Reliability of ultrasound for measurement of selected foot structures

G Crofts¹, S Angin², K J Mickle¹,³, S Hill⁴, C J Nester¹

¹ - School of Health Sciences, University of Salford, Salford, Manchester, M6 6PU, United Kingdom
² - School of Physical Therapy and Rehabilitation, Dokuz Eylul University, Imciralti, 35340, Izmir, Turkey
³ - Biomechanics Research Laboratory, University of Wollongong, Australia.
⁴ - Podiatry Department, Singapore General Hospital, Singapore

Corresponding Author:
Dr Gillian Crofts
Allerton Building,
School of Health Sciences
University of Salford, Salford
M6 6PU,
United Kingdom
g.crofts@salford.ac.uk

Keywords:
1. Ultrasound
2. Foot structures
3. Reliability

Research highlights
1. We measured soft tissue structures in the foot and lower leg using ultrasound
2. Inter-operator reliability was good for tissue thickness and cross sectional area
3. Variation in measures was less than expected changes in structure due to disease or ageing
Abstract

Background: Understanding the relationship between the lower leg muscles, foot structures and function is essential to explain how disease or injury may relate to changes in foot function and clinical pathology. The aim of this study was to investigate the inter-operator reliability of an ultrasound protocol to quantify features of: rear, mid and forefoot sections of the plantar fascia (PF); flexor hallucis brevis (FHB); flexor digitorum brevis (FDB); abductor hallucis (AbH); flexor digitorum longus (FDL); flexor hallucis longus (FHL); tibialis anterior (TA); and peroneus longus and brevis (PER).

Methods: A sample of 6 females and 4 males (mean age 29.1 ± 7.2 years, mean BMI 25.5 ± 4.8) was recruited from a university student and staff population. Scans were obtained using a portable Venue 40 musculoskeletal ultrasound system (GE Healthcare UK) with a 5-13MHz wideband linear array probe with a 12.7 mm x 47.1 mm footprint by two operators in the same scanning session.

Results: Intraclass Correlation Coefficients (ICC) values for muscle thickness (ICC range 0.90-0.97), plantar fascia thickness (ICC range 0.94-0.98) and cross sectional muscle measurements (ICC range 0.91-0.98) revealed excellent inter-operator reliability. The limits of agreement, relative to structure size, ranged from 9.0%-17.5% for muscle thickness, 11.0%-18.0% for plantar fascia, and 11.0%-26.0% for cross sectional area measurements.

Conclusions: The ultrasound protocol implemented in this work has been shown to be reliable. It therefore offers the opportunity to quantify the structures concerned and better understand their contributions to foot function.
Introduction

Understanding the relationship between the lower leg muscles, foot structures and function is essential to explain how disease or injury may relate to changes in foot function and clinical pathology. Furthermore, characterisation of individual foot structures is required to explain their separate contributions to foot function. For example, explanations for differences between cavus and planus feet have primarily focused on external and bony morphology of the foot (e.g. foot posture index, arch height) [1] whereas soft tissue contributions have received only scant attention [2]. Measuring structural features of the foot can be challenging however, because the structures are relatively small and contact with the ground largely prevents measurement via the plantar surface. Measurement challenges can lead to low reliability of measures and thus hamper appropriate characterisation of how disease, ageing or other processes affect foot structures.

MRI is widely regarded as the gold standard method to quantify soft tissues in the foot [3]. However, access to MRI is often limited and it is not suitable for many clinical and community based studies, especially where repeated measures are required, such as longitudinal and intervention studies. In contrast, real-time ultrasound (US) is relatively inexpensive and portable and has been used previously to quantify lower limb muscle morphology [4, 5] and various foot soft tissue structures. The primary limitation of ultrasound is its operator dependency. Unlike other imaging tools, the site of imaging is entirely dependent upon the operator, and identification of key features, such as the area where muscle is thickest is subjective. Despite this, ultrasound has been shown to be comparable to MRI for cross-sectional area measurements [6] and muscle thickness measured from ultrasound can be used to predict muscle volume obtained from MRI [7].

Of the numerous functional soft tissues of the foot the plantar fascia has perhaps received most attention due to its association with heel pain and plantar fasciitis. Morphologically,
Increased fascial thickness has been associated with increased stiffness of the fascia [8] and reported in cases of heel pain and/or plantar fasciitis [9]. The measurement of the plantar fascia has primarily been limited to thickness at the site of insertion on the calcaneus, with only two studies reporting either inter-rater or intrarater reliability [10] or average bias for repeated measurements [11]. However, the thickness of the plantar fascia varies along its length, perhaps with some functional implications [12] and therefore the prior focus on the site of insertion seems rather limited. Thus, a reliable means of quantifying facial thickness in the rear, mid and forefoot could be advantageous.

Less attention has been given to the measurement of intrinsic foot muscles and most literature has focused on muscle atrophy in people with diabetes [3, 13]. However, atrophy of foot muscles may have wider clinical importance as toe flexor weakness has been found to be associated with toe deformity, and an increased risk of falling in older people [14, 15]. Measurement of intrinsic muscles is challenging because of their arrangement in a complex of four interwoven layers, and differentiating one muscle from another can be difficult. Ultrasound offers particular advantages in this case [16], because passive or active use of the muscles (e.g. toe flexion) might aid identification of structural boundaries [17]. Verhulst et al [18] reported reference values for abductor hallucis and extensor digitorum brevis, but did not consider the operator dependency of their ultrasound technique. This limits the value of their data set as a basis for comparison to groups with pathology, especially in clinical research where data might be captured over multiple sites and operators.

The posterior and anterior leg muscles have far greater muscle mass than the intrinsic foot muscles and are therefore important determinants of foot function. For example, atrophy in tibialis posterior (TP) and compensatory hypertrophy in flexor hallucis longus (FHL) muscle have been detected in pes planus foot types [19]. Flexor hallucis longus also plays an important role in the coupling between the rearfoot and 1st MTP joint kinematics [20] and toe
flexors shown to affect pressure under the forefoot [21]. The peroneus longus and brevis muscles are likely to have a role in maintaining the transverse foot arch, although there are only a few reports of their structural features [18].

Thus, there is a need to measure various foot and lower leg structures that relate to foot function using ultrasound as a flexible measurement approach, but there are either limited protocols (e.g. plantar fascia) or only reports of intra operator reliability[22]. The aim of this study was to investigate the inter-operator reliability of an ultrasound protocol to quantify features of: rear, mid and forefoot sections of the plantar fascia (PF); flexor hallucis brevis (FHB); flexor digitorum brevis (FDB); abductor hallucis (AbH); flexor digitorum longus (FDL); flexor hallucis longus (FHL); tibialis anterior (TA); and peroneus longus and brevis (PER).

**Material and methods**

A sample of 6 females and 4 males (mean age 29.1 ± 7.2 years, mean BMI 25.5 ± 4.8) was recruited from a university student and staff population. Participants were over the age of 18 years and had no self-reported lower limb disorders or systemic disease affecting the neuromusculoskeletal system (e.g. diabetes, rheumatoid arthritis). Written informed consent was obtained from each participant and ethics approval was obtained from the University's Research Ethics Panel.

**Data collection**

Ultrasound scans were performed by one experienced (8 years) and one inexperienced operator (the inexperienced operator had attended an intensive training period in ultrasound scanning of the foot and ankle over a four week period). Each participant had their right and left feet assessed using the Foot Posture Index (FPI) [1] as part of a larger study. If an
individual’s foot was classified as pronated or supinated, the foot with the highest score (furthest from zero) was scanned. If an individual was classified with a normal foot type, the foot with the lowest FPI score (closest to zero) was scanned. If the scores were equal for both sides, then the right side was selected for scanning.

A portable Venue 40 musculoskeletal ultrasound system (GE Healthcare, UK) with a 5-13 MHz wideband linear array probe with 12.7 mm x 47.1 mm footprint area was used for scanning. The scans were performed independently by each operator according to the scan protocol within the same session. Operator order was dependent upon logistics. Good contact was maintained between probe and skin without applying excessive pressure, three assessments were taken at each site with the probe removed between each recording. Each subject lay in the prone position for scanning PF, FHB and FDB muscles, and in the supine position for scanning the AbH, FDL, FHL, TA and PER muscles.

**Scanning protocol**

A summary of the probe position and sample images for all measures are detailed in figure 1. The plantar fascia (PF) was scanned in a longitudinal direction and images captured in three different regions. Firstly, the probe was placed over the PF at its insertion onto the calcaneus. The long axis of the probe was positioned on the longitudinal line between the medial calcaneal tubercle and the second toe for capturing the calcaneal end of the PF as described elsewhere [23]. Secondly, to assess the middle region of the PF the probe was placed on the same scanning line but at the level of the navicular tubercle. Thirdly, to image the metatarsal end of the PF the probe was placed on the same longitudinal scanning line but slightly proximal to the second metatarsal head.
To locate the muscles of the primary toe flexor muscles (FHB, FDB, AbH, FHL and FDL) the protocol described by Mickle et al [22] was followed. Briefly, the thickness of the FHB was measured longitudinally along the shaft of the 1st metatarsal at the thickest portion of the muscle, and then the probe was rotated 90° to obtain CSA of the muscle. The thickness of FDB was measured longitudinally along a line from the medial tubercle of the calcaneus to the third toe at the thickest portion of the muscle and then the probe was then rotated through 90° to measure the CSA. The thickest part of the AbH was located on a scanning line between the muscle’s origin on the medial calcaneal tuberosity and the navicular tuberosity in longitudinal section. The CSA was acquired on a scanning line drawn perpendicular to the long axis of the foot at the anterior aspect of the medial malleolus. Cross sectional area of FDL was imaged on a transverse line drawn at 50% of the distance between the medial tibial plateau and inferior border of the medial malleolus on the medio-posterior aspect of the tibia. On the same line the probe was then rotated through 90° to measure the thickness of the muscle in longitudinal section. The CSA of the FHL was measured on the same transverse line as above, but after manoeuvring the probe posteriorly. The probe was rotated 90° for capturing the longitudinal image of the muscle.

Peroneus longus and brevis were scanned together transversely to obtain the cross-sectional area. This was captured at a line 50% between fibular head and the inferior border of the lateral malleolus. Rotating the probe 90° at this point provided the longitudinal image of PER muscles where the thickness measurement was taken.

Tibialis anterior was scanned longitudinally at a 20% of the distance between fibular head and the inferior border of the lateral malleolus to obtain a thickness measurement. Due to the footprint size of the transducer, it was not possible to image the cross-sectional area of the muscle in its entirety.

**Image measurement**
All saved ultrasound images were decoded and measured by the same operator [SA] who was blinded to operator and participant information. Measurements were taken from still images of each structure using Image J software (National Institute for Health, Bethesda, MD, USA).

Data Analysis
The mean of three measurements, as reported elsewhere [22, 24], for each structure and operator were used to calculate Intraclass Correlation Coefficients (ICC 3,1). ICC >0.8 were classed as moderate, ICC >0.9 as excellent [25]. Bruton et al. [26] suggest that ICC values should not be interpreted in isolation as they can offer overly positive indicator of reliability. They also fail to contextualise the reliability in terms that might aid application of the measurements being tested in subsequent research. Therefore, limits of agreement (LoA) were also used to assess inter-operator reliability [27], which allows variation between operators to be put into the context of the measurements being taken, and likely changes in the measures due to disease, aging or injury. MedCalc® software (trial version 12.2.10.0, http://www.medcalc.org) was used for all data analysis.

Results
The foot posture index showed a median score of 0.5 with a range from -2 to 11, which represented a normal range of foot types. Descriptive information of the CSA and thickness measures for each operator are listed in Table 1 and Table 2. For CSA, the inter-operator ICC was excellent (ICC range 0.91-0.98). The limits of agreement were between 11.0%-26.0% of the CSA (Table 1). The CSA for flexor hallucis longus had the lowest ICC (0.91) (mean CSA measurements 4.12 cm² and 3.94cm² for tester 1 and tester 2 respectively) and flexor digitorum brevis had the highest ICC (0.98) (mean CSA measurements 1.82cm² and 1.85cm² for tester 1 and tester 2 respectively).
The inter-operator ICC for muscle thickness was excellent (ICC range 0.90-0.97). The limits of agreement were between 9.0%-17.5% of the average muscle thickness (Table 2). The anterior tibialis thickness had the lowest ICC (0.90), with mean thickness measurements of: 2.32cm and 2.33cm for tester 1 and tester 2 respectively. In contrast the flexor hallucis brevis and peroneus longus and brevis had the highest ICC (0.97) for muscle thickness values. For plantar fascia thickness, the inter-operator ICC was excellent (ICC range 0.94-0.98). The limits of agreement were between 11.0%-18.0% of the average thickness at each of the rear, mid and forefoot sites (Table 2).

Discussion
Ultrasound is a noninvasive imaging method whose key role is to define soft tissue structure, shape and size. The aim of this study was to investigate the inter-operator reliability of an ultrasound protocol to quantify features of structures related to foot function. Understanding the relationship between the lower leg muscles and intrinsic foot structures will provide opportunities to better understand how disease or injury may relate to changes in foot function. By regulated ultrasonographer standards in the human biomechanics field little mention is made of the factors that contribute to production and interpretation of ultrasound images. This can lead to an overly critical view of ultrasound and the potential to under value its role in soft tissue morphology assessment. This study has shown an ultrasound protocol can be reliable for the assessment of selected lower limb and foot muscles and plantar fascia, and thus we can be confident of using this protocol to provide information relevant to foot structure and function. However, reliability must be viewed as a relative concept, since the measurements need only be sufficiently consistent between operators (and over time) for their intended purpose, often to identify changes due to ageing, disease or intervention.
Mean measures (0.29cm) of plantar fascia at the calcaneal insertion were within the expected range (0.21-0.47cm) [28] and reliable (ICC; 0.98, 0.96, 0.94) at all three measurement sites (Table 2), which compares favourably with an ICC range 0.76-0.87 for longitudinal assessment of plantar fascia [10]. And the average bias for repeated measures of 0.01cm with limits of agreement ranging from 0.05cm to 0.07cm has been reported [11]. The good reliability of PF thickness, especially distal measures, coupled with good reliability in measures of AbH, FHB and FDB thickness suggests that we can investigate how these structures relate to foot function with confidence. Abductor hallucis has a large cross sectional area compared to other intrinsic muscles [29] and is an important contributor to medial long arch support and acquired flat foot in cases of tibialis posterior tendon rupture [30].

The limits of agreement in relation to the percentage of muscle thickness were within the ranges that we may expect for changes in muscle size due to ageing, disease or intervention. For example, in the region of 40-60% reduction has been reported in selected foot and leg muscle thickness [31], which is at least three time the size for the vast majority of our measures.

The LoA for the peroneal muscles were the poorest for those muscles tested (Table 2). Despite the superficial position of the peroneal muscles they were quite challenging to image concisely and the angle of the ultrasound beam in relation to the lateral borders meant that beam attenuation could have led to the reduced LoA. Using a lower frequency curvilinear probe may be a preferred option in the future because of the reduced effects of beam attenuation and the wider field of view provided by a narrower probe footprint.

The LoA for plantar fascia measurement was marginally worse than for the peroneal muscles (Table 2). This mid foot portion of the plantar fascia was the most challenging part of this structure to image due to the complex emergence of the digital slips. The similar acoustic
impedance of the individual slips limits the ability of ultrasound to differentiate these tissues easily.

A relatively low level of agreement (i.e. large Limit of Agreement) was observed for the CSA of FDL (26%) (Table 1). Compared to other muscles FDL is in close proximity to the tibia, and large bones are known attenuators of ultrasound. This would limit the ability to detect deep muscle borders close to the tibia. FDB, FHB and FHL have shown high reliability for CSA and thickness (Table 1 and 2). Previous work has measured the four toe flexor muscles in intrinsic (FDB and FHB together) and extrinsic (FDL and FHL together) groups [32]. This assumes that hallux and lesser digit function are coupled and the muscles have identical function. However, combinations of lesser and hallux toe deformities are highly variable and the hallucis and lesser toe muscles have different moment arms around the midfoot joints. Therefore, as per our protocol, it is therefore preferable to investigate the individual structures and their contributions to foot function separately.

The high ICC values should also be interpreted in the context of contrasting levels of experience between operator 1 and 2. Mickle et al [22] reported ICC values of 0.89-0.99 and relative Limits of Agreement of 8-28% for intra-operator measurements of the toe flexor muscles. These values are almost identical to the current study. This suggests that with structured but minimal training from an experienced operator, reliable data can be measured by relatively new operators. Coupled with its accessibility, this is a further advantage of ultrasound over MRI.

In summary, the ultrasound protocol implemented in this work has been shown to be reliable. It therefore offers the opportunity to quantify the structures concerned and better understand their contributions to foot function. Whilst we observed good reliability between operators of varying experience, we would still advocate structured education and training in operator skills.
References


Table 1: Reliability and limit of agreement results for CSA of foot and lower limb muscles

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Tester 1 (Mean)±(SD)</th>
<th>Tester 2 (Mean)±(SD)</th>
<th>ICC (3.1)</th>
<th>95% LoA (cm²) Lower</th>
<th>95% LoA (cm²) Upper</th>
<th>LoA (%) Average muscle size</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbH CSA (cm²)</td>
<td>3.03 (0.44)</td>
<td>3.06 (0.45)</td>
<td>0.91</td>
<td>-0.52</td>
<td>0.46</td>
<td>16</td>
</tr>
<tr>
<td>FDB CSA (cm²)</td>
<td>1.82 (0.54)</td>
<td>1.85 (0.52)</td>
<td>0.98</td>
<td>-0.34</td>
<td>0.28</td>
<td>17</td>
</tr>
<tr>
<td>FDL CSA (cm²)</td>
<td>2.64 (0.70)</td>
<td>2.54 (0.55)</td>
<td>0.92</td>
<td>-0.56</td>
<td>0.77</td>
<td>26</td>
</tr>
<tr>
<td>FHB CSA (cm²)</td>
<td>3.17 (0.50)</td>
<td>3.31 (0.50)</td>
<td>0.95</td>
<td>-0.69</td>
<td>0.40</td>
<td>17</td>
</tr>
<tr>
<td>FHL CSA (cm²)</td>
<td>4.12 (0.40)</td>
<td>3.94 (0.37)</td>
<td>0.91</td>
<td>-0.28</td>
<td>0.61</td>
<td>11</td>
</tr>
<tr>
<td>PER CSA (cm²)</td>
<td>3.98 (1.29)</td>
<td>4.09 (1.25)</td>
<td>0.97</td>
<td>-0.91</td>
<td>0.68</td>
<td>19.7</td>
</tr>
</tbody>
</table>
Table 2: Reliability and limit of agreement results for muscle thickness of foot, lower limb, and plantar fascia

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Tester 2 (Mean) (SD)</th>
<th>Tester 2 (Mean) (SD)</th>
<th>ICC (3.1)</th>
<th>95% LoA (cm)</th>
<th>LoA (% Average muscle size)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbH T (cm)</td>
<td>1.27 (0.14)</td>
<td>1.27 (0.18)</td>
<td>0.92</td>
<td>-0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>FDB T (cm)</td>
<td>1.05 (0.19)</td>
<td>1.03 (0.18)</td>
<td>0.96</td>
<td>-0.12</td>
<td>0.16</td>
</tr>
<tr>
<td>FDL T (cm)</td>
<td>1.85 (0.28)</td>
<td>1.82 (0.23)</td>
<td>0.92</td>
<td>-0.25</td>
<td>0.31</td>
</tr>
<tr>
<td>FHB T (cm)</td>
<td>1.59 (0.29)</td>
<td>1.59 (0.28)</td>
<td>0.97</td>
<td>-0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>FHL T (cm)</td>
<td>2.14 (0.29)</td>
<td>2.09 (0.25)</td>
<td>0.95</td>
<td>-0.17</td>
<td>0.29</td>
</tr>
<tr>
<td>PER T (cm)</td>
<td>1.39 (0.22)</td>
<td>1.35 (0.19)</td>
<td>0.97</td>
<td>-0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>TA T (cm)</td>
<td>2.32 (0.38)</td>
<td>2.33 (0.35)</td>
<td>0.90</td>
<td>-0.21</td>
<td>0.20</td>
</tr>
<tr>
<td>PF1 T (cm)</td>
<td>0.29 (0.05)</td>
<td>0.29 (0.05)</td>
<td>0.98</td>
<td>-0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>PF2 T (cm)</td>
<td>0.19 (0.03)</td>
<td>0.20 (0.03)</td>
<td>0.96</td>
<td>-0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>PF3 T (cm)</td>
<td>0.13 (0.01)</td>
<td>0.13 (0.01)</td>
<td>0.94</td>
<td>-0.017</td>
<td>0.011</td>
</tr>
</tbody>
</table>