Effect of osteoarthritis on the repeatability of patella tendon angle measurement in dogs

Laura M Homer[1] BVM&S, MRCVS

Bárbara AJ Gomes[1] CertAVP(VDI) MRCVS

Megan C Murphy[1] DVM MRCVS

Gawain JC Hammond[1] MA VetMB MVM CertVDI DipECVDI FHEA MRCVS

Tim DH Parkin[1] BSc, BVSc, PhD, DipECVPH, FHEA, MRCVS

Cameron JA Broome[1] BVSc(Hons) DVCS FANZCVS

[1] Small Animal Hospital, School of Veterinary Medicine, University of Glasgow, Glasgow, UK

The authors declare no grant or other financial support for this report.

The authors declare no financial or other conflict of interest related to a company or product used in this report.

Corresponding author: Laura M Homer, Fitzpatrick Referrals, Halfway Lane, Eashing, Surrey, UK, GU7 2QQ. LHomer@fitzpatrickreferrals.co.uk

ABSTRACT

Objective: To evaluate the influence of osteoarthritis on the measurement of patella tendon angle (PTA) and determine intra- and inter-observer variability.

Study Design: Retrospective clinical study.

Sample Population: 87 medio-lateral radiographs obtained prior to tibial tuberosity advancement.

Methods: Radiographic osteoarthritis was scored by two observers, using guidelines derived from the International Elbow Working Group Protocol. PTA was measured by 3 observers on three occasions, with at least seven days between measurements. The data was statistically analysed via Weighted Kappa and Kruskal-Wallis testing.

Results: A fair strength of agreement was found between observers scoring osteoarthritis, with the same grades in 48% of radiographs. The intra-observer average bias between PTA measurements 1 and 3 ranged from -0.38° to -0.94°. Inter-observer bias in angle measurement ranged from -0.92° to -2.00°. Observer 1 had the narrowest range of PTA differences (12.1°), and observer 3 the highest (23.5°). Observer 2 had the lowest mean bias (-0.38°). The mean bias was lowest between observers 1 & 2 (-0.92°) and highest between 1 & 3 (-2.0°). The mean intra-observer standard deviation of the PTA measurement differences was 2.90° and inter-observer was 2.26°. The degree of osteoarthritis did not influence PTA measurements, nor their variability.

Conclusion: The current study did not find evidence of an influence of osteoarthritis on PTA, nor the repeatability of measurements.

Clinical Significance: Our findings suggest that osteoarthritis should not affect the radiographic planning for TTA surgery. The high variances in PTA measurement in less experienced observers may influence the clinical outcome of surgery.

INTRODUCTION

Cranial cruciate ligament disease is one of the most common causes of hindlimb lameness in dogs.[1] Surgical stabilisation of the stifle is recommended over conservative management due to improved outcomes, especially in larger breed dogs.[2–4] The tibial tuberosity advancement (TTA) technique aims to position the patellar ligament perpendicular to the tibial plateau by advancing the tibial tuberosity cranially.[5] The benefit of this advancement is that it theoretically reduces the tibiofemoral shear force to zero. As a consequence, the need for a functional cranial cruciate ligament is eliminated.[6]

Biomechanical studies have shown that neutralisation of tibiofemoral shear forces occurs at a PTA of $90.3 \pm 9.0°$.[7] In contrast to tibial plateau levelling osteotomy (TPLO), TTA has been shown to avoid alteration to the alignment of the femorotibial-articulating surfaces. TTA has also been shown to restore femorotibial contact mechanics to normal after surgery in-vitro.[8–10] In-vivo studies have shown a high proportion of persistent tibial subluxation postoperatively, but most dogs returned to good limb function.[11] Objective studies have documented a return of approximately 90% of normal function after TTA.[12]

However, TTA has also been shown to have a significantly higher rate of major complications and subsequent meniscal tears when compared to TPLO or the Tight Rope procedure (a modification of the lateral fabellotibial suture technique).[13] A study comparing TTA to the TPLO and lateral fabellotibial suture extracapsular repair (ECR) techniques found a lower degree of early postoperative lameness in the TTA group.[14] TTA and TPLO groups achieved normal function at the walk, but TPLO attained this earlier. Overall, the TPLO group was the only technique to achieve normal function at the trot.[14]

Preoperative planning is crucial to the TTA procedure with the requirement to assess medio-lateral radiographs of the stifle in extension at 135°. It is from these radiographs that the patellar tendon angle (PTA) is calculated. There are two main methods of measuring the PTA; (i); the conventional tibial plateau method; and (ii); the common tangent method.

The conventional tibial plateau method calculates the angle between a line representing the cranial border of the patellar ligament and a line passing through both the origins of the cranial and caudal cruciate ligaments known as the tibial plateau. The amount of tibial tuberosity advancement required to bring the patella perpendicular to the tibial plateau can then be calculated.[5]

The common tangent method defines the tibiofemoral contact point by drawing circles that correspond to the articular surfaces of both the femoral condyles and tibial plateau. A first line is drawn between the centres of these circles and a second line is then drawn perpendicular to that first line within the tibiofemoral joint space. The second line represents the common tangent and the PTA is calculated between the common tangent and a line representing the cranial border of the patella tendon.[1,15]

Studies have conflicting information regarding the most valuable measurement method. The conventional method was seen to be more reliable with better intra-observer and inter-observer reliability in a study by Millet et al.[16] There was also poor agreement between methods. The common tangent method was seen to be below anatomical measurement whereas the conventional method was found to be above anatomical measurement in a study by Bismuth et. al, giving an overall poor validity of both methods.[15] More variation was discovered with the conventional method in a study by Hoffman et al.[17] Additionally, the common tangent method was seen to be less influenced by the stifle angle. At our institution, the standard method for measurement of the PTA was the conventional method.

To the authors' knowledge there are few studies documenting the intra and inter-observer variation in tibial tuberosity advancement surgical planning. Previous studies looked at the accuracy of measurement of the tibial plateau angle for TPLO surgery within and between observers. These documented intra-observer variability of ±3.4° and inter-observer variability of ±4.8°.[18] Another study looked at the standard deviation of mean measurements and discovered an intra-observer variability of 1.5° and inter-observer variability of 0.8°.[19] We therefore hypothesized that there would be variation between measurements of the PTA both between and within observers of different experience levels and our aim was to define this level.

The amount of osteoarthritis at the caudal aspect of the tibial plateau has been found to correlate with the variation in defining the tibial plateau, affecting the planning of tibial plateau levelling osteotomies.[19] We therefore hypothesized that there would be significant, quantifiable variation between measurements of the patellar tendon angle both between and within observers of different experience levels.

Therefore, the objectives of our study were to: (i) evaluate the effect of the degree of osteoarthritis on the measurement of the PTA in dogs and (ii) determine the intra-observer and inter-observer variability of measurement of the PTA between observers of differing experience levels.

MATERIALS AND METHODS

Ethical approval for the study was granted from the institution's Research Ethics Committee prior to commencement. Pre-operative medio-lateral radiographs that were taken at the referral institution between 2008-2014 for TTA surgical planning were accessed from the institution's PACS database and viewed using DICOM viewing software (Clear Canvas, Synaptive Medical, Toronto, Canada). The radiographs were scrutinised and images which met the following inclusion criteria were kept within the study: standing stifle angle of 135° ±5° (measured via the anatomic axis method); radiograph centred on the stifle joint; and femoral condyles non-superimposed by <2mm.[20] Eighty-seven radiographs in total were selected, after a power calculation was performed to detect a 5-degree difference in angles; a standard deviation of 6-degrees; and a power of 80%.

The radiographic images were first assessed by a diplomate of the European College of Veterinary Diagnostic Imaging (Observer A) and a first-year diagnostic imaging resident (Observer B). Each radiograph was given an osteoarthritis score. The osteophytes were measured on the medio-lateral radiographs at the distal pole of the patella; femoral trochlear ridges; insertion of the cranial cruciate ligament; cranial and caudal aspects of the tibial plateau; and the fabellae. The overall degree of osteoarthritis was graded using a modified International Elbow Working Group Protocol (0 = normal - no evidence of osteophytes; 1 = mild - osteophytes of less than 2mm; 2 = moderate - osteophytes 2-5mm; and 3 = severe - osteophytes >5mm).[1,21] The observers scoring the osteoarthritis were unaware of the PTA measurements.

The radiographs were then assessed by three different observers (senior surgery clinician with a Fellowship of the Australian and New Zealand College of Veterinary Scientists [Observer 1], surgical intern [Observer 2] & first-year diagnostic imaging resident [Observer 3]). PTA was measured via the conventional tibial plateau method, on three occasions, with at least 7 days between repeated measurements (**Figure 1**). These observers were masked to the osteoarthritis score given for each radiograph, and although an impression of the degree of osteoarthritis could be estimated from viewing the radiographs, the observers made no attempt to measure or quantify this finding.

Data was analyzed with commercially available statistical software (Minitab, MiniTab Incorporated, Coventry, United Kingdom; STATA SE 12.1, College Station, TX). The osteoarthritis scores were analysed and total values for each score were calculated. Percentage agreement between observers was determined. A weighted Kappa was performed via the construction of a two-way table. K-values were interpreted via the parameters documented by Landis & Koch.[22]

Analysis of the first and third PTA observations was then performed to compare the two observations furthest apart in time. Means, standard deviations and 95% confidence intervals (CI) between the PTA recorded on measurements 1 and 3 were calculated for each observer and between different observers. Intra- and inter-observer agreement between repeated independent readings was determined via the use of Bland-Altman plots. The differences between each observer’s mean PTA measurement of 1 and 3 was plotted against the mean of the measurements. Plots were then analysed within and between the observers.

Kruskal-Wallis tests were then performed comparing the mean of the measurements of 1 and 3 and the difference between the two measurements. The aim was to identify if radiographs with higher osteoarthritis scores had different PTA and if higher osteoarthritis scores had any influence on the repeatability of the PTA angle measurement respectively. P-values <0.05 were considered significant.

RESULTS

A total of 42 of the 87 radiographs were given the same osteoarthritis score by both observers, giving a percentage agreement of 48%. A total of 39 scores had a difference of 1-point between them (45%) and a total of six had a difference of 2-points between the observers (7%). Observer B tended to grade more radiographs as grade 0 or 3 compared to Observer A, who scored the majority at grade 2. Observer A was the most experienced of the observers and graded only one radiograph as grade 0 (no osteoarthritis signs). By contrast, observer B, the least experienced observer, graded 10 radiographs as 0 (**Table 1**). A weighted Kappa was performed on the results to account for agreement occurring by chance. The K value was calculated at 0.2689. Interpreting this value with reference to the ranges put forth by Landis & Koch, the strength of agreement between observers was fair ($0.21 \leq K \leq 0.4$).[22]

The average difference (or bias) between measurement one and measurement three was calculated for each observer: Observer 1: -0.74°; Observer 2: -0.38°; and Observer 3: -0.94°. Standard deviation from the mean bias within observers ranged from 2.13° in observer 1, to 3.76° in observer 3, with an overall average of 2.90° (**Table 2**).

The mean PTA for each observer was obtained from measurement 1 & 3 and was used to calculate the difference in measurements between the observers. Observer 1 and Observer 2 produced a mean difference of -0.92°; Observer 2; and Observer 3 = -1.08°, Observer 1 and Observer 3 = -2.00°. Standard deviation from the mean bias between observers ranged from 1.82° between observers 1 and 2, to 2.65° between observers 2 & 3, with an overall average of 2.26° (**Table 3**).

Considering Bland Altman plots for each observer (Figure 2-7), the most experienced observer (Observer 1) had the lowest intra-observer variation of 12.1° and the lowest single difference of 5.2°. Observer 2 had a highest single difference between measurement 1 & 3 PTA of 11.4°. Observer 3 had the highest range of differences at 23.5° and the highest single difference of 14.6°.

A Kruskal-Wallis test of each observer's mean PTA compared to the osteoarthritis score resulted in P-values ranging from 0.224 – 0.511. A second Kruskal-Wallis test on the difference between PTA measurements and the osteoarthritis score resulted in P-values ranging from 0.108 – 0.752.

DISCUSSION

Kruskal-Wallis testing resulted in non-significant P-values when looking at each observer's mean PTA of measurements 1 & 3 compared to the osteoarthritis scores given. Therefore, we can conclude that there was no evidence of a difference in the PTA calculated with regards to the degree of osteoarthritis present. In other words, all PTAs measured were around the same value, regardless of the osteoarthritis score. Kruskal-Wallis testing also resulted in non-significant P-values when looking at the difference between each observer's PTA measurements 1 & 3 compared to the osteoarthritis scores. Therefore, we can also conclude that there was no evidence that the osteoarthritis score affected the repeatability of the PTA calculated from the radiographs.

The standard deviation of the mean PTA was similar between and within all three observers. All observers were therefore calculating angles which were within a similar range. The mean intra-observer standard deviation of the PTA measurement differences was 2.90°. The mean inter-observer standard deviation of the PTA measurement differences was 2.26°. Therefore, there was similar intra- and inter-observer deviation. These figures are higher than a similar study which documented standard deviation of tibial plateau angle measurements for TPLO surgery of 1.5° for intra-observer and 0.8° inter-observer.[19] That study included board certified surgeons, surgical residents and radiology residents. Experience levels were similar, although the surgical resident would have had more experience of making the measurements than the surgical intern in our study. They also included a total of eleven observers compared to our three, which may explain the observed differences.

Pre-surgical planning is an important aspect of the safe and accurate performance of the TTA procedure. Our study highlights the differences in angles calculated, both within and between observers of differing experience levels and speciality. Observer 1, the most experienced, demonstrated the smallest range of values and lowest standard deviation of bias. With experience, observers can more reliably and repeatedly depict the correct points on radiographs to measure the PTA. Observer 2 and Observer 3 had higher differences between angles measured, with 11.4° and 14.6° respectively.

The inter-observer variability analysis also revealed a higher range of values from the mean values when comparing the less experienced observers. These less experienced observers may have been more likely to make errors during the measurement process, which may have detrimental consequences for the surgical procedure and may have influenced our results. The variances in the PTA (5.2° to 14.6°) calculated pre-operatively are higher than expected and highlights the inconsistency and inaccuracy of the measurement process. The variance documented is likely to have consequences on the post-operative PTA, which may limit the clinical outcome of the procedure.

Observer 3 had the highest mean bias and highest standard deviation of bias. As a first-year diagnostic imaging resident, this observer would have had less clinical experience of the surgical planning technique for TTA surgery. It would be interesting to assess the change in mean bias as experience is gained over the measurement process. It has been documented that for the TTA procedure there was a learning curve of 22 procedures to gain clinical surgical competency. With the pre-surgical planning having a high influence on the final outcome, measurement of PTA may improve with experience as well.[23]

It was interesting to note the variance of osteoarthritic scores calculated for the two osteoarthritis observers. It can be postulated that with higher experience levels, an observer can depict a more subtle degree of osteoarthritic changes on radiographs. Overall agreement between observers' measurements was fair when analysed by a weighted Kappa, which is lower than expected.

A limitation of our study was the lack of reproducibility of the osteoarthritis scores between observers, which highlights the complications of such a scoring system. When the osteophytes are on the borderline of intermediate grades, this is of the highest significance, where a very mild variation of measurement could lead to a different grade. Our results may have been different with a more reproducible or detailed osteoarthritis scoring system. In addition, we only performed the osteoarthritis scoring on a single occasion with each observer. It would be worthwhile to look at repeated scoring, to see if this influences the results. Future research could investigate how the PTA measured affected the surgical planning and implant sizes chosen for the procedure to determine clinical relevance. Additionally, the common tangent method could be analysed in a similar manner to compare the degree of variation to the conventional method used in this study. Finally, this study only took into account the effect of a single variable - the mean osteoarthritis score - on the measurement of the PTA obtained. There could potentially be several factors with an influence on the angle calculated, such as breed, conformation and radiographic positioning.

In conclusion, our study showed evidence of a variation in both the osteoarthritis score and PTA measured by different observers. There was no statistically significant evidence to show that a difference in angle or its repeatability correlated with the increase in osteoarthritis score. Overall the degree of osteoarthritis did not appear to affect the variability of the PTA measured. Our findings suggest that osteoarthritis should not affect the radiographic planning of PTA measurement for TTA surgery. The high variances in PTA measurement between observers, especially in those less experienced, may influence the clinical outcome of surgery. Further clinical studies are required to investigate this.

DISCLOSURE STATEMENT

The authors declare no conflict of interest related to this report.

REFERENCES

1. Kowaleski, MP, Boudrieau, RJ, Pozzi A. *Stifle Joint. In: Tobias KM & Johnson SA, Ed. Veterinary Surgery Small Animal, Volume One.* St. Louis, Missouri: Elsevier Saunders; 2012:906-1229.

2. Vasseur PB. Clinical resuIts foIlowing nonoperative management for rupture of the cranial cruciate ligament in dogs. *Vet Surg*. 1984;13(4):234-246.

3. Pond MJ, Campell JR. The canine stifle joint I . Rupture of the anterior cruciate ligament. An assessment of conservative and surgical treatment. *J Small Anim Pr*. 1972;13(1):1-10.

4. Wucherer, KL; Conzemius, MG; Evans, R; Wilke VL. Short-term and long-term outcomes for overweight dogs with cranial cruciate ligament rupture treated surgically or nonsurgically. *J Am Vet Med* Assoc. 2013;242(10):1364-1372.

5. Tepic, S, Montavon PM. Is cranial tibial advancement relevant in the cruciate deficient stifle? *Proc 12th ESVOT Congr Munich Ger*. 2004:132-133.

6. Montavon PM, Damur DM, Tepic S. Advancement of the tibial tuberosity for the treatment of cranial cruciate deficient canine stifle. *Proc 1st World Orthop Vet Congr Munich Ger*. 2002:152.

7. Apelt D, Kowaleski MP, Boudrieau RJ. Effect of tibial tuberosity advancement on

cranial tibial subluxation in canine cranial cruciate-deficient stifle joints: An In Vitro Experimental Study. *Vet Surg*. 2007;36(2):170-177.

8. Kim SE, Pozzi A, Banks SA, Conrad BP, Lewis DD. Effect of tibial tuberosity advancement on femorotibial contact mechanics and stifle kinematics. *Vet Surg*. 2009;38(1):33-39.

9. Kim SE, Pozzi A, Banks SA, Conrad BP, Lewis DD. Effect of cranial cruciate ligament deficiency, tibial plateau leveling osteotomy, and tibial tuberosity advancement on contact mechanics and alignment of the stifle in flexion. *Vet Surg*. 2010;39(3):363-370.

10. Guerrero TG, Pozzi A, Dunbar N, et al. Effect of tibial tuberosity advancement on the contact mechanics and the alignment of the patellofemoral and femorotibial joints. *Vet Surg*. 2011;40(7):839-848.

11. Skinner OT, Kim SE, Lewis DD, Pozzi A. In vivo femorotibial subluxation during weight-bearing and clinical outcome following tibial tuberosity advancement for cranial cruciate ligament insufficiency in dogs. *Vet J*. 2013;196(1):86-91.

12. Voss K, Damur DM, Guerrero T, Haessig M, Montavon PM. Force plate gait analysis to assess limb function after tibial tuberosity advancement in dogs with cranial cruciate ligament disease. *Vet Comp Orthop Traumatol*. 2008;21(3):243-

249.

13. Christopher SA, Beetem J, Cook JL. Comparison of long-term outcomes associated with three surgical techniques for treatment of cranial cruciate ligament disease in dogs. *Vet Surg*. 2013;42(3):329-334.

14. Krotscheck U, Nelson SA, Todhunter RJ, Stone M, Zhang Z. Long term functional outcome of tibial tuberosity advancement vs. tibial plateau leveling osteotomy and extracapsular repair in a heterogeneous population of dogs. *Vet Surg*. 2016;45(2):261-268.

15. Bismuth C, Ferrand FX, Millet M, et al. Comparison of radiographic measurements of the patellar tendon-tibial plateau angle with anatomical measurements in dogs: Validity of the common tangent and tibial plateau methods. *Vet Comp Orthop Traumatol*. 2014;27(3):222-229.

16. Millet M, Bismuth C, Labrunie A, et al. Measurement of the patellar tendon-tibial plateau angle and tuberosity advancement in dogs with cranial cruciate ligament rupture. *Vet Comp Orthop Traumatol*. 2013;26(6):469-478.

17. Hoffmann DE, Kowaleski MP, Johnson KA, Evans RB, Boudrieau RJ. Ex vivo biomechanical evaluation of the canine cranial cruciate ligament-deficient stifle with varying angles of stifle joint flexion and axial loads after tibial tuberosity

advancement. *Vet Surg*. 2011;40(3):311-320.

18. Caylor KB, Zumpano CA, Evans LM, Moore RW. Intra- and interobserver measurement variability of tibial plateau slope from lateral radiographs in dogs. *J Am Anim Hosp Assoc*. 2001;37(3):263-268.

19. Fettig AA, Rand WM, Sato AF, Solano M, McCarthy RJ, Boudrieau RJ. Observer variability of tibial plateau slope measurement in 40 dogs with cranial cruciate ligament-deficient stifle joints. *Vet Surg.* 2003;32(5):471-478.

20. Osmond CS, Marcellin-Little DJ, Harrysson OLA, Kidd LB. Morphometric assessment of the proximal portion of the tibia in dogs with and without cranial cruciate ligament rupture. *Vet Radiol Ultrasound*. 2006;47(2):136-141.

21. International elbow working group protocol. *Vet Radiol Ultrasound*. 1995;36(2):172-173.

22. Landis JR, Koch GG. The measurement of observer agreement for categorical data *International Biometric Society*. 1977;33(1):159-174.

23. Proot JLJ, Corr SA. Clinical audit for the tibial tuberosity advancement procedure. *Vet Comp Orthop Traumatol*. 2013;26(4):280-284.

FIGURE LEGENDS

Figure 1: Measurement of the patellar tendon angle via the conventional tibial plateau method. Line A corresponds to the tibial plateau (a line passing through both the origins of the cranial and caudal cruciate ligaments) and line B represents the cranial margin of the patellar ligament. The angle between these lines (PTA) is the patellar tendon angle.

Figures 2-4: Bland-Altman plots displaying the intra-observer variation of PTA between measurements 1&3. The mean of the two measurements is on the X-axis and the difference between the two measurements is on the Y-axis.

Figures 5-7: Bland-Altman plots displaying the inter-observer variation of the mean PTA of measurements 1&3. The mean of the measurements between observers is on the X-axis and the difference of the mean between the observers is on the Y-axis.

TABLES

Table 1: The osteoarthritis (OA) scores of Observer A compared to Observer B.

		Observer A OA SCORE			
Observer B OA SCORE		**0**	**1**	**2**	**3**
	0	1	3	6	0
	1	0	6	15	0
	2	0	9	34	2
	3	0	0	10	1

Table 2: Intra-observer agreement for patellar tendon angle measurements based on mean PTA, range of and total differences, 95% limits and mean bias.

Observer	Mean PTA ± Standard Deviation (°)	Range of difference (°)	Total range of difference (°)	95% Limits agreement (°)	Mean Bias (°) ± Standard Deviation (°)
1	102.75 ± 5.29	-6.9 – 5.2	12.1	-4.92; 3.44	-0.74 ± 2.13
2	103.58 ± 5.43	-6.7 – 11.4	18.1	-5.88; 5.11	-0.38 ± 2.80
3	104.74 ± 5.53	-14.6 – 8.9	23.5	-8.31; 6.43	-0.94 ± 3.76

Table 3: Inter-observer agreement for patellar tendon angle measurements based on mean PTA, range of and total differences, 95% limits and mean bias.

Observer	Mean PTA ± Standard Deviation (°)	Range of difference (°)	Total range of difference (°)	95% Limits agreement (°)	Mean Bias (°) ± Standard Deviation (°)
1 vs 2	103.06 ± 5.31	-5.5 – 5.4	10.9	-4.48; 2.65	-0.92 ± 1.82
2 vs 3	104.06 ± 5.20	-8.6 – 6.2	14.8	-6.27; 4.11	-1.08 ± 2.65
1 vs 3	103.60 ± 5.16	-7.5 – 4.9	12.4	-6.52; 2.52	-2.00 ± 2.31