No Differences in Hip Joint Space Measurements Between Weightbearing or Supine Anteroposterior Pelvic Radiographs

Felipe S. Bessa, M.D., Brady T. Williams, M.D., Evan M. Polce, B.S., Bhargavi Maheshwer, B.S., Joel C. Williams, M.D., Shane J. Nho, M.D., M.S., and Jorge Chahla, M.D., Ph.D.

Purpose: To to assess whether there are any significant differences in hip joint space width (JSW) between weight-bearing versus supine pelvic radiographs. **Methods:** Standing and supine anteroposterior pelvic radiographs of 86 patients (146 hips) were included. Sample size was sufficiently powered to assess for equivalence between standing and supine films for JSW measurements made at the medial, lateral, and central aspects of the sourcil line. Measurements were made by 2 independent reviewers blinded to patient positioning. Each reviewer repeated a subset of the measurements to assess intra-rater reproducibility. Mean differences in joint space measurements between standing and supine radiographs were reported for each point of the sourcil. Intraclass correlation coefficients (ICCs) for inter and intra-rater reliability were also calculated. **Results:** There were no significant differences between JSW measurements made on standing and supine pelvic radiographs ($P = .468$). Furthermore, equivalence testing demonstrated statistical equivalence between standing and supine JSW measurements made based on an equivalence threshold of ±0.5 mm. Inter-rater reliability demonstrated good agreement with an overall ICC of 0.775 (95% confidence interval [CI] 0.734-0.809). Intra-rater reliability also demonstrated good agreement with ICCs of 0.84 (95% CI 0.758-0.889) and 0.798 (95% CI 0.721-0.851) for the 2 reviewers, respectively. **Conclusions:** JSW measurements on standing and supine pelvic radiographs were not significantly different, and their inter-rater agreement and intra-rater reproducibility demonstrated good reliability and repeatability. Therefore, either may be used to assess JSW, including measurements that may impact treatment decisions for hip arthroscopy. **Level of Evidence:** Level III; retrospective comparative study

Over the past 2 decades, the understanding of hip dysplasia and femoroacetabular impingement syndrome (FAIS) with regards to its pathomorphology, diagnosis, and treatment has improved dramatically, with numerous studies reporting excellent outcomes after arthroscopic and open treatments.[1-5] However, outcomes are not uniformly positive. This has led to better-defined indications and validation of preoperative prognostic factors. Through these efforts, multiple preoperative predictors of poor surgical outcome have been identified, including presence of osteoarthritis (OA), joint space narrowing, and a lateral center-edge angle <20°. The preoperative threshold joint space width (JSW) <2 mm has been found to be an important risk factor and predictor of poor outcomes after hip arthroscopy.[1-3,6] Patients with FAIS and labral pathology with less than 2 mm of joint space have significantly greater rates of conversion to total hip

From the Section of Young Adult Hip Surgery, Division of Sports Medicine, Department of Orthopedic Surgery (F.S.B.), Midwest Orthopaedics at Rush (B.T.W., E.M.P., B.M., S.J.N., J.C.), and Hip Preservation and Orthopedic Trauma (J.C.W.), Rush University Medical Center, Chicago, Illinois, U.S.A.; and Instituto Brasil de Tecnologias da Saúde (IBTS), Rio de Janeiro, Brazil (F.S.B.).

The authors report the following potential conflicts of interest or sources of funding: J.C. reports other from Arthrex and Smith & Nephew, outside the submitted work, and is an unpaid consultant for Arthrex, Smith & Nephew, and Conmed Linvatec. J.C.W. reports personal fees and other from Synthes, other from DePuy Synthes Sales, other from Medcial Device Business Service, and other from DePuy Orthopaedics, outside the submitted work. S.J.N. reports other from AlloSoucre, Arthrex, Athletico, DJ Orthopaedics, Linvatec, Miomed, Ossur, Smith & Nephew, Springer, and Stryker, outside the submitted work. Full ICMJE author disclosure forms are available for this article online, as supplementary material.

Received February 24, 2020; accepted July 3, 2020.

Address correspondence to Jorge A. Chahla, Department of Orthopaedics, Rush University Medical Center, Midwest Orthopaedics at Rush, 1611 W Harrison St., Chicago, IL 60612. E-mail: jachahla@msn.com

© 2020 by the Arthroscopy Association of North America

0749-8063/20212/$36.00

https://doi.org/10.1016/j.arthro.2020.07.009

arthroplasty following hip arthroscopy. Specifically, these patients have an odds ratio for conversion to total hip arthroplasty ranging from 6 to 9.29 (95% confidence interval [CI] 2.79-19.9) following hip arthroscopy when compared with patients with greater than 2 mm of hip JSW.[1]

Given this critical threshold, it is imperative that preoperative joint space measurements are accurate. Currently, the impact of patient positioning and weight-bearing during radiographic imaging (supine vs standing radiographs) on joint space measurements is not well established in the literature. If significant differences exist, this could subsequently mislead a surgeon in the decision-making process. In this regard, studies have reported decreased joint space in standing radiographs,[7] decreased joint space only in osteoarthritic hips in standing radiographs,[8] decreased joint space only in male patients on standing radiographs,[9] increased joint space in standing images,[10] and no difference in JSW between standing and supine anteroposterior (AP) pelvic radiographs.[11] However, some of the studies may not have been performed in patient samples with results that are generalizable to patients considered for hip arthroscopy, such as dysplastic hips previously treated by closed reduction.[11] Given this controversy and limitations of previous patient samples, the purpose of this study was to assess whether there are any significant differences in hip JSW between weight-bearing versus supine pelvic radiographs. It was hypothesized that there would be no significant difference in joint space measurements between weight-bearing and supine pelvic radiographs.

Methods

Eligibility Criteria

This study received institutional review board approval by the Rush University Institutional Review Board (ORA# 19103001). Patients who received weight-bearing (standing) and non-weightbearing (supine) AP pelvic radiographs, both performed on the same day of service, for the evaluation of hip pain and/or postoperative follow-up of hip surgeries (periacetabular osteotomies and proximal femur and acetabular fractures) between August 2016 and August 2019 at a single academic institution (Rush University Medical Center) were included in the study. Patients undergoing surgery for FAIS were not included. Hips with acetabular surgical hardware (e.g., periacetabular osteotomy); advanced OA (Tönnis grades 2 and 3[12]); Crowe II, III, and IV developmental hip dysplasia[13]; joint incongruency, radiographs without calibration; and radiographs of insufficient quality or resolution (as determined by consensus) were excluded, as these factors would have made measurements more difficult and unreliable, potentially resulting in erroneous values for JSW.

Radiograph Technique

All radiographs were performed at the same institution with the same image acquisition protocol. For both weight-bearing and supine AP pelvic radiographs, the X-ray tube was positioned at a distance of 40 inches from the film, and the patient was positioned such that the superior aspects of the iliac crest were included in the view. Patients were positioned with the hips in 15° of internal rotation to compensate for femoral neck anteversion. Digital x-ray images were visualized, and measurements were made using a picture archiving and communication system, Viztek Opal-RAD (Konica Minolta, Wayne, NJ).

Measurement Technique

Joint space measurements were performed in accordance with previous studies by Terjesen and Gunderson[11] and Mannava et al.[14] Measurements were performed at 3 distinct points along the acetabular sourcil line, perpendicular to the surface of the femoral head. The medial point was measured at the most medial point of the acetabular sourcil line, the central point was measured at the most proximal aspect of the femoral head, and the lateral point was measured from the most lateral aspect of the acetabular sourcil line (Fig 1). Additional measurements, including medial joint space measurements, were considered. However, due to concerns of reproducibility and susceptibility of medial measurements to error due the presence of the fovea and cotyloid fossa, these measurements were not included.

All measurements were performed by 2 independent authors with varying levels of medical training (resident and orthopaedic hip surgeon; B.T.W. and F.S.B.) under the guidance of the senior author (J.C.). Reviewers were blinded to patient positioning (standing vs supine) and to the measurements of the other reviewer for each radiograph. Both reviewers subsequently repeated a subset of measurements (30 patients, 47 hips) separated by a minimum of 2-weeks for the purpose of assessing intra-rater agreement and reproducibility.

Statistical Analysis

Sample size was determined a priori using RStudio (Boston, MA). Sample size calculations were performed for two one-sided tests to test for equivalence. Assuming a Type 1 error rate of 5%, power of 90%, and a standardized mean difference (Cohen's d) equivalence range of −0.3 to 0.3, a sample size of 121 hips (~61 AP pelvic radiographs) was calculated. Accounting for an approximately 20% prevalence of surgical hardware, advanced OA, and significant dysplasia that would result in exclusion, a target of 72 AP pelvic radiographs was used. A total of 86 AP pelvic radiographs were eligible for inclusion before exclusion criteria of surgical hardware, OA, and dysplasia were applied. The equivalence threshold and corresponding standardized

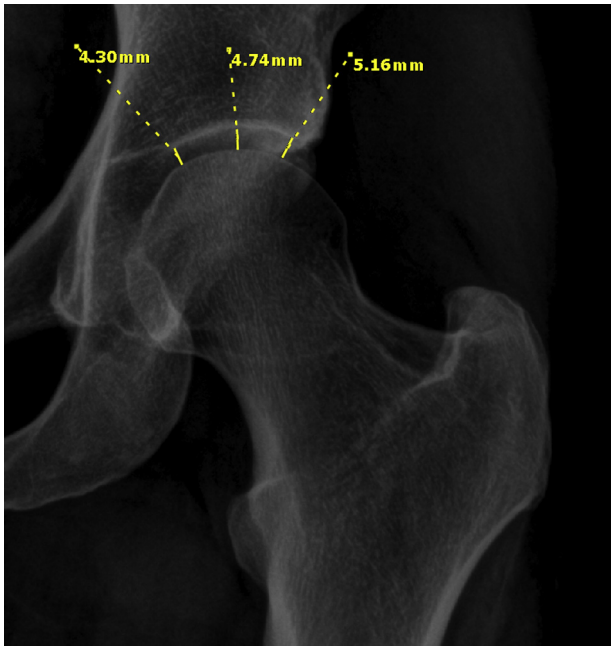

Fig 1. Left hip radiograph demonstrating the 3 distinct points of joint space measurement: (1) medial point at the most medial aspect of the acetabular sourcil line to the femoral head, (2) central point at the most central aspect of the acetabular sourcil line to the most proximal aspect of the femoral head, and (3) lateral point at the most lateral aspect of the acetabular sourcil line to the femoral head. All measurements were made perpendicular to the surface of the femoral head.

mean difference were based on previous literature that measured JSW on both standing and supine radiographs.[7] With a threshold of equivalence of ±0.5 mm and a standard deviation of 1.6 mm, an equivalence standard mean difference (Cohen's d) range of ±0.3 was determined and used in sample size calculations. The JSW in weightbearing and supine pelvic radiographs were compared at the lateral, central, and medial points of the acetabular sourcil. An overall difference in standing versus supine measurements was also reported pooling all measurements from the 3 distinct points.

Intraclass correlation coefficients (ICCs) were calculated to compare the measurements of both authors (inter-rater reliability), according to the guidelines by Koo and Li.[15] To assess measurement reproducibility (intra-rater reliability), ICCs were calculated using the repeated measurements from each author. For the purpose of reporting results, JSW values measured by the author with the highest overall intra-rater reliability ICC were used.

Equivalence testing determined whether JSW in standing and supine pelvic radiographs could be considered equivalent. Both 90% and 95% CIs were constructed for the lateral, central, and medial points and an overall pooled measurement with a type I error rate of 0.05. The calculated CIs were then compared with a minimal clinically important threshold of ±0.5 mm. If the CIs did not exceed the minimal clinically important threshold, then the measurements in standing and supine pelvic radiographs were considered statistically equivalent. Normality was assessed using the Shapiro–Wilk test. The non-parametric Mann–Whitney *U* test and Kruskal–Wallis test were used to compare differences in weight-bearing and supine JSW with respect to the all measurements and at the 3 distinct positions along the acetabular sourcil line. Statistical analysis was performed using the computing software R (R version 1.2.1335; R Foundation for Statistical Computing, Vienna, Austria). Equivalence testing was computed using the TOSTER package.[16]

Results

A total of 832 consecutive AP pelvic radiographs were identified, of which 86 (10.33%, 172 hips) were obtained in both weight-bearing and supine positions. Twenty-six hips were excluded: 4 presented with advanced hip developmental dysplasia, 6 presented with advanced osteoarthritis, 2 contained hardware from previous acetabular fracture, and 14 contained hardware from previous periacetabular osteotomy. Ultimately, 146 hips were included in the study for analysis.

The overall inter-rater reliability ICC was 0.775 (95% CI 0.734-0.809), demonstrating good reliability.[15] The overall intra-rater ICCs were 0.840 (95% CI 0.758-0.889) and 0.798 (95% CI 0.721-0.851), for the 2 reviewers. The measurements with the greatest inter-rater reliability were the central measurement on standing films (ICC 0.842) and the lateral point on supine films (ICC = 0.826). The medial joint space measurements demonstrated the lowest agreement between reviewers on both standing and supine films (ICC = 0.730, and 0.687, respectively). The overall ICC and the ICC for each distinct point of the acetabular sourcil for inter-rater and intra-rater reliabilities are presented in Table 1. The measurements made by the author with the highest overall ICC (0.840) were used to report the final results of this study.

There were no statistically significant differences between joint space measurements in standing and supine pelvic radiographs overall ($P = .468$) or when separately compared for the 3 distinct points along the sourcil line ($P = .903$). The median values of JSW for lateral, central, and medial points in the standing position were 4.49 mm, 4.14 mm, and 4.2 mm, respectively. In the supine position, the median values of JSW were 4.51 mm, 4.29 mm, and 4.11 mm, respectively. A boxplot was constructed to depict the median, interquartile range, and outlier values for the JSW measurements with respect to orientation (Fig 2). In addition, JSW measurements obtained from

Table 1. Inter- and Intra-Rater Reliability With 95% CI for Overall, Lateral, Central, and Medial Points for Both Standing and Supine Radiographs

	Inter-Rater Reliability ICC (95% CI)	Intra-Rater Reliability (B.T.W.) ICC (95% CI)	Intra-Rater Reliability (F.S.B) ICC (95% CI)
Overall	0.775 (0.734-0.809)	0.840 (0.758-0.889)	0.798 (0.721-0.851)
Standing, lateral point	0.75 (0.669-0.814)	0.795 (0.655-0.882)	0.814 (0.688-0.893)
Standing, central point	0.842 (0.762-0.892)	0.840 (0.658-0.92)	0.671 (0.477-0.803)
Standing, medial point	0.730 (0.635-0.802)	0.816 (0.649-0.902)	0.791 (0.548-0.896)
Supine, lateral point	0.826 (0.765-0.872)	0.846 (0.738-0.912)	0.784 (0.617-0.88)
Supine, central point	0.786 (0.705-0.845)	0.831 (0.607-0.919)	0.789 (0.648-0.878)
Supine, medial point	0.687 (0.586-0.766)	0.825 (0.641-0.91)	0.723 (0.512-0.846)

CI, confience interval; ICC, intraclass correlation coefficient.

weight-bearing and supine radiographs demonstrated statistical equivalence at the lateral, central, and medial positions, and when pooled for overall comparison (all $P < .001$) as determined by two one-sided tests (Fig 3).

Discussion

The main finding of this study was that there was no significant difference between JSW measurements made on weight-bearing and supine AP pelvic radiographs in patients without significant hip deformity. The study further demonstrated that these measurements are reliable and have good agreement both between reviewers and across repeated trials. Given the importance of preoperative joint space measurements and their role in decision-making of which patients are candidates for hip arthroscopy and which are not,[3,6] this study provides valuable information for clinical care.

The findings of the present study are in contrast to what has been reported by Fuchs-Winkelmann et al.,[7] who reported decreased joint space on standing pelvic radiographs compared with supine films. In their study, 61 patients (71 hips) who had undergone previous closed reduction for developmental dislocation of the hip were assessed, showing a significant decrease ($P < .001$) in JSW in standing radiographs (2.24 ± 1.49 mm) when compared with supine images (2.73 ± 1.63 mm). This difference could be explained by the fact that hips evaluated by Fuchs−Winkelmann et al. already presented with some degree of degenerative changes, whereas the hips in the present study did not. This is similar to the findings of Conrozier et al.,[8] who reported decreased JSW in osteoarthritic hips, and no difference in non-osteoarthritic hips when comparing standing with supine pelvic radiographs. In the study by

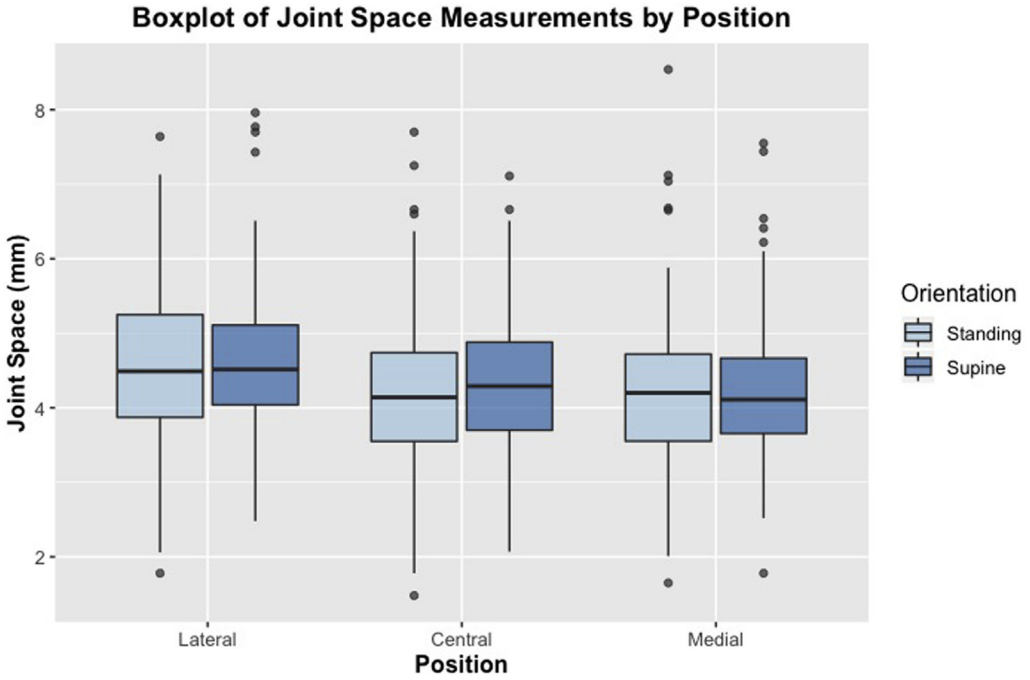

Fig 2. Boxplot of the median, interquartile range, and outlier values of joint space width measurements for the lateral, central, and medial points for both supine and standing radiographs.

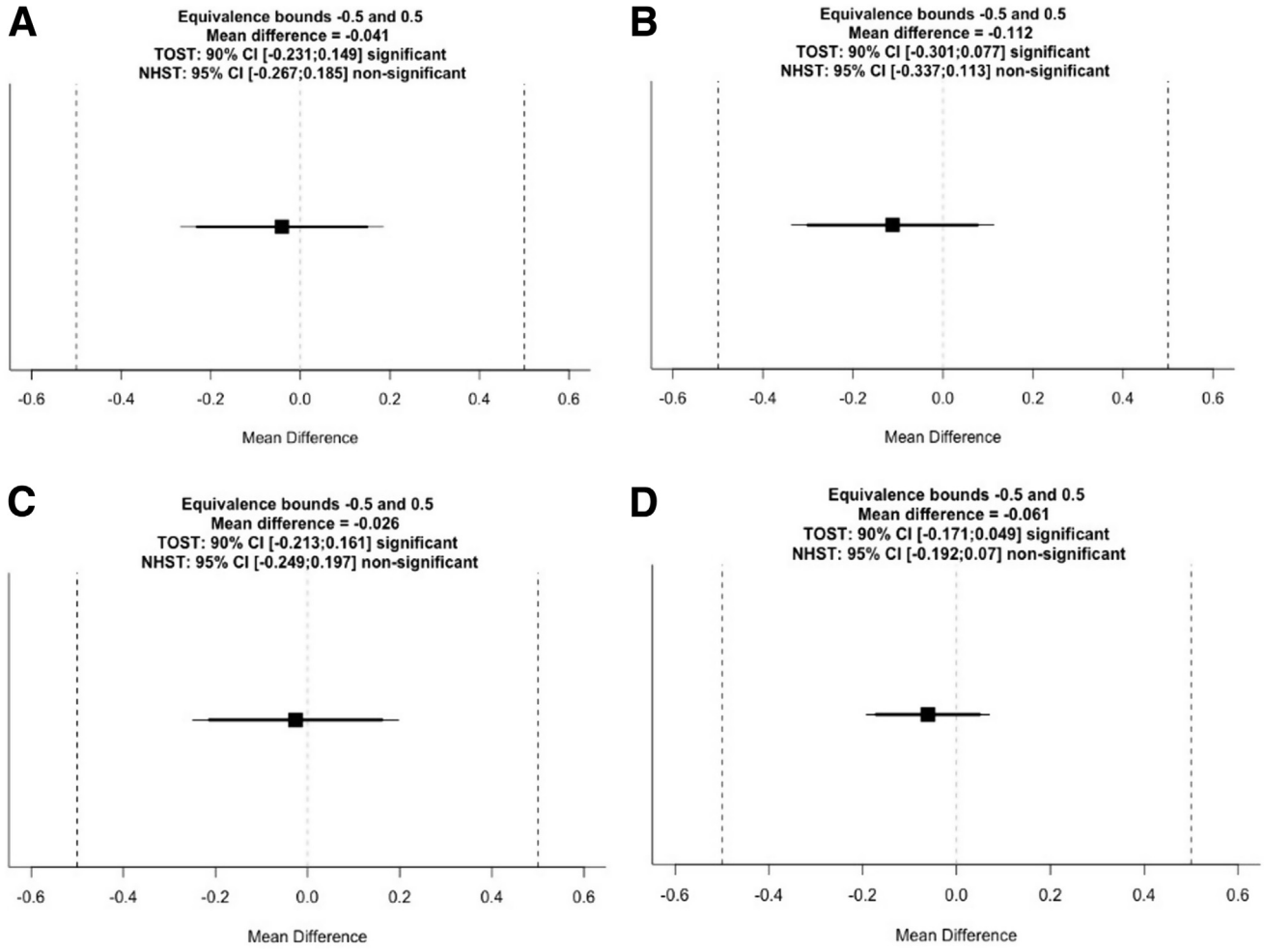

Fig 3. Equivalence plots comparing mean difference in millimeters at the (A) lateral, (B) central and (C) medial positions and (D) the overall pooled mean differences between weightbearing and supine pelvic radiographs. The bold and light horizontal lines represent the 90% and 95% confidence intervals, respectively. The black square indicates mean difference. The vertical dashed lines constitute the minimal clinically important threshold (±0.5 mm). (CI, confidence interval; NHST, null hypothesis significance testing; TOST, two one-sided tests.)

Fuchs—Winkelmann et al., the measurements were made twice by 2 independent reviewers, with good intra-rater reliability (ICCs = 0.9 and 0.86) and inter-rater reliability (ICC = 0.78), but there was no mention to whether reviewers were blinded or not to the radiographic position, which could be a bias generator. In the second study, by Conrozier et al., the measurements were performed by a single reviewer blinded to the x-ray positions. The findings of the present study are corroborated by Terjesen and Gunderson,[11] who reported no differences between JSW measured on both standing and supine pelvic radiographs in 63 dysplastic hips, treated by closed reduction during childhood (mean age at reduction: 19 months). However, a power analysis was not performed and investigators were not blinded to radiograph position, which could be sources of bias to the study. These are specific limitations that the current study improves upon including appropriately powered equivalence testing, 2 independent reviewers blinded to radiographic position, and most importantly, a patient sample without significant anatomic abnormalities such as dysplasia or significant joint degeneration.

The importance of being able to accurately and reproducibly measure hip joint space has been demonstrated by several clinical studies. Many studies have reported poor outcomes and increased risk for conversion to total hip arthroplasty after arthroscopic hip surgeries in patients with a joint space <2 mm. Philippon et al.[6] followed a series of 153 patients older than the age of 50 years undergoing hip arthroscopy for FAIS. In addition to lower modified Harris Hip Scores (<60), patients with preoperative joint space < 2mm were at significantly greater risk of conversion to total hip arthroplasty (adjusted odds ratio 14.6, 95% CI 5.1-41.83) using multivariate regression analysis. Similarly, Larson et al.[2] followed 210 patients who underwent hip arthroscopy for FAIS, 154 patients (169 hips) without joint space narrowing (Tönnis

grades 0 and 1), and 56 patients (58 hips) with joint space narrowing (Tönnis grades 2 and 3). Greater failure rates were observed in patients with joint space narrowing compared with those with preserved joint space (52% vs 12%, respectively).[2] Lebus et al.[3] followed 317 patients who underwent arthroscopic labral reconstruction, reporting joint space ≤2 mm as an important risk factor for conversion to THR (no odds ratio reported) and joint space >2 mm as a predictor of postoperative better outcomes and increased satisfaction. Given the important prognostic ability of preoperative joint space measurements, the findings of the present study have high clinical relevance for the orthopaedic surgeon when deciding whether to perform a hip-preserving surgery in a patient with a borderline JSW. Specifically, the findings suggest that joint space measurements can be performed on either standing or supine films, without concern that patient position will result in an under or overestimation of JSW.

Limitations

The results of the current study should be interpreted in the context of a few limitations. Foremost, radiographs were limited to subjects that had both standing and supine films. This study included patients with a variety of hip pathologies; however, radiographs were screened and excluded for pathologies (e.g., dysplasia, advanced OA) and surgical hardware (e.g., periacetabular osteotomy) that would potentially confound JSW measurements. In addition, while the measurements were digitally performed and presented good reliability, they were conducted manually, thus potentially contributing to variability in the measured values. This was particularly evident in the measurement of the JSW at the medial point of the sourcil line. The medial point was not as well defined on the acetabular side, and potentially obscured on the femoral side by structures including the fovea. As a consequence, the lowest inter-rater reliability was seen for these measurements on standing and supine films.

Conclusions

Joint space width measurements on standing and supine pelvic radiographs were not significantly different, and their inter-rater agreement and iinter-rater reproducibility demonstrated good reliability and repeatability. Therefore, either may be used to assess joint space width, including measurements that may impact treatment decisions for hip arthroscopy.

References

1. Degen RM, Nawabi DH, Bedi A, Kelly BT. Radiographic predictors of femoroacetabular impingement treatment outcomes. *Knee Surg Sports Traumatol Arthrosc* 2017;25:36-44.
2. Larson CM, Giveans MR, Taylor M. Does arthroscopic FAI correction improve function with radiographic arthritis? *Clin Orthop Relat Res* 2011;469:1667-1676.
3. Lebus GF, Briggs KK, Dornan GJ, McNamara S, Philippon MJ. Acetabular labral reconstruction: Development of a tool to predict outcomes. *Am J Sports Med* 2018;46:3119-3126.
4. Nwachukwu BU, Rebolledo BJ, McCormick F, Rosas S, Harris JD, Kelly BT. Arthroscopic versus open treatment of femoroacetabular impingement: A systematic review of medium- to long-term outcomes. *Am J Sports Med* 2016;44:1062-1068.
5. Öhlin A, Ahldén M, Lindman I, et al. Good 5-year outcomes after arthroscopic treatment for femoroacetabular impingement syndrome. *Knee Surg Sports Traumatol Arthrosc* 2019;28:1311-1316.
6. Philippon MJ, Schroder ESBG, Briggs KK. Hip arthroscopy for femoroacetabular impingement in patients aged 50 years or older. *Arthroscopy* 2012;28:59-65.
7. Fuchs-Winkelmann S, Peterlein CD, Tibesku CO, Weinstein SL. Comparison of pelvic radiographs in weightbearing and supine positions. *Clin Orthop Relat Res* 2008;466:809-812.
8. Conrozier T, Lequesne MG, Tron AM, Mathieu P, Berdah L, Vignon E. The effects of position on the radiographic joint space in osteoarthritis of the hip. *Osteoarthritis Cartilage* 1997;5:17-22.
9. Troelsen A, Jacobsen S, Rømer L, Søballe K. Weightbearing anteroposterior pelvic radiographs are recommended in DDH assessment. *Clin Orthop Relat Res* 2008;466:813-819.
10. Auleley GR, Rousselin B, Ayral X, Edouard-Noel R, Dougados M, Ravaud P. Osteoarthritis of the hip: Agreement between joint space width measurements on standing and supine conventional radiographs. *Ann Rheum Dis* 1998;57:519-523.
11. Terjesen T, Gunderson RB. Reliability of radiographic parameters in adults with hip dysplasia. *Skeletal Radiol* 2012;41:811-816.
12. Kovalenko B, Bremjit P, Fernando N. Classifications in brief: Tönnis classification of hip osteoarthritis. *Clin Orthop Relat Res* 2018;476:1680-1684.
13. Crowe JF, Mani VJ, Ranawat CS. Total hip replacement in congenital dislocation and dysplasia of the hip. *J Bone Joint Surg Am* 1979;61:15-23.
14. Mannava S, Geeslin AG, Frangiamore SJ, et al. Comprehensive clinical evaluation of femoroacetabular impingement: Part 2, plain radiography. *Arthrosc Tech* 2017;6:e2003-e2009.
15. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155-163.
16. Lakens D. Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Soc Psychol Personal Sci* 2017;8:355-362.