

## Inter- and Intraobserver Reliability of Two-dimensional CT Scan for Total Knee Arthroplasty Component Malrotation

Beau Konigsberg MD, Ryan Hess MD,  
Curtis Hartman MD, Lynette Smith MS,  
Kevin L. Garvin MD

Published online: 28 June 2013  
© The Association of Bone and Joint Surgeons® 2013

### Abstract

**Background** Rotational malalignment of total knee arthroplasty (TKA) has been correlated with patellofemoral maltracking, knee instability, and stiffness. CT is the most accurate method to assess rotational alignment of prosthetic components after TKA, but inter- and intraobserver reliability of CT scans for this use has not been well documented.

**Questions/purposes** The objective of this study was to determine the inter- and intraobserver reliability and the repeatability of the measurement of TKA component rotation using two-dimensional CT scans.

**Methods** Fifty-two CT scans of TKAs being evaluated for revision surgery were measured by three different physicians. An orthopaedic resident and attending measured the

same scans twice (more than 2 weeks apart) and a musculoskeletal radiologist measured them once. To assess interobserver reliability, intraclass correlation coefficients (ICCs) with two-way mixed-effects analysis of variance models as well as 95% confidence intervals for each were done. The repeatability coefficient was calculated as well, which is defined as the difference in measurements that include 95% of the values. This indicates the magnitude of variability among measurements in the same scale, which in this study is degrees.

**Results** The interobserver ICC measurement for the femoral component was 0.386 (poor), and it was 0.670 (good) for the tibial component. The interobserver ICC for the combined rotation measurements was 0.617 (good). The intraobserver ICC for the femoral component was 0.606 (good), and it was 0.809 (very good) for the tibial component. The intraobserver ICC for combined rotation was 0.751 (good). The intraobserver repeatability coefficient for the femoral component was  $0.49^\circ$ ,  $10.64^\circ$  for the tibial component, and  $12.29^\circ$  for combined rotation.

**Conclusions** In this study, the inter- and intraobserver reliability, and the repeatability, of TKA component rotation were variable. This raises concern about whether CT scan is diagnostic in the assessment of component malrotation after TKA.

**Level of Evidence** Level IV, diagnostic study. See Guidelines for Authors for a complete description of levels of evidence.

---

One or more of the authors (CH, KLG) have received funding that is not related to the content of this manuscript. The funding source includes Biomet (Warsaw, IN, USA) (KLG) in the amount of USD 100,000 to USD 1,000,000 and Smith & Nephew (Memphis, TN, USA) (CH) in the amount of USD 10,000 to USD 100,000. The institution of the authors has received funding from TRAK Surgical Inc (Omaha, NE, USA), Smith & Nephew, Biomet, Vanguard (Valley Forge, PA, USA), Exponent (Menlo Park, CA, USA), Gruppo Bioimplanti (St Paul, MN, USA), and Renovis (Redlands, CA, USA). All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research* editors and board members are on file with the publication and can be viewed on request. Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

---

B. Konigsberg, R. Hess, C. Hartman, L. Smith,  
K. L. Garvin (✉)  
University of Nebraska Medical Center,  
Omaha, NE 68198-1080, USA  
e-mail: kgarvin@unmc.edu

### Introduction

The majority of patients undergoing TKA have excellent clinical results [14, 18, 19, 21]. A minority of patients may experience knee pain, stiffness, and/or patellar instability

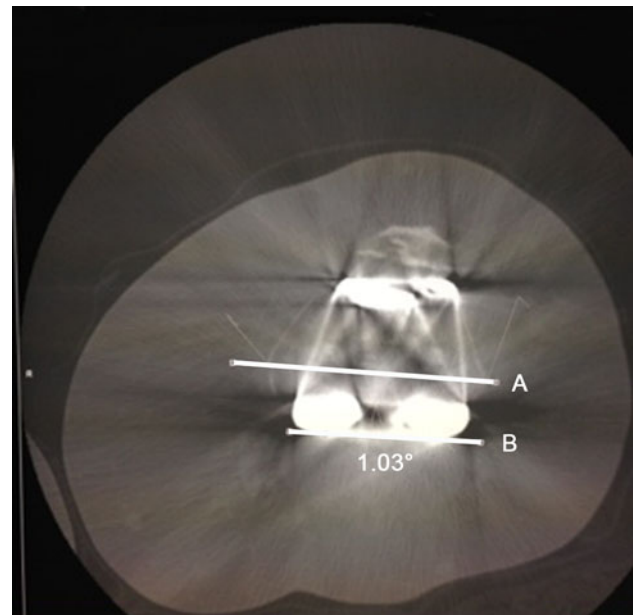
postoperatively [2, 3, 20]. Some authors have suggested that rotational malalignment of the femoral and/or tibial components causes these problems, and because of this belief, malrotation is a major cause of revision TKA [4, 21]. Others have written that rotational component alignment can cause rotational knee instability and even posterior dislocations of the knee after TKA [1, 17]. If revision surgery is proposed, the indication for surgery such as malrotation should be identified. Also, if poor outcomes are attributed to malrotation, it is important for surgeons to have a reliable and reproducible way to measure it to confirm the diagnosis before recommending revision surgery.

Before the wide adoption of CT, component rotation was measured clinically or on plain radiography [6, 10, 11, 13, 16]. Although these methods could demonstrate gross malalignment, it was difficult to quantify and probably not adequate to identify more subtle cases of malrotation. Jazwari et al. [15] demonstrated the accuracy of CT in measuring rotation of TKA components on cadavers. Berger et al. [4] described a CT protocol to measure component rotation in patients undergoing TKA and proposed that the amount of combined malrotation is directly correlated with the severity of patellofemoral complications. Specifically, Berger et al. describe that patients with lateral patella tracking and tilting have combined component excessive internal rotation of  $1^\circ$  to  $4^\circ$ . Those with patella subluxation have  $3^\circ$  to  $8^\circ$  of excessive internal rotation, patella dislocation  $7^\circ$  to  $16^\circ$ , and component failure at  $8^\circ$  to  $17^\circ$  [4]. However, to attribute these complications to component malrotation with any confidence, in particular if the result of these measurements might cause a surgeon to recommend revision arthroplasty, one must know the reliability of the measurement. Unfortunately, there is a lack of adequate intra- and interobserver reliability data in the published literature regarding two-dimensional (2-D) CT scan measurements of TKA components.

The primary purpose of our study, therefore, was to determine the inter- and intraobserver reliability of 2-D CT scan measurements of TKA component rotation. The second purpose was to determine the repeatability coefficient of this test, which is defined as the difference in measurements that include 95% of the values using 2 SDs from the mean. This measurement identifies the magnitude of variability by the same reviewer in degrees of rotation measured.

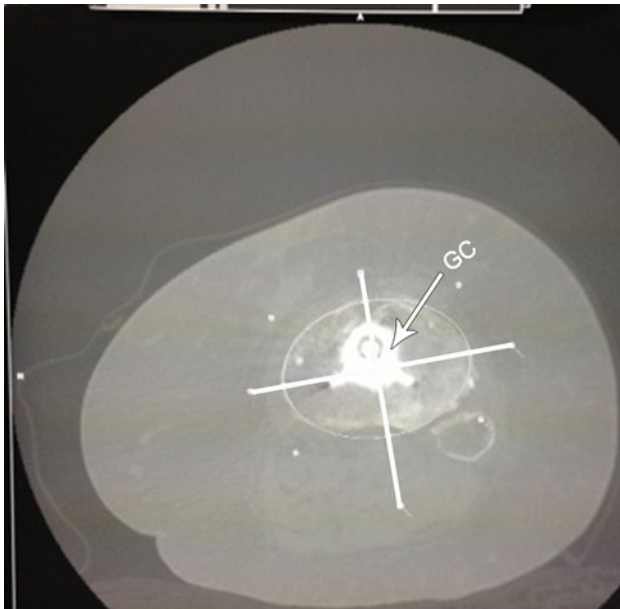
## Patients and Methods

Fifty patients (52 knees) had 2-D CT scans using the McKesson PACS system (McKesson Corporate

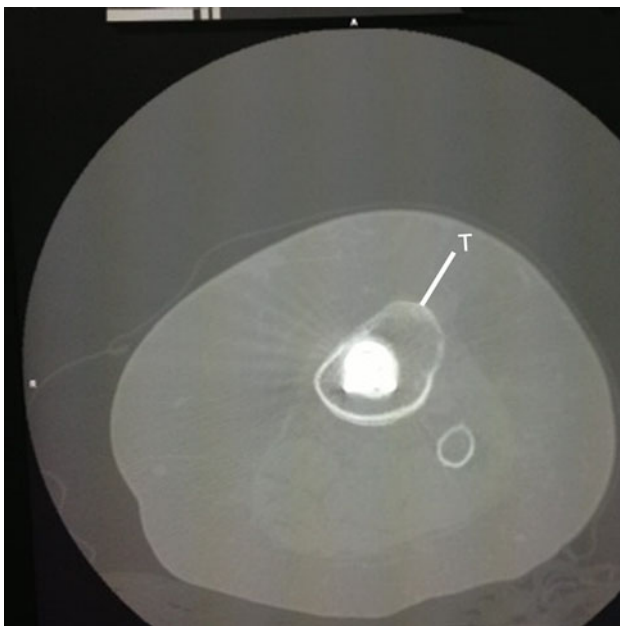


**Fig. 1** The CT scan is an image at the level of the femoral medial and lateral epicondyles. Line A represents the epicondylar axis of the femur and Line B represents the femoral component posterior condylar surface and the component rotational axis. In this patient, the femoral component is  $1.03^\circ$  internally rotated.

Headquarters, San Francisco, CA, USA). There were 22 men and 28 women. Average age at time of CT scan was 64 years with a range of 46 to 88 years. All patients were being evaluated for possible revision as the indication for ordering the CT scan. These patients were being evaluated for one or more of the following diagnoses: patellar subluxation and anterior knee pain, patellar instability, knee instability with suspected malalignment, and knee stiffness (ROM  $< 70^\circ$  of flexion) [9]. All 52 scans were measured according to the protocol Berger et al. described [4]. This protocol uses 1.5-mm thick axial images obtained at four locations of the knee: through the femoral epicondylar axis, through the tibial component tray, the proximal tibial plateau, and the tibial tubercle. The femoral component rotation is determined by the angle comparing the anatomic epicondylar axis with the posterior condylar surface of the component. Native rotation from the posterior condylar angle is  $0.3^\circ$  ( $\pm 1.2^\circ$ ) internal rotation for females and  $3.5^\circ$  ( $\pm 1.2^\circ$ ) internal rotation for males relative to the anatomic epicondylar axis [4] (Fig. 1). The tibial component rotation is first calculated by identifying the geometric center of the tibial plateau. Next, the lines for the tibial component axis and the tibial anatomic axis (the line from the geometric center to the center of the tibial tubercle) are drawn. The angle between the two lines is the rotation of the tibial component. Normal rotation using this method is  $18^\circ$  ( $\pm 2.5^\circ$ ) [4] (Figs. 2, 3, 4). The evaluators included an adult reconstruction surgeon (BK), an orthopaedic chief resident (RH), and a musculoskeletal radiologist

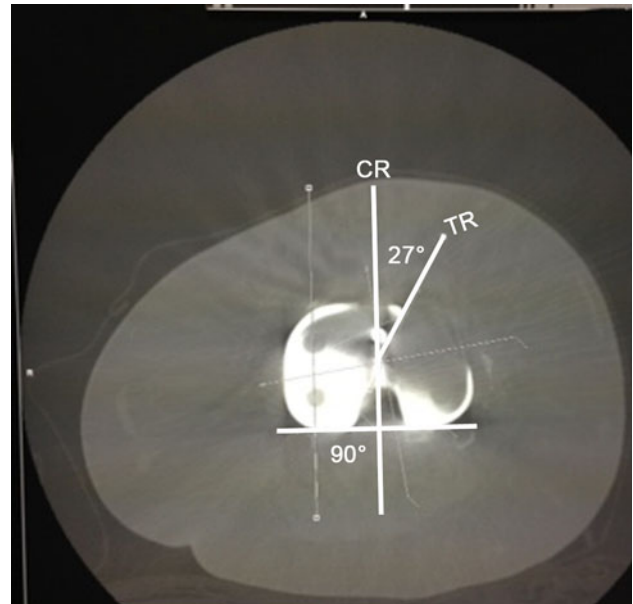


**Fig. 2** The intersection of the lines shown on the tibia represents the geometric center of the knee (GC).



**Fig. 3** The line is to the center of the tibial tubercle (T).

(MM) to determine interobserver reliability. The scans were measured twice, more than 2 weeks apart, by the orthopaedic resident and attending surgeon to determine intraobserver reliability. A sample size was conducted to ensure adequate power for the study. We determined that 50 CT scans read twice by the same observer will achieve 80% power to detect an intraclass correlation coefficient (ICC) of 0.78 compared with an ICC of 0.60 under the null hypothesis with a



**Fig. 4** Line “TR” represents the tibial anatomic axis for rotation and line “CR” represents the tibial component axis for rotation. The angle formed by the intersection of these lines represents the component rotation. In this patient, the tibial component is internally rotated 27°.

**Table 1.** Intraclass correlation coefficient

Value	Definition
0.81 to 1*	Very good
0.61 to 0.8	Good
0.41 to 0.6	Fair to moderate
0.40 and below	Poor

\* Perfect reliability (ICC value of 1).

significance level of 0.05. For interobserver reliability with 50 CT scans read by each of the reviewers, we had 80% power to detect a ICC of 0.51 compared with an ICC 0.29 [12] under the null hypothesis. To assess intraobserver reliability, ICC values were calculated using a one-way analysis of variance (ANOVA) model for agreement and for interobserver reliability, we calculated ICC values with two-way mixed-effects ANOVA models using SPSS Version 13.0 (SPSS Inc, Chicago, IL, USA); 95% confidence intervals for ICC values are provided. An ICC value of 1 indicates perfect reliability, 0.8 to 1 is very good, 0.61 to 0.8 is good, 0.41 to 0.6 is moderate, and less than or 0.4 is poor [12] (Table 1). We also calculated the average difference in degrees between the measurements we expect 95% of the differences to fall within 2 SDs of zero, which is the repeatability coefficient [7, 8]. As an additional metric, we calculated the proportion of the time that the same measurement by different observers and by the same observer would fall within 5° of each other, called the  $\pm$  margin of equivalency [7, 8].

**Table 2.** Interobserver reliability

Observer	Measure 1	Measure 2	Intraclass correlation coefficient	95% confidence interval	
				Lower bound	Upper bound
Femoral	Orthopaedic resident	Musculoskeletal radiologist	0.681	0.502	0.804
	Orthopaedic resident	Orthopaedic faculty	0.156	-0.123	0.410
	Musculoskeletal radiologist	Orthopaedic faculty	0.328	0.061	0.551
	Overall		0.386	0.214	0.556
Tibial	Orthopaedic resident	Musculoskeletal radiologist	0.755	0.122	0.909
	Orthopaedic resident	Orthopaedic faculty	0.615	0.414	0.759
	Musculoskeletal radiologist	Orthopaedic faculty	0.618	0.260	0.799
	Overall		0.670	0.474	0.800
Combined	Orthopaedic resident	Musculoskeletal radiologist	0.726	0.131	0.892
	Orthopaedic resident	Orthopaedic faculty	0.532	0.306	0.701
	Musculoskeletal radiologist	Orthopaedic faculty	0.573	0.275	0.754
	Overall		0.617	0.430	0.757

**Table 3.** Intraobserver reliability

Observer	Component	Intraclass correlation coefficient	95% confidence interval		Average difference (measure 1-2)	SD difference (measure 1-2)
			Lower bound	Upper bound		
Resident	Femoral	0.842	0.740	0.906	-0.060	1.692
	Tibial	0.896	0.825	0.939	-0.917	4.199
	Combined	0.861	0.771	0.918	-0.977	4.839
Faculty	Femoral	0.293	0.027	0.522	0.365	3.029
	Tibial	0.654	0.467	0.785	-2.340	6.205
	Combined	0.577	0.365	0.733	-1.975	7.231
Overall	Femoral	0.606	0.462	0.710		
	Tibial	0.809	0.732	0.867		
	Combined	0.751	0.652	0.832		

## Results

The interobserver reliability of rotational measurement ranged from fair to good. The overall ICC measurement for interobserver reliability was 0.39 for the femoral component, 0.67 for the tibial component, and 0.62 for the combined rotation. The 95% confidence intervals were 0.21 to 0.56 for the femoral component, 0.47 to 0.80 for the tibial component, and 0.43 to 0.76 for the combined measurement (Table 2).

The intraobserver reliability of rotational measurement ranged from good to very good. The overall intraobserver ICC for the orthopaedic resident and faculty member was

0.61 for the femoral component, 0.81 for the tibial component, and 0.75 for the combined measurement. The respective 95% confidence intervals were 0.46 to 0.71 for the femoral component, 0.73 to 0.86 for the tibial component, and 0.66 to 0.83 for combined rotation (Table 3).

The intraobserver overall  $\pm 5^\circ$  margin of equivalency for the femoral component was 95%; for the tibial component, it was 69%; and for the combined rotation, there was a 54% chance that all the measurements taken would fall within  $5^\circ$  of each other (Table 4). The intraobserver repeatability coefficient for the femoral component was  $5^\circ$ , tibial component  $11^\circ$ , and the combined rotation  $12^\circ$  (Table 4).

**Table 4.** Intraobserver agreement and test repeatability

Observer	Component	Percent agree (within $\pm 5$ )	95% confidence interval		Average difference (measure 1–2)	Repeatability coefficient (2 SD)
			Lower	Upper		
Resident	Femoral	100%	–	–	–0.060	$\pm 3.38$
	Tibial	75.0%	61.1%	86.0%	–0.917	$\pm 8.40$
	Combined	61.5%	47.0%	74.7%	–0.977	$\pm 9.68$
Faculty	Femoral	90.4%	79.0%	96.8%	0.365	$\pm 6.06$
	Tibial	63.5%	49.0%	76.4%	–2.340	$\pm 12.41$
	Combined	46.2%	32.2%	60.5%	–1.975	$\pm 14.46$
Overall	Femoral	95.2%	89.1%	98.4%	0.153	$\pm 4.90$
	Tibial	69.2%	59.4%	77.9%	–1.629	$\pm 10.64$
	Combined	53.9%	43.8%	63.7%	–1.476	$\pm 12.29$

## Discussion

Attention to implant rotation after TKA has increased following the contribution of Berger et al. [4]. Additionally, Bedard et al. [3] demonstrated the benefit of revision surgery to correct malrotation. In this study, the authors showed improved Knee Society score, pain scores, and ROM with revision surgery to correct internal rotation of TKA components and confirmed the correction with pre- and postrevision 2-D CT scans. Unfortunately, there is a lack of established literature evaluating the reliability of measuring rotation by 2-D CT scan. Our study evaluated the 2-D CT scans of TKA patients' components to determine the inter- and intraobserver reliability and repeatability coefficient of our measurements for component rotation. The interobserver reliability measurement was good (0.62 combined) but much better for the tibial component than the femur. The 95% confidence intervals were rather wide for all measurements and better for the tibia than the femur. We found that the intraobserver reliability ICC measurements were good (0.75 combined) with the tibial component again having greater reliability than the femur. Also important were the results of the  $\pm 5^\circ$  equivalency calculations. Although it was not consistent for the measurement of the femoral component to be the same when done twice by the same reviewer, or another, the measurements were closely grouped within a few degrees. Therefore, it is likely that a given measurement will be within  $5^\circ$  of another on the femoral side. This is much less likely on the tibial side or for the combined rotation. According to this study, there is only a 54% chance that if the measurements are done twice, the combined rotations will be within  $5^\circ$  of each other. Perhaps most interesting were the results of the repeatability coefficient. The magnitude of variability by the same reviewer was high for all three measurements and significantly

higher for the tibial component ( $11^\circ$ ) than the femur ( $5^\circ$ ). With the same reviewer reading the same scan twice, there was a  $12^\circ$  difference in combined rotation. This variability could significantly alter a surgeon's evaluation and treatment plan. There are two limitations to our study worth considering. First, this is a relatively small sample (52 knees) and we used only three reviewers. Although our statistical power was adequate, larger studies with more patients or reviewers, and perhaps other practice settings, would help to determine how well our findings might generalize to others. Second, there is no validation with known controls to these measurements as would be done in most studies. It is possible that our measurements are significantly different from others and a known control value would identify this bias. Hirschmann et al. evaluated 30 knees in 29 patients who had undergone TKA with 2-D and three-dimensional (3-D) CT to determine which method would be more accurate [12]. They found the measurements on 3-D CT (intraobserver reliability ICC 0.91) were statistically better compared with 2-D (intraobserver ICC 0.29). Based on their findings, they recommended 3-D CT scans for determining component orientation of TKA.

The strength of this study is that the measurements are objective readings by three blinded reviewers using an established technique and protocol [5]. The technique by Berger and Rubash remains the common practice for measuring TKA component malrotation, and its contribution to our understanding of this complication should not be overlooked. A technique for using a 3-D CT scan has been described and may gain wide acceptance [12]. Our study demonstrates that the 2-D CT protocol most commonly used to assess rotation of TKA components has good but not excellent intra- and interobserver reliability and that there can be a significant magnitude of variability in degrees measured by the same reviewer. Our findings highlight the difficulty measuring component malrotation



with results that are precise and repeatable. The clinical use of this technique must be carefully considered therefore if the attribution of symptoms or recommendation for revision surgery is based on the CT measurement of malrotation.

**Acknowledgments** We thank Melissa Manzer MD, for her contributions to this study by her radiographic measurements of component rotation.

## References

- Bargren JH. Total knee dislocation due to rotatory malalignment of tibial component: A case report. *Clin Orthop Relat Res.* 1980;147:271–274.
- Barrack RL, Schrader T, Bertot AJ, Wolfe MW, Myers L. Component rotation and anterior knee pain after total knee arthroplasty. *Clin Orthop Relat Res.* 2001;392:46–55.
- Bedard M, Vince KG, Redfern J, Collen SR. Internal rotation of the tibial component is frequent in stiff total knee arthroplasty. *Clin Orthop Relat Res.* 2011;8:2346–2355.
- Berger RA, Crossett LS, Jacobs JJ, Rubash HE. Malrotation causing patellofemoral complications after total knee arthroplasty. *Clin Orthop Relat Res.* 1998;356:144–153.
- Berger RA, Rubash HE. Rotational instability and malrotation after total knee arthroplasty. *Orthop Clin North Am.* 2001;4:639–47, ix.
- Berger RA, Rubash HE, Seel MJ, Thompson WH, Crossett LS. Determining the rotational alignment of the femoral component in total knee arthroplasty using the epicondylar axis. *Clin Orthop Relat Res.* 1993;286:40–47.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;8476:307–310.
- British Standards Institution. Precision of test methods. In: *Guide for the Determination and Reproducibility for a Standard Test Method.* London, UK: London BSI; 1979.
- Christensen CP, Crawford JJ, Olin MD, Vail TP. Revision of the stiff total knee arthroplasty. *J Arthroplasty.* 2002;4:409–415.
- Eckhoff DG, Metzger RG, Vandewalle MV. Malrotation associated with implant alignment technique in total knee arthroplasty. *Clin Orthop Relat Res.* 1995;321:28–31.
- Eckhoff DG, Piatt BE, Gnadinger CA, Blaschke RC. Assessing rotational alignment in total knee arthroplasty. *Clin Orthop Relat Res.* 1995;318:176–181.
- Hirschmann MT, Konala P, Amsler F, Iranpour F, Friederich NF, Cobb JP. The position and orientation of total knee replacement components: A comparison of conventional radiographs, transverse 2D-CT slices and 3D-CT reconstruction. *J Bone Joint Surg Br.* 2011;5:629–633.
- Incavo SJ, Wild JJ, Coughlin KM, Beynon BD. Early revision for component malrotation in total knee arthroplasty. *Clin Orthop Relat Res.* 2007:131–136.
- Insall JN, Kelly M. The total condylar prosthesis. *Clin Orthop Relat Res.* 1986;205:43–48.
- Jazrawi LM, Birdzell L, Kummer FJ, Di Cesare PE. The accuracy of computed tomography for determining femoral and tibial total knee arthroplasty component rotation. *J Arthroplasty.* 2000;6:761–766.
- Kanekasu K, Kondo M, Kadoya Y. Axial radiography of the distal femur to assess rotational alignment in total knee arthroplasty. *Clin Orthop Relat Res.* 2005;434:193–197.
- Nagamine R, White SE, McCarthy DS, Whiteside LA. Effect of rotational malposition of the femoral component on knee stability kinematics after total knee arthroplasty. *J Arthroplasty.* 1995;3:265–270.
- Ranawat CS. The patellofemoral joint in total condylar knee arthroplasty. pros and cons based on five- to ten-year follow-up observations. *Clin Orthop Relat Res.* 1986;205:93–99.
- Rand JA, Coventry MB. Ten-year evaluation of geometric total knee arthroplasty. *Clin Orthop Relat Res.* 1988;232:168–173.
- Torga-Spak R, Parikh SN, Stuchin SA. Anterior knee pain due to biplanar rotatory malalignment of the femoral component in total knee arthroplasty. case report. *J Knee Surg.* 2004;2:113–116.
- Vince KG, Insall JN, Kelly MA. The total condylar prosthesis. 10- to 12-year results of a cemented knee replacement. *J Bone Joint Surg Br.* 1989;5:793–797.