Assessment of novel digital and smartphone goniometers for measurement of canine stifle joint angles

This is a summary of an article by Drs. Freund, Kieves, Hart, Foster, Jeffery and Duerr published in the American Journal of Veterinary Research 2016;77(7):749-55.1

Take Home Message

The universal goniometer (UG) had the lowest coefficient of variation (CV) and correlated most strongly with radiographic angles. However, none of the devices accurately represented radiographically measured stifle joint angles. Additional veterinary studies are indicated prior to the use of novel goniometers. Further investigation into a superior goniometer for the veterinary market is warranted.

Introduction

Goniometry is frequently used as an objective outcome measure during rehabilitation to assess the clinical progression of orthopedic patients in veterinary and human medicine.2,3 Goniometry is considered a simple, reliable, low-cost, and time-efficient outcome measurement readily available for practical clinical application.4,5

Currently used tools for measurement of joint range of motion include radiography, the commonly used plastic UG, and more specialized technologies such as electrogoniometers, bubble inclinometers, fluoroscopy, and kinematic gait analysis.5–8 Regardless of the device used, these methods require time to obtain measurements and some require additional equipment that may not be widely available to clinicians.5,9 The challenges inherent to veterinary practice require devices and methods that allow observers to accurately, reliably, and efficiently measure joint angles in a broadly diverse set of patients.6

Smartphone-based applications provide clinicians with alternative goniometers that are readily available, easily transportable, can allow for sharing of patient data, and can be less expensive than some other specialized devices used for range of motion measurement. Such applications are increasingly considered acceptable alternatives for various indications in human medicine.5,11,12 However, to our knowledge, the use of smartphone-based applications has not yet been evaluated in dogs in the peer-reviewed literature.

The purpose of the study reported here was to evaluate the reliability and accuracy of 3 novel goniometers (2 smartphone-based applications and a novel digital goniometer) for angle measurements of the stifle joint in dogs and compare the results with those obtained by use of a UG, with radiographic measurements used as the gold standard. We hypothesized that there would be no differences in these variables among the 4 goniometry devices.

Methods

Eight canine cadaver hind limbs from medium-sized dogs euthanized for reasons unrelated to this study were utilized. Limbs were excluded if palpable orthopedic abnormalities of the hind limbs were present.

Figure 1: Drawing illustrating the cadaveric set up: (A) Model design as seen when viewed from above; (B) Model design as seen when viewed from the side.
The limbs were individually mounted on wooden platforms, simulating lateral recumbency, in three random stifle angles within the published normal range of motion, resulting in a total of 24 angles. This design allowed the limb to be raised in a lateral position off the wooden platform allowing circumferential palpation of the limb mimicking a clinical setting (Figure 1). The lateral incisions were cosmetically sutured in a simple continuous subcuticular pattern.

Radiographs were obtained of the limbs at each individual angle prior to goniometric measurements being performed. The limbs remained in each position/angle until all observers completed their measurements of the respective angle. The nuts were then loosened from the bolts affixing the tibia to the wooden platform, the tibia was moved to a different angle and the nuts were tightened. The limbs were then radiographed and goniometric measurements were obtained in their new position. This sequence of limb fixation, radiography, and goniometric measurement was performed on each limb in each of the three random stifle angles per limb.

Four goniometers were used in this investigation (see Figure 2). All applications were installed on a single iphone that was used throughout the duration of the study. Three evaluators, trained in the use of each novel goniometer and experienced in utilizing the UG in daily practice, performed all measurements. The evaluators were given written and verbal instructions on the use of the methods and allowed the opportunity to practice with the devices prior to beginning measurements. The order in which the evaluators used each goniometer was designed such that no device was used repeatedly for measurement of subsequent angles. After completing a set of measurements with one device and prior to measuring with the next, the presentation of the limbs was randomly rearranged by one of the authors that did not perform any measurements. This was performed out of view of the evaluators, in an attempt to eliminate any recall of previously measured angles. This resulted in 96 measurements that were independently performed by three evaluators in triplicate fashion, for a total of 864 measurements.

For the assessment of bias, the mean of the three replicates of the radiographic angle was treated as the true value, and the mean of the three replicates of each angle measured by each observer as the measured value. In order to determine the appropriate statistical method for estimating bias, the relationship between measured and radiographic values was assessed for each device using simple linear regression. The correlation coefficient (R) was <0.99 for all goniometers, therefore Deming regression was used to assess the relationship between the angle measured by each device, and the radiographic value. Constant bias and proportional bias were considered present if the confidence interval for the y intercept or the slope of the regression line did not overlap 0 or 1 respectively. The regression equation, measured value = (slope*radiographic value) + y-intercept, was used to predict the measured angle when the radiographic value was 50 or 100 degrees. Bias at these angles was then calculated using bias(%) = (measured value-radiographic value)/radiographic value*100. Total error observed (TEo) was calculated at 50 and 100 degrees using the equation TEo(%) = 2CV(%)+bias (%), and compared to a total allowable error
(TEa) set by the authors at 5% based on previously reported differences in prognosis after tibial plateau leveling osteotomy.\(^4\)

**Results**

**Reliability** - The UG was the most and HALO the least reliable device, with mean CV of 4.88% (range 2.24-8.34) and 12.71% (range 3.11-24.63) respectively. Mean CV was similar for DrGoniometer (7.37%, range 2.44-20.40) and IHandy (7.57%, range 3.15%-14.46%).

**Correlation** – The UG correlated best with the radiographic measurements, with a correlation coefficient (R) of 0.97. HALO had the lowest correlation with R of 0.78. DrGoniometer and IHandy had an R of 0.97 and 0.93, respectively.

**Bias** – Deming regression showed evidence of proportional and constant bias for all the devices except DrGoniometer for which the confidence interval of the slope and y intercept overlapped 1 and 0 respectively. Both UG and IHandy showed a positive constant bias and HALO a constant negative bias. DrGoniometer had the lowest bias at both 50 and 100 degrees, HALO had the greatest bias at 100 degrees, and IHandy at 50 degrees.

**Performance Assessment** - All four devices exceeded the TEa of 5%, and therefore failed to meet the authors’ requirements for acceptable accuracy to represent radiographic measurements. UG performed the best at 100 degrees (TEo 18.51%) and DrGoniometer the best at 50 degrees (TEo 24.69%). HALO performed the least well of the four devices at both 100 and 50 degrees (TEo 36.90 and 46.46% respectively).

**Conclusions**

When evaluating whether a device is useful to determine a clinical measurement, several approaches can be considered. Reliability is frequently evaluated by means of the coefficient of variation. Accuracy of a device can be assessed in many different ways, including the correlation coefficient and bias. Calculation of total error integrates both CV and bias and comparison of total error with the total allowable error determines if the device meets the needs clinicians or researchers. R describes the strength of a linear association between devices, and is therefore not an ideal assessment of actual accuracy.\(^5\) Bias, an expression of inaccuracy, is an assessment of the closeness of a measurement with that of the true value.\(^6\) The evaluation of bias is beneficial in that it can be used to predict the extent of inaccuracy for different magnitudes of measurement. If proportional bias were present, this would indicate that the degree of overestimation or underestimation varies for large angles versus small angles. Constant bias, where the degree of error remains independent of the measured true value, would indicate that a device consistently under or overestimates the true value.\(^6,7\)

The results of this study demonstrate that no tested novel goniometer performed better than the UG. The UG demonstrated a low proportional and constant bias, percent bias, as well as TEo when evaluated at acute and obtuse angles. For clinical evaluations of progression after implemented therapy, the relative lack of variability is likely the most important factor to consider. The UG also had the highest correlation with the true angle as measured on radiographs, estimated by R.

Goniometry is most frequently used to assess clinical progression, such as evaluating the success of a treatment by comparing serial stifle joint angles.\(^4\) Given this purpose, the comparison of a goniometer to the radiographic measurement is likely not the most essential outcome measure for device assessment. Rather, how much a device varies between repeated measurements, between patients, and between other devices has the greater potential effect on outcome measures. Therefore, while DrGoniometer had the lowest percent bias, or highest measured accuracy when compared to the radiographic measurement, this does not suggest that DrGoniometer would be the best alternative to the UG. For the novel goniometers, IHandy demonstrated a similar proportional bias to the UG, while HALO resulted in the largest proportional bias of all devices tested. While the UG remains the preferred goniometer, IHandy would produce the most similar results to the UG in clinical use.

**Acknowledgements**

This study was supported in full by the Young Investigator Grant Program in the Center for Companion Animal Studies, Colorado State University.
References


5. Sedgwick P. Limits of agreement (Bland-Altman method), 2013.


