Clinical Research

Tibial Tubercle-Trochlear Groove Distance Is a Reliable and Accurate Indicator of Patellofemoral Instability

Giampietro L. Vairo PhD, ATC, Joaquin Moya-Angeler MD, Michael A. Siorta MD, Ashley H. Anderson MD, Paul S. Sherbondy MD

Abstract

Background Tibial tubercle-trochlear groove (TT-TG) distance is a measurement generally made on CT scans that is commonly used to quantify the risk of patellofemoral instability (PFI); however, its interrater reliability and accuracy as an indicator of PFI in patients is poorly characterized.

Questions/purposes The purposes of our study were (1) to primarily analyze interrater reliability of the TT-TG distance among orthopaedists with varied experience as measured by MRI, (2) to secondarily compare TT-TG distances between PFI and control groups, and (3) to determine an accurate TT-TG distance threshold indicative of PFI.

Methods The electronic medical records of a senior fellowship-trained orthopaedic sports medicine surgeon were surveyed between 2012 and 2016 for patients who had experienced at least one episode of patellar subluxation or dislocation, who reported no other knee-related history, and who underwent MRI due to persistent PFI signs and symptoms. The records of 48 PFI patients (23 males, 25 females; 19 ± 4 years of age) were compared with 83 controls (60 males, 23 females; 31 ± 8 years of age) having no history of PFI, presenting with an isolated meniscal lesion as determined from MRI and treated by the same orthopaedist during this time. All records meeting study criteria were consecutively included to offset selection bias of the retrospective analysis. Two sports medicine fellows, one who had just completed orthopaedic residency training, and another with a year of experience after residency, and a sports medicine subspecialist with more than 15 years of experience in practice independently recorded TT-TG distance, indicative of tibial tubercle lateralization relative to the femoral trochlea, to the nearest millimeter (mm) in a blinded and randomized fashion. Intraclass correlation coefficient computed interrater reliability accompanied by standard error of measurement (SEM); a one-tailed, two-sample t-test analyzed group differences with accompanying effect size per Cohen's d; receiver operating characteristic (ROC) curve determined accuracy and threshold for PFI risk. A p value < 0.05 denoted statistical significance.

G. L. Vairo, Departments of Kinesiology and Orthopaedics & Rehabilitation, Colleges of Health & Human Development, and Medicine, The Pennsylvania State University, University Park, PA, USA

J. Moya-Angeler, M. A. Siorta, A. H. Anderson, P. S. Sherbondy, Department of Orthopaedics & Rehabilitation, College of Medicine, The Pennsylvania State University, University Park, PA, USA

G. L. Vairo (✉), 146 Recreation Building, Departments of Kinesiology, and Orthopaedics & Rehabilitation, Colleges of Health & Human Development, and Medicine, The Pennsylvania State University, 146 Recreation Building, University Park, PA 16802, USA, Email: glv103@psu.edu

Results Interrater reliability was excellent, at 0.93 (95% confidence interval [CI], 0.84–0.97; SEM = 0.6 mm) for PFI patients and 0.95 (95% CI, 0.91–0.97; SEM = 0.4 mm) for controls. Distance was greater (95% CI, 2–5; $p < 0.001$) in PFI patients (14 ± 4 mm; range = 7–24 mm) than controls (10 ± 3 mm; range = 3–19 mm) with an effect size of 1 (95% CI, 0.3–2). Area under the ROC curve was 0.75 (95% CI, 0.66–0.83) and threshold was 13 mm (sensitivity = 0.52, 1-specificity = 0.25), suggesting the measure is a fairly accurate indicator of risk and values of 13 mm or greater are better suited to rule in PFI. Respective positive and negative likelihood ratios of 2 and 0.6 at this threshold confirm that this distance yields a small increase in probability for PFI and a minimal decrease in probability for risk; specifically, a 13-mm TT-TG distance is two times more likely to be found in patients with PFI. Furthermore, this threshold is estimated to increase a correct PFI diagnosis by approximately 15%.

Conclusions The MRI-specific TT-TG distance, based on a single measurement using cartilaginous-tendon landmarks within a standardized trochlear range, is reliable as performed by orthopaedists of varied experience. Patients with PFI display a 4-mm greater distance than controls, which may represent a difference large enough for clinicians to discern in practice using MRI. A 13-mm TT-TG distance is two times more likely seen in patients with PFI. However, this threshold increases a correct PFI diagnosis by only about 15%; therefore, clinical decision-making should not be influenced by this criterion alone and instead used in conjunction with other relevant variables.

Level of Evidence Level IV, diagnostic study,

Introduction

Patellofemoral instability (PFI) is a multifactorial condition often associated with recurrent patellar dislocations. The diagnosis of PFI is made through a comprehensive clinical evaluation, often incorporating the use of MRI or CT scan [21, 26, 32, 65]. An important measurement made on these imaging tests is the tibial tubercle-trochlear groove (TT-TG) distance [3, 4, 15, 26, 69], which can help determine whether nonsurgical management may be appropriate [83], and if not, can help guide the surgeon toward an appropriate surgical approach, which might involve reconstruction of the medial patellofemoral ligament or tibial tubercle osteotomy [6, 22, 48, 82, 83]. The TT-TG distance assesses lateralization of the tibial tubercle as originally described with the use of planar radiography [33].

The accuracy of the TT-TG distance measurement on a CT scan depends on the localization of bony landmarks as reference points [23, 73], and CT has been the preferred method because of its ability to offset radiographic imaging artifact [77]. The use of a CT scan for quantifying TT-TG

distance is reported to be reliable [41, 66, 69, 86] and has been proposed to be associated with PFI risk [18, 23, 35, 67, 81] when the measurement has been made by musculoskeletal radiologists, physicians, and various other health care personnel. Despite the advantages of a CT scan, its applicability by contemporary clinical practice standards is limited given that MRI is more routinely performed subsequent to PFI episodes because of its ability to evaluate the integrity of related soft tissue structures while reducing radiation exposure [39]. In addition, the ability of MRI to image soft tissues facilitates identification of more applicable cartilaginous and tendinous reference points for calculating TT-TG distance. While this has caused some authors to assert that MRI is the better test to use, the fact that TT-TG distance values measured on MRI and CT scan appear to differ (with MRI values generally being lower [14, 39]) has resulted in considerable ambiguity for how the measure should be used in practice [7, 14, 38, 41, 55, 58, 80, 84]. In addition, although intrarater reliability for the measure has been reported [80, 86], interrater reliability estimates have varied or are incompletely described [1, 10, 12, 13, 16, 25, 35-39, 68, 80], making interpretation and application to practice still more challenging. Finally, the standardized difference of MRI-based TT-TG distance between PFI patients and controls as well as accuracy of this specific measurement technique to gauge PFI risk is currently inconclusive [39, 65, 80].

Therefore, we sought (1) to primarily analyze interrater reliability of the TT-TG distance among orthopaedists with varied experience as measured by MRI, (2) to secondarily compare TT-TG distances between PFI and control groups, and (3) to determine an accurate TT-TG distance threshold indicative of PFI.

Methods and Materials

A retrospective cohort experimental design was conducted at our university bone and joint institute. After institutional review board approval, a musculoskeletal radiologist surveyed the electronic medical records of a senior investigator (PSS) between 2012 and 2016 for patients who experienced at least one episode of patellar subluxation or dislocation, who reported no other knee-related history, and who underwent MRI due to persistent PFI signs and symptoms after an initial bout of conservative treatment. The MRI studies included preceded potential surgical intervention at the knee. Patients who sustained a knee ligament injury or who underwent surgery on the affected knee before the time of MRI were excluded. This generated the records of 48 PFI patients (23 males, 25 females; 19 ± 4 years) for analysis. A control group consisting of the records of 83 individuals (60 males, 23 females; 31 ± 8 years) who had an isolated meniscal lesion as determined

from MRI and who were treated by the same orthopaedist during the same period were identified for inclusion. Controls who had a documented patellar dislocation, patellofemoral pathology, or underwent knee surgery before MRI were excluded. With records occasionally lacking patient height, weight, self-reported physical activity level or Kujala questionnaire scores, appropriate matching of cases to controls was not possible. While an inability to control for these factors limits the internal validity of our study, it does lend to its external validity, and thus generalizability of our outcomes. Based on the retrospective nature of our analysis, all records meeting study criteria in this timeframe were consecutively included to offset potential selection bias. Although this resulted in an unbalanced array of samples between groups and distribution of sex in the controls, evaluations of data normality, randomness and variance were conducted to ensure our statistical conclusion validity.

Measurement Technique

MRI was performed with an Optima™ MR450w, 70 cm, 1.5-T MR System (GE Healthcare, Chicago, IL, USA), per the technical specifications of Hinckel et al. [39]. Patients were imaged in a standardized supine position with the knee maximally extended, the quadriceps muscles relaxed, and the patella oriented anteriorly [13, 23, 58, 86]. Measurement of TT-TG distance was calculated to the nearest millimeter (mm) using cartilaginous-tendon landmarks as described by Schoettle et al. [66] (Fig. 1) using Advantage Workstation for Diagnostic Imaging (GE Healthcare). Cartilaginous-tendon landmarks were used to measure TT-TG distance given their anatomic basis and related biomechanical implications as well as representing the more contemporary application for associated analyses [38, 58, 66, 75, 80]. To identify an ideal range for performing such measures, we quantified TT-TG distance within an interval of 6 mm through 9 mm below the most proximal groove point of the trochlea [86]. The specific measurement parameters described were adopted to facilitate a standardized technique, and thus precision, among raters.

Three orthopaedists independently performed all measurements in a blinded and randomized fashion (so that the diagnosis was unknown); two were sports medicine fellows, one of whom (AHA) had just completed orthopaedic residency training, and the other (MAS) had a year of experience after residency, and one (PSS) was a senior sports medicine subspecialist with more than 15 years of experience in practice. The raters performed one measure of TT-TG distance for each imaging study as a means to reflect the mode in which the measurement technique is pragmatically performed in clinical applications. Thereafter, a separate author (JMA) organized the data per
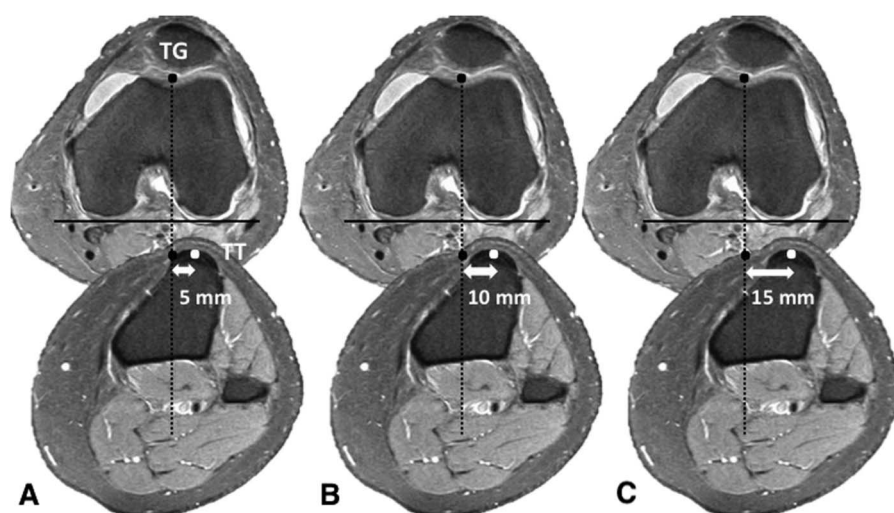

Fig. 1 A-C The MRI measurement technique for TT-TG distance is demonstrated. The black dot represents orientation of the trochlear groove (TG) designated by the deepest cartilaginous point on the image. The tibial tubercle (TT) is denoted by the white dot as determined from the central insertion site of the patellar tendon. The horizontal continuous line illustrates the posterior intercondylar line. The vertical dashed line renders the intersection of the intercondylar line per position of the TG, which is used to calculate distance in reference to the TT. (**A-B**) This is a typical range for what constitutes a normal TT-TG distance [58, 75, 83]; (**C**) this resembles an abnormal TT-TG distance that may receive limited benefit from rehabilitation in the presence of PFI [83].

PFI patient and control in a manner to gauge reliability among raters for each group. Subsequently, the data of all raters were averaged per PFI patients and controls for the purpose of comparing differences between groups, evaluating accuracy of TT-TG distance to determine PFI risk and identifying an optimal-related threshold value as measured by MRI.

Statistical Analyses

The lead author (GLV) performed all statistical analyses. An a priori sample size calculation was based on our primary aim for analyzing reliability [30]. Per the results of Hinckel et al. [39], a minimum of 46 PFI patient and 48 control records were necessary to satisfy the following indices: significance level (α) of 0.05; power (1-β) of 80%; acceptable reliability (ρ_0) of 70%; and expected reliability (ρ_1) of 80% using an intraclass correlation coefficient (ICC) model. Coincidentally, these sample sizes were sufficient to also adequately power our secondary aims evaluating differences between PFI patients and controls [2], and accuracy [56] of the TT-TG distance to determine PFI. IBM® SPSS Statistics (IBM Corp, Armonk, NY, USA) software was used to determine if the data met necessary assumptions for the subsequent statistical analyses described. The ICC (3,1) was computed for determining interrater reliability [8, 11, 17, 47, 51] and was interpreted per Portney and Watkins [63]; the ICC (3,1) statistic denotes that every patient in our study was assessed by each rater; furthermore, each rater performed a single measure of the TT-TG distance. This was accompanied by a calculated standard error of measurement (SEM); smaller values of this metric constitute greater precision [79]. Mean and SD were calculated for PFI and control groups. A homogeneity of variance test determined that equal variance existed between PFI and control groups to pool SDs for hypothesis testing and mitigate Type I error. A one-tailed, two-sample t-test analyzed between-group differences accompanied by effect size (Cohen's d) to indicate the standardized difference between means [50, 57, 74] and interpreted accordingly [20]. A ROC curve was plotted to examine the area under the curve (AUC) for quantifying accuracy of the TT-TG distance measure for determining PFI risk [49] using standard operational definitions [34]. From the ROC curve data, an optimal threshold value for risk of PFI was computed per the index of union (IU) method. The IU method identifies a cut point on the ROC curve that maximizes sensitivity and specificity while being both simultaneously close to the AUC value and minimizing the difference between sensitivity and specificity [76]. This method was identified as an optimal threshold given the premise for its use in attempting to classify most of the sample population correctly and,

thus, the least incorrectly [62]. To bolster statistical inference for the optimal threshold, and thus, clinical interpretation, we used sensitivity and specificity indices to calculate likelihood ratios [59] and interpreted them accordingly [43, 54]. A p value < 0.05 denoted a statistically significant difference a priori.

Results

Interrater reliability for measurement of TT-TG distance on MRI among our three surgeons was excellent with assessment of both the PFI (0.93; 95% confidence interval [CI], 0.84–0.97; SEM = 0.6 mm) and control (0.95; 95% CI, 0.91–0.97; SEM = 0.4 mm) groups.

Distance was greater (95% CI, 2–5; p < 0.001) in PFI patients (14 ± 4 mm; range, 7–24 mm) than controls (10 ± 3 mm; range, 3–19 mm) with an effect size of 1 (95% CI, 0.3–2).

Accuracy of the TT-TG distance was fair for determining PFI risk and the associated optimal threshold value was 13 mm with corresponding indices of sensitivity = 0.52 and 1-specificity = 0.25 (Fig. 2), suggesting that values of 13 mm or greater are better suited to rule in, rather than rule out, PFI. Respective positive and negative likelihood ratios of 2 and 0.6 (Table 1) at this
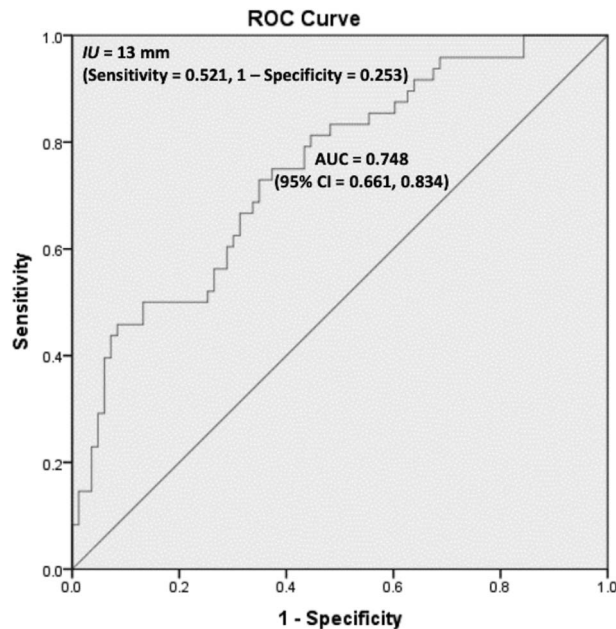
Fig. 2 Accuracy and optimal threshold value of TT-TG distance to determine PFI risk. Accuracy is reflected by the AUC. The optimal threshold, IU, is obtained from the equation IU(c) = (| sensitivity(c) − AUC | + | specificity(c) − AUC |) where c denotes cut points on the ROC curve; mathematically, the cut point minimizing the IU(c) function and the | sensitivity(c) − specificity(c) | difference is designated as an optimal value [76].

Wolters Kluwer

Table 1. Relative indices for a TT-TG distance of 13 mm as an optimal threshold for determining PFI risk

Statistic	Formula	Value (95% CI)
Sensitivity	TP/(TP + FN) x 100%	52% (37%-67%)
Specificity	TN/(FP + TN) x 100%	75% (64%-84%)
Positive likelihood ratio	Sensitivity/1-specificity	2 (1-3)
Negative likelihood ratio	1-sensitivity/specificity	0.6 (0.5-0.9)

TT-TG = tibial tubercle-trochlear groove; PFI = patellofemoral instability; CI = confidence interval; TP = true-positive; FN = false-negative; FP = false-positive; TN = true-negative.

threshold confirm the 13-mm distance yields a small increase in probability for PFI and a minimal decrease in probability for risk. This suggests a 13-mm TT-TG distance is twice as likely to be found in patients with PFI; moreover, this threshold is estimated to increase a correct PFI diagnosis by approximately 15%.

Discussion

Differences in CT scan and MRI yield dissimilar TT-TG distance measures that result in varying clinical interpretations and, thus, implications. While previous authors [1, 10, 12, 13, 16, 25, 35-39, 68, 80] have reported outcomes on diverse reliability measures of MRI-based TT-TG distance, these data are often limited in scope, lack essential experimental design and statistical elements, and thus, require replication for corroboration or refined interpretation [30, 49, 50]. Furthermore, an analysis of TT-TG distance accuracy and threshold value for determining PFI risk as measured by MRI is warranted to determine the measure's utility by current clinical practice standards. We found that using MRI to quantify TT-TG distance was reliable even among observers of disparate levels of experience and that patients with PFI displayed a 4-mm greater distance than controls, which may represent a clinical difference. We also discovered that MRI-based TT-TG distance measures are accurate in determining PFI risk and a threshold value of 13 mm is better suited to rule in PFI, but is estimated to only increase likelihood of a correct diagnosis by about 15%.

This study had several limitations, the chief being that our results are based on static measures of TT-TG distance. There is evidence to suggest that this kind of screening does not predict dynamic lateral patellar tracking [15]; moreover, our measurements were performed with the knee maximally extended and unloaded, which may not be consistent with outcomes garnered from varying positions of knee flexion or when weightbearing [25, 42]. In light of these limitations, our findings demonstrate that the static TT-TG distance is an indicator of PFI risk, particularly when appraised in a systematic, standardized manner yielding an absolute measure [53, 72], but our findings may

not correlate perfectly with dynamic patellar maltracking. Further, patellofemoral instability is linked to a myriad of influential variables that span a spectrum of modifiable to nonmodifiable factors, which were not controlled for in this study; because of this, we cannot account for the potential contributing effects of such things like soft tissue constraints, articular shape of patella and trochlea, patellar orientation, skeletal alignment, muscular strength and control as well as level of physical activity. Based on our retrospective design, we could not account for missing anthropometric and demographic information from medical records such as height [61] and knee size [40] that have been suggested to impact TT-TG distance. We note that contrasting evidence proposes that height may not correlate with TT-TG distance [72], or that it may only potentially do so in a pediatric population whose body size varies drastically depending on prepubescent age, developmental rate and skeletal maturity [28]. Furthermore, knee size has been suggested to be a highly fluctuating variable that does not affect TT-TG distance across the population [28]. These findings complement those suggesting that sex-based assumptions of anthropometry should not influence the interpretation of TT-TG distance [40, 85, 71]; nonetheless, such variables warrant clinical consideration and require future study given that we could not confirm or refute their influence through our analyses. Moreover, clinical utility of TT-TG distance may be increased when considered in combination with other individual-specific anatomical contributing factors, such as trochlear dysplasia, which is known to markedly influence PFI [5, 16, 70].

Although we met most of Standards for Reporting Diagnostic Accuracy (STARD) 2015 criteria, which lists the essential items for reporting diagnostic accuracy in studies, in disseminating our work [9], we did not account for the number of instability episodes in categorizing PFI patients, which represents a limitation since individuals with a recurrent history are documented to have a greater TT-TG distance than those with only a single occurrence [85]; nonetheless, our 13-mm threshold for PFI risk closely parallels the identified 13.5-mm value found to increase relative risk for recurrent instability episodes [85]. Moreover, although raters assessed MRI records in a blinded and randomized fashion, the inspection of signs like bony

edema or patellofemoral ligamentous injury may have introduced observational bias in our study [9], which is typical of similarly designed studies and can be difficult to control for [28].

Our reliability results contrast with those of Charles et al. [16], who noted only a fair degree of interrater reliability, which may stem from apparent inconsistencies in how distances were accounted for in their study; for instance, an experienced musculoskeletal radiologist only performed about 12% of the measures with the remaining allotment completed by individuals who were described as two medical students. Another possible source of difference may be associated with the heterogeneity of populations studied given that Charles et al. [16] focused on a sample of patients with recurrent dislocation of whom nearly 70% had trochlear dysplasia, which has been documented to influence related reliability measures [29]. In contrast, every orthopaedic surgeon in our study performed ratings for all records in a standardized fashion that did not account for trochlear dysplasia. By contrast, our findings complement most comparable studies [1, 3, 10, 12-14, 25, 31, 35-39, 58, 66, 68, 80, 86], although our results are original when specifically describing reliability from a single measurement using a robust statistical approach [44] absent in the preceding analyses. This is noteworthy given that it reflects proficiency of orthopaedists with varying levels of experience to use the measure in a manner that mimics MRI-based clinical applications with a high degree of precision across a broad patient spectrum [39].

We identified TT-TG distance for PFI patients to be greater than controls, which was consistent with published results [2, 3, 10, 46, 61] that conducted this comparison using experimental methods similar to ours. Specifically, we found patients with PFI display a 4-mm greater distance than controls, which may represent a clinical difference per our effect size that parallels most comparable studies [3, 10, 46, 61] assessed with a post-hoc analysis (Table 2). Failing

to control for MRI resolution may underpin those instances where contrasting effect sizes have been found. In related investigations [2, 3], data were collected on both 1.5-T and 3-T modalities, which confound results since values garnered with 3-T imaging are suggested to be approximately 2 mm less than 1.5 T [24]. A degree of fluctuation noted in comparing our results with previous findings may also be linked to diverse sample sizes and heterogeneity of the studied populations. Consequently, none of the prior pertinent experiments provide a sample size calculation, indicating previously published results may be insufficiently powered, whereas ours was more than adequately powered. An additional factor potentially compromising previous investigations may be the trochlear level at which TT-TG distance was calculated. Yin et al. [86] found that measures were most consistent when performed proximally, which is why we conducted them in a range 6 to 9 mm below the most proximal groove point of the trochlea. Although Brady et al. [10] is the only prior group to indicate levels of the trochlea at which measures were taken, they report no differences between the proximal and distal levels; however, their measurement intervals were not as discrete as Yin et al. [86], which suggests additional study is warranted to definitively determine whether this variable is important.

We noted accuracy of the MRI-based TT-TG distance as an indicator of PFI risk to be less than that reported for CT scan using bony landmarks [64], but the CT study included only Korean patients [64]; therefore, it may be important to study the anatomy we are evaluating in a variety of populations. Specifically, we found a TT-TG distance of 13 mm to be the best cutoff value to use for PFI diagnosis in our series. This finding suggests that imaging and landmark-specific thresholds are necessary to correctly detect risk in patients [14, 41], especially since few PFI patients evaluated with MRI reach the 15- to 20-mm cutoff range that prior studies have suggested when using a CT scan [1, 7, 16, 18, 19, 27, 37, 45, 64, 75]. Coincidentally,

Table 2. Comparative TT-TG distances (mm) between PFI cases, and controls as measured by MRI with cartilaginous-tendon landmarks that include post-hoc analyses of percent difference, and effect size

Authors	PFI cases (N; M:F)	Controls (N; M:F)	% difference	d (95% CI)
Balcareck et al . [2]	13 ± 4 (73; 40:33)	11 ± 4 (73; 40:33)	18	0.6 (-0.1, 1.2)
Balcareck et al. [3]	13 ± 5 (43; 43:0)	11 ± 4 (78; 78:0)	14	0.4 (-0.4, 1.1)
	14 ± 4 (50; 0:50)	10 ± 4 (71; 0:71)	34	1 (0.5, 1.8)*
Brady et al. [10]†	14 ± 5 (129; 30:99)	10 ± 5 (109; 69:40)	36	0.9 (0.3, 1.5)*
Kohlitz et al. [46]	15 ± 6 (186; 79:101)	10 ± 4 (186; 79:105)	39	1 (0.5, 1.5)*
Pennock et al. [61]	16 ± 5 (45; 23:22)	12 ± 5 (180; 104:76)	33	0.9 (0.3, 1.6)*

*Denotes statistical significance.
†Measurements collapsed across all levels of the trochlear range given no differences between proximal, and distal TT-TG distance levels; values are mean ± SD; PFI = patellofemoral instability; N = total count of sample population; M:F = male:female numbers enrolled; % Diff = percent difference; d = Cohen's effect size; CI = confidence interval.

unlike Dejour et al. [23], our threshold was found to be 1 mm less than our PFI patients' mean values and slightly less than those of most comparable studies, which proposes that relying on descriptive statics alone may limit efficacious clinical decision-making [52, 60, 78]. While the MRI cutoff we identified is helpful, the value provided only 52% sensitivity and 75% specificity, suggesting it is better suited to rule in PFI rather than rule it out. The associated positive likelihood ratio suggests a 13-mm TT-TG distance is twice as likely to be found in PFI patients. Subsequently, the cutoff's negative likelihood ratio suggests a distance below this value is found in approximately six of 10 PFI patients. When collectively considered, this indicates a 13-mm TT-TG distance yields a small increase in probability for PFI and a minimal decrease in risk [43]. Simplifying the interpretation of likelihood ratios suggests this threshold is estimated to increase a correct PFI diagnosis by approximately 15% [54]; therefore, clinical decision-making should not be influenced by this criterion alone and instead used in conjunction with other relevant variables.

In conclusion, we found interrater reliability of measuring TT-TG distance using 1.5T MRI to be excellent even among orthopaedic surgeons of different levels of experience, and that patients with PFI display a 4 mm greater clinical difference than controls. A 13-mm cutoff is the best for identifying patients who might be at risk for PFI, and based on the 50% sensitivity and 75% specificity, this value is better used to rule in PFI than to rule it out. Furthermore, clinicians should consider that a 13-mm TT-TG distance is twice as likely to be found in PFI patients, and that use of this cutoff is estimated to increase the probability of a correct PFI diagnosis by only about 15%. Future studies might look to repeat this experiment using relative measures of the TT-TG distance with focus on its relation to influential variables like trochlear dysplasia.

References

1. Anley CM, Morris GV, Saithna A, James SL, Snow M. Defining the role of the tibial tubercle-trochlear groove and tibial tubercle-posterior cruciate ligament distances in the work-up of patients with patellofemoral disorders. *Am J Sports Med*. 2015;43:1348-1353.
2. Balcarek P, Ammon J, Frosch S, Walde TA, Schuttrumpf JP, Ferlemann KG, Lill H, Strumer KM, Frosch K-H. Magnetic resonance imaging characteristics of the medial patellofemoral ligament lesion in acute lateral patellar dislocations considering trochlear dysplasia, patella alta, and tibial tuberosity–trochlear groove distance. *Arthroscopy*. 2010;26:926-936.
3. Balcarek P, Jung K, Ammon J. Anatomy of lateral patellar instability: trochlear dysplasia and tibial tubercle-trochlear groove distance is more pronounced in women who dislocate the patellar. *Am J Sports Med*. 2010;38:2320-2327.
4. Balcarek P, Jung K, Frosch K-H, Sturmer KM. Value of the tibial tuberosity-trochlear groove distance in patellar instability in the young athlete. *Am J Sports Med*. 2011;38:1756-1761.
5. Balcarek P, Oberthur S, Hopfensitz S, Frosch S, Walde TA, Wachowski MM, Schuttrumpf JP, Strurmer KM. Which patellae are likely to redislocate? *Knee Surg Sports Traumatol Arthrosc*. 2014;22:2308-2314.
6. Balcarek P, Walde TA, Frosch S, Schuttrumpf JP, Wachowski MM, Sturmer KM, Frosch K-H. Patellar dislocations in children, adolescents and adults: a comparative MRI study of medial patellofemoral ligament injury patterns and trochlear groove anatomy. *Eur J Radiol*. 2011;79:415-420.
7. Berruto M, Ferrua P, Carimati G, Uboldi F, Gala L. Patellofemoral instability: classification and imaging. *Joints*. 2013;1:7-13.
8. Bland JM. *An Introduction to Medical Statistics*. Oxford University Press, Oxford, United Kingdom; 1987.
9. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Limer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF, STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology*. 2015;277:826-832.
10. Brady JM, Sullivan JP, Nguyen J, Mintz D, Green DW, Strickland S, Stein BES. The tibial tubercle-to-trochlear groove (tt-tg) distance is reliable in the setting of trochlear dysplasia, and superior to the tibial tubercle-to-posterior cruciate ligament (tt-pcl) distance when evaluating coronal malalignment in patellofemoral instability. *Arthroscopy*. 2017;33:2026-2034.
11. Bruton A, Conway JH, Holgate ST. Reliability: what is it, and how is it measured? *Physiotherapy*. 2000;86:94-99.
12. Camp CL, Heidenreich MJ, Dahm DL, Bond JR, Collins MS, Krych AJ. A simple method of measuring tibial tubercle to trochlear groove distance on MRI: description of a novel and reliable technique. *Knee Surg Sports Traumatol Arthrosc*. 2016;24:879-884.
13. Camp CL, Heidenreich MJ, Dahm DL, Stuart MJ, Levy BA, Krych AJ. Individualizing the tibial tubercle-trochlear groove distance; patellar instability ratios that predict recurrent instability. *Am J Sports Med*. 2015;44:393-399.
14. Camp CL, Stuart MJ, Krych AJ, Levy BA, Bond JR, Collins MS, Dahm DL. CT and MRI measurements of tibial tubercle-trochlear groove distances are not equivalent in patients with patellar instability. *Am J Sports Med*. 2013;41:1835-1840.
15. Carlson VR, Sheehan FT, Shen A, Yao L, Jackson JN, Boden BP. The relationship of static tibial tubercle-trochlear groove measurement and dynamic patellar tracking. *Am J Sports Med*. 2017;45:1856-1863.
16. Charles MD, Haloman S, Chen L, Ward SR, Fithian D, Afra R. Magnetic resonance imaging-based topographical differences between control and recurrent patellofemoral instability patients. *Am J Sports Med*. 2013;41:374-384.
17. Chinn S. Statistics in respiratory medicine. 2. Repeatability and method comparison. *Thorax*. 1991;46:454-456.
18. Christensen TC, Sanders TL, Pareek A, Mohan R, Dahm DL, Krych A. Risk factors and time to recurrent ipsilateral and contralateral patellar dislocations. *Am J Sports Med*. 2017;45:2105-2110.
19. Clark D, Walmsley K, Schranz P, Mandalia V. Tibial tuberosity transfer in combination with medial patellofemoral ligament reconstruction: surgical technique. *Arthroscopy*. 2017;6:r591-e597.
20. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
21. Colvin AC, West RV. Patellar instability. *J Bone Joint Surg Am*. 2008;90:2751-2762.
22. Dejour D, Le Coultre B. Osteotomies in patello-femoral instabilities. *Sports Med Arthrosc*. 2007;15:39-46.

23. Dejour D, Walch H, Nove-Josserand L, Guier C. Factor of patellar instability: an anatomic radiographic study. *Knee Surg Sports Traumatol Arthosc*. 1994;2:19-26.

24. Dickens AJ, Morrell NT, Doering A, Tandberg D, Treme G. Tibial tubercle-trochlear groove distance: defining normal in a pediatric population. *J Bone Joint Surg Am*. 2014;96:318-324.

25. Dietrich TH, Betz M, Pfirrmann CWA, Koch PP, Fucentese SF. End-stage extension of the knee and its influence on tibial tuberosity-trochlear groove distance (TTTG) in asymptomatic volunteers. *Knee Surg Sports Traumatol Arthrosco*. 2014;22:214-218.

26. Dietrich TH, Fucentese SF, Pfirrmann CW. Imaging of individual anatomical risk factors for patellar instability. *Semin Musculoskelet Radiol*. 2016;20:65-73.

27. Diks MJF, Wymenga AB, Anderson PG. Patients with lateral tracking of patella have better pain relief following CT-guided tuberosity transfer than patients with unstable patella. *Knee Surg Sports Traumatol Arthrosc*. 2003;11:384-388.

28. Dornacher D, Reichel H, Kappe T. Does tibial tuberosity-trochlear groove distance (TT-TG) correlate with knee size or body height? *Knee Surg Sports Traumatol Arthrosc*. 2016;24:2861-2867.

29. Dornacher D, Reichel H, Lippacher S. Measurement of tibial tuberosity-trochlear groove distance: evaluation of inter- and intraobserver correlation dependent on the severity of trochlear dysplasia. *Knee Surg Sports Traumatol Arthrosc*. 2014;22:2382-2387.

30. Farrokhyar F, Reddy D, Poolman RW, Bhandari M. Why perform a priori sample size calculation? *Can J Surg*. 2013;56:207-213.

31. Ferlic PW, Runer A, Dirisamer F, Balcarek P, Giesinger J, Biedermann R, Liebensteiner MC. The use of tibial tuberosity-trochlear groove indices based on joint size in lower limb evaluation. *Inter Orthop*. 2018;42:995-1000.

32. Fiqueroa D, Novoa F, Melean P, Calvo R, Vaisman A, Fiqueroa F, Delgado G. Usefulness of magnetic resonance imaging in the evaluation of patellar alignment. *Rev Esp Cir Ortop Taumatol*. 2014;58:19-23.

33. Goutallier D, Bernageau J, Lecudonnec B. The measurement of the tibial tuberosity: patella groove distanced technique and results. *Rev Chir Orthop Reparatrice Appar Mot*. 1978;64:423-428.

34. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29-36.

35. Heidenreich MJ, Camp CL, Dahm DL, Stuart MJ, Levy BA, Krych AJ. The contribution of the tibial tubercle to patellar instability: analysis of tibial tubercle-trochlear groove (TT-TG) and tibial tubercle-posterior cruciate ligament (TT-PCL) distances. *Knee Surg Sports Traumatol Arthrosc*. 2017;25:2347-2351.

36. Heidenreich MJ, Sanders TL, Hevesi M, Johnson NR, Wu IT, Camp CL, Dahm DL, Krych AJ. Individualizing the tibial tubercle to trochlear groove distance to patient specific anatomy improves sensitivity for recurrent instability. *Knee Surg Sports Traumatol Arthrosc*. 2018;9:2858-2864.

37. Hinckel BB, Gobbi RG, Filho EN, Pécora JR, Camanho GL, Rodriguez MB, Demange MK. Are the osseous and tendinous-cartilaginous tibial tuberosity-trochlear groove distances the same on CT and MRI? *Skeletal Radiol*. 2015;44:1085-1093.

38. Hinckel BB, Gobbi RG, Kihara Filho EN, Demange MK, Pécora JR, Camanho GL. Patellar tendon-trochlear groove angle measurement. A new method for patellofemoral rotational analyses. *Orthop J Sports Med*. 2015;3:1-7.

39. Hinckel BB, Gobbi RG, Kihara Filho EN, Demange MK, Pécora JR, Rodriguez MB, Camanho GL. Why are bone and soft tissue measurements of the TT-TG distance on MRI different in patients with patellar instability? *Knee Surg Sports Traumatol Arthosc*. 2017;25:3053-3060.

40. Hinglebaum S, Best R, Huth J, Wagner D, Bauer G, Mauch F. The TT-TG index: a new knee size adjusted measure to determine the TT-TG distance. *Knee Surg Sports Traumatol Arthrosc*. 2014;22:2388-2395.

41. Ho CP, James EW, Surowiec RK, Gatlin CC, Ellman MB, Cram TR, Dornan GJ, LaPrade RF. Systematic technique-dependent differences in CT versus MRI measurement of the tibial tubercle-trochlear groove distance. *Am J Sports Med*. 2015;43:675-682.

42. Izadpanah K, Weitzel E, Vicari M, Hennig J, Weigel M, Sudkamp NP, Niemeyer P. Influence of knee flexion angle and weight bearing on the tibial tuberosity-trochlear groove distance for evaluation of patellofemoral alignment. *Knee Surg Sports Traumatol Arthrosc*. 2014;22:2655-2661.

43. Jaeschke R, Guyatt GH, Sackett DL. Users; guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients. *J Am Med Assoc*. 1994;271:703-707.

44. Keating J, Matyas T. Unreliable inferences from reliable measurements. *Aust J Physiother*. 1998;44:5-10.

45. Koeter S, Diks MJF, Anderson PG, Wymenga AB. A modified tibial tubercle osteotomy for patellar maltracking. *J Bone Joint Surg Br*. 2007;89:180-185.

46. Kohlitz T, Scheffler S, Jung T, Hiburg Am Vollnberg B, Wiener E, Diederichs G. Prevalence and patterns of anatomical risk factors in patients after patellar dislocation: a case control study using MRI. *Eur Radiol*. 2013;23:1067-1074.

47. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155-162.

48. Laidlaw MS, Diduch DR. Current concepts in the management of patellar instability. *Indian J Orthop*. 2017;51:493-504.

49. Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. *J Biomed Inform*. 2005;38:404-415.

50. Lee DK. Alternatives to P-value: confidence interval and effect size. *Korean J Anesthesiol*. 2016;69:555-562.

51. Lee KM, Lee J, Chung CY, Ahn S, Sung KH, Kim TW, Lee HJ, Park MS. Pitfalls and important issues in testing reliability using intraclass correlation coefficients in orthopaedic research. *Clin Orthop Surg*. 2012;4:149-155.

52. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5:1315-1316.

53. Marquez-Lara A, Andersen J, Lenchik L, Ferguson CM, Gupta P. Variability in patellofemoral alignment measures on MRI: influence on knee position. *Am J Roentgenol*. 2017;208:1097-1102.

54. McGee S. Simplifying likelihood ratios. *J Gen Intern Med*. 2002;17:647-650.

55. Moezzi D. Editorial Commentary: Elevated tibial tubercle-trochlear groove distance and patellar instability. *Arthroscopy*. 2016;32:62.

56. Obuchowski NA. Fundamentals of clinical research for radiologists: ROC analysis. *AJR Am J Roentgenol*. 2005;184:364-372.

57. Page P. Beyond statistical significance: clinical interpretation of rehabilitation research literature. *Inter J Sports Phys Ther*. 2014;9:726-736.

58. Pandit S, Frampton C, Stoddart J, Lynskey T. Magnetic resonance imaging assessment of tibial tuberosity-trochlear groove

distance: normal values for males and females. *Int Orthop*. 2011; 35:1799-1803.

59. Parikh R, Parikh S, Arun E, Thomas R. Likelihood ratios: clinical application in day-to-day practice. *Indian J Ophthamol*. 57: 217-221.

60. Pauker SG, Kassirer JP. The threshold approach to clinical decision-making. *N Engl J Med*. 1980;302:1109-1117.

61. Pennock AT, Alam M, Bastrom T. Variation in tibial tubercle-trochlear groove measurement as a function of age, sex, size, and patellar instability. *Am J Sports Med*. 2013;42:389-393.

62. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cut-points using two ROC based criteria. *Am J Epidemiol*. 2006;163: 670-675.

63. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. Upper Saddle River, New Jersey: Prentice Hall; 2000.

64. Prakash J, Seon J-K, Woo SH, Jin C, Song EK. Comparison of radiological parameters between normal and patellar dislocation groups in Korean population A rotational profile CT –based study. *Knee Surg Relat Res*. 2016:28:302-311.

65. Ridley TJ, Hinckel BB, Kruckberg BM, Agel J, Arendt EA. Anatomical patella instability risk factors on MRI show sensitivity without specificity in patients with patellofemoral instability: a systematic review. *J ISAKOS*. 2016;1:141-152.

66. Schoettle PB, Zanetti M, Seifert B, Pfirrmann CWA, Fucentese SF, Romero J. The tibial tuberosity-trochlear groove distance: a comparative study between CT and MRI scanning. *Knee*. 2006; 13:26-31.

67. Schueda MA, Astur DC, Bier RS, Bier DS, Astur N, Cohen M. Use of computer tomography to determine the risk of patellar dislocation in 921 patients with patellar instability. *Open Access J Sports Med*. 2015;6:66-62.

68. Shakespeare D, Fick D. Patellar instability-can the TT-TG distance be measured clinically? *Knee*. 2005;12:201-204.

69. Skelley N, Friedman M, McGinnis M, Smith C, Hillen T, Matava M. Inter- and intraobserver reliability in the MRI measurement of the tibial tubercle-trochlear groove distance and trochlea dysplasia. *Am J Sports Med*. 2015;43:873-878.

70. Smith TO, Davies L, Toms AP, Hing CB, Donell ST. The reliability and validity of radiological assessment for patellar instability: a systematic review and meta-analysis. *Skeletal Radiol*. 2011;40:399-414.

71. Sobhanardekani M, Sobhan MR, Moghadam RN, Nabavinejad S, Kazem S, Ratki R. The normal value of tibial tubercle trochlear groove distance in patients with normal knee examinations using MRI. *Acta Med Iran*. 2017;55:573-577.

72. Soumalainen JS, Regalado G, Joukainen A, Kääriäinen T, Könönen M, Manninen H, Sipola P, Kokki H. Effects of knee flexion and extension on the tibial tuberosity-trochlear groove (TT-TG) distance in adolescents. *J Exp Orthop*. 2018;5:31.

73. Steensen RN, Bentley JC, Trinh TQ, Backes JR, Wiltfong RE. The prevalence and combined prevalences of anatomic factors associated with recurrent patellar dislocation. A magnetic resonance imaging study. *Am J Sports Med*. 2015;43:921-927.

74. Sullivan GM, Feinn R. Using effect size-or why the P value is not enough. *J Grad Med Educ*. 2012;4:279-282.

75. Tsavalas N, Katonis P, Karantanas AH. Knee joint anterior malalignment and patellofemoral osteoarthritis: an MRI study. *Eur Radiol*. 2012; 22:418-428.

76. Unal I. Defining an optimal cut-point value in ROC analysis: an alternative approach. *Comput Math Methods Med*. 2017;2017: 3762651.

77. Wagenaar FC, Koeter S, Anderson PG, Wymenga AB. Conventional radiography cannot replace CT scanning in detecting tibial tubercle lateralisation. *Knee*. 2007;14:51-54.

78. Watson PF, Petrie A. Method agreement analysis: a review of correct methodology. *Theriogenology*. 2010;73:1167-179.

79. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005; 19:231-240.

80. Wilcox JJ, Snow BJ, Aoki SK, Hung M, Burks RT. Does landmark selection affect the reliability of tibial tubercle-trochlear groove measurements using MRI? *Clin Orthop Rel Res*. 2012; 470:2253-2260.

81. Williams AA, Elias JJ, Tanaka MJ, Thawait GK, Demehri S, Carrino JA, Cosgarea AJ. The relationship between tibial tuberosity-trochlear groove and abnormal patellar tracking in patients with unilateral patellar instability. *Arthroscopy*. 2016;32: 55-61.

82. Wittstein JR, Bartlett EC, Easterbrook J, Bryd JC. Magnetic resonance imaging evaluation of patellofemoral malalignment. *Arthroscopy*. 2006;22:643-649.

83. Wittstein JR, O'Brien SD, Vinson EN, Garrett WE Jr. MRI evaluation of anterior knee pain: predicting response to nonoperative treatment. *Skeletal Radiol*. 2009;38:895-901.

84. Yao L, Gai N, Boutin RD. Axial scan orientation and the tibial tubercle-trochlear groove distance: error analysis and correction. *AJR Am J Roentgenol*. 2014;202:1291-1296.

85. Yeoh CSN, Lam KY. Tibial tubercle to trochlear groove distance and index in children with one-time versus recurrent patellar dislocation: a magnetic resonance imaging study. *J Orthop Surg*. 2016;24:253-257.

86. Yin L, Chen C, Duan X, Deng B, Xiong R, Wang F, Yang L. Influence of the image levels of the distal femur on the measurement of tibial tubercle-trochlear groove distance – a comparative study. *J Orthop Surg Res*. 2015;16:174.