Original Clinical Article

JOURNAL OF CHILDREN'S ORTHOPAEDICS

The lateral edge and sourcil acetabular indices for surgical decision-making in developmental dysplasia of the hip

Alpesh Kothari[1]
Saqib Noor[1]
Connor L. Maddock[1]
Jan H.H. Vanderstappen[2]
Catharine S. Bradley[1]
Simon P. Kelley[1]

Abstract

Purpose The acetabular index (AI) is a radiographic measure that guides surgical decision-making in developmental dysplasia of the hip (DDH). Two AI measurement methods are described; to the lateral edge of the acetabulum (AI-L) and to the lateral edge of the sourcil (AI-S). The purpose of this study was to determine the level of agreement between AI-L and AI-S on the diagnosis and degree of acetabular dysplasia in DDH.

Methods A total of 35 patients treated for DDH with Pavlik harness were identified. The AI-L and AI-S were measured on radiographs (70 hips) at two and five years of age. AI-L and AI-S were then transformed relative to published normative data (tAI-L and tAI-S). Bland-Altman plots, linear regression and heat mapping were used to evaluate the agreement between tAI-L and tAI-S.

Results There was poor agreement between tAI-S and tAI-L on the Bland-Altman plots with wide limits of agreement and no proportional bias. The two AI measurements were in agreement as to the presence and severity of dysplasia in only 63% of hips at two years of age and 81% at five years of age, leaving the remaining hips classified as various combinations of normal, mildly and severely dysplastic.

Conclusion AI-L and AI-S have poor agreement on the presence or degree of acetabular dysplasia in DDH and cannot be used interchangeably. Clinicians are cautioned to prudently evaluate both measures of AI in surgical decision-making.

[1] The Hospital for Sick Children, Toronto, Canada
[2] Orthokliniek, Department of Orthopaedic Surgery, AZ Nikolaas, Sint-Niklaas, Belgium

Correspondence should be sent to Simon Kelley, The Hospital for Sick Children, Division of Orthopaedic Surgery, 555 University Avenue, Toronto, ON, Canada M5G 1X8.
E-mail: simon.kelley@sickkids.ca

Level of evidence: I

Cite this article: Kothari A, Noor S, Maddock CL, Vanderstappen JHH, Bradley CS, Kelley SP. The lateral edge and sourcil acetabular indices for surgical decision-making in developmental dysplasia of the hip. *J Child Orthop* 2020;14. DOI: 10.1302/1863-2548.14.200199

Keywords: DDH; acetabular index; radiographic measurement; diagnosis

Introduction

Hip dysplasia in adulthood is associated with premature hip osteoarthritis.[1-3] As such, radiographic monitoring of hip development in children with developmental dysplasia of the hip (DDH) is critical to guide the timing of surgical interventions to correct residual dysplasia in an attempt to ensure a durable pain free hip and to mitigate the need for early hip arthroplasty. The most common surgical procedure used to correct residual dysplasia in children is the acetabular osteotomy, which is typically undertaken prior to five years of age, as after this age remodelling is less assured.[4] Reported indications for this procedure are acetabular dysplasia requiring 15° to 20° of correction,[5] yet defining acetabular dysplasia in the context of surgical decision-making remains controversial in the literature.[6,7]

Plain pelvic radiographs serve as an inexpensive and readily available method to appraise developing hip morphology. A number of radiographic measures are used to assess hip dysplasia in children 18 months to six years with one of the most common being the acetabular index (AI).[8] First described by Hilgenreiner in 1925,[8] AI gives an indication of acetabular roof inclination and is classically measured from the triradiate cartilage to the lateral most edge of the acetabulum (AI-L). Normative reference values for this measure of AI have been previously published by Tönnis and Brunken.[9]

More recently, using cross-sectional imaging techniques, it has been hypothesized that a better representation of the degree of acetabular dysplasia is obtained by an alternative measure of AI using the lateral edge of the sourcil (AI-S) which characterizes dysplasia at the mid-superior portion rather than the anterolateral edge of the bony acetabulum as defined by the AI-L.[6,10] Normative reference values for this measure of AI-S have recently been published by Novais et al.[7]

While both measurement methods are described as the AI and both have been demonstrated to have excellent inter-rater and intra-rater reliability with intra-class correlation coefficient (ICC) values of 0.94, 0.93 for AI-L and 0.91, 0.90 for AI-S respectively,[11] there are clear differences in their normative reference value ranges which has important implications for surgical decision-making. In the clinical setting, surgeons will confidently intervene if the AI is greater than two SDS from the normative mean based on age, sex and laterality. However, which measure, or if both should be considered in this context, remains unclear. Thus, the aim of this study was to evaluate the relationship between AI to the lateral edge (AI-L) and AI to the sourcil (AI-S) in a population of young children with DDH as a means to further inform surgical decision-making.

Our hypotheses were that there would be no difference between AI-L and AI-S when placed in the context of published normative data, that both AI-L and AI-S would strongly agree on the diagnosis and degree of dysplasia, and as such one could use either measure of AI when considering surgical management of residual acetabular dysplasia.

Patients and Methods

Study design

The imaging of a cohort of paediatric patients identified through an existing departmental, prospectively collected DDH database was evaluated in this study.

Inclusion criteria

Patients were included in the study if they had previously been managed for unilateral or bilateral DDH with a Pavlik harness prior to six months of age at our institution. All patients had undergone standard anteroposterior pelvic radiographs at two and five years of age. In addition, the decision was made to include only female patients as being female alone, in the absence of other risk factors accounts for > 75% of those with DDH,[12] and thus allows for the best representation of DDH at the time of this study. Families had previously provided consent by agreeing to be included in the DDH database.

Exclusion criteria

Patients were excluded if they had a teratologic dislocation or any neuromuscular diagnosis, any surgical intervention for DDH or any intervention for DDH outside of our institution. Radiographs were assessed for quality by the research team based on the established standards described by Tönnis[13] and patients with poor quality radiographs at either the two- or five-year-old mark were excluded.

Sample size

A required sample size of 64 hips was derived using an alpha level of 0.05, power of 0.8 and an effect size of 0.5 (medium effect) based on Cohen's sample size tables for Student's *t*-tests.[14]

Data collection

Radiographic measurements of both AI-L and AI-S on standard anteroposterior pelvic radiographs at two and five years of age were completed by one member of the research team (AK), a fellowship-trained paediatric orthopaedic surgeon. The methods for measuring these angles and a depiction of the measures being distinct, replicated from a previous publication, are shown in Figure 1.[11] The previously established reliability of both measures obviated the need for multiple raters in this study.[11] All AI-L and AI-S values were then referenced to their published normative data sets[7,9] and redefined as the transformed AI (tAI) in terms of SDS away from their respective published mean based on sex, age and side. For example, an AI-L of 17.5° in a left hip of a two-year-old female patient would be transformed to a tAI-L of + 0.56, which means 0.56 SD above the AI-L population mean based on age, sex and laterality (mean 15.2° (SD 4.1°)).

Statistical analysis

This was undertaken using SPSS v 22.0 (IBM, Chicago, Illinois). Differences between the absolute values of AI-L and AI-S at two and five years of age were assessed using paired two-tailed Student's *t*-tests ($p < 0.05$). Bland-Altman plots were used to evaluate the agreement between tAI-L and tAI-S at the two and five year time points.[15,16] The Bland-Altman method is considered the benchmark for comparing two measures of the same variable and is a simple way to evaluate bias between the mean differences between measurement techniques and to estimate an agreement interval within which 95% of the differences of one method, compared with the other, exist.[15] However, the Bland-Altman plot only defines the limits of agreement (LOA) and does not indicate if the limits are clinically acceptable. This must be determined *a priori*, thus, for the purposes of this study, an *a priori* range of LOA of 1 SD was deemed as the maximum allowed difference to demonstrate agreement between tAI-L and tAI-S. Linear regression and heat mapping were also used to more clearly delineate agreement between the two measures.

Results

The resulting patient cohort comprised 35 female patients (70 hips) treated for DDH with a Pavlik harness between 1

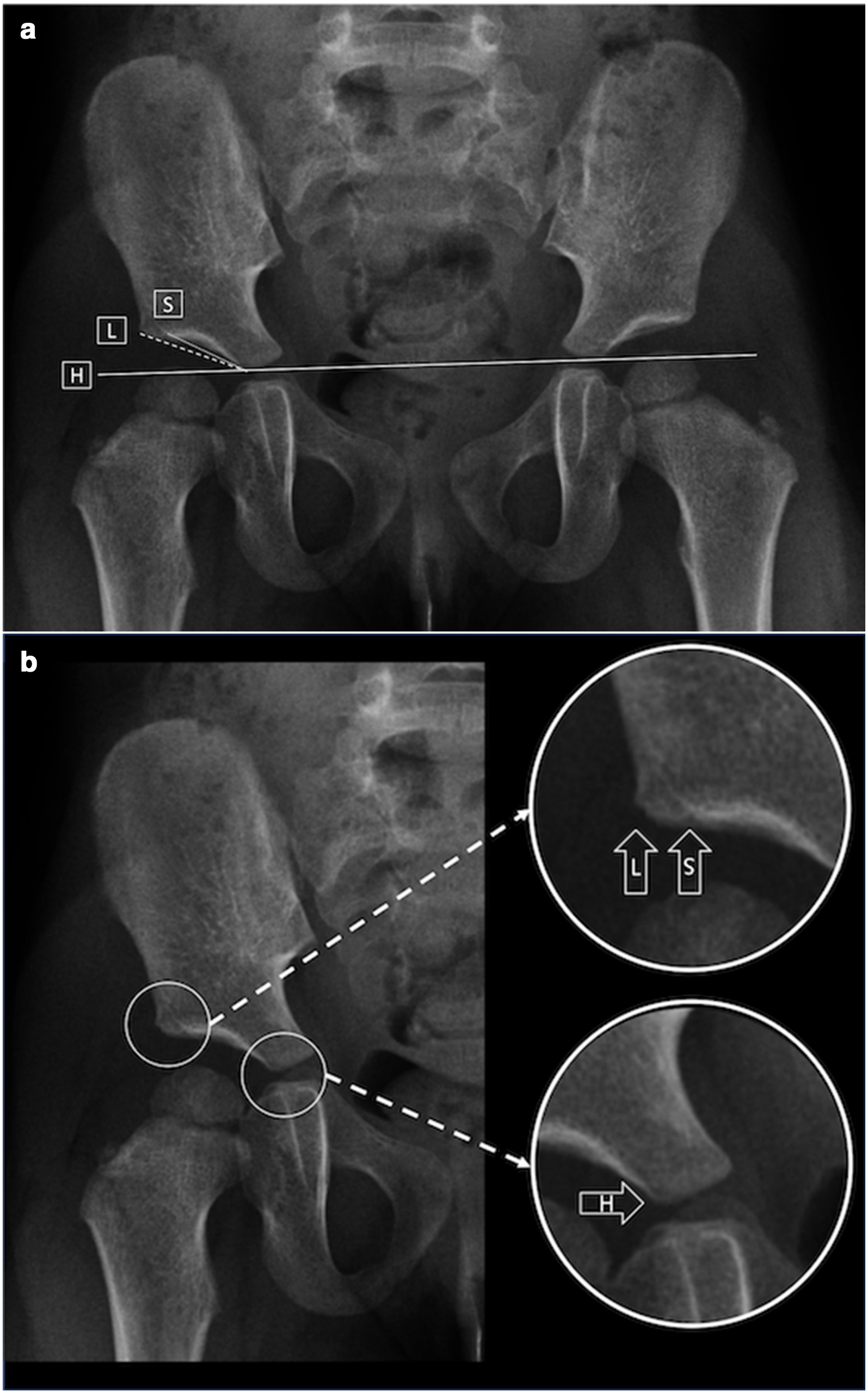

Fig. 1 a) Method of measuring acetabular indices and depiction of lateral edge of the sourcil (AI-S) and lateral edge of the acetabulum (AI-L) being distinct measures (H, Hilgenreiner's line; S, line from which AI-S is measured; L, line from which AI-L is measured); **b)** exploded view of important anatomical landmarks for defining lines used to measure acetabular indices (H, most inferior point of bony ilium at triradiate cartilage from which Hilgenreiner's line is drawn; S, lateral edge of the sourcil from which AI-S is measured; L, lateral edge of the acetabulum from which AI-L is measured). Reproduced from Maddock CL, Noor S, Kothari A, Bradley CS, Kelley SP. Reliability of the sourcil method of acetabular index measurement in developmental dysplasia of the hip. *J Child Orthop* 2019;13:167-171.[11]

January 2012 and 31 December 2014. When Pavlik harness treatment was instituted, of the 70 hips, 21 were regarded as normal, 37 dysplastic, seven dislocatable and five dislocated based on clinical and ultrasound assessment.

Following confirmation that the data were normally distributed with no kurtosis (minimal outliers), *t*-testing revealed a significant difference between absolute values of AI-L and AI-S, with AI-L being lesser in magnitude at two years of age (mean difference 2.42° (SD 3.54; 95% confidence interval (CI) 1.58 to 3.27); p < 0.001) and at five years of age (mean difference 4.56° (SD 3.16; 95% CI 3.81 to 5.31); p < 0.001). This finding stands to reason, as by definition the sourcil can never be more lateral than the lateral edge of the acetabulum, so when distinct to the lateral edge the sourcil will always be more medial and thus result in a higher AI value.

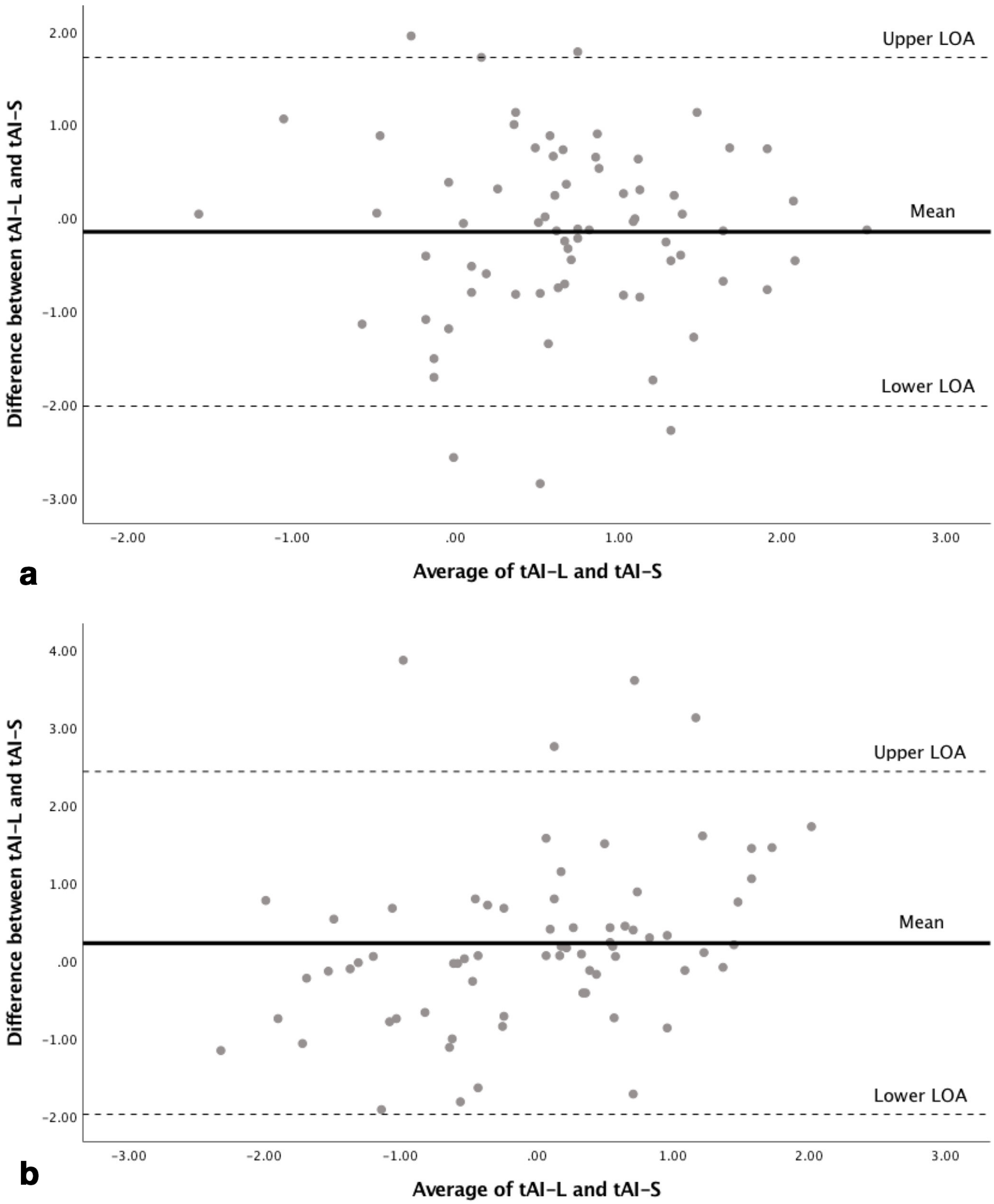

Fig. 2 a) Bland-Altman plot at two years of age; **b)** Bland-Altman plot at five years of age (LOA, limits of agreement; AI-L, lateral edge of the acetabulum; AI-S, lateral edge of the sourcil).

The Bland-Altman plots (Fig. 2) demonstrated poor agreement between tAI-L and tAI-S at both two and five years of age. This is evidenced by an even horizontal spread of data points across the full range of AI values, meaning that there was no relationship between the difference and the mean. There is also marked vertical spread in the differences between the two measures with the LOA at two years being almost ± 2 SDs in the AI values and ≥ ± 2 SDs at five years of age, both of which are greater than what was deemed acceptable (± 1 SD) at the outset of the study. There was also no evidence of proportional bias in the plot at both two and five years of age as the differences between tAI-L and tAI-S did not get larger as their average values increased.

The linear regression models confirmed there was a weak relationship between tAI-L and tAI-S at two years of age ($R^2 = 0.174$) and a slightly stronger, albeit still weak, relationship at five years of age ($R^2 = 0.299$) (Fig. 3).

Heat mapping in Fig. 4 and Fig. 5 demonstrate the classification of each hip as normal (< 1 SD), mildly dysplastic (1 to 2 SD) or severely dysplastic (> 2 SD) by tAI-L and tAI-S at two and five years. At two years of age, the tAI-L and tAI-S measurements were in agreement as to the presence and severity of dysplasia in only 44 of 70 (63%) hips, leaving 26 of 70 (37%) hips being classified as conflicting combinations of normal, mildly dysplastic and severely dysplastic, with the most extreme situation being four (6%) hips classified as severely dysplastic on one measure

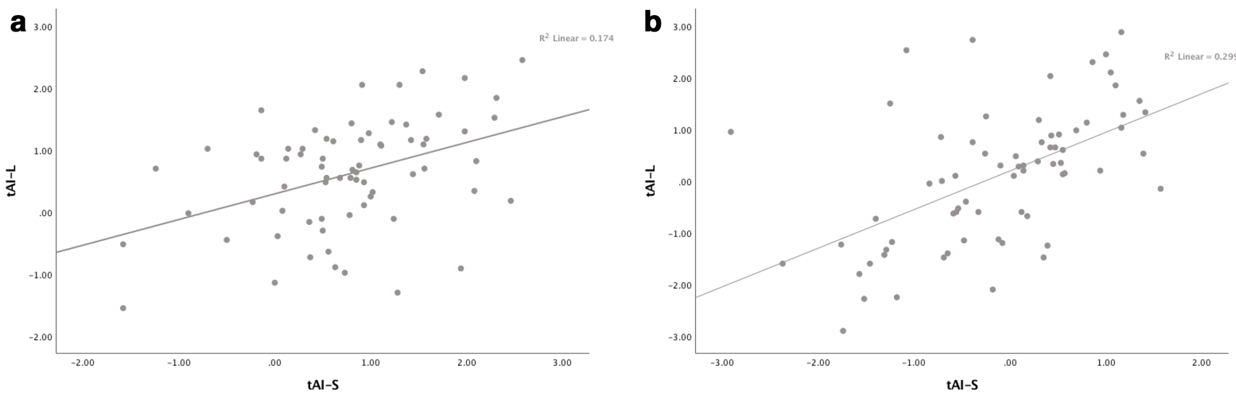

Fig. 3 a) Linear regression plot for two-year-old data; **b)** linear regression plot for five-year-old data (AI-L, lateral edge of the acetabulum; AI-S, lateral edge of the sourcil).

		tAI-S (number of hips)		
		< 1 SD (normal)	**1 SD to 2 SD (mild)**	**> 2 SD (severe)**
tAI-L (number of hips)	< 1 SD (normal)	34	7	3
	1 SD to 2 SD (mild)	10	9	2
	> 2 SD (severe)	1	3	1

Fig. 4 Heat map showing relative clinical agreement between transformed lateral edge of the acetabulum (tAI-L) and transformed lateral edge of the sourcil (tAI-S) at age two years

		tAI-S (number of hips)		
		< 1 SD (normal)	1 SD to 2 SD (mild)	> 2 SD (severe)
tAI-L (number of hips)	< 1 SD (normal)	52	2	0
	1 SD to 2 SD (mild)	4	5	0
	> 2 SD (severe)	4	3	0

Fig. 5 Heat map showing relative clinical agreement between transformed lateral edge of the acetabulum (tAI-L) and transformed lateral edge of the sourcil (tAI-S) at age five years

of AI but classified as normal on the other. At five years of age, the tAI-L and tAI-S measurements showed improved agreement with 57 of 70 (81%) hips being classified the same by each measure, but four (6%) hips classified as severely dysplastic on one measure of AI while classified as normal on the other.

Discussion

A detailed understanding of developmental hip morphology in patients treated for DDH is critical for informing clinical progress, guiding surgical decision-making and for future research. The evaluation of hip dysplasia on plain radiographs in young children based on AI measurement and its deviation from published norms often inform surgeons in their surgical decision-making. However, there are two reported methods for measuring AI and to date, it had been unclear if they could be used interchangeably.[6,7]

In this study we investigated the relationship between AI-L and AI-S in a population of children with DDH. We used two- and five-year-old follow-up radiographs as these are considered clinically relevant time points for the assessment of residual dysplasia following nonoperative treatment with a view to determining the need for surgical intervention.[17]

The working hypotheses were that if the distinction between AI-L and AI-S was purely academic then there would be no differences between the two measures when placed in the context of published normative data, that they would strongly agree on the diagnosis of dysplasia, and as such one could use either measure of AI when considering surgical management of residual acetabular dysplasia.

Our results refuted the hypotheses, suggesting that surgeons must heed caution when using these measures as part of surgical decision-making. At both two and five years of age, poor agreement was shown between tAI-L and tAI-S as demonstrated on the Bland-Altman plot. Having LOA between the two measures of ± 2 SD is far beyond even the most generous maximum acceptable limit and is certainly well above the *a priori* acceptable limit of this study of ± 1 SD.

In addition, it was determined that a hip that reads as normal on one radiographic measure of AI could be deemed dysplastic using the alternative AI measure. This finding was evidenced when hips were classified by heat mapping as normal, mildly dysplastic and severely dysplastic based on SD from the mean using AI-L and AI-S reference values. At two years of age there was agreement in only 44/70 hips and at five years of age there was agreement in only 57/70 hips. This is highly problematic as interchangeably using AI-L and AI-S to assess residual dysplasia may inadvertently lead to unnecessary surgical intervention in young children. Moreover, these findings suggest that if a hip is only moderately dysplastic at two years of age, it would be prudent to wait until five years of age before considering surgical intervention.

Most of the current literature relates to reporting AI-L. Li et al[18] demonstrated that the AI-L was the best predictor of final radiographic outcome in DDH treated with closed reduction. These findings are echoed in the work by numerous others, with a relative consensus that an AI-L of 25° or more at the age of four years would be an indication to consider surgical intervention.[13,17,19-20] There remains to be any long-term data about the AI-S in DDH and its relationship to hip osteoarthritis. Our study is the first to report AI-S as an independent measure of dysplasia in a DDH population which is as of yet incompletely understood. The prognostic implications of an abnormal AI-S will, therefore, be an important focus of future research, particularly in light of the recent findings that it is more anatomically relevant to hip dysplasia than AI-L as it represents the mid-superior aspect of the load bearing portion of the acetabulum. Until that time, the current recommendation is that surgeons should consider both AI-L and AI-S measurements independently when monitoring acetabular remodelling and during surgical decision-making.

The primary limitations of this study are the intended omission of male patients and the inclusion of normal hips at presentation. Future replication of this study in male patients is certainly feasible, but we have no reason to suspect that the outcomes would differ. Normal hips were included to allow for a full spectrum of hip anatomy that must be considered by surgeons during clinical practice. Replication of this study using more severe pathology is unlikely necessary as these hips tend to be most obvious when moving ahead with surgical intervention and the more challenging decisions are based on the types of hips included in the current study.

The findings of this study come to one conclusion: AI-L and AI-S have poor agreement concerning the presence or degree of acetabular dysplasia in DDH, and thus cannot be used interchangeably. Clinicians are cautioned to prudently evaluate both measures of AI independently as they monitor a hip over time and in the context of surgical decision-making.

Received 25 September 2020; accepted after revision 7 November 2020.

COMPLIANCE WITH ETHICAL STANDARDS

FUNDING STATEMENT

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

ETHICAL STATEMENT

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

Informed consent: It was deemed that informed consent was not required due to the nature of the study.

ICMJE CONFLICT OF INTEREST STATEMENT

None declared.

AUTHOR CONTRIBUTIONS

AK: Contributed significantly to the study design, data acquisition, data analysis and interpretation, drafting and editing of the manuscript.

SN: Contributed significantly to the study design, data acquisition, drafting and editing of the manuscript.

CLM: Contributed significantly to the study design, data acquisition, data analysis and interpretation, drafting and editing of the manuscript.

JHHV: Contributed significantly to the study design, data acquisition, drafting and editing of the manuscript.

CSB: Contributed significantly to the study design, data acquisition, data analysis and interpretation, drafting and editing of the manuscript.

SK: Contributed significantly to the study design, data acquisition, data analysis and interpretation, drafting and editing of the manuscript.

REFERENCES

1. **Terjesen T.** Residual hip dysplasia as a risk factor for osteoarthritis in 45 years follow-up of late-detected hip dislocation. *J Child Orthop* 2011;5:425-431.

2. **Reijman M, Hazes JM, Pols HA, Koes BW, Bierma-Zeinstra SM.** Acetabular dysplasia predicts incident osteoarthritis of the hip: the Rotterdam study. *Arthritis Rheum* 2005;52:787-793.

3. **Lane NE, Lin P, Christiansen L, et al.** Association of mild acetabular dysplasia with an increased risk of incident hip osteoarthritis in elderly white women: the study of osteoporotic fractures. *Arthritis Rheum* 2000;43:400-404.

4. **Ponseti IV.** Growth and development of the acetabulum in the normal child. Anatomical, histological, and roentgenographic studies. *J Bone Joint Surg [Am]* 1978;60-A:575-585.

5. **Salter RB, Dubos JP.** The first fifteen year's personal experience with innominate osteotomy in the treatment of congenital dislocation and subluxation of the hip. *Clin Orthop Relat Res* 1974;98:72-103.

6. **Kim HT, Kim JI, Yoo CI.** Diagnosing childhood acetabular dysplasia using the lateral margin of the sourcil. *J Pediatr Orthop* 2000;20:709-717.

7. **Novais EN, Pan Z, Autruong PT, Meyers ML, Chang FM.** Normal percentile reference curves and correlation of acetabular index and acetabular depth ratio in children. *J Pediatr Orthop* 2018;38:163-169.

8. **Hilgenreiner H.** Classic. Translation: Hilgenreiner on congenital hip dislocation. *J Pediatr Orthop* 1986;6:202-214.

9. **Tönnis D, Brunken D.** Differentiation of normal and pathological acetabular roof angle in the diagnosis of hip dysplasia. Evaluation of 2294 acetabular roof angles of hip joints in children. *Arch Orthop Unfallchir* 1968;64:197-228.

10. **Ogata S, Moriya H, Tsuchiya K, et al.** Acetabular cover in congenital dislocation of the hip. *J Bone Joint Surg [Br]* 1990;72-B:190-196.

11. **Maddock CL, Noor S, Kothari A, Bradley CS, Kelley SP.** Reliability of the sourcil method of acetabular index measurement in developmental dysplasia of the hip. *J Child Orthop* 2019;13:167-171.

12. **Bache CE, Clegg J, Herron M.** Risk factors for developmental dysplasia of the hip: ultrasonographic findings in the neonatal period. *J Pediatr Orthop B* 2002;11:212-218.

13. **Tönnis D.** Normal values of the hip joint for the evaluation of X-rays in children and adults. *Clin Orthop Relat Res* 1976;119:39-47.

14. **Cohen J.** *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillside, NJ: Lawrence Erlbaum Associates, 1988.

15. **Bland JM, Altman DG.** Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-310.

16. **Bland JM, Altman DG.** Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135-160.

17. **Gotoh E, Tsuji M, Matsuno T, Ando M.** Acetabular development after reduction in developmental dislocation of the hip. *Clin Orthop Relat Res* 2000;378:174-182.

18. **Li Y, Guo Y, Li M, et al.** Acetabular index is the best predictor of late residual acetabular dysplasia after closed reduction in developmental dysplasia of the hip. *Int Orthop* 2018;42:631-640.

19. **Li Y, Xu H, Li J, et al.** Early predictors of acetabular growth after closed reduction in late detected developmental dysplasia of the hip. *J Pediatr Orthop B* 2015;24:35-39.

20. **Albinana J, Dolan LA, Spratt KF, et al.** Acetabular dysplasia after treatment for developmental dysplasia of the hip. Implications for secondary procedures. *J Bone Joint Surg [Br]* 2004;86-B:876-886.