

The reliability of four widely used patellar height ratios

Dennis van Duijvenbode¹ · Michel Stavenuiter¹ · Bart Burger¹ · Cees van Dijke² ·
Jacco Spermon³ · Marco Hoozemans⁴

Received: 15 May 2015 / Accepted: 24 June 2015 / Published online: 16 July 2015
© SICOT aisbl 2015

Abstract

Purpose The objective of this study was to evaluate the inter-observer reliability and the intra-observer reliability of four patellar height ratios: Insall-Salvati (IS), modified Insall-Salvati (MIS), Blackburne-Peel (BP) and Caton-Deschamps (CD).

Methods The patellar height ratios were assessed by four independent examiners using weight-bearing lateral knee radiographs in 30° flexion. Intra-class correlation coefficients and Fleiss' kappa's were determined.

Results The inter-observer reliability was excellent for the IS and moderate for the other ratios. When the ratio values were categorized, the inter-observer reliability was strong for the IS, moderate for the MIS and BP, and poor for the CD. The intra-observer reliability was excellent for the IS, MIS and CD, and strong for the BP. When the ratio values were categorized, the intra-observer reliability was strong for the IS and MIS, and moderate for the other ratios.

Conclusion Although the IS showed best reliability, we advise to use the MIS as it showed the second best reliability but is, according to the literature, associated with better validity.

Keywords Patellar height · Patella alta · Reliability · Variability · Reproducibility

Introduction and rationale

Measurement of patellar height is important in evaluating knee conditions and can be important in planning treatment [1–4]. Patellar height can be assessed on radiographs using numerous methods and corresponding ratios [5–8], however none is perceived as the gold standard [5]. These methods either relate the position of the patella to the femur (direct assessment) or relate the position of the patella to the tibia (indirect assessment). The direct assessment methods are not widely applied because they proved to be too complex to apply [5]. The methods that use an indirect assessment, such as the Insall-Salvati (IS) [9], modified Insall-Salvati (MIS) [10], Blackburne-Peel (BP) [11] and Caton-Deschamps (CD) [12] are the most widely applied radiographic assessment methods.

With respect to the widely used indirect assessment methods, there is no consensus about which method should be routinely used to assess patellar height. Few studies analysed and compared inter-observer reliability or intra-observer reliability of the most widely applied patellar height ratios [6–8, 13–17]. One of these studies included patients with total knee arthroplasty [15], one study [14] included skeletally immature patients, and in one study [16] it was not clear if the entire study cohort was skeletally mature. Both knee arthroplasty and skeletal immaturity can affect the landmarks used by the ratios. Furthermore, one of the studies [13] did not report whether they used a digital system for evaluating the radiographs, which is customary in current clinical practice. In two studies [15, 17] it was unclear if the measurements were performed using a digital system. It also was unclear what the status was of the assessors (e.g., orthopaedic

✉ Dennis van Duijvenbode
d.c.duijvenbode@mca.nl

¹ Department of Orthopaedic Surgery, Medisch Centrum Alkmaar/
Centre for Orthopaedic Research Alkmaar (CORAL), Alkmaar, the
Netherlands
² Department of Radiology, Medisch Centrum Alkmaar, Alkmaar, the
Netherlands
³ Department of Radiology, Ziekenhuis Gelderse Vallei, Ede, the
Netherlands
⁴ Faculty of Human Movement Sciences, MOVE Research Institute
Amsterdam, Amsterdam, Netherlands

surgeon, radiologist or resident) in half of the studies [7, 13, 14, 17]. In total, two studies [7, 8] reported reliability of patellar height ratios for mature native knees, using a digital system, for evaluating routine radiographs. Of these studies only one study [7] compared all four ratios, but this study proposed to use their newly introduced direct method to measure patellar height. None of these studies reported the inter-observer and intra-observer reliability of the classification of the patellar height ratios into ‘patella baja’, ‘normal patellar height’ and ‘patella alta’ of the four patellar height ratios.

The objective of this study was to evaluate the inter-observer and intra-observer reliability of four patellar height ratios (based on IS, MIS, BP, CD) using a digital system for evaluating standard radiographs of mature native knees. The secondary objective was to evaluate the inter-observer and intra-observer reliability of the classification of the patellar height ratios into ‘patella baja’, ‘normal patellar height’ and ‘patella alta’ of the four patellar height ratios.

Material and methods

A retrospective study was conducted. Institutional review board approval was obtained. Given the retrospective design, an informed patients’ consent was not deemed necessary.

Patient selection

We collected radiographs from patients aged 15–40 years, registered in the hospital information system for visiting our outpatient clinic for patellofemoral pain syndrome between January 2011 and July 2012. Inclusion criteria were: availability of standard weight bearing lateral knee radiograph in minimal 30° flexion, closed physis, availability of an MRI report compiled within one year before or after the radiograph, excluding knee deformities, meniscus injury, anterior cruciate ligament injury, collateral ligament injury or other traumatic injury. Technically poor radiographs, i.e., the with landmarks obscured, were excluded.

Image acquisition and assessment

Standard weight bearing lateral knee radiographs in approximately 30° flexion were used to measure the patellar height ratios. A goniometer was not used when obtaining standard weight bearing lateral knee radiographs.

The patellar height was assessed in millimetres on the lateral knee radiographs using the built-in ruler of our picture archiving and registration system (IMPAX 6.4, Agfa HealthCare N.V.). The diagnostic images were independently interpreted in an identical sequence by four experts (two radiologists specialized in the musculoskeletal system and two orthopaedic surgeons specialized in knee surgery) who were blinded to the patients’ data and the other examiners results. The

measurements were assessed independently at two different moments with an interval of at least six weeks. The sequence of the radiographs was different for the second assessment.

Figure 1 shows the four methods for assessing patellar height [9–12]. Patellar length (A) and patellar tendon length (B) were measured to obtain the IS ratio (B/A). For the MIS ratio (D/C), the length of the patellar articular surface (C) and the distance from the inferior edge of the patellar articular surface to the insertion of the patellar tendon (D) were assessed. The length of the patellar articular surface (E) and the perpendicular distance from the inferior edge of the patellar articular surface to the line from the tibial plateau surface (F) were used for the BP ratio (F/E). The length of the patellar articular surface (G) and the distance from the inferior edge of the patellar articular surface to the anterosuperior angle of the tibial plateau surface (H) were measured for the CD ratio (H/G). Depending on the ratio used, patellar height was classified into ‘patella baja’, ‘normal patellar height’ and ‘patella alta’ according to the original [9–12] cutoff points (Table 1).

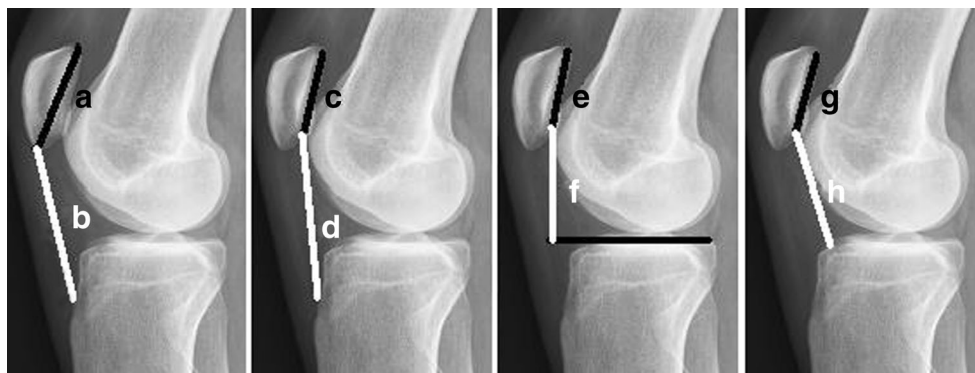
Statistical analysis

Descriptive statistics were calculated according to standard methods, including frequencies, means, medians, standard deviations and ranges when appropriate.

The inter-observer reliability and the intra-observer reliability of the four ratios were determined using intraclass correlation coefficients (ICC) for single measures [18]. ICCs and 95 % confidence intervals were based on a two-way random model utilizing absolute agreement using IBM SPSS Statistics 20.0 (IBM Corporation, Armonk, NY, USA). Inter-observer reliability was calculated with use of the data from the first acquisition session only. Intra-observer reliability was calculated with use of the data from both acquisitions for each observer. The individual ICCs were averaged subsequently. Scores were interpreted on the basis of the values suggested by Shrout and Fleiss [18] with a score of 0–0.4 indicating poor reliability, 0.4–0.75 indicating moderate reliability and a score of more than 0.75 indicating excellent reliability.

Based on the four different ratios, Fleiss’ kappa’s and confidence intervals were calculated for the categories “patella baja”, “normal patellar height” or “patella alta”. Fleiss’ kappa and confidence intervals were calculated using <http://www.statstodo.com> (Fleiss’s kappa from rating scores). Inter-observer reliability was calculated with use of the data from the first acquisition session only. Intra-observer reliability was calculated with use of the data from both acquisitions for each observer. The individual Fleiss’ kappa’s were averaged subsequently. A kappa of <0.2 is considered poor agreement, 0.21–0.4 fair agreement, 0.41–0.6 moderate agreement, 0.61–0.8 strong agreement and more than 0.8 near complete agreement [19].

Fig. 1 From left to right: Insall-Salvati (b/a), modified Insall-Salvati (d/c), Blackburne-Peel (f/e), Caton-Deschamps (h/g)



Results

Between January 2011 and July 2012, 269 patients visited our outpatient clinic for patellofemoral pain syndrome. Of the 269 patients 45 patients were included in this study. Mean age at the time of the radiograph was 24 (SD 8) years and 73 % was female. When the measurements were averaged for all observers and both assessment moments (360 measurements), the average ratios were 1.1 (SD 0.2) for the IS, 1.7 (SD 0.2) for the MIS, 0.9 (SD 0.2) for the BP and 1.1 (SD 0.2) for the CD. All average ratios indicate normal patellar height.

The inter-observer reliability of the first measurement for the four ratios is presented in Table 2. The IS showed excellent reliability, the other ratios showed moderate reliability. When the measurements were categorized in “patella baja”, “normal patellar height” and “patella alta”, the IS showed a strong inter-observer reliability, the MIS and BP had moderate inter-observer reliability, and the CD showed poor inter-observer reliability.

The intra-observer variability for the four ratios is presented in Table 3. The IS, MIS and CD had excellent reliability, and the BP had strong reliability. When the ratio values were categorized the IS and MIS had strong intra-observer reliability, while the BP and the CD showed moderate intra-observer reliability.

Discussion

In this study the IS showed to have excellent reliability and has the best reliability of all the ratios. The MIS showed to have moderate to excellent reliability and has the second best

reliability, except for the inter-observer ICC where it had the third best score. There was no evident difference between the intra-observer reliability of the orthopaedic surgeons and the radiologists. The high reliability of the IS might be explained by the patellar and tibial landmarks used by the ratio. The superior and inferior patellar pole, as used by IS, are easy to identify on plain radiographs. However, the shape of the inferior pole can be affected by Sinding-Larsen-Johansson disease. The inferior ridge of the patellar articular surface as used by the MIS, BP and CD is less easy to determine than the inferior patellar pole [5]. It is not always clear where the articular surface starts and ends distally. It is especially hard to identify the distal articular border when the radiograph is not perfectly lateral. In our opinion, radiographs are frequently not perfectly lateral in daily practice and therefore the patellar height assessment methods should allow a certain degree of imperfection in the quality of the radiographs when assessing patellar height. The tibial tuberosity, as used by IS and MIS, is easily determined, because the patellar tendon is usually visible on contemporary radiographs. However, the shape of the tibial tuberosity and the patellar morphology can be affected in Osgood Schlatter’s disease [20]. The horizontal line projected anteriorly from the tibial plateau in the BP ratio (Fig. 1) is less easily determined. When the radiograph is not obtained with a perfectly lateral view, the medial and lateral plateau will not be parallel on the picture and it will be difficult to determine where to place the line. The antero-superior articular tibial margin in the CD ratio is not always easily determined and not found in 10 % [21]. Again, the identification of this

Table 1 The cutoff points of the four patellar height ratios

Ratio ^a	Patella baja	Normal patella	Patella alta
Insall-Salvati (B/A)	<0.8	0.8–1.2	>1.2
Modified Insall-Salvati (D/C)	n.a.	n.a.	>2.0
Blackburne-Peel (F/E)	<0.8	0.8–1.0	>1.0
Caton-Deschamps (H/G)	<0.6	0.6–1.3	>1.3

^a See Fig. 1

Table 2 The inter-observer reliability of patellar height ratios

Ratio	ICC	Fleiss’s Kappa ^a
Insall-Salvati	0.87 (0.81–0.92)	0.69 (0.58–0.81)
Modified Insall-Salvati	0.60 (0.40–0.075)	0.47 (0.35–0.58)
Blackburne-Peel	0.74 (0.62–0.084)	0.45 (0.37–0.54)
Caton-Deschamps	0.55 (0.33–0.72)	0.19 (0.07–0.31)

ICC intra-class correlation coefficient (95 % confidence interval)

^a Fleiss’ kappa (95 % confidence interval) based on the categorized ratios: “patella baja”, “normal patellar height” and “patella alta”

Table 3 The intra-observer reliability of patellar height ratios

	Observer 1		Observer 2		Observer 3		Observer 4		Mean ^a	
	ICC	FK	ICC	FK	ICC ^a	FK	ICC	FK	ICC	FK
IS	0.94 (0.90–0.97)	0.85 (0.56–1.14)	0.72 (0.75–0.84)	0.63 (0.35–0.91)	0.94 (0.90–0.67)	0.74 (0.44–1.03)	0.97 (0.94–0.99)	0.68 (0.38–0.97)	0.90 ±0.12	0.72 ±0.10
MIS	0.86 (0.75–0.92)	1.00 (0.71–1.29)	0.72 (0.52–0.84)	0.33 (0.03–0.62)	0.87 (0.77–0.93)	0.79 (0.50–1.01)	0.76 (0.60–0.86)	0.63 (0.34–0.92)	0.80 ±0.08	0.69 ±0.28
BP	0.61 (0.40–0.77)	0.65 (0.44–0.87)	0.58 (0.34–0.74)	0.62 (0.39–0.85)	0.80 (0.67–0.89)	0.54 (0.32–0.76)	0.82 (0.69–0.90)	0.51 (0.29–0.73)	0.70 ±0.13	0.58 ±0.07
CD	0.92 (0.85–0.95)	1.00 (0.71–1.29)	0.55 (0.30–0.72)	0.32 (0.03–0.62)	0.93 (0.87–0.96)	0.36 (0.07–0.66)	0.67 (0.47–0.81)	0.48 (0.19–0.77)	0.76 ±0.19	0.54 ±0.31

IS insall-salvati, MIS modified insall-salvati, BP blackburne-peel, CD caton-deschamps

ICC Intra Class Correlation Coefficient (95 % confidence interval)

FK Fleiss's Kappa (95 % confidence interval) based on the categorized ratio's ("patella baja", "normal patellar height" and "patella alta")

^a Mean of individual scores±standard deviation

landmark is highly dependent on the quality of the radiograph. However, this landmark is not always clearly identifiable due to bone morphology. Considering the arguments above, we believe that the anatomical landmarks of the IS ratio are most easily determined in a patella without patellofemoral pathology.

Two studies reported on the reliability of patellar height ratios for evaluating routine radiographs of mature native knees using a digital system [7, 8]. In accordance with the present study, Nizić et al. [7] reported the highest reliability for the IS for both the inter-observer reliability and the intra-observer reliability when assessing the IS, MIS, BP, and CD. However, the reliability of the IS, MIS, BP and CD reported in this study differed only marginally, and according to these findings the reliability of all ratios could be considered as excellent. Portner et al. [8] also reported the highest inter-observer and the intra-observer reliability for the IS when compared to the BP and CD. The MIS was not assessed by Portner et al. [8]. Chareancholvanich et al. [6] also reported the reliability of the four most widely used patellar height ratios. However, they used standardized, non-weight-bearing radiographs. In accordance with the present study, Chareancholvanich et al. [6] reported the highest inter-observer reliability for the IS and the second best reliability for the MIS.

Ideally, patellar height ratios use a distal femoral reference point to represent the true patellofemoral articulation. However, all ratios evaluated in this study related the patellar height to a reference point on the tibia. Anatomic variability in the position of the landmarks used to calculate the patellar height when using a tibial reference point can compromise the validity of the ratio. Grelsamer et al. [21] defined three types of sagittal morphologies for the patella. These types represent different proportions of the inferior pole related to the articular surface. Both lines used by IS (Fig. 1) are influenced by these proportions. The distribution of the morphologic types differs

between healthy subjects and patients with patellofemoral symptoms [21]. Grelsamer et al. [21] showed that patients with patellofemoral symptoms had a significantly higher percentage of type II patella's, i.e., these patients showed a patella with a short distance between the cranial and the inferior articulating surface. However, the patella was normal when measured from the cranial to the inferior pole. Because the IS uses the inferior pole as an anatomic landmark, the IS does not give a good representation of the position of the articulating surface of the patella relative to the trochlea in "long nosed" patella's and thus has poor validity, especially in patients with patellofemoral symptoms. In our opinion, patellar height is mainly assessed in patients with patellofemoral symptoms and, therefore, should not be used despite its excellent reliability. Furthermore, the inclination of the tibia plateau varies considerably among patients [22] and this influences the "horizontal line". Minor errors in determining the horizontal line over the tibial plateau or difference in slope of the "horizontal line" will lead to significant differences for the point where the horizontal line and the vertical line from the patella will meet. This will influence the length of the vertical line (Fig. 1, line F) and thus the ratio. Even if the BP would be highly reliable, the variation of the tibia plateau inclination makes the validity questionable. The validity of patellar height ratios is also dependent on the use of proper reference values. These reference values should be based on asymptomatic knees because knee symptoms can be due to abnormalities of the patellofemoral relationship [5]. Only the CD had reference values that were based on a healthy population [5, 9–12]. New reference values could be extracted from Chareancholvanich et al. [6], because they only included non-symptomatic volunteers and calculated normal values for the IS, MIS, BP, and CD. However, this was in a non-weight-bearing situation. It is also important that the ratios maintain accuracy in varying degrees of knee flexion.

Chareancholvanich et al. [6] showed that knee angles between 0 and 60° of flexion did not influence the measurement of the four patellar height ratios in non-weight-bearing in healthy subjects. Yiannakopoulos et al. [23] showed that weight-bearing radiographs in 30° resulted in higher patellar height in ratios in healthy subjects when compared with non-weight-bearing radiographs. Furthermore, biomechanical analysis of weight-bearing healthy subjects showed that the patellar tendon length increased significantly between 0 and 30° of flexion, and did not further increase after 30° [24]. In patients with patella alta, the effect of weight-bearing and knee flexion might even be greater due to a potentially longer tendon which might be flabby in the extended knee or a potentially more stretchable patellar tendon. We would advise to obtain the lateral knee radiograph in some degree of flexion and in weight-bearing position to guarantee a tensed patellar tendon (i.e., be at full length).

The main limitation of this study was the moderate sample size [25]. However, we do not expect that the inclusion of more patients would have influenced our results. A strength of this study was the assessment of the radiographs by both orthopaedic surgeons and musculoskeletal radiologists, because these clinicians would be the assessors of the radiographs in daily practice. Furthermore, we consider the use of imaging obtained as part of the normal diagnostic evaluation as a strength. These radiographs were of varying quality (e.g., not always perfectly lateral). In our opinion, the use of these radiographs give a good representation of the imaging a clinician will encounter in daily practice, therefore the reliability obtained using these radiographs is more valid than the reliability obtained from perfectly lateral radiographs made for study purposes only.

In conclusion, when comparing the most widely applied ratios to assess patellar height, the IS showed to be most reliable. However, although it showed the second best reliability, we advise to use the MIS because it has—according to the literature—better validity.

Conflict of interest The authors declare that they have no competing interests.

References

- AL-Sayyad MJ, Cameron JC (2002) Functional outcome after tibial tubercle transfer for the painful patella alta. *Clin Orthop Relat Res* 396:152–162
- Monk AP, Doll HA, Gibbons CL, Ostlere S, Beard DJ, Gill HS, Murray DW (2011) The patho-anatomy of patellofemoral subluxation. *J Bone Joint Surg (Br)* 93(10):1341–1347
- Neyret P, Robinson AH, Le CB, Lapra C, Chambat P (2002) Patellar tendon length—the factor in patellar instability? *Knee* 9(1): 3–6
- Stefanik JJ, Zhu Y, Zumwalt AC, Gross KD, Clancy M, Lynch JA, Frey Law LA, Lewis CE, Roemer FW, Powers CM, Guermazi A, Felson DT (2010) Association between patella alta and the prevalence and worsening of structural features of patellofemoral joint osteoarthritis: the multicenter osteoarthritis study. *Arthritis Care Res* 62:1258–1265
- Phillips CL, Silver DAT, Schranz PJ, Mandalia V (2010) The measurement of patellar height: a review of the methods of imaging. *J Bone Joint Surg (Br)* 92:1045–1053
- Chareancholvanich K, Narkbunnam R (2012) Novel method of measuring patellar height ratio using a distal femoral reference point. *Int Orthop* 36:749–753
- Nizic D, Pervan M, Kovacevic B (2014) A new reference line in diagnosing a high-riding patella on routine digital lateral radiographs of the knee. *Skelet Radiol* 43(8):1129–37
- Portner O, Pakzad H (2011) The evaluation of patellar height: a simple method. *J Bone Joint Surg Am* 93:73–80
- Insall J, Salvati E (1971) Patella position in the normal knee joint. *Radiology* 101(1):101–104
- Grelsamer RP, Meadows S (1992) The modified Insall-Salvati ratio for assessment of patellar height. *Clin Orthop Relat Res* 282:170–176
- Blackburne JS, Peel TE (1977) A new method of measuring patellar height. *J Bone Joint Surg (Br)* 59:241–242
- Caton J, Deschamps G, Chambat P, Lerat JL, Dejour H (1982) Patella infera. A propos of 128 cases. *Rev Chir Orthop Reparatrice Appar Mot* 68:317–325
- Berg EE, Mason SL, Lucas MJ (1996) Patellar height ratios. A comparison of four measurement methods. *Am J Sports Med* 24: 218–221
- Lee P, Chalian M, Carrino J, Eng J, Chhabra A (2012) Multimodality correlations of patellar height measurement on X-ray, CT, and MRI. *Skelet Radiol* 41:1309–14
- Rogers BA, Thornton-Bott P, Cannon SR, Briggs TWR (2006) Interobserver variation in the measurement of patellar height after total knee arthroplasty. *J Bone Joint Surg (Br)* 88:484–488
- Smith TO, Cogan A, Patel S, Shakokani M, Toms AP, Donell ST (2013) The intra- and inter-rater reliability of X-ray radiological measurements for patellar instability. *Knee* 20:133–138
- Seil R, Muller B, Georg T, Kohn D, Rupp S (2000) Reliability and interobserver variability in radiological patellar height ratios. *Knee Surg Sports Traumatol Arthrosc* 8:231–236
- Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86:420–428
- Fleiss JL (1981) *Statistical methods for rates and proportions*. Wiley, New York
- Visuri T, Pihlajamäki HK, Mattila VM, Kiuru M (2007) Elongated patellae at the final stage of Osgood-Schlatter disease: a radiographic study. *Knee* 14:198–203
- Grelsamer RP, Proctor CS, Bazos AN (1994) Evaluation of patellar shape in the sagittal plane. A clinical analysis. *Am J Sports Med* 22: 61–66
- Matsuda S, Miura H, Nagamine R, Urabe K, Ikenoue T, Okazaki K, Iwamoto Y (1999) Posterior tibial slope in the normal and varus knee. *Am J Knee Surg* 12:165–168
- Yiannakopoulos CK, Mataragas E, Antonogiannakis E (2008) The effect of quadriceps contraction during weight-bearing on four patellar height indices. *J Bone Joint Surg (Br)* 90:870–873
- Defrate LE, Nha KW, Papannagari R, Moses JM, Gill TJ, Li G (2007) The biomechanical function of the patellar tendon during in-vivo weight-bearing flexion. *J Biomech* 40:1716–1722
- de Vet HCW, Terwee CB, Mokkink LB, Knol DL (2011) *Reliability*. In: *Measurement in medicine*. Cambridge University Press, Cambridge, pp 96–147