Clinical Research on Foot & Ankle

McPoil et al., Clin Res Foot Ankle 2013, 1:1
DOI: 10.4172/2329-910X.1000104

Research Article Open Access

# The Optimal Method to Assess the Vertical Mobility of the Midfoot: Navicular Drop versus Dorsal Arch Height Difference?

**Thomas G. McPoil[1*], Mark W. Cornwall[2], Monica G. Abeler[1], Kelsey J. Devereaux[1], Laura J. Flood[1], Shannon E. Merriman[1], Sara Sullivan[2], Michael J. van Der Laan[1], Tiffany A. Villadiego[1] and Karie Wilson[2]**

[1]*School of Physical Therapy, Rueckert-Hartman College for Health Professions, Regis University, Denver, CO, USA*
[2]*Department of Physical Therapy and Athletic Training, College of Health and Human Services, Northern Arizona University, Flagstaff, AZ, USA*

## Abstract

**Background:** Clinical measurements that assess the vertical mobility of the midfoot are utilized to assist the clinician in understanding general foot mobility as well as when prescribing foot orthoses and footwear. The primary purpose of this study was to determine the level of intra-rater and inter-rater reliability for navicular drop and dorsal arch height difference using both experienced and inexperienced raters at two geographical locations. In addition, the relationship between navicular drop and the dorsal arch height difference was assessed and normative values for these measurements were provided.

**Methods:**One hundred and ninety-two healthy participants, from two geographical locations, volunteered to participate in the study. Six raters performed the foot measurements required to calculate navicular drop and dorsal arch height difference. These measurements were assessed on 107 female and 85 male participants. Twenty participants from each geographical location were assessed in two sessions, separated by one week, to determine the reliability of the measurements.

**Results:** All foot measurements were shown to have high levels of intra-rater and inter-rater reliability. The strength of the correlation between navicular drop and the dorsal arch height difference was poor indicating that the two measures are not equivalent and cannot be used interchangeably by clinicians when assessing the vertical mobility of the midfoot.

**Conclusion:** While both measures have high levels of reliability, the dorsal arch height difference requires a special non-weight bearing platform as well as patient/client feedback to properly position the platform. As a result, the authors recommend the use of navicular drop as the method of choice for the assessment of the vertical mobility of the midfoot.

**Keywords:** Reliability; Navicular drop; Arch height; Normative data

## Introduction

The clinician often assesses foot mobility to assist in determining the appropriate prescription for footwear as well as foot orthoses. Based on the foot mechanism model described by Huson [1], the measurement of the vertical mobility of the midfoot can also be an indicator of rearfoot or hindfoot mobility because of the constrained tarsal mechanism formed by the talus, calcaneus, navicular, and the cuboid bones. One of the first methods described to assess the vertical mobility of the midfoot was navicular drop. In his 1982 paper describing techniques to evaluate the injured runner, Brody stated that the assessment of navicular drop was helpful in evaluating the amount of foot pronation [2]. The procedure described by Brody for performing the navicular drop assessment required the patient to stand barefoot on a firm surface with equal weight on each foot. After the navicular tuberosity was identified and marked using palpation, the patient's foot was placed in subtalar joint neutral position by palpating the talus in relation to the navicular bone. With the patient standing in subtalar joint neutral position, the height of the navicular tuberosity was marked on an index card that was placed on the medial side of the foot. The patient was then instructed to relax his/her feet and the position of the navicular tuberosity was marked again on the index card. The distance between the navicular tuberosity in subtalar joint neutral position was then subtracted from the navicular tuberosity resting position to determine the amount of navicular drop. Brody indicated that a normal navicular drop was approximately 10 mm and that values greater than 15 mm were considered abnormal. In his paper, however, Brody provided no information on the reliability of the measurement or the number of patients/subjects that were assessed to develop the normal and abnormal values he described for navicular drop [2].

Since the publication of Brody's paper, numerous studies have been conducted to determine the reliability of the navicular drop measurement. These studies have reported high levels of within-rater or intra-rater reliability for the measurement of navicular drop [3-13]. Of the nine studies that have assessed between-rater or inter-rater reliability, four studies have reported high levels of reliability [5,9,12,13] and five studies have reported low levels of reliability [4,7,8,10,11]. In all of these studies both experienced and inexperienced raters have been used, which suggests that the numbers of years a clinician has performed the navicular drop measurement does not influence the level of inter-rater reliability. Other reasons that have been proposed for the low levels of inter-rater reliability associated with navicular drop include difficulty in locating the navicular tuberosity as a result of anatomical variations among individuals as well as the in-consistency

***Corresponding author:** Thomas G. McPoil, PT, PhD, School of Physical Therapy, Rueckert-Hartman College for Health Professions, Regis University, 3333 Regis Blvd, G-4, Denver, CO 80221, USA, Tel: 303-964-5137; E-mail: tmcpoil@regis.edu

**Received** January 14, 2013; **Accepted** March 27, 2013; **Published** April 02, 2013

**Citation:** McPoil TG, Cornwall MW, Abeler MG, Devereaux KJ, Flood LJ, et al. (2013) The Optimal Method to Assess the Vertical Mobility of the Midfoot: Navicular Drop versus Dorsal Arch Height Difference? Clin Res Foot Ankle 1: 104. doi:10.4172/2329-910X.1000104

of palpating subtalar joint neutral position [11]. The findings from these studies indicate that an individual clinician could consistently measure navicular drop as long as they were the only individual performing the measurements. However, when multiple clinics and different clinicians are required to measure navicular drop as part of multicenter outcome studies or injury prediction research, the between-rater consistency of performing the navicular drop measurement is questionable. It is also important to note that of the eleven studies that have assessed navicular drop reliability, the maximum number of subjects in any of these studies has been 60 with the majority reporting data on 20 to 30 subjects. Thus, normative data for navicular drop on a large, single cohort of healthy individuals in which reliability of the measure was assessed has not been previously published. Even with these issues, navicular drop is still widely used by clinicians to assess the vertical mobility of the midfoot.

In an attempt to improve the clinical reliability of the assessment of vertical movement of the midfoot, McPoil et al reported on the use of photographic images obtained with the foot in weight-bearing and non-weight bearing [14]. While the photographic method described by these authors did demonstrate high levels of intra- and inter-rater reliability, the methodology required for the clinician to utilize this technique proved to be too time intensive and inefficient. In 2009, McPoil and colleagues described another method for assessing the vertical mobility of the midfoot [15]. This measurement, termed the dorsal arch height difference, was determined by subtracting the dorsal arch height of the foot measured in non-weight bearing from the dorsal arch height measured in weight bearing. The dorsal arch height for both measurements was assessed at 50% of the total foot length. In order to perform the non-weight bearing assessment of dorsal arch height, a portable measurement platform that could easily be placed under the plantar surface of the patients or clients foot was specially-built. These authors reported high levels of intra-rater and inter-rater reliability for the dorsal arch height difference and provided normative data on 345 healthy subjects that consisted of 211 females and 134 males [15]. Although the dorsal arch height difference has been shown to have high levels of reliability, the placement of the portable measurement platform to obtain the non-weight bearing dorsal arch height measurement requires that the patient/client provide feedback to the clinician while the platform is positioned under the plantar surface of the foot. While not an issue with healthy individuals, this could be a problem when collecting data on individuals with decreased plantar surface sensation associated with diabetes or peripheral vascular disease.

As previously noted, high levels of inter-rater measurement reliability are critical for multicenter outcome or injury prediction studies that include foot measurements performed by numerous clinicians or raters. In the only multicenter study to date, Piva et al utilized two pairs of clinicians to assess navicular drop at two geographical locations and reported high levels of both intra-rater and inter-rater reliability [12]. The four clinicians utilized in the study by Piva et al, were all experienced with time in practice ranging from 2 to 10 years [12]. A multicenter assessment for the dorsal arch height difference has not been conducted.

Based on the inconsistency in the previous studies that have assessed the inter-rater reliability of navicular drop and the importance of assessing the level of reliability for the dorsal arch height difference in more than one clinical setting, we designed this study for the following purposes: 1) to determine the level of intra-rater and inter-rater reliability for navicular drop and the dorsal arch height difference using both experienced and inexperienced raters at two geographical locations; 2) to determine the relationship between navicular drop and the dorsal arch height difference; and 3) to provide normative values on the same subject population for navicular drop and the dorsal arch height difference.

## Methods

### Participant characteristics

One hundred and ninety-two participants (107 females and 85 males) volunteered to participate in the study. Participants were from two geographic locations: 1) the Regis University population and surrounding Denver, Colorado community (Regis-CO); and 2) the Northern Arizona University population and surrounding Flagstaff, Arizona community (NAU-AZ). All participants met the following inclusion criteria: 1) no history of congenital deformity in the lower extremity or foot; 2) no previous history of lower extremity or foot fractures; 3) no systemic diseases that could affect lower extremity or foot posture; 4) no visible signs of foot pathology in either foot, including non-reducible claw or hammer toes, hallux valgus, hallux limitus, or hallux rigidus; and 5) no history of trauma or pain to either foot, lower extremity, or lumbosacral region at least 6 months prior to the start of the investigation. The total number of participants at the Denver site was 102 and 90 participants at the Flagstaff site. The number of female and male participants was 107 and 85, respectively. The mean age of the 192 participants was 26.3 $\pm$ 4.2 years with a range of 20 to 48 years. The mean age of the female and male participants was 25.6 $\pm$ 3.9 and 27.2 $\pm$ 4.3 years, respectively. The Institutional Review Boards of Regis University and Northern Arizona University approved the protocol for data collection and all participants provided written informed consent prior to participation.

### Instrumentation

Two instruments were manufactured for the study to permit the measurement of dorsal arch height. The weight bearing arch height gauge consisted of a digital caliper (Model #700-126, Mitutoyo America Corp, Aurora, IL 60502) with the fixed point attached to a 1.2 × 5.0 × 10.0 cm plastic block to hold the device in a vertical position. A sliding metal rod was attached to the moving point of the caliper to permit the assessment of arch height (see Figure 1). Non-weight bearing arch height was measured with a second identical digital caliper (Model #700-126, Mitutoyo America Corp, Aurora, IL 60502) mounted to a 0.5 × 12.0 × 41.0 cm plastic portable platform (see Figure 2). The plastic block attached to the fixed point of the caliper was attached to the portable platform so that it could be moved in order to permit proper alignment of the sliding metal rod to different foot lengths. To enhance the participant's awareness of the platform touching the plantar surface of their foot, 80-grit sandpaper was taped to the superior surface of

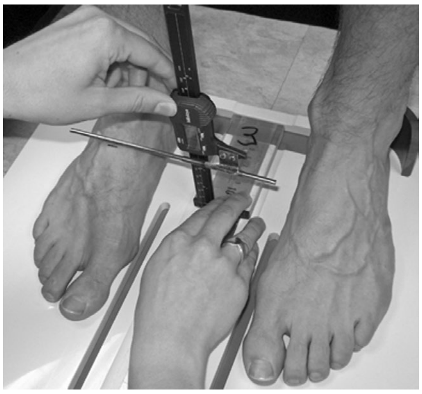

**Figure 1**: Digital gauge used to measure the dorsal arch height in weight bearing.

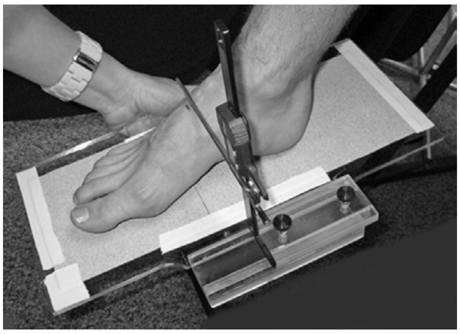

**Figure 2:** Portable platform with digital gauge used to measure non-weight bearing arch height.

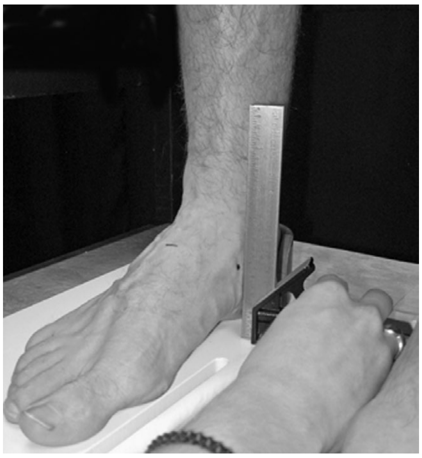

**Figure 3:** Measurement of navicular height using the angle ruler.

the portable platform. The measurement of navicular height was made using a commercially available angle ruler (Model #2117562, Ace Hardware Corp, Oak Brook, IL 60523) with a millimeter scale (see Figure 3).

## Procedures

Each subject was asked to stand on a previously described foot measurement platform so that total foot length, weight bearing dorsal arch height, as well as navicular height in both resting standing posture and subtalar joint neutral position could be measured in bilateral lower limb weight bearing (see Figure 4) [15]. Prior to obtaining the foot measurements, each subject was asked to stand on the foot measurement platform with both heels placed in left and right heel cups that were positioned 15.24 cm apart. Once the subject was properly positioned on the platform, the subject was instructed to place equal weight on both feet so that the weight bearing measurements could be obtained. Total foot length was first measured by placing the sliding bar on the centered metal ruler attached to the platform and moving the bar to just touch the longest toe, usually the hallux, of each foot (see Figure 4). Next, the dorsal arch height at 50% of total foot length was measured bilaterally using the weight bearing arch height gauge previously described. To determine the point of 50% of total foot length, the previously measured total foot length was divided in half and the dorsum of both feet were marked at the 50% length point using a water-soluble ink pen. The sliding metal rod of the weight bearing height gauge was then positioned over the 50% length mark and the vertical height from the top of the platform to the dorsum of the foot (DAH-Rest) was measured bilaterally (see Figure 1). Next, the navicular tuberosity was identified using palpation and marked with a water-soluble ink pen. The vertical height from the top of the platform to the ink mark on the navicular tuberosity in resting standing posture (NH-Rest) was measured bilaterally (see Figure 3). The rater then placed each foot in subtalar joint neutral position by asking the participant to elevate and lower the medial longitudinal arch of one foot followed by the other foot while the rater palpated the medial and lateral aspect of the head of the talus in relation to the navicular bone. When the rater felt congruency between the head of the talus and the navicular bone (subtalar joint neutral position) in both feet, the participant was instructed to maintain that position. The vertical height from the top of the platform to the ink mark on the tuberosity of the navicular (NH-SJN) was measured bilaterally.

Following the completion of the weight bearing measurements, each subject was asked to sit on the end of a table so that both lower legs were hanging in a perpendicular position to the floor with the feet non-weight bearing and the ankles slightly plantar-flexed. In this position, the non-weight bearing dorsal arch height was measured. The rater positioned the portable platform under, but without touching, the plantar surface of the right foot for each participant. As the portable platform was then moved upward to make contact with the plantar surface of the foot, the participant was instructed to state when they sensed the portable platform "just touching" the plantar surface of the heel, lateral forefoot and medial forefoot of the right foot simultaneously (see Figure 5). The participant was told to indicate to the rater if they felt that the portable

**Figure 4:** Measurement of total foot length on foot measurement platform.

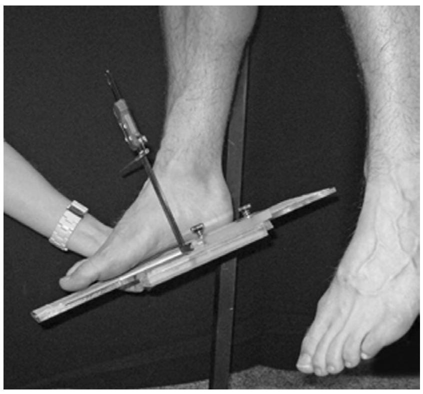

**Figure 5:** Placement of the portable platform under the plantar surface of the foot.

platform was forcibly pushing their foot into ankle dorsiflexion. If this did happen, the procedure was stopped and repeated so that the participant only sensed that the portable platform was just touching the plantar surface of the right foot. When the participant indicated that the portable platform was "just touching" the plantar surface of their right foot, the vertical digital caliper attached to the portable platform was positioned so that the sliding metal rod could be placed over the 50% foot length mark on the dorsum of the foot (see Figure 4). Once the sliding metal rod of the vertical digital caliper was positioned over the 50% foot length mark, the vertical height from surface of the portable platform to the dorsum of the foot was measured (DAH-NonWB). The same procedure was then repeated for the left foot.

To determine dorsal arch height difference (DiffAH), the dorsal arch height measured in weight bearing was subtracted from the dorsal arch height recorded in non-weight bearing. To determine navicular drop (NavDp) the height of the navicular tuberosity measured in resting standing posture was subtracted from the height of the navicular tuberosity recorded in subtalar joint neutral position.

### Determination of reliability

To establish intra-rater and inter-rater reliability for the measurements, three raters at two geographical locations (NAU-AZ and Regis-CO) were asked to assess the left and right feet of 20 randomly selected participants. At the NAU-AZ location, three raters were a physical therapist with 20 years of experience performing the measures of vertical mobility of the midfoot foot utilized in this study and two graduate physical therapy students with no previous experience. At the Regis-CO location, the three raters were a physical therapist with 21 years of experience performing the measures of vertical mobility of the midfoot foot utilized in this study, a Certified Athletic Trainer with five years of experience performing only the measures used to assess navicular drop, and a graduate physical therapy student with no previous experience. To insure consistency in the measurement procedures at each data collection site, a video recording that provided visual and verbal instructions for performing each of the foot measurements was made by the lead investigator. All raters were asked to watch the video at least once prior to attending two separate one-hour training sessions to practice the techniques to ensure that they were taking the measurements correctly. At each site, reliability data collection consisted of two sessions, one-week apart, in which each group of three raters performed the measurements on all 20 participants (40 participants total). During each session, all raters performed all the foot measurements on both feet of each participant twice with at least 10 minutes separating the two sets of measurements. In addition to inter-rater reliability, intra-rater reliability was assessed for both within-session and between-session for each rater. The marks placed over the dorsum and navicular tuberosity of each foot were removed after each set of measurements to prevent subsequent rater bias. The left and right feet for all 20 participants at each geographic location were treated as independent observations so that the analysis of reliability was conducted on 40 feet and reported separately for each geographic location.

### Data analysis

Intraclass Correlation Coefficients (ICC) were calculated to determine the consistency of each rater to repeatedly perform the measurements individually (intra-rater; $ICC_{3,1}$) for both within-session and between-session as well as in comparison to the other raters (inter-rater; $ICC_{2,3}$) [16,17]. For the between-session reliability assessment, the values for the four measurements collected twice in each session were averaged. The level of reliability for the ICC was classified using the characterizations reported by Landis and Koch [18]. These characterizations were: *slight*, if the correlation ranged from 0.00 to 0.20; *fair*, if the correlation ranged from 0.21 to 0.40; *moderate*, if the correlation ranged from 0.41 to 0.60; *substantial*, if the correlation ranged from 0.61 to 0.80; and *almost perfect*, if the correlation ranged from 0.81 to 1.00. In addition to ICC values, the Standard Error of the Measurement (SEM) was also calculated as another index of rater reliability. The SEM is in the same units as the original measurement and represents how the foot posture measurements would vary if measured more than once by each rater [19]. In addition to descriptive statistics, t-tests were performed to determine if differences existed between the left and right feet for the female and male participants. Pearson product-moment correlation coefficients were also performed to assess the relationship between NavDp and DiffAH. The alpha level was set at 0.05 and StatView, version 5.0.1 (SAS Institute Inc., Cary, N.C.) was used for all statistical analyses.

## Results

The within-session ICC values for the three Regis-CO raters ranged from 0.87 to 0.99 and from 0.88 to 0.99 for the three NAU-AZ raters for NH-Rest, NH-SJN, DAH-Rest, and DAH-NonWB, with SEM values ranging from 0.6 to 1.8 mm for both sets of raters. The between-session and inter-rater reliability ICC and SEM values for each set of three raters at both sites are shown in Tables 1 and 2. The between-session ICC values for all four measurements ranged from 0.83 to 0.96 for the three Regis-CO raters and from 0.85 to 0.98 for the three NAU-AZ

| **REGIS-CO** | **Rater 1 (experience 0 years)** | | | **Rater 2 (experience 5 years)** | | | **Rater 3 (experience 37 years)** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **ICC** | **Mean** | **SEM** | **ICC** | **Mean** | **SEM** | **ICC** | **Mean** | **SEM** |
| **NH Resting** | 0.91 | 45.75 | 1.80 | 0.95 | 45.85 | 1.30 | 0.96 | 44.14 | 1.10 |
| **NH SJN** | 0.83 | 48.06 | 2.30 | 0.93 | 50.29 | 1.40 | 0.91 | 49.07 | 1.60 |
| **DAH Resting** | 0.95 | 63.80 | 1.30 | 0.95 | 62.63 | 1.20 | 0.96 | 63.62 | 1.00 |
| **DAH NWB** | 0.93 | 71.91 | 1.60 | 0.94 | 72.10 | 1.50 | 0.94 | 73.40 | 1.50 |
| **NAU-AZ** | **Rater 1 (experience 0 years)** | | | **Rater 2 (experience 0 years)** | | | **Rater 3 (experience 35 years)** | | |
| | **ICC** | **Mean** | **SEM** | **ICC** | **Mean** | **SEM** | **ICC** | **Mean** | **SEM** |
| **NH Resting** | 0.95 | 45.85 | 1.60 | 0.92 | 45.53 | 2.00 | 0.92 | 45.62 | 1.80 |
| **NH SJN** | 0.88 | 50.50 | 2.10 | 0.88 | 49.33 | 1.90 | 0.85 | 48.94 | 1.90 |
| **DAH Resting** | 0.96 | 62.82 | 1.00 | 0.95 | 64.32 | 1.10 | 0.98 | 64.99 | 0.80 |
| **DAH NWB** | 0.95 | 74.13 | 1.30 | 0.93 | 73.57 | 1.50 | 0.97 | 74.89 | 1.10 |

**Note**: NH: Navicular Height; SJN: Subtalar Joint Neutral; DAH: Dorsal Arch Height; NWB: Non-Weight Bearing

**Table 1:** Between-session intra-rater reliability coefficients (ICC) and standard error of the means (SEM).

| | | REGIS -CO | | | NAU - AZ | | |
|---|---|---|---|---|---|---|---|
| | | ICC | Mean | SEM | ICC | Mean | SEM |
| Session 1 | NH Resting | 0.92 | 45.39 | 1.60 | 0.90 | 45.97 | 2.10 |
| | NH SJN | 0.89 | 49.28 | 1.70 | 0.92 | 49.95 | 1.60 |
| | DAH Resting | 0.96 | 63.87 | 1.10 | 0.95 | 66.77 | 1.20 |
| | DAH NWB | 0.94 | 73.39 | 1.40 | 0.97 | 77.11 | 1.00 |
| | | ICC | Mean | SEM | ICC | Mean | SEM |
| Session 2 | NH Resting | 0.93 | 45.20 | 1.60 | 0.94 | 45.56 | 1.70 |
| | NH SJN | 0.90 | 48.90 | 1.80 | 0.88 | 49.46 | 1.90 |
| | DAH Resting | 0.94 | 63.99 | 1.30 | 0.97 | 66.38 | 0.90 |
| | DAH NWB | 0.98 | 73.31 | 0.90 | 0.94 | 76.65 | 1.50 |

**Note**: NH: Navicular Height; SJN: Subtalar Joint Neutral; DAH: Dorsal Arch Height; NWB: Non-Weight Bearing

**Table 2:** Inter-rater reliability coefficients (ICC) and standard error of the means (SEM).

| | LEFT | | | RIGHT | | |
|---|---|---|---|---|---|---|
| **FEMALES** | **Mean (in mm)** | **Standard Deviation** | **Range (in mm)** | **Mean (in mm)** | **Standard Deviation** | **Range (in mm)** |
| NH Resting | 43.62 | 6.09 | 26 – 57 | 44.26 | 5.71 | 30 - 56 |
| NH SJN | 47.80 | 4.92 | 34 - 59 | 48.28 | 4.28 | 39 – 58 |
| DAH Resting | 61.08 | 4.41 | 46 - 73 | 62.05 | 4.27 | 45 - 71 |
| DAH NWB | 70.03 | 4.77 | 53 - 81 | 71.01 | 4.31 | 55 – 80 |
| NavDp | 4.26 | 3.06 | 0 - 16 | 4.14 | 3.11 | 0 – 14 |
| DiffAH | 9.04 | 2.68 | 1 - 15 | 9.04 | 2.96 | 1 - 15 |
| **MALES** | **Mean (in mm)** | **Standard Deviation** | **Range (in mm)** | **Mean (in mm)** | **Standard Deviation** | **Range (in mm)** |
| NH Resting | 46.43 | 7.48 | 29 - 63 | 47.63 | 7.15 | 34 - 64 |
| NH SJN | 51.46 | 5.63 | 39 - 66 | 52.62 | 5.09 | 40 – 66 |
| DAH Resting | 67.06 | 4.31 | 55 - 77 | 68.07 | 4.49 | 58 – 77 |
| DAH NWB | 76.45 | 4.38 | 64 - 88 | 77.62 | 4.61 | 68 – 89 |
| NavDp | 5.06 | 3.21 | 0 – 14 | 5.06 | 3.32 | 0 – 13 |
| DiffAH | 9.36 | 2.63 | 2 - 17 | 9.53 | 2.63 | 3 - 15 |

**Note**: NH: Navicular Height; SJN: Subtalar Joint Neutral; DAH: Dorsal Arch Height; NWB: Non-Weight Bearing

**Table 3:** Descriptive statistics for the female (n=107) and the male (n=85) participants.

raters. SEM values ranged from 1.1 to 2.3 mm for the Regis-CO raters and from 0.8 to 2.1 mm for the NAU-AZ raters. The inter-rater ICC values ranged from 0.88 to 0.97 for all four measurements for both sites with SEM values ranging from 0.9 to 1.9 mm.

The means and standard deviations for NH-Rest, NH-SJN, DAH-Rest, DAH-NonWB, DiffAH, and NavDp for the female and male participants are listed in Table 3. The results of the t-tests were not significant for NH-Rest (p=0.431), NH-SJN (p=0.452), DAH-Rest (p=0.103), DAH-NonWB (p=0.115), DiffAH (p=0.993), and NavDp (p=0.768) between the left and right feet for the female participants. The results of the t-tests were also not significant for NH-Rest (p=0.288), NH-SJN (p=0.159), DAH-Rest (p=0.136), DAH-NonWB (p=0.091), DiffAH (p=0.662), and NavDp (p=0.997) between the left and right feet for the male participants.

Pearson correlation coefficients between DiffAH and NavDp for the male participants were r=0.389 ($r^2$=0.152) for all feet (n=170), r=0.439 ($r^2$=0.193) for the left feet (n=85), and r=0.118 ($r^2$=0.344) for the right feet. Pearson correlation coefficients between DiffAH and NavDp for the female participants were r=0.376 ($r^2$=0.141) for all feet (n=170), r=0.332 ($r^2$=0.110) for the left feet (n=85), and r=0.416 ($r^2$=0.173) for the right feet.

Although the assessment of NavDp is the relative change between resting posture and subtalar joint neutral position while standing, previous researchers have discussed the importance of standardizing the measurement to the individual's foot length [13]. To assess if this influenced the correlations, both NavDp and DiffAH were standardized to foot length. The correlation coefficient between DiffAH and NavDp for the male participants after being standardized for foot length was r=0.411 ($r^2$=0.169) for the left feet and r=0.103 ($r^2$=0.321) for the right feet. For the female participants, the correlation coefficient between DiffAH and NavDp after being standardized for foot length was r=0.319 ($r^2$=0.102) for the left feet and r=0.414 ($r^2$=0.171) for the right feet.

## Discussion

The primary purpose our study was to determine the reliability of navicular drop and the dorsal arch height difference using experienced and inexperienced raters at two geographic locations. As noted in the introduction, previous studies have consistently reported high levels of intra-rater reliability but results for inter-rater reliability have been mixed. The results from the current study indicate excellent levels of intra-rater reliability, within-session and between-session, as well as inter-rater reliability for the foot measurements that were assessed in this study. The amount of rater experience or geographic location did no influence the level of reliability. Regardless of geographic location, the ICC values for the three inexperienced raters in this study ranged from 0.83 to 0.96 and the ICC values for the three experienced raters ranged from 0.85 to 0.98. Based on the characterization of ICC values proposed by Landis and Koch [18], the ICC values in this study would all be classified as "almost perfect". Although there are no standardized ways to assess the level of the SEM, the fact that all of the SEM values in the study were less than 5% of the mean values for all measures of foot posture is another indication of the consistency of the measurements.

The findings in the current study are similar to two of the most recent studies that have assessed intra-rater and inter-rater reliability of NavDp. Piva et al. reported high levels of reliability for NavDp that were assessed at two geographical locations [12]. Unlike Piva et al. who only used experienced raters, Barton et al also reported high levels of reliability for NavDp using two experienced raters with different years of experience and a third inexperienced rater with no clinical background [13]. Shultz et al conducted the only other study that assessed the reliability of NavDp with the same number of raters as in the current study [11]. In their study, six raters having varying amounts of clinical experience reported high levels of intra-rater reliability but low levels of inter-rater reliability. In the Shultz et al study, the inexperienced raters were trained by a more experienced rater and the training sessions consisted of a two hour practice session in which the raters practiced NavDp as well as two other lower extremity measurements [11]. The high levels of inter-rater reliability obtained in the current study could be attributed to an increased amount of time to practice performing NavDp as well as the video recording that provided visual and verbal instructions on the foot measurements which the raters could view at anytime prior to the start of the study.

While the reliability of DiffAH has not previously been assessed at different geographical locations, the results of this study are in agreement with the findings of McPoil and colleagues' who also reported "almost perfect" intra-rater and inter-rater reliability for AH-Rest and AH-NonWB [15]. Based on the high levels of reliability obtained in our study, we concluded that the consistency of the measurement techniques was at a high enough level to warrant further statistical analysis to evaluate DiffAH and NavDp.

The other purposes of this study were to determine the relationship between NavDp and DiffAH and to provide normative values for both measures of vertical mobility of the midfoot. Although no significant differences were found between the left and right feet for either the female or male subjects, the authors of the current study decided against combining the left and right feet to create a larger data set since several issues have been raised regarding this practice. Menz has noted that the counting of the left and right feet as single independent observations artificially increases the data set by counting the same subject twice [20]. Furthermore, Menz states that it could be problematic to conduct research on individual feet rather than people since the manner in which an individual foot functions is at least partly dependent on the person to whom the foot is attached [20]. Unfortunately, previous studies that have assessed navicular drop in large cohorts of subjects have failed to report their findings based on gender and/or extremity. In the current study, the mean for NavDp for all 192 subjects (384 feet) was 4.6 mm with a standard deviation of 3.18 mm. As previously noted, of the eleven studies that have assessed NavDp inter-rater reliability the largest single cohort assessed was 60 subjects. In that study by Evans et al, the mean for NavDp was 7.21 mm with no standard deviation reported [7]. Two more recent studies have provided normative values for NavDp on larger cohorts of subjects. Jonely et al. assessed 92 healthy subjects (184 feet) and reported similar values to those of Evans et al with the mean NavDp 7.00 mm and a standard deviation of 5.00 mm [21]. Most recently, Rathleff et al. evaluated NavDp for the left foot only in 79 subjects and reported a mean NavDp of 3.30 mm with a standard deviation of 0.50 mm [22]. Rathleff also assessed two-dimensional movement of the navicular bone in the same study and found that the dynamic amount of NavDp during walking was 5.40 mm [22]. These findings are in close agreement with Cornwall and McPoil who assessed three-dimensional movement of the navicular bone in 106 healthy subjects and reported 5.90 mm of vertical movement of the navicular bone during walking [23]. Thus the NavDp mean of 4.56 mm for the 384 feet in the current study, which represents normative data on the largest cohort of healthy individuals to date, is situated between the NavDp means previously reported in the literature and would appear to be representative of the amount of dynamic movement of the navicular bone during walking. In the current study, the mean for DiffAH for all 192 participants (384 feet) was 9.22 mm with a standard deviation of 2.74 mm. These values are less than the DiffAH values reported by McPoil et al, but are within plus or minus one and one-half standard deviations of the means for the 211 female and 134 male participants that were assessed in their study [15].

The results of the Pearson correlation coefficients indicate that DiffAH and NavDp are poorly related even though both measures assess the vertical mobility of the midfoot. For the male participants, whether considering all 170 feet or the 85 left and right feet, NavDp could only explain less than 20% of DiffAH. For the female participants, irrespective of considering all 214 feet or the 107 left and right feet, NavDp could also only explain less than 20% of DiffAH. In addition, standardizing NavDp and DiffAH by foot length did not improve the strength of the correlations for either the female or male participants. Based on these results, NavDp and DiffAH are not equivalent and cannot be used interchangeably by clinicians when assessing the vertical mobility of the midfoot. These findings suggest that the clinician must select either NavDp or DiffAH at the time of the initial examination and then continue to use the same method for all future re-assessments.

It is important to note that NavDp and DiffAH have been used to assess the vertical mobility of the midfoot in various patient populations. NavDp and DiffAH have both been shown to be effective methods to demonstrate increased foot mobility in individuals with patellofemoral pain in comparison to healthy controls [12,24]. Bennett et al. [25] and Raissi et al. [26] have also utilized NavDp to document increased foot mobility in athletes with medial tibial stress syndrome. While few studies have assessed the vertical mobility of the midfoot in patient populations with foot deformities, Schrader et al. demonstrated that NavDp could be used to reliably assess foot mobility in adults with rheumatoid arthritis [9].

Thus irrespective of whether NavDp or DiffAH is selected, the assessment of the vertical mobility of the midfoot should be considered as a component of the physical examination for certain patient populations. While further research examining the relationship between NavDp and DiffAH is always warranted, the fact that the assessment of the DiffAH requires a special non-weight bearing platform as well as feedback from the patient/client while the clinician is positioning the portable measurement platform under the foot, the authors would recommend the use of NavDp as the method of choice for the assessment of the vertical mobility of the midfoot.

A limitation of this study is that all participants were asymptomatic individuals. Thus, the normative values reported in this study may or may not be representative of individuals who have had an injury or a systemic disease such as rheumatoid arthritis. Another limitation of the current study is the assumption by the authors that when participants were asked to stand and place equal weight on both feet, that the participant was actually placing 50% of their body weight on each foot. While other methodologies, such as having one foot positioned on a scale, could be used to ensure that each participant was placing 50% of his/her body weight on each foot, the methodology used in this study can be easily replicated by the clinician. Tessem et al. has previously reported that the degree of asymmetry in the distribution of body weight between extremities in relaxed standing is 4% or less in healthy subjects [27]. Furthermore, the high level of intra-rater and inter-rater reliability as well as the degree of similarity for the foot posture measurements for the left and right feet would suggest that any degree of asymmetry in body weight distribution between extremities was minimal.

## Conclusions

In summary, the findings of the current study indicate that NavDp and DiffAH both have high levels of intra-rater and inter-rater reliability irrespective of rater experience or geographic location. The data provided on 192 participants represent normative values for these measures of the vertical mobility of the midfoot on one of the largest cohorts of healthy individuals to date. While both NavDp and DiffAH have high levels of reliability, since the DiffAH requires a special non-weight bearing platform as well as feedback from the patient/client while the clinician is positioning the portable measurement platform under the plantar surface of the foot, the authors recommend the use of NavDp as the method of choice for the assessment of the vertical mobility of the midfoot.

### References

1. Huson A (2000) Biomechanics of the tarsal mechanism. A key to the function of the normal human foot. J Am Podiatr Med Assoc 90: 12-17.
2. Brody DM (1982) Techniques in the evaluation and treatment of the injured runner. Orthop Clin North Am 13: 541-558.
3. Mueller MJ, Host JV, Norton BJ (1993) Navicular drop as a composite measure of excessive pronation. J Am Podiatr Med Assoc 83: 198-202.
4. Picciano AM, Rowlands MS, Worrell T (1993) Reliability of open and closed kinetic chain subtalar joint neutral positions and navicular drop test. J Orthop Sports Phys Ther 18: 553-558.
5. Sell KE, Verity TM, Worrell TW, Pease BJ, Wigglesworth J (1994) Two measurement techniques for assessing subtalar joint position: a reliability study. J Orthop Sports Phys Ther 19: 162-167.
6. McPoil TG, Cornwall MW (1996) The relationship between static lower extremity measurements and rearfoot motion during walking. J Orthop Sports Phys Ther 24: 309-314.
7. Evans AM, Copper AW, Scharfbillig RW, Scutter SD, Williams MT (2003) Reliability of the foot posture index and traditional measures of foot position. J Am Podiatr Med Assoc 93: 203-213.
8. Sutlive TG, Mitchell SD, Maxfield SN, McLean CL, Neumann JC, et al. (2004) Identification of individuals with patellofemoral pain whose symptoms improved after a combined program of foot orthosis use and modified activity: a preliminary investigation. Phys Ther 84: 49-61.
9. Shrader JA, Popovich JM Jr, Gracey GC, Danoff JV (2005) Navicular drop measurement in people with rheumatoid arthritis: interrater and intrarater reliability. Phys Ther 85: 656-664.
10. Vinicombe A, Raspovic A, Menz HB (2001) Reliability of navicular displacement measurement as a clinical indicator of foot posture. J Am Podiatr Med Assoc 91: 262-268.
11. Shultz SJ, Nguyen AD, Windley TC, Kulas AS, Botic TL, et al. (2006) Intratester and intertester reliability of clinical measures of lower extremity anatomic characteristics: implications for multicenter studies. Clin J Sport Med 16: 155-161.
12. Piva SR, Fitzgerald K, Irrgang JJ, Jones S, Hando BR, et al. (2006) Reliability of measures of impairments associated with patellofemoral pain syndrome. BMC Musculoskelet Disord 7: 33.
13. Barton CJ, Bonanno D, Levinger P, Menz HB (2010) Foot and ankle characteristics in patellofemoral pain syndrome: a case control and reliability study. J Orthop Sports Phys Ther 40: 286-296.
14. McPoil TG, Cornwall MW, Medoff L, Vicenzino B, Forsberg K, et al (2008) Arch height change during sit-to-stand: an alternative for the navicular drop test. J Foot Ankle Res 1: 3.
15. McPoil TG, Vicenzino B, Cornwall MW, Collins N, Warren M (2009) 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 2: 6.
16. Portney LG, Watkins MP (1993) Statistical measures of reliability. In: Foundations of Clinical Research: Applications to Practice. Mcgraw-Hill/ appleton and Lange, Connecticut, United States.
17. Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86: 420-428.
18. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33: 159-174.
19. Rothstein JM (1985) Measurement and clinical practice, the theory and application. In: Measurement in Physical Therapy, Volume 7 of Clinics in Physical Therapy Series. Churchill Livingstone, New York.
20. Vinicombe A, Raspovic A, Menz HB (2001) Reliability of navicular displacement measurement as a clinical indicator of foot posture. J Am Podiatr Med Assoc 91: 262-268.
21. Jonely H, Brismée JM, Sizer PS Jr, James CR (2011) Relationships between clinical measures of static foot posture and plantar pressure during static standing and walking. Clin Biomech (Bristol, Avon) 26: 873-879.
22. Rathleff MS, Nielsen RG, Kersting UG (2012) Navicula drop test ad modum Brody: does it show how the foot moves under dynamic conditions? J Am Podiatr Med Assoc 102: 34-38.
23. Cornwall MW, McPoil TG (1999) Relative movement of the navicular bone during normal walking. Foot Ankle Int 20: 507-512.
24. McPoil TG, Warren M, Vicenzino B, Cornwall MW (2011) Variations in foot posture and mobility between individuals with patellofemoral pain and those in a control group. J Am Podiatr Med Assoc 101: 289-296.
25. Bennett JE, Reinking MF, Pluemer B, Pentel A, Seaton M, et al. (2001) Factors contributing to the development of medial tibial stress syndrome in high school runners. J Orthop Sports Phys Ther 31: 504-510.
26. Raissi GR, Cherati AD, Mansoori KD, Razi MD (2009) The relationship between lower extremity alignment and Medial Tibial Stress Syndrome among non-professional athletes. Sports Med Arthrosc Rehabil Ther Technol 1: 11.
27. Tessem S, Hagstrøm N, Fallang B (2007) Weight distribution in standing and sitting positions, and weight transfer during reaching tasks, in seated stroke subjects and healthy subjects. Physiother Res Int 12: 82-94.