Inter-observer variability and its correlation to experience in measurement of lower limb mechanical axis on long leg radiographs

Raju Vaishya *, Vipul Vijay [1], Vikas P. Birla [1], Amit Kumar Agarwal [1]

Department of Orthopaedics, Joint Replacement & Arthroscopy, Indraprastha Apollo Hospital, New Delhi 110076, India

A R T I C L E   I N F O

A B S T R A C T

Introduction: Restoration of mechanical axis (MA) is one of the most important aims of treatment of knee arthritis. The measurement of MA is often done on hip knee ankle (HKA) radiographs (LLR), but its measurement is often difficult and variable. We studied to check if inter-observer variability in the measurement of MA is dependent on the experience of the observer.
Materials and methods: 48 patients (70 knees) underwent measurement of MA on HKA radiographs. The measurement was done by five observers of different experience. All the results were tested for inter-observer variability.
Results: The overall intra-class correlation was 0.70. The two full time consultants had good agreeability among them ($p < 0.456$). All the surgeons who had less than five years of experience among them also had good agreeability amongst them (all p values >0.005). All the other groups (except senior consultant and senior registrar) had statistically significant difference amongst them (all p values <0.005).
Conclusion: There is an inter-observer variability in the measurement of MA on the HKA radiographs. The agreeability between the observers increases as the experience of the personnel increases. Although long leg radiographs can be used for assessment for HKA, this variability should be kept in mind while using this as a tool for planning management of the arthritic knee.

© 2016 Delhi Orthopedic Association. All rights reserved.

1. Introduction

Osteoarthritis (OA) is one of the most common reasons for knee pain in patients of elderly age groups. The progression of OA is closely associated with increasing malalignment of the knee. Knee alignment is in fact now considered the major determinant in the progression of knee OA.[1] Varus or valgus alignment is associated with medial or lateral knee joint OA respectively.[2]

Total knee arthroplasty [TKA] is commonly accepted mode of management for advanced OA of the knee after the failure of conservative methods. Again, the success of TKA is dependent on the achievement of accurate alignment of mechanical axis [MA] of lower limb postoperatively.[3] Thus, analysis of preoperative MA of the leg is necessary for the planning of the surgery.[4] Moreover, the calculation of the pre-operative mechanical alignment can help the surgeon in making a decision about offering a patient joint preservation surgery (like high tibial osteotomy) or joint replacement.

Hip-knee-ankle [HKA] radiographs of the lower limb are regarded as the gold standard for assessment of MA in patients suspected to have malalignment or patients planned for surgeries (TKA, unicompartmental replacement, osteotomy).[5–7] Thus, any variation or error in the assessment of HKA may reflect the outcome of management.

Traditionally measurement of HKA axis on radiograph was done using printed paper. With the advent of Picture Archiving and Communication Systems (PACS), the measurement is made with the help of computers, avoiding the need for bulky hard copies.[8,9] The PACS has also avoided not only the drawbacks of hard copies such as efforts and but also the cost of printing and storing bulky hard copies.[10–12] Also, using magnification and contrast adjustment facilities with PACS measurement can be more accurate.[13]

The topic of current debate is whether the calculation of the HKA axis with PACS is reproducible with different observers and to calculate the inter-observer reliability in the measurement of MA on HKA radiographs. Very few studies have been done to examine the inter-observer variation and the changes that may occur depending on the experience.

* Corresponding author. Tel.: +91 11 29871210.
E-mail addresses: raju.vaishya@gmail.com (R. Vaishya),
dr_vipulvijay@yahoo.com (V. Vijay), gyankj1@gmail.com (V.P. Birla),
amitorthopgi@yahoo.com (A.K. Agarwal).
[1] Tel.: +91 11 29871210.

The purpose of the present study was to assess the inter-observer variability in the measurement of HKA axis using the PACS and try to evaluate its correlation with the experience of the staff measuring the axis.

2. Materials and methods

All the patients who came to our clinic with a history of knee pain and advanced degenerative knee OA were included in this study, from October 2013 to August 2015. The study group comprised of 28 females and 20 males. Among this 26 patients were unilateral and 22 patients were bilateral. The patients in whom unilateral measurement was done had a previous TKA on the contralateral side. The average age of the individuals was 62 years (range 55–76 years).

Standard full weight bearing whole lower limb (including Hip, Knee, and Ankle) radiograph in standing position were taken. Knees were kept in maximum extension with both feet parallel to each other and the patella facing forwards. Radiographs were focused from a distance of 2.5 m to include hip to ankle and were recorded on 35 cm × 43 cm cassettes.

Five doctors of the same unit were involved in the measurement of mechanical femoral–tibial angle (MFT angle) from HKA radiographs of study population done in standing position. All the measurements were done on a computer using PACS available in the hospital. All five observers were of different ranks on qualification and experience (Table 1). At the start of the study, a formal training session was done to let the observers know about the way to identify and mark the anatomical landmarks. The measurements were done individually and at different times using the same set of radiographs. The results of one observer were not revealed to the other surgeons and the results were blinded between observers.

The MFT angle was measured (Fig. 1) by drawing two lines on HKA radiograph done in standing weight bearing position. The first line was drawn from centre of hip to centre of the knee that denotes mechanical axis of the femur.[14] The second line was from centre of the knee to centre of the ankle that indicates the mechanical axis of the tibia.[14] The angle between these two lines is MFT angle.[15,16] A calculation of the angles was done after identifying the centre of hip joint by using the Mose circles.[17] The centre of the knee joint was determined by finding the midpoint of the tibial spines halfway between the intercondylar notch. The centre of the ankle was identified by the midpoint of the superior talar facet (talar dome). All five observer measured MFT angle from 70 HKA radiographs using PACS. On assessing the results, there was inter-observer variability in the measurement of the HKA axis.

The Friedman test was used to ascertain the differences in results from multiple test attempts by different personnel. All the observers were ranked by the Friedman's test (Table 2). The results were tested for statistical significance for consistency among the observers. Post hoc tests were run on the combination of the related group (observer one vs. observer 2, observer one vs.

observer three, etc.). In total, ten combinations were assessed for statistically significant differences. Bonferroni adjustments were done for the Wilcoxon signed rank tests and the data were found to be significant if the p-value was <0.005.

Bland–Altmann assessment of the measurements was done to identify the limits between which 95% of the values will fall. It was agreed that if the values of limits were small, then the agreement was acceptable. This data were presented as mean and standard deviation.

3. Results

The average value of the MFT of the different observers could be 175.50 (SD – 2.97), 175.660 (SD – 3.01), 174.940 (SD – 3.19), 174.670 (SD – 3.31), and 175.010 (SD – 3.40) with 70 HKA radiographs measured by each observer (Fig. 2). The results were checked for inter-observer variability using the statistical analysis as described and the results were interpreted.

By this statistical analysis, it was found that the two full-time consultants working in the department of adult reconstructive surgery had agreeability among them ($p = 0.456$, not significant). Similarly, the surgeons who were less than five years of experience in the branch also had an agreement among them (all $p > 0.005$). The statistically different readings were found to be between the senior consultant and the post graduate trainee ($p = 0.001$, statistically significant). Similar statistical significance difference was found to be between the senior consultant and the clinical fellow ($p = 0.002$, statistically significant). The other full-time consultant also had statistically significant difference with all the younger orthopaedic surgeons with less than five years of post PG experience ($p < 0.005$) (Table 1).

Ninety-five percent of the differences between measurements performed by other observers and ideal (RV) were between -2.2 and $3.1°$ (MFT angle) as calculated using the Bland–Altman method (Fig. 3). The intraclass correlation coefficient (ICC) for the values was found to be 0.70.

4. Discussion

The determination of the MA of the lower limb is essential in various surgeries of lower limb like corrective osteotomy, unicompartmental knee replacement, TKA. The intraoperative decision for bone cuts depends on the preoperative assessment of MA of the lower limb. Also, it has been widely accepted that the long-term survival of the TKA depends on the achievement of the accurate postoperative mechanical axis.[1] Many authors have also suggested that the total knee replacement that has varus alignment are more likely to fail.[18] Thus, an accurate measurement of the MFT is of utmost importance in planning the TKR. The amount of correction needed in a patient requiring high tibial osteotomy is also calculated by assessment of the pre-operative MFT.[19] Thus, the evaluation of the mechanical axis of paramount importance in any surgery involving restoration or realignment surgery around the knee. Thus, assessment of MA is critical, and

Table 1
The hierarchy of the observers with their values along with their descriptive statistics.

|	Hierarchy	Experience (in years) [0 being the year of passing training exam]	N	Mean	Std. deviation	Mean rank
Observer 1	Senior Adult Reconstructive Surgeon	25	70	175.5143	2.97386	3.59
Observer 2	Junior Adult Reconstructive Surgeon	7	70	175.6657	3.01662	3.75
Observer 3	Clinical Fellow	5	70	174.9429	3.19436	2.59
Observer 4	Registrar	1	70	174.6700	3.31913	2.66
Observer 5	Postgraduate trainee	−1	70	175.0171	3.40358	2.41
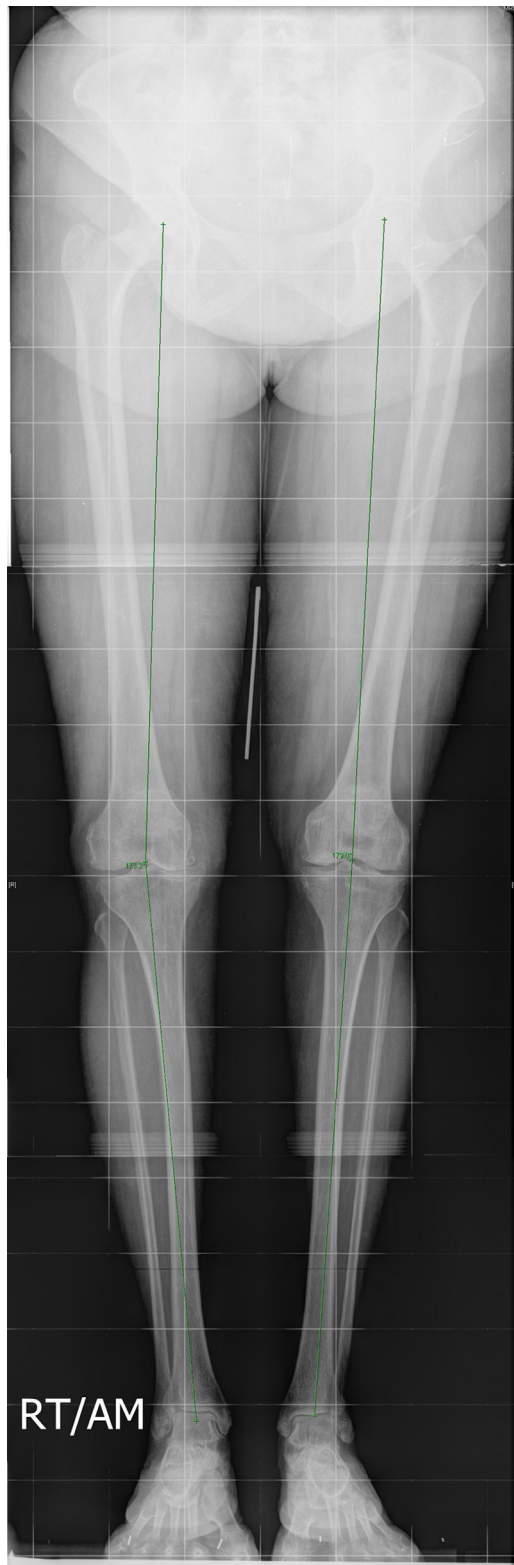
Fig. 1. Image depicting the measurement of the mechanical axis using the picture archiving system on standing long leg radiographs.

error in measurement may affect the long-term outcomes of joint preserving as well as joint replacement surgery of the knee.

Various methods have been used for measurement of the mechanical axis, but PACS has been shown to be an acceptable way due to the ease of analysis and storage.[13] Marx et al. showed higher inter-observer correlation coefficient on measurements done using PACS compared to analysis done using hard copies.[20] Hence, they suggested PACS should be used for assessment of MA instead of hard copies. However, they failed to comment on inter-observer variation in the measurement using PACS. Since then, there have been studies which have suggested good inter-observer reliability of using PACS for measurement of MA.[6,8]

The results of the present study were similar to other studies which showed a good correlation of the measurement between different observers. Hankemeier et al. reported good intra-observer reliability with the use of the computer-assisted measurement of the lower limb alignment and also cited the advantages of ease of storage and transfer of the Images.[8] Similarly, Specogna et al. reported that the estimates of error for an individual measurement of the mechanical axis were 1.50.[15] They suggested that the measurement error was minimal with good reliability. They also suggested that even though the error is small, it should always be considered when planning the clinical options. Rozzanigo et al. also reported a good accuracy of computer-assisted measurement of the mechanical axis.[21] Intra-operatively, accurate alignment of the lower limb in patients planned for TKA can also be achieved using computer navigation. Computer navigation has shown, in multiple studies, to improve the accuracy of lower limb alignment for arthroplasty surgeons.[22] Cip et al., in their randomized prospective trial, reported better overall implant positioning in the femur as well as tibia and also demonstrated better restoration of the mechanical axis in TKAs performed using computer navigation.[23] In the present study, we were able to achieve similar results with good correlation (ICC – 0.70) and the difference between the senior author and other observers was an acceptable −2.20 to 3.10.

None of the studies discussed in the literature have tried to correlate the reliability of measurement based on the experience of the observer. In our study inter-observer variation in measurement using PACS was analyzed using Friedman test for inter-observer reliability and was assessed for its correlation with experience. The two full-time consultants working in the adult reconstruction unit did not have a statistically significant difference in the measurement of the HKA (p-value 0.456). Both these consultants were involved in a full-time joint reconstructive unit and had an experience of 7 years and 25 years respectively. The results of the two consultants were statistically different from all the junior staff, including the fellow, registrar and the PG trainee (all p values <0.005). The only exception was the difference between a senior surgeon and the senior registrar.

We found that the results of measurement of the mechanical axis had a better agreement between the observers as the experience of the surgeons increased. The findings of the present study can be attributed to the better identification of the anatomical landmarks as the experience of the surgeon increases. The calculation of the mechanical axis involves the identification of three anatomical landmarks, on each at the hip, knee and ankle. Since these anatomical landmarks are not fixed points, they are calculated geometrically by finding the centre of the hip, knee and

Table 2
The statistical analysis between the various observer groups and the p-value (based on the Wilcoxon signed rank test after Bonferroni adjustment).

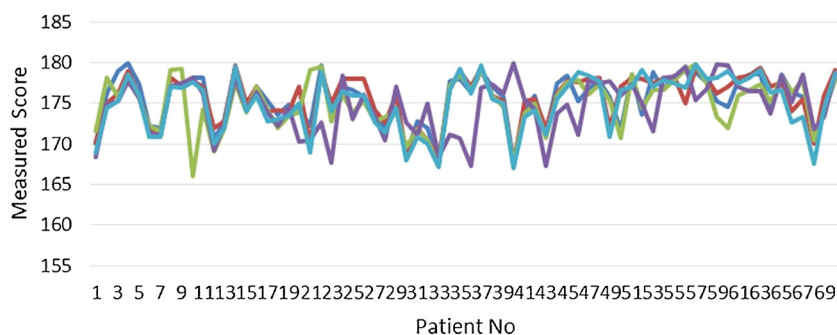|	RV-VV	RV-TCA	RV-VK	RV-VPB	VV-TCA	VV-VK	VV-VPB	VK-TCA	VPB-TCA	VPB-VK
Z	−.746	−3.158	−2.736	−3.223	−3.071	−3.080	−5.172	−.723	−.015	−.706
Asymp. Sig. (2-tailed)	.456	.002	.006	.001	.002	.002	.000	.470	.988	.480

Fig. 2. Scatter diagram showing the distribution of the values of all the observers.
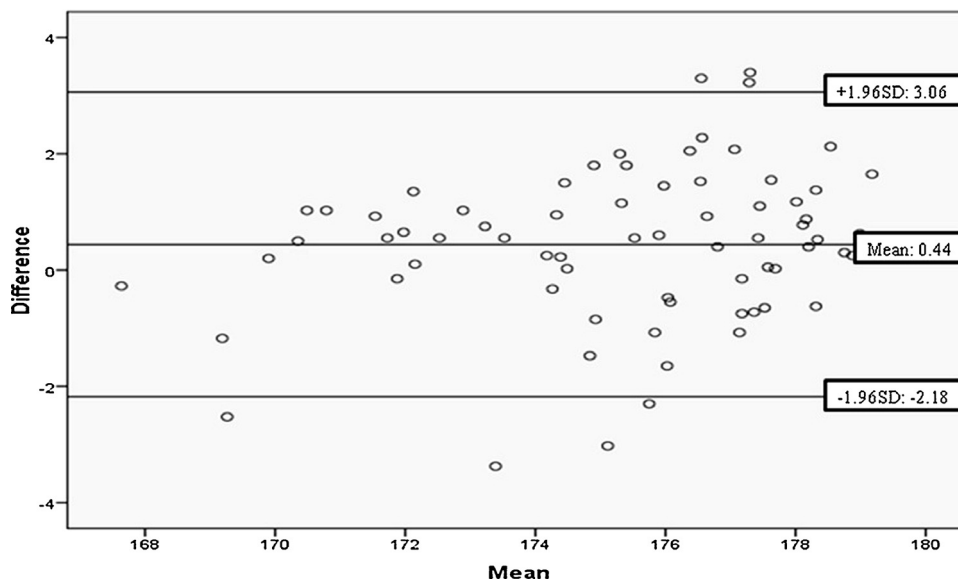

Fig. 3. Bland–Altman plot of alignment measurement by other observers and ideal (RV) for the mechanical axis (MFT angle).

ankle. There can be individual variations in finding the centre of these anatomic sites, and this can be attributed as the reason for inter-observer variability. Moreover, the identification of the tibial spines in the arthritic knee is difficult, which may improve with the experience of the observer. Identification of these anatomic landmarks grows with the experience of the surgeon and hence the two senior full-time consultants had no statistical difference between their angles. Even though all the observers were trained for identification of the anatomical landmarks before taking part in the study, the presence of a statistically significant difference in the mechanical axis may be due to a difference in the perception of the landmarks in the hip, knee and ankle.

The calculation of the mechanical axis by standing radiographs and PACS is an easy and less time-consuming way of planning the management of the arthritic knee. The inter-observer variability is an issue and can be attributed to the difference in identification of the centre of the hip, knee and ankle. The measurement shows good intra-class correlation. Any differences which remain in the measurement of the angle keep on decreasing as the experience of the observer increases. This difference can be due to a better understanding of the anatomy and better interpretation of the radiographs with experience.

There are certain limitations of the present study. The sample size was small.

Conflicts of interest

The authors have none to declare.

References

1. Teichtahl AJ, Cicuttini FM, Janakiramanan N, Wluka AE. Static knee alignment and its association with radiographic knee osteoarthritis. *Osteoarthr Cartil.* 2006;14: 958–962.
2. Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. *JAMA.* 2001;286:188–195.
3. Longstaff LM, Sloan K, Stamp N, Scaddan M, Beaver R. Good alignment after total knee arthroplasty leads to faster rehabilitation and better function. *J Arthroplasty.* 2009;24(4):570–578.
4. Holme TJ, Henckel J, Hartshorn K, Cobb JP, Hart AJ. Computed tomography scanogram compared to long leg radiograph for determining axial knee alignment. *Acta Orthopaed.* 2015;86(4):440–443.
5. Lonner JH, Laird MT, Stuchin SA. Effect of rotation and knee flexion on radiographic alignment in total knee arthroplasties. *Clin Orthop Relat Res.* 1996;331: 102–106.
6. Rauh MA, Boyle J, Mihalko WM, et al. Reliability of measuring long-standing lower extremity radiographs. *Orthopedics.* 2007;30:299–303.
7. Swanson KE, Stocks GW, Warren PD, Hazel MR, Janssen HF. Does axial limb rotation affect the alignment measurements in deformed limbs? *Clin Orthop Relat Res.* 2000;371:246–252.
8. Hankemeier S, Gosling T, Richter M, Hufner T, Hochhausen C, Krettek C. Computer-assisted analysis of lower limb geometry: higher intraobserver reliability compared to conventional method. *Comput Aided Surg.* 2006;11: 81–86.
9. Sailer J, Scharitzer M, Peloschek P, Giurea A, Imhof H, Grampp S. Quantification of axial alignment of the lower extremity on conventional and digital total leg radiographs. *Eur Radiol.* 2005;15:170–173.
10. Chao EYS, Sim FH. Computer-aided preoperative planning in knee osteotomy. *Iowa Orthop J.* 1995;15:4–18.
11. Prakash U, Wigderowitz CA, McGurty DW, Rowley DI. Computerised measurement of tibiofemoral alignment. *J Bone Joint Surg Am.* 2001;83:819–824.
12. Takahashi T, Yamanaka N, Komatsu M, Ogawa Y, Yoshida S, Yamamoto H. A new computer-assisted method for measuring the tibio-femoral angle in patients with osteoarthritis of the knee. *Osteoarthr Cartil.* 2003;12:256–259.

13. Sled EA, Sheehy LM, Felson DT, et al. Reliability of lower limb alignment measures using an established landmark-based method with a customized computer software program. *Rheumatol Int.* 2011;31:71–77.

14. Chang CB, Choi JY, Koh IJ, Seo ES, Seong SC, Kim TK. What should be considered in using standard knee radiographs to estimate mechanical alignment of the knee? *Osteoarthr Cartil.* 2010;18:530–538.

15. Specogna AV, Birmingham TB, DaSilva JJ, et al. Reliability of lower limb frontal plane alignment measurements using plain radiographs and digitized images. *J Knee Surg.* 2004;17(October (4)):203–210.

16. Hagstedt B, Norman O, Olsson TH, Tjörnstrand B. Technical accuracy in high tibial osteotomy for gonarthrosis. *Acta Orthop Scand.* 1980;51:963–970.

17. Mose K. Methods of measuring in Legg–Calve–Perthes disease with special regard to the prognosis. *Clin Orthop Relat Res.* 1980;150:103–109.

18. Liu HX, Shang P, Ying XZ, Zhang Y. Shorter survival rate in varus-aligned knees after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc.* 2015. [Epub ahead of print].

19. Lee YS, Kim MG, Byun HW, Kim SB, Kim JG. Reliability of the imaging software in the preoperative planning of the open-wedge high tibial osteotomy. *Knee Surg Sports Traumatol Arthrosc.* 2013. [Epub ahead of print].

20. Marx RG, Grimm P, Lillemoe KA, et al. Reliability of lower extremity alignment measurement using radiographs and PACS. *Knee Surg Sports Traumatol Arthrosc.* 2011;19(October (10)):1693–1698.

21. Rozzanigo U, Pizzoli A, Minari C, Caudana R. Alignment and articular orientation of lower limbs: manual vs computer aided measurements on digital radiograms. *Radiol Med.* 2005;109:234–238.

22. Goradia VK. Computer-assisted and robotic surgery in orthopedics: where we are in 2014. *Sports Med Arthrosc.* 2014;22(December (4)):202–205. http://dx.doi.org/10.1097/JSA.0000000000000047.

23. Cip J, Widemschek M, Luegmair M, Sheinkop MB, Benesch T, Martin A. Conventional versus computer-assisted technique for total knee arthroplasty: a minimum of 5-year follow-up of 200 patients in a prospective randomized comparative trial. *J Arthroplasty.* 2014;29(September (9)):1795–1802. http://dx.doi.org/10.1016/j.arth.2014.04.037.