
ORIGINAL RESEARCH

INTRA-RATER AND INTER-RATER RELIABILITY OF THE FIVE IMAGE-BASED CRITERIA OF THE FOOT POSTURE INDEX-6

Masafumi Terada, MS, ATC¹
Ara M. Wittwer, MS, ATC¹
Phillip A. Gribble, PhD, ATC¹

ABSTRACT

Purpose/Background: The Foot Posture Index-6 (FPI-6) is considered a simple quantification tool to assess static foot alignment. Palpation of the foot is required for assessment of one of the six criteria that comprise the FPI-6; the remaining five criteria may be evaluated using still-frame photographs. Using only the image-based criteria may allow multiple clinicians to evaluate large groups of patients quickly. Reliability using only these five image-based criteria has not been established. The purposes of the current study were to establish the inter- and intra-rater reliability using five image-based criteria from the Foot Posture Index-6 (FPI-6) as well as to examine the agreement between the raters in identifying foot type using the composite five FPI scores.

Methods: Forty participants (23 females, 17 males; 23.67 ± 8.49 years; 64.59 ± 14.43 kg; 166.07 ± 11.79 cm) volunteered for this study. An investigator took three photos with a digital camera of the medial longitudinal arch, posterior ankle, and of the talonavicular joint approximately 45° from the posterior calcaneus for both right and left feet. Two investigators assessed the five image-based criteria of the FPI-6 for both feet of 40 participants on three occasions separated by a day. Inter- and intra-rater reliability were assessed with Intraclass Correlation Coefficients ($ICC_{3,2}$). The amount of agreement for classification of foot posture type between the two raters was assessed with Cohen's kappa coefficient. Significance was set a priori at $P < 0.05$.

Results: The inter-rater reliability was poor to moderate for all three sessions ($ICC_{3,2} = 0.334-0.634$). For the foot posture classification, the amount of agreement between two raters was poor for left ($\kappa = 0.12$) and right ($\kappa = 0.19$) feet. The intra-rater reliability was excellent for left ($ICC_{3,2} = 0.956$) and right feet ($ICC_{3,2} = 0.959$).

Conclusions: Excellent intra-rater and poor to moderate inter-rater reliability was found using only the five image-based criteria of the FPI-6. However, the classification of foot posture did not improve the amount of agreement between raters. Therefore, caution is needed when interpreting FPI scores from five image-based criteria.

Levels of Evidence: 3b

Keywords: Foot posture, reliability, static postural assessment

¹ University of Toledo, Toledo, OH, USA

Conflict of Interests

We certify that no party having a direct interest in the results of the research supporting this article has or will confer a benefit on us or on any organization with which we are associated. No conflicts of interest are directly relevant to the content of this study.

Ethical approval

The methods used in this study have been approved by the internal review board at the University of Toledo.

CORRESPONDING AUTHOR

Masafumi Terada, MS, ATC
2801 W Bancroft Street, Mailstop#119
University of Toledo
Toledo OH 43606
Tel: 419-530-2691
Fax: 419-530-2477
E-mail: mterada@rockets.utoledo.edu

INTRODUCTION

It is important for sports-related injury prevention to identify potential risk factors. Foot posture is one factor considered when examining risk of sports-related injury.¹⁻³ Static foot alignment is frequently assessed to identify foot posture in clinical practice for evaluating injury risk and determining a course of therapeutic intervention to decrease the risk.

The Foot Posture Index-6 (FPI-6) is considered a fast, simple, inexpensive, and multisegmental clinical quantification tool to assess static foot alignment in all three planes and to classify foot posture types.⁴ The FPI-6 consists of six individual criteria (Table 1), each having a 5-point scale that ranges from -2 to +2, with negative numbers indicating a more supinated foot posture and positive numbers indicating a more pronated foot posture.⁵ A composite FPI-6 score ranges from -12 to +12, and a foot type is classified as a highly pronated posture with a score of 10 or greater, a pronated posture with scores of +6 to +9, normal posture with scores of 0 to +5, a supinated posture with scores of -4 to -1, or a highly supinated posture with ≤ -5 .^{4,6,7} The FPI-6 has demonstrated good construct validity as a clinical instrument,⁶ and has been used in clinical and research settings to identify risk factors for sports-related injuries.^{1,8,9}

Direct patient contact is required for assessment of one of the six criteria that comprise the FPI-6, specifically palpation of the head of the talus to assess rearfoot alignment. The remaining five criteria of the FPI-6 are evaluated based on visual observations of the rearfoot, midfoot, and forefoot of a patient. When evaluating an individual patient in a clinical setting, the time needed to complete the direct visual observation criteria is typically available to the patient and clinician. However, when screening large groups of physically active individuals to identify injury risk, it is critical for selected screening measures to be much more time efficient. While direct visual observation has typically been performed for the FPI-6 assessment, if the five image-based criteria could be evaluated effectively with a digital image after the physical screening is complete, it would allow multiple clinicians to evaluate large groups of individuals more efficiently, as well as patients not to be present excessive time during the assessment.

The reliability of clinicians' ratings of postural evaluation is crucial for assessment and the interpretation of examination findings in clinical practice and research.¹⁰ If using only the five-imaged criteria of the FPI-6 is reliable, this could introduce a foot posture assessment into a pre-participation examination for injury prediction

Table 1. The five image-based criteria of the Foot Posture Index (FPI-6) Criteria^{4,5}

| FPI-Subscale | Description |
|---|---|
| Visual Observation Criteria | |
| Supralateral and infralateral malleolar curvature (assessed from the posterior of participants) | -2 = No concavity below the lateral malleolus compared to the curvature above the lateral malleolus 0 = Equal curvature below and above the lateral malleolus +2 = More concavity below the lateral malleolus compared to the curvature above the lateral malleolus |
| Frontal plane alignment and position of the calcaneus (assessed from the posterior of participants) | -2 = Greater than 5° estimated position of the calcaneus in the frontal plane 0 = Neutral (0°) calcaneal position between inversion and eversion +2 = less than 5° estimated position of the calcaneus in the frontal plane |
| Bulging in region of the talonavicular joint (assessed from the posteromedial side of participants) | -2 = Marked concavity in the region of the talonavicular joint 0 = Flat in the region of the talonavicular joint +2 = Marked convexity in the region of the talonavicular joint |
| Congruence of the medial longitudinal arch (assessed from the medial side of participants) | -2 = a very high medial longitudinal arch with a steeper posterior wall of the arch 0 = Normal arch height with natural curve +2 = a very low and flat medial longitudinal arch contacting the ground or floor |
| Abduction/Adduction of the forefoot on the rearfoot (assessed from the posterior of participants) | -2 = Medial toes visible, but not lateral toes 0 = Both medial and lateral toes equally visible +2 = Lateral toes visible, but not medial toes |

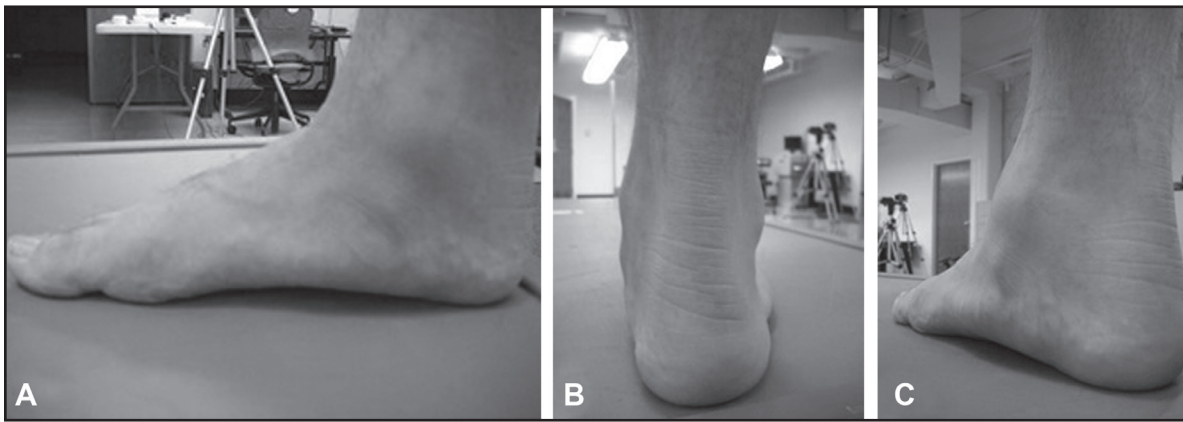


Figure 1. Three photos of the medial longitudinal arch (A), posterior ankle (B), and talonavicular joint (C).

and determination of appropriate intervention, without introducing excessive time demands on the patient, clinician, or researcher during the assessment. Although previous authors^{8,11-14} have reported that the FPI-6 has a moderate to excellent inter- and intra-rater reliability, reliability using only these five image-based criteria has not been established. Therefore, the purposes of the current study were to establish the inter- and intra-rater reliability using the five image-based criteria from the FPI-6 as well as examine the agreement between the raters in identifying foot type using the composite five FPI scores.

METHODS

Design

Using a descriptive laboratory study design, the main outcome variables of the total score using the five image-based criteria of the FPI-6 and a foot posture type classified based on the composite scores were assessed across sessions (day 1, day 3, day 5) and raters (rater 1, rater 2).

Participants

Forty volunteers (23 females, 17 males, age = 23.67 ± 8.49 years, body mass = 64.59 ± 14.43 kg, height = 166.07 ± 11.79 cm) participated in this study. Participants signed an informed consent form approved by the University of Toledo Institutional Review Board.

Procedures

Two investigators served as raters of the five image-based criteria of the FPI-6 for both feet of 40 participants on three occasions each separated by a day. Rater 1 and 2 were certified athletic trainers with

postgraduate clinical experiences. Neither rater had previous experience using the FPI-6 in their clinical practice. The raters were exposed to the use of the FPI-6 during pilot work on 15 participants that were not included in this study's sample. During the pilot work, the raters performed the FPI-6 assessment using the original protocol⁴ and were allowed to have open discussion about the index criteria.

Participants were asked to take several steps in place and stand in a relaxed stance with double-limb support, arms by the side, and looking straight ahead. Participants maintained this comfortable stance position while an investigator took three photos of the medial longitudinal arch (Figure 1-A), posterior ankle (Figure 1-B), and talonavicular joint (Figure 1-C) for both right and left feet with a digital camera (Kodak EasyShare M853; Rochester, New York). The camera was positioned approximately 30cm from the medial side of the foot for the medial longitudinal arch, approximately 25 cm from the posterior aspect of the calcaneus for the posterior ankle view, as well as 25 cm from the talonavicular joint and approximately 45° from the posterior calcaneus for the talonavicular joint view. Using the photographs, the two raters independently scored both feet of each of the 40 participants based on the five image-based criteria of the FPI-6 on the same day that the photographs were taken. Two and four days later, the raters assessed the same photographs again in a random order and scored static foot alignment on each photo with the five image-based criteria of the FPI-6 again (Table 1). Each criterion was scored on a scale of -2 to +2, with a total FPI score ranging from -10

to +10. Each participant's foot was classified based on an established version (6 criteria) of the composite FPI score as "highly pronated" with a score of ≥ 10 , "pronated" with scores of 6 to 9, "normal" with scores of 0 to 5, "supinated" with scores of -1 to -4, or "highly supinated" with scores of ≤ -5 .^{4,6}

Statistical Analyses

To assess inter-rater reliability, Intraclass Correlation Coefficients ($ICC_{3,2}$) with 95% confidence intervals (CI) were calculated for the total FPI scores for two raters for each individual assessment session. The intra-rater reliability ($ICC_{3,3}$) of the average FPI scores of two raters [(Rater 1 score + Rater 2 score)/2] for each posture assessment session was determined. The intra-rater reliability of the composite score of five image-based criteria of FPI-6 was also assessed for each rater. Intraclass Correlation Coefficients were interpreted as poor (0.0-0.50), moderate (0.51-0.75), good (0.76-0.90), or excellent (0.91-1.00).¹⁵ Cohen's kappa coefficient (κ) was used to examine the amount of agreement for classification of foot posture type between the two raters.¹⁶ The degree of agreement for classification was inter-

preted as poor ($\kappa \leq 0$), slight ($0.01 \leq \kappa \leq 0.20$), fair ($0.21 \leq \kappa \leq 0.40$), moderate ($0.41 \leq \kappa \leq 0.60$), substantial ($0.61 \leq \kappa \leq 0.80$), almost perfect ($0.81 \leq \kappa \leq 1.00$).¹⁷⁻²⁰ All analyses were completed for the left and right feet of the models, but no side-to-side comparisons were performed. Significance was set a priori at $P < 0.05$. All statistical tests were performed using SPSS 17.0 (SPSS, Inc, Chicago, IL).

RESULTS

The means of the total of all FPI scores using five-imaged criteria across three assessments and two raters were 2.84 ± 2.06 for left and 1.47 ± 2.26 for right feet. The means of the total of all FPI scores using five-imaged criteria for two raters for each assessment session are provided in Table 2. The inter-rater reliability for each session was poor to moderate with large 95% CIs (Table 3). For the classification of foot posture based on raw scores, as well as the amount of agreement between two raters was poor for both left ($\kappa = 0.12$) and right ($\kappa = 0.19$) feet. The number of foot posture types classified by two raters based on the raw Foot Posture Index scores is provided in Table 4.

Table 2. The Average Total Scores (Mean \pm SD) from Five Imaged-based Criteria of Foot Posture Index-6

| | | All raters | Rater 1 | Rater 2 |
|---|-------|-----------------|-----------------|-----------------|
| Mean for FPI scores for the three assessments | Left | - | 3.54 ± 1.95 | 2.14 ± 2.57 |
| | Right | - | 2.58 ± 1.82 | 0.36 ± 3.01 |
| Assessment 1 | Left | 2.69 ± 2.13 | 3.45 ± 2.34 | 1.93 ± 2.63 |
| | Right | 1.54 ± 2.53 | 2.63 ± 2.17 | 0.45 ± 3.31 |
| Assessment 2 | Left | 2.79 ± 2.12 | 3.33 ± 2.06 | 2.25 ± 2.74 |
| | Right | 1.46 ± 2.10 | 2.60 ± 2.04 | 0.33 ± 3.02 |
| Assessment 3 | Left | 3.05 ± 1.97 | 3.85 ± 1.99 | 2.25 ± 2.68 |
| | Right | 1.41 ± 2.19 | 2.53 ± 1.77 | 0.30 ± 3.31 |

Table 3. Inter-Rater Reliability Data.

| | | ICC (95% CI) | Standard Error of Measurement |
|--------------|-------|----------------------|-------------------------------|
| Assessment 1 | Left | 0.455 (0.171, 0.670) | 3.14 |
| | Right | 0.634 (0.405, 0.788) | 3.06 |
| Assessment 2 | Left | 0.527 (0.261, 0.718) | 2.92 |
| | Right | 0.334 (0.029, 0.582) | 3.43 |
| Assessment 3 | Left | 0.398 (0.103, 0.629) | 3.06 |
| | Right | 0.492 (0.217, 0.695) | 3.12 |

Table 4. The number of foot posture types classified by two raters based on the raw Foot Posture Index scores (80 feet)

| | | Rater 2 | | | | |
|---|-------------------|---|-----------|--------|----------|-----------------|
| | | Highly Supinated | Supinated | Normal | Pronated | Highly Pronated |
| Rater 1 | Left (Total =40) | | | | | |
| | Highly Supinated | 0 | 0 | 0 | 0 | 0 |
| | Supinated | 0 | 0 | 1 | 0 | 0 |
| | Normal | 0 | 4 | 26 | 2 | 0 |
| | Pronated | 0 | 0 | 5 | 2 | 0 |
| Highly Pronated | 0 | 0 | 0 | 0 | 0 | |
| | | Cohen's kappa coefficient = 0.12 | | | | |
| Rater 1 | Right (Total =40) | | | | | |
| | Highly Supinated | 0 | 0 | 0 | 0 | 0 |
| | Supinated | 0 | 2 | 0 | 0 | 0 |
| | Normal | 0 | 12 | 22 | 2 | 0 |
| | Pronated | 0 | 0 | 1 | 1 | 0 |
| Highly Pronated | 0 | 0 | 0 | 0 | 0 | |
| | | Cohen's kappa coefficient = 0.19 | | | | |
| The shadow areas indicate where agreement between the two raters would be expected. | | | | | | |

The intra-rater reliability was excellent for both left ($ICC_{3,3} = 0.956$, 95% CI; 0.925, 0.975, standard error of measurement [SEM] = 1.25) and right feet ($ICC_{3,3} = 0.959$, 95% CI; 0.931, 0.977, SEM = 1.33) when the average of the total FPI scores for two raters for each posture assessment session was analyzed between the sessions. Rater 1 had excellent reliability for left ($ICC_{3,3} = 0.900$, 95% CI; 0.832, 0.944, SEM = 1.85) and good for right ($ICC_{3,3} = 0.895$, 95% CI; 0.822, 0.941, SEM = 1.77) between the sessions. Rater 2 demonstrated excellent reliability for both left ($ICC_{3,3} = 0.952$, 95% CI; 0.819, 0.973, SEM = 1.69) and right ($ICC_{3,3} = 0.950$, 95% CI; 0.915, 0.972, SEM = 2.02) between the sessions.

DISCUSSION

Focusing on the image-based visual observation may allow clinicians to save assessment time, and multiple clinicians to evaluate large groups of patients quickly. Establishing the reliability of clinicians' rating digital, rather than live, observations from this assessment tool may be important for foot posture assessments and the interpretation of the findings in clinical and research settings. This study showed poor to moderate inter-rater reliability and excellent intra-rater reliability when using only the five image-based criteria of the FPI-6. The measurements of reliability for the five image-based criteria of the FPI-6 are similar to previously published reli-

ability of the measurements using all criteria of the FPI-6.^{8,11} However, the classification of foot posture did not improve the amount of agreement between raters. This finding of the categorization based on the raw FPI score using Kappa coefficient analysis contradicted the results of Morrison and Ferrari.¹³

While reliability of the FPI-6 has been reported previously, to the authors' knowledge, this is the first study to establish reliability only of the five observation-based criteria of the FPI-6. Cornwall et al¹¹ reported excellent intra-rater reliability of the FPI-6 among three clinicians who had different levels of clinical experience. Cain et al⁸ and Evans et al¹⁴ also demonstrated excellent intra-rater reliability of the FPI-6 in pediatric population among clinicians with various levels of clinical experiences. In our study, both raters had similar experience with the FPI-6 and demonstrated excellent intra-rater reliability when using only the five-image-based criteria of the FPI-6. The current results, coupled with these previous reports, seem to support the FPI as a reliable clinical tool for assessing static foot alignment, regardless of clinical experience with the tool.

However, caution is needed when interpreting FPI scores from five image-based criteria. The authors of the current study found poor to moderate inter-rater reliability for raw FPI scores with large 95% CIs and poor agreement between raters when the

average FPI scores for each rater for each assessment were used to classify each foot using the scales presented by Redmond et al⁴ Cornwall et al¹¹ also reported moderate agreement among clinicians for both the FPI-6 raw scores and classification of foot posture based on the raw scores. In contrast, using the FPI-6 in pediatric population, Morrison and Ferrari¹³ reported almost perfect agreement between two observers who were considered to have similar experience with using the FPI-6 when the FPI-6 scores were categorized. Barton et al²¹ demonstrated good inter-rater reliability among three raters in individuals with knee pain. The wide variety of results may be attributable to differences in the populations studied across these studies, as well as observing and rating static foot alignment with individual criteria of the FPI-6. Because several age-related changes in function influences foot posture and structure,²² the current results obtained from a young adult population cannot be assumed to be equally reliable when the FPI is applied to other populations. While direct visual observations of each participant's foot were performed for this assessment in the study of Morrison and Ferrari,¹³ the raters in the current study observed static foot alignment using photographs of the ankle and foot in order to score on the FPI. Additionally, the two raters in our current study had no experience with the index criteria. While the raters were exposed to the use of the FPI-6 and the rating scale for each of the criteria during pilot work with open discussion about the index criteria, the two raters did not receive standardized training from a clinician who has previous experiences with the FPI-6. Therefore, the lower inter-rater reliability may be associated with the level of two raters' experiences with FPI assessment.

LIMITATIONS

The primary limitation of this study is that the authors' used the categories of foot posture that were proposed for the actual FPI-6 by Redmond et al.⁴ The actual FPI-6 has six individual criteria which are summated to yield a composite score (-12 to +12) that is then used to classify foot posture based upon cutoff scores established by a previous study.⁴ However, the raters observed and rated static foot posture with incorporation using only the five image-based criteria of the FPI-6. This current study used an

established composite score ranging from -12 to +12 and the cutoff scores previously established for the actual FPI-6 to classify each participant's foot. The established composite FPI-6 score from -12 to -5 was classified as highly supinated, -1 to -4 as supinated, 0 to +5 as neutral, +6 to +9 as pronated, and +10 to +12 as highly pronated.^{4,6} However, we used only the five imaged-based criteria of the FPI-6 and the actual composite scores could only range from -10 to +10. It is possible that the use of the FPI composite score as the reference cutoff score negatively influenced the classification of foot posture in this study. Therefore, different cut-off values for foot posture classification may be required when using only the five imaged based criteria of the FPI-6 to screen foot posture types. Future research should investigate what cut-off score is best for using only the five image-based criteria of FPI-6. Furthermore, the palpation measure provides information regarding the orientation of the head of the talus. It is unknown how taking out the talar head palpation criterion from the original FPI-6 influences validity of static foot posture assessment. Further investigation is necessary to determine validity of the five image-based criteria of the FPI-6 and the correlation between the palpation and observation criteria of the FPI-6.

While previous studies demonstrated good internal construct validity of the actual FPI-6,^{5,6,23} the validity of the FPI incorporating only the five observation-based criteria remains unknown, resulting in agreement between raters that was perhaps underestimated and not generalizable.²⁰ Establishing validity of this assessment may be necessary in order to provide accurate evaluation of the foot posture. Furthermore, the sample size used in this study may lead to underestimation of Cohen's Kappa analysis.²⁰ Although thirty cases with two raters was deemed appropriate for this study, larger sample sizes are mathematically more likely to produced very small confidence intervals, which may lead to more precise estimates of agreement. Further investigation with a larger sample size may be necessary to establish normative values for the FPI composite score based upon only the five image-based criteria.

This image-based visual observation is important and useful to clinicians because this may allow them to save assessment time, and multiple clinicians to

evaluate large groups of patients quickly. However, scoring of the FPI appears to be influenced by soft tissue morphology,²³ and slight changes in angle for observation may lead to inconsistent observations of static foot posture between clinicians. While the investigators in this current study made efforts to take the photographs as consistently as possible, it is possible that minor inconsistencies in the photographs of each participant could have negatively influenced the quality of the assessment of each image-based criterion. However, the current results still demonstrated moderate to excellent associated reliability.

Finally, the participants were from a sample of convenience; therefore, there was no control for the distribution of foot types among participants. It will be important in the future investigations to consider recruiting equal distributions of foot type among the participants enrolled.

CONCLUSION

The FPI-6 is a clinical evaluation tool used to assess static foot posture in all three planes, which includes five multisegmental visually observable criteria and an additional criterion of talar head palpation. Using only the five image-based criteria of the FPI-6 may be beneficial for clinicians to save assessment time and evaluate large groups of patients quickly. The results of the current study indicate that the use of only the five image-based criteria of the FPI-6 demonstrates strong intra-rater but poor to moderate inter-rater reliability. Additionally, the visual attempt at classification of foot posture demonstrated poor agreement between raters. Therefore, caution is needed when interpreting FPI scores from the five image-based criteria. It is important for future investigations to establish appropriate cutoff values for the foot posture classification system using only the five image-based criteria, as well as normative values for the FPI composite scores, with larger sample sizes, in order to effectively assess static foot posture.

REFERENCES

1. Burns J, Keenan AM, Redmond A. Foot type and overuse injury in triathletes. *J Am Podiatr Med Assoc.* 2005;95(3):235-241.
2. Cowan DN, Jones BH, Robinson JR. Foot morphologic characteristics and risk of exercise-related injury. *Arch Fam Med.* 1993;2(7):773-777.
3. Sommer HM, Vallentyne SW. Effect of foot posture on the incidence of medial tibial stress syndrome. *Med Sci Sports Exerc.* 1995;27(6):800-804.
4. Redmond AC, Crosbie J, Ouvrier RA. Development and validation of a novel rating system for scoring standing foot posture: The foot posture index. *Clin Biomech (Bristol, Avon).* 2006;21(1):89-98.
5. Teyhen DS, Stoltenberg BE, Eckard TG, et al. Static foot posture associated with dynamic plantar pressure parameters. *J Orthop Sports Phys Ther.* 2011;41(2):100-107.
6. Keenan AM, Redmond AC, Horton M, Conaghan PG, Tennant A. The foot posture index: Rasch analysis of a novel, foot-specific outcome measure. *Arch Phys Med Rehabil.* 2007;88(1):88-93.
7. Redmond AC, Crane YZ, Menz HB. Normative values for the foot posture index. *J Foot Ankle Res.* 2008;1(1):6.
8. Cain LE, Nicholson LL, Adams RD, Burns J. Foot morphology and foot/ankle injury in indoor football. *J Sci Med Sport.* 2007;10(5):311-319.
9. Hreljac A, Marshall RN, Hume PA. Evaluation of lower extremity overuse injury potential in runners. *Med Sci Sports Exerc.* 2000;32(9):1635-1641.
10. Sim J, Wright CC. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Phys Ther.* 2005;85(3):257-268.
11. Cornwall MW, McPoil TG, Lebec M, Vicenzino B, Wilson J. Reliability of the modified foot posture index. *J Am Podiatr Med Assoc.* 2008;98(1):7-13.
12. Barton CJ, Levinger P, Crossley KM, Webster KE, Menz HB. Relationships between the foot posture index and foot kinematics during gait in individuals with and without patellofemoral pain syndrome. *J Foot Ankle Res.* 2011;4:10.
13. Morrison SC, Ferrari J. Inter-rater reliability of the foot posture index (fpi-6) in the assessment of the paediatric foot. *J Foot Ankle Res.* 2009;2:26.
14. Evans AM, Rome K, Peet L. The foot posture index, ankle lunge test, beighton scale and the lower limb assessment score in healthy children: A reliability study. *J Foot Ankle Res.* 2012;5(1):1.
15. Portney LG, Watkins MP. *Foundations of clinical research : Applications to practice.* 3rd ed. Upper Saddle River, N.J.: Pearson/Prentice Hall; 2009.
16. Cohen J. Citation-classic - a coefficient of agreement for nominal scales. *Cc/Soc Behav Sci.* 1986(3):18-18.
17. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychological Bulletin.* 1971;76(5):378-382.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.

-
19. Kundel HL, Polansky M. Measurement of observer agreement. *Radiology*. 2003;228(2):303-308.
 20. Crewson PE. Fundamentals of clinical research for radiologists - reader agreement studies. *Am J Roentgenol*. 2005;184(5):1391-1397.
 21. Barton CJ, Bonanno D, Levinger P, Menz HB. Foot and ankle characteristics in patellofemoral pain syndrome: A case control and reliability study. *J Orthop Sports Phys Ther*. 2010;40(5):286-296.
 22. Endo M, Ashton-Miller JA, Alexander NB. Effects of age and gender on toe flexor muscle strength. *J Gerontol A Biol Sci Med Sci*. 2002;57(6):M392-397.
 23. Menz HB, Munteanu SE. Validity of 3 clinical techniques for the measurement of static foot posture in older people. *J Orthop Sports Phys Ther*. 2005;35(8):479-486.