The Reliability of Three Techniques for Measuring Foot Position

SHARON WEINER-OGILVIE, BSc (Hons), DPodM, SRCh*
KEITH ROME, MSc, BSc (Hons), DPodM, SRCh†

The reliability of three commonly used techniques for measuring foot position—valgus index, navicular height, and arch height—was evaluated in a study involving 20 healthy subjects. The results demonstrated significant differences ($P < .05$) between two observers for all three techniques, although there were no significant differences between two visits for the same observer ($P > .05$). Secondary analysis demonstrated that navicular height yielded the highest degree of intraobserver and interobserver agreement. The results suggest that there is a wide variation in foot position in the general population, and that measurement error may result from difficulties in defining foot position, techniques used, and instrumentation. (J Am Podiatr Med Assoc 88(8): 381-386, 1998)

Clinical measurement of foot position is an essential part of routine foot assessment for evaluating foot function, prescribing and fabricating foot orthoses, and treating numerous pathologies, including hallux abducto valgus, knee pain, and shin splints.[1] In order to categorize an abnormally positioned foot, a reliable and valid measurement technique must be used. With regard to clinical assessment, reliability can be defined as the consistency of a measurement, and represents the degree of agreement of measurements taken from the same entity by different observers or repeated measurements by the same observer over time.[2] Validity can be defined as evidence that a clinical measuring technique actually measures what it is supposed to measure.[2] Unfortunately, clinical measurement of foot position is inherently difficult because of the complex interactions of the joints of the foot. Foot position involves eversion of the calcaneus, adduction and plantarflexion of the talus, abduction of the forefoot, lowering of the longitudinal arch, and internal rotation of the tibia.[3] However, each of these component positions varies across individuals, and quantification of foot position could involve measurement of any one of these components.

Previous studies have used traditional techniques focusing on the frontal-plane component of foot position, such as the relaxed calcaneal stance position, which has been reported to be the most commonly used technique to assess position.[4, 5] Intraobserver reliability of this technique was found to be good,[6, 7] but interobserver reliability[7] was reported to be poor when inexperienced observers carried out the tests. However, when observers were experienced or when training sessions and specific guidelines were provided, interobserver reliability was good.[7-11]

Alternative techniques for measuring foot position, such as navicular height, valgus index, and arch height, have recently been described in the literature. Navicular height is a measurement of sagittal-plane position, and previous studies have reported good intraobserver reliability and moderate interobserver reliability of this technique.[12] Sell et al[11] reported excellent intraobserver and interobserver reliability when measurements of navicular height were recorded. McCrory et al[13] suggested that navicular height can be reliably determined with radiographic measure-

*Postgraduate research student, Physiotherapy Department, School of Health, University of Teesside, Middlesbrough TS1 3BA, England, United Kingdom.

†Senior Lecturer, Physiotherapy Department, School of Health, University of Teesside, Middlesbrough, England, United Kingdom.

ments; this method reduces the errors involved in palpation, skin movement, and soft-tissue distribution. Although radiographic measurements may be more reproducible, studies have shown that they also exhibit significant intraobserver and interobserver variability.[14, 15]

Arch height is a measure of the distance from the highest point of the soft-tissue margin of the medial longitudinal arch to the ground. Hawes et al[6] reported excellent intraobserver and interobserver reliability with asymptomatic patients, and Saltzman et al,[12] examining orthopedic patients, also reported good intraobserver reliability and acceptable interobserver reliability. Cowan et al,[8] however, found unacceptable interobserver variability and stressed the need for more objective standards in evaluating morphology.

The valgus index is a general indicator of transverse- and frontal-plane position of the foot relative to the malleoli. Thompson[16] reported an error value of ±2° when three successive measurements were obtained from a single subject, and suggested that valgus index is a less subjective and more sensitive measure of foot position than resting calcaneal stance position. Song et al[17] reported excellent intraobserver and interobserver reliability for this technique.

Although previous studies have examined the reliability of individual techniques for assessing position, limited research has been conducted comparing intraobserver and interobserver variability using different clinical techniques. Sell et al[11] reported increased reliability of navicular-height measurement compared with relaxed calcaneal stance position. Intraobserver reliability was good for both navicular height and relaxed calcaneal stance position, and interobserver reliability measurements were excellent and moderate, respectively. Reliability studies often report reliability using simple or intraclass correlation.[17] The magnitude of the error component of measurements in clinical terms is rarely reported. The purpose of the present study was to assess the intraobserver and interobserver reliability for three different types of measurement and to determine the amount of measurement error.

Methods

The investigation involved the determination of the valgus index from static footprints; measurement of the arch height from a weightbearing static position; and the computation of navicular height as a component of foot position. Twenty healthy subjects (13 females and 7 males) participated in this study (Table 1).

Table 1. Characteristics of the Study Cohort (N = 20)

Characteristic	Mean ± SD	Range
Age (years)	25.4 ± 6.00	19.4 – 43.7
Height (m)	1.72 ± 0.09	1.58 – 1.90
Weight (kg)	67.84 ± 11.06	46.3 – 82.0

Subjects were informed about the procedure to be performed, and informed consent was obtained from each subject prior to the start of the study. The subjects were volunteers who satisfied the following criteria: 1) pain-free motion available at both the ankle and subtalar joints; 2) no history of arthritic or neuromuscular disorders; 3) no history of major trauma to the ankle or subtalar joints (including fractures, dislocations, and sprains); and 4) no apparent pain or discomfort during walking.

Measurements were carried out by two experienced clinicians (mean of 9 years in practice). Prior to data collection, a pilot study was conducted so that examiners could become proficient in specific measurement techniques. All measurements were taken of the right foot and were recorded by an independent observer. Lighting, temperature, and sound were monitored and kept consistent. Subjects were randomly assigned and measured on two separate visits, with a 24-hour interval between them. Each subject visit consisted of three repetitions of each technique. All measurements were taken in the morning, with initial markings made by the examiners removed by the independent observer after each procedure.

For the determination of arch height, the subject stood in relaxed calcaneal stance position on a 1-m raised platform, as suggested by Cowan et al.[8] The platform was designed to facilitate full visualization of the arch region. The examiner marked the top of the arch region with a black felt-tip pen. A modified Mitutoyo®[1] digital caliper was used to measure arch height from the ground to the top of the arch region. The Mitutoyo digital caliper is reported to have a mean accuracy of ±0.9 mm when digitizing a calibration device.

To obtain a valgus index, a Harris-Beath mat was used. The 3-mm-thick rubber mat is engraved with a series of ridges at three different heights arranged in two planes. A thin layer of printer's water-soluble ink was evenly distributed on the Harris-Beath mat with a printer's roller. The Harris-Beath mat was placed next to another rubber mat of the same thickness and the subject was positioned standing.[16] The sub-

®[1] Mitutoyo (UK) Ltd, Andover, England.

ject was then instructed to step onto the mats one foot at a time and asked to look straight ahead while a measurement was taken. To record the malleolar position, a modified set square was placed against the most prominent aspect of each malleolus. The modified set square was placed next to the most prominent portion of the medial malleolus, and was positioned adjacent to the lateral malleolus. A black marking pen was placed in notches perpendicular to both set squares and used to mark the blank paper. To determine the valgus index from the footprints, the projection of the lateral malleolus was labeled A and that of the medial malleolus was labeled B. The center of the heel print was located using a pair of compasses to find the circle that best conformed with the margins of the heel print.[16] The center of the third toe was used as the forefoot reference and a single line drawn between this point and the center of the heel. The point of intersection was labeled C. The lines AB and AC were measured and the valgus index determined as follows:

$$\text{Valgus Index} = \frac{.5AB - AC}{AB} \times 100$$

To calculate the navicular height, the most prominent palpable portion of the navicular tuberosity was marked with a dot using a fine black felt-tip pen when the subject was in the prone position. The subject stood on the platform in a bilateral stance position. The modified Mitutoyo digital caliper was used to measure navicular height. The caliper was set to zero prior to taking the measurements and placed vertically on the inner side of the foot just anterior to the navicular. The distance (in millimeters) from the fixed end of the caliper to the mark on the navicular tuberosity was measured by the examiner and recorded by the observer.

Data Analysis

Primary analysis of the results used multifactorial repeated-measures analysis of variance (ANOVA), with the factors of observer, visit, and technique being analyzed ($P < .05$). Secondary analysis was carried out to assess the size of the error in clinical terms. To assess intraobserver agreement, the difference between the measurements on the two visits for each observer was calculated and recorded as a percentage of the mean measurement. For example, if observer 1 measured the navicular height of a subject as 38 mm on visit 1 and 42 mm on visit 2, the difference of 4 mm was recorded as a 10% difference. This calculation was carried out for all subjects for each observer separately. To assess interobserver agreement, differences between the two observers for each visit were calculated and recorded as a percentage of the mean measurement. For example, if observer 1 measured valgus index as 12 for a given subject and observer 2 measured it as 8, the difference of 4 was recorded as a 40% difference. This calculation was carried out for each subject at each visit.

Results

The results of the multifactorial ANOVA are presented in Table 2. The results showed significant differences between subjects, within subjects and observers, and within subjects and techniques ($P < .05$). Table 3 provides descriptive statistics for each technique used. There was no significant difference within subjects and visits ($P > .05$). Figure 1 shows the mean percent difference between visits for each observer. Figure 2 demonstrates the mean percent difference between the two observers for each visit.

Discussion

The results demonstrated a high degree of variability between subjects, which suggests a wide diversity of

Table 2. Major Results of Multifactorial ANOVA: Two Observers, Two Visits, Three Techniques

Variable	*P* Value	Significance
Between subjects	.000	Significant
Within subjects/observers	.000	Significant
Within subjects/visits	.248	Not significant
Within subjects/techniques	.000	Significant
Within subjects/observers by visit	.134	Not significant

Table 3. Descriptive Statistics for the Techniques Used in the Study

Technique	Observer	Mean ± SD	Range
Arch height	1	18.57 ± 5.05	10.18 – 32.90
	2	10.00 ± 4.03	1.71 – 19.68
Navicular height	1	32.89 ± 8.09	15.38 – 49.04
	2	34.23 ± 7.66	22.46 – 50.86
Valgus index	1	5.03 ± 3.69	–3.21 – 13.00
	2	3.28 ± 3.94	–2.42 – 12.85

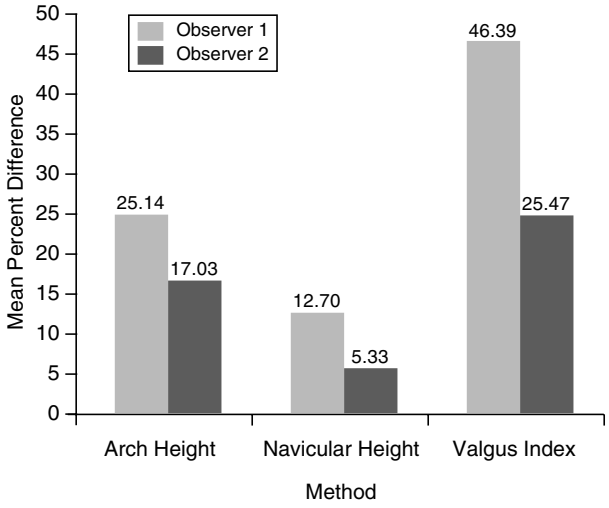

Figure 1. Intraobserver agreement: mean percent difference between visits for each observer.

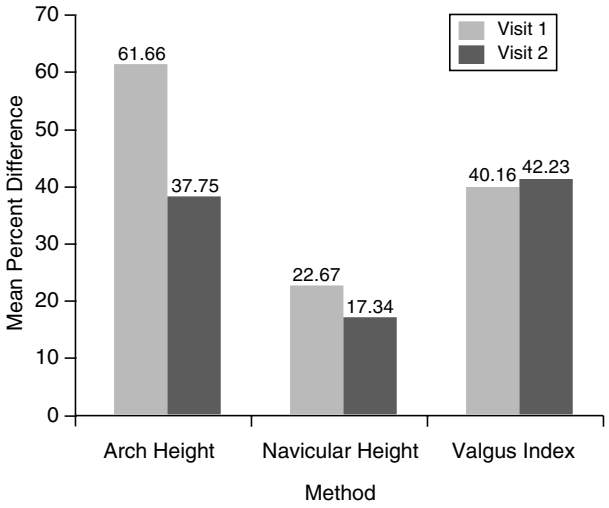

Figure 2. Interobserver agreement: mean percent difference between observers for each visit.

foot position within the sample cohort (Table 2). Previous studies have also reported variability of foot position, which suggests that there is considerable variation in foot position in the general population. There was no significant difference between the two visits (Table 2), which suggests that foot position did not significantly change from one visit to another over 24 hours, and therefore that it is unnecessary to measure foot position on different visits for healthy individuals. The two observers differed significantly in their measurements of foot position (Table 2). This is not surprising, as interobserver reliability of lower-limb measurements is often reported as moderate or poor.[5, 9] Measurement error has been attributed to difficulties in defining foot position, techniques used, instrumentation, and population variations.[18] Differences were observed between the techniques, which indicates that each technique measures a different component of foot position and therefore has a different range of values. For example, navicular height relates to sagittal-plane position and has a relatively large range of values, whereas valgus index relates to transverse- and frontal-plane position and has a relatively restricted range of values.

Secondary analysis was performed in order to further analyze agreement between and within observers. In general, intraobserver agreement was found to be higher than interobserver agreement (Figs. 1 and 2). These results may contradict those of other studies that reported good or acceptable interobserver reliability.[8, 12] However, those studies used intraclass correlations as their primary analysis and did not analyze measurement differences between and within observers. The problem arising from using correlation coefficients as a measure of reliability has been already discussed by others.[19, 20] Bland and Altman[20] reported that data that seem to be in poor agreement can result in a high degree of correlation, and thus that correlation should not be used as a measure of agreement.

Navicular-height measurements demonstrated the smallest intraobserver and interobserver differences. Mean differences between the two observers were 17% to 22%. When the same observer took the measurements, mean differences between the two visits were 5% to 12%. These findings are in agreement with those of other studies.[21-24] Nevertheless, it is difficult to compare the present findings to those of the other studies, as they reported intraclass correlation values.

A mean difference of 40% was observed between the two observers when calculating the valgus index, and a similar difference was observed when the same observer took the measurement on two different visits. Mean differences between the two observers were 37% to 61% when measuring arch height. When the same observer took the measurements on two visits, mean differences were 17% to 25%. These may be perceived as unduly high in a clinical setting. Further work is required on the acceptable range of difference within and between clinicians in clinical measurement of foot position.

Intraobserver and interobserver mean differences

were smallest for navicular height, but mean standard deviation values were generally large. This suggests that differences between the two observers or between two different visits of the same observer are larger than the mean difference. Mean standard deviations of the valgus index and arch height were larger than that of navicular height, indicating an even larger percent difference between observers or between two different visits of the same observer. However, the clinical significance of these results requires further clarification.

Navicular height and arch height have similar types and ranges of values. Navicular height showed smaller intraobserver and interobserver differences than arch height. This may be a result of the difficulty of finding the highest point along the arch in the absence of any anatomic landmarks. These findings are not in agreement with those of previous studies that used intraclass correlation.[7, 22] However, in subjects in whom the navicular could not be easily palpated, agreement was notably reduced. This further suggests that clinical measurements may demonstrate variability within the population of choice. The reader also should note the limited number of subjects used in this study. The mean values reported for each of the measurement variables should not be interpreted as normal values. Future investigators should conduct studies with appropriate sample sizes to allow for generalization of the results. Although the present study used mean differences expressed as percentages as a technique for observing measurement variability, the technique may be inappropriate in the case of small values such as those for valgus index and navicular height. The use of other statistical techniques such as Bland and Altman's level of agreement has been suggested as a means of overcoming these problems.[20]

Conclusion

The error associated with measurements was evaluated for three commonly used clinical techniques for assessing foot position. The results demonstrated significant differences ($P < .05$) between two observers for all techniques, although there were no significant differences between the two visits for the same observer ($P > .05$). Secondary analysis demonstrated that measurement of navicular height yielded the highest intraobserver and interobserver agreement. The clinical significance of the percentage differences should be further evaluated. Further research will help establish the variation and acceptable range of normative values for clinical measurements of foot position.

References

1. Menz H: Clinical hindfoot measurement: a critical review of the literature. Foot **5:** 57, 1995.
2. Gajdosik RL, Bohannon RW: Clinical measurement of range of motion: review of goniometry emphasizing reliability and validity. Phys Ther **67:** 1867, 1987.
3. Rodgers MM: Dynamic foot biomechanics. J Orthop Sports Phys Ther **21:** 306, 1995.
4. Diamond JE, Mueller MJ, Delitto A, et al: Reliability of a diabetic foot evaluation. Phys Ther **69:** 797, 1989.
5. Elveru RA, Rothstein JM, Lamb RL: Goniometric reliability in a clinical setting: subtalar and ankle joint measurements. Phys Ther **68:** 672, 1988.
6. Hawes MR, Nachbauer W, Sovak D, et al: Footprint parameters as a measure of arch height. Foot Ankle **13:** 22, 1992.
7. Freeman AC: A study of the interobserver and intraobserver reliability in the measurement of resting calcaneal stance position and neutral calcaneal stance position. Aust Podiatr **6:** 10, 1990.
8. Cowan DN, Robinson JR, Jones BH, et al: Consistency of visual assessments of arch height among clinicians. Foot Ankle **15:** 213, 1994.
9. Donatelli R, Hurlbert C, Conaway D: Biomechanical foot orthotics: a retrospective study. J Orthop Sports Phys Ther **10:** 205, 1988.
10. Jonson SR, Gross MT: Intraexaminer reliability, interexaminer reliability, and mean values for nine lower extremity skeletal measures in healthy naval midshipmen. J Orthop Sports Phys Ther **25:** 253, 1997.
11. Sell KE, Verity TM, Worrell TW, et al: Two measurement techniques for assessing subtalar joint position: a reliability study. J Orthop Sports Phys Ther **19:** 162, 1994.
12. Saltzman CI, Nawoczenski DA, Talbort KD: Measurement of the medial longitudinal arch. Arch Phys Med Rehabil **76:** 45, 1995.
13. McCrory JL, Young MJ, Boulton AJM, et al: Arch index as a predictor of arch height. Foot **7:** 79, 1997.
14. Plant MJ, Saklatvala J, Jones PW, et al: Prediction of radiographic damage in hands and feet in rheumatoid arthritis by clinical evaluation. Clin Rheumatol **13:** 487, 1994.
15. Rhodes DW, Mansfield ER, Bishop PA, et al: Comparison of leg inequality measurement methods as estimators of the femur head height difference on standing x-ray. J Manipulative Physiol Ther **18:** 448, 1995.
16. Thompson CE: An investigation into the reliability of the valgus index and its reliability as a clinical measurement. Foot **4:** 191, 1994.
17. Song J, Hillstrom HJ, Secord D, et al: Foot type biomechanics: comparison of planus and rectus foot types. JAPMA **86:** 16, 1996.
18. Ball P, Johnson GR: Technique for the measurement of hindfoot inversion and eversion and its use to study a normal population. Clin Biomech **11:** 165, 1996.
19. Shrout PE, Fleiss JL: Intraclass correlation: uses in assessing rater reliability. Psychol Bull **86:** 420, 1979.
20. Bland JM, Altman DG: Statistical techniques for assessing agreement between two methods of clinical measurement. Lancet **1:** 307, 1986.
21. McPoil TG, Cornwall MW: Relationship between three static angles of the rearfoot and the pattern of rearfoot motion during walking. J Orthop Sports Phys Ther **23:** 370, 1996.
22. Mueller MJ, Host JV, Norton BJ: Navicular drop as a

composite measure of excessive position. JAPMA **83:** 198, 1993.
23. WEINER-OGILVIE S, RANDALL GC, ABBOUD RJ: Reliability of open kinetic chain subtalar joint measurement. Foot **7:** 128, 1997.
24. SMITH-ORICCHIO K, HARRIS BA: Interrater reliability of subtalar neutral, calcaneal inversion and eversion. J Orthop Sports Phys Ther **12:** 10, 1990.