



University of Southern Denmark

Ultrasonographic assessment of patellar tendon thickness at 16 clinically relevant measurement sites – a study of intra- and interrater reliability

Holm, Pætur M.; Skou, Søren T.; Olesen, Jens L.; Mølbak, Sissel; Jørgensen, Heidi; Skjoldager, Morten; Christensen, Steffan W.

Published in:
Journal of Bodywork and Movement Therapies

DOI:
10.1016/j.jbmt.2019.02.004

Publication date:
2019

Document version:
Accepted manuscript

Document license:
CC BY-NC-ND

Citation for published version (APA):
Holm, P. M., Skou, S. T., Olesen, J. L., Mølbak, S., Jørgensen, H., Skjoldager, M., & Christensen, S. W. (2019). Ultrasonographic assessment of patellar tendon thickness at 16 clinically relevant measurement sites – a study of intra- and interrater reliability. *Journal of Bodywork and Movement Therapies*, 23(2), 344-351.
<https://doi.org/10.1016/j.jbmt.2019.02.004>

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use

This work is brought to you by the University of Southern Denmark.
Unless otherwise specified it has been shared according to the terms for self-archiving.
If no other license is stated, these terms apply:

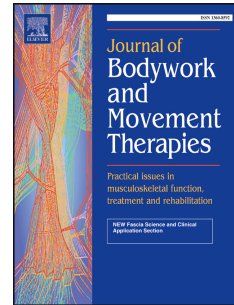
- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.
Please direct all enquiries to puresupport@bib.sdu.dk

Accepted Manuscript

Ultrasonographic assessment of patellar tendon thickness at 16 clinically relevant measurement sites – a study of intra- and interrater reliability

Pætur M. Holm, Søren T. Skou, Jens L. Olesen, Sissel Mølbak, Heidi Jørgensen, Morten Skjoldager, Steffan W. Christensen



PII: S1360-8592(19)30075-0

DOI: <https://doi.org/10.1016/j.jbmt.2019.02.004>

Reference: YJBMT 1771

To appear in: *Journal of Bodywork & Movement Therapies*

Received Date: 29 January 2019

Accepted Date: 2 February 2019

Please cite this article as: Holm, P.M., Skou, S.T., Olesen, J.L., Mølbak, S., Jørgensen, H., Skjoldager, M., Christensen, S.W., Ultrasonographic assessment of patellar tendon thickness at 16 clinically relevant measurement sites – a study of intra- and interrater reliability, *Journal of Bodywork & Movement Therapies*, <https://doi.org/10.1016/j.jbmt.2019.02.004>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title**Ultrasonographic assessment of patellar tendon thickness at 16 clinically relevant measurement sites – a study of intra- and interrater reliability****Authors**

Pætur M. Holm (PMH)^{1,2,*}, Søren T. Skou (STS)^{1,2}, Jens L. Olesen (JLO)³, Sissel Mølbak (SM)⁴, Heidi Jørgensen (HJ)⁴, Morten Skjoldager (MS)⁵, Steffan W. Christensen (SWC)^{4,6}

¹ Department of Physiotherapy and Occupational Therapy, Næstved-Slagelse-Ringsted Hospitals, Denmark

² Research unit for Musculoskeletal Function and Physiotherapy, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark

³ Research Unit for General Practice in Aalborg, Department of Clinical Medicine, Aalborg University, Denmark

⁴ Department of Physiotherapy, University College of Northern Denmark, Aalborg, Denmark

⁵ Arkadens Fysioterapi, Aalborg, Denmark

⁶ SMI, Department of Health Science and Technology, Aalborg University, Denmark

Academic degrees

PMH: PhD fellow, MSc., Physiotherapist

STS: Associate Professor, PhD, MSc., Physiotherapist

JLO: Professor, PhD, MD

SM: Physiotherapist

HJ: Physiotherapist

MS: Physiotherapist

SWC: Assistant Professor, PhD, MPhty., Physiotherapist

*Corresponding author: Pætur M. Holm.

Department of Physiotherapy and Occupational Therapy, Næstved-Slagelse-Ringsted Hospitals, Fælledvej 7, DK-4200 Slagelse. Tel.: +45 61 14 48 33.

E-mail address: pamh@regionsjaelland.dk

Title**Ultrasonographic assessment of patellar tendon thickness at 16 clinically relevant measurement sites – a study of intra- and interrater reliability****Authors**

Pætur M. Holm (PMH)^{1,2,*}, Søren T. Skou (STS)^{1,2}, Jens L. Olesen (JLO)³, Sissel Mølbak (SM)⁴, Heidi Jørgensen (HJ)⁴, Morten Skjoldager (MS)⁵, Steffan W. Christensen (SWC)^{4,6}

¹ Department of Physiotherapy and Occupational Therapy, Næstved-Slagelse-Ringsted Hospitals, Denmark

² Research unit for Musculoskeletal Function and Physiotherapy, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark

³ Research Unit for General Practice in Aalborg, Department of Clinical Medicine, Aalborg University, Denmark

⁴ Department of Physiotherapy, University College of Northern Denmark, Aalborg, Denmark

⁵ Arkadens Fysioterapi, Aalborg, Denmark

⁶ SMI, Department of Health Science and Technology, Aalborg University, Denmark

Academic degrees

PMH: PhD fellow, MSc., Physiotherapist

STS: Associate Professor, PhD, MSc., Physiotherapist

JLO: Professor, PhD, MD

SM: Physiotherapist

HJ: Physiotherapist

MS: Physiotherapist

SWC: Assistant Professor, PhD, MPhty., Physiotherapist

*Corresponding author: Pætur M. Holm.

Department of Physiotherapy and Occupational Therapy, Næstved-Slagelse-Ringsted Hospitals, Fælledvej 7, DK-4200 Slagelse. Tel.: +45 61 14 48 33.

E-mail address: pamh@regionsjaelland.dk

1 **Title**2 **Ultrasonographic assessment of patellar tendon thickness at 16 clinically relevant**
3 **measurement sites – a study of intra- and interrater reliability**4 **Abstract**

5 **Objectives:** To determine intra- and interrater reliability of ultrasonographic imaging (USI) measurements of
6 patellar tendon (PT) thickness using 16 measurement sites covering the entire tendon.

7 **Design:** Reliability study

8 **Setting:** Physiotherapy outpatient clinic

9 **Participants:** Twenty healthy and physically active volunteers (9 women). Mean age: 24 years (SD \pm 2.73).
10 Mean body mass: 75.8 kg (SD \pm 11.8).

11 **Main outcome measures:** Intraclass correlation coefficient (ICC) and 95% limits of agreement (LOA) in cm
12 and in percentage relative to the mean PT thickness.

13 **Results:** Intrarater reliability ranged from 0.59 to 0.87 and 0.59 to 0.93 for examiner I and II, respectively.
14 Interrater reliability ranged from 0.37 to 0.89. Measurement precision for examiner I ranged from 0.05 to
15 0.09 cm (17.5% to 26.7%) while ranging from 0.04 to 0.13 cm (13.3% to 38.7%) for examiner II. Interrater
16 measurement precision ranged from 0.07 to 0.15 cm (19.1% to 42.5%).

17 **Conclusion:** In an attempt to replicate daily clinical USI practice, this was the first study extensively
18 assessing reliability throughout the full range of the patellar tendon - revealing a considerable variation in
19 intra- and interrater reliability as well as measurement precision throughout the 16 individual PT sites. In a
20 clinical context, the low interrater reliability and precision found at the proximal tendon insertion site may
21 have implications for USI of the symptomatic PT, as this is the site mainly associated with underlying
22 pathologic changes. Further reliability studies are needed to clarify the region-specific reliability of the full
23 length PT.

24

25 *Keywords: Knee; patellar tendon; reliability; ultrasonography; musculoskeletal disorders.*

26

INTRODUCTION

27

28 The use of musculoskeletal ultrasound imaging (USI) in the assessment of soft tissue structures is a
29 widely used, inexpensive imaging modality in both research and clinical settings (Finnoff, 2016).

30 With unique dynamic properties and high-resolution imaging of soft tissue structures,
31 musculoskeletal USI has become a valuable tool in the clinical examination of human tendons
32 (Grassi et al., 2000). The patellar tendon (PT) is a body part that has been of particular interest in
33 the literature due to two main reasons. Firstly, due to the prevalence of overuse injuries in the PT
34 which regularly leads to pain and disabling symptoms (Miller, 2013) in both athlete and non-athlete
35 populations (Lian et al., 2005), (Zwerver et al., 2011). Secondly, due to its size, linearity and
36 superficial location above the anterior aspect of the patella, the patellar tendon (PT) is an ideal fit
37 for diagnostic USI (Miller, 2013). Sonographically, the healthy tendon appears hyperechoic due to a
38 strong fibrillar bundle with parallel superficial and deep surfaces, meaning that the healthy parts of
39 the tendon appear bright on a screen as the dense fibrillar bundle reflects a high degree of sound
40 waves from the USI transducer (Miller, 2013).

41 In contrast, the pathologic PT often appears hypoechoic and thickened, meaning that the pathologic
42 part of the tendon appears dark on the screen as it reflects nearly no sound waves from the USI
43 transducer due to loss of the fibrillar pattern and swelling (Miller, 2013), (Kainberger et al., 1997).

44 Since reduction in PT thickness might predict successful treatment outcomes following PT
45 tendinopathy (Fredberg et al., 2004) and since correlations have been found between reduced
46 tendon thickness and a decrease in pain (Mahowald et al., 2011), evaluation of tendon thickness is a
47 particularly important clinical measure.

48 Due to the operator-dependent nature of diagnostic USI (Wakefield et al., 2005), reliability of USI
49 measurements of the PT has been the source of attention in several studies in recent years (Black et

50 al., 2004; Ekizos et al., 2013; Gellhorn and Carlson, 2013; Skou and Aalkjaer, 2013; Sunding et al.,
51 2014). However, previous studies on reliability of PT measurements have focused exclusively on
52 one (Black et al., 2004; Gellhorn and Carlson, 2013; Skou and Aalkjaer, 2013; Sunding et al., 2014)
53 to three (Ekizos et al., 2013) sites, even though full-length PT USI is the common procedure in
54 clinical practice. This leaves not only a big gap in the existing knowledge on reliability of USI
55 measurements when applied on different parts of the PT with different regional characteristics, but
56 also goes against the clinical guidelines for diagnostic USI which recommend scanning the entire
57 PT for pathology (Martinoli, 2010). Thus, keeping in line with clinical recommendations, it is of
58 specific importance to provide USI reliability data covering the various regional aspects of the PT
59 to aid the clinician with region-specific reliability values when scanning the PT for pathology.

60 The aim of this study was to determine intra- and interrater reliability of USI assessment of PT
61 thickness using clinically relevant measurement sites (16 in total) covering the entire PT.

62 MATERIALS & METHODS

63 Participants

64 Twenty healthy and physically active volunteers (9 women) were recruited using flyers posted at
65 University College of Northern Denmark and at a physiotherapy outpatient clinic where the study
66 took place. Mean age was 24 years ($SD \pm 2.73$) with a mean body mass of 74.8 kg ($SD \pm 11.8$).

67 The study was carried out in a single session lasting approximately 90 min per subject (see study
68 protocol for more details).

69 Participants were recruited with exclusion criteria being current or prior lower extremity pain within
70 the past 6 weeks leading up to the test session, previous knee surgery as well as sports activities at
71

72 elite-level. Furthermore, participants were told to refrain from lower limb strength training on the
73 day of testing.

74 This study was based on data collected for a clinical assessment study in a physiotherapy outpatient
75 clinic; hence, approval from the Danish Data Protection Agency was not needed. Written informed
76 consent was obtained from all participants on forms provided by the local ethics committee and the
77 study was conducted in accordance with the Helsinki Declaration.

78 Examiners

79 USI measurements were performed with a SonoSite S-MSK (SonoSite, Inc., Bothell, WA, USA)
80 with an HFL38x 13-6 MHz linear transducer. Two experienced examiners familiar with using the
81 specific USI device in daily clinical practice performed the USI measurements.

82 Both examiners had followed and were teaching on a formal musculoskeletal USI education for
83 doctors, physiotherapists and other healthcare professionals, where one part of the course has a
84 specific focus on the knee. Examiner I had 9 years of experience performing USI measurements
85 while examiner II had 4 years of experience. The examiners were not otherwise involved in
86 collection and synthesis of data.

87 Prior to the start of the study both examiners were carefully instructed in the test-setup and the
88 specific protocol which was based on the previously mentioned education where scanning the entire
89 patella tendon is an inherent part of the education. The examiners had two training sessions each,
90 with one examiner scanning while the other observed the scan.

91 The aim of the test session was to ensure that both examiners adhered to the protocol and were in
92 agreement of where to place cursors in order to perform the measurements. Total time spent on
93 training of the protocol was approximately 8 hours. Throughout each test-session during the study
94 they were under supervision of a research assistant to ensure adherence to the study protocol.

95 Study protocol

96 The USI measurement protocol was set up in accordance with the recommendations from the
97 European Society of Musculoskeletal Radiology (Martinoli, 2010).

98 Subjects were positioned in supine position on an examination table with both knees flexed to
99 approximately 30° knee-flexion supported by a firm cushion underneath the popliteal area. By
100 keeping the target knee in slight flexion (30°) as opposed to fully extended, potential concave
101 anisotropies from the PT were avoided (Martinoli, 2010).

102 Two successive bilateral measurements were performed in random order on each subject by both
103 examiners resulting in one of the following orders: (right-left-right-left) or (left-right-left-right).

104 This order of measurements would force the examiner to reposition the transducer between the first
105 and second unilateral measurement. By placing tape over the measurement values on the screen, the
106 examiners were blinded to the results while these were still visible for the research assistant who
107 recorded each value. Four longitudinal clinically relevant PT sites were identified for analysis; apex
108 patella, 1 cm. under apex patella, tibial tuberosity and 1 cm. over tibial tuberosity.

109 >>>TABLE 1 HERE<<<

110 >>>FIGURE 1 HERE<<<

111 At each site, one longitudinal (central placement) and three transversal measurements were
112 performed, giving a total of 16 individual sites (table 1 & figure 1). The scan depth was kept
113 constant at 1.8 cm and the measurement of PT thickness was performed by the built-in software
114 (figure 2, 3 & 4). Each test session lasted approximately 45 minutes and was performed
115 simultaneously on two subjects by the two examiners in different rooms.

116 After 45 minutes, the two examiners switched rooms and thereby subjects. The straight succession
117 of each examiners measurements on the same subject ensured that no activity-induced effects on
118 tendon thickness was present between measurements.

119 >>>FIGURE 2 HERE<<<

120 >>>FIGURE 3 HERE<<<

121 >>>FIGURE 4 HERE<<<

122 Statistical analysis

123 All analyses were performed using IBM SPSS Statistics Version 24.0. Only the participant's
124 dominant leg was chosen for statistical analysis. Data on non-dominant leg are available as a
125 supplementary appendix (appendix A). Visual inspection of QQ-plots was conducted to ensure
126 normal distribution of data. Changes in PT thickness between the first and second measurement was
127 investigated using a Paired samples *t* test. To investigate potential learning effects over the range of
128 measurements of the 20 subjects, the randomly generated blocks of participants received
129 consecutive identification numbers.

130 Evaluation of intrarater reliability for examiner I and II was conducted via two-way random effects,
131 single measure model (2,1), absolute agreement type intraclass correlation coefficient (ICC).

132 Interrater reliability using the mean of first and second measurements was analyzed via two-way
133 random effects, average measure model (2, *k*), absolute agreement type ICC.

134 Since mean ratings have shown to improve measurement precision compared to single ratings
135 (Skou and Aalkjaer, 2013), (Rathleff et al., 2011), only means from examiner I and II, respectively
136 were derived for interrater reliability analysis. A Bland-Altman plot was constructed to graphically
137 assess agreement between the two examiners. Measurement precision was evaluated by plotting

138 1.96 standard deviations (SD) above and below the mean of the difference scores, respectively. The
139 confidence interval ranging from 1.96 SD above the mean of the difference scores to 1.96 SD below
140 represents 95% limits of agreement (LOA). LOA was also presented in percentage relative to the
141 mean PT thickness (LOA-%).

142 RESULTS

143 Results for PT thickness measurements for examiner I and II are found in table 2 and 3,
144 respectively.

145 For examiner II, the measurement sites 3b ($p=0.025$) and 4a ($p=0.033$) were significantly different
146 between the two measurements. Hence, these measurement sites were excluded from further
147 reliability testing for examiner II. There was no consistent evidence of learning effects over the
148 range of the 16 measurement sites when plotting the difference in PT thickness between each
149 examiners first and second measurements against the consecutive identification number of subjects.

150 >>>TABLE 2 HERE<<<

151 >>>TABLE 3 HERE<<<

152 Reliability

153 Results for intra- and interrater reliability for examiner I and II are presented in table 4.

154 For illustrative purposes, the ICC results are categorized as low (0.0 – 0.50), moderate (0.50 – 0.75)
155 and good (>0.75) (Portney and Watkins, 2014). However, we are well aware of the difficulties in
156 determining appropriate quality cut-offs for reliability (Weir, 2005). As such, the current
157 categorization should be viewed as crude estimates only.

158 The ICC with 95% confidence interval (95% CI) for examiner I ranged from moderate to good
159 (0.59 to 0.87) for all 16 measurements. The ICC with 95% CI for examiner II ranged from moderate

160 to good (0.59 to 0.93) for 14 out of 16 measurements with two measurement sites (3b & 4a)
161 excluded due to a significant difference between the first and second measurement. The interrater
162 ICC for examiner I and II ranged from low to good (0.37 to 0.89) for all 14 interrater reliability
163 tested measurements.

164 >>>TABLE 4 HERE<<<

165 Measurement precision

166 Results for intra- and interrater LOA and LOA-% are presented in table 5.

167 Visual inspection of the intrarater Bland Altman plots for both examiners revealed no proportional
168 differences over the measurement range. For examiner I, the LOA (LOA-%) ranged from 0.05 to
169 0.09 cm (17.5% to 26.7%) for intrarater reliability over the range of the 16 PT measurement sites.
170 For examiner II, the LOA (LOA %) for intrarater reliability ranged from 0.04 to 0.13 cm. (13.3% -
171 38.7%) over the range of the 14 PT measurement sites. LOA (LOA-%) for interrater reliability
172 ranged from 0.07 to 0.15 cm (21% - 42.5%) over the range of 14 PT measurement sites.

173 >>>TABLE 5 HERE<<<

174 Visual inspection of the interrater Bland Altman plots suggested a proportional difference of
175 measurements for “AP” and “1uAP”, however when fitted into a linear regression analysis, normal
176 distribution of the difference scores was confirmed ($p>0.05$) (figure 5). The remaining Bland
177 Altman plots revealed no systematic graphical differences, represented by figure 6.

178 >>>FIGURE 5 HERE<<<

179 >>>FIGURE 6 HERE<<<

180 A summary of results for measurement reliability and precision is presented in table 6.

181 >>>TABLE 6 HERE<<<

182

DISCUSSION

183 This is the first study evaluating intra- and interrater reliability and measurement precision when
184 measuring PT thickness of the entire tendon using USI. Intrarater reliability ranged from 0.59 to
185 0.93 for both examiners over the range of 14 and 16 sites, respectively. Results for interrater
186 reliability were within a wider range (0.37 to 0.89) for the 14 assessed sites. Precision for intrarater
187 measurements varied from 0.04 cm to 0.13 cm (13.3% to 38.7%) while ranging from 0.06 cm to
188 0.15 cm (19.1% to 42.5%) for interrater measurements.

189 Previous intrarater- and interrater USI studies on muscle- and tendon thickness reveal a cumulative
190 ICC range from 0.64 to 0.97 (Bentman et al., 2010; Cheng et al., 2012; Costa et al., 2009; Craig et
191 al., 2008; Gellhorn and Carlson, 2013; Koppenhaver et al., 2009; Liang et al., 2007; O'Sullivan et
192 al., 2007; Rathleff et al., 2011; Skou and Aalkjaer, 2013; Wallwork et al., 2007) and from 0.40 to
193 0.97, respectively (Bentman et al., 2010; Cheng et al., 2012; Gellhorn and Carlson, 2013;
194 O'Sullivan et al., 2007; Rathleff et al., 2011; Skou and Aalkjaer, 2013; Wallwork et al., 2007).

195 Previous results on measurement precision (LOA-%) reveal a cumulative range from 1.8% to 53%
196 for intrarater (Bentman et al., 2010; Bjordal et al., 2003; Costa et al., 2009; Koppenhaver et al.,
197 2009; O'Connor et al., 2004; O'Sullivan et al., 2007; Rathleff et al., 2011; Skou and Aalkjaer, 2013;
198 Springer et al., 2006; Wallwork et al., 2007; Ying et al., 2003) and 15.8% to 49% for interrater
199 (Bentman et al., 2010; O'Sullivan et al., 2007; Rathleff et al., 2011; Skou and Aalkjaer, 2013;
200 Wallwork et al., 2007; Ying et al., 2003) for USI-derived measures of muscle- and tendon
201 thickness. The considerable variation in reliability and measurement precision in USI studies is in
202 part reflective of the different structures being measured, with proximity to bone, depth and
203 adjacent soft tissue, to a varying degree, influencing the quality of the sonographic image. When
204 compared to deeper and irregular soft-tissue structures, the PT is considered relatively feasible for

205 USI examinations (Henderson et al., 2015). Yet, the wide reliability- and precision range found in
206 our study indicates some degree of uncertainty.

207 In a study on PT measurement precision in healthy adults, O'Connor et al. measured LOA-% 1 cm
208 above the tendons insertion onto the tibia tubercle corresponding to "1 o. TT" in our study and
209 found an intrarater LOA-% of 19% and 22% for measures of long transverse axis width and 24%
210 and 32% for short axis transverse diameter. Interrater LOA-% was 22% and 27% for long axis and
211 short axis, respectively (O'Connor et al., 2004). The present findings on transverse axis PT
212 thickness at this measurement point revealed a similar trend with intrarater LOA-% values from
213 20.2% to 23.9% (examiner I) and from 14% to 22.5% (examiner II) and interrater LOA-% from
214 21.5% to 22.1%. When studying intra- and interrater reliability of PT measurements, Gellhorn et al.
215 found an intrarater reliability from 0.87 to 0.96 and interrater reliability of 0.90 and 0.92, when
216 measuring the cross-sectional area of the tendon 1 cm distal to apex patella (Gellhorn and Carlson,
217 2013). In our study, the longitudinal measurements 1 cm distal to apex patella (1 u. AP) revealed
218 considerably lower intra- and interrater reliability of 0.59 to 0.84 and 0.50, respectively. The highly
219 standardized measurement protocol including strapwires as external markers in the Gellhorn study
220 might partially explain the higher reliability values found in this study. However, such extensive
221 measures of standardization might not adequately reflect the use of USI in daily clinical practice.
222 Still, using a similar examination protocol as in the current study, Skou et al. found an intrarater
223 reliability of 0.89-0.94 and interrater reliability of 0.78 with intrarater LOA of 0.07 cm. and
224 interrater LOA of 0.10 cm. for two examiners measuring PT thickness in a longitudinal plane 1 cm
225 distal to apex patella (Skou and Aalkjaer, 2013). The findings in the Skou study reveal somewhat
226 higher ICC reliability scores for intrarater (0.89-0.94 vs. 0.59-0.84) and interrater (0.78 vs. 0.50)
227 when compared to our findings. This might indicate that the task of identifying several

228 measurement sites in one scanning session may compromise specificity at each individual site and
229 thereby reduce the reliability of measurements.

230 When comparing longitudinal and transverse plane findings at the same site in our study, ICC and
231 LOA results are largely consistent (Table 6). However, there is a notable exception at the
232 measurement site 1 cm distal to Apex Patella (1uAP), with this site showing a considerably higher
233 interrater ICC for the transverse plane measurements (0.87 vs. 0.50).

234 As such, transverse axis scans at 1uAP seems to provide more consistent PT thickness
235 measurements in our study. Interestingly, since the proximal PT region is aligned in a cone shaped
236 structure, transversal thickness measurements without a metric tendon reference point at the
237 transversal axis, naturally becomes more difficult for the examiner to reproduce than longitudinal
238 measurements, where thickness can be measured with the aid of metric reference points.

239 Coincidentally, higher measurement variability has been found for transverse axis PT thickness
240 scans (Fredberg et al., 2008) - making the present findings somewhat surprising.

241 As mentioned earlier, this is the first USI reliability study using measurement sites covering the
242 entire PT, making the results more clinically relevant. Previously, when studying the reliability of
243 more than one PT measurement site, Ekizos et al. found an average interrater reliability of 0.59,
244 combining findings on PT cross-sectional area from three examiners on the proximal, and distal PT
245 borders as well as the metric midpoint (Ekizos et al., 2013).

246 In particular, measurements at the proximal portion of the tendon showed the highest variability in
247 the study by Ekizos and colleagues (root mean square range from $7.9 \pm 3.9 \text{ mm}^2$ to $16.1 \pm 11.3 \text{ mm}^2$).
248 In our study, reliability values at the proximal border (AP, 1a, 1b & 1c) revealed a mean interrater
249 reliability for both measurement planes ranging from 0.37 to 0.64 with an LOA range of 0.14 - 0.15
250 cm. (37.7% to 40.4%) (Table 6). In comparison, the corresponding measurements at the distal

251 border (TT, 3a & 3c) revealed an interrater reliability ranging from 0.64 to 0.71 with LOA ranging
252 from 0.10 to 0.11 (25.9% to 27.5%). However, the significant difference between first and second
253 measurement for 3b indicates a certain degree of measurement variability at the distal border as
254 well and may point to difficulties in reproducing measurements near the tendon insertion sites.

255 Consequently, the frequently reported pathological findings at the posterior, proximal aspect of the
256 PT (Helland et al., 2013; Johnson et al., 1996; Khan et al., 1996) might be limited by the reliability
257 of measurements. One possible contributing factor to the limited interrater reliability at various PT
258 points in our study might be that the mean was derived from two measurements as opposed to three.
259 Previously, a mean consisting of three measurements as opposed to two has resulted in improved
260 intra- and interrater reliability as well as measurement precision when applied in the evaluation of
261 plantar fascia (Rathleff et al., 2011) and transversus abdominis and lumbar multifidi muscles
262 (Koppenhaver et al., 2009). However, Skou et al. found no further improvement on intra- and
263 interrater reliability or measurement precision when using a mean of three measurements compared
264 to two in the evaluation of PT thickness (Skou and Aalkjaer, 2013). Yet, regional PT characteristics
265 might influence the reliability of the mean of measurements to a varying degree. This might
266 especially hold true for measurements at the proximal PT region. As such, inclusion of means of
267 two and three measurements in this study would have clarified this relationship.

268 Both examiners were experienced USI examiners and produced largely comparable intrarater
269 reliability scores over the range of measurements. However, examiner II had two sites (3b & 4a)
270 with statistically significant different PT thickness at the first compared to the second measurement.
271 This might be suggestive of a need for a more standardized examination protocol and should be
272 considered in future studies.

273 Put into a clinical context, this study provides region-specific reliability data covering the entire PT.
274 This is important information for the clinician when assessing the different parts of the PT for
275 pathology.

276 The limited agreement between examiners when assessing the proximal portion of the PT suggests
277 some degree of uncertainty and should be taken into consideration when assessing the frequently
278 found pathologic changes in this region. Future studies should explore the reliability of
279 measurements retrieved at the proximal portion of PT in patients with specific overuse injuries as
280 this might be of clinical relevance when considering treatment effects.

281 It is important to note that the current study is limited by the fact that only healthy tendons were
282 studied, hence the findings cannot be extrapolated fully to pathological or degenerated tendons.
283 However, since pathological and degenerated tendons often display a heterogeneous tendon
284 structure with blurred tendon margins (Grassi et al., 2000), reliability may be even more
285 compromised. Further studies testing the protocol used in this study on pathological and
286 degenerated PT are warranted and will aid with important clinical insights into potential reliability
287 issues.

288 In conclusion, USI reliability assessment of PT thickness, using 16 measurement points covering
289 the entire PT revealed contrasting degrees of reliability and measurement precision. Especially
290 interrater reliability and agreement fluctuated throughout the range of measurements.

291 Further reliability studies on the different aspects of PT are needed to clarify the uniformity of
292 region-specific PT examinations using USI.

293 Ethical approval

294 Obtainment of informed and written consent from all participants on forms provided by the local
295 ethical committee was sufficient for the scope of this study.

296 Funding

297 This research did not receive any specific grant from funding agencies in the public, commercial, or
298 not-for-profit sectors.

299 Conflict of interests

300 The authors have no conflict of interests

ACCEPTED MANUSCRIPT

REFERENCES

- Bentman, S., O'Sullivan, C., and Stokes, M. (2010). Thickness of the middle trapezius muscle measured by rehabilitative ultrasound imaging: description of the technique and reliability study. *Clin. Physiol. Funct. Imaging* 30, 426–431.
- Bjordal, J.M., Demmink, J.H., and Ljunggren, A.E. (2003). Tendon Thickness and Depth from Skin for Supraspinatus, Common Wrist and Finger Extensors, Patellar and Achilles Tendons. *Physiotherapy* 89, 375–383.
- Black, J., Cook, J., Kiss, Z.S., and Smith, M. (2004). Intertester reliability of sonography in patellar tendinopathy. *J. Ultrasound Med. Off. J. Am. Inst. Ultrasound Med.* 23, 671–675.
- Cheng, J.-W., Tsai, W.-C., Yu, T.-Y., and Huang, K.-Y. (2012). Reproducibility of sonographic measurement of thickness and echogenicity of the plantar fascia. *J. Clin. Ultrasound JCU* 40, 14–19.
- Costa, L.O.P., Maher, C.G., Latimer, J., Hodges, P.W., and Shirley, D. (2009). An investigation of the reproducibility of ultrasound measures of abdominal muscle activation in patients with chronic non-specific low back pain. *Eur. Spine J. Off. Publ. Eur. Spine Soc. Eur. Spinal Deform. Soc. Eur. Sect. Cerv. Spine Res. Soc.* 18, 1059–1065.
- Craig, M.E., Duffin, A.C., Gallego, P.H., Lam, A., Cusumano, J., Hing, S., and Donaghue, K.C. (2008). Plantar fascia thickness, a measure of tissue glycation, predicts the development of complications in adolescents with type 1 diabetes. *Diabetes Care* 31, 1201–1206.
- Ekizos, A., Papatzika, F., Charcharis, G., Bohm, S., Mersmann, F., and Arampatzis, A. (2013). Ultrasound does not provide reliable results for the measurement of the patellar tendon cross sectional area. *J. Electromyogr. Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.* 23, 1278–1282.
- Finnoff, J.T. (2016). The Evolution of Diagnostic and Interventional Ultrasound in Sports Medicine. *PM&R* 8, S133–S138.
- Fredberg, U., Bolvig, L., Pfeiffer-Jensen, M., Clemmensen, D., Jakobsen, B.W., and Stengaard-Pedersen, K. (2004). Ultrasonography as a tool for diagnosis, guidance of local steroid injection and, together with pressure algometry, monitoring of the treatment of athletes with chronic jumper's knee and Achilles tendinitis: a randomized, double-blind, placebo-controlled study. *Scand. J. Rheumatol.* 33, 94–101.
- Fredberg, U., Bolvig, L., Andersen, N.T., and Stengaard-Pedersen, K. (2008). Ultrasonography in evaluation of Achilles and patella tendon thickness. *Ultraschall Med. Stuttg. Ger.* 1980 29, 60–65.
- Gellhorn, A.C., and Carlson, M.J. (2013). Inter-rater, intra-rater, and inter-machine reliability of quantitative ultrasound measurements of the patellar tendon. *Ultrasound Med. Biol.* 39, 791–796.
- Grassi, W., Filippucci, E., Farina, A., and Cervini, C. (2000). Sonographic imaging of tendons. *Arthritis Rheum.* 43, 969–976.

- Helland, C., Bojsen-Møller, J., Raastad, T., Seynnes, O.R., Moltubakk, M.M., Jakobsen, V., Visnes, H., and Bahr, R. (2013). Mechanical properties of the patellar tendon in elite volleyball players with and without patellar tendinopathy. *Br. J. Sports Med.* *47*, 862–868.
- Henderson, R.E.A., Walker, B.F., and Young, K.J. (2015). The accuracy of diagnostic ultrasound imaging for musculoskeletal soft tissue pathology of the extremities: a comprehensive review of the literature. *Chiropr. Man. Ther.* *23*, 31.
- Johnson, D.P., Wakeley, C.J., and Watt, I. (1996). Magnetic resonance imaging of patellar tendonitis. *J. Bone Joint Surg. Br.* *78*, 452–457.
- Kainberger, F., Mittermaier, F., Seidl, G., Parth, E., and Weinstabl, R. (1997). Imaging of tendons--adaptation, degeneration, rupture. *Eur. J. Radiol.* *25*, 209–222.
- Khan, K.M., Bonar, F., Desmond, P.M., Cook, J.L., Young, D.A., Visentini, P.J., Fehrmann, M.W., Kiss, Z.S., O'Brien, P.A., Harcourt, P.R., et al. (1996). Patellar tendinosis (jumper's knee): findings at histopathologic examination, US, and MR imaging. Victorian Institute of Sport Tendon Study Group. *Radiology* *200*, 821–827.
- Koppenhaver, S.L., Parent, E.C., Teyhen, D.S., Hebert, J.J., and Fritz, J.M. (2009). The effect of averaging multiple trials on measurement error during ultrasound imaging of transversus abdominis and lumbar multifidus muscles in individuals with low back pain. *J. Orthop. Sports Phys. Ther.* *39*, 604–611.
- Lian, O.B., Engebretsen, L., and Bahr, R. (2005). Prevalence of jumper's knee among elite athletes from different sports: a cross-sectional study. *Am. J. Sports Med.* *33*, 561–567.
- Liang, H.-W., Wang, T.-G., Chen, W.-S., and Hou, S.-M. (2007). Thinner plantar fascia predicts decreased pain after extracorporeal shock wave therapy. *Clin. Orthop.* *460*, 219–225.
- Mahowald, S., Legge, B.S., and Grady, J.F. (2011). The correlation between plantar fascia thickness and symptoms of plantar fasciitis. *J. Am. Podiatr. Med. Assoc.* *101*, 385–389.
- Martinoli, C. (2010). Musculoskeletal ultrasound: technical guidelines. *Insights Imaging* *1*, 99–141.
- Miller, T.T. (2013). The patellar tendon. *Semin. Musculoskelet. Radiol.* *17*, 56–59.
- O'Connor, P.J., Grainger, A.J., Morgan, S.R., Smith, K.L., Waterton, J.C., and Nash, A.F.P. (2004). Ultrasound assessment of tendons in asymptomatic volunteers: a study of reproducibility. *Eur. Radiol.* *14*, 1968–1973.
- O'Sullivan, C., Bentman, S., Bennett, K., and Stokes, M. (2007). Rehabilitative ultrasound imaging of the lower trapezius muscle: technical description and reliability. *J. Orthop. Sports Phys. Ther.* *37*, 620–626.
- Portney, L.G., and Watkins, M.P. (2014). *Foundations of clinical research: applications to practice* (Harlow: Pearson).
- Rathleff, M.S., Moelgaard, C., and Lykkegaard Olesen, J. (2011). Intra- and interobserver reliability of quantitative ultrasound measurement of the plantar fascia. *J. Clin. Ultrasound JCU* *39*, 128–134.

Skou, S.T., and Aalkjaer, J.M. (2013). Ultrasonographic measurement of patellar tendon thickness-- a study of intra- and interobserver reliability. *Clin. Imaging* 37, 934–937.

Springer, B.A., Mielcarek, B.J., Nesfield, T.K., and Teyhen, D.S. (2006). Relationships among lateral abdominal muscles, gender, body mass index, and hand dominance. *J. Orthop. Sports Phys. Ther.* 36, 289–297.

Sunding, K., Fahlström, M., Werner, S., Forssblad, M., and Willberg, L. (2014). Evaluation of Achilles and patellar tendinopathy with greyscale ultrasound and colour Doppler: using a four-grade scale. *Knee Surg. Sports Traumatol. Arthrosc. Off. J. ESSKA.*

Wakefield, R.J., Balint, P.V., Szkudlarek, M., Filippucci, E., Backhaus, M., D'Agostino, M.-A., Sanchez, E.N., Iagnocco, A., Schmidt, W.A., Bruyn, G.A.W., et al. (2005). Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J. Rheumatol.* 32, 2485–2487.

Wallwork, T.L., Hides, J.A., and Stanton, W.R. (2007). Intrarater and interrater reliability of assessment of lumbar multifidus muscle thickness using rehabilitative ultrasound imaging. *J. Orthop. Sports Phys. Ther.* 37, 608–612.

Weir, J.P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J. Strength Cond. Res. Natl. Strength Cond. Assoc.* 19, 231–240.

Ying, M., Yeung, E., Li, B., Li, W., Lui, M., and Tsoi, C.-W. (2003). Sonographic evaluation of the size of Achilles tendon: the effect of exercise and dominance of the ankle. *Ultrasound Med. Biol.* 29, 637–642.

Zwerver, J., Bredeweg, S.W., and van den Akker-Scheek, I. (2011). Prevalence of Jumper's knee among nonelite athletes from different sports: a cross-sectional survey. *Am. J. Sports Med.* 39, 1984–1988.

TABLES & FIGURESTABLE 1**Table 1** | Measurement sites (n=16) of the patellar tendon.

Site	Longitudinal	Transversal
AP	AP	1a: Thickest middle third
		1b: Thickest medial third
		1c: Thickest lateral third
1 u. AP	1 u. AP	2a: Thickest middle third
		2b: Thickest medial third
		2c: Thickest lateral third
TT	TT	3a: Thickest middle third
		3b: Thickest medial third
		3c: Thickest lateral third
1 o. TT	1 o. TT	4a: Thickest middle third
		4b: Thickest medial third
		4c: Thickest lateral third

Thickness (cm) was measured anteroposterior.

AP: Apex Patella. **1 u. AP:** 1 cm. under Apex Patella.

TT: Tibial Tuberosity. **1 o. TT:** 1 cm over Tibial Tuberosity.

TABLE 2

Table 2 | Patellar tendon thickness (cm) on dominant leg from all 16 measurement sites for examiner I. Results presented in mean \pm SD.

Sites (n=16)	Men (n=11) mean (SD)	Women (n=9) mean (SD)	Total (n=20) mean (SD)
AP	0.47 (0.08)	0.40 (0.06)	0.44 (0.08)
1a	0.33 (0.05)	0.30 (0.06)	0.32 (0.06)
1b	0.35 (0.06)	0.32 (0.04)	0.34 (0.05)
1c	0.32 (0.04)	0.29 (0.03)	0.31 (0.04)
1 u. AP	0.39 (0.07)	0.34 (0.05)	0.37 (0.06)
2a	0.33 (0.06)	0.29 (0.05)	0.31 (0.06)
2b	0.33 (0.05)	0.30 (0.03)	0.32 (0.05)
2c	0.33 (0.04)	0.29 (0.04)	0.31 (0.04)
TT	0.45 (0.07)	0.42 (0.07)	0.44 (0.07)
3a	0.34 (0.06)	0.30 (0.05)	0.32 (0.06)
3b	0.34 (0.05)	0.32 (0.05)	0.33 (0.05)
3c	0.34 (0.03)	0.31 (0.06)	0.33 (0.05)
1 o. TT	0.37 (0.06)	0.31 (0.05)	0.34 (0.06)
4a	0.33 (0.05)	0.30 (0.05)	0.31 (0.05)
4b	0.33 (0.04)	0.31 (0.04)	0.32 (0.04)
4c	0.32 (0.03)	0.29 (0.04)	0.31 (0.04)

TABLE 3

Table 3 | Patellar tendon thickness (cm) on dominant leg from all 16 measurement sites for examiner II. Results presented in mean \pm SD.

Sites (n=16)	Men (n=11) mean (SD)	Women (n=9) mean (SD)	Total (n=20) mean (SD)
AP	0.32 (0.06)	0.35 (0.08)	0.33 (0.07)
1a	0.37 (0.09)	0.29 (0.04)	0.33 (0.08)
1b	0.38 (0.10)	0.30 (0.04)	0.35 (0.09)
1c	0.35 (0.09)	0.28 (0.04)	0.32 (0.08)
1 u. AP	0.30 (0.05)	0.31 (0.04)	0.30 (0.05)
2a	0.33 (0.06)	0.28 (0.03)	0.31 (0.06)
2b	0.33 (0.06)	0.28 (0.03)	0.31 (0.06)
2c	0.35 (0.05)	0.27 (0.03)	0.31 (0.06)
TT	0.42 (0.06)	0.38 (0.07)	0.40 (0.06)
3a	0.38 (0.06)	0.33 (0.05)	0.36 (0.06)
3b	0.36 (0.05)	0.31 (0.05)	0.34 (0.06)*
3c	0.39 (0.06)	0.32 (0.04)	0.36 (0.06)
1 o. TT	0.36 (0.05)	0.32 (0.05)	0.34 (0.05)
4a	0.34 (0.05)	0.29 (0.05)	0.31 (0.05)*
4b	0.33 (0.04)	0.29 (0.04)	0.31 (0.05)
4c	0.35 (0.05)	0.28 (0.05)	0.32 (0.06)

*Significant different patellar tendon thickness between first and second measurements ($p < 0.05$).

TABLE 4

Table 4 | Intra- & Interrater reliability results for both examiners. ICC, (95% CI).

	Examiner I	Examiner II	Examiner I + II
AP	▲ 0.86 (0.68 - 0.94)	▲ 0.81 (0.59 - 0.92)	▼ 0.37 (-0.11 - 0.58)
1a	▲ 0.80 (0.57 - 0.92)	▲ 0.79 (0.55 - 0.91)	▶ 0.66 (0.15 - 0.86)
1b	▲ 0.85 (0.57 - 0.92)	▶ 0.72 (0.42 - 0.88)	▶ 0.62 (0.02 - 0.85)
1c	▶ 0.75 (0.47 - 0.89)	▶ 0.71 (0.40 - 0.88)	▶ 0.65 (0.09 - 0.86)
1 u. AP	▲ 0.84 (0.65 - 0.93)	▶ 0.59 (0.21 - 0.81)	▶ 0.50 (-0.25 - 0.82)
2a	▲ 0.83 (0.62 - 0.93)	▲ 0.82 (0.61 - 0.93)	▲ 0.89 (0.73 - 0.96)
2b	▲ 0.77 (0.51 - 0.90)	▲ 0.93 (0.84 - 0.97)	▲ 0.87 (0.67 - 0.95)
2c	▶ 0.66 (0.32 - 0.85)	▲ 0.83 (0.62 - 0.93)	▲ 0.86 (0.65 - 0.94)
TT	▶ 0.75 (0.47 - 0.89)	▲ 0.88 (0.72 - 0.95)	▶ 0.71 (0.24 - 0.89)
3a	▲ 0.87 (0.71 - 0.95)	▲ 0.86 (0.68 - 0.94)	▲ 0.79 (0.09 - 0.94)
3b	▶ 0.68 (0.36 - 0.86)	*	*
3c	▶ 0.59 (0.21 - 0.82)	▲ 0.84 (0.64 - 0.93)	▼ 0.48 (-0.15 - 0.78)
1 o. TT	▲ 0.84 (0.65 - 0.94)	▲ 0.78 (0.53 - 0.91)	▲ 0.86 (0.65 - 0.95)
4a	▶ 0.72 (0.43 - 0.88)	*	*
4b	▶ 0.72 (0.42 - 0.88)	▲ 0.89 (0.74 - 0.95)	▲ 0.80 (0.50 - 0.92)
4c	▶ 0.71 (0.41 - 0.87)	▲ 0.81 (0.58 - 0.92)	▲ 0.84 (0.61 - 0.94)

▲ (ICC > 0.75) = good reliability

▶ (ICC 0.50 – 0.75) = moderate reliability

▼ (ICC 0.0 – 0.50) = low reliability

*Excluded from the reliability analysis due to significant difference in PT thickness between first and second measurement.

TABLE 5

Table 5 | Intra- & interrater reliability presented as 95% LOA (cm) and as percentage of the mean PT thickness (%).

	Examiner I		Examiner II		Examiner I + II	
	LOA (cm)	LOA (%)	LOA (cm)	LOA (%)	LOA (cm)	LOA (%)
AP	0.09	19.5	0.09	26.1	0.15	37.7
1a	0.07	22.1	0.11	32.6	0.14	41.9
1b	0.05	16.1	0.13	38.7	0.15	42.5
1c	0.05	17.6	0.12	37.4	0.12	36.9
1 u. AP	0.07	19.2	0.09	28.8	0.10	28.6
2a	0.07	21.0	0.07	22.3	0.07	21.7
2b	0.07	20.7	0.04	13.3	0.07	21.9
2c	0.07	22.5	0.07	21.2	0.07	21.3
TT	0.09	21.4	0.06	15.5	0.11	25.9
3a	0.06	17.5	0.06	17.8	0.07	21.0
3b	0.08	23.4	*	*	*	*
3c	0.09	26.7	0.07	20.2	0.12	33.9
1 o. TT	0.07	20.2	0.07	21.1	0.08	22.5
4a	0.08	23.9	*	*	*	*
4b	0.07	20.5	0.04	14.0	0.07	21.5
4c	0.06	20.2	0.07	22.5	0.07	22.1

*Excluded from measurement precision analysis due to significant difference in PT thickness between first and second measurements.

TABLE 6

Table 6 | Intra- and interrater reliability and precision of the four longitudinal plane measurement sites.

	Longitudinal			Transversal		
	EX I	EX II	EX I + II	EX I	EX II	EX I + II
AP						
ICC	0.86	0.81	0.37	0.80	0.74	0.64
LOA	0.09	0.09	0.15	0.06	0.12	0.14
LOA-%	19.5	26.1	37.7	18.6	36.2	40.4
IuAP						
ICC	0.84	0.59	0.50	0.75	0.86	0.87
LOA	0.07	0.09	0.10	0.07	0.06	0.07
LOA-%	19.2	28.8	28.6	21.4	18.9	21.6
TT						
ICC	0.75	0.88	0.71	0.71	0.79*	0.64*
LOA	0.09	0.06	0.11	0.08	0.07*	0.10*
LOA-%	21.4	15.5	25.9	22.5	19.0*	27.5*
IoTT						
ICC	0.84	0.78	0.86	0.72	0.85*	0.82*
LOA	0.07	0.07	0.08	0.07	0.06*	0.07*
LOA-%	20.2	21.1	22.5	21.5	18.3*	21.8*

Results are divided between longitudinal axis scans and transverse axis scans. Transverse axis results are presented as the mean of the transverse measurements: a (middle third), b (medial third) and c (lateral third) at each of the four sites. EX: Examiner.

*Since there was a significant difference in PT thickness between first and second measurement on 3b and 4a for examiner II, only two transversal sites were used for the derivation of means at these sites.

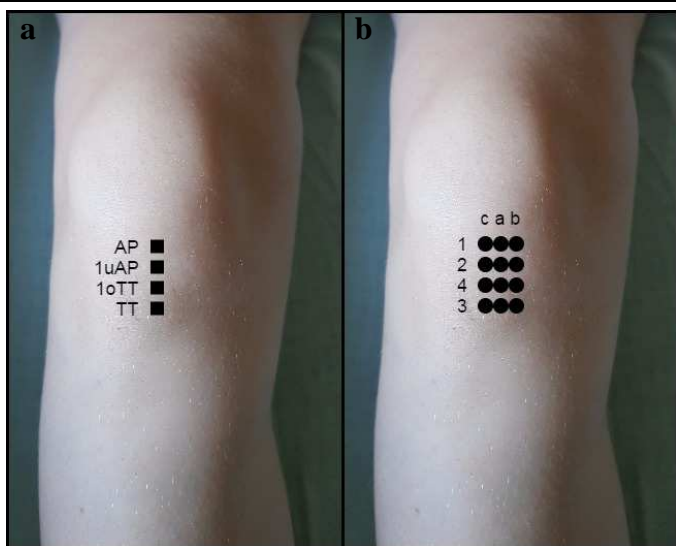
FIGURE 1**Figure 1** | Images of the knee with longitudinal (a) and transversal (b) measurement sites depicted.

Image 1a: Longitudinal plane thickness measurements (■) were conducted at Apex Patella (AP), 1 cm. distal from Apex Patella (1uAP), 1 cm. proximal from the Tibial Tuberosity (1oTT) and at the Tibial tuberosity (TT).

Image 1b: Transversal plane thickness measurements (●) were conducted at the thickest part of the middle- (a), medial- (b) and lateral (c) third of the tendon at; AP (1), 1uAP (2), TT (3) and 1oTT (4).

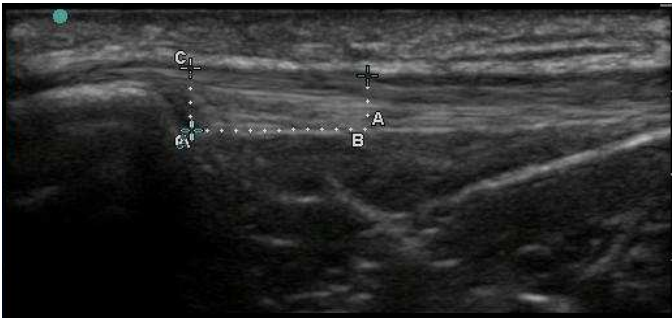
FIGURE 2**Figure 2** | Sonographic image of measurement site (AP & 1uAP).

Image above: Longitudinal anteroposterior thickness measurement of AP (C) and 1uAP (A), respectively. 1uAP was found by tracing 1 cm. distally from the posterior border of AP parallel with the tendon fibres (B).

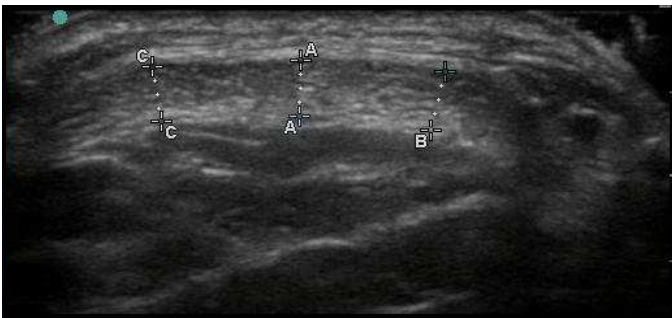
FIGURE 3**Figure 3** | Sonographic image of measurement site (Transversal)

Image above: Transversal anteroposterior thickness measurements of the patellar tendon. C represents the thickest lateral third, A represents the thickest middle third and B represents the thickest medial third.

Transversal thickness measurements were performed at the following sites:

- Apex Patella, represented by the above image (1a, 1b & 1c)
- 1 cm. distal to Apex Patella (2a, 2b & 2c)
- Tibial tuberosity (3a, 3b & 3c)
- 1 cm. proximal to the Tibial tuberosity (4a, 4b & 4c).

FIGURE 4

Figure 4 | Sonographic image of measurement site (TT & 1oTT)

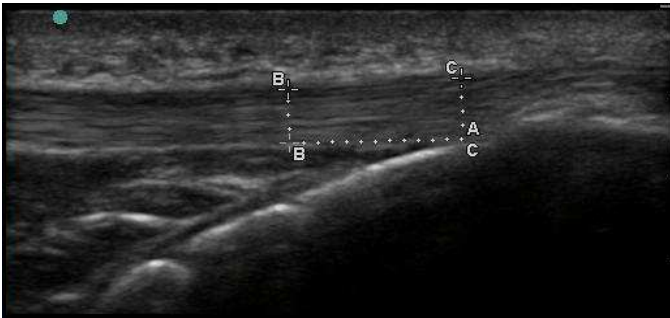
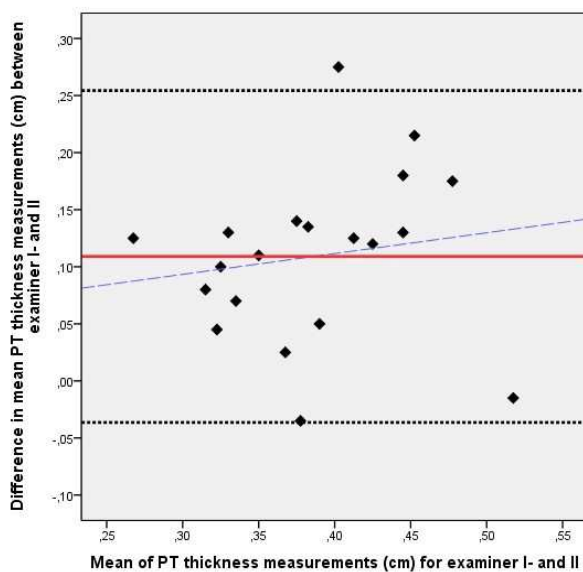


Image above: Longitudinal anteroposterior thickness measurement of the TT (C) and 1oTT (B). 1oTT was found by tracing 1 cm. proximally from the posterior border of TT parallel with the tendon fibres (B).

FIGURE 5

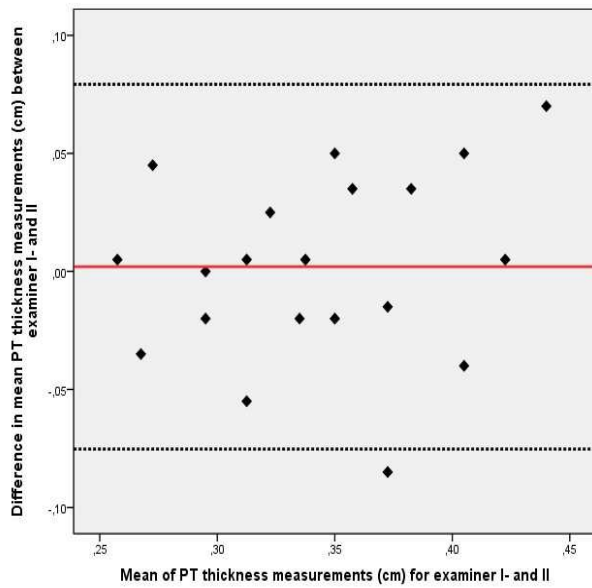
Figure 5 | Interrater Bland Altman plot with 95% limits of agreement (dotted black lines) for the measurement site AP.



The regression analysis (dotted blue line) revealed no systematic different distribution of scores ($p=0.519$). The other measurement site (1 cm. u. AP) with a visually suggested proportional biased trend of difference scores was similarly non-significant in the regression analysis ($p=0.073$) (plot not shown).

FIGURE 6

Figure 6 | Interrater Bland Altman plot with 95% limits of agreement (dotted black lines) of the measurement site IoTT.



This representative Bland Altman plot showed no systematic different distribution of measurements. Bland Altman plots of the other measurement sites had a similar appearance, revealing no systematic difference over the measurement range (plots not shown).

Reliability and precision for non-dominant leg

Table 1 | Patellar tendon thickness (cm) on non-dominant leg from all 16 measurement points for examiner I. Results presented in mean \pm SD.

PT measurement points	Men (n=11) Mean (SD)	Women (n=9) Mean (SD)	Total (n=20) Mean (SD)
AP	0.45 (0.08)	0.42 (0.05)	0.43 (0.07)*
1a	0.35 (0.06)	0.32 (0.04)	0.33 (0.06)
1b	0.36 (0.06)	0.31 (0.04)	0.34 (0.06)
1c	0.35 (0.05)	0.31 (0.03)	0.33 (0.05)
1 u. AP	0.38 (0.07)	0.34 (0.06)*	0.36 (0.07)*
2a	0.34 (0.07)	0.31 (0.03)	0.32 (0.06)
2b	0.34 (0.05)	0.30 (0.03)	0.32 (0.05)
2c	0.33 (0.06)	0.32 (0.04)	0.32 (0.05)
TT	0.45 (0.07)	0.40 (0.06)	0.43 (0.07)
3a	0.35 (0.05)	0.30 (0.05)	0.33 (0.06)
3b	0.34 (0.06)	0.29 (0.04)	0.32 (0.06)
3c	0.35 (0.07)	0.33 (0.04)	0.34 (0.06)
1 o. TT	0.35 (0.05)	0.31 (0.05)	0.34 (0.06)
4a	0.34 (0.07)	0.30 (0.04)	0.32 (0.06)
4b	0.34 (0.05)*	0.29 (0.05)	0.32 (0.06)*
4c	0.33 (0.05)	0.30 (0.03)	0.31 (0.05)

*Significantly different patellar tendon thickness between first and second measurement ($p < 0.05$).

Table 3 | Intra- & Interrater reliability on non-dominant leg for both examiners. ICC, (95% CI).

	Examiner I	Examiner II	Examiner I + II
AP	*	▶ 0.68 (0.37 - 0.86)	*
1a	▶ 0.61 (0.25 - 0.83)	▲ 0.80 (0.56 - 0.91)	▲ 0.85 (0.62 - 0.94)
1b	▶ 0.71 (0.39 - 0.87)	*	*
1c	▶ 0.61 (0.25 - 0.82)	▲ 0.83 (0.61 - 0.93)	▶ 0.74 (0.35 - 0.90)
1 u. AP	*	▲ 0.77 (0.51 - 0.90)	*
2a	▲ 0.87 (0.70 - 0.94)	*	*
2b	▲ 0.76 (0.50 - 0.90)	▶ 0.70 (0.40 - 0.87)	▲ 0.89 (0.72 - 0.96)
2c	▶ 0.75 (0.47 - 0.89)	▲ 0.82 (0.61 - 0.93)	▲ 0.84 (0.61 - 0.94)
TT	▶ 0.72 (0.42 - 0.88)	▲ 0.80 (0.56 - 0.92)	▶ 0.58 (0.07 - 0.83)
3a	▲ 0.90 (0.77 - 0.96)	▲ 0.90 (0.78 - 0.96)	▶ 0.72 (0.20 - 0.89)
3b	▶ 0.74 (0.47 - 0.89)	▶ 0.75 (0.48 - 0.89)	▲ 0.85 (0.36 - 0.95)
3c	▲ 0.80 (0.57 - 0.92)	▶ 0.66 (0.31 - 0.85)	▶ 0.67 (0.15 - 0.87)
1 o. TT	▲ 0.83 (0.62 - 0.93)	▲ 0.78 (0.53 - 0.91)	▲ 0.78 (0.45 - 0.91)
4a	▲ 0.89 (0.74 - 0.96)	▲ 0.82 (0.61 - 0.93)	▲ 0.91 (0.76 - 0.96)
4b	*	▶ 0.73 (0.44 - 0.88)	*
4c	▲ 0.78 (0.52 - 0.90)	▲ 0.80 (0.56 - 0.92)	▲ 0.83 (0.53 - 0.94)

▲ (ICC > 0.75) = good reliability

▶ (ICC 0.50 - 0.75) = moderate reliability

▼ (ICC 0.0 - 0.50) = low reliability

*Excluded from reliability due to a significant difference in PT thickness between first and second measurement.

Table 2 | Patellar tendon thickness (cm) on non-dominant leg from all 16 measurement points for examiner II. Results presented in mean \pm SD.

PT measurement points	Men (n=11) Mean (SD)	Women (n=9) Mean (SD)	Total (n=20) Mean (SD)
AP	0.34 (0.06)	0.35 (0.06)	0.35 (0.06)
1a	0.39 (0.08)	0.31 (0.05)	0.36 (0.08)
1b	0.42 (0.10)	0.33 (0.06)*	0.38 (0.09)*
1c	0.38 (0.09)	0.29 (0.04)	0.34 (0.08)
1 u. AP	0.33 (0.07)	0.32 (0.07)	0.32 (0.07)
2a	0.34 (0.06)	0.29 (0.03)	0.32 (0.06)*
2b	0.34 (0.05)	0.28 (0.03)	0.32 (0.05)
2c	0.34 (0.04)	0.28 (0.03)	0.31 (0.05)
TT	0.42 (0.06)	0.35 (0.05)	0.39 (0.07)
3a	0.40 (0.06)	0.33 (0.05)	0.37 (0.07)
3b	0.37 (0.05)	0.32 (0.04)	0.35 (0.05)
3c	0.40 (0.05)	0.34 (0.03)	0.37 (0.05)
1 o. TT	0.38 (0.05)	0.32 (0.04)*	0.35 (0.05)
4a	0.36 (0.05)	0.29 (0.04)	0.33 (0.06)
4b	0.35 (0.06)	0.28 (0.03)	0.32 (0.06)
4c	0.36 (0.06)	0.30 (0.03)	0.33 (0.06)

*Significantly different patellar tendon thickness between first and second measurement ($p < 0.05$).

Table 4 | Intra- & interrater reliability for non-dominant leg presented as 95% LOA and as percentage of the mean PT thickness (LOA %)

	Examiner I		Examiner II		Examiner I + II	
	LOA (cm)	LOA (%)	LOA (cm)	LOA (%)	LOA (cm)	LOA (%)
AP	*	*	0.10	30,3	*	*
1a	0,11	32,4	0,10	28,5	0,09	25,9
1b	0,09	25,5	*	*	*	*
1c	0,09	27,6	0,10	29,6	0,12	35,5
1 u. AP	*	*	0,10	29,9	*	*
2a	0,06	18,9	*	*	*	*
2b	0,07	21,3	0,09	27,5	0,06	20,2
2c	0,08	23,9	0,06	18,8	0,07	22,2
TT	0,12	28,0	0,09	22,3	0,15	35,6
3a	0,05	15,4	0,06	16,2	0,10	29,5
3b	0,09	28,6	0,07	21,1	0,06	19,1
3c	0,08	22,1	0,10	26,6	0,10	28,7
1 o. TT	0,07	20,3	0,08	21,6	0,09	26,5
4a	0,06	18,7	0,07	22,5	0,07	21,9
4b	*	*	0,09	28,4	*	*
4c	0,06	20,7	0,08	22,6	0,07	22,1

*Excluded from measurement precision analysis due to a significant Difference in PT thickness between first and second measurement.

Declarations of interests

None

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT