

Rehabilitative ultrasound imaging for the assessment of selected morphometric parameters of the flexor hallucis longus muscle in healthy individuals: An inter- and intra-rater reliability study

Type

Research paper

Keywords

ultrasound, reliability, flexor hallucis longus, rater

Abstract

Introduction

Ultrasound imaging (USI) is useful to evaluate structures of the foot to guide treatment, but the reliability of USI technique needs to be clarified. The goal of the study was to evaluate the intra- and inter-examiner reliability of USI image capture, and measurement of the cross-sectional area (CSA) and thickness of the flexor hallucis longus (FHL) for experienced and novice examiners

Material and methods

FHL images were captured for 20 healthy adults. Reliability of image capture was evaluated between images repeated at 10-min interval for an experienced and a novice examiner. Reliability of image-based measurements was evaluated for one experienced and one novice rater, using all images. The intra-class correlation coefficient (ICC) and the standard error of measurement (SEM) were calculated

Results

Intra-examiner reliability of image capture for the FHL muscle examined by USI was excellent for both thickness (ICC_{3,1}, 0.944–0.976; SEM, 6.8%–10.0%) and CSA (ICC_{3,1}, 0.954–0.979; SEM, 10.8%–16.5%), with no effect of examiner experience. Reliability was also excellent for measurement of thickness (ICC_{3,1}, 0.954–0.972; SEM, 1.2%–9.6%) and CSA (ICC_{3,1}, 0.961–0.986; SEM, 9.2%–14.1%), with no effects of experience.

Conclusions

Reliability of image capture and image-based measurements developed by USI of CSA and thickness for the FHL muscle in healthy individuals was excellent, independent of the examiner experience.

1 **Rehabilitative ultrasound imaging for the assessment of selected morphometric**
2 **parameters of the flexor hallucis longus muscle in healthy individuals: An inter- and**
3 **intra- rater reliability study**

4 **Abstract:** Background: Ultrasound imaging (USI) is useful to evaluate structures of the
5 foot to guide treatment, but the reliability of USI technique needs to be clarified. The goal
6 of the study was to evaluate the intra- and inter-examiner reliability of USI image capture,
7 and measurement of the cross-sectional area (CSA) and thickness of the flexor hallucis
8 longus (FHL) for experienced and novice examiners. Methods: FHL images were captured
9 for 20 healthy adults. Reliability of image capture was evaluated between images repeated
10 at 10-min interval for an experienced and a novice examiner. Reliability of image-based
11 measurements was evaluated for one experienced and one novice rater, using all images.
12 The intra-class correlation coefficient (ICC) and the standard error of measurement (SEM)
13 were calculated. Results: Intra-examiner reliability of image capture was excellent for both
14 thickness (ICC_{3,1}, 0.944–0.976; SEM, 6.8%–10.0%) and CSA (ICC_{3,1}, 0.954–0.979; SEM,
15 10.8%–16.5%), with no effect of examiner experience. Reliability was also excellent for
16 measurement of thickness (ICC_{3,1}, 0.954–0.972; SEM, 1.2%–9.6%) and CSA (ICC_{3,1},
17 0.961–0.986; SEM, 9.2%–14.1%), with no effects of experience. Conclusion: Reliability of
18 image capture and image-based measurements of CSA and thickness for the FHL muscle in
19 healthy individuals was excellent, independent of the examiner experience.

20 .

21 **Keywords:** flexor hallucis longus, rater, reliability, ultrasound

22 **Level of evidence:** II.b

23

24

25

26

27

28

29

30 **Introduction**

31 The flexor hallucis longus is the primary plantar flexor of the first metatarsophalangeal
32 (MTP) and interphalangeal (IP) joints of the great toe; it also restrains the passive dorsiflex-
33 ion at the first MTP joint and contributes to ankle plantar flexion [1]. FHL tendinopathy is
34 one of the most common pathological conditions of the ankle and foot among active people
35 [2].

36 Ultrasound imaging (USI) is considered as a safe, non-invasive, and valid method to
37 measure muscle morphology to inform diagnosis, and to compare the cross-sectional area
38 (CSA) and thickness of the FHL in healthy subjects [3], subjects with pes planus [4,5] , and
39 dancers [6]. Extending the application of USI in clinical practice would require assessing the
40 reliability of measurements obtained.

41 Mickleet al. [7] evaluated the intra-examiner reliability of USI to measure the
42 morphology of the primary toe flexor muscles, while Crofts et al. [8] analyzed the inter-
43 examiner reliability of USI for selected foot structures. Both of these studies suggest that USI
44 offers the opportunity to quantify the structures of the foot for a better understanding of their
45 functional contribution. Researchers have identified the need for determining the reliability
46 of USI not only for measurement (USI-based calculation) but also for image capture (patient
47 positioning), both of which might be influenced by the experience of the examiner. Valera-
48 Calero et al. [9] have previously investigated the reliability of US image capture and
49 measurement of the CSA of the cervical multifidus muscle as a function of the examiner's
50 experience in asymptomatic individuals, with excellent reliability reported.

51 **The objective of this study** was to determine intra- and inter-examiner reliability of US
52 image capture (probe assessment/subject positioning) and measurement (scan/image
53 assessment) for the CSA and thickness of the FHL as a function of the examiner's experience.
54 Understanding the architecture of the FHL has implications for invasive physiotherapeutic
55 therapies, **as most of them use ultrasound guidance to guarantee accuracy** [10,11].

56 **Materials and Methods**

57 *Participants*

58 Twenty healthy volunteers were recruited from private physiotherapy clinics between
59 April and May 2020. **The procedures were approved by the local Ethics Committee (1046-**
60 **N-19) in accordance with the current national and international laws and regulations**
61 **governing the use of human subjects. Before participating in the study, the subjectss were**

62 fully informed about the protocol and written informed consent was obtained from each
63 participant before testing. Enrolled participants were females over the age of 18 years and
64 with no history of ankle pain over the previous year. The exclusion criteria were as follows:
65 acute lower limb injury within the previous six months; prior ankle surgery; history of low
66 back pain; history of peripheral nerve impairment; congenital/acquired foot deformities; any
67 other medical condition or treatment affecting muscle tone.

68 *Imaging capturing and FHL measurement protocol*

69 For ultrasound evaluation of the FHL muscle, the participants lay in the prone position,
70 with the feet unsupported at the edge of the bed. The prone position is normally used for
71 invasive treatment techniques [10,11]. Using a previously described procedure [10,11], the
72 FHL muscle was located at the midpoint of the distance between the fibular head and the
73 inferior border of the lateral malleolus on the posterior aspect of the fibula, using the Logiq
74 US machine (GE Healthcare, USA; Fig 1). Images were captured using a 12 MHz linear
75 transducer, in the longitudinal and transverse plane of the muscle. Image-based measures
76 were performed using ImageJ (version 1.53a; Bethesda, MD, USA).

77 [Figure 1]

78 *Reliability protocol*

79 Reliability of US image capture was evaluated for one novice (1 year of practice) and
80 one experienced (13 years of practice) examiner, both using the prone patient positioning as
81 previously described. Each examiner captured two images of the FHL muscle, with a 10-min
82 period between the two image-captures, with participant positioning required for each
83 assessment. Reliability of image-based measurements was evaluated, again, for one
84 experienced and a novice rater, with the same years of experience as the examiners previously
85 described. Measurements were performed in blinded fashion (no identification if images
86 were obtained by the novice or experienced examiner), with images coded using numerical
87 codes.

88 [Figure 1]

89 *Reliability of image-based measurements*

90 The intra-rater reliability of image-based measurements of the CSA and thickness of the
91 FHL was evaluated for two rater (experienced/novice), and using images obtained from the

92 experienced and novice operators. Inter-rater reliability was evaluated for measures of CSA
93 and thickness using the images obtained during the first imaging assessment from both the
94 experienced and novice examiner. We also evaluated inter-rater reliability of measurement
95 of the FHL CSA and thickness for the experienced and novice examiners for all images
96 obtained at once. For intra-examiner reliability, each examiner repeated measurements of
97 CSA and thickness at a 1-week interval.

98 *Statistical analysis*

99 Statistical analysis was performed using SPSS (v.22, IBM; Armonk, NY: IBM Corp.),
100 with an α error of 0.05 and a β error of 0.2 for both intra- and inter-examiner/rater reliability.
101 The Shapiro-Wilk test was employed for the normality assumption. A 2-way, mixed-model,
102 consistency-type, intra-class correlation coefficients (ICC) was calculated to test the intra-
103 and inter-rater reliability. Reliability was defined as fair ($ICC < 0.50$), moderate
104 ($0.50 < ICC < 0.75$), good ($0.75 < ICC < 0.90$), or excellent ($ICC > 0.90$) [12]. In addition, the ICC
105 and associated standard deviation (SD) were used to calculate the standard error measurement
106 as a measure of the precision of measurement, as follows: $SEM (\%) = (SD \times \sqrt{1 - ICC}) \times 100$.

107 **Results**

108 The age and relevant anthropometric data for the study group are summarized in Table
109 1.

110 [Table 1]

111 Reliability results for image capture and FHL thickness is reported in Table 2. The
112 results showed in this table demonstrate that reliability was excellent, overall, with ICC
113 values ranging between 0.944 to 0.976, with SEM values of 6.8% to 10.0%. There was no
114 difference between the experienced and novice examiner, with inter-examiner ICCs ranging
115 between 0.983 and 0.994, with SEM values of 0.1% to 5.3%.

116 [Table 2]

117 Reliability of image-based measurement of FHL thickness is reported in Table 3. The
118 intra-rater reliability was excellent, with ICCs ranging between 0.954 and 0.972 and SEM
119 values of 1.2% to 9.6%. There was no difference in reliability between the experienced and
120 novice rater, with inter-rater ICCs of 0.950 to 0.980 and SEM values of 6.2% to 8.5%.

121 [Table 3]

122 Reliability of image capture for CSA is reported in Table 4. The reliability was again
123 excellent, with ICCs of 0.954 to 0.979 and SEM values of 10.8% to 16.5%. There was no
124 difference between the experienced and novice examiner, with inter-rater ICC values of 0.986
125 to 0.984 and SEM values of 8.9% to 15.0%.

126 [Table 4]

127 Reliability of image-based measurement of CSA is reported in Table 5. The intra-rater
128 reliability was excellent, with ICCs of 0.961 to 0.986 and SEM values of 9.2% to 14.1%.
129 Again, there was no different between the experienced and novice rater, with inter-rater ICCs
130 of 0.961 to 0.991 and SEM values of 6.9% to 15.1%.

131 [Table 5]

132 Discussion

133 Our findings indicate high reliability of image capture of the CSA and thickness of the
134 FHL, as well as for image-based measurements of CSA and thickness. This reliability would
135 be independent of the examiner experience. To the authors' knowledge, this is the first study
136 that has evaluated the reliability of both USI capture (positioning) and measurement for the
137 FHL muscle considering the experience of the examiner. Our findings are consistent with
138 previous studies which have reported ICC values >0.90 [7,8], and underscore the feasibility
139 of using USI. Despite the high prevalence of FHL tendinopathy, there is limited evidence of
140 the role of the FHL in foot mechanics, and of the impact of an injury in this muscle [13].
141 Injuries to the FHL muscle and its tendon are often overlooked. In specific populations,
142 however, appropriate assessment of this muscle is essential, such as in dancers who rely on
143 the FHL for dynamic stability of the foot during movements that require alternation between
144 extreme plantar flexion and dorsiflexion [14]. Dancers are high-performance athletes who
145 are particularly susceptible to lower extremities -specially foot- injuries [15–19]. By
146 demonstrating the reliability of USI of the FHL, this study may offer a novel approach to
147 assess and quantify the injury status of the FHL, as well as providing the image guidance
148 needed for invasive therapies (such as needle insertion) targeting this muscle. The impact of
149 any injury in professional athletes justifies the requirement of improving assessment and
150 treatment techniques and guarantee their reliability [20,21]. Crofts et al. [7] and Mickle et al.
151 [8] placed individuals in a supine position to locate the FHL. However, other studies [6,22],

152 and in particular those targeting invasive therapies such as US-guided percutaneous
153 neuromodulation, have used a prone position for patient placement, with feet unsupported at
154 the end of the plinth. In the present study, prone position was selected, as it has been reported
155 to be more comfortable for the patient, provided greater lower limb stability, and facilitated
156 handling of the probe with one hand and the possible needle with the other [9,10]. Due to
157 these facts, assessing USI reliability in this position was considered to have more implications
158 for potential use in the field of invasive techniques.

159 The authors acknowledge the limitations of this study. Foremost, the small sample size
160 of the study group. The characteristics of the sample, with median age of 21 years, healthy
161 BMI, no previous history of FHL injury and no lower limb impairments, might be not
162 representative of the range of patients with FHL tendinopathy. However this results may be
163 useful in future studies in rehabilitation medicine in other populations. Future lines of
164 research might include repeated assessment of reliability in patients with tendinopathy, as
165 well as in athletes with suspected impairments in FHL function. Further research with a
166 higher sample size is needed.

167 **Conclusions**

168 USI examination of the FHL muscle is highly reliable for the evaluation of the CSA and
169 thickness of the muscle in healthy participants, not depending on the experience of the
170 examiner. Continued assessment of the reliability of US image capture and image-based
171 measurements of FHL morphometric parameters would be important to extend the
172 importance of USI in clinical practice.

173

174

175 **References**

- 176 1. Michelson J, Dunn L. Tenosynovitis of the flexor hallucis longus: A clinical study of
177 the spectrum of presentation and treatment. *Foot Ankle Int* 2005;26(4):291–303.
- 178 2. Khan K, Brown J, Way S, Vass N, Crichton K, Alexander R, et al. Overuse Injuries in
179 Classical Ballet. Vol. 19, *Sports Medicine* 1995. p. 341–57.
- 180 3. Latey PJ, Burns J, Nightingale EJ, Clarke JL, Hiller CE. Reliability and correlates of
181 cross-sectional area of abductor hallucis and the medial belly of the flexor hallucis
182 brevis measured by ultrasound. *J Foot Ankle Res* 2018;11(1).
- 183 4. Angin S, Crofts G, Mickle KJ, Nester CJ. Ultrasound evaluation of foot muscles and
184 plantar fascia in pes planus. *Gait Posture*. 2014;40(1):48–52.
- 185 5. Angin S, Mickle KJ, Nester CJ. Contributions of foot muscles and plantar fascia
186 morphology to foot posture. *Gait Posture* 2018;61:238–42.
- 187 6. De-la-Cruz-Torres B, Barrera-García-Martín I, De la Cueva-Reguera M, Bravo-
188 Aguilar M, Blanco-Morales M, Navarro-Flores E, et al. Does Function Determine the
189 Structure? Changes in Flexor Hallucis Longus Muscle and the Associated
190 Performance Related to Dance Modality: A Cross-Sectional Study. *Medicina*
191 (Kaunas) 2020;56(4).
- 192 7. Mickle KJ, Nester CJ, Crofts G, Steele JR. Reliability of ultrasound to measure
193 morphology of the toe flexor muscles *J Foot Ankle Res*. 2013;6(1).
- 194 8. Crofts G, Angin S, Mickle KJ, Hill S, Nester CJ. Reliability of ultrasound for
195 measurement of selected foot structures *Gait Posture*. 2014;39(1):35–9.
- 196 9. Valera-Calero JA, Sánchez-Jorge S, Álvarez-González J, Ortega-Santiago R, Cleland
197 JA, Fernández-de-Las-Peñas C, et al. Intra-rater and inter-rater reliability of
198 rehabilitative ultrasound imaging of cervical multifidus muscle in healthy people:
199 Imaging capturing and imaging calculation. *Musculoskelet Sci Pract*
200 2020;48:102158.
- 201 10. de-la-Cruz-Torres B, Barrera-García-Martín I, Romero-Morales C. Comparative
202 Effects of One-Shot Electrical Stimulation on Performance of the Flexor Hallucis
203 Longus Muscle in Professional Dancers: Percutaneous Versus Transcutaneous?
204 Neuromodulation. 2019;
- 205 11. de la Cruz-Torres B, Barrera-García-Martín I, Albornoz-Cabello M. Immediate effects
206 of ultrasound-guided percutaneous neuromodulation versus physical exercise on
207 performance of the flexor hallucis longus muscle in professional dancers: a
208 randomised clinical trial *Acupunct Med*. 2019;37(2):91–7.

- 209 12. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation
210 Coefficients for Reliability Research. *J Chiropr Med* 2016;15(2):155–63.
- 211 13. Vosseller JT, Dennis ER, Bronner S. Ankle Injuries in Dancers. *J Am Acad Orthop
212 Surg* 2019;27(16):582–9.
- 213 14. Hamilton WG. Foot and ankle injuries in dancers. *Clin Sports Med* 1988;7(1):143–
214 73.
- 215 15. De-la-Cruz-Torres B, Barrera-García-Martín I, Valera-Garrido F, Minaya-Muñoz F,
216 Romero-Morales C. Ultrasound-Guided Percutaneous Needle Electrolisis in Dancers
217 with Chronic Soleus Injury: A Randomized Clinical Trial. *Evid Based Complement
218 Alternat Med* 2020;2020:4156258.
- 219 16. Bronner S, Novella T, Becica L. Management of a delayed-union sesamoid fracture in
220 a dancer. *J Orthop Sports Phys Ther* 2007;37(9):529–40.
- 221 17. Bronner S, Ojofeitimi S, Rose D. Repair and rehabilitation of extensor hallucis longus
222 and brevis tendon lacerations in a professional dancer. *J Orthop Sports Phys Ther*
223 2008;38(6):362–70.
- 224 18. Filipa A, Barton K. Physical Therapy Rehabilitation of an Adolescent Preprofessional
225 Dancer Following Os Trigonum Excision: A Case Report. *J Orthop Sport Phys Ther*
226 2017;48(3):194–203.
- 227 19. Ojofeitimi S, Bronner S, Becica L. Conservative Management of Second
228 Metatarsophalangeal Joint Instability in a Professional Dancer: A Case Report. *J
229 Orthop Sport Phys Ther* 2016;46(2):114–23.
- 230 20. Hamilton WG. Posterior ankle pain in dancers. *Clin Sports Med* 2008;27(2):263–77.
- 231 21. Luk P, Thordarson D, Charlton T. Evaluation and management of posterior ankle pain
232 in dancers. *J Danc Med Sci Off Publ Int Assoc Danc Med Sci* 2013;17(2):79–83.
- 233 22. de-la-Cruz-Torres B, Barrera-García-Martín I, Cueva-Reguera M de la, Bravo-Aguilar
234 M, Abuin-Porras V, Romero-Morales C. Ultrasound imaging features of the Achilles
235 tendon in dancers. Is there a correlation between the imaging and clinical findings?
236 A cross-sectional study. *Phys Ther Sport Off J Assoc Chart Physiother Sport Med*
237 2020;43:181–7.
- 238
- 239

Table 1. Sociodemographic data

Data	Total (n = 20)
Age (years)	21.05 ± 2.3
Weight (kg)	56.30 ± 5.6
Height (m)	1.65 ± 0.0
BMI (kg/m ²)	20.46 ± 0.6

Abbreviations: body mass index, BMI

Preprint

Table 2. Intra- and inter-rater reliability of imaging for the FHL thickness capturing

	ICC (95% CI)	SD	SEM
FHL Thickness (cm)			
<i>Intra- Rater (A)</i>			
Experienced Rater	0.957 (0.892- 0.983)	0.435	6.8%
Novice Rater	0.969 (0.922- 0.988)	0.432	7.6%
<i>Intra- Rater (B)</i>			
Experienced Rater	0.963 (0.909- 0.985)	0.426	8.1%
Novice Rater	0.976 (0.938- 0.990)	0.426	6.5%
<i>Intra- Rater (C)</i>			
Experienced Rater	0.964 (0.909- 0.986) *	0.463	8.7%
Novice Rater	0.969 (0.922- 0.988)	0.450	7.9%
<i>Intra- Rater (D)</i>			
Experienced Rater	0.978 (0.945- 0.991)	0.450	6.6%
Novice Rater	0.944 (0.856- 0.978)	0.426	10.0%
Inter- Rater (A)	0.991 (0.976- 0.996)	0.418	0.1%
Inter- Rater (B)	0.983 (0.957- 0.993)	0.412	5.3%
Inter- Rater (C)	0.994 (0.986- 0.998)	0.452	0.1%
Inter- Rater (D)	0.991 (0.977- 0.996)	0.435	0.1%

Abbreviations: SEM, standard error measurement; SD, standard deviation

Table 3. Intra- and inter-rater reliability of imaging for the FHL thickness calculation

	ICC (95% CI)	SD	SEM
FHL thickness (cm)			
<i>Intra- Rater (A)</i>			
Experienced Rater	0.968 (0.920- 0.987)	0.455	8.1%
Novice Rater	0.969 (0.922- 0.988)	0.432	7.6%
<i>Intra- Rater (B)</i>			
Experienced Rater	0.963 (0.909- 0.985)	0.432	8.3%
Novice Rater	0.972 (0.930- 0.989)	0.442	7.3%
<i>Intra- Rater (C)</i>			
Experienced Rater	0.954 (0.885- 0.982)	0.450	9.6%
Novice Rater	0.966 (0.913- 0.986)	0.443	8.1%
<i>Intra- Rater (D)</i>			
Experienced Rater	0.973 (0.931- 0.989)	0.454	1.2%
Novice Rater	0.968 (0.920- 0.987)	0.431	7.7%
Inter- Rater (A)	0.958 (0.897- 0.983)	0.430	7.8%
Inter- Rater (B)	0.967 (0.915- 0.987)	0.415	7.5%
Inter- Rater (C)	0.961 (0.902- 0.985)	0.435	8.5%
Inter- Rater (D)	0.980 (0.950- 0.992)	0.444	6.2%

Abbreviations: SEM, standard error measurement; SD, standard deviation

Table 4. Intra- and inter-rater reliability of imaging for the FHL CSA capturing

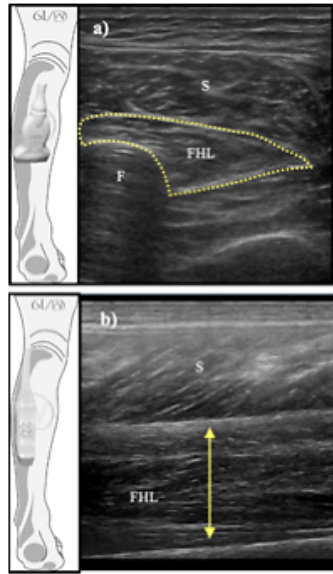
	ICC (95% CI)	SD	SEM
FHL CSA (cm²)			
<i>Intra- Rater (A)</i>			
Experienced Rater	0.967 (0.913- 0.987)	0.733	13.3%
Novice Rater	0.971 (0.928- 0.988)	0.736	12.5%
<i>Intra- Rater (B)</i>			
Experienced Rater	0.979 (0.948- 0.992)	0.749	10.8%
Novice Rater	0.962 (0.904- 0.985)	0.738	14.3%
<i>Intra- Rater (C)</i>			
Experienced Rater	0.967 (0.917- 0.987)	0.694	12.6%
Novice Rater	0.954 (0.882- 0.982)	0.700	16.5%
<i>Intra- Rater (D)</i>			
Experienced Rater	0.970 (0.925- 0.988)	0.744	12.8%
Novice Rater	0.962 (0.907- 0.985)	0.772	15.0%
Inter- Rater (A)	0.984 (0.953- 0.994)	0.710	8.9%
Inter- Rater (B)	0.964 (0.909- 0.986)	0.760	14.4%
Inter- Rater (C)	0.984 (0.960- 0.994)	0.737	9.3%
Inter- Rater (D)	0.965 (0.913- 0.986)	0.732	13.6%

Abbreviations: SEM, standard error measurement; SD, standard deviation

Table 5. Intra- and inter-rater reliability of imaging for the FHL CSA calculation.

	ICC (95% CI)	SD	SEM
FHL CSA (cm²)			
<i>Intra- Rater (A)</i>			
Experienced Rater	0.964 (0.910- 0.986)	0.727	13.7%
Novice Rater	0.964 (0.909- 0.986)	0.723	13.7%
<i>Intra- Rater (B)</i>			
Experienced Rater	0.979 (0.947- 0.991)	0.748	10.8%
Novice Rater	0.972 (0.930- 0.989)	0.757	12.6%
<i>Intra- Rater (C)</i>			
Experienced Rater	0.986 (0.965- 0.994)	0.781	9.2%
Novice Rater	0.961 (0.898- 0.985)	0.714	14.1%
<i>Intra- Rater (D)</i>			
Experienced Rater	0.974 (0.932- 0.990)	0.734	11.8%
Novice Rater	0.966 (0.915- 0.986)	0.769	14.1%
Inter- Rater (A)	0.967 (0.910- 0.988)	0.713	12.9%
Inter- Rater (B)	0.967 (0.918- 0.987)	0.715	12.9%
Inter- Rater (C)	0.991 (0.977- 0.996)	0.735	6.9%
Inter- Rater (D)	0.961 (0.902- 0.985)	0.765	15.1%

Abbreviations: SEM, standard error measurement; SD, standard deviation



Ultrasound imaging of the flexor hallucis longus thickness and CSA

Preprint