

Original Article

Low measurement reliability of emergency department radiographs of patellar dislocation

Misk Ghassan Farik Muhammad¹, Abdullah Ghassan Farik Muhammad², Anders Bøvling³, Rune Bech⁴, Bertel Understrup³ & Bjarke Viberg³

1) Department of Plastic Surgery and Burns Treatment, Copenhagen University Hospital – Rigshospitalet, 2) Department of Surgery, Zealand University Hospital, Roskilde, 3) Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, 4) Department of Orthopaedic Surgery, Zealand University Hospital, Køge, Denmark

Dan Med J 2026;73(11):A11250912. doi: 10.61409/A11250912

ABSTRACT

INTRODUCTION. Trochlear dysplasia and patella alta are key markers for recurrent patellar dislocation. However, these measurements are developed for true lateral X-rays, and their reproducibility on routine ED X-rays with malrotation remains uncertain. This study assessed the intra- and inter-rater reliability of measurements under these conditions.

METHODS. In this single-centre reliability study, 114 adults with primary patellar dislocation were retrospectively identified from the ED registry at Odense University Hospital (2018–2021), Denmark. Four raters (a medical student, two residents and a knee consultant) assessed each X-ray twice, ≥ 14 days apart. Trochlear dysplasia was evaluated using the Dejour classification, anterior translation and trochlear depth. Patella height was measured with the Insall-Salvati, Blackburne-Peel and Caton-Deschamps indices. Reliability was estimated using unweighted kappa statistics.

RESULTS. The Insall-Salvati index showed the highest reproducibility (intra-rater $\kappa = 0.63$ – 0.85 ; inter-rater $\kappa = 0.59$ – 0.70). Trochlear dysplasia measurements showed variable inter-rater reliability: $\kappa = 0.05$ – 0.39 for the Dejour classification, -0.08 – 0.60 for anterior translation, and 0.13 – 0.70 for trochlear depth. The Blackburne-Peel and Caton-Deschamps indices showed low reliability (inter-rater $\kappa \leq 0.33$ and ≤ 0.58).

CONCLUSIONS. The Insall-Salvati index demonstrated the highest and most consistent reproducibility on routine ED X-rays. Other indices and trochlear dysplasia measurements showed low agreement, likely reflecting the challenges of applying these measurements to X-rays with variable malrotation. These findings underline the importance of correct application and interpretation of X-ray measurements.

FUNDING. None.

TRIAL REGISTRATION. Not relevant.

Primary patellar dislocation most commonly occurs in adolescents and young adults, with an incidence estimated at 29–43 per 100,000 person-years, particularly affecting adolescents [1].

Following acute dislocation, ED X-rays are routinely obtained to evaluate predisposing anatomical factors. Among the most critical markers associated with recurrent patellar instability are trochlear dysplasia and patella alta [2, 3]. Trochlear dysplasia is characterised by a shallow or flattened trochlear groove that reduces osseous containment and facilitates lateral patellar displacement [2]. Patella alta refers to an abnormally high patella, often due to patellar tendon elongation, which delays patellofemoral engagement during flexion and increases

the risk of instability [3, 4].

Patellar height is commonly assessed using the Insall-Salvati [5], Blackburne-Peel [6] and Caton-Deschamps [7] indices on lateral knee X-ray. These measurements are used to identify patellofemoral malalignment [3]. However, their reproducibility has been questioned due to variability in anatomical landmarks and X-ray positioning [5]. The Dejour classification depends on subtle X-ray features and may show limited reproducibility [2, 8].

These X-ray measurements were developed and validated on true lateral X-rays, where femoral condyles are superimposed. However, in routine clinical practice, particularly in the ED, X-rays are often obtained under suboptimal conditions without strict rotational alignment control. Even minor deviations from a true lateral projection may alter the appearance of anatomical landmarks and affect measurement accuracy.

Despite the clinical relevance of these measurements, there is limited evidence evaluating their reproducibility in patients with primary patellar dislocation. Previous studies have reported limitations of X-ray assessment of patella height and trochlear dysplasia, largely based on standardised imaging protocols or specialised clinical settings [2, 5, 8]. In contrast, early evaluation after primary patellar dislocation is often based on routine ED X-rays, where data on measurement reproducibility remain limited.

This study aimed to assess the intra- and inter-rater reproducibility of commonly used X-ray measurements for trochlear dysplasia and patellar height in patients with primary patellar dislocation when applied to routine ED knee X-rays with non-true lateral positioning.

Methods

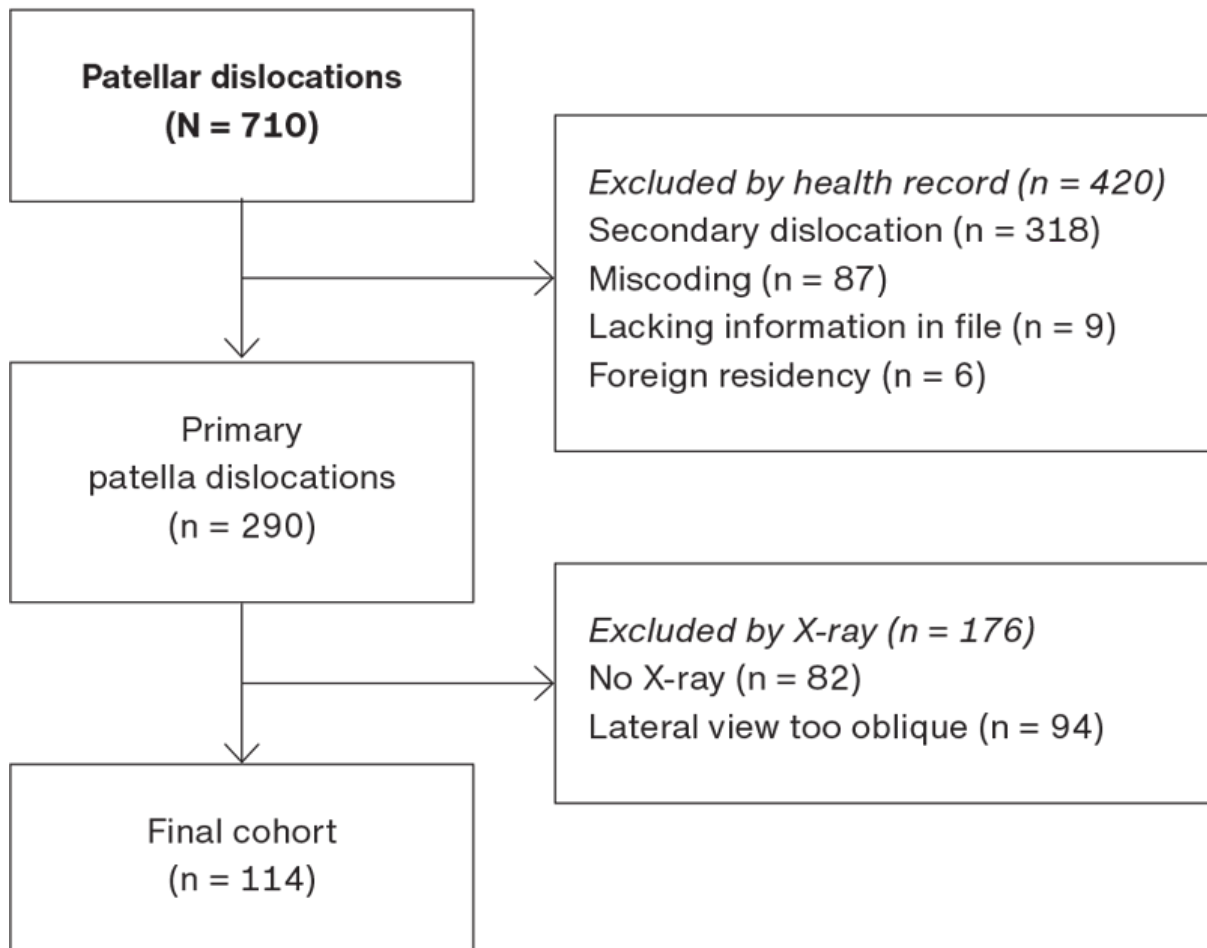
Design

This single-centre reliability study of X-ray measurements in adult patients with primary patellar dislocation was conducted at the Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, Denmark. The Danish Data Agency approved the study (case number 14/15156), and reporting followed the Guidelines for Reporting Reliability and Agreement Studies [9]. All procedures involving human participants were conducted in accordance with the ethical standards of the Danish Data Protection Agency and with the 1964 Helsinki Declaration and its subsequent amendments. As this was a retrospective analysis of anonymised X-rays, informed consent was waived by the regional ethics committee.

Patients

A retrospective search of the hospital administrative system was performed for patients with patellar dislocation treated in the ED at Odense University Hospital using the International Statistical Classification of Diseases and Related Health Problems, tenth revision (ICD-10) diagnosis code for patellar dislocation (DS830) over a three-year period. Health records and X-rays were reviewed for inclusion. A total of 710 patients with patellar dislocations were identified; 420 patients were excluded based on health records, primarily due to secondary dislocation ($n = 318$), miscoding ($n = 87$), insufficient information ($n = 9$), or foreign residency ($n = 6$). Of the remaining 290 patients with primary patellar dislocation, 176 were excluded because no X-ray was available ($n = 82$) or the lateral radiograph was too oblique for assessment ($n = 94$). The final cohort consisted of 114 patients (Figure 1).

FIGURE 1 Flow chart of patient inclusion and exclusion.



Measurements

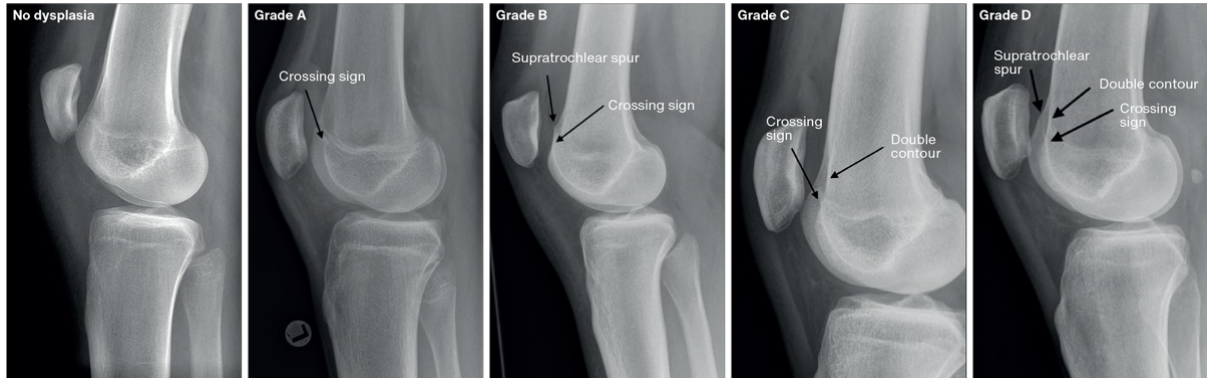
All X-rays were obtained as part of routine ED evaluation following acute patellar dislocation. Standard anteroposterior and lateral knee X-rays were acquired according to local protocols. These protocols do not include fluoroscopic guidance or dedicated patellofemoral positioning to ensure strict rotational alignment. Consequently, greater variation in knee rotation may occur compared with standardised patellofemoral imaging protocols used in specialised outpatient settings.

The measurements were performed by four raters at different stages of training: one medical student, two residents and one consultant in knee arthroscopy. Each performed two measurements ≥ 14 days apart, and the measurements were conducted independently by the raters. The raters were aware that their measurements would be compared with those of the other raters. All measurements were assessed in the Sectra Web Picture Archiving and Communication System (version 10-02-005).

Trochlear dysplasia was assessed by the Dejour classification [10], Anterior translation [11] and Trochlear Depth [12]. The Dejour classification [10] was assessed on lateral examination of the X-ray and categorised into Grade A (crossing sign), Grade B (A + spur), Grade C (A + double contour) and Grade D (A + B + C) (Figure 2). Anterior translation and Trochlear depth were dichotomised as dysplasia present or absent to reflect their potential

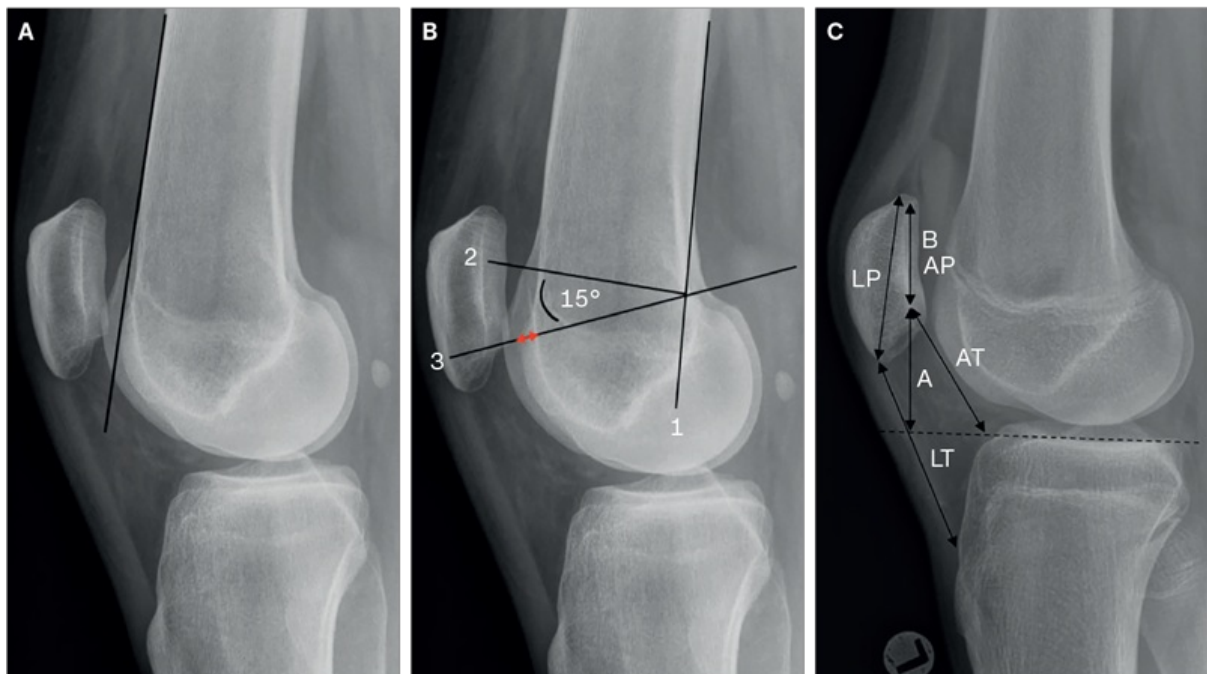
clinical use as screening measurements in the ED.

FIGURE 2 Dejour classification for trochlear dysplasia.



Anterior translation [11] was assessed by depicting a line along the anterior femur. If the trochlea crossed that line, anterior translation [11] was positive (Figure 3 A).

FIGURE 3 A. Anterior translation. A line along the anterior femur; dysplasia is present if the trochlear groove crosses the line. **B.** Trochlear depth. A line (1) along the posterior femur, then placing a line (2) perpendicular to line 1 that runs to the anterior-proximal part of the trochlea. Then a final line (3) is 15 degrees inferior, and measures trochlear depth (red double arrow). Dysplasia is present if trochlear depth < 4 mm. **C.** Patella height measurements. Insall-Salvati: LP:LT; Blackburne-Peel: B:A; Caton-Deschamps: AP:AT.



Trochlear depth [12] was evaluated by placing a line (1) along the posterior femur, then placing a line (2) perpendicular to line 1, running to the anterior-proximal part of the trochlea. A final line (3) was then placed 15 degrees inferior to line 2, and trochlear depth was measured along line 3 (Figure 3 B). Trochlear dysplasia was defined as a depth of less than four millimetres.

Patella height was assessed by Insall-Salvati [13], Blackburne-Peel [6] and Caton-Deschamps [7]. The Insall-

Salvati [13] ratio was measured as the length of the patellar tendon divided by the length of the patella (Figure 3 C). The Blackburne-Peel index [6] was measured as the distance from a horizontal line drawn at the level of the proximal tibial articular surface to the inferior aspect of the articular surface of the patella divided by the length of the articular surface of the patella (Figure 3 C).

The Caton-Deschamps index was measured as the distance from the inferior border of the patellar articular surface to the anterior edge of the tibial plateau divided by the length of the articular surface of the patella (Figure 3 C).

Patella alta was defined as greater than 1.2 for the Insall-Salvati ratio and Caton-Deschamps index, and greater than 1.0 for the Blackburne-Peel index. Patella baja was defined as less than 0.8 for the Insall-Salvati ratio, less than 0.5 for the Blackburne-Peel index [6] and less than 0.6 for the Caton-Deschamps index [7].

Statistics

EpiData software was used for data entry and management.

For statistical purposes, the Dejour classification was grouped as low (Grade A) or high (Grade B-D), since there were very few patients in grades B to D. A sample size calculation performed using an expected kappa of 0.6, precision ± 0.15 , a proportion of outcome 0.5, a confidence level of 95% and an expected drop-out rate of zero per cent yielded a minimum of 110 participants [14].

Statistical analyses were performed using unweighted kappa statistics for both intra- and inter-rater measurements, as all measurements were categorised: Dejour classification as grade A versus B-D, anterior translation positive/negative, trochlear dysplasia by depth positive/negative and all patella alta measurements as yes/no. The dichotomisation was performed for clinical relevance. However, dichotomising continuous variables may reduce statistical precision and lead to information loss, which should be considered when interpreting the findings.

For inter-rater analyses, first measurements were used. For descriptive purposes, the proportion of pairwise inter-rater comparisons with $\kappa \leq 0.40$ was calculated for each measurement method. Guidelines proposed by Landis & Koch [15] were used to categorise intra- and inter-rater Kappa agreement: slight agreement (0-0.20), fair agreement (0.21-0.40), moderate agreement (0.41-0.60), substantial agreement (0.61-0.80) and almost perfect agreement (0.81-0.99). A kappa value > 0.60 was considered to indicate significant reproducibility. All statistical analyses were conducted using STAT/IC14.2 (Stata Corp., College Station, TX, USA).

Trial registration: not relevant.

Results

Trochlear dysplasia

The intra-rater reliability for the Dejour classification ranged from 0.26 to 0.53; -0.01 to 0.70 for anterior translation; and from 0.29 to 0.85 for trochlear depth (Table 1). The combined kappa rating was 0.15 for the Dejour classification, 0.23 for anterior translation and 0.26 for trochlear depth. For the medical student, the intra-rater measurements ranged from 0.26 to 0.85; -0.01 to 0.74 for the residents, and 0.3 to 0.46 for the consultant.

TABLE 1 Intra- and inter-raters (Grade A vs grade B-D).

	MS	R1	R2	CK
<i>For Dejour classification</i>				
MS	0.26	0.11	0.39	0.05
R1	-	0.39	0.18	0.15
R2	-	-	0.53	0.13
CK	-	-	-	0.30
<i>For anterior translation</i>				
MS	0.62	-0.04	0.60	0.47
R1	-	-0.01	-0.05	-0.08
R2	-	-	0.70	0.44
CK	-	-	-	0.46
<i>For trochlear depth</i>				
MS	0.85	0.18	0.70	0.23
R1	-	0.29	0.14	0.13
R2	-	-	0.74	0.28
CK	-	-	-	0.46
<i>For the Caton-Deschamps index</i>				
MS	0.12	0.10	0.58	0.15
R1	-	0.37	0.10	0.36
R2	-	-	0.70	0.26
CK	-	-	-	0.32
<i>For the Insall-Salvati index</i>				
MS	0.85	0.69	0.63	0.70
R1	-	0.63	0.59	0.59
R2	-	-	0.76	0.64
CK	-	-	-	0.78
<i>For the Blackburn-Peel index</i>				
MS	0.24	0.28	0.16	0.30
R1	-	0.48	0.07	0.33
R2	-	-	0.57	0.08
CK	-	-	-	0.51

CK = consultant in knee arthroscopy; MS = medical student;
R1/R2 = resident 1 and resident 2.

Inter-rater reliability for the Dejour classification ranged from 0.05 to 0.39; for anterior translation, from -0.08 to 0.60; and for trochlear depth, from 0.13 to 0.70 (Table 1). Among the six pairwise inter-rater comparisons, 100% had κ values ≤ 0.40 for the Dejour classification, compared with 50% for anterior translation and 83% for trochlear depth.

Patella alta

The intra-rater reliability for the Insall-Salvati index ranged from 0.63 to 0.85; for the Caton-Deschamps index, from 0.12 to 0.70; and for the Blackburn-Peel index, from 0.24 to 0.57 (Table 1). The combined kappa ratings

were 0.66 for Insall-Salvati, 0.24 for Caton-Deschamps and 0.19 for Blackburn-Peel.

For the medical student, intra-rater measurements for patella alta ranged from 0.12 to 0.85; for residents, 0.37 to 0.76; and for the consultant, 0.32 to 0.78. The inter-rater reliability for the Caton-Deschamps index ranged from 0.10 to 0.58; for the Insall-Salvati index, from 0.59 to 0.70; and for the Blackburn-Peel index, from 0.07 to 0.33 (Table 1). In summary, among the six pairwise inter-rater comparisons, 83% had κ values ≤ 0.40 for the Caton-Deschamps index and 100% for the Blackburn-Peel index. None of the comparisons for the Insall-Salvati index had κ values ≤ 0.40 .

Discussion

In the present study, only the Insall-Salvati index demonstrated acceptable reproducibility when applied to routine ED X-rays. The reduced reproducibility observed for the remaining measurements likely reflects the challenges of applying them to X-rays with variable rotational malalignment, rather than inherent shortcomings of the measurement methods themselves. Routine ED X-rays are not optimised for rotational alignment, and variations in projection may alter anatomical landmarks, particularly in measurements of trochlear morphology.

These findings are consistent with prior literature.

Verhulst et al. reported good intra- and inter-rater reliability for the Insall-Salvati ratio across conventional radiography, CT and MRI, with inter-rater ICC values of 0.80, 0.75 and 0.78, respectively, whereas the other indices demonstrated inferior agreement [5]. Similarly, Seil et al. found that although all patellar height indices could be calculated from X-rays, their agreement varied widely and classification could differ depending on the method used [16]. In a study by Konrads et al., the Caton-Deschamps-derived index demonstrated moderate inter-rater agreement in the context of knee arthroplasty [7].

Measurements of trochlear dysplasia on lateral X-rays have also been criticised for their limited reliability. The Dejour classification depends on the presence of subtle X-ray features such as the crossing sign, supratrochlear spur and double contour, which are all subject to interpretation and variability in X-ray technique [2, 10]. This is reflected in our data, which showed poor agreement for the Dejour classification and related parameters, echoing findings from Diederichs et al. and Jibri et al., who both emphasised the role of MRI over X-rays for accurate assessment of trochlear morphology [8, 12].

Increased rater seniority was not consistently associated with improved agreement. Instead, measurement complexity and the constraints of routine ED X-rays, where optimal rotational alignment is rarely achieved, likely contributed to variability. Consequently, methodological and imaging-related limitations, rather than clinical expertise per se, appear to drive measurement inconsistency, highlighting the need for assessment methods based on clearly defined and reproducible anatomical landmarks.

Our findings suggest that the Insall-Salvati ratio may be a more reproducible option when applied to routine ED X-rays. Conversely, measurements such as the Blackburn-Peel, Caton-Deschamps, and Dejour classification should be interpreted with caution. Where uncertainty remains, careful interpretation and consideration of X-ray limitations are essential, particularly when images deviate from true lateral positioning [8, 17].

This study has several limitations. X-rays were obtained as part of routine ED care without standardised rotational alignment, and the degree of rotational malrotation was not quantified. Although reflecting clinical practice, it introduces variability that may affect measurement reliability. The evaluated measurements were originally developed for true lateral X-rays, and their application to non-true lateral projections represents a deviation from their intended use. Finally, the retrospective single-centre design may limit generalisability.

Conclusions

When applied to routine ED knee X-rays, the Insall-Salvati index demonstrated the highest and most consistent reliability. In contrast, the Blackburne-Peel index, Caton-Deschamps index and trochlear dysplasia measures showed poor reproducibility, reflecting the challenges of applying X-ray indices to images with variable rotational malalignment. These findings emphasise the importance of adopting an appropriate X-ray technique and of careful interpretation when using such measurements in clinical practice.

Correspondence *Misk Ghassan Farik Muhammad*. E-mail: misk.gfm@gmail.com

Accepted 14 August 2026

Published 15 October 2026

Conflicts of interest none. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. These are available together with the article at ugeskriftet.dk/dmj

References can be found with the article at ugeskriftet.dk/dmj

Cite this as Dan Med J 2026;73(11):A11250912

doi 10.61409/A11250912

Open Access under Creative Commons License [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)

REFERENCES

1. Fithian DC, Paxton EW, Stone ML, et al. Epidemiology and natural history of acute patellar dislocation. *Am J Sports Med*. 2004;32(5):1114-1121. <https://doi.org/10.1177/0363546503260788>
2. Bollier M, Fulkerson JP. The role of trochlear dysplasia in patellofemoral instability. *J Am Acad Orthop Surg*. 2011;19(1):8-16. <https://doi.org/10.5435/00124635-201101000-00002>
3. Neyret P, Robinson AHN, Le Coultre B, et al. Patellar tendon length-the factor in patellar instability? *Knee*. 2002;9(1):3-6. [https://doi.org/10.1016/S0968-0160\(01\)00136-3](https://doi.org/10.1016/S0968-0160(01)00136-3)
4. Ward SR, Terk MR, Powers CM. Patella alta: association with patellofemoral alignment and changes in contact area during weight bearing. *J Bone Joint Surg Am*. 2007;89(8):1749-1755. <https://doi.org/10.2106/JBJS.F.00508>
5. Verhulst FV, van Sambeek JDP, Olthuis GS, et al. Patellar height measurements: Insall-Salvati ratio is most reliable method. *Knee Surg Sports Traumatol Arthrosc*. 2020;28(3):869-875. <https://doi.org/10.1007/s00167-019-05531-1>
6. Blackburne JS, Peel TE. A new method of measuring patellar height. *J Bone Joint Surg Br*. 1977;59B(2):241-242. <https://doi.org/10.1302/0301-620X.59B2.873986>
7. Konrads C, Grosse LC, Ahmad SS, et al. Reliability of a Caton-Deschamps-derived patella height index for knee arthroplasty. *Int Orthop*. 2021;45(8):2001-2005. <https://doi.org/10.1007/s00264-020-04931-0>
8. Diederichs G, Issever AS, Scheffler S. MR imaging of patellar instability: injury patterns and assessment of risk factors. *Radiographics*. 2010;30(4):961-981. <https://doi.org/10.1148/rg.304095755>
9. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64(1):96-106. <https://doi.org/10.1016/j.jclinepi.2010.03.002>
10. Kazley JM, Banerjee S. Classifications in brief: the Dejour classification of trochlear dysplasia. *Clin Orthop Relat Res*. 2019;477(10):2380-2386. <https://doi.org/10.1097/CORR.0000000000000886>
11. Batailler C, Neyret P. Trochlear dysplasia: imaging and treatment options. *EFORT Open Rev*. 2018;3(5):240-247. <https://doi.org/10.1302/2058-5241.3.170058>
12. Jibri Z, Jamieson P, Rakhra KS, et al. Patellar maltracking: an update on the diagnosis and treatment strategies. *Insights Imaging*. 2019;10(1):65. <https://doi.org/10.1186/s13244-019-0755-1>
13. Insall J, Salvati E. Patella position in the normal knee joint. *Radiology*. 1971;101(1):101-104.

<https://doi.org/10.1148/101.1.101>

14. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. *Stat Methods Med Res.* 2004;13(4):251-271. <https://doi.org/10.1191/0962280204sm365ra>
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174. <https://doi.org/10.2307/2529310>
16. Seil R, Müller B, Georg T, et al. Reliability and interobserver variability in radiological patellar height ratios. *Knee Surg Sports Traumatol Arthrosc.* 2000;8(4):231-236. <https://doi.org/10.1007/s001670000121>
17. Paul RW, Brutico JM, Wright ML, et al. Strong agreement between magnetic resonance imaging and radiographs for Caton-Deschamps index in patients with patellofemoral instability. *Arthrosc Sports Med Rehabil.* 2021;3(6):e1621-e1628. <https://doi.org/10.1016/j.asmr.2021.07.017>