

# Assessment of Patellofemoral Relationships Using Kinematic MRI: Comparison Between Qualitative and Quantitative Methods

Samuel R. Ward, PT,<sup>1</sup> Frank G. Shellock, PhD,<sup>2,3</sup> Michael R. Terk, MD,<sup>2,4</sup> Gretchen B. Salsich, PhD, PT,<sup>5</sup> and Christopher M. Powers, PhD, PT<sup>1\*</sup>

**Purpose:** To compare the level of agreement between quantitative and qualitative methods in determining patellofemoral relationships, since controversy exists regarding the use of quantitative vs. qualitative criteria to interpret images of the patellofemoral joint (PFJ) obtained using kinematic magnetic resonance (MR) imaging.

**Materials and Methods:** One hundred twenty mid-patellar axial plane images obtained using kinematic MR imaging from fifteen subjects were randomly selected for analysis. MR images represented various knee flexion angles ranging from 0 to 60 degrees. Quantitative analysis (bisect offset and patellar tilt angle) was performed by two examiners using a computer-assisted software program. Based on data from previously published literature, MR images were characterized as demonstrating normal, medial, or lateral patellar subluxation, and/or normal, medial, or lateral tilt. Using similar categories, two different examiners experienced in reading MR images of the PFJ then applied qualitative criteria to the same images.

**Results:** The average agreement between the quantitative and qualitative assessments of horizontal patellar displacement and patellar tilt ranged from poor to moderate (Kappa coefficient values of 0.27 and 0.45, respectively). Quantitative and qualitative techniques demonstrated acceptable intra- and inter-observer reliability.

**Conclusion:** These findings indicate that the use of quantitative criteria does not compare well to qualitative criteria in the analysis of kinematic MR images of the PFJ. One

explanation for this discrepancy relates to the fundamental difference between the techniques. That is, quantitative measurements are based on the use of osseous landmarks, while the qualitative assessments tend to rely on a description of patellofemoral relationships based on joint surfaces.

**Key Words:** magnetic resonance imaging; patellofemoral joint; subluxation; kinematic; alignment

**J. Magn. Reson. Imaging 2002;16:69–74.**

© 2002 Wiley-Liss, Inc.

THE USE OF STANDARD radiography for diagnostic evaluation of the patellofemoral joint (PFJ) has many recognized limitations and, as such, has questionable clinical value (1). In 1988, a kinematic magnetic resonance imaging (MRI) procedure was developed to provide diagnostic information pertaining to patellar alignment during the initial increments of joint flexion, when subtle position-related abnormalities are the most apparent (2). Since then, reports have indicated that kinematic MRI of the PFJ is a sensitive and useful technique for assessment and characterization of aberrant positions of the patella (1,3,4).

Interpretation of images obtained from kinematic MRI examinations may be accomplished either using qualitative or quantitative techniques (3–6). Qualitative assessments have been used in the vast majority of clinical reports utilizing kinematic MRI of the PFJ (3,5,7–9). Quantitative interpretation, however, is the preferred method for research purposes and typically includes the use of established PFJ indices such as the Merchant angle, lateral patellofemoral angle, and bisect offset (5,7,10). The preference for quantitative measures in experimental studies is fairly obvious, as the use of objective criteria minimizes experimenter bias associated with qualitative assessment, and provides numerical values for statistical analyses. Objective measurements are rarely used clinically, however, as they are time consuming, difficult to apply in cases of patella alta, or patellofemoral dysplasia, and often are difficult to interpret (5).

Despite the fact that qualitative and quantitative techniques are used for assessment of kinematic MRI of

<sup>1</sup>Musculoskeletal Biomechanics Research Laboratory, Department of Biokinesiology and Physical Therapy, University of Southern California, Los Angeles, California.

<sup>2</sup>Keck School of Medicine, University of Southern California, Los Angeles, California.

<sup>3</sup>Shellock R&D Services, Los Angeles, California.

<sup>4</sup>LAC/USC Imaging Science Center, Los Angeles, California.

<sup>5</sup>Department of Physical Therapy, Saint Louis University, St. Louis, MO.

\*Address reprint requests to: C.M.P., Department of Biokinesiology and Physical Therapy, 1540 East Alcazar Street, CHP 155, Los Angeles, California 90089-4500. E-mail: powers@hsc.usc.edu

Received November 14, 2001; Accepted March 19, 2002.

DOI 10.1002/jmri.10124

Published online in Wiley InterScience (www.interscience.wiley.com).

the PFJ, it is not known if both methods yield similar interpretations. In fact, evidence exists suggesting that the two methods may give different results. For example, using qualitative criteria from images obtained using kinematic MRI, Shellock, et al (6) reported 76% of PFJs examined had a correction or improvement in patellar subluxation after the application of a realignment brace. Utilizing the same realignment brace and kinematic MRI protocol, a recent study by Powers, et al (11) found no difference in patellar tracking in subjects with patellofemoral pain when objective measurements were employed. Although the results of these two studies may be related to differences in patient populations, it is also possible that the conflicting findings were the result of the contrasting methods used to assess the images (i.e., quantitative vs. qualitative criteria).

Given the potential discrepancy between qualitative assessment and quantitative measurement of PFJ relationships, the purpose of this investigation was to compare the level of agreement between these two methods. A secondary purpose of this study was to establish the inter- and intra-observer reliability of qualitative and quantitative assessment of PFJ alignment.

## MATERIALS AND METHODS

### MR Images

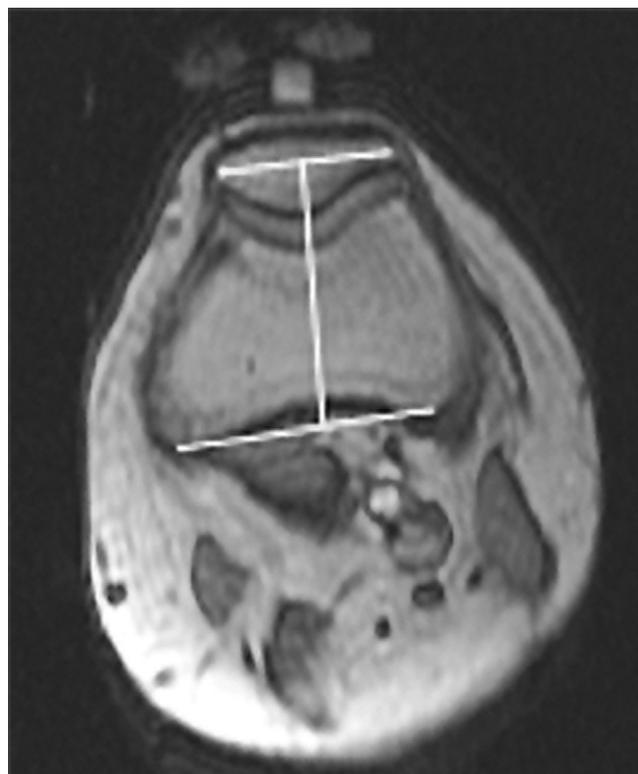
One hundred twenty mid-patellar, axial plane MR images of the PFJ were randomly selected from a pool of 546 images. Using previously described kinematic techniques (6,11), data were obtained from 15 subjects (eight male and seven female, mean age  $26.1 \pm 8.1$  years). Nine of these subjects had a diagnosis of patellofemoral pain and six subjects were pain-free. The selected images represented various knee flexion angles ( $0^\circ$  to  $60^\circ$ ) and were obtained during either weight-bearing knee flexion and extension (i.e., squatting against body weight) or non-weightbearing knee extension (i.e., extending the knee against gravity in the seated position) (12).

A 0.5-T vertically open MR system was used, allowing images to be acquired in the standing or seated positions (5). A fast multiplanar spoiled gradient-echo (FSPGR) pulse sequence was used to acquire axial images of the PFJ. Imaging parameters were: time to repeat (TR) 10.3 msec, echo time (TE) 2.7 msec,  $40^\circ$  flip angle,  $35 \times 18$  cm field of view (FOV),  $256 \times 128$  matrix size, number of excitations (NEX) 2, and 12 mm slice thickness. Fifteen images were acquired in 20 seconds (0.75 seconds per image).

### Procedures

Quantitative analysis of the selected images was made by two researchers experienced in measuring PFJ relationships. A custom macro written for NIH Image software (National Institute of Health, Bethesda, MD) was used for all measurements.

Medial and lateral patellar displacement was assessed using the bisect offset measurement as described by Brossmann, et al (7). The bisect offset index was measured by drawing a line connecting the poste-

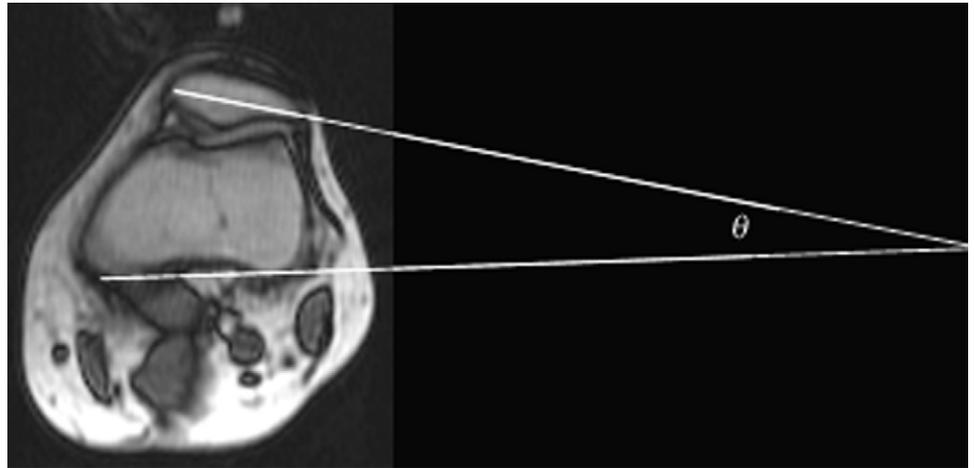


**Figure 1.** The line spanning the width of the patella is divided into two sections by the line perpendicular to the posterior femoral condyles and projecting through the deepest portion of the trochlear groove. Bisect off-set index is the percent of the patella that lies lateral to the deepest portion of the trochlear groove.

rior femoral condyles and projecting a line anteriorly through the deepest point (apex) of the trochlear groove. This line intersected the patellar width line, which connected the widest points of the patella (Fig. 1) (7). To obtain data when the trochlear groove was flattened, the perpendicular line was projected anteriorly from the bisection of the posterior femoral condylar line (13). The bisect offset was representative of the extent of the patella lateral to midline and was expressed as a percentage of total patellar width.

Medial/lateral patellar tilt was measured using a modification of the technique described by Powers, et al (13). The patellar tilt angle was reported as the angle formed by the lines joining the maximum width of the patella and the line joining the posterior femoral condyles (Fig. 2) (13). All tilt measurements were reported in degrees.

Qualitative assessment of the selected images were made by two investigators, each with over 10 years of clinical and research experience in reading axial plane MR images obtained using kinematic methods. These two examiners used subjective criteria previously described by Shellock, et al (5,6,14), and were asked to characterize each image with respect to horizontal patellar displacement (medial, normal, and lateral) and patellar tilt (medial, normal, and lateral). Briefly, normal patellar alignment was qualitatively defined as be-



**Figure 2.** The patellar tilt angle ( $\theta$ ) lies at the intersection of a line projecting through the maximum patellar width and a second line connecting the posterior femoral condyles.

ing present when the median ridge of the patella was positioned in the femoral trochlear groove, without transverse displacement of the medial or lateral facets. Medial or lateral subluxation of the patella was considered to be present when the median ridge of the patella was displaced medially or laterally relative to the femoral trochlea, and the medial or lateral facet was displaced relative to its respective anterior femoral condyle. Medial or lateral tilt of the patella was considered to be present when the anterior aspect of the patella was oriented medially or laterally. For purposes of establishing intra- and inter-observer reliability, both qualitative assessments and quantitative measures were made on 50 images on two different days approximately one week apart. All investigators were blinded to previous measurements and to identifying information on the MR images. In all cases, images were presented in random order.

### Data Analysis

Comparison of continuous numeric data (quantitative measurements) and categorical data (qualitative measures) required subsetting the numerical data into distinct categories. The establishment of these categories was based on previously published data obtained from healthy individuals using techniques similar to those employed in the current study (13).

With respect to medial/lateral displacement (Fig. 1), an image was considered “normal” if the bisect offset value fell between 0.44 and 0.64. Based on the work of Powers, et al (13), this range of patellar displacement represents the mean patellar displacement of pain-free individuals ( $0.54 \pm 2 \text{ SD } (0.05)$ ), and should encompass 95% of the healthy population. Therefore, a bisect offset value greater than 0.64 was considered indicative of lateral patellar displacement. A bisect offset value less than 0.44 was considered indicative of medial patellar displacement.

With respect to medial/lateral patellar tilt (Fig. 2), an image was considered “normal” if the patellar tilt angle was found to fall between  $-2.5$  (negative value indicating medial tilt) and  $13.5$  degrees. This range of tilt represents the mean patellar tilt angle of pain-free individ-

uals subjects ( $5.5 \text{ degrees} \pm 2 \text{ SD } (4.0 \text{ degrees})$  (13), and should encompass 95% of the this healthy subject population (15). Therefore, a patellar tilt angle value greater than  $13.5$  degrees was considered indicative of lateral tilt. A patellar tilt angle value less than  $-2.5$  degrees was considered indicative of medial tilt.

### Statistical Analyses

Interclass correlation coefficients (ICC) 2,1 (the standard equation number) were used to determine intra- and inter-observer reliability for the quantitative measurements (continuous numeric data). Kappa coefficients were used to determine intra- and inter-observer reliability for the qualitative assessments (categorical data).

Once the quantitative data were characterized categorically, the Kappa coefficient was used to assess the level of agreement between the qualitative and quantitative assessments of patellar alignment. Separate analyses were performed for horizontal patellar displacement and patellar tilt and were performed using SPSS statistical software (SPSS Inc., Chicago, Illinois).

## RESULTS

The intra-observer reliability of the two examiners performing the quantitative measurements was excellent for both horizontal patellar displacement (ICCs: 0.92 and 0.99; Table 1) and patellar tilt (ICCs: 0.99 and 0.96; Table 2). In addition, there was excellent agreement (inter-observer reliability) between the two quantitative examiners for both horizontal patellar displacement (ICC = 0.90; Table 1) and patellar tilt (ICC = 0.90; Table 2).

The intra-observer reliability of the two examiners performing the qualitative measures ranged from substantial to excellent for both horizontal patellar displacement (Kappa coefficients: 0.67 and 0.80; Table 1) and patellar tilt (Kappa coefficients: 0.80 and 0.94; Table 2). The inter-observer reliability of the qualitative measures was moderate for horizontal patellar displacement (Kappa = 0.54; Table 1) and patellar tilt (Kappa = 0.51; Table 2).

Table 1  
Interclass Correlation Coefficients and Kappa Coefficients for Medial/Lateral Patellar Displacement

	Quantitative examiner 1	Quantitative examiner 2	Qualitative examiner 3	Qualitative examiner 4
Quantitative examiner 1	ICC 0.99 <sup>a</sup>			
Quantitative examiner 2	ICC 0.90 <sup>b</sup>	ICC 0.92 <sup>a</sup>		
Qualitative examiner 3	K 0.45 <sup>d</sup>	K 0.29 <sup>d</sup>	K 0.67 <sup>a</sup>	
Qualitative examiner 4	K 0.37 <sup>d</sup>	K 0.28 <sup>d</sup>	K 0.54 <sup>c</sup>	K 0.80 <sup>a</sup>

<sup>a</sup>Intratester reliability.

<sup>b</sup>Intertester reliability (quantitative).

<sup>c</sup>Intertester reliability (qualitative).

<sup>d</sup>Agreement between quantitative and qualitative.

ICC = Interclass correlation coefficient; K = Kappa Coefficient.

The level of agreement between qualitative and quantitative assessment techniques was poor to moderate for horizontal patellar displacement (Kappas ranging from 0.28 to 0.45; Table 1). Similarly, the level of agreement between qualitative and quantitative assessment techniques was poor for patellar tilt (Kappas ranging from 0.27 to 0.31; Table 2).

## DISCUSSION

The primary result of this study is the poor agreement between quantitative and qualitative methods in characterizing PFJ relationships. This finding suggests that care must be made when comparing data obtained from these two methods. One likely explanation for this discrepancy may be related to differences in landmarks used by these two techniques. Quantitative measurements are typically made by identifying specific osseous landmarks, while qualitative assessments are made by visually describing the relationship of articulating joint surfaces relative to one another.

Another possible explanation for the poor agreement between the qualitative and quantitative methods is reflected by the fact that quantitative measurements may be more precise than qualitative assessments. For example, a measurement of 13.6 degrees of lateral patellar tilt would fall into the normal quantitative category but could easily be considered lateral based on qualitative criteria (Fig. 3). This suggests that although discrepancies may not occur when evaluating images with obvious malalignment, disagreement would more likely occur when evaluating "borderline images".

Most of the observed differences between the two methods were seen with qualitative assessment categorizing horizontal patellar displacement as being lateral while quantitative measurement categorized the patella as being normal (Table 3). The same trend was evident with patellar tilt. The fact that many of the disagreements were "borderline" images illustrates the difficulty in comparing data from these two methods (Fig. 2). Expanding the "normal" range would have improved agreement between the two assessment techniques; however, a data point that is 2 SD above the population mean typically represents a value that would be considered outside of the normal range (15).

The secondary finding of this study was that intra-observer reliability for both qualitative assessments and quantitative measurements were acceptable, with the intra-observer reliability of the quantitative measurements being slightly greater than that of the qualitative assessments. The higher degree of intra-observer reliability found in the quantitative measurements may be attributed to the precise landmark criteria, and the fact that measurements were made using a computer-aided system. Additionally, inter-observer reliability was higher for quantitative measurements. It stands to reason that quantitative measurements would be reliable as long as the precise anatomic criteria for the measurement is well understood by all examiners.

The moderate inter-observer reliability found with the qualitative examiners was somewhat unexpected; however, this result should be interpreted with caution. The qualitative examiners agreed on more than 70% of the images when categorizing horizontal patellar displace-

Table 2  
Interclass Correlation Coefficients and Kappa Coefficients for Medial/Lateral Patellar Tilt

	Quantitative examiner 1	Quantitative examiner 2	Qualitative examiner 3	Qualitative examiner 4
Quantitative examiner 1	ICC 0.99 <sup>a</sup>			
Quantitative examiner 2	ICC 0.90 <sup>b</sup>	ICC 0.96 <sup>a</sup>		
Qualitative examiner 3	K 0.31 <sup>d</sup>	K 0.33 <sup>d</sup>	K 0.94 <sup>a</sup>	
Qualitative examiner 4	K 0.27 <sup>d</sup>	K 0.34 <sup>d</sup>	K 0.50 <sup>c</sup>	K 0.80 <sup>a</sup>

<sup>a</sup>Intratester reliability.

<sup>b</sup>Intertester reliability (quantitative).

<sup>c</sup>Intertester reliability (qualitative).

<sup>d</sup>Agreement between quantitative and qualitative.

ICC = Interclass correlation coefficient; K = Kappa Coefficient.



**Figure 3.** Image illustrating borderline disagreement between assessment techniques. Quantitatively, this image was characterized as demonstrating normal tilt (8.0 degrees) while qualitatively, the image was characterized as demonstrating lateral tilt.

ment, and agreed in 80% of the cases when categorizing patellar tilt. These percentages were similar to the percentage of agreements found between quantitative examiners, which produced ICC values in excess of 0.90. The inconsistency between the reliability statistic and the percentage of agreements can be explained by the computational formula of the Kappa coefficient. The Kappa coefficient weights the number of agreements between examiners by a proportion of chance agreement. In this study, qualitative examiners categorized the majority of patellae as normal or laterally displaced, and normal or laterally tilted. Therefore, the possibility of chance agreement between qualitative examiners was relatively high, thus reducing the magnitude of the Kappa value. A sample with a more even distribution of

medial, normal, and lateral patellar displacement and tilt would have reduced the proportion of chance agreement, therefore improving the reliability of the qualitative examiners. Hence, the relatively low agreement between qualitative examiners compared to the quantitative examiners should be interpreted with caution.

Based on the results of this comparison study, care must be taken in making a determination as to which technique (qualitative or quantitative) should be used when assessing PFJ relationships. In order to make such a decision, both methods would have to be compared to a known “gold standard” with the underlying assumption being whichever method more closely approximated the “gold standard” would likely provide the best estimate of true patellofemoral alignment. To date, such a “gold standard” does not exist. However, an argument could be made that assessing PFJ relationships using osseous landmarks (as is commonly done with quantitative methods) may not be valid because cartilagenous surfaces, and not bony surfaces, are interacting at the joint. This premise is highlighted in a recent study by Staubli, et al (16), who found that patellofemoral relationships based on osseous landmarks do not coincide with patellofemoral relationships defined by cartilagenous surfaces. Unfortunately, articular cartilage is not readily visualized using fast gradient-echo pulse sequences typically utilized for kinematic MRI of the PFJ. Therefore, the use of such indices would not be possible using kinematic imaging techniques. Although most (if not all) quantitative methods to assess patellofemoral relationships use osseous landmarks, it is possible that quantitative indices that use cartilagenous interfaces could have a higher degree of agreement with the qualitative assessment methods used in the current study. Further research is necessary to test this hypothesis.

In conclusion, although quantitative and qualitative methods demonstrated moderate to excellent intra- and inter-observer reliability, there was poor agreement between qualitative and quantitative methods. The most common area of disagreement occurred when the patella was characterized as being laterally displaced or tilted based on qualitative assessment, as opposed to being characterized as “normal” based on quantitative measurements. While quantitative and qualitative methods are likely used in different circumstances (i.e., research vs. clinical), these results indicate that care

**Table 3**  
Example Measurement Comparison Matrix Between Quantitative and Qualitative Examiners

Tilt				Displacement			
Qualitative examiner #3				Qualitative examiner #3			
	Medial	Normal	Lateral		Medial	Normal	Lateral
Quantitative examiner #1				Quantitative examiner #1			
Medial	0 <sup>a</sup>	6	1	Medial	2 <sup>a</sup>	9	0
Normal	2	71 <sup>a</sup>	17 <sup>b</sup>	Normal	8	58 <sup>a</sup>	18 <sup>b</sup>
Lateral	0	6	15 <sup>a</sup>	Lateral	0	7	16 <sup>a</sup>

<sup>a</sup>Agreements between examiners.

<sup>b</sup>Category of most frequent disagreement.

must be taken when comparing data obtained from these two methods.

## REFERENCES

1. Brossmann J, Muhle C, Bull CC, et al. Cine MR imaging before and after realignment surgery for patellar tracking: comparison with axial radiographs. *Skeletal Radiol* 1995;24:191–296.
2. Shellock FG, Mink JH, Fox JM. Patellofemoral joint: kinematic MR imaging to assess tracking abnormalities. *Radiology* 1988;168:551–553.
3. Brown SM, Bradley WG. Kinematic magnetic resonance imaging of the knee. *Magn Reson Imaging Clin N Am* 1994;2:441–449.
4. Kujala UM, Osterman K, Kormanen M, Komu M, Schlenzka D. Patellar motion analyzed by magnetic resonance imaging. *Acta Orthop Scand* 1989;60:13–16.
5. Shellock FG. Kinematic MRI of the patellofemoral joint. In: Shellock FG, Powers CM, editors. *Kinematic MRI of the joints: functional anatomy, kinesiology, and clinical applications*. Boca Raton: CRC Press LLC; 2001. p 165–202.
6. Shellock FG, Mink JH, Deutsch AL, et al. Effect of a patellar realignment brace on patellofemoral relationships: evaluation with kinematic MR imaging. *J Magn Reson Imaging* 1994;4:590–594.
7. Brossmann J, Muhle C, Schroder C, et al. Patellar tracking patterns during active and passive knee extension: evaluation with motion-triggered cine MR imaging. *Radiology* 1993;187:205–212.
8. Shellock FG, Mink JH, Deutsch AL, Fox JM. Patellar tracking abnormalities: clinical experience with kinematic MR imaging in 130 patients. *Radiology* 1989;172:799–804.
9. Shellock FG, Mullen M, Stone K, Coleman M, Crues JV. Kinematic MRI evaluation of the effect of bracing on patellar positions: qualitative assessment using an extremity MR system. *J Athletic Training* 2000;35:44–49.
10. Powers CM, Shellock FG, Pfaff M. Quantification of patellar tracking using MRI. *J Magn Reson Imaging* 1998;8:724–732.
11. Powers CM, Shellock FG, Beering TV, Garrido DE, Goldbach RM, Molnar T. Effect of bracing on patellar kinematics in patients with patellofemoral joint pain. *Med Sci Sports Exerc* 1999;31:1714–1720.
12. Ward SR, Powers CM, Fredericson M, Guillet M, Cooper C, Shellock FG. The effect of bracing on patellar kinematics during weightbearing and non-weightbearing movements. *Med Sci Sports Exerc* 2001;33:S128.
13. Powers CM, Shellock FG, Pfaff M, Perry J. Quantification of patellar tracking using kinematic resonance imaging. *J Magn Reson Imaging* 1995;8:724–732.
14. Shellock FG, Stone KR, Crues JV. Development and clinical application of kinematic MRI of the patellofemoral joint using an extremity MR system. *Med Sci Sports Exerc* 1999;31:788–791.
15. Portney LG, Watkins MP. Concepts of measurement. In: Portney LG, Watkins MP, editors. *Foundations of clinical research: applications to practice*. Stamford: Appleton and Lange; 1993. p 39–47.
16. Staubli HU, Durrenmatt U, Porcellini B, Rausching W. Anatomy and surface geometry of the patellofemoral joint in the axial plane. *J Bone Joint Surg Br* 1999;81B:452–458.