The Reliability of Commonly Used Radiographic Parameters in the Evaluation of the Pre-Arthritic Hip

A Systematic Review

Dominic S. Carreira, MD

Brendan R. Emmons, BS

Investigation performed at Peachtree Orthopedics, Atlanta, Georgia

Abstract

Background: Radiographs play a central role in the evaluation of the extent of arthritic change and the quantification of morphological parameters characteristic of femoroacetabular impingement (FAI) and/or dysplasia. Reporting of these parameters is widespread in the literature, occasionally as an instrument for data pooling in collaborative research endeavors. Yet, the metrological properties of commonly used radiographic descriptors of acetabular and femoral architecture relevant to the field of hip preservation have not been reviewed systematically. The purpose of the present review was twofold: (1) to identify the current body of data available on the reliability of commonly used radiographic descriptors of the pre-arthritic hip and (2) to collate best practices and methodological pearls and pitfalls from the existing investigations that have explored the reliability of these radiographic measures.

Methods: Article screening methodology involved an extensive search of 2 databases (PubMed Central, Embase). Metrological reliability data and best practices and frequencies with which these practices have been reported in existing investigations were extracted from the methods of the included studies. The radiographic parameters that were chosen as a focus for the present review were based on a discussion among a number of experienced hip-preservation surgeons.

Results: Forty-three articles met all inclusion criteria and were included in this review. Eleven recurring components of study design were isolated from the included investigations as being critical for minimizing bias in the assessment of the reliability of these radiographic morphological parameters. Sixteen (37.2%) of the included studies reported the use of at least 1 quality-control measure for ensuring radiographic integrity. The radiographic parameters with the most reliability data included the lateral center-edge angle (22 studies), the Tönnis angle (14 studies), and the alpha angle (18 studies). Categorical markers of acetabular retroversion and the extent of osteoarthritis have been less-frequently studied with respect to their reliability, with the exception of the crossover sign, which was evaluated in 11 of the included studies.

Conclusions: The radiographic parameter with the most metrological evidence in support of its reliability was the lateral center-edge angle (with

Disclosure: No external funds were received in support of this work. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work (http://links.lww.com/JBJSREV/A422).

22 [100%] of 22 studies reporting acceptable reliability), whereas the Tönnis grade of osteoarthritis consistently demonstrated subpar interrater and intrarater reliability. Side-arm investigations for testing reliability may be needed by individual studies to prove the validity of data-pooling practices as collaborative endeavors continue to become more widespread.

In recent years, there has been a growing recognition of the pathophysiology of intra-articular structural damage of the hip in patients without advanced osteoarthritic changes. These pre-arthritic abnormalities often consist of more focal soft-tissue lesions such as chondrolabral damage and ligamentum teres injury. Morphological osseous impingement lesions and structural instability often are associated with the development of these intra-articular injuries and may explain the pathomechanism behind the development of osteoarthritis. As our ability to treat these structural abnormalities has grown, surgical intervention for the pre-arthritic hip has become increasingly widespread.

A comprehensive clinical evaluation of the pre-arthritic hip consists of assessment with use of several modalities, including patient-reported symptoms, history, physical examination, and imaging[1-3]. While advanced imaging tools such as computed tomography (CT) and magnetic resonance imaging/arthrography (MRI/MRA) commonly play a role as secondary diagnostic modalities, the low cost, widespread availability, and speed of acquisition of radiographs allow for their central place as a primary diagnostic instrument in standard practice[4]. Radiographs permit evaluation of the extent of arthritic change and the quantification of morphological parameters characteristic of femoroacetabular impingement (FAI) and/or dysplasia. Additionally, when surgery is performed, radiographs often play a role in preoperative planning and surgical decision-making, given the well-understood consequences of inadequate treatment of both FAI and dysplasia[5,6].

The role of arthroscopy versus periacetabular osteotomy for the treatment of recalcitrant pain in the setting of instability is often weighed preliminarily on the basis of radiographic findings[7]. Given the implications of these radiographic readings, commonly used radiographic quantitative and categorical descriptors of pelvic morphology should be reproducible both from one provider to the next and by a single provider. To our knowledge, the metrological properties of commonly used radiographic descriptors of acetabular and femoral architecture relevant to the field of hip preservation have not been reviewed in a systematic way. Additionally, given the ever-growing role of collaborative and multicentered efforts in broadening our understanding of FAI, hip instability, and orthopaedics more broadly, there is a need for both (1) a repository of metrological data on the subject of reliability and (2) a comprehensive effort to collate best practices and pitfalls in the assessment of reliability in the radiographic evaluation of the pre-arthritic hip[8-13]. Many of these multicentered investigations have relied on pooling of patient data that are later categorized into subgroups on the basis of quantitative and qualitative radiographic parameters. Appropriately designed side-arm investigations for testing reliability may be needed to prove the validity of these data-pooling practices by individual studies.

The purpose of the present review was twofold: (1) to identify, collate, and evaluate the current body of data available on the reliability of commonly used radiographic descriptors of the pre-arthritic hip and (2) to collate best practices and methodological pearls and pitfalls from the existing investigations that have explored the reliability of these radiographic measures.

Materials and Methods

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement was used to design the study, and article selection proceeded in accordance with the multi-tiered system defined by that statement[14]. Articles were derived from a search of 2 databases: PubMed Central and Embase. The search strategy (see Appendix) was employed on January 12, 2018.

In order to capture parameters thought to be important in the evaluation of the pre-arthritic hip in the setting of impingement or dysplasia, an ongoing multicenter study group of high-volume hip arthroscopists (defined as those performing >100 hip arthroscopies annually and having performed a minimum of 300 hip arthroscopies) were queried as to the radiographic measures that they believed to be essential to a comprehensive evaluation. After multiple discussions, the following parameters were chosen and are a focus for the present review: lateral center-edge angle (Wiberg and Ogata et al.)[15,16], anterior center-edge angle on a false-profile radiograph[17], Tönnis angle[18], neck-shaft angle[19], alpha angle (as assessed on a modified Dunn, Dunn, anteroposterior, cross-table lateral, frog-leg lateral, or false-profile radiograph)[20], crossover sign[21], posterior wall sign[21], ischial spine sign[22], and Tönnis grade of osteoarthritis[18].

In order to be included in the present review, a study had to meet all of the following criteria: (1) peer-reviewed analysis, (2) availability of the full text of the study in the English language, (3) evaluation of radiographs that were made no earlier than 2001 (as radiographs made prior to that time did not meet Digital Imaging and Communications in Medicine [DICOM] standards), (4) assessment of interrater and/or intrarater reliability for at least 1 of the predetermined parameters on radiographs, and, when applicable, (5) assessment of interrater and/or intrarater reliability for continuous, quantitative metrics with use of the intraclass correlation coefficient (ICC). Studies that failed to meet any of the inclusion criteria were excluded. Additionally, studies were excluded if (1) the reliability of the predetermined parameter(s) was assessed in a pediatric population with cerebral palsy or Legg-Calvé-Perthes disease or in a population with advanced osteoarthrosis (Tönnis grade >2), a previous hemiarthroplasty, or a previous total hip arthroplasty, (2) the mean age of the study population was <16 years, (3) the reliability of making a diagnosis (i.e., cam impingement, pincer impingement, dysplasia) was analyzed without defining the reproducibility of a specific radiographic measurement or sign, (4) the reliability of a number of different radiographic measurements/signs was tested but the range of the values was reported without assigning a quantity to each individual measurement, (5) the intermodality correlation of a measurement was evaluated (i.e., CT vs. radiography) without reporting on the reproducibility of the measurement within a single imaging modality.

A pilot review of several studies exploring the reliability of radiographic parameters in this pre-arthritic population suggested that the ICC was the most widely used tool for analyzing the reliability of continuous, quantitative measures[23-26]. While other methods have been described in the literature (e.g., Bland-Altman plots, mean interobserver differences, Pearson correlation coefficients, concordance correlation coefficients), only including studies reporting ICCs was deemed helpful for facilitating qualitative synthesis and comparison.

Two reviewers reached complete consensus on the inclusion and exclusion criteria and ultimately the list of studies that were included in the present review. Studies that were evaluated for eligibility in this review were not limited to those that were returned by the combined database search output, as the references from the included articles were also eligible for full-text review. All articles that met the inclusion criteria were evaluated for data points thought to be critical to a systematic presentation of interrater and intrarater reliability data. In extracting the "level of experience" of the reviewers participating in radiographic measurements in each of the included studies, investigator initials were used as a starting point for further research into investigator background when level of experience was not explicitly stated in the article. While there is no applicable and validated quality-assessment tool for the quantification of risk of bias in studies directed toward reporting on the reliability of a quantitative or categorical variable, factors thought to be important to a well-designed study of reliability were extracted from the methodologies of individual articles and were collated for presentation.

Quantitative pooling of the reliability data was desired but was deemed inappropriate given the lack of detail in many of the included studies with regard to the finer nature of the statistical methods employed (2-way random model ICC versus 1-way random model ICC, Cohen kappa coefficient versus weighted Cohen kappa coefficient, etc.). In order to facilitate qualitative synthesis of study data, despite its obvious limitations as a series of arbitrary divisions, an ICC or Cohen kappa coefficient of 0 to 0.20 was deemed low; 0.21 to 0.40,

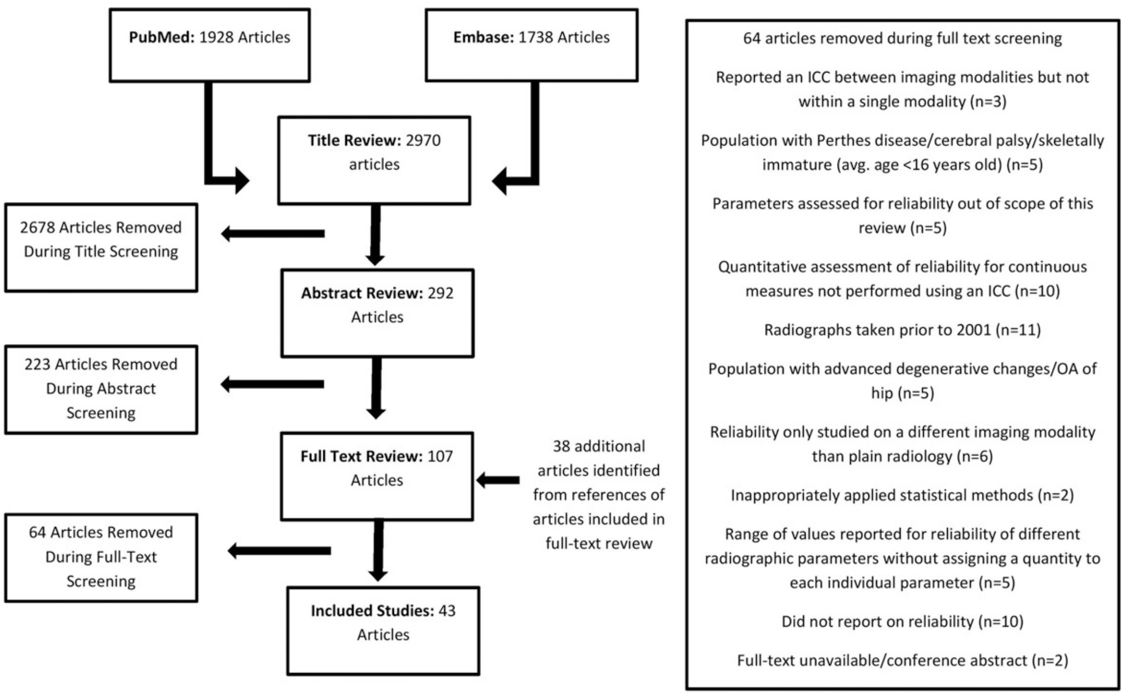

Fig. 1
PRISMA diagram illustrating the screening procedure for study inclusion. OA = osteoarthritis.

TABLE I Compilation of Best Practices Modeled by the Included Studies

Study Feature	No. of Studies Reporting
1. Removal of identifying information from radiographs/blinding of reviewer to clinical history of patient[23,25,29-31,33-37,41,44,45,47,63,65]	16 (37.2%)
2. Reviewers uninvolved in recruitment of the subjects whose radiographs are to be studied[29,31]	2 (4.7%)
3. Selection of radiographs for reproducibility calculations should be via randomization or from a consecutive series of patients so as to minimize selection bias[23,31,34,35,43,44,46,53,54,56,60,62]	12 (27.9%)
4. Review of radiographs at different time points for intrarater reliability tabulation should be separated by a minimum of 1 week[23-25,29-31,33-36,39,41-44,46,52,54,55,59-63,65,67]	26 (81.3%)*
5. Order of images at the time of second viewing for the intrarater reliability tabulation should be randomized with respect to the initial order of viewing[23,25,29,31,33-35,67]	8 (25.0%)*
6. Reviewers operate independently of one another, blinded to the measurements/grades of other study participants[29,30,32-35,37,39,41,43-45,50-52,54,57,60-62,64,66,67]	23 (53.5%)
7. Reviewers participating in a second review for assessment of intrarater reliability are not permitted to access measurements/grades obtained at the time of the first evaluation[29,34,35]	3 (9.4%)*
8. Radiographs included in the study are best obtained at a single center by a single radiographic technician operating in accordance with a standardized imaging protocol[41,63]	2 (4.7%)
9. If multiple reviewers participate in assessment at >1 time point, interrater reliability is best assessed at the time of initial viewing so as to mitigate against practice-effects bias[31,34-36]	4 (16.7%)†
10. Hold a consensus-building discussion prior to measurement where standard procedures/measuring techniques can be discussed[25,31,34,35,41,44,45,50]	8 (18.6%)
11. In the assessment of intraobserver reliability, more than a single rater ought to continue on to make the repeat set of measurements[23,24,29-31,33-39,41,43,46,49,53,54,58-60,62,63,66]	24 (75.0%)*

*Thirty-two of the 43 studies assessed intrarater reliability; thus, the percentage is based on a denominator of 32. †Twenty-four of the 43 studies assessed reliability with multiple reviewers at multiple time points; thus, the percentage is based on a denominator of 24.

fair; 0.41 to 0.60, moderate; 0.61 to 0.80, substantial; and 0.81 to 1.00, almost perfect[27,28]. A threshold for an ICC or Cohen kappa coefficient of ≥0.60 was utilized for consolidating and interpreting the vast array of quantitative data. The percentages of investigations attaining this threshold for each measurement were reported.

Results

Search Results

After the removal of duplicate articles, 2,970 articles were returned from the combined database search, of which 69 were deemed eligible for full-text review. At the time of full-text review of the reference sections of these articles, an additional 38 studies that had not been captured by the combined database search were identified, and, on the basis of title, were thought to potentially satisfy the review's inclusion criteria. Consequently, 107 articles were evaluated in their entirety in order to assess their concordance with the predetermined inclusion criteria. Ultimately, 43 articles met all inclusion criteria and were included in this review (Fig. 1)[23-26,29-67].

Compilation of Best Practices

The methodologies of the 43 included articles were highly variable with respect to both the level of detail and the rigor of design. Eleven recurring components of study design were isolated from the included investigations as being critical for minimizing bias in the assessment of the reliability of these radiographic morphological parameters. The most commonly reported consideration was separating the initial and second viewings for the assessment of intrarater reproducibility by a minimum of 1 week, reported in 26 (81.3%) of the 32 articles assessing intrarater reliability. The remaining considerations thought to be critical to a well-designed study for assessing reliability and the frequency with which these considerations were reported are presented for reference in Table I.

Screening Protocols for Ensuring the Integrity of Radiographs

Quality control of the included radiographs was reported as a component of several of the included investigations. In order to account for the effects of pelvic tilt and rotation on the assessment of morphological characteristics of the pelvis on an anteroposterior view, a number of screening protocols were employed, the most common of which was an evaluation of the distance between the sacrococcygeal joint or the tip of the coccyx and the superior border of the pubic symphysis. There was some disagreement with regard to the appropriate interval for these anatomical intervals, but the most commonly used lower boundary for this coccyx tip-symphysis interval was 1.0 cm and the upper limit reported generally spanned

TABLE II Commonly Used Protocols to Screen for Radiographic Integrity

1. Screening for pelvic tilt by measuring distance between pubic symphysis and coccygeal tip (reported interval spans 1.0-3.0 cm) or between sacrococcygeal joint and pubic symphysis (reported interval spans 2.0-4.7 cm)[24,26,36,39,46,51,56,64,66]
2. Screening for pelvic rotation by evaluating for symmetry of the obturator foramina (obturator index, 0.8-1.2), trochanters, iliac wings, and/or acetabular teardrops[38,41,50,56,64,66]
3. Screening for pelvic rotation by measuring offset of sacrum-pubic symphysis vertical axis (0.5-1.5 cm of lateral shift generally acceptable)[26,51,60]
4. Calibrating for rotational offset based on a measurement of distance between central sacral line and a second line down the center of the pubic symphysis[37,54]
5. Screening out frog-leg lateral/modified Dunn radiographs that do not clearly show the greater trochanter (obscured by femoral neck)[49]

2.0 to 3.0 cm. The lower boundary for the sacrococcygeal joint-symphysis interval was reported as 2.0 cm, and the upper interval ranged from 3.0 to 4.7 cm. Notably, Kutty et al. suggested that this interval should be sex-dependent given native differences in anatomy, reporting an upper limit of 4.7 cm for women and an upper limit of 3.2 cm for men[39]. A number of additional quality-control measures for radiographic integrity have been proposed (Table II). Several of these quality-control measures are illustrated graphically in Figure 2. Sixteen (37.2%) of the included studies disclosed employing at least 1 of these control measures in their respective protocols[24,26,36-39,41,46,49-51,54,56,60,64,66].

Characteristics of Included Studies

Of the 43 investigations included in this review, 26 (60.5%) involved an assessment of reliability of ≥1 of the predetermined radiographic parameters as one of the primary stated purposes of its investigation. The remaining 17 studies (39.5%) reported metrological data on the reliability of ≥1 radiographic parameter but as a "side-arm" calculation rather than as a stated purpose of the investigation. The most common demographic group among the included studies consisted of patients with impingement and/or instability presenting to a hip-preservation specialist, with several studies including asymptomatic volunteers from the general population. Four studies consisted of making measurements using cadaveric specimens (Table III)[24,32,52,58]. The maximum number of observers involved in some component of the included reliability studies averaged 3.7 (range, 1 to 39), but, commonly, only a single observer would continue on to make second readings for the tabulation of intrarater reliability[25,42,44,52,55,61,65,67]. Twenty studies (46.5%) involved >2 observers participating at some stage. The modality used to make measurements was reported explicitly in 31 of the included studies, with manual measurements on a digital platform (90.3%; 28 of 31 studies) being employed most regularly and with computer-driven analysis based on pelvic landmark delineation (6.5%; 2 of 31 studies)[31,63] and hand-derived measurements with use of assistive "protractor-like" devices (6.5%; 2 of 31 studies)[47,60] being used less commonly. While reliability was most frequently tested between attending orthopaedic surgeons and/or residents or fellows, in

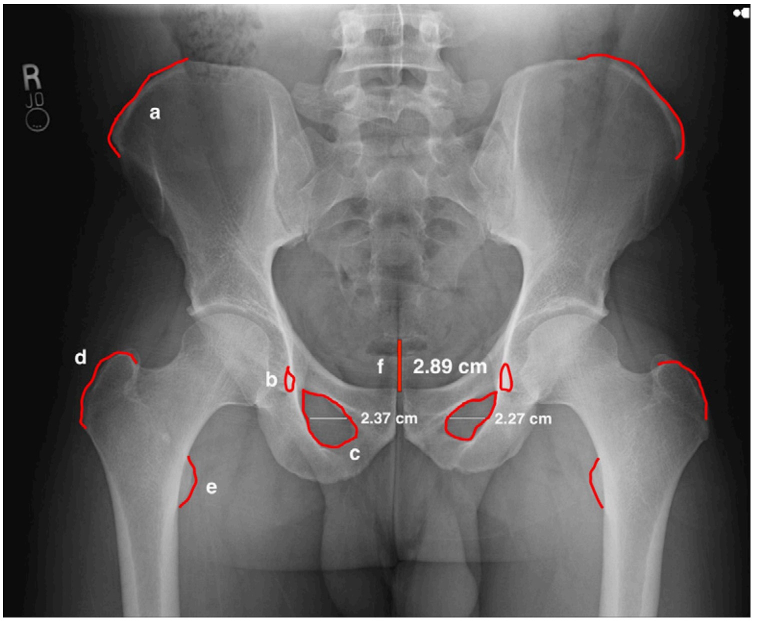

Fig. 2

Radiograph showing common findings used to screen for anteroposterior radiographic integrity, including (a) symmetrical iliac wings, (b) level acetabular teardrops, (c) symmetrical obturator foramina (obturator index of 1.04), (d) clearly distinguishable greater trochanters, (e) clearly distinguishable lesser trochanters, and (f) negligible lateral shift of the sacrum-pubic symphysis axis.

TABLE III Study Populations, Radiographic Measures Tested, and Number and Background of Reviewers Involved*

Study	No. of Patients/ No. of Hips	No. of Reviewers	Population Descriptors	Identity of Reviewers
Cadet et al.[23] (2016)	8/8	4	8 consecutive patients with FAI pathology slated to undergo surgical correction	2 musculoskeletal radiologists with expertise in FAI surgery, 2 orthopaedic surgeons with expertise in FAI surgery
Jamali et al.[24] (2007)	43/84	2	Pelvic specimens from the Institute for Anthropologic Research (Aesch, Switzerland)	1 orthopaedic surgeon fellowship-trained in hip-preservation surgery, 1 orthopaedic surgeon
Ratzlaff et al.[25] (2016)	49/49	2	40 randomly selected volunteers from a population-based sample and 10 additional patients with a confirmed FAI diagnosis	1 third-year medical student with no prior experience with interpreting imaging for FAI, 1 fellowship-trained musculoskeletal radiologist
Philippon et al.[26] (2010)	58/58	2	58 patients undergoing hip arthroscopy for FAI with rim trimming performed at time of surgery	2 orthopaedic surgeons
Valera et al.[29] (2016)	NR/117	3	60 hips were candidates for hip-preservation surgery, 57 hips (control) were free of any abnormality and any history of pain/disease	2 experienced hip surgeons, 1 fourth-year orthopaedic resident
Bouttier et al.[30] (2013)	NR/51	2	Patients with hip pain and/or hip range of motion limitations without severe arthrosis	2 "experienced observers"
Nepple et al.[31] (2014)	70/140	7†	70 consecutive patients: 55 patients undergoing hip- preservation surgery for FAI, 15 patients with dysplasia undergoing hip-preservation surgery	7 experienced hip-preservation surgeons (>10 years in practice)
Zingg et al.[32] (2009)	4/8	3	4 cadaveric specimens preserved in formalin; hips were prescreened and none of them showed evidence of arthrosis	NR
Anderson et al.[33] (2011)	30/30	4	30 young adult patients with hip dysplasia	NR
Carlisle et al.[34] (2011)	45/45	6	Preop. images of 15 patients undergoing surgical intervention for acetabular dysplasia, 15 patients undergoing surgical intervention for FAI, and 15 patients evaluated clinically and found to have no evidence of hip disease	1 adult reconstruction fellow, 2 third-year orthopaedic surgery residents, 1 attending orthopaedic surgeon specializing in sports medicine of knee and shoulder, 2 attending physiatrists
Clohisy et al.[35] (2009)	77/77	6	Preop. images of 27 patients undergoing a periacetabular osteotomy for acetabular dysplasia, 25 patients undergoing open dislocation or combined arthroscopy/open osteochondroplasty for FAI, 25 patients with asymptomatic hips	5 hip specialists with >5 years of practice in treatment of hip dysplasia and FAI, 1 fellow with interest in hip surgery
Schottel et al.[36] (2014)	55/55	4	20 patients with FAI as a diagnosis in principal investigator's institutional database, 20 patients with acetabular dysplasia (AD) as diagnosis, 15 patients without abnormal hip morphology; all FAI/AD patients were >18 years old	1 attending orthopaedic hip surgeon, 1 attending musculoskeletal radiologist, 1 orthopaedic sports fellow, 1 third-year orthopaedic surgery resident

continued

TABLE III (*continued*)

Study	No. of Patients/ No. of Hips	No. of Reviewers	Population Descriptors	Identity of Reviewers
Tannast et al.[37] (2008)	50/100	2	100 hips in a series of 51 consecutive patients in outpatient clinic: 65 hips with FAI, 19 with developmental dysplasia and anterior focal overcoverage, 16 asymptomatic hips	1 orthopaedic surgery fellow, 1 orthopaedic surgeon attending
Hansen et al.[38] (2012)	34/34	2	18 patients with no morphological abnormality of the hip, 16 patients with hip pain and a diagnosis of acetabular retroversion	1 orthopaedic surgeon attending, 1 orthopaedic surgery resident
Kutty et al.[39] (2012)	19/19	2	19 patients with pincer or mixed cam/pincer impingement as determined by 2 radiologists	2 orthopaedic surgeons
Hellman et al.[40] (2015)	45/45	2	45 patients who eventually underwent hip arthroscopy for FAI	2 orthopaedic surgeons
Laborie et al.[41] (2011)	NR/350	2	Subset of individuals enrolled in a population-based randomized trial on developmental dysplasia of the hip	1 pediatric musculoskeletal radiologist (25 years of experience), 1 resident in radiology (1 year of experience)
Pollard et al.[42] (2010)	20/40	2	Asymptomatic individuals from general population and no evidence of arthrosis	2 orthopaedic surgeons
Wilson et al.[43] (2011)	15/30	3	Patients with radiographs on file at a hospital on a single, randomly selected day (no more than "subtle" arthritic change)	3 orthopaedic surgeons (all with minimum 2 years training)
Lee et al.[44] (2011)	36/36	3	No patients had prior hip surgery or pelvic trauma leading to obvious deformity	3 orthopaedic surgeons (4 years, 5 years, and 6 years of experience)
Goldman et al.[45] (2017)	5/10	2	No evident posttraumatic deformity or surgical hardware	1 third-year orthopaedic surgery resident, 1 musculoskeletal radiologist (9 years of experience)
Kappe et al.[46] (2011)	20/40	5	20 randomly selected patients with non-arthritic hip pain	5 orthopaedic surgeons (all surgeons at a high-volume center for hip-preservation surgery)
Khan et al.[47] (2015)	NR/50	4	Young adults with groin-based pain presenting to a single hip arthroscopist	2 orthopaedic surgeons (minimum 5 years in practice), 2 musculoskeletal fellowship-trained radiologists (minimum 5 years in practice)
Malviya et al.[48] (2016)	10/10	39	10 patients being evaluated for FAI/ dysplasia	5 orthopaedic trainees, 9 hip surgeons with no experience performing hip-preservation surgery, 12 hip surgeons with experience in hip-preservation surgery, 10 high-volume periacetabular osteotomy/ arthroscopic hip surgeons, 3 NR
Hooper et al.[49] (2016)	177/177‡	3	Preop. radiographs of adolescents who underwent hip arthroscopy for treatment of chondrolabral damage	1 fellowship-trained musculoskeletal radiologist, 1 fourth-year medical student, 1 orthopaedic surgery resident

continued

TABLE III (*continued*)

Study	No. of Patients/ No. of Hips	No. of Reviewers	Population Descriptors	Identity of Reviewers
Park et al.[50] (2016)	118/132	2	Patients who had a clinical history of hip pain and were between 20 and 86 years old; patients with a history of hip fracture, prior surgery, septic/ inflammatory arthritis, osteonecrosis, metastasis were excluded	2 musculoskeletal radiologists (12 and 7 years of experience)
Stubbs et al.[51] (2011)	30/60	3	30 skeletally mature patients without advanced arthritic changes (Tönnis 0-1)	3 hip-specialist orthopaedic surgeons
Bittersohl et al.[52] (2013)	1/1§	2	A single dry, human cadaveric pelvis moved through sequential positions of sagittal tilt and axial rotation	NR
Domb et al.[53] (2013)	50/NR	1	Patients undergoing hip arthroscopy. None had a history of traumatic hip injury, prior hip abnormality, prior hip arthroscopy, or advanced arthrosis	1 orthopaedic surgeon
Mast et al.[54] (2011)	39/NR	2	Patients undergoing hip arthroscopy with impingement and labral/ cartilage disorders; no patients had severe dysplasia with subluxation	1 third-year resident in orthopaedics, 1 second-year hip fellow
Wensaas et al.[55] (2012)	36/43	3	Patients returning at long term (avg. 37 years) following surgical intervention for slipped upper femoral epiphyses	2 orthopaedic surgeons, 1 musculoskeletal radiologist
Chadayammuri et al.[56] (2015)	NR/25	2	Non-arthritic patients presenting to a hip-preservation surgeon for possible treatment of instability or impingement. Any patients with severe anatomical deformity (slipped capital femoral epiphysis, Perthes, osteochondromatosis, post-dislocation syndrome) were excluded	NR
Hanson et al.[57] (2015)	188/188#	2	Adult patients (>18 years old) who subsequently underwent hip arthroscopy	1 orthopaedic physician's assistant, 1 second-year medical student
Colvin et al.[58] (2011)	5/10	3	Formalin-preserved cadavers	3 orthopaedists (various levels of training)
Barton et al.[59] (2011)	68	2	Patients undergoing hip-preservation surgery with unilateral hip pain, impingement, and no radiographic signs of dysplasia. No patient had prior hip surgery.	NR
Lequesne et al.[60] (2004)	30/60	2	Patients free of hip pain, >18 years old requiring radiographs for monitoring of metastasis, bone marrow disease, renal failure, long-term steroid treatment, or trauma. Hips with osteophytes, osteoarthritis, or cysts were excluded	NR
Espié et al.[61] (2014)	75/96	3	All patients underwent surgical intervention for resection of a cam lesion	NR

continued

TABLE III (continued)

Study	No. of Patients/ No. of Hips	No. of Reviewers	Population Descriptors	Identity of Reviewers
Espié et al.[62] (2016)	100/NR	3	50 patients who ultimately underwent surgery for decompression of a cam lesion; 50 asymptomatic, age-matched controls	NR
Engesæter et al.[63] (2012)	95/190	3	Patients with dysplasia at skeletal maturity (all 18-19 years old): ~2/3 of patients were "subjectively normal", ~1/3 had "mild dysplasia," 9 had "significant" dysplasia	NR (possibly "research fellows")
Levy et al.[64] (2015)	391/438	2	Female patients with hip pain, a diagnosis of impingement, and no evidence of advanced arthrosis (all patients had Tönnis grade <2)	2 senior resident physicians
Clohisy et al.[65] (2007)	80/85	2	61 hips with FAI ultimately treated surgically, 24 age and sex-matched control hips without intra-articular hip disease	NR
Farkas et al.[66] (2015)	20/NR	2	Symptomatic FAI patients scheduled for arthroscopic surgery	2 fellowship-trained orthopaedic surgeons
Konan et al.[67] (2010)	10/10	2	Patients treated for FAI by arthroscopic surgery with cam lesion noted preoperatively	NR

*NR = not reported. †The study involved a total of 7 observers, but interrater agreement was analyzed in groups of 4 observers for anteroposterior-derived measurements. ‡Only 124 of 177 radiographs were available for determination of the lateral center-edge angle, 122 were available for determination of the Tönnis angle, and 105 were available for determination of the alpha angle. §The study included 1 cadaver, but 13 measurements were performed (in different positions of sagittal tilt and axial rotation). #Only 133 of 188 false-profile radiographs were available for measurement of the anterior center-edge angle.

a number of investigations, reproducibility was also evaluated outside the immediate realm of orthopaedic surgery, with medical students, musculoskeletal radiologists, and physiatrists being involved in the assessment (Table III)[23,25,34,36,41,45,47,49,50,55,57].

Radiographic Parameters

In the included investigations, the radiographic parameters with the most reliability data include the lateral center-edge angle (22 studies), the Tönnis angle (14 studies), and the alpha angle (18 studies). Categorical markers of acetabular retroversion and the extent of osteoarthritis have been less frequently studied with respect to their reliability, with the exception of the crossover sign, which was evaluated in 11 of the included studies (Table IV). Radiographic measurements, categorical signs, and their associated definitions are illustrated for clarity in Figures 3, 4, and 5.

The proportions of investigations reporting an ICC or Cohen kappa coefficient value attaining the predetermined threshold of ≥ 0.60 are presented for each radiographic parameter in Table V. Data specific to the individual studies included in this review can be found in the Appendix.

Discussion

The principal aim of the present study was to summarize the metrological evidence in support of or against the reliability of commonly used radiographic measures of acetabular and proximal femoral morphology in the assessment of the non-arthritic hip. Another aim was to collate best practices and the frequencies with which these practices have been reported in existing investigations presenting data on the reliability of these measures. With regard to the principal objective, we found that the lateral center-edge angle, regardless of whether it is measured to the lateral osseous edge or to the lateral margin of the weight-bearing sourcil, had most consistently demonstrated a high level of reliability as a measure of acetabular coverage in the dysplastic, pincer, and "normal" hip. All 22 studies (100%) that evaluated the lateral center-edge angle demonstrated acceptable interrater and intrarater reliability of this measure regardless of whether it was investigated as a continuous measure (ICC) or a categorical proxy for a diagnosis of dysplasia or pincer impingement (Cohen kappa coefficient). A similarly robust body of evidence has been reported in the literature on the reliability of the alpha angle, with 18 of the included articles reporting on the reliability of the alpha angle on ≥ 1 radiographic view. However, the

TABLE IV Radiographic Parameters and Associated Volume of Evidence from the Included Investigations*

LCEA (Wiberg/Ogata) (22 Studies)	ACEA (6 Studies)	NSA (6 Studies)	PWS (5 Studies)	TA (14 Studies)	AA (18 Studies)	COS (11 Studies)	ISS (4 Studies)	TG (7 Studies)
Philippon et al.[26] (2010)	Zingg et al.[32] (2009)	Nepple et al.[31] (2014)	Nepple et al.[31] (2014)	Nepple et al.[31] (2014)	Nepple et al.[31] (2014)	Nepple et al.[31] (2014)	Nepple et al.[31] (2014)	Nepple et al.[31] (2014)
Nepple et al.[31] (2014)	Bouttier et al.[30] (2013)	Bouttier et al.[30] (2013)	Laborie et al.[41] (2011)	Bouttier et al.[30] (2013)	Cadet et al.[23] (2016)	Cadet et al.[23] (2016)	Kappe et al.[46] (2011)	Valera et al.[29] (2016)
Cadet et al.[23] (2016)	Carlisle et al.[34] (2011)	Wilson et al.[43] (2011)	Kappe et al.[46] (2011)	Carlisle et al.[34] (2011)	Ratzlaff et al.[25] (2016)	Ratzlaff et al.[25] (2016)	Malviya et al.[48] (2016)	Carlisle et al.[34] (2011)
Ratzlaff et al.[25] (2016)	Tannast et al.[37] (2008)	Park et al.[50] (2016)	Malviya et al.[48] (2016)	Hansen et al.[38] (2012)	Carlisle et al.[34] (2011)	Jamali et al.[24] (2007)	Mast et al.[54] (2011)	Clohisy et al.[35] (2009)
Bouttier et al.[30] (2013)	Anderson et al.[33] (2011)	Mast et al.[54] (2011)	Tannast et al.[37] (2008)	Goldman et al.[45] (2017)	Schottel et al.[36] (2014)	Clohisy et al.[35] (2009)		Schottel et al.[36] (2014)
Carlisle et al.[34] (2011)	Hanson et al.[57] (2015)	Lequesne et al.[60] (2004)		Hooper et al.[49] (2016)	Hellman et al.[40] (2015)	Schottel et al.[36] (2014)		Mast et al.[54] (2011)
Schottel et al.[36] (2014)				Stubbs et al.[51] (2011)	Pollard et al.[42] (2010)	Laborie et al.[41] (2011)		Malviya et al.[48] (2016)
Hansen et al.[38] (2012)				Domb et al.[53] (2013)	Khan et al.[47] (2015)	Kappe et al.[46] (2011)		
Kutty et al.[39] (2012)				Mast et al.[54] (2011)	Hooper et al.[49] (2016)	Malviya et al.[48] (2016)		
Lee et al.[44] (2011)				Lequesne et al.[60] (2004)	Mast et al.[54] (2011)	Mast et al.[54] (2011)		
Goldman et al.[45] (2017)				Engesæter et al.[63] (2012)	Barton et al.[59] (2011)	Tannast et al.[37] (2008)		
Hooper et al.[49] (2016)				Tannast et al.[37] (2008)	Espié et al.[61] (2014)			
Stubbs et al.[51] (2011)				Bittersohl et al.[52] (2013)	Espié et al.[62] (2016)			
Domb et al.[53] (2013)				Lee et al.[44] (2011)	Levy et al.[64] (2015)			
Mast et al.[54] (2011)					Wensaas et al.[55] (2012)			
Chadayammuri et al.[56] (2015)					Clohisy et al.[65] (2007)			
Colvin et al.[58] (2011)					Farkas et al.[66] (2015)			
Lequesne et al.[60] (2004)					Konan et al.[67] (2010)			
Engesæter et al.[63] (2012)								
Tannast et al.[37] (2008)								
Anderson et al.[33] (2011)								
Hanson et al.[57] (2015)								

*LCEA = lateral center-edge angle (Wiberg or Ogata), ACEA = anterior center-edge angle on false-profile radiograph, NSA = neck-shaft angle, PWS = posterior wall sign, TA = Tönnis angle, AA = alpha angle, COS = crossover sign, ISS = ischial spine sign, and TG = Tönnis grade of osteoarthritis.

evidence is less consistent with regard to the alpha angle as a reliable measure for the assessment of proximal femoral morphology, with the ICC spanning −0.08 to 0.96 for interrater reliability and spanning 0.34 to 0.987 for intrarater reliability. A qualitative synthesis of the evidence suggests that the intrarater reliability of this measure may be greater than its reliability among raters, with 4 studies failing to demonstrate substantial interrater reliability (ICC $\geq$ 0.60) of the alpha angle then demonstrating much improved intrarater reliability[31,34,47,55].

Reliable identification of the center of the femoral head is fundamental to a reproducible center-edge or alpha angle, and a simple shift of the point of localization by 3 mm laterally in the horizontal plane has been reported to alter the center-edge angle by 6°[33].

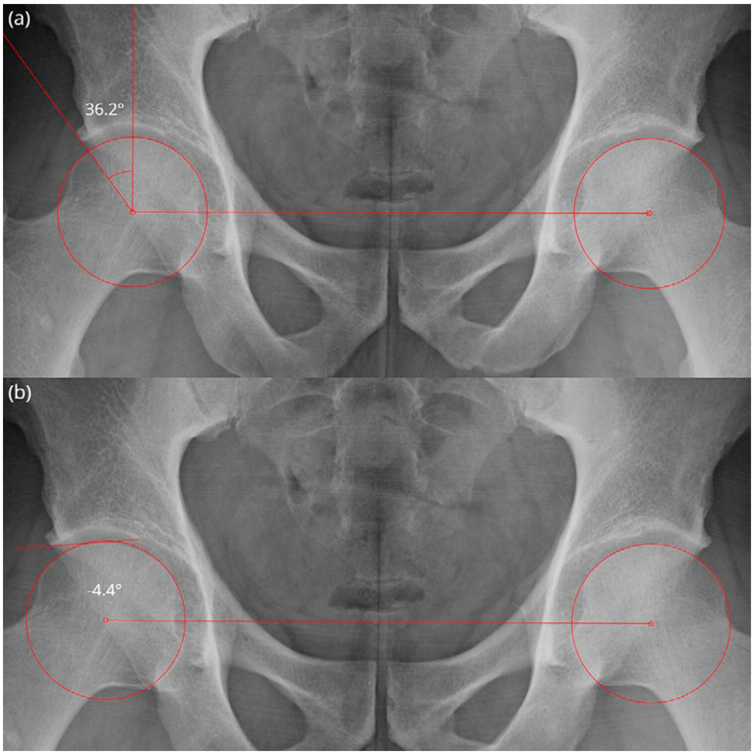

Fig. 3

Fig. 3-A Radiograph showing the lateral center-edge angle of Wiberg, which is formed by the perpendicular to the line connecting the center of the femoral heads and the line from the center of the head to the lateral acetabular rim. Pelvic obliquity is accounted for by connecting the femoral head centers. **Fig. 3-B** Radiograph showing the Tönnis angle, which is formed by the line connecting the femoral head centers (horizontal pelvic plane) and the line connecting the medial and lateral edges of the acetabular sourcil.

Localization of the center of the femoral head is sensitive to whether the asphericity of the cam deformity is included in the placement of the spherical template, and, in the setting of cam impingement, Anderson et al. demonstrated a degradation of reliability among observers in measuring anterior and lateral center-edge angles as the alpha angle increases[33]. We agree that differences in mapping of the spherical contour of the femoral head in the setting of an anterolateral cam

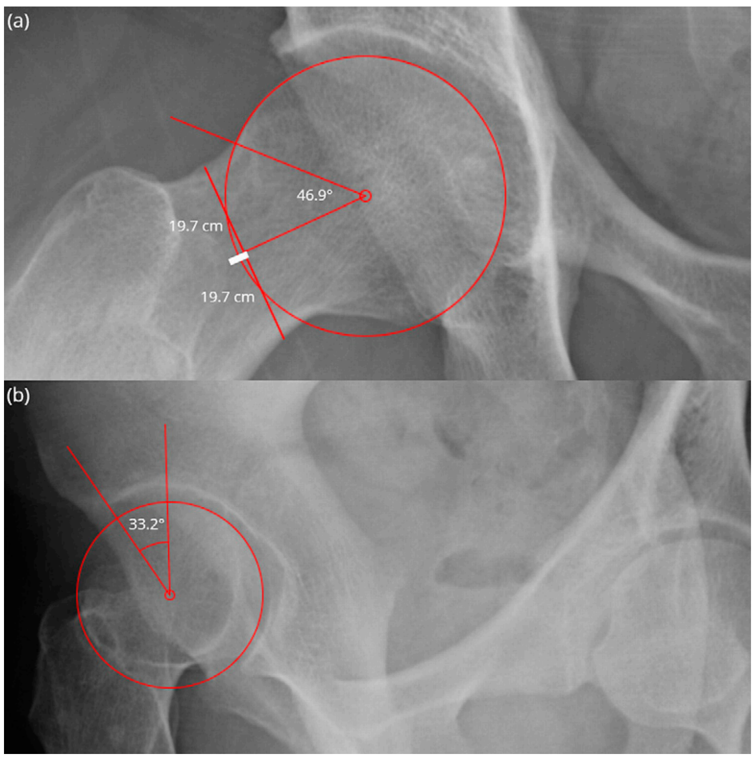

Fig. 4

Fig. 4-A Radiograph showing the alpha angle, which is formed by a line extending down the axis of the femoral neck and a line from the femoral head center to the point at which the contour of the head exceeds its radius; the midpoint of the femoral neck is defined quantitatively. **Fig. 4-B** Radiograph showing the anterior center-edge angle, which is formed by a vertical line and a line extending from the center of the femoral head to the tip of the anterior acetabular sourcil.

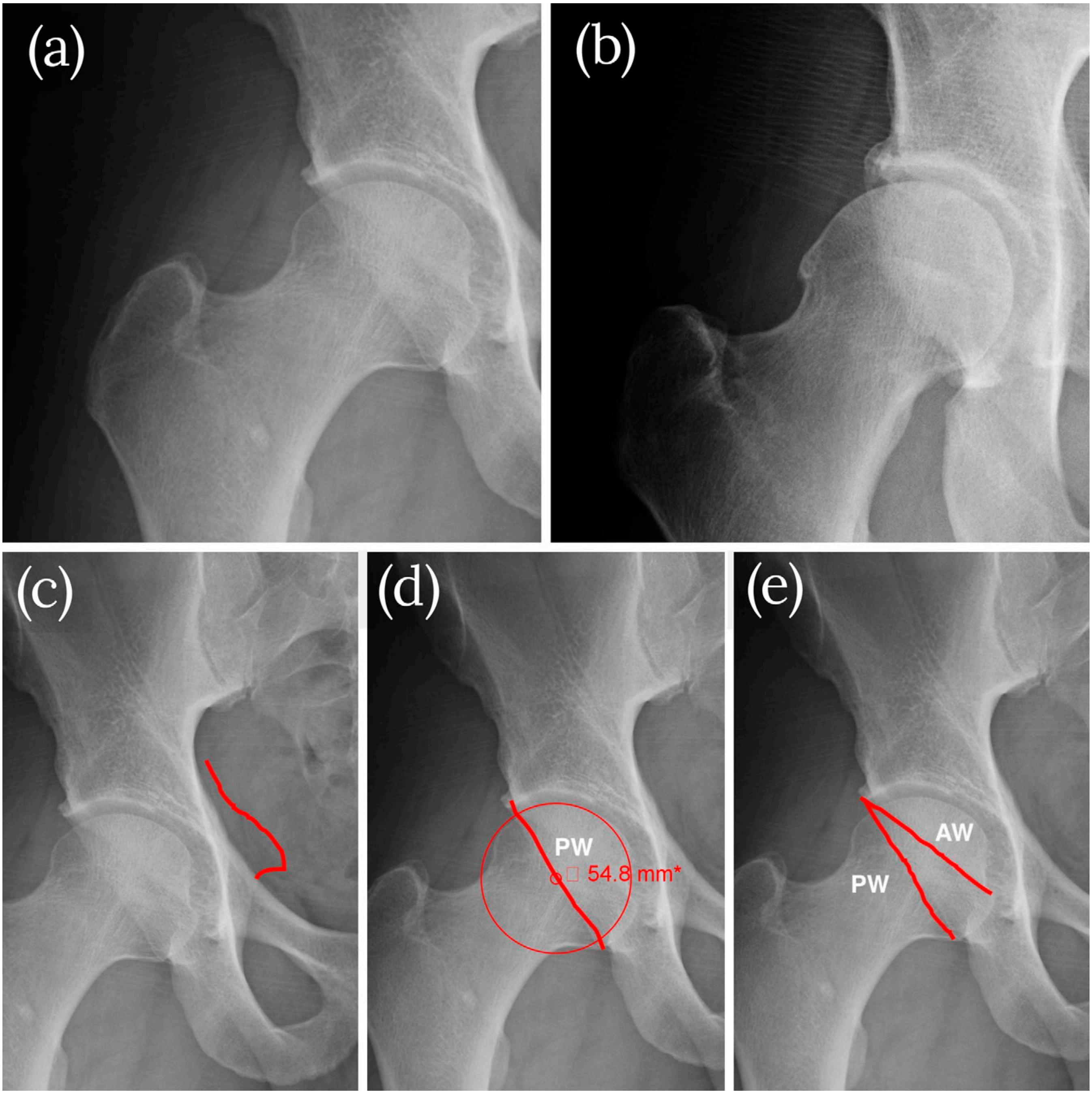

Fig. 5

Fig. 5-A Radiograph showing a Tönnis grade-0 hip. Figs. 5-B through 5-E Radiographs of a Tönnis grade-1 hip, showing increased acetabular sclerosis and slight loss of head sphericity with a lateral femoral osteophyte (Fig. 5-B), a prominent ischial spine (**Fig. 5-C**), a negative posterior wall sign (indicated by the posterior acetabular wall [PW] intersecting, but not being medial to, the femoral head center [small circle]) (**Fig. 5-D**), and a negative crossover sign (with the anterior wall [AW] and posterior wall [PW] of the acetabulum not intersecting medial to the lateral acetabular roof) (**Fig. 5-E**).

lesion and consequent localization of the center of the head are responsible for some of the discord among reviewers in the assessment of the alpha angle, and we hypothesize that individualized interpretations of the protocol for identifying the center of the head may reflect its increased intraobserver reliability. While holding a consensus-building discussion prior to measurement is recommended as a standard component of study design for the assessment of reliability (Table I), this may not reflect the reality of standard practice, in which consensus among providers may not be readily facilitated. Although it is unclear why this difficulty did not appear to influence the interrater reliability of the measurement of the lateral center-edge angle in the studies included in the present review, we suggest that the

TABLE V Percentage of Investigations Reporting ICC or Cohen Kappa Coefficient for Radiographic Measurements*†

Radiographic Measure	Percentage of Studies Attaining ICC or Cohen Kappa Coefficient of ≥0.60: Interrater Reliability	Intrarater Reliability
Lateral center-edge angle	100.0% (24 of 24)	100.0% (18 of 18)
Anterior center-edge angle	83.3% (5 of 6)	100.0% (4 of 4)
Neck-shaft angle	83.3% (5 of 6)	100.0% (5 of 5)
Posterior wall sign	80.0% (4 of 5)	100.0% (4 of 4)
Tönnis angle	78.6% (11 of 14)	90.9% (10 of 11)
Alpha angle	69.6% (16 of 23)	90.5% (19 of 21)
Crossover sign	54.5% (6 of 11)	60.0% (6 of 10)
Ischial spine sign	50.0% (2 of 4)	100.0% (3 of 3)
Tönnis grade	14.3% (1 of 7)	33.3% (2 of 6)

*When a combined calculation for intrarater reliability was not reported when multiple assessors were involved, the average ICC or Cohen kappa coefficient across raters was tabulated for the purposes of screening the study's data against the 0.60 threshold. †If a study included a value for ICC or Cohen kappa coefficient for multiple views (alpha anteroposterior, alpha cross-table lateral) or multiple variations of a measurement (lateral center-edge angle of Wiberg or Ogata), the study was counted once for each view/variation of the measurement.

interpretation of the break in femoral head sphericity is more variable than the interpretation of the lateral margin of the acetabulum. Additionally, cam lesions often are not visualized on an anteroposterior radiograph, so identification of the center of the femoral head may be less problematic when making a lateral center-edge angle measurement. Furthermore, drawing an axis down the midpoint of the femoral neck, parallel to the neck axis, is challenging to perform reproducibly in the absence of a quantitative approach to delineating the center of the neck. The error associated with this component of the measurement is only further compounded by the error associated with reliably identifying the break in sphericity of the femoral head. In the case of the lateral center-edge angle, differences in understanding with regard to what constitutes the lateral boundary of the acetabulum in the setting of calcified labra and lateral osteophytes do threaten reliability. Additional 3-dimensionally-based imaging studies are needed to clarify what truly constitutes femoral coverage in these settings, such that clinically relevant end points for the lateral delineation can be recommended[56].

The overall body of evidence evaluating the reliability of the Tönnis angle and the crossover sign is reasonably robust, with the Tönnis angle demonstrating greater evidence of reproducibility than the crossover sign. Difficulties with the crossover sign as a categorical proxy for acetabular retroversion have been reported in the literature, particularly in the setting of a low-lying anterior inferior iliac spine, which can complicate identification of the anterior acetabular wall[68]. While an assessment of lateral acetabular coverage on radiographs with use of the lateral center-edge angle of Wiberg does not require differentiating superimposed bone, such determinations are required during the identification of the presence or absence of a crossover or posterior wall sign. Pelvic bone outside of the acetabular margin may be mistaken for the acetabulum, leading to variability in assessment. Additionally, of great importance in the assessment of acetabular retroversion by way of the binary categorical measures studied in this review are the influences of pelvic tilt and rotation on the appearance of acetabular morphology. Sixteen of the studies that were included in this review reported the use of some sort of quality-control procedure to account for the influence of pelvic tilt/and or rotation on the integrity of anteroposterior radiographs. Correlations have been reported in the literature between positive crossover signs and pelvic inclination and also between positive posterior wall signs and pelvic inclination[69,70]. These checks on radiographic integrity need to be in place not only to facilitate appropriate diagnosis and surgical management but also to promote concordance among providers who may interpret these signs differently, depending on whether or not malpositioning is noted. Given the inconsistency in the metrological data on the reliability of these binary proxies for acetabular version and the known confounders to the consistency of the crossover and posterior wall signs as valid instruments, 3-dimensional imaging should continue to be considered in cases in which acetabular retroversion and/or posterior insufficiency is suspected[71].

Multicentered, collaborative efforts in hip preservation allow for data-pooling practices, and the emergence of these collaborative efforts has underscored the importance of the interrater reliability of radiographic measures. Measurements often become the basis for study group classification (i.e., dysplasia versus normal) and comparative analysis. The present review was intended (1) to serve as a repository for the data that can be used by these investigations and (2) to highlight some of the best practices for minimizing bias and improving the strength of the findings that proceed from the efforts of individual study groups to demonstrate the reliability of their data. We compiled 11 methodological considerations modeled by the investigations included in this review that we felt to be most important to keep in mind (Table I). The frequency with which these practices were reported varied across individual studies. Although potentially of less importance than the lateral center-edge angle or the alpha angle as a marker for the diagnosis of abnormality (i.e., dysplasia, cam impingement) at the time of data pooling, the anterior center-edge angle, neck-shaft angle, ischial spine sign, and posterior wall sign have more limited evidence in support of their reliability and may be worth the focus of further investigative efforts.

Despite having been evaluated in only 7 of the included studies, the Tönnis grade of osteoarthritis stood out as consistently demonstrating subpar interrater and intrarater reliability. While the Tönnis system originally was used to classify the extent of osteoarthrosis in a dysplastic population, the use of this system has become widespread among hip-preservation specialists[72,73]. In this setting, Tönnis grade-0 and 1 hips constitute the vast majority of treatment populations. Underrepresentation of patients with more severe degenerative changes likely contributes to some of the difficulties in demonstrating reliability in this pre-arthritic population. The difference between Tönnis grade-0 and 1 hips can be difficult to ascertain, and the classification system in its current form is complicated by its subpar construction. In particular, the designation of a Tönnis grade-1 hip as a hip with "no or slight loss of head sphericity" is inconsistently regarded by practitioners, as patients with a cam deformity but an absence of degenerative changes have been designated either way. The designation of "slight joint-space narrowing" as being indicative of a Tönnis grade-1 hip is also limited as these designations often are made without measurement, and reference to the contralateral hip for an indication of a "normal" joint space may be complicated in patients with bilateral, early degenerative change. Incorporating a quantitative cutoff for what constitutes relative joint-space narrowing or a percentage of narrowing relative to the contralateral hip might help to improve the reproducibility of this classification scheme in the setting of hip preservation. In addition, the importance of being able to make such a distinction is not entirely clear. The current evidence is mixed with regard to whether the outcomes for patients with minimal radiographic osteoarthritic changes are comparable with those for patients with no signs of degenerative change[74-76].

Limitations

The present review had a number of limitations. Several of the included investigations presented reliability data as part of a side-arm investigation to establish the import and validity of a series of other conclusions. The aggregate metrological reliability data could be artificially inflated by some selection bias, as investigators demonstrating adequate reliability might be more inclined to report these data. Additionally, only 46.5% (20) of the investigations involved >2 observers participating at some stage of the reliability assessment, and often these observers were intricately involved with study design and treatment. A subgroup of studies involving >2 readers was previously discussed by Nepple et al., who reported that 2-reader studies generally provide higher estimates of reliability than do more robust, multireader (>2) efforts[31].

Furthermore, ICCs were rarely reported by way of model, type, and measure used, and ICC values can change notably depending on whether the model appropriate to the individual context is used. Values were collated and tested against an arbitrary threshold in order to synthesize a large amount of data, but the limitations of this practice are clear. In addition, the present review does not include any data regarding the validity of these radiographic measures. Last, while the reliability of a given radiographic measure may be influenced by the level of experience of the reviewers and the modality used for making measurements (i.e., computer-assisted measurements versus manual measurements), the extent to which these factors play a role in reliability would benefit from further investigation.

Appendix

Tables showing the search strategy and data from the individual studies included in the present review are available with the online version of this article as a data supplement at jbjs.org (http://links.lww.com/JBJSREV/A423).

Dominic S. Carreira, MD[1],
Brendan R. Emmons, BS[1]

[1]Peachtree Orthopedics, Atlanta, Georgia

E-mail address for D.S. Carreira: dcarreira@gmail.com

ORCID iD for D.S. Carreira: 0000-0003-0887-7424
ORCID iD for B.R. Emmons: 0000-0003-3563-6124

References

1. Ranawat AS, Schulz B, Baumbach SF, Meftah M, Ganz R, Leunig M. Radiographic predictors of hip pain in femoroacetabular impingement. HSS J. 2011 Jul;7(2):115-9. Epub 2011 Jan 8.

2. Frangiamore S, Mannava S, Geeslin AG, Chahla J, Cinque ME, Philippon MJ. Comprehensive clinical evaluation of femoroacetabular impingement: part 1, physical examination. Arthrosc Tech. 2017 Oct 30;6(5):e1993-2001.

3. Haldane CE, Ekhtiari S, de Sa D, Simunovic N, Ayeni OR. Preoperative physical examination and imaging of femoroacetabular impingement prior to hip arthroscopy-a systematic review. J Hip Preserv Surg. 2017 Jun 15; 4(3):201-13.

4. Smith KM, Gerrie BJ, McCulloch PC, Lewis BD, Mather RC, Van Thiel G, Nho SJ, Harris JD. Arthroscopic hip preservation surgery practice patterns: an international survey. J Hip Preserv Surg. 2016 Dec 7;4(1):18-29.

5. Larson CM, Giveans MR, Samuelson KM, Stone RM, Bedi A. Arthroscopic hip revision surgery for residual femoroacetabular impingement (FAI): surgical outcomes compared with a matched cohort after primary arthroscopic FAI correction. Am J Sports Med. 2014 Aug;42(8):1785-90. Epub 2014 May 29.

6. Mei-Dan O, McConkey MO, Brick M. Catastrophic failure of hip arthroscopy due to iatrogenic instability: can partial division of the ligamentum teres and iliofemoral ligament cause subluxation? Arthroscopy. 2012 Mar; 28(3):440-5. Epub 2012 Feb 1.

7. Wylie JD, Kapron AL, Peters CL, Aoki SK, Maak TG. Relationship between the lateral center-edge angle and 3-dimensional acetabular coverage. Orthop J Sports Med. 2017 Apr 12;5(4): 2325967117700589.

8. Hilibrand AS, Spindler K, O'Keefe RJ. Demonstrating the value of orthopaedic surgery through multicenter trials: AOA critical issues. J Bone Joint Surg Am. 2015 Apr 1;97(7):e35.

9. Matsuda DK, Khatod M, Antounian F, Burchette R, Bini S, Anthony FF, Harris J, Calipusan C. Multicenter outcomes of arthroscopic surgery for femoroacetabular impingement in the community hospital setting. J Hip Preserv Surg. 2016 Jul 17;3(4):318-24.

10. Hevesi M, Krych AJ, Johnson NR, Redmond JM, Hartigan DE, Levy BA, Domb BG. Multicenter analysis of midterm clinical outcomes of arthroscopic labral repair in the hip: minimum 5-year follow-up. Am J Sports Med. 2018 Feb; 46(2):280-7. Epub 2017 Oct 24.

11. Matsuda DK, Wolff AB, Nho SJ, Salvo JP Jr, Christoforetti JJ, Kivlan BR, Ellis TJ, Carreira DS; Multicenter Arthroscopic Study of the Hip (MASH) Study Group. Hip dysplasia: prevalence, associated findings, and procedures from large multicenter arthroscopy study group. Arthroscopy. 2018 Feb;34(2):444-53. Epub 2017 Nov 13.

12. Lund B, Mygind-Klavsen B, Grønbech Nielsen T, Maagaard N, Kraemer O, Hölmich P, Winge S, Lind M. Danish Hip Arthroscopy Registry (DHAR): the outcome of patients with femoroacetabular impingement (FAI). J Hip Preserv Surg. 2017 Apr 4;4(2):170-7.

13. Larson CM, Clohisy JC, Beaulé PE, Kelly BT, Giveans MR, Stone RM, Samuelson KM; ANCHOR Study Group. Intraoperative and early postoperative complications after hip arthroscopic surgery: a prospective multicenter trial utilizing a validated grading scheme. Am J Sports Med. 2016 Sep;44(9):2292-8. Epub 2016 Jun 16.

14. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009 Jul 21;6(7): e1000097. Epub 2009 Jul 21.

15. Wiberg G. Studies on dysplastic acetabula and congenital subluxation of the hip joint. Acta Chir Scand. 1939 Jul;83(Suppl 58):28-38.

16. Ogata S, Moriya H, Tsuchiya K, Akita T, Kamegaya M, Someya M. Acetabular cover in congenital dislocation of the hip. J Bone Joint Surg Br. 1990 Mar;72(2):190-6.

17. Lequesne M, de Seze S. [False profile of the pelvis. A new radiographic incidence for the study of the hip. Its use in dysplasias and different coxopathies]. Rev Rhum Mal Osteoartic. 1961 Dec;28:643-52.

18. Tönnis D. Congenital dysplasia and dislocation of the hip in children and adults. Berlin: Springer; 1987.

19. Bardakos NV, Villar RN. Predictors of progression of osteoarthritis in femoroacetabular impingement: a radiological study with a minimum of ten years follow-up. J Bone Joint Surg Br. 2009 Feb;91(2):162-9.

20. Nötzli HP, Wyss TF, Stoecklin CH, Schmid MR, Treiber K, Hodler J. The contour of the femoral head-neck junction as a predictor for the risk of anterior impingement. J Bone Joint Surg Br. 2002 May;84(4):556-60.

21. Reynolds D, Lucas J, Klaue K. Retroversion of the acetabulum. A cause of hip pain. J Bone Joint Surg Br. 1999 Mar;81(2):281-8.

22. Kalberer F, Sierra RJ, Madan SS, Ganz R, Leunig M. Ischial spine projection into the pelvis : a new sign for acetabular retroversion. Clin Orthop Relat Res. 2008 Mar;466(3):677-83. Epub 2008 Feb 10.

23. Cadet ER, Babatunde OM, Gorroochurn P, Chan AK, Stancato-Pasik A, Brown M, Johnson S, Kaiser PB, Gardner TR, Ayeni OR. Inter- and intra-observer agreement of femoroacetabular impingement (FAI) parameters comparing plain radiographs and advanced, 3D computed tomographic (CT)-generated hip models in a surgical patient cohort. Knee Surg Sports Traumatol Arthrosc. 2016 Jul;24(7):2324-31. Epub 2014 Sep 26.

24. Jamali AA, Mladenov K, Meyer DC, Martinez A, Beck M, Ganz R, Leunig M. Anteroposterior pelvic radiographs to assess acetabular retroversion: high validity of the "cross-over-sign". J Orthop Res. 2007 Jun;25(6):758-65.

25. Ratzlaff C, Zhang C, Korzan J, Josey L, Wong H, Cibere J, Prlic HM, Kopec JA, Esdaile JM, Li LC, Barber M, Forster BB. The validity of a non-radiologist reader in identifying cam and pincer femoroacetabular impingement (FAI) using plain radiography. Rheumatol Int. 2016 Mar; 36(3):371-6. Epub 2015 Oct 3.

26. Philippon MJ, Wolff AB, Briggs KK, Zehms CT, Kuppersmith DA. Acetabular rim reduction for the treatment of femoroacetabular impingement correlates with preoperative and postoperative center-edge angle. Arthroscopy. 2010 Jun; 26(6):757-61. Epub 2010 Apr 3.

27. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998 Oct;26(4):217-38.

28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 Mar;33(1):159-74.

29. Valera M, Ibañez N, Sancho R, Tey M. Reliability of Tönnis classification in early hip arthritis: a useless reference for hip-preserving surgery. Arch Orthop Trauma Surg. 2016 Jan; 136(1):27-33. Epub 2015 Nov 14.

30. Bouttier R, Morvan J, Mazieres B, Rat AC, Ziegler LE, Fardellone P, Fautrel B, Guillemin F, Pouchot J, Roux C, Coste J, Saraux A. Reproducibility of radiographic hip measurements in adults. Joint Bone Spine. 2013 Jan;80(1):52-6. Epub 2012 Mar 23.

31. Nepple JJ, Martell JM, Kim YJ, Zaltz I, Millis MB, Podeszwa DA, Sucato DJ, Sink EL, Clohisy JC; ANCHOR Study Group. Interobserver and intraobserver reliability of the radiographic analysis of femoroacetabular impingement and dysplasia using computer-assisted measurements. Am J Sports Med. 2014 Oct;42(10): 2393-401. Epub 2014 Aug 18.

32. Zingg PO, Werner CM, Sukthankar A, Zanetti M, Seifert B, Dora C. The anterior center edge angle in Lequesne's false profile view: interrater correlation, dependence on pelvic tilt and correlation to anterior acetabular coverage in the sagital plane. A cadaver study. Arch Orthop Trauma Surg. 2009 Jun;129(6):787-91. Epub 2008 Jul 16.

33. Anderson LA, Gililland J, Pelt C, Linford S, Stoddard GJ, Peters CL. Center edge angle measurement for hip preservation surgery: technique and caveats. Orthopedics. 2011 Jan 1;34(2):86.

34. Carlisle JC, Zebala LP, Shia DS, Hunt D, Morgan PM, Prather H, Wright RW, Steger-May K, Clohisy JC. Reliability of various observers in determining common radiographic parameters of adult hip structural anatomy. Iowa Orthop J. 2011;31:52-8.

35. Clohisy JC, Carlisle JC, Trousdale R, Kim YJ, Beaule PE, Morgan P, Steger-May K, Schoenecker PL, Millis M. Radiographic evaluation of the hip has limited reliability. Clin Orthop Relat Res. 2009 Mar;467(3):666-75. Epub 2008 Dec 2.

36. Schottel PC, Park C, Chang A, Knutson Z, Ranawat AS. The role of experience level in radiographic evaluation of femoroacetabular impingement and acetabular dysplasia. J Hip Preserv Surg. 2014 Aug 22;1(1):21-6.

37. Tannast M, Mistry S, Steppacher SD, Reichenbach S, Langlotz F, Siebenrock KA, Zheng G. Radiographic analysis of femoroacetabular impingement with Hip2Norm-reliable and validated. J Orthop Res. 2008 Sep;26(9):1199-205.

38. Hansen BJ, Harris MD, Anderson LA, Peters CL, Weiss JA, Anderson AE. Correlation between radiographic measures of acetabular morphology with 3D femoral head coverage in patients with acetabular retroversion. Acta Orthop. 2012 Jun;83(3):233-9. Epub 2012 May 4.

39. Kutty S, Schneider P, Faris P, Kiefer G, Frizzell B, Park R, Powell JN. Reliability and predictability of the centre-edge angle in the assessment of pincer femoroacetabular impingement. Int Orthop. 2012 Mar;36(3):505-10. Epub 2011 Jul 1.

40. Hellman MD, Mascarenhas R, Gupta A, Fillingham Y, Haughom BD, Salata MJ, Nho SJ. The false-profile view may be used to identify cam morphology. Arthroscopy. 2015 Sep;31(9): 1728-32. Epub 2015 May 6.

41. Laborie LB, Lehmann TG, Engesæter IØ, Eastwood DM, Engesæter LB, Rosendahl K. Prevalence of radiographic findings thought to be associated with femoroacetabular impingement in a population-based cohort of 2081 healthy young adults. Radiology. 2011 Aug;260(2):494-502. Epub 2011 May 25.

42. Pollard TC, Villar RN, Norton MR, Fern ED, Williams MR, Simpson DJ, Murray DW, Carr AJ. Femoroacetabular impingement and classification of the cam deformity: the reference interval in normal hips. Acta Orthop. 2010 Feb;81(1):134-41.

43. Wilson JD, Eardley W, Odak S, Jennings A. To what degree is digital imaging reliable? Validation of femoral neck shaft angle measurement in the era of picture archiving and communication systems. Br J Radiol. 2011 Apr;84(1000):375-9. Epub 2010 Dec 15.

44. Lee YK, Chung CY, Koo KH, Lee KM, Kwon DG, Park MS. Measuring acetabular dysplasia in plain radiographs. Arch Orthop Trauma Surg. 2011 Sep;131(9):1219-26. Epub 2011 Mar 1.

45. Goldman AH, Hoover KB. Source-to-detector distance and beam center do not affect radiographic measurements of acetabular morphology. Skeletal Radiol. 2017 Apr;46(4):477-81. Epub 2017 Jan 22.

46. Kappe T, Kocak T, Neuerburg C, Lippacher S, Bieger R, Reichel H. Reliability of radiographic signs for acetabular retroversion. Int Orthop. 2011 Jun;35(6):817-21. Epub 2010 May 10.

47. Khan M, Ranawat A, Williams D, Gandhi R, Choudur H, Parasu N, Simunovic N, Ayeni OR. Relationship between the alpha and beta angles in diagnosing cam-type femoroacetabular impingement on frog-leg lateral radiographs. Knee Surg Sports Traumatol Arthrosc. 2015 Sep;23(9):2595-600. Epub 2014 Jul 22.

48. Malviya A, Raza A, Witt JD. Reliability in the diagnosis of femoroacetabular impingement and dysplasia among hip surgeons: role of surgeon volume and experience. Hip Int. 2016 May 16;26(3):284-9. Epub 2016 Mar 30.

49. Hooper P, Oak SR, Lynch TS, Ibrahim G, Goodwin R, Rosneck J. Adolescent femoroacetabular impingement: gender differences in hip morphology. Arthroscopy. 2016 Dec;32(12):2495-502. Epub 2016 Aug 8.

50. Park S, Lee HY, Cuong PM, Won YY, Ji HM, Yoon SH, Kwack KS. Supine versus standing radiographs for detecting ischiofemoral impingement: a propensity score-matched analysis. AJR Am J Roentgenol. 2016 Jun;206(6): 1253-63. Epub 2016 Apr 12.

51. Stubbs AJ, Anz AW, Frino J, Lang JE, Weaver AA, Stitzel JD. Classic measures of hip dysplasia do not correlate with three-dimensional computer tomographic measures and indices. Hip Int. 2011 Sep-Oct;21(5):549-58.

52. Bittersohl B, Freitas J, Zaps D, Schmitz MR, Bomar JD, Muhamad AR, Hosalkar HS. EOS imaging of the human pelvis: reliability, validity, and controlled comparison with radiography. J Bone Joint Surg Am. 2013 May 1;95(9):e58.

53. Domb BG, Martin DE, Botser IB. Risk factors for ligamentum teres tears. Arthroscopy. 2013 Jan;29(1):64-73.

54. Mast NH, Impellizzeri F, Keller S, Leunig M. Reliability and agreement of measures used in radiographic evaluation of the adult hip. Clin Orthop Relat Res. 2011 Jan;469(1):188-99. Epub 2010 Jul 2.

55. Wensaas A, Gunderson RB, Svenningsen S, Terjesen T. Femoroacetabular impingement after slipped upper femoral epiphysis: the radiological diagnosis and clinical outcome at long-term follow-up. J Bone Joint Surg Br. 2012 Nov;94(11):1487-93.

56. Chadayammuri V, Garabekyan T, Jesse MK, Pascual-Garrido C, Strickland C, Milligan K, Mei-Dan O. Measurement of lateral acetabular coverage: a comparison between CT and plain radiography. J Hip Preserv Surg. 2015 Oct 22; 2(4):392-400.

57. Hanson JA, Kapron AL, Swenson KM, Maak TG, Peters CL, Aoki SK. Discrepancies in measuring acetabular coverage: revisiting the anterior and lateral center edge angles. J Hip Preserv Surg. 2015 Jun 13;2(3):280-6.

58. Colvin AC, Koehler SM, Bird J. Can the change in center-edge angle during pincer trimming be reliably predicted? Clin Orthop Relat Res. 2011 Apr;469(4):1071-4.

59. Barton C, Salineros MJ, Rakhra KS, Beaulé PE. Validity of the alpha angle measurement on plain radiographs in the evaluation of cam-type femoroacetabular impingement. Clin Orthop Relat Res. 2011 Feb;469(2):464-9.

60. Lequesne M, Malghem J, Dion E. The normal hip joint space: variations in width, shape, and architecture on 223 pelvic radiographs. Ann Rheum Dis. 2004 Sep;63(9):1145-51.

61. Espié A, Chaput B, Murgier J, Bayle-Iniguez X, Elia F, Chiron P. 45°-45°-30° frog-leg radiograph for diagnosing cam-type anterior femoroacetabular impingement: reproducibility and thresholds. Orthop Traumatol Surg Res. 2014 Dec;100(8):843-8. Epub 2014 Nov 6.

62. Espié A, Elia F, Murgier J, Chiron P, Chaput B. Modified head-neck offset for diagnosing anterior femoro-acetabular impingement. Int Orthop. 2016 Apr;40(4):687-95. Epub 2015 Jul 2.

63. Engesæter IØ, Laborie LB, Lehmann TG, Sera F, Fevang J, Pedersen D, Morcuende J, Lie SA, Engesæter LB, Rosendahl K. Radiological findings for hip dysplasia at skeletal maturity. Validation of digital and manual measurement techniques. Skeletal Radiol. 2012 Jul;41(7): 775-85. Epub 2011 Sep 25.

64. Levy DM, Hellman MD, Harris JD, Haughom B, Frank RM, Nho SJ. Prevalence of cam morphology in females with femoroacetabular impingement. Front Surg. 2015 Dec 1;2:61.

65. Clohisy JC, Nunley RM, Otto RJ, Schoenecker PL. The frog-leg lateral radiograph accurately visualized hip cam impingement abnormalities. Clin Orthop Relat Res. 2007 Sep; 462:115-21.

66. Farkas GJ, Cvetanovich GL, Rajan KB, Espinoza Orías AA, Nho SJ. Impact of femoroacetabular impingement morphology on gait assessment in symptomatic patents. Sports Health. 2015 Sep-Oct;7(5):429-36. Epub 2015 Jun 25.

67. Konan S, Rayan F, Haddad FS. Is the frog lateral plain radiograph a reliable predictor of the alpha angle in femoroacetabular impingement? J Bone Joint Surg Br. 2010 Jan; 92(1):47-50.

68. Zaltz I, Kelly BT, Hetsroni I, Bedi A. The crossover sign overestimates acetabular retroversion. Clin Orthop Relat Res. 2013 Aug; 471(8):2463-70.

69. Siebenrock KA, Kalbermatten DF, Ganz R. Effect of pelvic tilt on acetabular retroversion: a study of pelves from cadavers. Clin Orthop Relat Res. 2003 Feb;407:241-8.

70. Wassilew GI, Heller MO, Diederichs G, Janz V, Wenzl M, Perka C. Standardized AP radiographs do not provide reliable diagnostic measures for the assessment of acetabular retroversion. J Orthop Res. 2012 Sep;30(9): 1369-76. Epub 2012 Mar 12.

71. Hashemi SA, Dehghani J, Vosoughi AR. Can the crossover sign be a reliable marker of global retroversion of the acetabulum? Skeletal Radiol. 2017 Jan;46(1):17-21. Epub 2016 Oct 18.

72. Byrd JWT, Bardowski EA, Jones KS. Influence of Tönnis grade on outcomes of arthroscopic management of symptomatic femoroacetabular impingement. Arthroscopy. 2018 Aug;34(8):2353-6.

73. Larson CM, Giveans MR, Taylor M. Does arthroscopic FAI correction improve function with radiographic arthritis? Clin Orthop Relat Res. 2011 Jun;469(6):1667-76. Epub 2010 Dec 22.

74. Gicquel T, Gédouin JE, Krantz N, May O, Gicquel P, Bonin N; SoFCOT. Function and osteoarthritis progression after arthroscopic treatment of femoro-acetabular impingement: a prospective study after a mean follow-up of 4.6 (4.2-5.5) years. Orthop Traumatol Surg Res. 2014 Oct;100(6):651-6. Epub 2014 Aug 23.

75. Domb BG, Chaharbakhshi EO, Rybalko D, Close MR, Litrenta J, Perets I. Outcomes of hip arthroscopic surgery in patients with Tönnis grade 1 osteoarthritis at a minimum 5-year follow-up: a matched-pair comparison with a Tönnis grade 0 control group. Am J Sports Med. 2017 Aug;45(10):2294-302. Epub 2017 Jun 13.

76. Byrd JWT, Jones KS, Bardowski EA. Influence of Tönnis grade on outcomes of arthroscopy for FAI in athletes: a comparative analysis. J Hip Preserv Surg. 2018 Apr 24;5(2):162-5.