SCIENTIFIC ARTICLE

CrossMark

Point vs. traditional method evaluation of hallux valgus: interreader reliability and intermethod performance using X-ray and MRI

Nathan Heineman [1] · Avneesh Chhabra [1,2,3] · Lihua Zhang [1] · Riham Dessouky [1,4] · Dane Wukich [2]

Abstract

Background The two most widely used measurements for diagnosing and assessing the severity of hallux valgus are the hallux valgus angle (HVA) and the intermetatarsal angle (IMA). Traditionally, these have been measured by using the midaxial lines approximating the axis of each bone. A new simpler point method has been recently suggested for measuring these angles by connecting points along the medial corners of each bone. Interreader reliability of these measurements on X-ray and MRI as well as intermethod and intermodality differences have not been assessed.

Methods A series of 56 consecutive patients between 18 and 100 years old with no history of foot trauma or orthopedic hardware in their feet were included. All had AP and lateral X-rays and MRI performed on the same foot between April 27, 2015 and March 9, 2016. Two readers measured HVA and IMA using both the traditional midaxial and new point methods. ICC correlations were obtained.

Results The interreader reliability for HVA was similar on point method (0.92) and traditional method (0.94). For the IMA, the ICC was 0.77 on point method versus 0.76 on traditional method. The intermodality agreement (between X-ray and MRI) was higher for HVA (ICC = 0.85, 0.88) as compared to IMA (0.58, 0.74), respectively on both methods. The mean difference between the methods was larger on traditional method = 5.5 for HVA and 2.5° for IMA.

Conclusions HVA is more reliable than IMA on both methods and modalities and a significant difference exists between the magnitudes of values obtained using the two methods.

Level of Clinical Evidence: 3

Introduction

Hallux valgus is a condition affecting approximately 25% of all adults in the United States and approximately 35% of adults over the age of 65 [1]. The prevalence increases with age and is also much higher in females (30%) as compared to

males (13%). It occurs when the hallux deviates laterally leading to a prominent metatarsal head [2]. The exact etiology is unknown, but a variety of factors has been associated with hallux valgus including abnormal foot anatomy, inflammatory disease, neurogenic factors, and genetic predisposition [3–5]. The primary presenting complaint is pain, particularly with ambulation and shoewear (Fig. 1).

There is a vast array of measurements available for assessing the severity of hallux valgus, two of the most common measurements used for diagnosing, grading severity, and assessing the response to surgery are the hallux valgus angle (HVA) and the intermetatarsal angle (IMA). Both angles have been shown to correlate well with the degree of deformity and are highly reproducible [6, 7]. Traditionally, when measuring the angles, a line is drawn bisecting each bone to encompass the longitudinal axis of the bone. Recently, Seo et al. reported a new and simpler point method, which demonstrated greater interreader reliability. It was reported to be better at capturing

✉ Avneesh Chhabra
avneesh.chhabra@utsouthwestern.edu

1   Radiology, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390-9178, USA

2   Orthopaedic Surgery, UT Southwestern Medical Center, Dallas, TX, USA

3   Adjunct Faculty, Johns Hopkins University, Baltimore, MD, USA

4   Radiology, Faculty of Medicine, Zagazig University, Zagazig, Egypt

Fig. 1 Measurement of HVA and IMA on X-ray and MRI using the traditional method. *Line a* passes through the midpoint of the first metatarsal head and base, *line b* passes through the midpoint of head and base of the first proximal phalanx, and *line c* passes through the midpoint of the second metatarsal head and base. HVA measures the acute angle between lines a and b while IMA measures the acute angle between lines a and c

the degree to which the bunion protrudes in cases where there is subluxation of the first metatarsophalangeal joint (MTPJ). Their study consisted of 60 feet from 57 patients with 20 classified as having mild hallux valgus, 20 as moderate, and 20 as severe. All patients underwent either a proximal or distal chevron osteotomy and both preoperative and postoperative measurements were taken using both the traditional and new methods. In the new method, lines were drawn connecting points at the medial corners of the bones to obtain the respective angles (Fig. 2). This approach eliminates some of the ambiguity associated with determining the midline as it uses clearly defined landmarks. The two methods were then compared according to severity, type of osteotomy, and congruency of the MTPJ [8].

MRI of forefoot-midfoot is being increasingly used for various pathologies due to higher soft tissue contrast and multiplanar capabilities [9]. MRI in most practices is however obtained in a non-weight-bearing position while X-rays are obtained in weight-bearing position for measurements of different angles to determine the hallux valgus presence and severity. The presence or absence of weight-bearing during imaging likely alters the degree of axial rotation of metatarsals and can affect these measurements. Although, many a times, X-rays are not available during the MRI read out. No information exists as to how these measures perform on X-ray versus MRI and whether there are intermodality or intermethod differences.

The goal of our study was to determine the interreader reliability of the point method versus the traditional method for both HVA and IMA on X-rays and MRI and assess intermethod differences. We hypothesized that both methods are reliable and intermethod measurement differences exist.

Materials and methods

Institutional review board approval was obtained for this retrospective study, and informed consent was waived. Primordial® software was used to select patients by retrospectively querying the database for 90 consecutive patients who had X-rays and MRIs performed on the same foot between April 27, 2015 and March 9, 2016. This provided a series of cases with a broad range of different degrees of HVA and IMA in the same feet on both modalities. It was not documented in the patients charts if they were clinically symptomatic of hallux valgus problems. To be included in the study, age criteria were between 18 and 100 years old and presence of weight-bearing AP and lateral X-rays of the same foot on which MRI was performed was required. The weight-bearing radiographs are obtained in a standardized manner in standing position in our institution. The patients are instructed to bear about 50% of their body weight on each foot while being imaged in anteroposterior and lateral projections. The MRIs are also obtained in a standardized manner with field of view from the transverse tarsal joint to the tip of the toes using a combination of fat-suppressed and non-fat-suppressed intermediate-weighted images with slice thickness of 3.5-4 mm. The coronal (horizontal long axis) plane is planned on the sagittal and
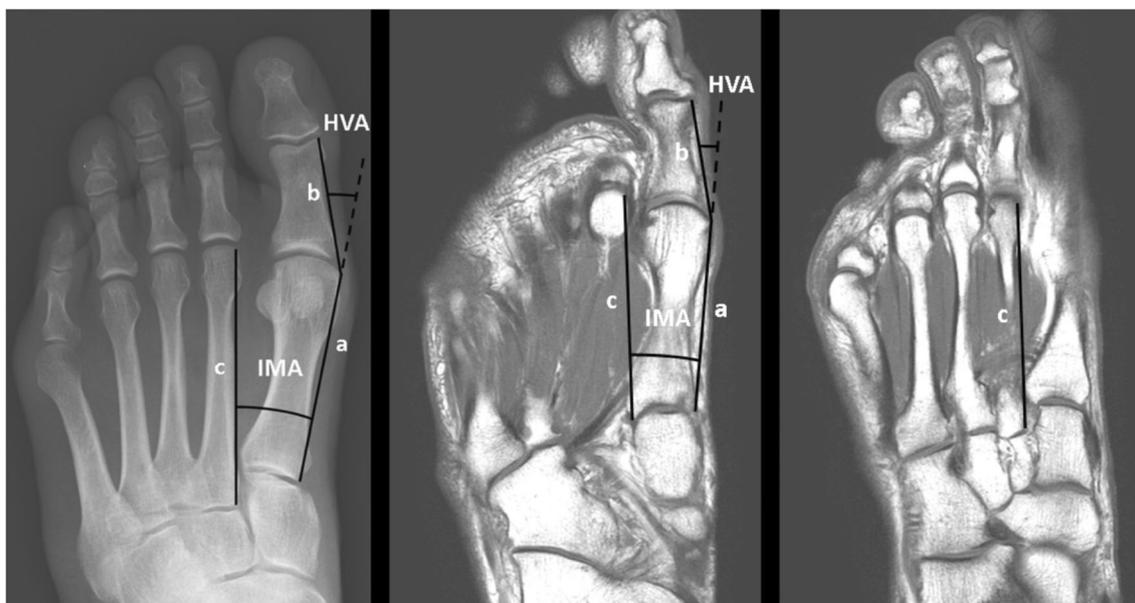
Fig. 2 Measurement of HVA and IMA on X-ray and MRI using the new point method. *Line a* connects the most medial prominent point of first metatarsal head and most medial prominent point of the first metatarsal base in the tarsometatarsal (TMT) joint. *Line b* connects the most medial prominent point of proximal phalanx in the first interphalangeal joint and most medial prominent point of the first metatarsal head. *Line c* connects the most medial prominent point of second metatarsal head and most medial prominent and sclerotic point of second metatarsal base in the TMT joint. HVA measures the acute angle between lines a and b while IMA measures the acute angle between lines a and c

axial localizers with acquisition along the metatarsal shafts. The 1–2 slices of the non-fat-suppressed sequence with visualization of majority of the 1st and 2nd metatarsal shafts were used for measurements. Nineteen patients who had a recent history of foot trauma or orthopedic hardware in their foot were excluded. Furthermore, 15 were excluded due to low-quality images or incomplete inclusion of the forefoot in the field of view resulting in a final sample of 56 patients.

Before beginning independent measurements, the reader 1 (NH- medical student) and reader 2 (LZ- radiologist with 8 years of experience) evaluated six training cases together to ensure the uniformity of their measurement technique. All cases were then read separately with each reader blinded to the other's results. We used the angle tool in Philips Intellispace® PACS to obtain each measurement. Each reader measured the HVA and IMA on both X-ray and MRI using both the traditional and point methods for a total of eight data points per patient per reader. The HVA measures the angle between the longitudinal axis of the first metatarsal and the longitudinal axis of the first proximal phalanx while the IMA measures the angle between the long axis of the first metatarsal and the long axis of the second metatarsal. For the traditional method, we used the technique outlined by Miller to determine the midline of each bone as shown in Fig. 1 [10]. For the point method, we used the technique outlined by Seo et al. as demonstrated in Fig. 2 [8].

The mean and standard deviation were calculated for all variables. A three-way mixed model was used for calculating ICC for reader agreements in HVA and IMA as well as for the agreement between traditional and point methods on X-ray and MRI, respectively; 95% bootstrap confidence intervals were calculated as well. Linear mixed model was used to detect the effect of modality and method of measurement on change in mean HVA and IMA. The average measurements of the two readers were used. A p value < 0.05 was considered as statistically significant. The Cicchetti scale (less than $0.40 =$ poor, 0.40 to $0.59 =$ fair, 0.60 to $0.74 =$ good, 0.75 to $1.00 =$ excellent) was used for categorization of intraclass correlation coefficient (ICC) [11]. Bland–Altman plots were created to assess the variability of mean differences as the magnitude of the angle increased.

Results

The mean age of the study group was 54.2 ± 15.4 (mean $\pm$ SD) years. The mean BMI was 28.7 ± 7.1 kg/m^2 and 71.4% (40/ 56) were females. Left feet comprised 53.6% of the study group (30/56) while right feet accounted for 46.4% (26/56) (Table 1). The interreader reliability (IRR) for HVA using the point method for both X-ray and MRI was excellent at 0.92 (95% CI 0.90–0.94) as opposed to 0.94 (95% CI 0.91–0.96) on the traditional method. For the IMA, the IRR was also excellent at 0.77 (95% CI 0.67–0.85) using the point method and 0.76 (95% CI 0.66–0.84) using the traditional method. The intermodality agreement (between X-ray and MRI) was excellent. These measured as 0.85 (95% CI 0.83–0.88) using the point method and 0.88 (95% CI 0.86–0.91) using the

Table 1 Study cohort characteristics ($N = 56$)

Sex	N patients (%)
Male	16 (28.6%)
Female	40 (71.4%)
Foot	
Left	30 (53.6%)
Right	26 (46.4%)
BMI (kg/m^2)	28.7 ± 7.0
Normal or underweight (< 25)	19 (33.9%)
Overweight (25–30)	15 (26.8%)
Obese (> 30)	20 (35.7%)
N/A	2 (3.6%)
Age (years)	54.2 ± 15.4

Table 3 Mean and standard deviations

			HVA		IMA	
			Mean	SD	Mean	SD
Reader 1	Point	MRI	21.16	8.29	11.02	2.77
		X-ray	20.59	9.37	11.43	3.08
	Traditional	MRI	15.77	9.48	9.32	3.54
		X-ray	13.50	8.92	8.55	3.09
Reader 2	Point	MRI	19.86	7.83	11.38	2.99
		X-ray	18.41	9.00	11.88	3.51
	Traditional	MRI	14.16	8.63	8.69	3.47
		X-ray	14.13	8.71	8.95	3.12
Combined	Point	MRI	20.51	7.88	11.2	2.49
		X-ray	19.5	9.03	11.65	3.08
	Traditional	MRI	14.96	8.82	8.97	3.1
		X-ray	13.81	8.66	8.75	2.87

traditional method for HVA. For IMA, it was good at 0.58 (95% CI 0.47–0.70) using the point method and excellent, 0.74 (95% CI 0.63–0.83) using the traditional method (Tables 2 and 3).

Of note, the mean HVA and IMA were larger on point method than the traditional method across both modalities. Mean differences between X-ray and MRI were calculated using both readers' data combined. The mean difference between the point and traditional method on MRI was 5.54 ± 0.52 for HVA ($p < 0.001$) and 2.22 ± 0.32 for IMA ($p < 0.001$). On X-ray, the mean difference was 5.69 ± 0.52 for HVA ($p < 0.001$) and 2.90 ± 0.32 ($p < 0.001$) for IMA (Table 4). Bland–Altman plots were generated for the difference between the point and traditional method. For MRI, 94.6% ($N = 53$) of data points for HVA and 96.4% ($N = 54$) of data points for IMA fell within two standard deviations of the mean (Fig. 3). For X-ray, 96.4% ($N = 54$) of data points for HVA and 98.2% ($N = 55$) for IMA fell within two standard deviations of the mean (Fig. 4).

Discussion

While HVA and IMA have traditionally been measured using the midline of each respective bone on radiographs, there exist a variety of novel methods for measuring these angles

Table 2 Intraclass correlation coefficients

		Reader 1 vs. Reader 2	X-ray vs. MRI
HVA	Point	0.92 (0.9, 0.94)	0.85 (0.83, 0.88)
	Traditional	0.94 (0.91, 0.96)	0.88 (0.86, 0.91)
IMA	Point	0.77 (0.67, 0.85)	0.58 (0.47, 0.7)
	Traditional	0.76 (0.66, 0.84)	0.74 (0.63, 0.83)

intended to make measurements cheaper, faster, or more straight-forward. Everything from footprints to digital photographs to mobile apps have been reported in the literature to have excellent reproducibility with interreader reliabilities upwards of 0.9 [12–14]. Seo et al. reported interreader reliability for HVA on X-rays as 0.91 for the new point method, being slightly superior to 0.89 for the traditional method while it was 0.86 and 0.82 for IMA, respectively. The authors also found that in cases where the MTPJ was subluxated as opposed to just deviated, there was a greater mean difference between pre- and post-operative HVA. This indicates that the new method yielded higher preoperative values for HVA when the MTPJ was subluxated, which may better reflect the severity of deformity as increased subluxation with similar HVA and IMA obtained using the traditional method is associated with worse clinical outcomes [8]. Similar IRR values have been reported by other publications using the traditional method alone with IRR for HVA ranging from 0.9 to 0.99 and for IMA ranging from 0.68 to 0.94 [6, 7, 15–19]. Our IRR of 0.92 when measuring HVA using the new point method was lower than the IRR of 0.94 on the traditional method, while IRR when measuring IMA was slightly higher using the new point method as compared to the traditional method (0.77 vs. 0.76). However, none of the differences were statistically significant and all were in excellent range, meaning that while the point method is highly reproducible, it cannot be concluded to be any more so than the traditional method.

We also measured the intermodality agreement for each angle using both the new point and traditional method. The traditional method was more consistent across X-ray and MRI than the new point method for both HVA (0.88 vs. 0.85) and IMA (0.74 vs. 0.58). The good agreement on IMA on point method as compared to excellent on traditional method might

Table 4 Mean difference between point and traditional methods

		HVA			IMA		
		Difference	Standard error	Adj. *p* value	Difference	Standard error	Adj. *p* value
MRI	Point vs. traditional	5.54	0.52	<.0001	2.22	0.32	<.0001
X-ray	Point vs. traditional	5.69	0.52	<.0001	2.90	0.32	<.0001

be explained by slice variations on MRI where the point of measurement is chosen. However, none of these differences were statistically significant. Therefore, we cannot conclude that the traditional method is any more consistent across modalities than the new point method.

Hypertrophy and/or cystic changes of the median eminence of the first metatarsal head are common with more advanced stages of hallux valgus deformity. These changes may affect the measurement of the angles, especially using the point method and another related measurement, the distal metatarsal angle (DMA). We did not study DMA as it was out of scope for this project. However, considering that the interreader reliability for HVA and IMA were similar using the point and traditional methods for both X-ray and MRI, it may suggest that the median eminence changes did not substantially impact the results.

We found that the mean difference between the two methods when measuring HVA was 5.54 ± 0.52 on MRI and 5.69 ± 0.52 on X-ray with $p < 0.001$ in both instances, reflecting a significant intermethod difference. The same was true for the IMA with the mean difference between methods being 2.22 ± 0.32 on MRI and 2.9 ± 0.32 on X-ray

and $p < 0.001$ in both instances. Traditionally, an HVA of less than 15° and IMA of less than 9° is considered normal and an HVA of greater than 40° or IMA of greater than 15 or 16 considered severe hallux valgus [20, 21]. This raises the question of what should be defined as normal values for HVA and IMA when using the point method as these values in combination with clinical symptoms are important for planning which procedure to use when correcting hallux valgus. The point method would tend to overestimate the angle and could lead to overcorrection if these differences were not accounted for during operative planning. The hallux valgus surgeries vary from distal to proximal correction procedures. With IMA of greater than 15°, it is generally agreed upon that proximal (Lapidus) procedure is warranted, which also addresses the 1st tarsometatarsal joint hypermobility. A change (overmeasurement) of 5–6° of HVA or 2–3° of IMA using the point method over the traditional method may potentially lead to more number of procedures, especially proximal procedures.

Further, it is important to determine whether the point method overestimates each angle by a constant value or if the value varies with the magnitude of the angle. To assess

Fig. 3 Bland–Altman plot: HVA. The *y*-axis represents the difference between the point and traditional method for each patient while the *x*-axis represents the average of the same two data points. The *solid line* represents the mean difference between all patients with the *dashed lines* defining two standard deviations to either side. For MRI, 94.6% (N = 53) of data points were within two standard deviations and 96.4% (N = 54) of data points were for X-ray
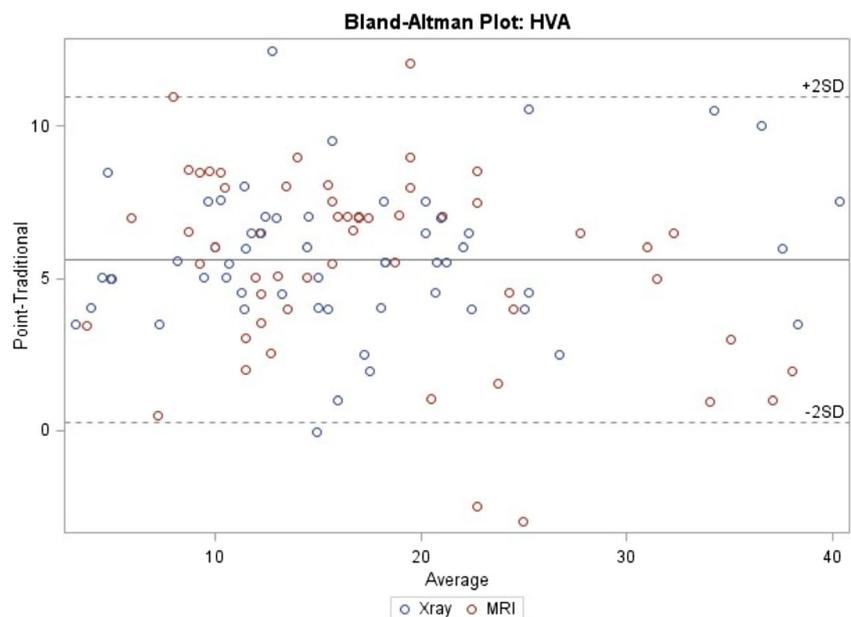
Fig. 4 Bland–Altman plot: IMA. The *y*-axis represents the difference between the point and traditional method for each patient while the *x*-axis represents the average of the same two data points. The *solid line* represents the mean difference between all patients with the *dashed lines* defining two standard deviations to either side. For MRI, 96.4% ($N = 54$) of data points fell within two standard deviations of the mean, and 98.2% ($N = 55$) did for X-ray
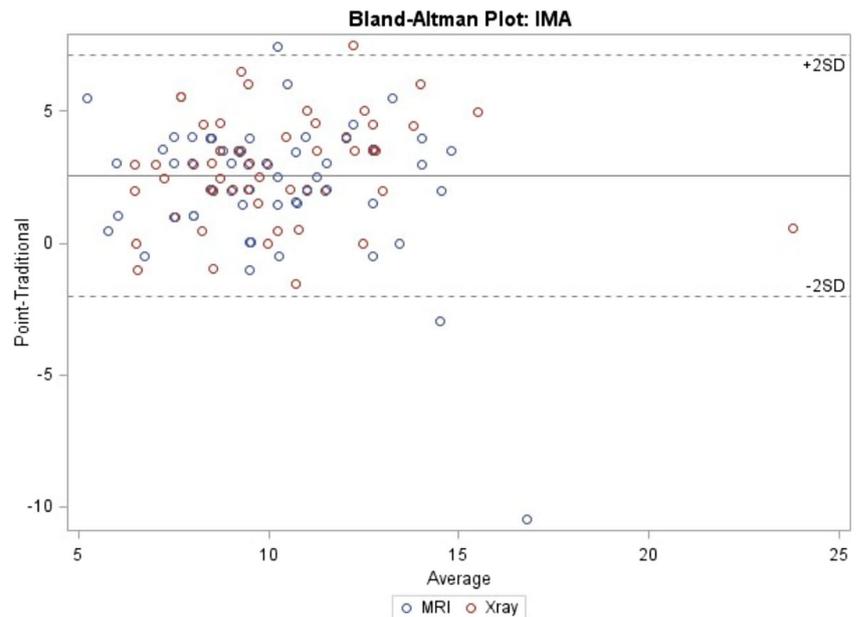
whether this difference was consistent across values or increased as the angle increased, we used Bland–Altman plots to plot the mean vs. the difference for each data point. These plots showed a relatively equal distribution of values across both sides of the mean from low values to high values for HVA and IMA on both X-ray and MRI. Thus, it was concluded that the mean differences did not differ significantly with the magnitude of the angle. Magnitude-dependent differences in the two methods cannot be entirely excluded as our data was mostly clustered in a relatively narrow range of values that may not have allowed us to pick up on the trend we would have seen if we included a wider distribution of values. Knowledge of these differences is important as a five-degree difference can easily exaggerate the severity of hallux valgus on traditional method, or vice versa if the point method was to be used.

The limitations of the study include the retrospective nature and our sample size, which was such that we did not have a significant number of patients with moderate or severe disease as defined by an HVA greater than 30°. We were, therefore, limited in our ability to assess the agreement of the two methods across the full spectrum of disease states. However, we wanted to make sure that we had a weight-bearing X-ray and MRI from the same patient. Finally, due to the retrospective nature of the study, X-rays and MRIs were not performed on the same day. Inclusion of X-ray followed by MRI within 3 months of X-ray study generation likely minimized the possibility of disease progression or regression in between the imaging sessions.

In conclusion, we found similar interreader reliability using point and traditional methods. We also found no significant intermodality differences between the two methods when comparing X-ray to MRI. However, larger values are obtained using the point method as compared to the traditional method. This represents a challenge using the new method in clinical practice as the existing cut offs for mild, moderate, and severe hallux valgus are based on the traditional method. It is, therefore, important to determine whether the mean values differ by a constant denominator or vary with the magnitude of the angle. This study could serve as the basis for future studies employing either method or modality to answer that question.

Compliance with ethical standards

References

1. Nix S, Smith M, Vicenzino B. Prevalence of hallux valgus in the general population: a systematic review and meta-analysis. J Foot Ankle Res. 2010;3:21.
2. Mann RACM. Adult Hallux Valgus. In: Coughlin MJMR, editor. Surgery of the foot and ankle. St. Louis: Mosby; 1999. p. 150–75.
3. Faber FW, Kleinrensink GJ, Verhoog MW, Vijn AH, Snijders CJ, Mulder PG, et al. Mobility of the first tarsometatarsal joint in relation to hallux valgus deformity: anatomical and biomechanical aspects. Foot Ankle Int. 1999;20(10):651–6.
4. Ferrari J, Malone-Lee J. The shape of the metatarsal head as a cause of hallux abductovalgus. Foot Ankle Int. 2002;23(3):236–42.
5. Kirkup JR, Vidigal E, Jacoby RK. The hallux and rheumatoid arthritis. Acta Orthop Scand. 1977;48(5):527–44.
6. Lee KM, Ahn S, Chung CY, Sung KH, Park MS. Reliability and relationship of radiographic measurements in hallux valgus. Clin Orthop Relat Res. 2012;470(9):2613–21.
7. Srivastava S, Chockalingam N, El Fakhri T. Radiographic angles in hallux valgus: comparison between manual and computer-assisted measurements. J Foot Ankle Surg. 2010;49(6):523–8.

8. Seo JH, Ahn JY, Boedijono D. Point-connecting measurements of the hallux valgus deformity: a new measurement and its clinical application. Yonsei Med J. 2016;57(3):741–7.

9. Ashman CJ, Klecker RJ. Yu JS. Forefoot pain involving the metatarsal region: differential diagnosis with MR imaging. Radiographics. 2001;21(6):1425–40.

10. Miller JW. Distal first metatarsal displacement osteotomy. Its place in the schema of bunion surgery. J Bone Joint Surg Am. 1974;56(5): 923–31.

11. Cicchetti D. Guidelines, Criteria, and rules of thumb for evaluating normed and standardized assessment instrument in psychology. Psychol Assess. 1994;6(4):284–90.

12. Nix S, Russell T, Vicenzino B, Smith M. Validity and reliability of hallux valgus angle measured on digital photographs. J Orthop Sports Phys Ther. 2012;42(7):642–8.

13. Pourhoseingholi E, Pourhoseingholi MA, Bagheri A, Esfandiar E, Saeb M. Footprint as an alternative to X-ray in hallux valgus angle measurement. Med J Islam Repub Iran. 2017;31:33.

14. Walter R, Kosy JD, Cove R. Inter- and intra-observer reliability of a smartphone application for measuring hallux valgus angles. Foot Ankle Surg. 2013;19(1):18–21.

15. Bryant A, Tinley P, Singer K. A comparison of radiographic measurements in normal, hallux valgus, and hallux limitus feet. J Foot Ankle Surg. 2000;39(1):39–43.

16. D'Arcangelo PR, Landorf KB, Munteanu SE, Zammit GV, Menz HB. Radiographic correlates of hallux valgus severity in older people. J Foot Ankle Res. 2010;3:20.

17. Lamm BM, Stasko PA, Gesheff MG, Bhave A. Normal foot and ankle radiographic angles, measurements, and reference points. J Foot Ankle Surg. 2016;55(5):991–8.

18. Saro C, Johnson DN, Martinez De Aragon J, Lindgren U, Fellander-Tsai L. Reliability of radiological and cosmetic measurements in hallux valgus. Acta Radiol. 2005;46(8):843–51.

19. Shima H, Okuda R, Yasuda T, Jotoku T, Kitano N, Kinoshita M. Radiographic measurements in patients with hallux valgus before and after proximal crescentic osteotomy. J Bone Joint Surg Am. 2009;91(6):1369–76.

20. Karasick D, Wapner KL. Hallux valgus deformity: preoperative radiologic assessment. AJR Am J Roentgenol. 1990;155(1):119–23.

21. Rippstein PF, Park YU, Naal FD. Combination of first metatarsophalangeal joint arthrodesis and proximal correction for severe hallux valgus deformity. Foot Ankle Int. 2012;33(5):400–5.