Arthroplasty Today 28 (2024) 101398

Contents lists available at ScienceDirect

Arthroplasty Today

journal homepage: http://www.arthroplastytoday.org/

Original Research

Radiographic Findings Associated With Mild Hip Dysplasia in 3869 Patients Using a Deep Learning Measurement Tool

Seong Jun Jang, MD[a, b, *], Daniel A. Driscoll, MD[b, c], Christopher G. Anderson, MD[d], Ruba Sokrab, MD[d], Dimitrios A. Flevas, MD, PhD[d], David J. Mayman, MD[c], Jonathan M. Vigdorchik, MD[c], Seth A. Jerabek, MD[c], Peter K. Sculco, MD[c]

[a] Weill Cornell College of Medicine, New York, NY, USA
[b] Department of Orthopedic Surgery, Hospital for Special Surgery, New York, NY, USA
[c] Adult Reconstruction and Joint Replacement Service, Hospital for Special Surgery, New York, NY, USA
[d] Stavros Niarchos Foundation Complex Joint Reconstruction Center, Hospital for Special Surgery, New York, NY, USA

ARTICLE INFO

Article history:
Received 4 October 2023
Received in revised form
10 March 2024
Accepted 2 April 2024

Keywords:
Hip dysplasia
Artificial intelligence
Deep learning
Osteoarthritis

ABSTRACT

Background: Hip dysplasia is considered one of the leading etiologies contributing to hip degeneration and the eventual need for total hip arthroplasty (THA). We validated a deep learning (DL) algorithm to measure angles relevant to hip dysplasia and applied this algorithm to determine the prevalence of dysplasia in a large population based on incremental radiographic cutoffs.

Methods: Patients from the Osteoarthritis Initiative with anteroposterior pelvis radiographs and without previous THAs were included. A DL algorithm automated 3 angles associated with hip dysplasia: modified lateral center-edge angle (LCEA), Tönnis angle, and modified Sharp angle. The algorithm was validated against manual measurements, and all angles were measured in a cohort of 3869 patients (61.2 ± 9.2 years, 57.1% female). The percentile distributions and prevalence of dysplastic hips were analyzed using each angle.

Results: The algorithm had no significant difference ($P > .05$) in measurements (paired difference: 0.3°-0.7°) against readers and had excellent agreement for dysplasia classification (kappa = 0.78-0.88). In 140 minutes, 23,214 measurements were automated for 3869 patients. LCEA and Sharp angles were higher and the Tönnis angle was lower ($P < .01$) in females. The dysplastic hip prevalence varied from 2.5% to 20% utilizing the following cutoffs: 17.3°-25.5° (LCEA), 9.4°-15.6° (Tönnis), and 41.3°-45.9° (Sharp).

Conclusions: A DL algorithm was developed to measure and classify hips with mild hip dysplasia. The reported prevalence of dysplasia in a large patient cohort was dependent on both the measurement and threshold, with 12.4% of patients having dysplasia radiographic indices indicative of higher THA risk.

Introduction

Hip osteoarthritis (OA) is the most common joint disease in which degenerative changes develop in the cartilage and surrounding bone, leading to pain, stiffness, loss of function, and limitations to daily activities [1]. Anatomical deformities in hip shape are considered key predisposing factors for the development of OA [2]. In particular, hip dysplasia is one of the prominent risk factors contributing to premature degenerative OA, and it involves an abnormality in the inclination, version, and volume of the acetabulum. This often results in a greater amount of hip joint reactive force concentrated on a smaller surface area. This focal wear ultimately leads to early degenerative changes in the hip joint and may necessitate the need for total hip arthroplasty (THA) [3].

Despite the risks associated with hip dysplasia and the potential need for THA, assessing hip dysplasia in adults remains unclear, especially in cases with mild acetabular deformity. Although rates between 3% and 5% are generally reported, the reported prevalence of hip dysplasia ranges from 1.7% to 20% in the general population based on different studies [4-6]. The wide range of prevalence may be attributed to variations in gender and ethnicity between cohorts

* Corresponding author. Department of Orthopaedic Surgery, Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021, USA. Tel: 1 703 851 4280.
E-mail address: seongjang22@gmail.com

https://doi.org/10.1016/j.artd.2024.101398

[6-8]. More importantly, it may also be a result of variations in the method of radiographic diagnosis. Several hip dysplasia radiographic indices, including the Sharp angle, the Tönnis angle, and the lateral center-edge angle (LCEA), are used in the literature to provide certain threshold values [9-11], but there exists considerable uncertainty regarding the cutoff for dysplasia classification with some studies even reporting ranges for "borderline dysplasia" [12] Furthermore, there are even differences in how these angles are measured between studies [13-16]. These inconsistencies ultimately contribute to the discrepancy in the assessment of hip dysplasia using these indices [3,17].

The differences in both the radiographic index and measurement method introduce subjectivity in defining and reporting the prevalence of hip dysplasia. Creating a validated, automated, and rapid measurement tool using emerging technology and applying it to a large cohort could uncover the exact variations inherent in dysplasia classification and prevalence reporting in large populations. Specifically, deep learning (DL) algorithms have already been used in several orthopaedic studies to investigate radiographic parameters, including those relevant to hip dysplasia [18-20]. Given that the radiographic evaluation of hip dysplasia can guide management decisions, these algorithms also have important clinical implications and applications. Namely, it may provide the rapid extraction of radiographic indices with the potential to aid in the early and accurate detection and evaluation of mild hip dysplasia.

The main aims of this study were to: 1) develop and validate a DL algorithm to automate angles relevant to hip dysplasia; and 2) apply this algorithm to a large cohort of patients to determine the prevalence of dysplasia radiographic findings based on varying radiographic parameters. We hypothesize 1) that the algorithm will be able to produce measurements with surgeon-level accuracy and 2) that the prevalence of dysplasia will vary substantially depending on the application of threshold values derived from different radiological parameters.

Material and methods

Patient and image selection

All patients from the Osteoarthritis Initiative (OAI) were included in this study. The OAI is a collection of publicly available data from 4796 patients who were prospectively enrolled in multiple institutions between 2004 and 2015 [21]. This analysis of a public database was exempt from Institutional Review Board review. From the OAI, patients with anteroposterior (AP) pelvis radiographs taken at the baseline enrollment were included. Any patient with a history of THA at the time of imaging was excluded. After inclusion and exclusion, the final cohort included 3869 pelvic images from independent patients. The mean age was 61.2 ± 9.2 years, and 57.1% of the cohort was female.

Dysplasia angles

Three angles associated with hip dysplasia were studied [14-16,22]. They included the modified Sharp angle, the Tönnis angle, and the modified LCEA. As described by Agus et al., the modified Sharp angle is the angle subtended by the line connecting the lateral sourcil edge to the ipsilateral teardrop and the line connecting the bilateral teardrops [14]. The Tönnis angle is the angle subtended by the line between the lateral and medial edges of the acetabular sourcil and the line connecting the bilateral pelvic teardrops. The modified LCEA, as described by Ogata et al., is the angle subtended by the line connecting the center of the femoral head to the lateral sourcil edge and the line approximating the vertical pelvic axis [16] (Fig. 1).

Image segmentation

Recent studies have leveraged the use of image segmentation with convolutional neural networks to measure clinically relevant orthopaedic parameters on radiographs including acetabular inclination and anteversion [23], hip joint center [24], and angles of dysplasia [18-20]. A convolutional neural network is a form of DL model used in computer vision tasks. Image segmentation's DL models produce annotated outputs of each pixel in an image as belonging to a specific object. We utilized image segmentation using a U-Net, a type of convolutional neural network, to automatically annotate relevant bony landmarks of interest as objects necessary for angle measurements (Fig. 2) (See Supplemental Methods for Model Creation and Metrics).

DL algorithm validation of angle measurements

A power analysis using an alpha of 0.05 and a power of 0.80 to detect a clinically relevant difference of 2 degrees between the DL algorithm and reader measurements produced a sample size of 73 hips. After algorithm creation, its accuracy in measuring relevant dysplasia angles was tested on 112 independent hips not used for model creation. Two trained readers, including a fellowship-trained orthopaedic surgeon, produced the ground-truth measurements for comparison. The model's and reader's classification for dysplastic hips based on literature-reported cutoffs [22] for the angles were compared to ensure accuracy in dysplasia detection. After validation, the model was then applied to the entire image cohort of 3869 pelvis radiographs.

Statistical analysis

Trained-reader measurements and DL measurements were compared using the interclass correlation coefficient (ICC) and an independent t-test after testing for normality. Bland-Altman plot

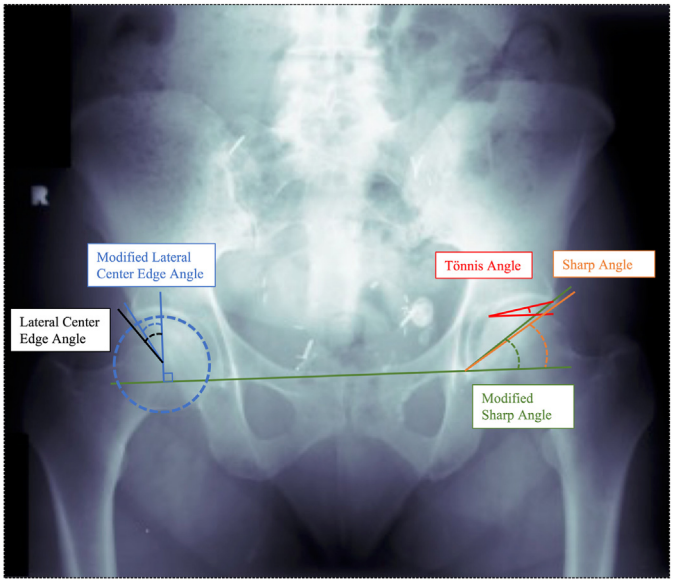

Figure 1. Modified LCEA, modified Sharp angle, and Tönnis angle. All angles were measured to the lateral edge of the sourcil.

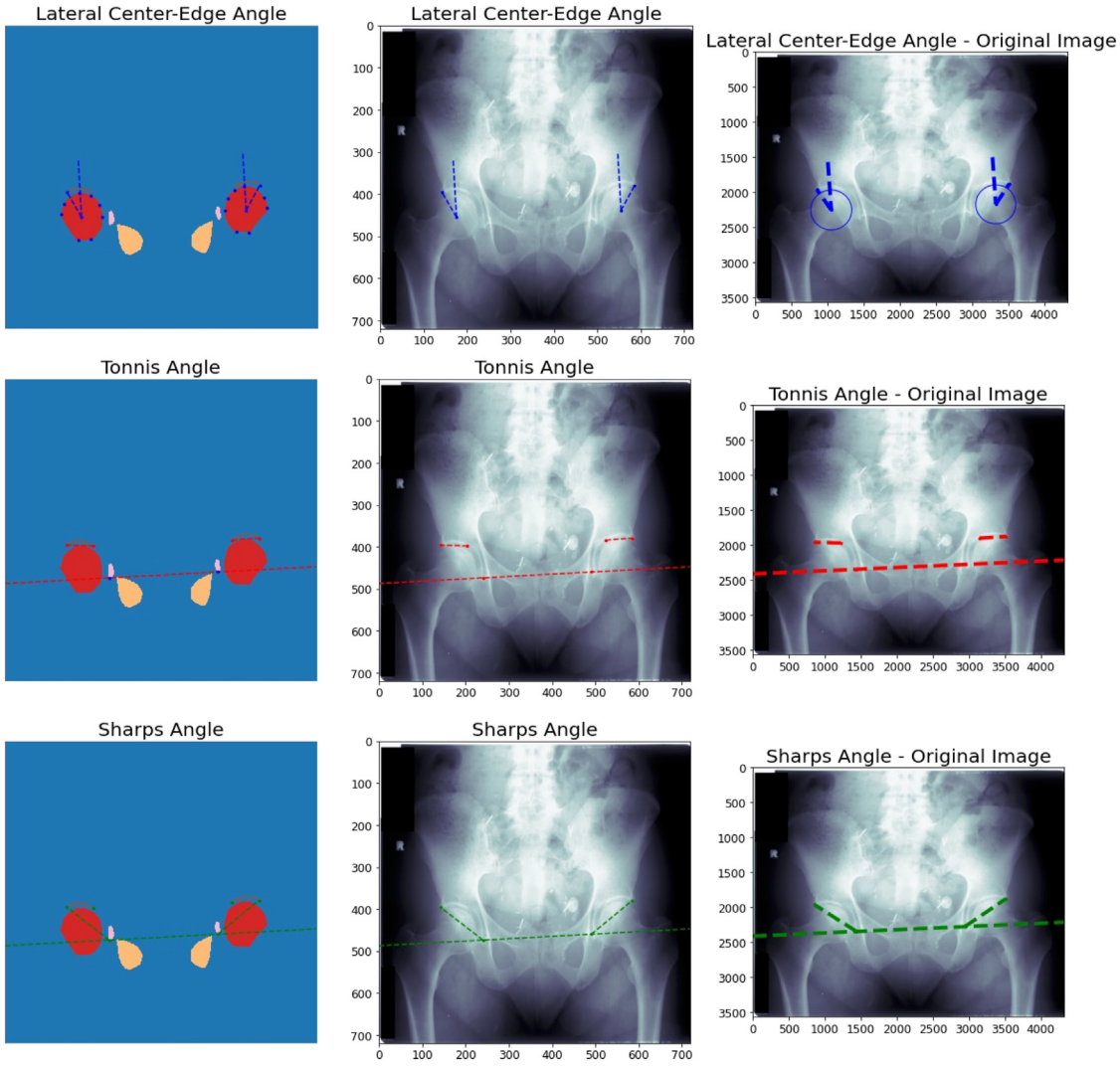

Figure 2. Deep learning segmentation outputs and automated angle measures. Column 1 = segmentation outputs and segmentation analysis. Column 2 = parameter generation. Column 3 = resizing and angle generation.

analyses were assessed for systematic errors and biases. Reader and algorithm agreement for classifying hips as dysplastic were assessed using Cohen's kappa. The Pearson correlation coefficient was also determined between the 3 angles.

In the entire cohort, percentile distributions of hip dysplasia based on incremental radiographic cutoffs for the 3 angles were calculated. The angles were also compared between males and females using an independent t-test.

All statistical analyses were conducted in Python using pingouin (V0.5.3), sklearn (V1.2.2), and SciPy (V1.10.1) packages.

Results

Deep learning algorithm performance

The DL model was optimized within 100 epochs of training. The optimized model had a multiclass dice segmentation coefficient of 0.89. When the DL algorithm was applied to the entire cohort of images, it measured 3 angles for each hip at a rate of 2.17 seconds. For the entire cohort of 3869 images (7738 hips), this totaled 23,214 measurements in 140 minutes.

Deep learning vs surgeon measurements

The ICC between the 2 trained readers was 0.81-0.91 for the 3 angles (Fig. 3). The ICC between the mean of the readers and the DL algorithm was higher with all ICCs >0.88 (Table 1). There was no significant difference between the readers and the DL-produced measurements on the subcohort of 112 hips ($P > .05$) for all 3 angles. The mean paired difference was highest for the LCEA (−0.7° standard deviation [SD]: 2.6°, interquartile range −2.3° to 0.9°). The Bland-Altman analysis revealed no significant bias in the DL measurements (Fig. 4).

Utilizing a cutoff of 22.5° for the modified LCEA, 10.0° for the Tönnis angle, and 40.5° for the modified Sharp angle, the kappa was 0.88, 0.78, and 0.81, respectively, between the algorithm and readers for classifying this subcohort of hips as dysplastic or non-dysplastic. Classification matrices are depicted in Figure 5.

Modified LCEA

The average LCEA was 31.1° (7.0°) (Table 2). There was a significant difference between males and females with females having a higher modified LCEA angle (31.6° (7.0°) vs 30.5° (7.0°), $P < .01$).

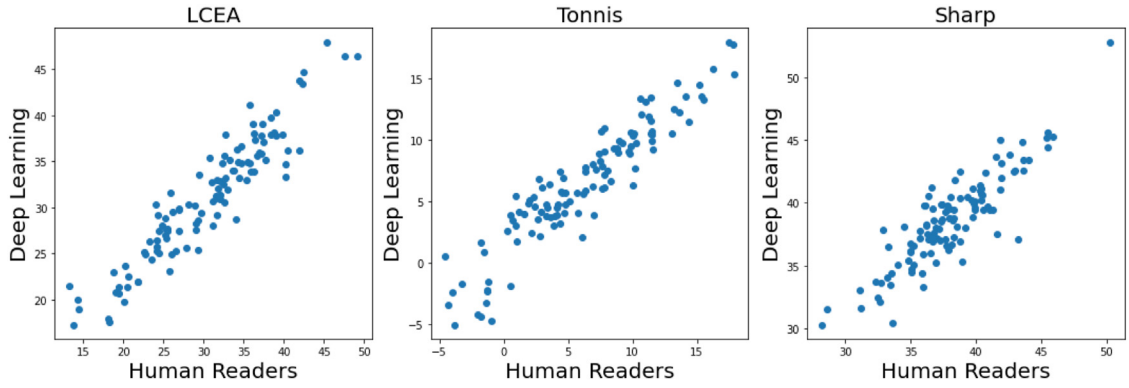

Figure 3. Deep learning algorithm measurement against reader means.

When considering the percentile cutoffs for the modified LCEA, a value of 25.5° indicated the 20th percentile (20% prevalence), whereas a value of 19.5° indicated the 5th percentile (5% prevalence) (Fig. 6).

Modified sharp angle

The average Sharp angle was 38.0° (4.0°) (Table 2). There was a significant difference between males and females with females having a higher modified Sharp angle (38.5° [4.1°] vs 37.4° [3.9°], $P < .01$). A value of 41.3° indicated the 80th percentile (20% prevalence), whereas a value of 44.5° indicated the 95th percentile (5% prevalence) (Fig. 7).

Tönnis angle

The average Tönnis angle was 5.1° (5.3°) (Table 2). Males had a higher Tönnis angle (5.3° [5.3°] vs 4.9° [5.3°], $P < .01$). A value of 9.4° indicated the 80th percentile whereas a value of 13.7° indicated the 95th percentile (Fig 8).

Angle relationships

The Pearson correlation coefficient was −0.79 ($P < .01$) between the modified LCEA and Tönnis angle, −0.67 ($P < .01$) between the modified Sharp and modified LCEA, and 0.59 ($P < .01$) between the modified Sharp and Tönnis angle.

Discussion

The main finding of this study was that a rapid imaging-based DL algorithm was developed and validated to automate angles relevant to hip dysplasia on pelvis radiographs and classify hips on the spectrum of hip dysplasia. The algorithm demonstrated analytical capabilities that were in accordance with landmarks and angles relevant to arthroplasty surgeons in practice. Furthermore, the application of this validated tool to a large, older patient cohort demonstrated that the reported prevalence of hip dysplasia was largely dependent on the type and threshold of radiographic indices used. The results for each radiographic marker in this large, elderly cohort were in concordance with previous studies investigating radiographic indices of hip dysplasia in younger populations. Thus, we suggest that radiographic markers of dysplasia may be consistent in populations at varying ages after skeletal maturity, even in a population more at-risk for OA and the need for eventual THA.

DL has been implemented in several medical disciplines, including orthopedic surgery, to investigate the propensity for automated algorithms to identify clinically relevant features in visual data including ultrasound images, radiographs, and advanced imaging [24-28]. In computer vision, DL functions by applying neural networks to images to experientially learn patterns associated with image features. Information derived from imaging-based analyses may then be further used for clinical applications or to provide insight into outcomes and overall prognosis [24,29,30]. The current study applied DL to automate various hip joint measurements relevant to dysplasia on AP pelvis radiographs with strong accuracy and consistency. The agreement (ICCs) between the readers and the DL algorithm was >0.88 for all angles, and there was no statistical difference in angle measurements. Critically, the time the algorithm needed to measure 3 angles for each hip was only 2.17 seconds. With the integration of this tool into hospital image systems, there is value in speeding up clinical and research workflows relevant to hip dysplasia and in decreasing the burden of manual measurements on clinicians and researchers for large cohort studies. To demonstrate this application, this study further utilized this algorithm on a cohort of 3869 images (7738 hips), and the algorithm produced 23,214 unique measurements in 140 minutes for analysis.

The analysis of these automated measures demonstrated that the prevalence of radiographic indices indicative of hip dysplasia is dependent on the measurement and cutoff used to define dysplasia. Indeed, despite the risks associated with dysplasia and the need for long-term THA, defining hip dysplasia in adults remains unstandardized, especially in less severe cases with mild acetabular deformity. The prevalence of hip dysplasia in the general population has been reported in the range of 1.7%-20% based on the type of study using various radiographic definitions [4-6].

Table 1
Comparison of measurements between algorithm and readers.

N = 112 hips	Reader mean (±SD)	Reader 1 and 2 ICC [95% CI]	DL algorithm mean (±SD)	Reader mean vs DL algorithm ICC [95% CI]	Paired difference mean (±SD, interquartile range)	*P* value
LCEA	30.3° (7.5°)	0.91 [0.87-0.94]	31.0° (6.6°)	0.93 [0.90-0.95]	−0.7° (2.6°, −2.3° to 0.9°)	.44
Tonnis	6.2° (5.3°)	0.87 [0.81-0.91]	6.5° (5.0°)	0.93 [0.90-0.95]	−0.3° (1.9°, −1.4° to 0.9°)	.67
Sharp	38.0° (3.7°)	0.81 [0.73-0.86]	38.6° (3.6°)	0.88 [0.83-0.92]	−0.6° (1.8°, −1.6° to 0.5°)	.24

CI, confidence interval.

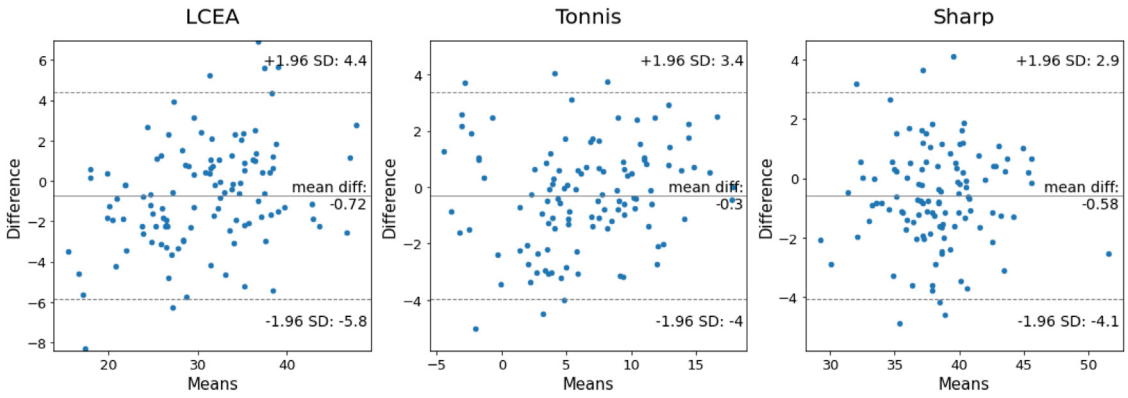

Figure 4. Bland-Altman analysis of measurements against reader means.

In this study, the prevalence of dysplastic hips in this cohort varied from 2.5%-20% utilizing the following cutoffs: 17.3°-25.5° for LCEA, 15.6°-9.4° for Tönnis angle, and 45.9°-41.3° for Sharps angle. These results are in close agreement with the findings of Laborie et al., which determined that cutoff values of 17°-18° for the modified LCEA, 15°-16° for the Tonnis angle, and 46°-47° for the modified Sharps angle resulted in a 2.5% prevalence of developmental hip dysplasia in a Norwegian population [8]. Interestingly, their study analyzed 2011 individuals at skeletal maturity with an average age of 18.6 (SD 0.6), and all measurements required the manual identification of landmarks on AP pelvis radiographs by an observer. This current study not only utilized DL to rapidly automate these measurements in a larger cohort without the need of an observer (n = 3869), but it also concurred with Laborie et al.'s [8] population analyses of these angles in an older cohort with a mean age of 61.2 (SD 9.2) years. This suggests that the radiographic markers of dysplasia may be consistent in populations at varying ages after skeletal maturity. Importantly, an exact threshold to determine the true prevalence of hip dysplasia is still debated [8]. Jang et al. recently analyzed the risk of long-term THA in this cohort of patients from the OAI and reported that a modified LCEA <22.4° and a Tönnis angle >10.8° conferred a higher risk of THA. Using these values as a clinically relevant threshold for classifying hip dysplasia, the prevalence of dysplasia associated with THA risk in this cohort would be 12.4% using either the LCEA or Tönnis angle. Jang et al. also reported that a modified LCEA <19.2° conferred an even higher risk of THA, which 4.5% of hips in this cohort had [31].

It is critical to note that this study used specific definitions of radiographic measurements of hip dysplasia (Fig. 1). The modified Sharp angle measured the femoral head coverage using the lateral margin of the sourcil, which was the subchondral bony condensation in the acetabular roof [14]. Similarly, the modified LCEA used the lateral margin of the sourcil as previously described by Ogata et al. [16] Many studies have utilized the LCEA described by Wilberg and the Sharp angle described by Sharp, which both measure to the lateral edge of the acetabular roof [13-16]. Studies comparing these methods have found that measuring to the edge of the acetabular roof may overestimate coverage and have proposed the lateral edge of the sourcil as a more reliable and clinically relevant landmark [15,32]. Therefore, this study investigated the modified LCEA and Sharp angles, and the reported findings are only relevant to these specific definitions.

In this study, we demonstrated the application and value of a DL tool to provide rapid radiographic measurements of dysplasia. The clinical value of this algorithm is that it is a potential screening tool for mild hip dysplasia in patients undergoing assessment for hip pathology. Critically, patients with hip OA complain of pain that can be a disabling symptom [1], and radiographs are routinely obtained as part of an orthopaedic evaluation [22]. Although 3-dimensional imaging, such as computed tomography scanning and magnetic resonance imaging, is available, plain radiographs remain the gold standard for the initial evaluation and a more economical screening option [3,33]. Anatomical deformities that can be identified on plain radiographs may represent an important predisposing factor for OA development [2,34], and previous studies have shown that

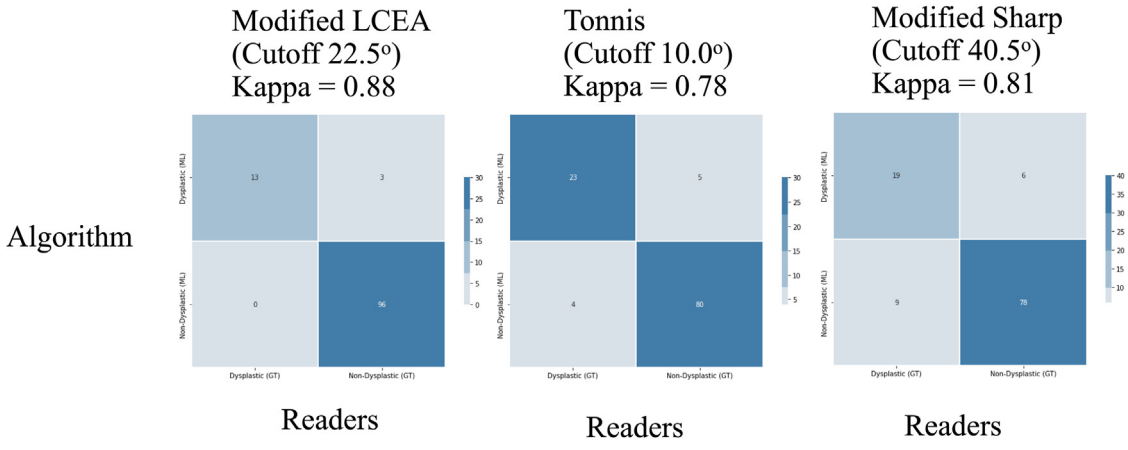

Figure 5. Classification matrices for classifying hips as dysplastic or nondysplastic based on literature-specified thresholds.

Table 2
Percentile distribution of dysplasia angles in cohort.

Measurement N = 3318 male hips, 4420 female hips	Degree	P (male vs female)	2.5 %tile	5.0 %tile	10 %tile	20 %tile	80 %tile	90 %tile	95.0 %tile	97.5 %tile
LCEA (Ogata)	31.1° (7.0°)	<0.01	17.3°	19.5°	22.3°	25.5°	36.9°	40.0°	42.5°	44.6°
Male	30.5° (7.0°)		16.6°	19.2°	21.6°	24.6°	36.4°	39.5°	41.9°	43.8°
Female	31.6° (7.0°)		17.7°	19.8°	23.1°	26.1°	37.3°	40.4°	42.9°	45.4°
Tonnis	5.1° (5.3°)	<0.01	−5.3°	−3.6°	−1.6°	0.7°	9.4°	11.8°	13.7°	15.6°
Male	5.3° (5.3°)		−4.9°	−3.2°	−1.3°	0.9°	9.8°	12.1°	13.8°	15.8°
Female	4.9° (5.3°)		−5.7°	−3.9°	−1.9°	0.6°	9.2°	11.5°	13.3°	15.5°
Sharp (Agus)	38.0° (4.0°)	<0.01	30.2°	31.3°	32.9°	34.7°	41.3°	43.1°	44.5°	45.9°
Male	37.4° (3.9°)		29.7°	30.9°	32.4°	34.2°	40.7°	42.2°	43.7°	45.0°
Female	38.5° (4.1°)		30.4°	31.8°	33.4°	35.2°	41.7°	43.5°	45.0°	46.5°

hip dysplasia, especially in borderline cases, is often overlooked radiologically and can lead to delayed diagnoses [18]. By having an automated tool for quantifying deformities relevant to hip dysplasia, even in subtle cases, patients that are predisposed to develop OA may be more readily identified in a standardized manner at clinical visits or in large patient registries. This study demonstrated that after setting cutoffs of 22.5° for the modified LCEA, 10.0° for the Tönnis angle, and 40.5° for the modified Sharp angle for hip dysplasia, the kappa was 0.88, 0.78, and 0.81, respectively, between the algorithm and observers. Thus, there is potential for leveraging DL for dysplasia screening in the setting of hip pathology. Nonetheless, for this application to be executed properly and further developed, the algorithms must be externally validated, radiographs must be screened for good quality, measurements must be standardized, and the ongoing possibility of measurement error and misinterpretation by the DL algorithm must be kept in mind.

This study had several limitations. First, all cohort analyses of radiographic indices were conducted on automated measurements using a DL algorithm. However, these measurements were all validated in a powered subcohort, and the final analyses were in concordance with previous studies. Furthermore, most of the OAI radiographs did not exhibit a large migration of the femoral head in the acetabulum. Therefore, this algorithm is only relevant to mild dysplasia where the acetabulum and sourcil are radiographically intact. Future studies would benefit from the inclusion of patients with higher-grade dysplasia to expand the clinical utility of the algorithm. Likewise, different images may have had varying degrees of hip flexion and extension (ie, more inlet and outlet views) across the database, and all radiographs were measured without any adjustment. Furthermore, other indices of hip dysplasia, including the alpha angle, the acetabular depth-width ratio, and the femoral head extrusion index, were not automated nor analyzed in this cohort but are of future interest. Finally, these measurements and cutoffs reported in this study may be specific only to populations with characteristics similar to those of the OAI study group, which may have a higher prevalence of OA.

Conclusions

The development of a radiographic DL algorithm for hip dysplasia parameters provides automation for the evaluation and assessment of an anatomical risk factor for OA. The algorithm demonstrated strong performance and reliability in measuring angles relevant to hip dysplasia, and its application to a large cohort revealed key variations in using these angles to determine hip dysplasia prevalence. Using literature thresholds to classify hip dysplasia with a higher risk of needing eventual THA, the prevalence would be 12.4% in this cohort. Further development and validation of this tool also hold clinical value in allowing for the rapid screening of patients with mild dysplasia.

Conflicts of interest

D. Mayman receives royalties from Smith and Nephew and OrthAlign, is a paid consultant for Stryker, has stock options in Cymedica, Imagen, MiCare Path, Wishbone, and OrthAlign, and is a board/committee member of the Hip and Knee Society. J. Vigdorchik receives royalties from Corin and Depuy Synthes, is a paid consultant for Stryker and Depuy Synthes, has stock options in Aware, Corin, Intellijoint, Motion Insights, Ortho AI, and Polaris, is an editorial board member of BJJ, and is a board/committee

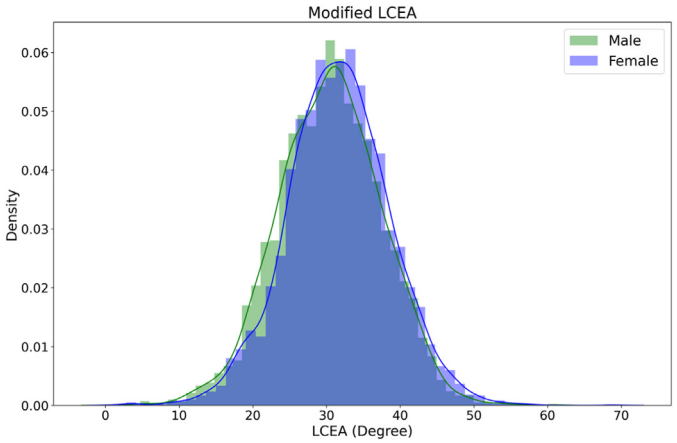

Figure 6. Histogram of the modified LCEA distribution (n = 7738 hips).

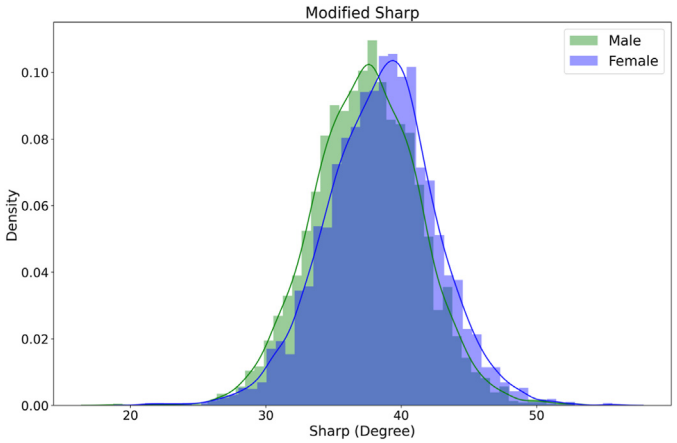

Figure 7. Histogram of the modified Sharp angle distribution (n = 7738 hips).

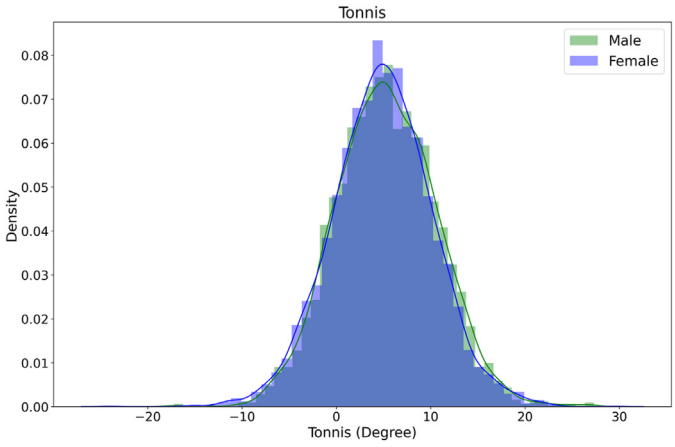

Figure 8. Histogram of the Tönnis angle distribution (n = 7738 hips).

member of AAHKS. S. Jerabek is a paid consultant for Stryker, receives royalties from Stryker, is a speaker for Stryker, has stock options in Imagen, receives research support from Stryker, and receives other financial support/royalties from Wolters Kluwer Health. P. Sculco is a speaker for DePuy, A Johnson & Johnson Company, EOS Imaging, and Intellijoint Surgical, is a paid consultant for DePuy, A Johnson & Johnson Company, EOS Imaging, and Intellijoint Surgical, Lima Corporate, and Zimmer, has stock options in Intellijoint Surgical and Parvizi Surgical Innovation, and receives research support from Intellijoint Surgical. All other authors declare no potential conflicts of interest.

For full disclosure statements refer to https://doi.org/10.1016/j.artd.2024.101398.

CRediT authorship contribution statement

Seong Jun Jang: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Daniel A. Driscoll:** Validation, Investigation, Formal analysis, Writing – original draft. **Christopher G. Anderson:** Conceptualization, Investigation, Methodology, Validation, Writing – original draft. **Ruba Sokrab:** Data curation, Formal analysis, Resources, Writing – original draft. **Dimitrios A. Flevas:** Investigation, Methodology, Validation, Writing – original draft. **David J. Mayman:** Conceptualization, Project administration, Supervision, Writing – review & editing. **Jonathan M. Vigdorchik:** Conceptualization, Project administration, Supervision, Writing – review & editing. **Seth A. Jerabek:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Peter K. Sculco:** Writing – review & editing, Supervision, Conceptualization, Project administration, Resources.

Acknowledgments

Data and/or research tools used in the preparation of this manuscript were obtained and analyzed from the controlled access datasets distributed from the Osteoarthritis Initiative (OAI), a data repository housed within the NIMH Data Archive (NDA). OAI is a collaborative informatics system created by the National Institute of Mental Health and the National Institute of Arthritis, Musculoskeletal and Skin Diseases (NIAMS) to provide a worldwide resource to quicken the pace of biomarker identification, scientific investigation, and OA drug development. Dataset identifier(s): 10.15154/1524216.

This manuscript was prepared using the OAI public use data set and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.artd.2014.12.004.

References

[1] Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. Lancet 2019;393:1745–59.
[2] Baker-Lepain JC, Lane NE. Relationship between joint shape and the development of osteoarthritis. Curr Opin Rheumatol 2010;22:538–43.
[3] 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;21:549–58.
[4] Gosvig KK, Jacobsen S, Sonne-Holm S, Palm H, Troelsen A. Prevalence of malformations of the hip joint and their relationship to sex, groin pain, and risk of osteoarthritis: a population-based survey. J Bone Joint Surg Am 2010;92:1162–9.
[5] Jacobsen S, Sonne-Holm S. Hip dysplasia: a significant risk factor for the development of hip osteoarthritis. A cross-sectional survey. Rheumatology 2005;44:211–8.
[6] Engesaeter IØ, Laborie LB, Lehmann TG, Fevang JM, Lie SA, Engesaeter LB, et al. Prevalence of radiographic findings associated with hip dysplasia in a population-based cohort of 2081 19-year-old Norweggians. Bone Joint Lett J 2013;95-B:279–85.
[7] Inoue K, Wicart P, Kawasaki T, Huang J, Ushiyama T, Hukuda S, et al. Prevalence of hip osteoarthritis and acetabular dysplasia in French and Japanese adults. Rheumatology (Oxford) 2000;39:745–8.
[8] Laborie LB, Engesæter IØ, Lehmann TG, Sera F, Dezateux C, Engesæter LB, et al. Radiographic measurements of hip dysplasia at skeletal maturity–new reference intervals based on 2,038 19-year-old Norwegians. Skeletal Radiol 2013;42:925–35.
[9] Sharp IK. Acetabular dysplasia. J Bone Joint Surg Br 1961;43-B:268–72.
[10] Tonnis D. Congenital dysplasia and dislocation of the hip in children and adults. Berlin, Germany: Springer-Verglag; 1987. p. 116–21. https://link.springer.com/book/10.1007/978-3-642-71038-4. [Accessed 6 May 2024].
[11] Tannast M, Hanke MS, Zheng G, Steppacher SD, Siebenrock KA. What are the radiographic reference values for acetabular under- and overcoverage? Clin Orthop Relat Res 2015;473:1234–46.
[12] Troelsen A, Rømer L, Kring S, Elmengaard B, Søballe K. Assessment of hip dysplasia and osteoarthritis: variability of different methods. Acta Radiologica 2010;51:187–93.
[13] Wiberg G. Studies on dysplastic acetabula and congenital subluxation of the hip joint: with special reference to the complication of osteoarthritis. Acta Chir Scand 1939;83:1–135.
[14] Agus H, Bicimoglu A, Omeroglu H, Tumer Y. How should the acetabular angle of Sharp be measured on a pelvic radiograph? J Pediatr Orthop 2023;22:228–31.
[15] 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;2:280–6.
[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;72:190–6.
[17] Clohisy JC, Carlisle JC, Trousdale R, Kim Y-J, Beaule PE, Morgan P, et al. Radiographic evaluation of the hip has limited reliability. Clin Orthop Relat Res 2009;467:666–75.
[18] Jensen J, Graumann O, Overgaard S, Gerke O, Lundemann M, Haubro M, et al. A deep learning algorithm for radiographic measurements of the hip in adults-A reliability and agreement study. Diagnostics 2022;12:2597.
[19] Park H, Jeon K, Cho Y, Kim S, Lee S, Choi G, et al. Diagnostic performance of a new convolutional neural network algorithm for detecting developmental dysplasia of the hip on anteroposterior radiographs. Korean J Radiol 2021;22:612–23.
[20] Li Q, Zhong L, Huang H, Liu H, Qin Y, Wang Y, et al. Auxiliary diagnosis of developmental dysplasia of the hip by automated detection of Sharp's angle on standardized anteroposterior pelvic radiographs. Medicine 2019;98:e18500.
[21] Lester G. The osteoarthritis initiative: a NIH public-private partnership. HSS J 2012;8:62–3.
[22] Mannava S, Geeslin AG, Frangiamore SJ, Cinque ME, Geeslin MG, Chahla J, et al. Comprehensive clinical evaluation of femoroacetabular impingement: part 2, plain radiography. Arthrosc Tech 2017;6:e2003–9.
[23] Rouzrokh P, Wyles CC, Philbrick KA, Ramazanian T, Weston AD, Cai JC, et al. A deep learning tool for automated radiographic measurement of acetabular component inclination and version after total hip arthroplasty. J Arthroplasty 2021;36:2510–2517.e6.
[24] Jang SJ, Kunze KN, Vigdorchik JM, Jerabek SA, Mayman DJ, Sculco PK. John charnley award: deep learning prediction of hip joint center on standard pelvis radiographs. J Arthroplasty 2022;37:S400–407.e1.

[25] Karnuta JM, Murphy MP, Luu BC, Ryan MJ, Haeberle HS, Brown NM, et al. Artificial intelligence for automated implant identification in total hip arthroplasty: a multicenter external validation study exceeding two million plain radiographs. J Arthroplasty 2023;38:1998–2003.e1.
[26] Arunachalam HB, Mishra R, Daescu O, Cederberg K, Rakheja D, Sengupta A, et al. Viable and necrotic tumor assessment from whole slide images of osteosarcoma using machine-learning and deep-learning models. PLoS One 2019;14:e0210706.
[27] Tang H, Sun N, Shen S. Improving generalization of deep learning models for diagnostic pathology by increasing variability in training data: experiments on osteosarcoma subtypes. J Pathol Inf 2021;12:30.
[28] Ghasseminia S, Lim AKS, Concepcion NDP, Kirschner D, Teo YM, Dulai S, et al. Interobserver variability of hip dysplasia indices on sweep ultrasound for novices, experts, and artificial intelligence. J Pediatr Orthop 2022;42:e315–23.
[29] Tolpadi AA, Lee JJ, Pedoia V, Majumdar S. Deep learning predicts total knee replacement from magnetic resonance images. Sci Rep 2020;10:6371.
[30] Lu L, Dercle L, Zhao B, Schwartz LH. Deep learning for the prediction of early on-treatment response in metastatic colorectal cancer from serial medical imaging. Nat Commun 2021;12:6654.
[31] Jang SJ, Fontana MA, Kunze KN, Anderson CG, Sculco TP, Mayman DJ, et al. An interpretable machine learning model for predicting 10-year total hip arthroplasty risk. J Arthroplasty 2023;38:S44–50.e6.
[32] Omeroglu H, Bicimoglu A, Agus H, Tumer Y. Measurement of center-edge angle in developmental dysplasia of the hip: a comparison of two methods in patients under 20 years of age. Skeletal Radiol 2002;31:25–9.
[33] Philippon MJ, Briggs KK, Carlisle JC, Patterson DC. Joint space predicts THA after hip arthroscopy in patients 50 years and older. Clin Orthop Relat Res 2013;471:2492–6.
[34] Murray RO. The aetiology of primary osteoarthritis of the hip. BJR 1965;38:810–24.

Appendix References

[1] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Cham: Springer International Publishing; 2015.
[2] Howard J, Gugger S. fastai: a layered API for deep learning. ArXiv 2020. abs/2002.04688.
[3] Zou KH, et al. Statistical validation of image segmentation quality based on a spatial overlap index. Acad Radiol 2004;11(2):178–89.
[4] Schwartz JT, et al. Deep learning automates measurement of spinopelvic parameters on lateral lumbar radiographs. Spine (Phila Pa 1976) 2021;46(12):E671–8.