The Knee 54 (2025) 217–221

Contents lists available at ScienceDirect

The Knee

journal homepage: www.elsevier.com/locate/thekne

Patellar tilt calculation utilizing artificial intelligence on CT knee imaging ☆

Johannes Sieberer [a,b,*], Albert Rancu [a], Nancy Park [a], Shelby Desroches [a], Armita R. Manafzadeh [a], Steven Tommasini [a], Daniel H. Wiznia [a,b], John Fulkerson [a]

[a] *Yale School of Medicine – Orthopaedics & Rehabilitation, 47 College Street, New Haven, CT, USA*

[b] *Yale School of Engineering and Applied Science – Department of Mechanical Engineering and Material Science, 17 Hillhouse, New Haven, CT, USA*

A R T I C L E I N F O

Article history:
Received 13 May 2024
Revised 11 February 2025
Accepted 15 February 2025

Keywords:
Patellar tilt
Artificial Intelligence
Three-dimensional (3D)

A B S T R A C T

Background: In the diagnosis of patellar instability, three-dimensional (3D) imaging enables measurement of a wide range of metrics. However, measuring these metrics can be time-consuming and prone to error due to conducting 2D measurements on 3D objects. This study aims to measure patellar tilt in 3D and automate it by utilizing a commercial AI algorithm for landmark placement.

Methods: CT-scans of 30 patients with at least two dislocation events and 30 controls without patellofemoral disease were acquired. Patellar tilt was measured using three different methods: the established method, and by calculating the angle between 3D-landmarks placed by either a human rater or an AI algorithm. Correlations between the three measurements were calculated using interclass correlation coefficients, and differences with a Kruskal-Wallis test. Significant differences of means between patients and controls were calculated using Mann-Whitney U tests. Significance was assumed at 0.05 adjusted with the Bonferroni method.

Results: No significant differences (overall: $p = 0.10$, patients: 0.51, controls: 0.79) between methods were found. Predicted ICC between the methods ranged from 0.86 to 0.90 with a 95% confidence interval of 0.77–0.94. Differences between patients and controls were significant ($p < 0.001$) for all three methods.

Conclusion: The study offers an alternative 3D approach for calculating patellar tilt comparable to traditional, manual measurements. Furthermore, this analysis offers evidence that a commercially available software can identify the necessary anatomical landmarks for patellar tilt calculation, offering a potential pathway to increased automation of surgical decision-making metrics.

Abbreviations: 2D, two-dimensional; 3D, three-dimensional; ICC, Inter class correlation; NMDID, New Mexico Decedent Image Database; KW, Kruskall-Wallis test; MWU, Mann-Whitney *U* test; STD, standard deviation.

☆ This article is part of a special issue entitled: 'AI and new technologies' published in The Knee.

* Corresponding author at: Yale School of Medicine – Orthopaedics & Rehabilitation, 47 College Street, New Haven, CT, USA.

E-mail addresses: johannes.sieberer@yale.edu (J. Sieberer), albert.rancu@yale.edu (A. Rancu), nancy.park@yale.edu (N. Park), shelby.desroches@yale.edu (S. Desroches), armita.manafzadeh@yale.edu (A.R. Manafzadeh), steven.tommasini@yale.edu (S. Tommasini), daniel.wiznia@yale.edu (D.H. Wiznia), john.fulkerson@yale.edu (J. Fulkerson).

https://doi.org/10.1016/j.knee.2025.02.019

1. Introduction

In the diagnosis and treatment of patellar instability, three-dimensional (3D) imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) facilitate the measurement of a number of informative metrics [1]. However, measuring metrics from 3D imaging is time-consuming and requires advanced training for the reading radiologist and orthopedic surgeon. In addition, many metrics manually measured from 3D imaging are known to have low interrater reliability [1].

Automated analysis of 3D imaging is a potential solution, eliminating the need for a specifically trained clinician and standardizing measurement methods. Patellar instability presents a unique opportunity for automation due to the abundance of existing metrics and their relative simplicity in calculation (e.g., angle between two lines, perpendicular distance between two points). One such potential automatable metric is patellar tilt, defined as the angle between the posterior condylar line and the line drawn through the maximum width of the patella as seen in the axial plane of a CT or MRI scan. Although known to have good interrater reliability with an ICC of 0.84 [1], the current methodology relies on the ability of the user to identify the correct 2D slices to make an accurate measurement (see Figure 2), thus requiring an adequately trained physician and their time. Given the relatively small size of the patella compared to the slice thickness generated by typical medical scanners, this task relies on individual judgement and can be prone to error without sufficient training.

The current study presents a methodology for calculating patellar tilt that avoids the subjective selection of correct 2D slice for measurement. Artificial intelligence (AI) was used to segment 3D CT scans for subjects with and without recurrent patellar dislocation and to derive patellar tilt from relevant surface landmarks. The results of the AI-derived measurements were compared to the traditional method and manual placement of landmarks done by one of the study authors.

2. Materials and methods

2.1. Participant selection

Thirty patients treated by the senior author (JPF) who demonstrated recurrent lateral patellofemoral instability, defined as at least two dislocation events, were identified. High-resolution single-knee CT scans that were originally obtained as part of standard clinical care were deidentified and exported as Digital Imaging and Communications in Medicine (DICOM) files. All patient imaging was taken before any surgical intervention at our institution. Patients who had undergone previous patellofemoral surgeries at other institutions, such as lateral release, were excluded. The control group consisted of 15 bilateral knee CT scans leading to 30 control knees obtained from patients with no history of patellofemoral disease from the New Mexico Decedent Image Database (NMDID) [3]. Our institution's Institutional Review Board (IRB) deemed this study exempt.

2.2. Image segmentation and landmark placement

Imaging datasets were automatically segmented using the Simpleware ScanIP (Sunnyvale, CA) AS ORTHO functionality. ScanIP in itself as a segmentation software is FDA 510(k) cleared, but the AS ORTHO functionality is not included in the 510(k) and its outputs need manual confirmation by a physician. This program relies on a proprietary machine learning algorithm to identify the major bony components of the patellofemoral and tibiofemoral joints including the femur, tibia, patella, and fibula. In addition, the AS ORTHO functionality automatically places key landmarks across the knee (see Figure 2), including on the posterior condyles of the femur and on the medial and lateral borders of the patella. These four landmarks are used here for calculation of patellar tilt. The algorithm is deterministic, meaning it will have the same output every time it is run regardless of the user and therefore has perfect interrater reliability. Separately, one author (ARM), a postdoctoral researcher in anatomy trained by the senior author, placed the four needed landmarks manually on each knee within the dataset. This process has been described in a study by Park et al. [4].

2.3. Manual patellar tilt calculation

Patellar tilt was measured manually by one of the authors (a fourth-year medical student [NP] doing a research year in sports medicine and trained by the senior author) using the traditional method [2]. Using axial 2D images taken from CT scans, lines were drawn across the posterior femoral condyles and through the maximum width of the patella. The angle between these lines was defined as the manual patellar tilt calculation (see Figure 1).

2.4. Three-dimensional patellar tilt calculation

Landmarks (as shown in Figure 2) were either placed manually by one of the authors or by a commercial algorithm included in the ScanIP software. Using the manually (3D-manual) and AI-placed (3D-automated) landmarks, patellar tilt was calculated by an algorithm developed by the study team as follows: vectors were drawn in 3 dimensions between the two femoral condylar landmarks and the two patella border landmarks; the projection of both vectors into the XY-plane was obtained, where the Z-axis was defined by the scanner's axial axis. The angle formed between these two lines

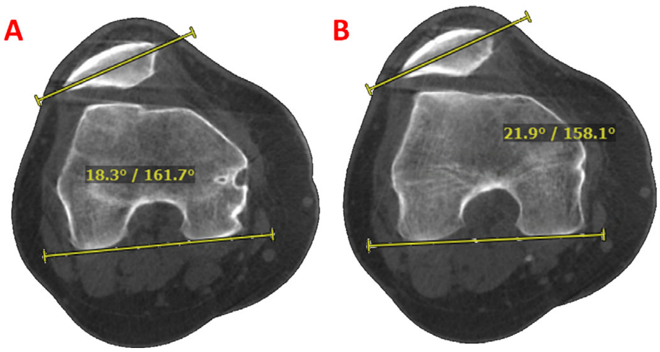

Figure 1. Established 2D Patellar Tilt Measurement – (A + B) Patellar tilt measurement according to Dejour [2] for the same knee and scan, measured on different slices of the CT scan, resulting in different values for the patellar tilt angle.

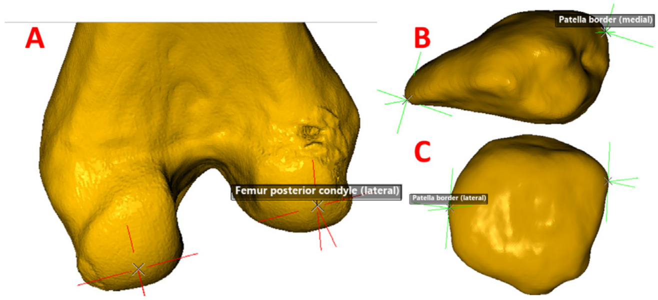

Figure 2. Placement of 3D Landmarks – 3D models of the posterior, distal femur (A) and the patella in axial (B) and coronal (C) view. The four landmarks (i.e., femoral posterior condyles and patellar borders) needed for patellar tilt calculation were placed on these models in the segmentation software, by AI and a human rater separately.

constituted the patellar tilt measurement. This process mirrors the traditional method of measuring patellar tilt. The algorithm and dataset developed to calculate patellar tilt from the acquired landmarks is open-source and has been uploaded in an online repository [5].

2.5. Statistical analysis

Differences between 2D, 3D-manual, and 3D-automated patellar tilt measurements were compared using the Kruskal-Wallis (KW) test for the whole dataset as well as the control and instability groups individually. Interclass correlation (ICC) between the three methods presented here were calculated using two-way, single score, absolute agreement ICCs (A,1) and their reliability discussed according to previous literature [6]. All analysis was conducted in MATLAB R2023b (Mathworks Inc, Natick, MA). The ICCs used here do not measure interrater or intrarater reliability, but agreement between different methods.

Differences in patellar tilt between the control and instability patient populations for each methodology used here were analyzed using the Mann-Whitney U (MWU) test. Statistically significant differences were defined as having a *p*-value less than 0.05 adjusted with the Bonferroni method.

3. Results

Patellar tilt was calculated using traditional methodology 3D manual, and 3D automated approaches for 30 patients' knees (19 female 11 male, age: 24.5 ± 10.8 years) and 30 control knees (16 female and 14 male, age: 26.8 ± 9.9 years, scans were bilateral with 15 control patients in total). The aggregated measurement results are presented in Table 1. Significant differences ($p < 0.001$) between the patient and control cohort were found with all three measurement methods. There were no significant differences in patellar tilt (all *p*-values greater than or equal to 0.10) between the measurement methods.

Interclass correlation coefficients between the different patellar tilt measurement methods are shown in Table 2. The agreement of the methods is regarded as good to excellent [6].

The AI algorithm was able to place landmarks in an appropriate location for 59 knees, while it failed to correctly identify medial and lateral for the patellar border landmarks in one knee, naming them "patella border (1)" and "patella border (2)" instead. Thus, the algorithm did not calculate the patellar tilt angle. These landmarks were manually renamed by one of the authors without moving the landmarks. Subsequently, the algorithm was used to calculate the angle.

Table 1
Mean and standard deviation for the complete dataset, the patient and control cohort. Significant differences between patients and controls were found. No significant differences between the methods were found.

Group	Overall (n = 60) Mean ± STD	Patients (n = 30) Mean ± STD	Control (n = 30) Mean ± STD	Difference between patients-controls p-value MWU
2D Measurement	18.3° ± 11.0°	26.5° ± 9.2°	10.1° ± 4.9°	<0.001
Manual 3D Measurement	19.9° ± 10.9°	27.8° ± 8.3°	12.0° ± 6.5°	<0.001
AI 3D Measurement	17.5° ± 11.0°	26.5° ± 6.6°	8.5° ± 6.1°	<0.001
KW p-value	0.10	0.51	0.79	–

Table 2
Inter class correlation (ICC) between the different methods. The agreement between methods were interpretated according to previous literature [6].

ICC (A,1)	2D Measurement	3D Measurement	AI Measurement
2D Measurement	–	0.87 [0.79–0.92]	0.86 [0.77–0.91]
3D Measurement	Good to excellent	–	0.90 [0.84–0.94]
AI Measurement	Good to excellent	Good to excellent	–

4. Discussion

The present study demonstrated a method for calculating patellar tilt using 3D anatomic landmarks placed by an author and by an automated algorithm. Both sets of results were compared to those generated using the traditional method. We found good to excellent agreement among the three methods. The agreement between the AI and traditional method is comparable to the traditional method's interrater reliability (0.84 [2]). As the AI-algorithm for placing the needed landmarks is deterministic, its interrater reliability would be, by definition, perfect (ICC: 1.00), and was therefore not tested. Each method had a statistically significant difference in measured patellar tilt between patients with a history of patellar instability and a control group. This outcome is in line with established literature [7–9] and validates the 3D approach. Notably, there were no significant differences in means between the 2D and 3D method of the measurement. Additionally, the agreement in ICCs was quite high. This shows that at least in large cohort studies the two methods will provide comparable results. Since the agreement is not perfect (i.e., ICC < 1.00), there might be differences in a case by case basis.

The results in this study demonstrate potential for the viability of complete automation of the patellar tilt metric utilizing a 3D landmark approach, ultimately enabling the reduction of workload for clinicians. The scripting functionality of the segmentation software we used and similar software packages enables broad implementation of this workflow. We believe this algorithm can be used on one hand in large scale research studies where the sheer amount of data will limit the possibility of a physician to manually measure each knee. On the other hand, larger hospital systems could use this or similar implementation to support their radiologist in achieving the necessary throughput.

Notably, of the 60 segmentations we carried out in total, there was one case (1.7%) where the software failed to correctly identify the lateral and medial side of the patella This error lead to the angle not being calculated, necessitating intervention by one author. We deem this to be acceptable as this would not adversely impact clinical decision making. Still, we suggest that such algorithms should not be relied on exclusively, but rather should be used as an aid and productivity tool in conjunction with physician input and validation. Verification of a physician is aided by displaying the landmarks on their respective anatomical features and is conducted by confirming the placement is correct.

5. Limitations

The reliability and accuracy of the traditional patellar tilt measurement is good but not perfect; therefore, we lack a gold standard to compare this approach to. Any measured difference to the current method could stem from inaccuracy or machine error in our method, or from inaccuracy or human error in the traditional method. Currently, we do not have a reliable method to distinguish where the error is coming from, therefore we can only say that our method is comparable to the traditional method of measuring patellar tilt. Finally, all CT scans were performed on supine patients, so we do not have any data on how the described methods measure patellar tilt on weightbearing knees in which the position of the patella would likely be changed.

The automatic landmarking algorithm is embedded in commercial software and was treated like a black box; as such, we do not have control over its internal model and any subsequent changes to it in the future. While AI models tend to improve each version, this may not necessarily be true. To ensure the validity of measurement results before deploying an algorithm update, the fully automated part of this study can be repeated as part of the clinical quality management system (QMS). In this study, we used version U-2022.12-SP2 of the Simpleware ScanIP software.

The patellar tilt angle is not enough to determine the need for surgery. Multiple factors (e.g., trochlea dysplasia or TT-TG) [10] need to be considered for clinical decision making. Therefore, this algorithm alone is not sufficient to remove the need for manual measurements done by radiologists.

6. Conclusion

The study offers an alternative, automated 3D approach to calculating patellar tilt with outputs comparable to those resulting from traditional, manual measurements. Furthermore, this approach offers a potential pathway to aiding radiologists in more efficiently providing surgical decision-making metrics.

CRediT authorship contribution statement

Johannes Sieberer: Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Albert Rancu:** Writing – original draft, Methodology. **Nancy Park:** Writing – original draft, Methodology, Data curation. **Shelby Desroches:** Writing – original draft, Validation, Methodology, Data curation. **Armita R. Manafzadeh:** Writing – review & editing, Methodology, Data curation. **Steven Tommasini:** Writing – review & editing. **Daniel H. Wiznia:** Writing – review & editing, Funding acquisition. **John Fulkerson:** Writing – review & editing, Supervision, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The study team did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The code and data used in this study is available in the provided repository [5]. Imaging was acquired from NMDID. The Free Access Decedent Database funded by the National Institute of Justice grant number 2016-DN-BX-0144.

References

[1] Fabricant PD, Heath MR, Mintz DN, Emery K, Veerkamp M, Gruber S, et al. Many radiographic and magnetic resonance imaging assessments for surgical decision making in pediatric patellofemoral instability patients demonstrate poor interrater reliability. Arthroscopy 2022;38(9):2702–13. doi: https://doi.org/10.1016/j.arthro.2022.03.033.

[2] Dejour H, Walch G, Nove-Josserand L, Guier C. Factors of patellar instability: an anatomic radiographic study. Knee Surg Sports Traumatol Arthrosc 1994;2(1):19–26. doi: https://doi.org/10.1007/bf01552649.

[3] Edgar HDB, Moes E, Adolphi NL, Bridges P, Nolte KB. New Mexico Decedent Image Database 2020. doi: https://doi.org/10.25827/5s8c-n515.

[4] Park N, Sieberer J, Manafzadeh A, Hackbarth R, Desroches S, Ghankot R, et al. Semiautomated three-dimensional landmark placement on knee models is a reliable method to describe bone shape and alignment. Arthrosc Sports Med Rehabilit 2024. doi: https://doi.org/10.1016/j.asmr.2024.101036.

[5] Sieberer JM. Code and data for: patellar tilt calculation utilizing artificial intelligence on CT knee imaging. Zenodo 2024. doi: https://doi.org/10.5281/zenodo.11053095.

[6] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155–63. doi: https://doi.org/10.1016/j.jcm.2016.02.012.

[7] Gobbi RG, Cavalheiro CM, Giglio PN, Hinckel BB, Camanho GL. Patellar tilt and patellar tendon-trochlear groove angle present the optimum magnetic resonance imaging diagnostic reliability for patients with patellar instability. Arthroscopy 2023;39(11):2339–51. doi: https://doi.org/10.1016/j.arthro.2023.04.005.

[8] Tan SHS, Kwan YT, Lee JZJ, Yeo LKP, Lim AKS, Hui JH. Patellar tilt, congruence angle, and tibial tubercle-trochlear groove distance are correlated with positive J-sign in adolescents. Phys Sportsmed 2024:1–5. doi: https://doi.org/10.1080/00913847.2024.2315012.

[9] Iacobescu G, Cirstoiu C, Cursaru A, Anghelescu D, Stanculescu D. Correlation between patellar tilt angle, femoral anteversion and tibial tubercle trochlear groove distance measured by computer tomography in patients with non-traumatic recurrent patellar dislocation. Maedica (Bucur) 2020;15(2):174–80. doi: https://doi.org/10.26574/maedica.2020.15.2.174.

[10] Dejour DH, Mesnard G, Giovannetti de Sanctis E. Updated treatment guidelines for patellar instability: "un menu à la carte". J Exp Orthop 2021;8(1):109. doi: https://doi.org/10.1186/s40634-021-00430-2.