

# Accuracy of Knee Range of Motion Assessment After Total Knee Arthroplasty

Carlos Lavernia, MD,\* Michele D'Apuzzo, MD, †  
Mark D. Rossi, PhD, PT, CSCS, ‡ and David Lee, PhD §

---

**Abstract:** Measurement of knee joint range of motion (ROM) is important to assess after total knee arthroplasty. Our objective was to determine level of agreement and accuracy between observers with different knowledge on total ROM after total knee arthroplasty. Forty-one patients underwent x-ray of active and passive knee ROM (gold standard). Five different raters evaluated observed and measured ROM: orthopedic surgeon, clinical fellow, physician assistant, research fellow, and a physical therapist. A 1-way analysis of variance was used to determine differences in ROM between raters over both conditions. Limit of agreement for each rater for both active and passive total ROM under both conditions was calculated. Analysis of variance indicated a difference between raters for all conditions (range,  $P = .004$  to  $P \leq .0001$ ). The trend for all raters was to overestimate ROM at higher ranges. Assessment of ROM through direct observation without a goniometer provides inaccurate findings.

**Key words:** range of motion, total knee arthroplasty, knee flexion.

© 2008 Published by Elsevier Inc.

---

Accurate measurement of knee joint range of motion (ROM) is extremely important for orthopedic surgeons (OSs). These measurements have prognostic significance preoperatively and are used to delineate postoperative rehabilitation. Sufficient knee joint flexion is required to safely complete activities of daily living such as stair climbing, walking, and rising from a chair [1,2]. Assessment of knee joint ROM before and after knee arthroplasty surgery is also used in outcome assessment measures such as knee scoring instruments [3]. In the

last 3 to 4 years, multiple high-flexion knee devices have been introduced in the marketplace by several manufacturers (ie, DePuy and Biomet). Most of these manufacturers claim that specific design features allow the devices to flex "safely" past 140°. In the next decade, accurate assessment of ROM will become even more important as we monitor the performance of these new devices.

Multiple authors have reported on the consistency of measuring knee joint ROM, within-tester and between-tester reliability [4-7]. Visual inspection of submaximal knee joint ROM using a goniometer has been reported to be very close to the gold standard (radiographic image) [8]. Edwards et al [9] reported that the intertester reliability (.91) was high between 3 different testers: an attending surgeon, senior orthopedic surgery resident, and physician assistant (PA). However, no study has evaluated the accuracy of assessing knee joint ROM between different health care providers (ie, surgeons, physical therapists [PTs], and fellows) and compared it with the gold standard, roentgenographic measurement. Our first objective was to determine level of agreement between raters on observed and measured total ROM (active and passive) with x-ray measurement after total knee arthroplasty (TKA). The second objective was to

---

*From the \*Orthopaedic Institute at Mercy Hospital, Miami, Florida; †Arthritis Surgery Research Foundation, Inc., Miami, Florida; ‡Department of Physical Therapy, Florida International University, Miami, Florida; and §Department of Epidemiology and Public Health, University of Miami, Miami, Florida.*

Submitted January 25, 2008; accepted May 11, 2008.

Conflict of interest statement: Financial support has been received by the corresponding author. Institutional support: Mercy Hospital and Mercy Foundation, Miami, Fla; Zimmer, Inc, Warsaw, Ind; and Arthritis Surgery Research Foundation, Inc, Miami, Fla.

Reprint requests: Carlos J. Lavernia, MD, Orthopaedic Institute at Mercy, 3659 S. Miami Ave Ste 4008, Miami, FL 33133.

© 2008 Published by Elsevier Inc.

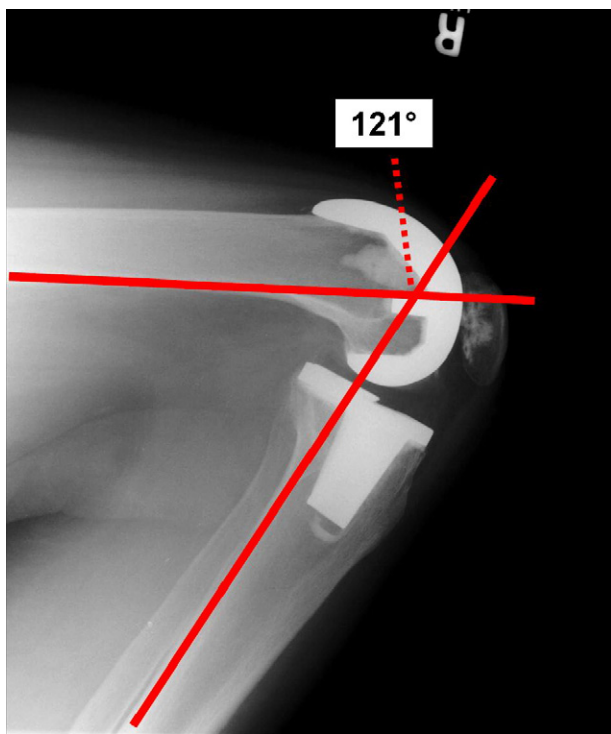
0883-5403/08/2306-0016\$34.00/0

doi:10.1016/j.arth.2008.05.019

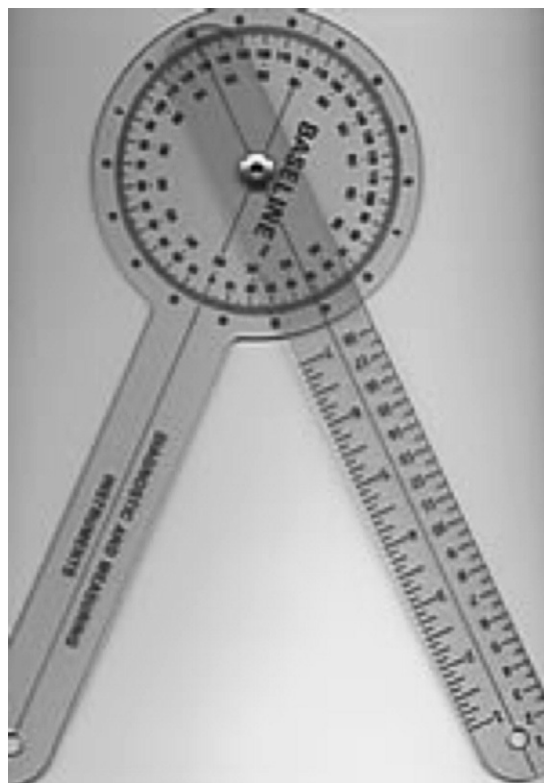
evaluate the accuracy of assessing active and passive observed and measured ROM between raters.

## Methods

Forty-one consecutive patients (25 females and 16 males), a minimum of 1 year post-TKA, volunteered and gave written consent to participate in this study. All patients gave written informed consent to participate. These patients had primary TKA performed by the senior author and received 8 different combinations of implant design and tibial insert. Fifteen patients received NexGen CR (Zimmer) implants, 13 received NexGen CR-Flex (Zimmer), 5 received anatomic modular knee (AMK, DePuy), 4 patients received low contact stress (LCS, DePuy), 3 patients received NexGen legacy knee constrained condylar knee (LCCK, Zimmer), and 1 patient received a Duracon (Howmedica) implant. Radiographic evaluation was done with the patient supine and the hip flexed on the x-ray table. A lateral x-ray of the knee was taken in passive and active flexion by a technician not involved in the clinical measurement phase of the study. The radiographs were taken using a digital radiograph x-ray system. All measurements were performed on a picture archiving and communication system (PACS, software E-Film, Milwaukee, Wisconsin) using the angle between lines drawn down the midshafts of the



**Fig. 1.** Radiograph showing measurement lines for knee flexion angle of 121°.



**Fig. 2.** Standard goniometer.

femur and tibia. The degree of flexion was obtained by measuring the angle between lines drawn down the posterior cortex of the femur and tibia (Fig. 1). The radiographic measurements were done with no knowledge of the clinical measurements.

Five different raters evaluated the ROM in all patients. These included an experienced fellowship-trained OS, a clinical fellow (CF), a PA, a research fellow (RF), and a clinical PT. Each performed 4 measurements in all patients: visual (observed) and measured (goniometer) assessments of passive and active motion were noted in each case. Evaluations by raters consisted of visually estimating (observed) total ROM from extension into flexion followed by actual measurement of ROM using a handheld 2-arm goniometer (26-cm arms with 2° markings; Fig. 2). Both observed and measured ROMs were assessed actively and passively. For measurement using the goniometer, 1 arm of the goniometer was placed parallel to the shaft of the femur lining up with the greater trochanter, and the other arm was placed parallel to the shaft of the lower leg lining up with the lateral malleolus of the fibula. The axis of the goniometer was placed over the approximate knee joint line, slightly below the lateral femoral epicondyle. The dependent variable was total ROM from full available extension to full available knee flexion. Each evaluator was blinded to all measurements by the other raters.

A 1-way analysis of variance was used to determine differences in ROM between raters for active and passive ROM over both conditions (observed and measured). Follow-up post hoc assessments using Tukey tests were completed when appropriate with level of significance set at .05. We calculated the mean difference for each rater, which was defined as the difference between the radiographic ROM (“true” ROM) and each ROM assessment. The 95% limit of agreement (mean difference  $\pm$  [1.96  $\times$  standard deviation of the differences] [10]) was also computed for each rater for both active and passive total ROM under both conditions (measured and observed).

Accuracy over a range of values for observed and measured active ROM was assessed by plotting the average of the active ROM obtained from x-ray and from the rater against the mean difference of the same 2 measures for active measurements over both conditions. For each rater, we calculated the average difference between the active observed measurement and the gold standard of the passive measurement taken by x-ray. We then calculated the average difference between the passive observed measurement and the passive measurement taken by x-ray. Finally, we calculated the average improvement or how “much closer” passive observed measurement came to the x-ray over and above the observed measure. We used paired *t* tests to determine if providing “over pressure” or passive motion significantly improved the measurement, bringing the value closer to the gold standard. SPSS software (Chicago, Ill) was used for all analyses.

## Results

### Clinician-Specific Average ROM Results

The analysis of variance indicated that there was a difference between raters for all conditions (range,  $P =$

.004 to  $P \leq .0001$ ) indicating that there were differences in each of the measured values across clinicians. Descriptives for all raters and all conditions are presented in Table 1.

### Mean Differences and Limits of Agreement Calculated Using X-Ray Values as the Gold Standard

Using x-ray values as the gold standard, the mean difference for the actively observed ROM ranged from  $-1.14^\circ$  for the RF to  $-9.41^\circ$  for the PT. Active measured ROM ranged from  $2.9^\circ$  for the RF to  $-7.4^\circ$  for the PT. For passive observed ROM, mean difference values ranged from  $-1.10^\circ$  for the RF to  $-14.42^\circ$  for the PT. For passive measured ROM, the mean difference ranged from  $-2.32^\circ$  for the RF to  $-12.65^\circ$  for the PT.

For all measures and conditions, the RF and CF were closest to “true values” as determined by the x-ray ROM measurement. The PT was the farthest away from true values for all measures and conditions. For all raters, the limits of agreement (LOA) was quite large, indicating substantial variability relative to the gold standard measure (x-ray) (eg, active observed ROM for the OS; 95% LOA,  $14.83^\circ$ – $22.87^\circ$ ).

### Assessment of Clinician-Specific Mean Differences

Figs. 3 and 4 provide a graphical representation of the accuracy of clinical assessment for the clinician-specific (observed) LOA over a range of average ROM values for active measures recorded by OS and PT, respectively. Graphs were similar for all other raters. This accuracy is represented by the dotted line of each graph. For example, if the clinician agreed perfectly with the x-ray-based findings, then the dotted line would be horizontal and would be centered a zero on the horizontal axis (ie, no mean difference between clinician and x-ray values across all ROM values). An upward sloping line would indicate that as the true ROM value increased, the clinician would overestimate this value; a downward sloping line would indicate the opposite trend.

### Active Observed ROM

Based on the scatter plot and using a mean difference of “zero degrees” as a reference, the trend for all raters was to overestimate ROM at higher ranges, more specifically; the RF underestimated knee joint ROM at low ranges and overestimated at ranges greater than  $110^\circ$  (mean difference,  $-1.14$ ; LOA,  $18.87^\circ$  to  $-16.59^\circ$ ). The CF underestimated ROM in the lower ranges with a slight overestimation at higher ranges, greater than  $110^\circ$  (mean difference,  $-2.12$ ; LOA,  $19.91^\circ$  to  $-24.15^\circ$ ). The PA (mean difference,  $-3.46$ ; LOA,  $10.33^\circ$  to  $-17.25^\circ$ ) and OS (mean difference,  $-4.02$ ; LOA,  $14.83^\circ$  to  $-22.87^\circ$ ; Fig. 3) exhibited a similar pattern, a general underestimation

**Table 1.** Patient Characteristics

Sex	
Female	25
Male	16
Race	
White	30
African American	4
Hispanic	4
Other	3
Ethnicity	
Cuban/Cuban American	30
Spanish NOS	5
South/Central American	5
Non-Spanish	1
BMI (kg./m <sup>2</sup> )	
Overweight (25-29.9)	11
Obese (>30)	30

NOS indicates not otherwise specified; BMI, body mass index.

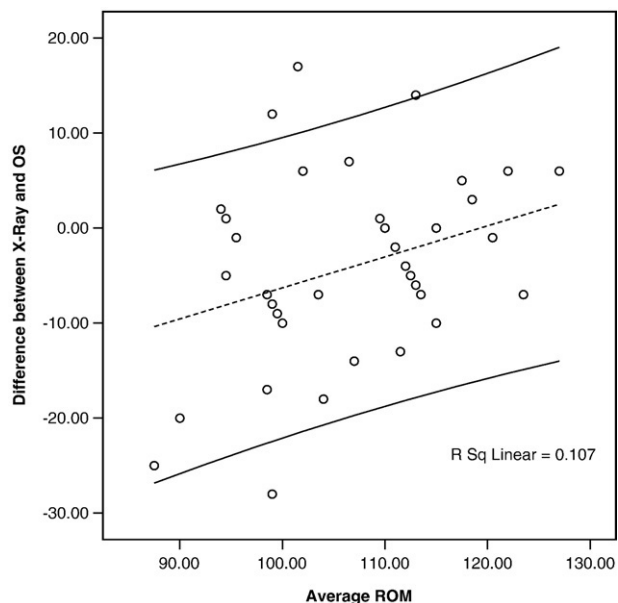
throughout most of the range with slight overestimation at higher ranges, greater than 115°. Although the trend line indicated an “upward” sloping trend toward overestimation at higher ranges, it should be noted that the PT (mean difference, -9.41; LOA, 5.97° to -24.79°; Fig. 4) underestimated ROM throughout the entire range.

**Active Measured ROM**

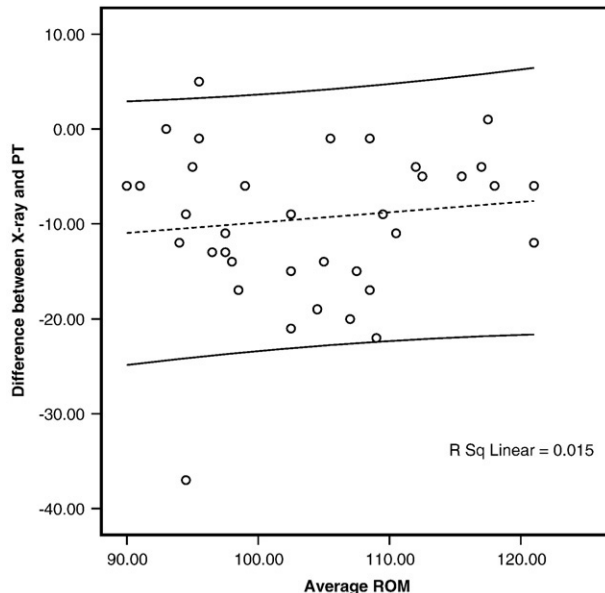
Using a zero degree mean difference as a reference, the clinical RF had a distribution with a large spread of scores from low to high ranges. The RF had higher variability of scores in the lower ranges (<105°) but had less variability in the higher ranges with a tendency for slight overestimation. The distribution of scores for the PA had low variability, and based on a “zero” reference, there was an underestimation of knee joint measurement over low to high scores. The PT had a high level of underestimation over low to high ranges. The OS had a low spread of scores with a fairly equal underestimation and overestimation of ROM over low to high ranges. In fact, measured scores for the OS were more accurate compared with other raters because values were much closer to the zero degree mean difference.

**Assessment of Net Change of ROM of Passive Assessment Over and Above the Observed Measure**

For all raters, there continued to be an underestimation of observed ROM even with the additional force provided by passive ROM. However, observed passive ROM



**Fig. 3.** Scatter plot for active observed ROM for OS. The dotted line indicates best line of fit and solid lines represent 95% confidence intervals.



**Fig. 4.** Scatter plot for active observed ROM for PT. The dotted line indicates best line of fit, and solid lines represent 95% confidence intervals.

significantly improved assessment for all raters over and above the active observed measure ( $P \leq .0001$ ). Observed passive measurement most improved the CFs assessment closer to the gold standard. The addition of overpressure or passive assessment brought the RF to within approximately 2° of the x-ray measurement, whereas the PTs assessment is still largely an underestimation with the measurement being greater than 14° from the x-ray. All descriptives, average differences, and  $P$  values are presented as Table 2.

**Discussion**

Knee joint ROM is important to assess before and after joint arthroplasty surgery. Range of motion will affect an individual's ability to safely and independently perform activities of daily living such as sitting or negotiating stairs. It also is an extremely important determinant of patient satisfaction. Assessment of joint ROM in knee surgery is a key factor in deciding when to encourage the patients to have the intervention early as well as modifying the postoperative rehabilitation regimen. A number of professionals are usually involved in the assessment of ROM of the knee before and after the surgical intervention. This measurement is important not only to researchers but also to the practicing clinician.

The therapists are the main line of information for surgeons making decisions on post-TKA patients after discharge from the hospital. Most surgeons rely on the PTs documentation to plan follow-up physician visits

**Table 2.** Change in ROM (Degrees  $\pm$  SEM) From Active to Passive ROM With Reference to X-Ray Measured Passively

Rater	Average X-Ray (Passive)	Average Observed (Active)	Average Difference Between x-ray and Active Observed	Average Observed (Passive)	Average Difference Between X-Ray and Passive Observed	Improvement in ROM From Active to Passive (Observed)
OS	117.53 $\pm$ 1.75	104.93 $\pm$ 1.89	12.58 $\pm$ 1.48	108.63 $\pm$ 1.97	8.89 $\pm$ 1.46	3.69 $\pm$ .46*
CF	117.53 $\pm$ 1.75	106.83 $\pm$ 2.08	10.65 $\pm$ 1.84	112.80 $\pm$ 2.04	4.72 $\pm$ 1.81	5.93 $\pm$ .94*
PA	117.53 $\pm$ 1.75	105.49 $\pm$ 1.76	12.03 $\pm$ 1.07	110.66 $\pm$ 1.78	6.86 $\pm$ 1.04	5.17 $\pm$ .55*
RF	117.53 $\pm$ 1.75	110.10 $\pm$ 2.20	7.42 $\pm$ 1.22	115.78 $\pm$ 2.40	1.74 $\pm$ 1.39	5.68 $\pm$ .56*
PT	117.53 $\pm$ 1.75	99.54 $\pm$ 1.58	17.96 $\pm$ 1.38	102.78 $\pm$ 1.72	14.74 $\pm$ 1.31	3.22 $\pm$ .37*

Extension degrees are degrees short of full extension.

\*  $P \leq .0001$ , significant improvement from active to passive measurement.

and the need for manipulations under anesthesia. The effectiveness of the rehabilitation is also assessed based on the documentation of PTs. There is a golden period after a primary procedure after which a manipulation under anesthesia is effective [11-13]. Our data clearly indicate that when compared with the surgeon and other medical providers, the PT tended to underestimate the ROM of the knee throughout the entire range. In fact, even with the addition of passive assessment, the PT continued to be far from the gold standard. In our study, the ROM being reported by the PT was the lower bound of actual motion. However, we cannot draw inferences about the underestimation by 1 PT to that of the entire population of PTs. What is important is that regardless of rater (clinician), there was a common underestimation of ROM. Range of motion assessment, when needed to define further procedures, should perhaps be done using x-ray rather than be clinician based. Our study data would suggest that patients within 10° of their trigger for changing rehabilitation protocol or manipulating the knee need to be evaluated by the surgeon or have an x-ray.

Compared with active measurements, passive forced motion assessments were much more variable with larger dispersion of scores for both observed and measured conditions for all raters. This probably reflects the difficulty in assessing small changes during passive assessment. This difficulty is present regardless of whether a goniometer is used and is probably due to individual variability. Our data demonstrate that regardless of rater, assessing passive or active ROM through visual observation without the use of a goniometer gives inaccurate scores with large variability over low to high ranges.

Range of motion assessment through direct observation without a goniometer provides inaccurate findings. More specifically, we are alerted to the fact that there is a tendency to overestimate ROM at low ranges in some clinicians. Less troubling, but important to recognize, is the underestimation of knee joint ROM at higher ranges.

We agree with Edwards et al [9] that clinical assessments of knee joint ROM should be routinely checked against an x-ray. Moreover, considering the number of

health care providers who assess and work with patients having knee arthroplasty and the importance of measuring knee joint ROM, we recommend that a standard approach for measuring knee joint ROM be used. Standardizing an approach could be achieved through continuing education across a variety of disciplines.

Researchers and reviewers should pay close attention to the methodology used to measure ROM in all studies. In addition, retrospective studies that rely on chart review for ROM should discard measurements done by visual observation. Moreover, a standardized approach should be used to assess ROM when many raters are involved, thus well defining operational procedures, that is, what position was the assessment made in? Except for the PT who consistently underestimated observed active ROM over all ranges, the remaining raters had similar findings; there was an underestimation of motion at lower ranges and an overestimation at higher ranges. This is an important finding for surgeons in large practices following patients clinically.

Measured scores using a goniometer provided an improved degree of accuracy, but results appear to be dependent on the clinician performing the measurement. The OS had better accuracy, followed by the PA who had a slight overestimation at high ranges. Involving a PA, senior orthopedic resident, and OS, Edwards et al reported that observed ROM ranged from underestimating range 15° to overestimating knee flexion 15° [9]. Contrary to our study, overestimation of ROM was reported at higher ranges. This is concerning because some clinicians are "giving" more range than does truly exist. Numerous studies have shown that the goniometric measuring technique, if held constant, has good to excellent (.79-.92) reproducibility within and between raters [4-7,9]. In our study, goniometric technique was reliable for all raters. Recording the proper motion on all clinic visits mandates the use of a long-armed goniometer. Although the technique for our study was adhered to by all raters, the difference in accuracy between raters could have been due to measurement parallax, level of education, and or years of experience.

Several investigators have reported on the effects of implant design in the final ROM after TKA [14-16]. Dennis et al [2] reported that posterior sacrificed knees had slight improvements in knee joint ROM over cruciate-sparing designs. Moreover, Banks et al [17] reported that there was more knee flexion ROM in individuals having fixed bearing prostheses compared to mobile bearing designs. Interestingly, Maloney and Schurman [18] noted that there was no difference in ROM between total condylar and posterior stabilized designs. However, regardless of design, studies have shown that preoperative ROM is the best predictor of postoperative ROM outcome [19,20]. None of the above authors used x-rays to assess ROM between designs. In the future, implant designers should use x-rays to get a better measurement of ROM.

The measured radiographic angulation between the long axis of the femur and the long axis of the tibia is considered the gold standard as the true ROM of the knee joint [8]. Our results show that accurate assessment of knee joint ROM after knee joint arthroplasty is difficult to achieve and that although not routinely done, the x-ray provides the best estimation of range of knee joint motion. This is extremely important in the studies where high flexion devices are compared with one another. The use of the gold standard x-ray is lacking in most series that study motion. To accurately make statements on the motion of these devices, an x-ray has to be taken and measurements done using digital radiographs.

A major weakness of this study is that only 1 professional from each field performed the measures and that our results may not apply or reflect the performance of more experienced raters thus, making generalizations difficult. However, our study does present new information that can help define further investigations. Furthermore, variability in passive ROM assessment may be related to the amount of force used to flex the knee by the examiner. In an effort to standardize the test and control force that flexes the knee, an alternative method such as measurement taken during weight-bearing flexion could have been used. Another limitation is that we studied individuals with a variety of implants and yet we did not assess ROM differences between implants. We also did not assess preoperative ROM with the same methodology and correlate with postoperative ROM; however, studies have shown that the best predictor of postoperative ROM is the preoperative measure [21,22].

Further studies should evaluate the effect of years of experience and level of education on the accuracy of assessing knee joint ROM across medical and allied health personnel. Considering the variability between raters evident in our study when assessing knee joint ROM, we are also concerned with the accuracy of measuring joint ROM of the other joints in the body and the potential variability across health care professionals.

## References

1. Lubenthal KN, Smidt GL, Kettlekamb DB. A quantitative analysis of knee motion during activities of daily living. *Phys Ther* 1972;52:34.
2. Dennis DA, Komistek RD, Stiehl JB, et al. Range of motion after total knee arthroplasty: the effect of implant design and weight-bearing conditions. *J Arthroplasty* 1998;13:748.
3. Insall JN, Dorr LD, Scott RD, et al. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res* 1989;248:13.
4. Brosseau L, Tousignant M, Budd J, et al. Intratester and intertester reliability and criterion validity of the parallelogram and universal goniometers for active knee flexion in healthy subjects. *Physiother Res Int* 1997;2:150.
5. Mayerson NH, Milano RA. Goniometric measurement reliability in physical medicine. *Arch Phys Med Rehabil* 1984;65:92.
6. Rothstein JM, Miller PJ, Roettger RF. Goniometric reliability in a clinical setting. *Elbow and knee measurements. Phys Ther* 1983;63:1611.
7. Boone DC, Azen SP, Lin CM, et al. Reliability of goniometric measurements. *Phys Ther* 1978; 58:1355.
8. Gogia PP, Braatz JH, Rose SJ, et al. Reliability and validity of goniometric measurements at the knee. *Phys Ther Res* 1987;67:192.
9. Edwards JZ, Greene KA, Davis RS, et al. Measuring flexion in knee arthroplasty patients. *J Arthroplasty* 2004;19:369.
10. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135.
11. Fox JL, Poss R. The role of manipulation following total knee replacement. *J Bone Joint Surg Am* 1981;63:357.
12. Esler CN, Lock K, Harper WM, et al. Manipulation of total knee replacements. Is the flexion gained retained? *J Bone Joint Surg Br* 1999;81:27.
13. Gandhi R, de Beer J, Leone J, et al. Predictive risk factors for stiff knees in total knee arthroplasty. *J Arthroplasty* 2006;21:46.
14. Coughlin KM, Incavo SJ, Doohen RR, et al. Kneeling kinematics after total knee arthroplasty: anterior-posterior contact position of a standard and a high-flex tibial insert design. *J Arthroplasty* 2007;22:160.
15. Klein GR, Restrepo C, Hozack WJ. The effect of knee component design changes on range of motion evaluation in vivo by a computerized navigation system. *J Arthroplasty* 2006;21:623.
16. Mahoney OM, McClung CD, dela Rosa MA, et al. The effect of total knee arthroplasty design on extensor mechanism function. *J Arthroplasty* 2002; 17:416.
17. Banks S, Bellemans J, Nozaki H, et al. Knee motions during maximum flexion in fixed and mobile-

- bearing arthroplasties. *Clin Orthop Relat Res* 2003;410:131.
18. Maloney WJ, Schurman DJ. The effects of implant design on range of motion after total knee arthroplasty. Total condylar versus posterior stabilized total condylar designs. *Clin Orthop Relat Res* 1992;278:147.
  19. Anouchi YS, McShane M, Kelly Jr F, et al. Range of motion in total knee replacement. *Clin Orthop Relat Res* 1996;331:87.
  20. Ritter MA, Harty LD, Davis KE, et al. Predicting range of motion after total knee arthroplasty. Clustering, log-linear regression, and regression tree analysis. *J Bone Joint Surg Am* 2003;85-A:1278.
  21. Ryu J, Saito S, Yamamoto K, et al. Factors influencing the postoperative range of motion in total knee arthroplasty. *Bull Hosp Jt Dis* 1993;53:35.
  22. Schurman DJ, Rojer DE. Total knee arthroplasty: range of motion across five systems. *Clin Orthop Relat Res* 2005:132.