were nearly equal in magnitude, but with opposing signs, and thus canceling out the effect of the other. Thus, the inclusion of 2 such variables would be counterintuitive. This was considered when creating the final models.

The final models were obtained using the different dependent variables ("Z Peak Diff," "Z Impulse Diff," and "Prop Peak Diff") that were previously thought to be useful in developing regression models for the questionnaire. Thus, 3 different sets of independent variables (questions) were obtained. An attempt was then made to narrow the questionnaire to 1 question from each principal component. While there was some overlap in the independent variables that were obtained for each model, there were definite differences in the models.

For each dependent model, it was found that 1 question from each component did not achieve the desired model fit. Thus, for each dependent variable there are 7 independent variables, 1 from each of the 6 components, plus an additional question from another component that seemed to aid in the fit of the model (See Table 12). For the dependent variables "Z Peak Diff" and "Z Impulse Diff" there was 1 behavior component that was represented by 2 questions (2a and 3a). Likewise, for the dependent variable "Prop Peak Diff" there was 2 questions from a different behavior component

(24a and 24b). These 2 sets of questions were significantly correlated to each other, and their relationship was analyzed to try and see if there were any outliers that were having a high influence on the model and requiring both to be in the model, as was the case with questions 3a and 3b. However, this phenomenon did not seem to be occurring with these 2 sets of questions. Thus, both sets of questions were included in the final models, and it was thought that their necessity may be due to their interpretation by the owner. For example, question 2a asks the owner to assess the dog's mood in the last month, whereas question 3a asks the owner to assess the dog's attitude in the last month. These 2 descriptors are thought to be synonymous by many people; however, their definitions are slightly different which could explain a difference in interpretation by the owners. Thus, both variables were included in the models using "Z Peak Diff" and "Z Impulse Diff" as the dependent variables. The relationship between questions 24a and 24b was also highly correlated, and could be interpreted differently by an owner. Question 24a asks the owner to quantify the stiffness displayed when rising for the day, whereas question 24b asks the owner to quantify stiffness of the dog at the end of the day. Question 24b could be incorporating the result of the dog's exertion throughout the day while question 24a is measuring the amount of stiffness after a night of rest and relaxation. Thus, both of these questions were also included in the model where "Prop Peak Diff" was the dependent variable. The additional representation of the behavior category in the final models is also consistent with the idea presented in Table 4 that the category of behavior questions does a fairly good job of depicting the force plate data.

There were some coefficients in the 3 final models that did not appear to be significant and had a negligible effect (See Table 13). However, they were forced into the model because I believed that each of the principal components should be represented in the final models. The lack of significance of these coefficients could be due to a variety of reasons. The lack of significance could be due to the effects of multicollinearity. The power of the questionnaire to detect significant coefficients may have also been limited by sample size. The questionnaire could also be measuring another aspect of pain that the model based on force plate measurements could not capture. Thus, the models obtained based entirely on force plate data may be slightly different than the models obtained using a dependent variable that measured all aspects of pain. Since the force plate was measuring lameness which is a subset of pain, it is conceivable that the model that predicts all aspects of pain would contain more questions which we may have inadvertently removed.

Interpretation of Final Models

While there were some commonalities between the final models, the differences may be due to the differences inherent in the dependent variables used and their interpretation. For example, the model using “Z Peak Diff” as the dependent variable is measuring the amount of change in lameness based on the maximum and minimum amount of force exerted on a limb in the Z direction (see Figure 1). The model using “Z Impulse Diff” as the dependent variable is assessing the changes in force over time in the Z direction, and thus is measuring more of an average force exerted by the limb. Thus, depending on how the dog redistributes forces when lame, and the stage of a condition (acute or chronic), the forces exerted on the limb may differ dramatically. Thus, by measuring multiple forces, some of these changes should be recorded. However, not all changes were recorded by a single force, and by recording multiple forces a better assessment of dog's lameness was recorded.

The questions from these 3 different models were found to be repeatable based on the test-retest assessment. The resultant models were also thought to be a measure of the concurrent criterion validity of the questionnaire. Based on the R^2 values, adjusted R^2 values, scatterplots, significance of the regression models, and significance of the coefficients, the models were thought to be valid.

Usefulness of the Final Models

Three different models were obtained which were thought to give an acceptable measure of the criterion validity of the questionnaire. While 3 different models were kept that adequately predicted force plate data, the application of these models to a clinical setting is still needed. Since we want a description of the total lameness of the dog for the clinician, and this can only be seen by identifying all forces, it is conceivable to create a composite questionnaire of all 3 models. This would result in a group of 11 visual analogue scale (VAS) questions that could hopefully aid the clinician in diagnosing the degree of lameness. See Table 14 for a list of these questions. There are now multiple questions from the same principal components, which work against each other in a regression model. However, the use of this composite model to accurately predict force plate values in a clinical setting is probably not possible using regression models. Multicollinearity among independent variables causes the variability of those coefficients to increase. This variability among the estimates causes more variability in the predictive ability of the model. Thus, an accurate estimate of the lameness may not be possible based on regression models.

However, the summation of the VAS scores in the composite questionnaire could be a practical method for the clinician to use this questionnaire in a clinical setting. As seen in Table 15, there were significant correlations between the 11 questions from the composite questionnaire and the 3 forces used in the modeling procedures. Thus, the composite questionnaire should aid the clinician in diagnosing lameness that is thought to be occurring. Ideally, the total score from these 11 questions would then be used to categorize an animal as either having mild, moderate, or severe lameness. However, when cases from our study were placed into discrete categories, there was a lack of correlation between the total VAS score and the force plate assessment. This was probably due to a small sample size in each of these categories. The dogs used in this study were all chosen based on a preliminary diagnosis of moderate lameness (see *Materials and Methods*). Further studies with more dogs across a broad range of lameness would be useful in showing that the composite questionnaire is sensitive to detecting different degrees of lameness. Nonetheless, the total VAS score was significantly correlated to the force plate assessment when all 44 dogs were considered. See *Appendix C* for a sample questionnaire based on these 11 questions. Note that there are actually 12 questions because 1 of the 11 questions refers to a listing of activities that the owner provided in the original questionnaire. This listing of activities was included in the shortened questionnaire because it may also aid the clinician, although, it was not used in calculating the total VAS score. The endpoints for questions 24a, 24b, 25, and 28 were also reversed in this questionnaire so that a higher score on all VAS questions were consistent with a sound dog.

Study Limitations

Since no single force can capture all aspects of lameness, it should be clear that when developing regression models, there will never be a perfect model in which the lameness can be completely quantified by a single force. The force plate analysis was used as the gold-standard to detect lameness because it could quantify the lameness more objectively than previous lameness scores and assessments.

One problem that may exist in this study is the difference between what the questionnaire is measuring and what the force plate is measuring. The questionnaire was designed to measure both pain and lameness, whereas the force plate only measures mechanical lameness of the dog. Thus, to say that a question is not providing useful information based on force plate analysis could be a misnomer. The question may simply be addressing another aspect of pain that is not evident in the force plate analysis. This discrepancy could account for some of the variability between the questionnaire and force plate measurements.

Another source of bias in this study could have occurred in the test-retest portion of the study. Questions 1a, 2a, and 3a asked the owner questions about occurrences in the last month. Since the second assessment was done at 2 weeks, there was a unique possibility that something had occurred 3-4 weeks prior to the first assessment which would not be captured by the second assessment. Thus, there could be a discrepancy between the responses to 1 of these "month" questions that would not be captured by the force plate data between the 2 assessments. This could have created the opportunity for these questions to appear as non-repeatable. However, these questions did turn out to have significant correlations in the test-retest assessment and were kept for the model building. However, the strength of their correlations may have been affected by this phenomenon. The use of the "month" questions may have a more important clinical use if there is an expected delay between the time of making an appointment and the physical exam by the veterinarian.

The use of force plate measurements to quantify mechanical lameness has been proven to be both valid and repeatable, and was used as a gold standard in our study. However, the calculation used to determine the amount of lameness depicted by the force plate (see Table 1) could have also been a source of error in this study. This technique was not a method found in the literature. However, it was derived based on the knowledge of redistribution of forces. There are inherent sources of error in this calculation. For example, a dog with a similar level of bilateral hindlimb or forelimb lameness may not show a marked amount of lameness using this calculation. In spite of this, a method was needed to create a global outcome variable from the force plate data, and this seemed to be the best way possible at the time of this study.

The sample size used in this experiment may have reduced the power of this study. Ideally, the principal component analysis would have been conducted on the entire questionnaire to show construct validity between those categories of questions that were predetermined and the principal components extracted. However, this was not possible due to sample size constraints. Also, the full regression models used in this study were comprised of 19 variables taken over 44 individual cases. This was adequate, but a larger sample size would have improved the power of the study. The sample size may also have affected the repeatability analysis of the questionnaire. Some questions were deemed not repeatable because there were not enough owners responding to a particular question. There were questions that were not relevant to all dogs (i.e. lifting a leg to urinate (question 20b) is predominately a male behavior). There were also questions referring to stairs (questions 21a, 21b, 22a, and 22b) that were not applicable to all owners. Thus, a larger sample size in the test-retest assessment of repeatability may have allowed for more questions to be considered repeatable. With more independent variables included in the model, we would also need a larger sample size to set up regression models containing all independent variables.

Recommendations for Future Studies

The use of force plate data to quantitate lameness has had an increasing acceptance in both human studies and animal studies as a gold standard to detect mechanical lameness. With the constant improvement of technology, it is necessary for us to update the means by which studies are validated. Methods of proving construct validity (showing that the theoretical constructs match those that were actually obtained) may have been used previously due to a lack of methods to assess criterion validity (using measurements from an instrument known to be valid). However, as methods for proving criterion validity become possible, they should be employed and used in conjunction with proving construct validity. To my knowledge, this is the first studies done in animals in which a questionnaire on lameness was validated using force plate analysis. Hopefully, others will follow and add to the body of knowledge until a better diagnostic tool is discovered.

Further validation of these 11 questions from the 3 final models would be ideal. Also, some questions that were removed in the repeatability analysis could be included in another repeatability assessment to see if they become repeatable when more animals are included in the study. A larger cohort of animals would definitely be advantageous. This would allow a better estimate of the coefficients in the model to be obtained and used to for a better prediction of the force plate. The correlation between these predicted values of the force plate and the actual measurements would then be improved. If a larger cohort of animals were used, a principal component analysis could also be done on the entire set of questions to show construct validity.

REFERENCES

1. Mathews KA. Pain assessment and general approach to management. *Vet Clin N Am Sm Anim Pract* 2000;30:729-755.

2. Association of Veterinary Teachers and Research Workers. Guidelines for the recognition and assessment of pain in animals. *Vet Rec* 1986;118:334-338.

3. Johnston SA. Overview of pain in the lame patient. *Vet Clin N Am Sm Anim Pract* 2001;31:39-53.

4. Renberg WC. Evaluation of the lame patient. *Vet Clin N Am Sm Anim Pract* 2001;31:1-16.

5. Wiseman ML, Nolan AM, Reid J, Scott EM. Preliminary study on owner-reported behaviour changes associated with chronic pain in dogs. *Vet Rec* 2001;149:423-424.

6. Morton DB, Griffiths PHM. Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. *Vet Rec* 1985;116:431-436.

7. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1:277-299.

8. Holton L, Reid J, Scott EM, Pawson P, Nolan A. Development of a behaviour-based scale to measure acute pain in dogs. *Vet Rec* 2001;148:525-531.

9. Poffenbarger EM, McCurnin DM. Physical diagnosis and history taking. In: McCurnin DM, Poffenbarger EM, eds. *Small Animal Physical Diagnosis*. Philadelphia: W.B. Saunders Co.; 1991:92.

10. Mathews KA. Nonsteroidal anti-inflammatory analgesics in pain management in dogs and cats. *Can Vet J* 1996;37:539-545.

11. Firth AM, Haldane SL. Development of a scale to evaluate postoperative pain in dogs. *J Am Vet Med Assoc* 1999;214:651-659.

12. Conzemius MG, Hill CM, Sammarco JL, Perkowski SZ. Correlation between subjective and objective measures used to determine severity of postoperative pain in dogs. *J Am Vet Med Assoc* 1997;210:1619-1622.

13. Holton LL, Scott EM, Nolan AM, Reid J, Welsh E, Flaherty D. Comparison of three methods used for assessment of pain in dogs. *J Am Vet Med Assoc* 1998;212:61-66.

14. Cambridge AJ, Tobias KM, Newberry RC, Sarkar DK. Subjective and objective measurements of postoperative pain in cats. *J Am Vet Med Assoc* 2000;217:685-690.

15. Welsh EM, Gettinby G, Nolan AM. Comparison of a visual analogue scale and a numerical rating scale for assessment of lameness, using sheep as a model. *Am J Vet Res* 1993;54(6):976-983.

16. Holton LL, Scott EM, Nolan AM, Reid J, Welsh E. Relationship between physiological factors and clinical pain in dogs scored using a numerical rating scale. *J Small Anim Pract* 1998;39:469-474.

17. Budsberg SC, Verstraete MC, Soutas-Little RW. Force plate analysis of the walking gait in healthy dogs. *Am J Vet Res* 1987;48:915-918.

18. Rumph PF, Lander JE, Kincaid SA, Baird DK, Kammermann JR, Visco DM. Ground reaction force profiles from force platform gait analyses of clinically normal mesomorphic dogs at the trot. *Am J Vet Res* 1994;55:756-761.

19. McLaughlin Jr. R, Roush JK. Effects of increasing velocity on braking and propulsion times during force plate gait analysis in Greyhounds. *Am J Vet Res* 1995;56(2):159-161.

20. Rumph PF, Kincaid SA, Baird DK, Kammermann JR, Visco DM, Goetze LF. Vertical ground reaction force distribution during experimentally induced acute synovitis in dogs. *Am J Vet Res* 1993;54:365-369.

21. Griffon DJ, McLaughlin Jr RM, Roush JK. Vertical ground reaction force redistribution during experimentally induced shoulder lameness in dogs. *VCOT* 1994;7:154-157.

22. Rumph PF, Kincaid SA, Visco DM, Baird DK, Kammermann JR, West MS. Redistribution of vertical ground reaction force in dogs with experimentally induced chronic hindlimb lameness. *Vet Surg* 1995;24:384-389.

23. O'Connor BL, Visco DM, Heck DC, Myers SL, Brandt KD. Gait alterations in dogs after transection of the anterior cruciate ligament. *Arthritis Rheum* 1989;32:1142-1147.

24. Dupuis J, Harari J, Papageorges M, Gallina A, Ratzlaff M. Evaluation of fibular head transposition for repair of experimental cranial cruciate ligament in dogs. *Vet Surg* 1994;23:1-12.

25. Budsberg SC, Verstraete MC, Soutas-Little RW, Flo GL, Probst CW. Force plate analyses before and after stabilization of canine stifles for cruciate injury. *Am J Vet Res* 1988;49:1522-1524.

26. Jevens DJ, DeCamp CE, Hauptman J, Braden TD, Richter M, Robinson R. Use of force-plate analysis of gait to compare two surgical techniques for treatment of cranial cruciate ligament rupture in dogs. *Am J Vet Res* 1996;57(3):389-393.

27. McLaughlin Jr RM, Miller CW, Taves CL, Hearne TC, Palmer NC, Anderson GI. Force plate analysis of triple pelvic osteotomy for the treatment of canine hip dysplasia. *Vet Surg* 1991;20:291-297.

28. Budsberg SC, Chambers JN, Van Lue SL, Foutz TL, Reece L. Prospective evaluation of ground reaction forces in dogs undergoing unilateral total hip replacement. *Am J Vet Res* 1996;57:1781-1785.

29. McLaughlin Jr. R, Roush JK. A comparison of two surgical approaches to the scapulohumeral joint in dogs. *Vet Surg* 1995;24:207-214.

30. Vasseur PB, Johnson AL, Budsberg SC, Lincoln JD, Toombs JP, Whitehair JG, Lentz EL. Randomized, controlled trial of the efficacy of carprofen, a

nonsteroidal anti-inflammatory drug, in the treatment of osteoarthritis in dogs. *J Am Vet Med Assoc* 1995;206 (6):807-811.

31. Jevens DJ, Hauptman JG, DeCamp CE, Budsberg SC, Soutas-Little RW. Contributions to variance in force-plate analysis of gait in dogs. *Am J Vet Res* 1993;54:612-615.

32. Rumph PF, Steiss JE, Montgomery RD. Effects of selection and habituation on vertical ground reaction force in Greyhounds. *Am J Vet Res* 1997;58:1206-1208.

33. Rumph PF, Steiss JE, West MS. Interday variation in vertical ground reaction force in clinically normal Greyhounds at the trot. *Am J Vet Res* 1999;60:679-683.

34. Riggs CM, DeCamp CE, Soutas-Little RW, Braden TD, Richter MA. Effects of subject velocity on force plate measured ground reaction forces in healthy greyhounds at the trot. *Am J Vet Res* 1993;54:1523-1526.

35. McLaughlin RM, Roush JK. Effects of subject stance time and velocity on ground reaction forces in clinically normal greyhounds at the trot. *Am J Vet Res* 1994;55:1666-1671.

36. Roush JK, McLaughlin RM. Effects of subject stance time and velocity on ground reaction forces in clinically normal greyhounds at the walk. *Am J Vet Res* 1994;55:1672-1676.

37. Renberg WC, Johnston SA, Ye K, Budsberg SC. Comparison of stance time and velocity as control variables in force plate analysis of dogs. *Am J Vet Res* 1999;60:814-819.

38. Last JM. *A Dictionary of Epidemiology*. Oxford: University Press; 2001.

39. Carmines EG, Zeller RA. *Reliability and Validity Assessment*. Beverly Hills, CA: Sage Publications; 1979.

40. Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). *Pain* 1985;23:345-356.

41. Andre M, Hedengren E, Hagelberg S, Stenstrom C. Perceived ability to manage juvenile chronic arthritis among adolescents and parents: development of a questionnaire to assess medical issues, exercise, pain, and social support. *Arthritis Care Res* 1999;12:229-237.

42. Wright FV, Kimber JL, Law M, Goldsmith CH, Crombie V, Dent P. The Juvenile Arthritis Functional Status Index (JASI): a validation study. *J Rheumatol* 1996;23:1066-1079.

43. Gragg RA, Rapoff MA, Danovsky MB, Lindsley CB, Varni JW, Waldron SA, Bernstein BH. Assessing chronic musculoskeletal pain associated with rheumatic disease: further validation of the pediatric pain questionnaire. *J Pediatr Psychol* 1996;21:237-250.

44. Idvall E, Hamrin E, Unosson M. Development of an instrument to measure strategic and clinical quality indicators in postoperative pain management. *J Adv Nurs* 2002;37:532-540.

45. Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire: an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *J Rheumatol* 1987;14:446-451.

46. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip of knee. *J Rheumatol* 1988;15:1833-1840.

47. Foster RL, Varni JW. Measuring the quality of children's postoperative pain management: initial validation of the child/parent Total Quality Pain Management (TPQM) instrument. *J Pain Symptom Manage* 2002;23:201-210.

48. Soetenga D, Frank J, Pellino TA. Assessment of the validity and reliability of the University of Wisconsin Children's Hospital Pain Scale for Preverbal and Nonverbal Children. *Pediatr Nurs* 1999;25:670-676.

49. Hooper DM, Morrissey MC, Drechlser WI, McDermott M, McAuliffe TB. Validation of the Hughston Clinic Subjective Knee Questionnaire using gait analysis. *Med Sci Sports Exerc* 2001;33:1456-1462.

50. Novacheck TF, Stout JL, Tervo R. Reliability and validity of the Gillette Functional Assessment Questionnaire as an outcome measure in children with walking disabilities. *J Pediatr Orthop* 2000;20:75-81.

51. Slater MR, Taylor L, Carroll GL, Hulse DA, Boothe DM. The reliability of a questionnaire to assess canine lameness and pain. *Conference of Research Workers in Animal Diseases* 78. Chicago: 1997.

52. Greenland S. Modeling and variable selection in epidemiologic analysis. *Am J Public Health* 1989;79:340-349.

53. Jolliffe IT. *Principal Component Analysis*. New York: Springer; 2002.

54. Ott RL, Longnecker M. More on multiple regression. In: *An Introduction to Statistical Methods and Data Analysis*. Pacific Grove, CA: Duxbury; 2001:705-826.

55. SPSS for Windows. Version 11.0.1. Chicago: SPSS Inc, 2001.

56. Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. Multiple regression II. In: *Applied Linear Statistical Models*. Boston: McGraw-Hill Companies, Inc.; 1996:260-326.

APPENDIX A

Canine Movement Assessment

Case Number _____________________ Owner's Name ________________

Clinician _________________________ Dog's Name __________________

Referring Veterinarian ______________ Evaluator _____________________

Please **Read Instructions First:**

- Reply to the questions by **<u>placing a vertical mark</u>** on the corresponding line. This vertical mark corresponds to a place between the two extremes. The distance between your mark and the left end will be measured to quantify your response.

- Please **notice the labeling** on the left and right sides before marking it.

- When assessing your dog over the past week (or month), mark down his/her **<u>usual condition.</u>**

- **Thank you!**

1a. How would you describe your overall assessment of your dog <u>in the last month</u>?

|____|
poor excellent

b. How would you describe your overall assessment of your dog <u>in the last week</u>?

|____|
poor excellent

2a. What kind of mood has your dog been <u>in the last month</u>?

|____|
bad good

b. What kind of mood has your dog been <u>in the last week</u>?

|____|
bad good

3a. How has your dog's attitude been in the last month?

|____|
negative positive

b. How has your dog's attitude been in the last week?

|____|
negative positive

4. How frequently does your dog display comfort or "happy dog" postures (for example, lying on back with toy in mouth? Not applicable: □

|____|
rarely frequently

5. How frequently do you get frustrated with your dog? Not applicable: □

|____|
less more

6. In the last month, how frequently have you offered special treats to your dog? Not applicable: □

|____|
less more

7. In the last month, how frequently have you massaged your dog? Not applicable: □

|____|
less more

8. In the last month, has the amount of physical interaction your dog seeks with family members changed? Not applicable: □

|____|
less more

9. In the last month, has the type of physical interaction your dog seeks with family members changed? Not applicable: □

|____|
no change mostly different activities

10. What type of daily activities does your dog engage in (ie., fetching newspapers, playing frisbee)?

____

____

11. Has your dog changed the frequency of these daily activities?

|____|
less more

12. Has your dog changed the amount of these activities?

|____|
less more

13. How frequently do you try to engage your dog in other activities?

|____|
rarely very often

14a. How willing is your dog to play voluntarily?

|____|
not at all very willingly

b. Has your dog's willingness to play voluntarily changed in the last month?

|____|
could not be less willing could not be more eager

15a. How much voluntary exercise does your dog get? Not applicable: □

|____|
none frequently

b. In the last month, has your dog changed his/her amount of voluntary exercise? Not applicable: □

|____|

could not decrease more — could not increase more

16. How often does your dog get exercise?

|____|

less than once — all day per day

17a. How difficult is it for your dog to get into the car/truck or jump on the furniture? Not applicable: □

|____|

must be picked up — easily by self

b. In the last month, has your dog changed the amount of assistance required to get into the car/truck or jump on the furniture? Not applicable: □

|____|

could not have gotten worse — could not have improved more

18a. How difficult is it for your dog to lie down?

|____|

difficult — easy

b. How difficult is it for your dog to sit down on his haunches?

|____|

difficult — easy

19a. How difficult is it for your dog to rise from a lying position?

|____|

difficult — easy

b. How difficult is it for your dog to rise from a sitting position?

|____|
difficult easy

20a. How difficult is it for your dog to squat to urinate/defecate?

|____|
difficult easy

b. How difficult is it for your dog to lift his leg to urinate? Not applicable: □

|____|
difficult easy

21a. How willing does your dog walk up stairs? Not applicable: □

|____|
unwilling eagerly

b. How willing did your dog used to walk up stairs? Not applicable: □

|____|
unwilling eagerly

22a. How willing does your dog walk down stairs? Not applicable: □

|____|
unwilling eagerly

b. How willing did your dog used to walk down stairs? Not applicable: □

|____|
unwilling eagerly

23. Dose your dog vocally indicate pain when touched?

|____|
rarely often

24a. How stiff is your dog when arising for the day?

|____|

not stiff — could not be more stiff

b. How stiff is your dog at the end of the day (post-activities)?

|____|

not stiff — could not be more stiff

25. Does your dog indicate any lameness at a walk?

|____|

rarely — always

26. Does your dog indicate any lameness at a trot?

|____|

rarely — always

27. Does your dog indicate any lameness at a run?

|____|

rarely — always

28. Does your dog indicate any pain when turning suddenly at a walk?

|____|

rarely — always

29. Now we would like you to think again about each of the aforementioned symptoms which you have just rated. The select three items which are most important to you, i.e., which you most hope the treatment your dog is about to receive will **improve**. Please indicate your selections by circling the appropriate item numbers.

Thank you for completing this questionnaire.

APPENDIX B

Spearman Rank Correlations

		1a-Overall (month)	1b-Overall (week)	2a-Mood (month)	2b-Mood (week)
Corr. Coeff.	1a-Overall (month)	1.00	0.96	0.48	0.54
Sig. (2-tailed)		.	0.00	0.00	0.00
Corr. Coeff.	1b-Overall (week)	0.96	1.00	0.50	0.60
Sig. (2-tailed)		0.00	.	0.00	0.00
Corr. Coeff.	2a-Mood (month)	0.48	0.50	1.00	0.91
Sig. (2-tailed)		0.00	0.00	.	0.00
Corr. Coeff.	2b-Mood (week)	0.54	0.60	0.91	1.00
Sig. (2-tailed)		0.00	0.00	0.00	.
Corr. Coeff.	3a-Attitude (month)	0.45	0.49	0.92	0.87
Sig. (2-tailed)		0.00	0.00	0.00	0.00
Corr. Coeff.	3b-Attitude (week)	0.49	0.55	0.86	0.92
Sig. (2-tailed)		0.00	0.00	0.00	0.00
Corr. Coeff.	4-Happy dog	0.33	0.40	0.39	0.40
Sig. (2-tailed)		0.03	0.01	0.01	0.01
Corr. Coeff.	12-Chg Amt activities	0.50[b]	0.46[b]	0.02[b]	0.04[b]
Sig. (2-tailed)		0.00	0.00	0.88	0.81
Corr. Coeff.	14a-Play Voluntarily	0.40	0.45	0.46	0.40
Sig. (2-tailed)		0.01	0.00	0.00	0.01
Corr. Coeff.	16-Often get Exercise	0.28[a]	0.32[a]	-0.08[a]	-0.08[a]
Sig. (2-tailed)		0.07	0.04	0.60	0.63
Corr. Coeff.	18b-Diff sitting	0.42[a]	0.44[a]	0.15[a]	0.22[a]
Sig. (2-tailed)		0.00	0.00	0.34	0.15
Corr. Coeff.	19a-Diff rise from lying	0.54	0.58	0.15	0.23
Sig. (2-tailed)		0.00	0.00	0.34	0.14
Corr. Coeff.	20a-Diff squat urine/def	0.62	0.67	0.38	0.46
Sig. (2-tailed)		0.00	0.00	0.01	0.00
Corr. Coeff.	24a-Stiff rising for day	-0.44	-0.44	-0.15	-0.16
Sig. (2-tailed)		0.00	0.00	0.33	0.28
Corr. Coeff.	24b-Stiff at end of day	-0.18	-0.16	-0.04	-0.06
Sig. (2-tailed)		0.23	0.29	0.79	0.68
Corr. Coeff.	25-Indic Lame (walk)	-0.30	-0.32	-0.19	-0.28
Sig. (2-tailed)		0.05	0.03	0.22	0.07
Corr. Coeff.	26-Indic Lame (trot)	-0.27	-0.31	-0.20	-0.31
Sig. (2-tailed)		0.07	0.04	0.20	0.04
Corr. Coeff.	27-Indic Lame (run)	-0.25[a]	-0.31[a]	-0.21[a]	-0.31[a]
Sig. (2-tailed)		0.10	0.05	0.17	0.04
Corr. Coeff.	28-Pain turning (walk)	-0.30	-0.33	-0.08	-0.17
Sig. (2-tailed)		0.05	0.03	0.59	0.27

44 pairwise comparisons were made for every question combination except where indicated.

[a]Only 43 pairwise comparisons possible, [b]Only 42 pairwise comparisons possible, [c]Only 41 pairwise comparisons possible

Spearman Rank Correlations cont.

		3a-Attitude (month)	3b-Attitude (week)	4-Happy dog	12-Chg Amt activities
Corr. Coeff. Sig. (2-tailed)	1a-Overall (month)	0.45 0.00	0.49 0.00	0.33 0.03	0.50[b] 0.00
Corr. Coeff. Sig. (2-tailed)	1b-Overall (week)	0.49 0.00	0.55 0.00	0.40 0.01	0.45[b] 0.00
Corr. Coeff. Sig. (2-tailed)	2a-Mood (month)	0.92 0.00	0.86 0.00	0.39 0.01	0.02[b] 0.88
Corr. Coeff. Sig. (2-tailed)	2b-Mood (week)	0.87 0.00	0.92 0.00	0.40 0.01	0.04[b] 0.81
Corr. Coeff. Sig. (2-tailed)	3a-Attitude (month)	1.00 .	0.95 0.00	0.38 0.01	0.00[b] 0.98
Corr. Coeff. Sig. (2-tailed)	3b-Attitude (week)	0.95 0.00	1.00 .	0.37 0.01	-0.02[b] 0.90
Corr. Coeff. Sig. (2-tailed)	4-Happy dog	0.38 0.01	0.37 0.01	1.00 .	0.13[b] 0.43
Corr. Coeff. Sig. (2-tailed)	12-Chg Amt activities	0.00[b] 0.98	-0.02[b] 0.90	0.13[b] 0.43	1.00 .
Corr. Coeff. Sig. (2-tailed)	14a-Play Voluntarily	0.40 0.01	0.34 0.02	0.55 0.00	0.08[b] 0.62
Corr. Coeff. Sig. (2-tailed)	16-Often get Exercise	-0.12[a] 0.45	-0.11[a] 0.49	0.15[a] 0.35	0.27[b] 0.08
Corr. Coeff. Sig. (2-tailed)	18b-Diff sitting	0.13[a] 0.41	0.18[a] 0.24	0.03[a] 0.86	0.19[c] 0.24
Corr. Coeff. Sig. (2-tailed)	19a-Diff rise from lying	0.13 0.41	0.19 0.21	0.17 0.27	0.44[b] 0.00
Corr. Coeff. Sig. (2-tailed)	20a-Diff squat urine/def	0.30 0.05	0.37 0.01	0.31 0.04	0.27[b] 0.09
Corr. Coeff. Sig. (2-tailed)	24a-Stiff rising for day	-0.22 0.15	-0.23 0.13	-0.19 0.21	-0.22[b] 0.17
Corr. Coeff. Sig. (2-tailed)	24b-Stiff at end of day	-0.02 0.91	-0.06 0.71	0.08 0.61	-0.18[b] 0.27
Corr. Coeff. Sig. (2-tailed)	25-Indic Lame (walk)	-0.22 0.16	-0.28 0.06	-0.19 0.22	-0.03[b] 0.83
Corr. Coeff. Sig. (2-tailed)	26-Indic Lame (trot)	-0.21 0.18	-0.28 0.07	-0.23 0.13	0.05[b] 0.75
Corr. Coeff. Sig. (2-tailed)	27-Indic Lame (run)	-0.20[a] 0.19	-0.29[a] 0.06	-0.16[a] 0.30	0.14[c] 0.38
Corr. Coeff. Sig. (2-tailed)	28-Pain turning (walk)	-0.07 0.66	-0.14 0.35	-0.06 0.71	0.03[b] 0.85

44 pairwise comparisons were made for every question combination except where indicated.

[a]Only 43 pairwise comparisons possible, [b]Only 42 pairwise comparisons possible, [c]Only 41 pairwise comparisons possible

Spearman Rank Correlations cont.

		14a-Play Voluntarily	16-Often get Exercise	18b-Diff sitting	19a-Diff rise from lying
Corr. Coeff. Sig. (2-tailed)	1a-Overall (month)	0.40 0.01	0.28[a] 0.07	0.42[a] 0.00	0.54 0.00
Corr. Coeff. Sig. (2-tailed)	1b-Overall (week)	0.45 0.00	0.32[a] 0.04	0.44[a] 0.00	0.58 0.00
Corr. Coeff. Sig. (2-tailed)	2a-Mood (month)	0.46 0.00	-0.08[a] 0.60	0.15[a] 0.34	0.15 0.34
Corr. Coeff. Sig. (2-tailed)	2b-Mood (week)	0.40 0.01	-0.08[a] 0.63	0.22[a] 0.15	0.23 0.14
Corr. Coeff. Sig. (2-tailed)	3a-Attitude (month)	0.40 0.01	-0.12[a] 0.45	0.13[a] 0.41	0.13 0.41
Corr. Coeff. Sig. (2-tailed)	3b-Attitude (week)	0.34 0.02	-0.11[a] 0.49	0.18[a] 0.24	0.19 0.21
Corr. Coeff. Sig. (2-tailed)	4-Happy dog	0.55 0.00	0.1[a] 0.35	0.03[a] 0.86	0.17 0.27
Corr. Coeff. Sig. (2-tailed)	12-Chg Amt activities	0.08[b] 0.62	0.27[b] 0.08	0.19[c] 0.24	0.44[b] 0.00
Corr. Coeff. Sig. (2-tailed)	14a-Play Voluntarily	1.00 .	0.04[a] 0.79	0.14[a] 0.38	0.17 0.28
Corr. Coeff. Sig. (2-tailed)	16-Often get Exercise	0.04[a] 0.79	1.00[a] .	0.15[b] 0.35	0.38[a] 0.01
Corr. Coeff. Sig. (2-tailed)	18b-Diff sitting	0.14[a] 0.38	0.15[b] 0.35	1.00[a] .	0.71[a] 0.00
Corr. Coeff. Sig. (2-tailed)	19a-Diff rise from lying	0.17 0.28	0.38[a] 0.01	0.71 0.00	1.00 .
Corr. Coeff. Sig. (2-tailed)	20a-Diff squat urine/def	0.42 0.01	0.33[a] 0.03	0.63[a] 0.00	0.64 0.00
Corr. Coeff. Sig. (2-tailed)	24a-Stiff rising for day	-0.13 0.39	-0.06[a] 0.70	-0.45[a] 0.00	-0.61 0.00
Corr. Coeff. Sig. (2-tailed)	24b-Stiff at end of day	0.01 0.97	0.20[a] 0.20	-0.34[a] 0.03	-0.33 0.03
Corr. Coeff. Sig. (2-tailed)	25-Indic Lame (walk)	-0.12 0.42	0.11[a] 0.48	-0.44 0.00	-0.35 0.02
Corr. Coeff. Sig. (2-tailed)	26-Indic Lame (trot)	-0.19 0.21	0.07[a] 0.65	-0.41[a] 0.01	-0.29 0.05
Corr. Coeff. Sig. (2-tailed)	27-Indic Lame (run)	-0.07[a] 0.67	0.00[b] 1.00	-0.45[b] 0.00	-0.40[a] 0.01
Corr. Coeff. Sig. (2-tailed)	28-Pain turning (walk)	-0.09 0.56	-0.24[a] 0.12	-0.45[a] 0.00	-0.33 0.03

44 pairwise comparisons were made for every question combination except where indicated.

[a]Only 43 pairwise comparisons possible, [b]Only 42 pairwise comparisons possible, [c]Only 41 pairwise comparisons possible

Spearman Rank Correlations cont.

		20a-Diff squat urine/def	24a-Stiff rising for day	24b-Stiff at end of day	25-Indic Lame (walk)
Corr. Coeff.	1a-Overall (month)	0.62	-0.44	-0.18	-0.30
Sig. (2-tailed)		0.00	0.00	0.23	0.05
Corr. Coeff.	1b-Overall (week)	0.67	-0.44	-0.16	-0.32
Sig. (2-tailed)		0.00	0.00	0.29	0.03
Corr. Coeff.	2a-Mood (month)	0.38	-0.15	-0.04	-0.19
Sig. (2-tailed)		0.01	0.33	0.79	0.22
Corr. Coeff.	2b-Mood (week)	0.46	-0.16	-0.06	-0.28
Sig. (2-tailed)		0.00	0.28	0.68	0.07
Corr. Coeff.	3a-Attitude (month)	0.30	-0.22	-0.02	-0.22
Sig. (2-tailed)		0.05	0.15	0.91	0.16
Corr. Coeff.	3b-Attitude (week)	0.37	-0.23	-0.06	-0.28
Sig. (2-tailed)		0.01	0.13	0.71	0.06
Corr. Coeff.	4-Happy dog	0.31	-0.19	0.08	-0.19
Sig. (2-tailed)		0.04	0.21	0.61	0.22
Corr. Coeff.	12-Chg Amt activities	0.27[b]	-0.22[b]	-0.18[b]	-0.03[b]
Sig. (2-tailed)		0.09	0.17	0.27	0.83
Corr. Coeff.	14a-Play Voluntarily	0.42	-0.13	0.01	-0.12
Sig. (2-tailed)		0.01	0.39	0.97	0.42
Corr. Coeff.	16-Often get Exercise	0.33[a]	-0.06[a]	0.20[a]	0.11[a]
Sig. (2-tailed)		0.03	0.70	0.20	0.48
Corr. Coeff.	18b-Diff sitting	0.63[a]	-0.45[a]	-0.34[a]	-0.44[a]
Sig. (2-tailed)		0.00	0.00	0.03	0.00
Corr. Coeff.	19a-Diff rise from lying	0.64	-0.61	-0.33	-0.35
Sig. (2-tailed)		0.00	0.00	0.03	0.02
Corr. Coeff.	20a-Diff squat urine/def	1.00	-0.47	-0.31	-0.37
Sig. (2-tailed)		.	0.00	0.04	0.01
Corr. Coeff.	24a-Stiff rising for day	-0.47	1.00	0.48	0.60
Sig. (2-tailed)		0.00	.	0.00	0.00
Corr. Coeff.	24b-Stiff at end of day	-0.31	0.48	1.00	0.37
Sig. (2-tailed)		0.04	0.00	.	0.01
Corr. Coeff.	25-Indic Lame (walk)	-0.37	0.60	0.37	1.00
Sig. (2-tailed)		0.01	0.00	0.01	.
Corr. Coeff.	26-Indic Lame (trot)	-0.37	0.44	0.31	0.90
Sig. (2-tailed)		0.01	0.00	0.04	0.00
Corr. Coeff.	27-Indic Lame (run)	-0.37[a]	0.37[a]	0.32[a]	0.73[a]
Sig. (2-tailed)		0.02	0.01	0.03	0.00
Corr. Coeff.	28-Pain turning (walk)	-0.48	0.37	0.45	0.65
Sig. (2-tailed)		0.00	0.01	0.00	0.00

44 pairwise comparisons were made for every question combination except where indicated. [a]Only 43 pairwise comparisons possible, [b]Only 42 pairwise comparisons possible, [c]Only 41 pairwise comparisons possible

Spearman Rank Correlations cont.

		26-Indic Lame (trot)	27-Indic Lame (run)	28-Pain turning (walk)
Corr. Coeff. Sig. (2-tailed)	1a-Overall (month)	-0.27 0.07	-0.25[a] 0.10	-0.30 0.05
Corr. Coeff. Sig. (2-tailed)	1b-Overall (week)	-0.31 0.04	-0.31[a] 0.05	-0.33 0.03
Corr. Coeff. Sig. (2-tailed)	2a-Mood (month)	-0.20 0.20	-0.21[a] 0.17	-0.08 0.59
Corr. Coeff. Sig. (2-tailed)	2b-Mood (week)	-0.31 0.04	-0.31[a] 0.04	-0.17 0.27
Corr. Coeff. Sig. (2-tailed)	3a-Attitude (month)	-0.21 0.18	-0.20 0.19	-0.07 0.66
Corr. Coeff. Sig. (2-tailed)	3b-Attitude (week)	-0.28 0.07	-0.29[a] 0.06	-0.14 0.35
Corr. Coeff. Sig. (2-tailed)	4-Happy dog	-0.23 0.13	-0.16[a] 0.30	-0.06 0.71
Corr. Coeff. Sig. (2-tailed)	12-Chg Amt activities	0.05[b] 0.75	0.14[c] 0.38	0.03[b] 0.85
Corr. Coeff. Sig. (2-tailed)	14a-Play Voluntarily	-0.19 0.21	-0.07[a] 0.67	-0.09 0.56
Corr. Coeff. Sig. (2-tailed)	16-Often get Exercise	0.07[a] 0.65	0.00[b] 1.00	-0.24[a] 0.12
Corr. Coeff. Sig. (2-tailed)	18b-Diff sitting	-0.41[a] 0.01	-0.45[b] 0.00	-0.45[a] 0.00
Corr. Coeff. Sig. (2-tailed)	19a-Diff rise from lying	-0.29 0.05	-0.40[a] 0.01	-0.33 0.03
Corr. Coeff. Sig. (2-tailed)	20a-Diff squat urine/def	-0.37 0.01	-0.37[a] 0.02	-0.48 0.00
Corr. Coeff. Sig. (2-tailed)	24a-Stiff rising for day	0.44 0.00	0.37[a] 0.01	0.37 0.01
Corr. Coeff. Sig. (2-tailed)	24b-Stiff at end of day	0.31 0.04	0.32[a] 0.03	0.45 0.00
Corr. Coeff. Sig. (2-tailed)	25-Indic Lame (walk)	0.90 0.00	0.73[a] 0.00	0.65 0.00
Corr. Coeff. Sig. (2-tailed)	26-Indic Lame (trot)	1.00 .	0.82[a] 0.00	0.70 0.00
Corr. Coeff. Sig. (2-tailed)	27-Indic Lame (run)	0.82[a] 0.00	1.00[a] .	0.55[a] 0.00
Corr. Coeff. Sig. (2-tailed)	28-Pain turning (walk)	0.70 0.00	0.55[a] 0.00	1.00 .

44 pairwise comparisons were made for every question combination except where indicated.

[a]Only 43 pairwise comparisons possible, [b]Only 42 pairwise comparisons possible, [c]Only 41 pairwise comparisons possible

APPENDIX C

Canine Movement Assessment

Case Number _____________________ Owner's Name ________________

Clinician _________________________ Dog's Name __________________

Referring Veterinarian ______________ Evaluator _____________________

Please **Read Instructions First:**

- Reply to the questions by **placing a vertical mark** on the corresponding line. This vertical mark corresponds to a place between the two extremes. The distance between your mark and the left end will be measured to quantify your response.

- Please **notice the labeling** on the left and right sides before marking it.

- When assessing your dog over the past week (or month), mark down his/her **usual condition.**

- **Thank you!**

1. How would you describe your overall assessment of your dog in the last month?

|____|
poor excellent

2. What kind of mood has your dog been in the last month?

|____|
bad good

3. How has your dog's attitude been in the last month?

|____|
negative positive

4. How frequently does your dog display comfort or "happy dog" postures (for example, lying on back with toy in mouth? Not applicable: □

|____|
rarely frequently

5. What type of daily activities does your dog engage in (ie., fetching newspapers, playing frisbee)?

____

____

6. Has your dog changed the <u>amount</u> of these activities?

|____|
less ... more

7. How willing is your dog to <u>play voluntarily</u>?

|____|
not at all ... very willingly

8. How often does your dog get exercise?

|____|
less than once per day ... all day

9. How stiff is your dog <u>when arising for the day</u>?

|____|
could not be more stiff ... not stiff

10. How stiff is your dog <u>at the end of the day</u> (post-activities)?

|____|
could not be more stiff ... not stiff

11. Does your dog indicate any lameness at a walk?

|____|
always ... rarely

12. Does your dog indicate any pain when turning suddenly at a walk?

|____|
always ... rarely

Thank you for completing this questionnaire.

VITA

Jonathan Thomas Hudson
132 Eagle Creek Ranch Blvd.
Floresville, TX 78114

EDUCATION

2003 Master of Science in Epidemiology, Texas A&M University

1999 Bachelor of Science in Biomedical Sciences, Minor in Business Administration, Texas A&M University

PROFESSIONAL EXPERIENCE

2002-2003 Program Coordinator, Aggie Feral Cat Alliance of Texas.

2002-2003 Teaching Assistant, Biomedical Anatomy, VAPH 305, Department of Veterinary Anatomy & Public Health, Texas A&M University.

2001-2002 Teaching Assistant, Laboratory Animal Management & Preventive Medicine, VAPH 418, Department of Veterinary Anatomy & Public Health, Texas A&M University.

2001-2002 Teaching Assistant, Meat Hygiene, VAPH 409, Department of Veterinary Anatomy & Public Health, Texas A&M University.

2001-2003 (Summers) Graduate/Research Assistant to Dr. Margaret Slater, Department of Veterinary Anatomy & Public Health, Texas, A&M University.

1998-2001 Surgical Technician, Texas A&M Small Animal Clinic.

PUBLICATIONS

Hudson J, Slater M, Taylor, L, Hulse D. The repeatability and analysis of a questionnaire to assess the degree of lameness in the canine. *Conference of Research Workers in Animal Diseases* 83. 2002.

A review of the literature intended for use by Ralston Purina on "Health effects of spaying/neutering dogs and cats" by Hudson and Slater, 2002.