Research Article

Reliability and validity of the Sit-To-Stand Test to assess Global Foot Mobility

Mark W Cornwall[1]*, Carissa Lane[2], Jennifer Norwood[3], Sara Patterson[4] and Daniel Strauss[4]

[1]Professor, Department of Physical Therapy and Athletic Training, Northern Arizona University, Flagstaff, Arizona 86004, USA

[2]Staff Physical Therapist, Healthsouth Rehabilitation Hospital, Mesa, Arizona 85206, USA

[3]Staff Physical Therapist, Therapeutic Associates Inc. Ballard Physical Therapy, Seattle, Washington 98107, USA

[4]Doctor of Physical Therapy Student, Department of Physical Therapy and Athletic Training, Northern Arizona University, Flagstaff, Arizona 86004, USA

*Address for Correspondence: Dr. Mark Westover Cornwall, Professor, Department of Physical Therapy and Athletic Training, Northern Arizona University, Flagstaff, Arizona 86004, USA, Email: mark.cornwall@nau.edu

SUMMARY

The Sit-to-Stand test (STST) involves comparing the change in a person's non-weight-bearing and weight-bearing foot posture to quickly classify a person's overall foot mobility. Despite the simplicity of the test, its reliability and validity has not been established. The purpose of this study is to determine the intra-rater and inter-rater reliability of the STST as well as its validity. Ninety-seven subjects with a mean age of 25 years (±3.7) participated in the study. Each subject's foot posture from non-weight-bearing to weight-bearing was evaluated by two different raters. Each rater classified each subject's change in foot posture as "Hypomobile", "Normal" or "Hypermobile". This same procedure was repeated approximately one week later without the raters being able to review what their original classification for that subject had been. The subjects also had their foot mobility quantified by measuring the height and width of their dorsal arch in both non-weight-bearing and weight-bearing. These quantitative measures of foot mobility were then classified as "Hypomobile", "Normal", or "Hypermobile" using quartiles. A series of Cohen's Kappa coefficients were used to assess the amount of agreement between the visual classifications by each rater as well as the classification between the observational and objective classifications. The between-day Kappa coefficients ranged from 0.613 to 0.719 and the inter-rater Kappa coefficients ranged from 0.473 to 0.531. The Kappa coefficients between the visual and quantitative classifications ranged from 0.281 to 0.436. The STST should therefore be used with caution because of its moderate between-rater reliability and validity.

INTRODUCTION

Either limited or excessive foot mobility, particularly that of the medial longitudinal arch, has been shown to influence lower extremity kinematics. Williams et al. reported that the runners in their study who had mobile arches demonstrated decreased internal rotation excursion of their tibia, a greater eversion-to-tibial internal rotation ratio, decreased second peak vertical ground reaction force, and decreased vertical loading rates compared to those with normal or limited mobility [1]. In an earlier study, Williams reported that decreased mobility of the arch in runners was related to an increased need for compliance at other lower extremity joints, such as the knee, and they theorized that arch mobility could therefore be related to running-related injuries [2]. In 2016, Wyndow and associates reported that foot mobility was significantly related to the frontal plane projection angle of the lower extremity during a single leg squat activity in healthy individuals [3]. Specifically, they reported that individuals with higher midfoot mobility had a greater frontal plane projection angle and recommended that the amount of foot mobility be considered in the clinical management of knee-related disorders.

In 2010, Barton and colleagues reported that individuals with patellofemoral pain had a more pronated foot posture as well as increased foot mobility compared to a control group [4]. These findings were further supported in a study by McPoil et al. the following year. In that study, foot mobility, as measured by the change in arch height between weight-bearing and non-weight-bearing, was four times more likely to be seen in individuals with patellofemoral pain compared to a control group [5]. Furthermore, Milles and associates reported that individuals with anterior knee pain who had increased midfoot mobility were more likely to experience a reduction in their symptoms when treated with pre-fabricated orthoses [6]. Foot hypermobility has also been associated with an increased risk of other injuries in sports, particularly the lower extremity [7]. Investigators have reported a relationship between foot mobility and such conditions as plantar fasciitis [8], lower extremity osteoarthritis [9], medial tibial stress syndrome [10,11] and anterior cruciate ligament injuries in females [12].

In the literature, foot mobility has been assessed with a variety of methods. One such method is that of the navicular drop test. Brody first described the navicular drop test in 1982. It consists of measuring the vertical change in the height of the navicular tuberosity between subtalar joint neutral position while standing and relaxed standing [13]. The test is therefore considered a measure of sagittal plane mobility of the midfoot. Because inter-rater reliability of the navicular drop test has been reported to range from poor to moderate [14-16], other methods of assessing foot mobility have been proposed. McPoil et al., described an alternative method of measuring vertical change of the arch by assessing the change in the height of the dorsum of the arch rather than the navicular tuberosity during weight-bearing and non-weight bearing. Using this method, they demonstrated good to excellent levels of intra-rater and inter-rater reliability [17]. Although reliability measures were not reported, assessment of foot mobility has also been described using the change between weight-bearing and non-weight-bearing of sagittal radiographic measures such as the calcaneal inclination angle and the calcaneal-first metatarsal angle [8].

In 2009, McPoil and colleagues described a method of assessing medial-lateral and vertical movement of the midfoot in both weight-bearing and non-weight-bearing that did not require palpation of the navicular tuberosity. Their study included 345 healthy individuals and they reported very high intra-rater and inter-rater reliability values for all of their measurements. In the same paper, they also described a measurement called the "Foot Mobility Magnitude", which represented the composite value for both the difference in dorsal arch height (vertical change in arch mobility) as well as the difference in midfoot width (change in medial-lateral midfoot mobility) [18].

Based on the relationship between foot mobility and lower extremity kinematics and injury, assessment of foot mobility should be included as part of a comprehensive physical examination for those individuals with foot-related injuries or disorders. Not only may such an assessment help clinicians to evaluate the person's overall foot function, but it may also assist in determining the appropriate footwear or foot orthoses prescription. While a variety of methods exist to assess foot mobility, they may not be suitable either because of marginal test reliability, the need for a radiographic image or limited time or a lack of equipment. Hoppenfeld described in his 1976 book, what he termed a "test for rigid or supple flat feet", based on observing the foot in sitting and then in standing [19]. The purpose of the test was to allow clinicians to quickly and easily determine the degree of foot mobility of an individual. In order for such a test to be clinically useful, however, it must demonstrate acceptable levels of within-rater and between-rater reliability and also have adequate validity. Although the test proposed by Hoppenfeld, sometimes referred to as the "Sit-to-Stand Test" (STST) is simple and quick, no data exists regarding its reliability nor its validity. Therefore, the purpose of this study was to assess the within-rater and between-rater reliability of the STST as well as determine if it is consistent with quantitative measures of overall foot mobility.

RESULTS

Subjects

A total of ninety-seven individuals (25 male, 72 female) between the age of 20 and 46 years participated in the study. Table 1 shows the demographic characteristics of the individuals used in this study such as height, weight, body mass index (BMI) and foot posture, measured using the Foot Posture Index (FPI).

Procedures

Reliability: Sixty-one individuals (14 male, 47 female) between the age of 22 and 46 years participated in the reliability phase of the study. Table 1 also contains the mean demographic information for the subjects used in this phase of the study. Each subject was instructed to sit on the edge of a table with their knees at 90 degrees and their feet dangling off the edge of the table and not touching the floor. While in this position, the overall shape and posture of their foot was observed by two different raters. The subject then stood with their feet comfortably apart and with a self-selected amount of lower extremity "toeing out". The subject's foot shape and posture was again observed. Rater 1 was an entry-level physical therapy student with minimal experience in evaluating or treating foot related conditions. Rater 2 was a licensed physical therapist with over 20 years of experience evaluating and treating foot related conditions. Based on the perceived change in foot posture from non-weight-bearing to weight-bearing, each subject's global foot mobility for each extremity was rated as "Hypomobile" (<25% change), "Normal" (25-75% change), or "Hypermobile" (>75% change) by both raters without knowledge of the other rater's classification. This same procedure was repeated approximately one week later without the raters being able to review what their original rating for that subject had been. In order to avoid biasing the raters, the FPI was performed after each rater had made their classification.

Although both feet were measured, only the right extremity was used for statistical analysis. In addition to descriptive statistics, a series of Cohen's Kappa coefficients, adjusted for both "prevalence" and "bias," [20], were calculated to determine the magnitude of intra-rater and inter-rater agreement for the STST.

As shown in table 2, the overall between-day agreement for Rater 1 and 2 was 74.2 and 81.3% respectively. The prevalence and bias adjusted between-day Kappa

Table 1: Demographic Characteristics of the Subjects Used in the Study. Values in Parentheses are Standard Deviations.

Variable	Reliability Phase			Validity Phase		
	Total n=61	Male n=14	Female n-47	Total n=97	Male n=25	Female n=72
Age (years)	25.3 (4.1)	25.8 (3.5)	25.1 (4.3)	25.0 (3.7)	25.7 (3.9)	24.8 (3.7)
Height (cm)	169.3 (7.8)	176.9 (6.7)	166.8 (6.4)	170.6 (8.1)	178.1 (6.6)	167.9 (6.9)
Weight (kg)	65.3 (13.0)	78.4 (12.1)	60.9 (10.1)	67.1 (13.6)	80.0 (11.7)	62.7 (11.3)
BMI (kg/m^2)	22.7 (3.4)	25.0 (3.1)	21.9 (3.2)	22.9 (3.4)	25.2 (2.9)	22.1 (3.3)
FPI	+3.2 (3.4)	+2.0 (3.4)	+3.6 (3.3)	+3.2 (3.2)	+3.0 (3.6)	+3.3 (3.1)

Table 2: Measures of Within-Rater and Between-Rater Agreement. Kappa Coefficient Values Have Been Adjusted For Bias and Prevalence [20].

	Agreement (%)	Kappa
Rater 1 (Between-Day)	74.2	0.613
Rater 2 (Between-Day)	81.3	0.719
Between-Rater (Day 1)	65.6	0.484
Between-Rater (Day 2)	68.8	0.531

coefficients for the two raters was 0.613 for Rater 1 and 0.719 for Rater 2. See table 2. Using the classification proposed by Landis et al. [21], such values would indicate "substantial" between-day reliability for each rater. Table 2 also shows the overall agreement and Kappa coefficients indicating the between-rater reliability of assessments on day 1 and 2. Overall agreement between Rater 1 and 2 was 65.6% and 68.8% with the Kappa coefficients being 0.484 for Rater 1 and 0.531 for Rater 2. The classification proposed by Landis et al. [21], would characterize such values as being "moderate". Table 3 shows the 3x3 tables used to determine the between-rater agreement.

Validity: All 97 subjects were included for the validity phase of the study. Each subject's dorsal arch height (DAH) and midfoot width (MFW) was measured at 50% of their overall foot length, first in non-weight-bearing and then again in weight-bearing using a digital caliper or linear gauge and the methodology described by McPoil et al. [18]. The vertical change (DiffDAH) and horizontal change (DiffMFW) of the foot between the non-weight-bearing and weight-bearing measurements was then calculated for each foot. A global foot mobility measure, called the mobility magnitude (MM), was then calculated for each subject using the following formula [18]: $MM = \sqrt{(DiffDAH)^2 + (DiffMFW)^2}$ Finally, each of the above measures were standardized to the subject's overall foot length [22]. The resulting normalized values, expressed as a percentage of foot length, were then classified as "Hypermobile", "Normal" or "Hypomobile" based on quartiles. The first or lowest quartile was designated as "Hypomobile", while the second and third quartiles were designated as "Normal". The fourth or highest quartile was designated as "Hypermobile". Again, although both feet were measured, only the right extremity was used for statistical analysis. In addition to descriptive statistics, a series of Cohen's Kappa coefficients, adjusted for bias and prevalence [20], were used to assess the amount of agreement between the visual classification of foot posture change from non-weight-bearing to weight-bearlng that was assigned by Rater 2 and the classification based upon quartiles from the quantitative measures of foot mobility.

The mean normalized values for DiffDAH, DiffMFW and MM for each visual classification by Rater 2 is shown in table 4. The resulting 3x3 tables for DiffDAH, DiffMFW and MM in addition to the percent agreement and Kappa coefficient values, adjusted for bias and prevalence, between Rater 2 and the classification based on the quartiles of the quantitative foot mobility measurements are shown in table 5. As can be seen, the amount of agreement between the visual classification of foot mobility by Rater 2 and the quantitative classification using quartiles varied depending on which

Table 3: The 3 x 3 Tables Used to Calculate Between-Rater Agreement on Each Day of Assessment. Kappa Coefficient Values Have Been Adjusted For Bias and Prevalence [20].

		Day 1			
		Rater 2 Classification			
		Hypermobile	Normal	Hypomobile	Total
Rater 1 Classification	Hypermobile	19	2	0	21
	Normal	9	19	0	28
	Hypomobile	0	11	4	15
	Total	28	32	4	64
Agreement = 65.6%; Kappa = 0.484					
		Day 2			
		Rater 2 Classification			
		Hypermobile	Normal	Hypomobile	Total
Rater 1 Classification	Hypermobile	21	2	0	23
	Normal	9	21	0	30
	Hypomobile	0	9	2	11
	Total	30	32	2	64
Agreement = 68.8%; Kappa = 0.531					

Table 4: Mean Quantitative Measures of Foot Mobility for Each Classification Determined by Rater 2. Values in Parentheses are Standard Deviations.

	Hypermobility (n=45)	Normal (n=49)	Hypomobility (n=3)
DiffDAH (% of Foot Length)	0.367 (0.177)	0.342 (0.138)	0.283 (0.079)
DiffMFW (% of Foot Length)	0.380 (0.143)	0.258 (0.152)	0.226 (0.045)
MM (% of Foot Length)	0.546 (0.180)	0.449 (0.153)	0.365 (0.079)

Table 5: The 3x3 Frequency Tables Used to Calculate Agreement Between Visual Classification and Classification Based on Each Quantitative Foot Mobility Measurement. Kappa Coefficient Values Have Been Adjusted For Bias and Prevalence [20].

DiffDAH					
		Quartile Classification			
		Hypermobile	Normal	Hypomobile	Total
Visual Classification	Hypermobile	16	17	12	45
	Normal	12	22	15	49
	Hypomobile	0	1	2	3
	Total	28	40	29	97

Agreement = 41.2%; Kappa = 0.281

DiffMFW					
		Quartile Classification			
		Hypermobile	Normal	Hypomobile	Total
Visual Classification	Hypermobile	14	25	6	45
	Normal	7	19	23	49
	Hypomobile	0	2	1	3
	Total	21	46	30	97

Agreement = 35.1%; Kappa = 0.374

MM					
		Quartile Classification			
		Hypermobile	Normal	Hypomobile	Total
Visual Classification	Hypermobile	15	23	7	45
	Normal	4	28	17	49
	Hypomobile	0	1	2	3
	Total	19	52	26	97

Agreement = 46.4%; Kappa = 0.436

quantitative measure was used. The agreement between the visual classification and classification based on DiffDAH had the lowest Kappa value (0.281), while the agreement between the visual classification and classification based on MM had the highest Kappa value (0.436). Based upon the suggested classification proposed by Landis, the Kappa values between visual classification and the classification based on either DiffDAH or DiffMFW would be considered "fair". The Kappa value between visual classification and classification based on MM would be considered "moderate" [21].

With respect to between-day reliability, rater experience does not appear to have a large effect. Values for Rater 2 were only slightly greater than those of Rater 1 and the categorical rating of each rater was "substantial". The larger between-day agreement values compared to between-rater agreement values seen in the current study is consistent with many other clinical tests used as part of a foot and ankle examination. These include such things as ankle dorsiflexion [23], subtalar joint neutral palpation [23,15], and navicular drop [15,16].

An analysis of the 3x3 tables used to calculate the percent agreement and Kappa coefficients between each rater shows that the two raters appear to agree more often with each other for those feet characterized as "Hypermobile" compared to those characterized as "Hypomobile". This might be related to the fact that with

"Hypermobile" feet, the flattening of the medial longitudinal arch and the concurrent widening of the foot is easier to see or is more pronounced compared to those with limited foot motion. In 2011, Cornwall et al., showed that the change in midfoot width had a stronger relationship to overall foot posture compared to arch height as measured by the FPI [24]. This dominant role in overall foot posture may therefore have contributed to the higher agreement by the raters.

Based on the findings of this study, the authors feel that the within-rater and between-rater agreement of the STST is sufficient for the test to be used as part of a comprehensive physical examination of the foot. Because the within-rater and between-rater agreement was only "moderate" or "substantial", the test should be used with caution, especially when comparing the rating of one clinician to that of another.

The validity of the STST was measured by comparing the classification of foot mobility by Rater 2 to that based on quartiles of the normalized quantitative values of foot mobility. This comparison yielded either "fair" or "moderate" agreement between the two classifications (Table 5). It is important to note that both DiffDAH and DiffMFW showed "fair" agreement between visual and quantitative classification, while MM demonstrated "moderate" agreement. MM is a calculated value based on both the vertical and horizontal change in foot posture and therefore represents a more comprehensive representation of the foot's change from non-weight-bearing to weight-bearing [18]. The greater agreement between visual classification and the quantitative classification using MM illustrates that Rater 2 did not focus on either vertical or horizontal change in foot posture, but rather the global change in foot posture that was observed when determining the classification.

Based on the resulting 3x3 tables used to calculate Kappa values, Rater 2 was better at accurately identifying "Hypermobility." Rater 2 agreed 57.1% and 63.6% of the time with the classification of "Hypermobility" using DiffDAH and DiffMFW respectively, but only 6.9% and 3.4% of the time with the classification of "Hypomobility." Using MM, Rater 2 agreed with the quantification classification of "Hypermobility" 78.9% of the time, but only 7.7% of the time with the quantification classification of "Hypomobility." As can be seen in table 5, a large number of feet that were classified as "Hypomobile" by the quantitative measure were classified as "Normal" by Rater 2. This would indicate that a change in foot posture of less than 25% is more difficult to visually discriminate compared to a change greater than 75%. It is possible that the sample of subjects used in the study did not have an adequate representation of those with limited foot mobility, therefore, such individuals would be wrongly classified as having normal rather than limited mobility. In consideration of such a possibility, the distribution of FPI values for the subjects in the study was analyzed. The mean FPI for the subjects in the current study was +3.2 with a standard deviation of 3.2 (Table 1). The distribution of these scores were normally distributed and 20% of the subjects had an FPI of zero or less, which is considered to be significantly supinated [25]. Since there is a statistically significant positive relationship between a people's FPI and the amount of foot mobility [24], it is therefore reasonable to assume that the current study had an adequate representation of individuals with limited foot mobility.

A limitation of the current study is the small number of males measured compared to females. Normalizing the quantification of foot mobility relative to each subject's foot length, however, increased measurement reliability and reduced the possible bias introduced by such a large proportion of females in the study and increased [22]. In addition, because there is no expectation that a rater would visually classify a male's foot mobility differently than a female's foot, the authors feel that the lack of more males in the study had minimal impact on the results. A replication of the study with more males would be able to confirm this.

The finding of "fair" to "moderate" validity of the STST does not preclude clinicians from using the test to quickly and easily classify a person's overall foot mobility, but its use does warrant caution, especially with regard to its interpretation. On the other hand, the STST has value beyond that of classifying foot mobility. For example, the test would provide valuable information regarding a person's willingness and ability to fully load their foot, especially if more extensive gait analysis is not warranted, not possible or is contraindicated. As such, clinicians may choose to use the STST without classifying a person's overall foot mobility.

The results of this study indicate that the STST has "substantial" between-day reliability and "moderate" between-rater reliability and that neither is influenced significantly by the experience of the rater. Further, "fair" to "moderate" validity of the STST was found when compared to quantitative measures of foot mobility. The authors believe that these findings are sufficient for clinicians to use the test as part of a comprehensive physical examination, but that it should be used with some caution. Such caution stems from the fact that the agreement between a visual classification of foot mobility and a classification based on quantitative measures yielded an overall agreement of between 35.1% and 46.4%. Such overall low agreement casts considerable doubt on the interpretation of the test's result. In addition, identifying those with limited foot mobility was poor. As such, the authors feel that clinicians that are interested in classifying individuals as "Hypermobile", "Normal" or "Hypomobile" would likely be better served by using more quantitative methods.

REFERENCES

1. Williams DSB, Tierney RN, Butler RJ. Increased medial longitudinal arch mobility, lower extremity kinematics, and ground reaction forces in high-arched runners. J Ath Train. 2014; 49: 290-296. **Ref.:** https://goo.gl/8FDw5R

2. Williams DS, Davis IM, Scholz JP, Hamill J, Buchanan TS. High-arched runner's exhibit increased leg stiffness compared to low-arched runners. Gait Posture. 2004; 9: 263-269. **Ref.:** https://goo.gl/BX6TQ3

3. Wyndow N, De Jong A, Rial K, Tucker K, Collins NJ, et al. The relationship of foot and ankle mobility to the frontal plane projection angle in asymptomatic adults. J Foot Ankle Res. 2016; 9: 1-7. **Ref.:** https://goo.gl/ffs41s

4. 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: 286-296. **Ref.:** https://goo.gl/vMqUB5

5. McPoil TG, Warren M, Vicenzino B, Cornwall MW. Variations in foot posture and mobility between individuals with patellofemoral pain and those in a control group. J Am Pod Med Assoc. 2011; 101: 289-296. **Ref.:** https://goo.gl/KhqixN

6. Mills K, Blanch P, Dev P, Martin M, Vicenzino B. A randomised control trial of short term efficacy of in-shoe foot orthoses compared with a wait and see policy for anterior knee pain and the role of foot mobility. Br J Sports Med. 2012; 46: 247-252. **Ref.:** https://goo.gl/NuV9Vd

7. Pacey V, Nicholson LL, Adams RD, Munn J, Munns CF. Generalized joint hypermobility and risk of lower limb joint injury during sport: a systematic review with meta-analysis. Am J Sports Med. 2010; 38: 1487-1497. **Ref.:** https://goo.gl/HKJsxP

8. Sahin N, Ozturk A, Atıcı T. Foot mobility and plantar fascia elasticity in patients with plantar fasciitis. Acta Orthop Traumatol Turc. 2010; 44: 385-391. **Ref.:** https://goo.gl/7qNPxS

9. Reilly KA, Reilly K, Barker KL, Barker K, Shamley D, et al. Influence of foot characteristics on the site of lower limb osteoarthritis. Foot Ankle Int. 2006; 27: 206-211. **Ref.:** https://goo.gl/EVSvd4

10. Bandholm T, Boysen L, Haugaard S, Zebis MK, Bencke J. Foot medial longitudinal-arch deformation during quiet standing and gait in subjects with medial tibial stress syndrome. J Foot and Ankle Surg. 2008; 47: 89-95. **Ref.:** https://goo.gl/DUti9C

11. Bennett JE, Reinking MF, Pluemer B, Pentel A, Seaton M, et al. Factors contributing to the development of medial tibial stress syndrome in high school runners. J Orthop Sports Phys Ther. 2001; 31: 504-510. **Ref.:** https://goo.gl/r5TJnp

12. Loudon JK, Jenkins W, Loudon KL. The relationship between static posture and ACL injury in female athletes. J Orthop Sports Phys Ther. 1996; 24: 91-97. **Ref.:** https://goo.gl/p7GHBH

13. Brody DM. Techniques in the evaluation and treatment of the injured runner. Orthop Clin North Am. 1982; 13: 541-558. **Ref.:** https://goo.gl/Cfjcsc

14. Evans AM, Copper AW, Scharfbillig RW, Scutter SD, Williams MT. Reliability of the foot posture index and traditional measures of foot position. J Am Podiatr Med Assoc. 2003; 93: 203-213. **Ref.:** https://goo.gl/vvLdQi

15. Picciano AM, Rowlands MS, Worrell T. Reliability of open and closed kinetic chain subtalar joint neutral positions and navicular drop test. J Orthop Sports Phys Ther. 1993; 18: 553-558. **Ref.:** https://goo.gl/THjZnp

16. Schultz S, Nguyen DM, Windley T, Kulas AS, Botic T, et al. Intratester and intertester reliability of clinical measures of lower extremity anatomic characteristics: Implications for muticenter studies. Clin J Sports Med. 2006; 16: 155-161. **Ref.:** https://goo.gl/hDNT8n

17. McPoil TG, Cornwall MW, Vicenzino B, Teyhen DS, Molloy JM, et al. Effect of using truncated versus total foot length to calculate the arch height ratio. Foot. 2008; 18: 220-227. **Ref.:** https://goo.gl/u6ZZeu

18. McPoil TG, Vicenzino B, Cornwall MW, Collins NJ, Warren M. Reliability and normative values for the foot mobility magnitude: a composite measure of vertical and medial-lateral mobility of the midfoot. J Foot Ankle Res. 2009; 2: 1-12. **Ref.:** https://goo.gl/qWcnXe

19. Hoppenfeld S. Physical examination of the spine and extremities. Appleton-Century Crofts. New York. 1976; 232.

20. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993; 46: 423-429. **Ref.:** https://goo.gl/tSr2PN

21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33: 159-174. **Ref.:** https://goo.gl/AVFipS

22. Williams DS, McClay IS. Measurements used to characterize the foot and the medial longitudinal arch: reliability and validity. Phys Ther. 2000; 80: 864-871. **Ref.:** https://goo.gl/xh412b

23. Elveru RA, Rothstein JM, Lamb RL. Goniometric reliability in a clinical setting. Subtalar and ankle joint measurements. Phys Ther. 1988; 68: 672-677. **Ref.:** https://goo.gl/yNn4Ft

24. Cornwall MW, McPoil TG. Relationship between static foot posture and foot mobility. J Foot Ankle Res. 2011; 4: 1-9. **Ref.:** https://goo.gl/JKH1be

25. Redmond AC, Crane YZ, Menz HB. Normative values for the Foot Posture Index. J Foot Ankle Res. 2008; 1: 1-9. **Ref.:** https://goo.gl/VSe9vK