Published in final edited form as:

Skeletal Radiol. 2017 February ; 46(2): 217–222. doi:10.1007/s00256-016-2539-8.

Test-Retest Reliability of Tibiofemoral Joint Space Width Measurements Made Using a Low-Dose Standing CT Scanner

Neil A. Segal, MD, MS¹, John Bergin, MA², Andrew Kern, MS², Christian Findlay, BFA², and Donald D. Anderson, PhD³

¹Professor, Department of Rehabilitation Medicine, The University of Kansas (Kansas City, KS)

²The University of Iowa (Iowa City, IA)

³Professor, Department of Orthopaedics and Rehabilitation, The University of Iowa (Iowa City, IA)

Abstract

Objective—To determine the test-retest reliability of knee joint space width (JSW) measurements made using standing CT (SCT) imaging.

Subjects and Methods—This prospective two-visit study included 50 knees from 30 subjects (66% female; mean $\pm$ SD age 58.2 ± 11.3 years; BMI 29.1 ± 5.6 kg/m²; 38% KL grade 0–1). Tibiofemoral geometry was obtained from bilateral approximately 20° fixed-flexed SCT images acquired at visits two weeks apart. For each compartment, the total joint area was defined as the area with a JSW <10mm. The summary measurements of interest were the percent of the total joint area with a JSW less than 0.5mm thresholds between 2.0mm–5.0mm in each tibiofemoral compartment. Test-retest reliability of the summary JSW measurements was assessed by Intraclass Correlation Coefficients (ICC 2,1) for the percent area engaged at each threshold of JSW and root-mean-square errors (RMSE) were calculated to assess reproducibility.

Results—The ICC were excellent for each threshold assessed, ranging from 0.95–0.97 for the lateral and 0.90–0.97 for the medial compartment. RMSE ranged from 1.1–7.2% for the lateral and from 3.1–9.1% for the medial compartment, with better reproducibility at smaller thresholds of JSW.

Conclusion—The knee joint positioning protocol used demonstrated high day-to-day reliability for SCT 3D tibiofemoral JSW summary measurements repeated 2 weeks apart. Low-dose SCT provides a great deal of information about the joint while maintaining high reliability, making it a suitable alternative to plain radiographs for evaluating JSW in people with knee OA.

Corresponding Author: Dr. Neil A. Segal, Department of Rehabilitation Medicine, University of Kansas Medical Center, 3901 Rainbow Boulevard, Mailstop 1046, Kansas City KS 66160, FAX: (913) 588-6765, segal-research@kumc.edu.

Compliance with Ethical Standards

Conflict of Interest: The corresponding author is named as an inventor on a patent application for the positioning frame used in this study, but has neither intellectual property rights nor financial rights to that patent and has not been employed or received honoraria, stock options, or other sources of financial support in relation to this work. The other authors report no conflicts of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Keywords

Knee osteoarthritis; imaging; computed tomography; reliability

Introduction

One of the principal pathognomonic features of knee osteoarthritis (OA) is loss of tibiofemoral joint space width (JSW) from damage to hyaline articular cartilage and meniscal tissue.[1] Measurement of JSW over time is the accepted standard for evaluation of knee OA progression.[2]

Plain radiography is considered the standard for imaging of OA because it is an inexpensive, rapid, and readily available technique.[1] Plain radiography also allows for a functional weight-bearing pose, which increases sensitivity to detection of narrowed JSW relative to unloaded imaging.[3] The tibiofemoral JSW on plain radiographic images is defined as the distance between the distal convex margin of the femur and the floor of the tibial plateau.[4] This radiographic measurement is the only structural outcome currently approved by the FDA for Phase III trials.[5]

However, plain radiographs are relatively insensitive, failing to detect narrowing of JSW for years after OA begins and requiring 2 or more years for detection of progression.[6] Radiographs require superimposition of the posterior and anterior edges of tibial plateaus to accurately determine JSW. Thus, the beam angle relative to the joint is of critical importance. JSW measurements can differ between radiographs taken during the same minute on the same knee imaged at slightly different angles.[7] Inappropriate beam angle can necessitate repetition of radiographs at multiple beam angles or fluoroscopic assistance. Even with repeated radiation exposures and calibration of the beam angle, reproducibility is negatively impacted by incorrect angulation or transverse rotation resulting from variability in patients and technologists.

Acquisition of fixed-flexion 3D images of the knees, using a standing CT (SCT) scanner with the same Relative Radiation Level as a set of knee radiographs, has been found to provide images unencumbered by overlapping anatomy[8] and may allow for more reliable JSW measurements, thereby enhancing capabilities to monitor disease activity. Earlier diagnosis of OA could lead to earlier intervention via medication or surgical management. However, there is a need to determine the reliability of JSW measurements using this new imaging modality prior to more widespread use.

The validity of JSW measurements using SCT in comparison with radiographic JSW and MRI assessments of cartilage and meniscal morphology has been reported.[9, 10] If this valid measurement also were found to have high reliability, then the weight-bearing 3D nature of SCT may be an attractive replacement for weight-bearing knee radiographs.[6, 8–10] Specifically, adequate reliability in weight-bearing JSW assessment would indicate that SCT imaging retains the advantages of knee radiography while circumventing its limitations. The purpose of this study was to determine the test-retest reliability of JSW measurements

using low-dose SCT imaging acquired 2 weeks apart in order to advance knowledge regarding the potential for improving assessment of tibiofemoral OA disease status.

Subjects and Methods

Subjects

Thirty subjects with a range of OA features (Kellgren/Lawrence grades 0–3) were recruited. Participants were age 40 years, had weight-bearing knee radiographs within the last four years, and reported ability to present for 2 SCT scans. No participants had knee replacements or were pregnant.

The investigators' Institutional Review Board (IRB) approved the study. The study was conducted between 5/28/2014 and 8/20/2014. All participants provided written informed consent following completion of an IRB-approved consent process.

SCT Image Acquisition

Tibiofemoral geometry was obtained from bilateral fixed-flexed SCT images acquired at visits two weeks apart. The SCT scanner was a modified version of an existing standing foot/ankle scanner (CurveBeam LLC, Warrington, PA) utilizing cone beam reconstruction and a PaxScan 3030D amorphous silicon flat-panel digital X-ray detector (Varian Medical Systems, Salt Lake City, Utah) over a 360° projection angle. The scans spanned 20cm with an isotropic resolution of 0.37mm and FOV of 200x350mm, using the manufacturer's standard settings (120 kVp, 5 mA). For the two visits, the estimated total effective radiation dose for each participant was 0.2 mSv, equivalent to the average environmental radiation of two weeks.

The acquisition protocol positioned the tips of the great toes, the patellae, and the anterior superior iliac spines coplanar to each other, resulting in knee flexion of approximately 20°. [11] This semiflexed pose, consistent with the Lyon-Schuss view, was chosen for its ability to maximize sensitivity and reliability.[11, 12]

Subjects removed footwear prior to stepping into a foot-positioning frame;[10] the toes and medial surface of the feet were placed against vertical platforms that externally rotated the feet by 10°. The skin over the patella and anterior thighs was in direct contact with the thigh-positioning plate. A coronal bar located anterior to the subject was positioned such that it made contact with the soft tissue overlying the anterior superior iliac spines. Additional parasagittal bars were placed on either side of the subject such that there was contact with the proximal lateral thighs. Participants were instructed to remain in contact with but not apply pressure to the bars in an attempt to prevent motion during the scan. Finally, subjects were instructed to remain motionless and focus on a single object, a sticker located at eye level on the wall in front of them, for the duration of the scan. After positioning was complete, each participant's safety and comfort was confirmed and the 2-minute scan was completed.

Image Analysis

DICOM data from each scan were imported into ITK-SNAP (www.itksnap.org)[13] where an operator, blinded to order of image acquisition, selected a 3D volume of interest around the tibiofemoral joint with a cropping tool. Tibial and femoral volumes were identified using a semi-automated watershed transform based algorithm implemented in MATLAB (The Mathworks, Natick, Massachusetts). Review ensured surfaces identified by the semi-automated algorithm corresponded with the bony outline and not radiographic artifacts. Discrepancies were corrected by manually tracing boundaries using either a computer mouse or an interactive pen display.[14]

A 3D triangulated surface mesh for each bone was exported from ITK-SNAP and imported into Geomagic Studio (Geomagic, Inc., Research Triangle Park, North Carolina). Angle deficiencies were removed to reduce voxelation artifact. Medial and lateral tibial plateaus were selected and saved as separate meshed surfaces to be analyzed independently. Plateaus were defined to be the area confined by the tibial spines and the anterior, posterior, lateral, and medial ridges of the tibia within the tibiofemoral space.

An algorithm for quantifying JSW was implemented in MATLAB. The center of each triangular element was located in 3D space for tibial plateau and femur articular surface meshes. For each center calculated on the tibial plateaus, the closest element center on the femur was located. The calculated distance between the two centers was defined as the JSW associated with that particular element on the tibial plateau. The area of each element on the tibial plateau was also calculated.

Statistical Analysis

For each tibiofemoral compartment, the calculated total area in apposition was defined as the area with a JSW less than 10mm. Test-retest reliability of the JSW measurements were assessed with Intraclass Correlation Coefficients (ICC 2,1) for the percent area in apposition at each 0.5mm threshold of JSW (from 2.0–5.0mm), using the Shrout-Fleiss single-score method,[15] with ICCs of less than 0.4 considered poor, ICCs between 0.4 and 0.75 considered fair to good and ICCs above 0.75 considered excellent. Precision of the ICC's was determined by calculation of 95% confidence intervals. Root-mean-square error was calculated to quantify measurement variability between the two imaging sessions in the measurements of % area with JSW less than or equal to each threshold. All statistical analyses were completed using SAS version 9.3 (SAS Inc., Cary, NC).

Results

Of the 60 knees enrolled, 50 knees (66% female; mean $\pm$ SD age 58.2 ± 11.3 years; BMI 29.1 ± 5.6 kg/m²) were used in final analyses. One subject did not return for follow-up (2 knees) and 8 knees were excluded for motion artifact that prevented accurate JSW measurements at either baseline or follow-up (2 knees due to Parkinsonian tremor, 2 knees due to body size mismatch with the SCT gantry, and 2 knees due to a subject moving during a scan). On pre-participation radiographs, 19 knees did not have radiologic evidence of OA (KL grade 0–1) and the remaining 31 had radiologic evidence of OA (KL grade 2–3).

Reliability and variability of the JSW measurement are presented in Table 1. A representative example of baseline and 2-week follow-up tibiofemoral JSW measurements is presented in Figure 1.

Discussion

Initial steps in determining whether SCT may be a suitable alternative to plain radiographs for evaluating knee OA include assessment of reliability, criterion validity, sensitivity, and specificity for detection of features of OA. This study focused on assessment of reliability of the JSW measurement when made in subjects with a continuum of knee OA features at visits two weeks apart. The knee joint positioning protocol used demonstrated high day-to-day reliability for measurement of 3D tibiofemoral JSW. The development of this 3D capability for reliably measuring JSW is significant in that low-dose SCT provides a great deal more information about both the tibiofemoral and patellofemoral joints than radiographs,[8] without a clinically significant increase in either acquisition time or radiation dose.

The use of CT imaging to circumvent the limitations of radiography for assessment of arthritis is not completely unique. Radiographic measurements of JSW for rheumatoid arthritis have similar limitations, leading some to develop more reliable positioning protocols and automated means of measurement using 3D CT imaging.[16] Because the JSW assessment used in our study was based upon the distribution of JSW over the whole joint, it was summarized as the percentage of the tibiofemoral area below each threshold for JSW. The reliability of this areal percentage, including approximately 30,000 JSW measurements per compartment, cannot be directly compared with the reliability for a single minimal JSW measurement on a 2D image such as a radiograph. However, the high reliability as well as reproducibility detected in this initial study suggests potential utility of this measurement and provides sufficient evidence to justify further exploration of 3D JSW assessment with SCT.

There are two principal factors that contributed to high reliability and reproducibility of the SCT assessment of knee JSW— careful subject positioning and identification of the bony margins. Our protocol involved positioning the knee in a similar manner to how it is rotated and flexed for standing radiographs. However, in addition to foot and knee positioning, we added anterior thigh, anterior pelvis, lateral hip and hand contact for stability and standardized positioning of the trunk and head (Figure 2A). While this can be completed in a matter of seconds, the greater body contact may have contributed to high reliability and reproducibility of the SCT assessment of knee JSW.

In contrast to these findings with SCT, Vignon et al have described a number of challenges in obtaining satisfactory reliability with radiographic measurement of knee JSW.[17] Overall, this group found that replication of medial tibial plateau alignment was poor, and that particularly with fixed-flexion radiographs, the poor replication “jeopardizes accurate evaluation of joint space narrowing.” While radiographic views suffer from difficulty in obtaining parallel alignment of the medial tibial plateau, SCT acquires a 3D dataset (Figure 2B–E) with an optimal view on every acquisition.

Narrowing of the JSW on radiographs is the current accepted means of assessing progression of knee OA in trials of disease modifying OA drugs. However, the limitations of this primary outcome measurement have contributed to clinical trials lasting greater than two years and requiring hundreds of subjects in each treatment arm, causing high expenses and difficulty in completing clinical trials. With additional development of the bone segmentation step, a summary measure of JSW, such as percent of joint space with a JSW < 2.5mm, could be automatically calculated at the time of an in-office scan, enabling efficient evaluation of joint space changes. For research studies of knee OA disease modification, this ability could provide advantages in evaluation of the tibiofemoral joint as well as providing an unencumbered view of the patellofemoral joint as well.

Radiation dose is always a consideration when selecting an imaging modality. A set of fixed-flexion knee radiographs, including two lateral images and a bilateral PA view has an effective dose of 0.04–0.05mSv and there is a potential need to repeat if beam angle requires correction. In comparison, the effective radiation dose for SCT of bilateral knees in our study was lower, at 0.024mSv. The incident air kerma value at the entrance (skin) plane for the SCT used in this study for two knees was 1.2 mGy, which compares favorably with both the 8–14 mGy absorbed dose at the center of the phantom reported for another upright cone beam CT scanner for a single knee and the absorbed dose at the center of a phantom reported for a conventional multidetector CT for one or two knees (27–40 mGy).[18] Evidence supports that SCT provides reliable imaging of weight-bearing knees with a similar radiation dose to radiographs. However, the potential for motion artifact could limit use in patients with tremor or inability to stand still for approximately 110 seconds in the lower body frame. Of note, following completion of this study, software was developed that reduces the potential for motion artifact.

Given the need to assess joint space narrowing, the next logical step will be assessment of responsiveness of SCT measurement of JSW to change over time. In comparison with fixed flexed radiographs, which have been reported to reproduce satisfactory positioning in only 11% of cases,[17] fluoroscopically-guided Lyon schuss radiographs have been found to be more responsive to loss of JSW over 12 months.[19] However, 3D SCT holds potential to improve responsiveness to change in JSW through improved reproduction of positioning, and could do so using less radiation than is required for fluoroscopic positioning.

Acknowledgments

The investigators appreciate the support of Ms. Maria Davis Hochstedler and Mr. Tom Baer for their help in facilitating key aspects of this research.

Funding: This work was supported by a National Heart, Lung and Blood Institute grant (T35HL007485). The sponsor had no role in the study design, collection, analysis and interpretation of data; in the writing of the manuscript; or in the decision to submit the manuscript for publication.

Funding: Research reported in this manuscript was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under award number P50AR055533. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. *Lancet*. 2011; 377(9783):2115–2126. [PubMed: 21684382]
2. Administration FaD. , editor. Clinical development programs for human drugs, biological products, and medical devices for the treatment and prevention of osteoarthritis. Food and Drug Administration; 1999.
3. Hellio Le Graverand MP, Mazzuca S, Duryea J, Brett A. Radiographic grading and measurement of joint space width in osteoarthritis. *Rheum Dis Clin North Am*. 2009; 35(3):485–502. [PubMed: 19931800]
4. Ravaud P, Giraudeau B, Auleley GR, Drape JL, Rousselin B, Paolozzi L, et al. Variability in knee radiographing: implication for definition of radiological progression in medial knee osteoarthritis. *Ann Rheum Dis*. 1998; 57(10):624–629. [PubMed: 9893575]
5. FDA. , editor. Federal Register. 2007 Aug 14. DOCKET NO. 1998D - 0077 (FORMERLY 98D - 0077) "Clinical Development Programs for Human Drugs, Biological Products, and Medical Devices for the Treatment and Prevention of Osteoarthritis"; p. 50-52.
6. Guermazi A, Roemer FW, Felson DT, Brandt KD. Motion for debate: osteoarthritis clinical trials have not identified efficacious therapies because traditional imaging outcome measures are inadequate. *Arthritis Rheum*. 2013; 65(11):2748–2758. [PubMed: 23861096]
7. Kinds MB, Vincken KL, Hoppinga TN, Bleys RL, Viergever MA, Marijnissen AC, et al. Influence of variation in semiflexed knee positioning during image acquisition on separate quantitative radiographic parameters of osteoarthritis, measured by Knee Images Digital Analysis. *Osteoarthritis Cartilage*. 2012; 20(9):997–1003. [PubMed: 22542633]
8. Segal NA, Nevitt MC, Lynch JA, Niu J, Torner JC, Guermazi A. Diagnostic performance of 3D standing CT imaging for detection of knee osteoarthritis features. *The Physician and sports medicine*. 2015; 43(3):213–220.
9. Segal NA, Frick E, Duryea J, Roemer F, Guermazi A, Nevitt MC, et al. Correlations of Medial Joint Space Width on Fixed-Flexed Standing Computed Tomography and Radiographs With Cartilage and Meniscal Morphology on Magnetic Resonance Imaging. *Arthritis Care Res (Hoboken)*. 2016; 68(10):1410–1416. [PubMed: 26991547]
10. Segal NA, Frick E, Duryea J, Nevitt MC, Niu J, Torner JC, et al. Comparison of tibiofemoral joint space width measurements from standing CT and fixed flexion radiography. *J Orthop Res*. 2016 In Press.
11. Vignon E, Piperno M, Le Graverand MP, Mazzuca SA, Brandt KD, Mathieu P, et al. Measurement of radiographic joint space width in the tibiofemoral compartment of the osteoarthritic knee: comparison of standing anteroposterior and Lyon schuss views. *Arthritis Rheum*. 2003; 48(2):378–384. [PubMed: 12571846]
12. Gossec L, Jordan JM, Mazzuca SA, Lam MA, Suarez-Almazor ME, Renner JB, et al. Comparative evaluation of three semi-quantitative radiographic grading techniques for knee osteoarthritis in terms of validity and reproducibility in 1759 X-rays: report of the OARSI-OMERACT task force. *Osteoarthritis Cartilage*. 2008; 16(7):742–748. [PubMed: 18417373]
13. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*. 2006; 31(3):1116–1128. [PubMed: 16545965]
14. McWalter EJ, Wirth W, Siebert M, von Eisenhart-Rothe RM, Hudelmaier M, Wilson DR, et al. Use of novel interactive input devices for segmentation of articular cartilage from magnetic resonance images. *Osteoarthritis Cartilage*. 2005; 13(1):48–53. [PubMed: 15639637]
15. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979; 86(2):420–428. [PubMed: 18839484]
16. Barnabe C, Buie H, Kan M, Szabo E, Barr SG, Martin L, et al. Reproducible metacarpal joint space width measurements using 3D analysis of images acquired with high-resolution peripheral quantitative computed tomography. *Med Eng Phys*. 2013; 35(10):1540–1544. [PubMed: 23684320]

17. Vignon E, Brandt KD, Mercier C, Hochberg M, Hunter D, Mazzuca S, et al. Alignment of the medial tibial plateau affects the rate of joint space narrowing in the osteoarthritic knee. *Osteoarthritis Cartilage*. 2010; 18(11):1436–1440. [PubMed: 20813195]
18. Carrino JA, Al Muhit A, Zbijewski W, Thawait GK, Stayman JW, Packard N, et al. Dedicated cone-beam CT system for extremity imaging. *Radiology*. 2014; 270(3):816–824. [PubMed: 24475803]
19. Le Graverand MP, Vignon EP, Brandt KD, Mazzuca SA, Piperno M, Buck R, et al. Head-to-head comparison of the Lyon Schuss and fixed flexion radiographic techniques. Long-term reproducibility in normal knees and sensitivity to change in osteoarthritic knees. *Ann Rheum Dis*. 2008; 67(11):1562–1566. [PubMed: 18258709]

Paired Tibiofemoral JSW Measurements for a Representative Knee

Fig. 1.
Paired Tibiofemoral JSW Measurements for a Representative Knee: JSW distributions plotted on the superior surface of the left tibia (anterior at top of figure)

Fig. 2. Positioning of Patient in Standing CT Scanner (A); 3D (B); Coronal (C); Sagittal (D); and Axial (E) images of a right knee.

Table 1

Test-Retest Reliability and Reproducibility of 3D JSW Measurements

Threshold (mm)	ICC for % of Compartment in Apposition with JSW Threshold	ICC 95%CI	RMS Error
Lateral Compartment			
5.0	0.97	0.94, 0.98	6.1%
4.5	0.97	0.94, 0.98	5.9%
4.0	0.95	0.92, 0.97	7.2%
3.5	0.95	0.91, 0.97	5.7%
3.0	0.97	0.94, 0.98	2.7%
2.5	0.97	0.94, 0.98	1.8%
2.0	0.97	0.95, 0.98	1.1%
Medial Compartment			
5.0	0.94	0.90, 0.97	7.7%
4.5	0.94	0.90, 0.96	9.1%
4.0	0.96	0.92, 0.97	7.8%
3.5	0.97	0.94, 0.98	6.6%
3.0	0.96	0.94, 0.98	5.6%
2.5	0.90	0.84, 0.94	5.8%
2.0	0.91	0.85, 0.95	3.1%

ICC=Intraclass Correlation Coefficient; 95%CI=confidence interval; RMS=root-mean-square