Qualitative visual assessment of the J-sign demonstrates high inter-rater reliability

Nicholas Walla, Toren Moore, Sarah Harangody, Sean Fitzpatrick, David C. Flanigan, Robert A. Duerr, Robert Siston, Robert A. Magnussen

Sports Medicine Research Institute, The Ohio State University, Columbus, OH, 43202, USA

ARTICLE INFO

Keywords:
J-sign
Patellar instability
Patellar tracking
Inter-rater reliability
Intra-rater reliability

ABSTRACT

Objectives: To assess the inter- and intra-rater reliability of the classification of the J-sign as “large” versus “small or none” as compared to another two-level system (“present” versus “absent”) and a three-level system (“large,” “small,” or “none”) and to identify anatomical and patient factors associated with the presence of a large J-sign.

Methods: Forty patients (40 knees) with recurrent patellar instability were prospectively enrolled and recorded on video actively extending their knee while seated. Four raters classified patellar tracking on two separate occasions using three systems: 1) two groups: J-sign versus no J-sign; 2) three groups: large J-sign, small J-sign, or no J-sign; and 3) two groups: large J-sign versus small or no J-sign. The intra- and inter-rater reliability of each system was assessed using kappa statistics. Anatomical (trochlear dysplasia, tibial tubercle-trochlear groove (TT-TG) distance, patellar height) and patient (Beighton score) factors as well as Knee Injury and Osteoarthritis Outcome Score (KOOS) subscales were compared between patients with a large J-sign and patients with a small or no J-sign.

Results: Inter- and intra-rater reliability were found to be highest with the two-level classification system of a large J-sign versus a small or no J-sign (inter-rater kappa = 0.76, intra-rater kappa = 0.75). Patients with a large J-sign had more severe trochlear dysplasia as assessed with the sulcus angle (p = 0.042) and were more likely to have a tight lateral retinaculum (p = 0.032) and an elevated Beighton score (p = 0.009). No significant differences in KOOS subscales were noted based on the presence of a large J-sign versus a small J-sign or no J-sign.

Conclusion: Qualitative visual assessment of patellar tracking with the J-sign demonstrates substantial inter- and intra-rater reliability, particularly when utilizing a two-group classification system to identify knees with a large J-sign. Patients with a large J-sign demonstrate an increased incidence of a tight lateral retinaculum, generalized ligamentous laxity, and trochlear dysplasia.

Level of evidence: Level III - cross-sectional study.

What is already known?

- Recurrent patellar instability is a debilitating clinical problem that results in decreased knee function, increased pain, and a long-term risk of osteoarthritis.
- Physical examination is a useful tool for the identification of patients with patellar instability and may correlate with disease severity.

What are the new findings?

- Visual assessment of patellar tracking with the J-sign demonstrates substantial inter- and intra-rater reliability.
- Patients with a large J-sign demonstrate an increased incidence of a tight lateral retinaculum, generalized ligamentous laxity, and trochlear dysplasia.
1. Introduction

Recurrent patellar instability is a debilitating clinical problem that results in decreased knee function, increased pain, and a long-term risk of osteoarthritis. There is an incidence of lateral patellar dislocations in 5–7 per 100,000 people; however, this incidence is as high as 29 per 100,000 in teenagers [1]. Numerous anatomical factors have been identified that predispose to patellar instability, drive treatment decisions, and may influence the outcomes of surgical treatments [2].

In addition to obtaining patient history and meticulously evaluating imaging [3,4], one important component of the physical examination includes a visual inspection of patellar tracking, including the presence of a J-sign [5]. A J-sign refers to a pathological lateral shift of the patella as the knee moves from flexion into full extension.

Surgical treatment with medial patellofemoral ligament (MPFL) reconstruction is the mainstay of treatment for recurrent patellar instability [6]. This procedure can be performed either in isolation or in combination with bony realignment procedures. While specific indications for the addition of an osteotomy to an MPFL reconstruction remain controversial, a number of recent studies have shown that isolated MPFL reconstruction can be successful in a large proportion of patients with recurrent instability [7,8]. However, the presence of a prominent J-sign has been identified by several authors as one of the predictors of failure in isolated MPFL reconstruction [9–11]. This finding could lead to the consideration of more invasive bony realignment procedures in this population of patients.

The use of the presence or absence of a J-sign as a surgical indication for bony procedures is problematic for several reasons. First, the inter-rater reliability of the J-sign is quite variable. One previous study on this subject noted only moderate inter-rater agreement (kappa = 0.53) on the presence of J-tracking [4]. Additionally, there is no agreement on the best method to classify the severity or magnitude of the J-sign. Donnell et al. described a detailed classification of patellar tracking patterns [12], while others may describe patients as having a “large” J-sign or a “subtle” J-sign (e.g. an obvious lateral shifting of the patella versus a faint lateral deviation). While differences in the severity of the J-sign may influence its ease of identification and impact on outcomes, these differences have not been formally studied.

The goals of this study are to determine the inter- and intra-rater reliability of visual assessment of patellar tracking, evaluate anatomical factors that may be associated with the presence of a large J-sign, and evaluate knee function as assessed by Knee Injury and Osteoarthritis Outcome Score (KOOS) subscales based on the presence or absence of a large J-sign. We hypothesise that the greatest inter- and intra-rater reliability of qualitative assessment of the J-sign will be achieved through a two-level classification of the J-sign as “large” versus “small” or none as compared to another two-level classification system (J-sign “present” versus “absent”) and a three-level classification system (J-sign “large,” “small,” or “none”). Further, we hypothesise that trochlear dysplasia, patellar alta, and elevated tibial tubercle-trochlear groove (TT-TG) distance are associated with the presence of a large J-sign and that patients with a large J-sign will demonstrate poorer knee function as assessed through the KOOS subscales.

2. Methods

This study was approved by our Institutional Review Board. Informed consent was not required. Forty patients (40 knees) with documented recurrent patellar dislocations and no prior knee surgery other than arthroscopy without lateral release were identified and prospectively enrolled in the study. Each patient completed the KOOS assessment, which is a 42-item questionnaire involving 5 separately scored subscales: pain, function in activities of daily living (ADL), function in sport and recreation (Sport/Rec), and knee-related quality of life (QOL) [13]. Patient demographics including age, sex, and body mass index (BMI), were recorded. Plain lateral radiographs and knee magnetic resonance imaging (MRI) were obtained in all patients. Plain films were reviewed by a single surgeon to assess patellar height (Caton-Deschamps index [CDI]) [14], the presence of a crossing sign and/or supra-trochlear spur, and trochlear boss height [2]. MRIs were reviewed, and TT-TG distance, sulcus angle, trochlear depth, patellar tilt, and patellar height (patello-trochlear index (PTI)) were assessed [15,16]. Patellar tracking was then recorded on video using a smartphone (Apple iPhone, Apple Inc., Cupertino, CA) from an anterior viewpoint as the knee was flexed and extended three times with the patient in the seated position. A physical examination was performed to assess whether there was a tight lateral retinaculum (defined as the inability to evert the patellar above the level of the horizon with the knee fully extended), and a Beighton score was calculated for each patient [17].

Four raters (two board-certified orthopaedic sports medicine surgeons and two sports medicine fellows) then watched the videos in random order and classified the patellar tracking according to three systems developed for this study: 1) two groups: J-sign present versus absent; 2) three groups: large J-sign, small J-sign, or none; and 3) two groups: large J-sign versus small J-sign or none. Each of the four raters reviewed the videos 6 weeks later in a new random order and classified patellar tracking again. The senior author discussed the proposed rating system with each rater prior to each rating session. No further specific instructions or examples were given.

Inter- and intra-rater reliability were calculated using kappa statistics. Reliability was assessed as poor (kappa <0.20), fair (kappa: 0.21–0.40), moderate (kappa: 0.41–0.60), substantial (kappa: 0.61–0.80), or near perfect (kappa: 0.81–0.99).

Patients were classified as having a large J-sign versus a small or absent J-sign (classification system number three as defined above) for evaluation of potential factors contributing to the presence of a large J-sign. For this portion of the study, patients were considered to have a large J-sign if they were identified as such by at least three of the four reviewers. Factors potentially contributing to patellar tracking including CDI, crossing sign, supratrochlear spur, boss height, TT-TG distance, sulcus angle, trochlear depth, tilt, patella trochlear index (PTI), tight lateral retinaculum, and Beighton score, as well as KOOS subscales, were compared between the groups using unpaired t-tests (continuous variables) and Fisher’s exact tests (categorical variables). Statistically significant differences were defined as p < 0.05.

3. Results

A total of 40 patients enrolled in the study, including 32 females (80%) and 8 males (20%). The median age at enrolment was 20 years (inter-quartile range: 17–24.3 years), and the median BMI was 26.4 kg/m² (inter-quartile range: 23.9–30.4 kg/m²).

The inter-rater reliability amongst the 4 reviewers was found to be substantial using all three classification systems: 1. J-sign present: yes or no with a kappa value of 0.64; 2. J-sign: large, small or none with a kappa value of 0.61; 3. J-sign: large versus small or none with a kappa value of 0.76 (Table 1). The most agreement was demonstrated with the third classification system: large J-sign versus small or no J-sign.

The intra-rater reliability amongst the 4 reviewers was also found to be substantial using the three classification systems: 1. J-sign present: yes or no, with a mean kappa value of 0.72; 2. J-sign: Large, small, or none

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Inter-rater reliability – 4 raters.</th>
</tr>
</thead>
<tbody>
<tr>
<td>J-sign: 2 categories: present versus absent</td>
<td>Kappa: 0.64</td>
</tr>
<tr>
<td>J-sign: 3 categories: large, small, none</td>
<td>Kappa: 0.61</td>
</tr>
<tr>
<td>J-sign: 2 categories: large versus small or none</td>
<td>Kappa: 0.76</td>
</tr>
</tbody>
</table>
with a mean kappa value of 0.65. J-sign: large versus small, or none with a mean kappa value of 0.75 (Table 2). The highest reliability was again noted in the third classification system: large J-sign versus small or no J-sign.

When comparing imaging and physical examination findings with J-signs, patients with large J-signs were more likely to have a tight lateral retinaculum and an elevated Beighton score than patients with no J-sign or a small J-sign (Table 3). The large J-sign group demonstrated an increased risk of trochlear dysplasia, as demonstrated by an increased sulcus angle ($p = 0.042$). Representative images from a subject rated as having a large j-sign by all reviewers and a subject rated as having no J-sign by all reviewers are shown (Fig. 1).

No significant differences in KOOS scores were noted based on the presence of a large J-sign versus a small J-sign or no J-sign (Table 4).

### 4. Discussion

This study demonstrates substantial inter- and intra-rater reliability of the visual assessment of the patella tracking through each of the classification systems used. The highest reliability was obtained using a two-level classification system of a large J-sign versus a small J-sign or no J-sign.

Other studies have reported mixed findings regarding the reliability of visual assessment of patellar tracking. Best et al. had 32 orthopaedic surgeons visually assess patellar tracking through quantification of the number of quadrants the patella translated laterally during range of motion (ROM) [18]. They defined maltracking as at least two quadrants of lateral translation and noted moderate inter-rater reliability (kappa = 0.45) for identification of maltracking with this standard. Inter-rater reliability was also assessed in a quantitative fashion, comparing the surgeons’ grade of lateral patellar translation (0–4 quadrants) against the same assessment using 4D computed tomography (CT) imaging, and moderate inter-rater reliability (kappa = 0.42) was again noted. The authors concluded that inter-rater reliability was inadequate using their methods of visual assessment of patellar tracking. Hijmestra et al. further highlighted the inadequacies of the inter-rater reliability of the J-sign, specifically when using the Donell and Quadrant classification schemes (inter-rater reliability mean kappa 0.49 and 0.51, respectively) [19].

Smith et al. evaluated the reliability of 5 orthopaedic surgeons’ assessments of 18 different physical examination tests in 10 knees with patellar instability [4]. Three of the 18 exams displayed a moderate level of inter-observer reliability, including the J-sign (kappa = 0.53). They defined a positive J-sign as an exaggerated lateral to medial translation of the patella in the trochlear groove as the knee ranges from extension into flexion. This assessment is similar to the two-level grading system that the current study found to have the highest inter-rater reliability (large J-sign versus small J-sign or no J-sign).

Due to a desire to more objectively quantify the J-sign, some authors have attempted to identify correlations between the relative position of the patella and trochlea on axial imaging and the presence of a J-sign on physical examination. Beckert et al. investigated patellar position relative to the trochlea using the modified lateral patellar edge measurement...
[20]. They assessed the modified lateral patellar edge measurement of the knee at 30 degrees of flexion and at full extension with an MRI and compared these values to clinical measurements. The average difference between clinical and MRI J-sign measurements was 4.3 mm. They found a relative underestimation of lateral patellar translation at 30 degrees of flexion clinically versus MRI but concluded that there was no statistically significant difference between the clinical and MRI measurements (p = 0.26). Xue et al. investigated whether several anatomical measurements on axial CT scans were correlated with a positive J-sign [21]. They concluded that increased lateral patellar translation on CT scan, as quantified by the bisect offset index (BOI), is correlated with the presence of a positive J-sign on clinical examination.

Recognizing the limitations of attempting to correlate measurements from static imaging studies with the dynamic nature of the J-sign, work has been published to attempt to better characterize patellar tracking using dynamic imaging techniques. Tanaka et al. investigated the use of dynamic kinetic computed tomography (DKCT) in 76 knees and devised a quantitative method to grade lateral tracking through dynamic range of motion (ROM) [22]. They found that they were able to accurately measure BOI and characterise patellar tracking patterns and severity using this methodology.

Because a positive J-sign has been associated with an independent risk factor in isolated MPFL reconstruction failure [9–11], a key question is which underlying anatomical factors are associated with the presence of a J-sign. The current study identified a relationship between the presence of trochlear dysplasia (as quantified by an increased sulcus angle) and a large J-sign but was likely underpowered to detect all potential anatomical relationships. Prior work has demonstrated correlations with patella alta as well as lower extremity rotational abnormalities and the presence of a large J-sign [11].

An interesting finding of this study is that both generalised ligamentous laxity (as defined by a Beighton score greater than 5) and a tight lateral retinaculum were more commonly seen in patients with a large J-sign. These findings suggest that soft tissue factors contribute to the effects of bony anatomy on patellar tracking. While prior work evaluating the influence of soft tissue factors on patellar tracking is limited, some authors have investigated the influence of these factors on the risk of patellar instability. Charles et al. compared patellar tilt on MRI between a control group and a group with recurrent patellar instability and found that patellar tilt was the most accurate measurement in delineating recurrent patellar instability knees versus controls [23]. Biyani et al. evaluated 7 paediatric patients with recurrent patellar instability with MRI and found that the strongest predictor of lateral patellar tracking was lateral trochlear inclination [24]. Elias et al. similarly identified lateral trochlear inclination as an anatomical risk factor for lateral patellar maltracking [25]. Simulation studies have confirmed this relationship and suggest that isolated MPFL reconstruction likely does not normalise patellar tracking in these situations [26,27].

Another interesting finding of this study was the lack of any correlation between the visual assessment of patellar tracking and patient-reported symptoms and knee function as assessed by the KOOS. This finding could suggest that patellar tracking is truly not related to patient perceptions of knee function or may reflect the inability of this study to identify such a relationship due to being underpowered. Further limitations of the tracking assessment or patient-reported outcome tool could lead to this finding.

This study did have a number of limitations. First, the study was small (40 knees) and underpowered to identify all possible correlations between anatomic measures and a large J-sign. Further, the correlation of patellar tracking with anatomical measures is limited by the anatomical measurements that were obtained. In particular, rotational assessments of the lower limbs were not performed. Prior work has suggested that this factor may be predictive of patellar tracking [11]. A further limitation was the use of a video assessment rather than an in-person assessment of tracking. There may be limitations to the video in terms of differences in viewing angle, shadowing, patient body habitus, and other factors that may not be applicable to an in-person examination. In addition, the reviewers (two board-certified orthopaedic sports medicine surgeons and two sports medicine fellows) may not be reflective of the orthopaedic community as a whole given their high interest in this area, potentially limiting the broad applicability of the findings.

5. Conclusions

Qualitative visual assessment of patellar tracking with the J-sign demonstrates substantial inter- and intra-rater reliability, particularly when utilising a two-group classification system to identify knees with a large J-sign. Patients with a large J-sign demonstrate an increased incidence of a tight lateral retinaculum, generalised ligamentous laxity, and trochlear dysplasia.

Ethics approval

This study was approved by the Biomedical Institutional Review Board of The Ohio State University.

Funding

This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interests

The authors declare the following financial interests/personal relationships, which may be considered potential competing interests: David Flanigan reports a relationship with Verticle Corporation that includes consulting or advisory services. David Flanigan reports a relationship with Smith and Nephew Inc. that includes consulting, advisory services and speaking and lecture fees. David Flanigan reports a relationship with Medical Device Business Services that includes consulting or advisory services.

References


N. Walla et al. Journal of ISAKOS xxx (xxxx) xxx


