Intermachine variation of ultrasound strain elastographic measures of the quadriceps and patellar tendons in healthy participants: implications for clinical practice

Dickson, Diane M.; Fawole, Henrietta O.; Hendry, Gordon J.; Smith, Stephanie L.

Published in:
Journal of Ultrasound in Medicine

DOI:
[10.1002/jum.15228](https://doi.org/10.1002/jum.15228)

Publication date:
2020

Document Version
Author accepted manuscript

[Link to publication in ResearchOnline](#)

Citation for published version (Harvard):
Dickson, DM, Fawole, HO, Hendry, GJ & Smith, SL 2020, 'Intermachine variation of ultrasound strain elastographic measures of the quadriceps and patellar tendons in healthy participants: implications for clinical practice', *Journal of Ultrasound in Medicine*, vol. 39, no. 7, pp. 1343-1353. <https://doi.org/10.1002/jum.15228>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please view our takedown policy at <https://edshare.gcu.ac.uk/id/eprint/5179> for details of how to contact us.

Title: Inter-machine variation of ultrasound strain elastography measures of the quadriceps and patellar tendon in healthy participants: implications for clinical practice.

Manuscript category: Original research

Diane M Dickson^{1*}, Henrietta O Fawole^{1,2}, Gordon J Hendry¹, Stephanie L Smith¹

¹Glasgow Caledonian University, School of Health and Life Sciences, Glasgow, UK.

²University of Benin, Physiotherapy Department, School of Basic Medical Sciences, Benin-City, Nigeria.

DM Dickson, Senior Lecturer, MSc

HO Fawole, Musculoskeletal Health PhD¹ student, MSc and Lecturer²

SL Smith, Musculoskeletal Health Research Fellow, PhD

GJ Hendry, Musculoskeletal Health Senior Research Fellow, PhD

***Reprints and Correspondence**

DM Dickson,

A366, Glasgow Caledonian University, School of Health and Life Science

70 Cowcaddens Road, Glasgow, Scotland, UK

G4 0BA

T: +44 (0)141 273 1396| F: +44 (0)141 273 1396 | E: Diane.Dickson@gcu.ac.uk

Running title: Inter-machine variation in strain elastography

Abstract

Objectives: To evaluate inter-machine variation and compare intra- and inter-operator agreement and repeatability characteristics of two ultrasound systems for strain elastography measurements of quadriceps and patellar tendons.

Methods: Forty tendons from 20 healthy participants were investigated by operators with different experience (operator 1; 12 years' ultrasound and >50 strain elastography (SE) examinations, operator 2; no ultrasound experience, 1 day SE training). Repeated measures were performed employing GE and esaote ultrasound systems. Percentage agreement, Cohen's kappa, intra-class correlation coefficient, and correlation tests assessed agreement, repeatability and associations of SE measures. Paired T-Test and Wilcoxon signed rank assessed differences in SE measures.

Results: The study participants included 5 males, 15 female volunteers; mean (range) age 29.3 (21-39) years). Better agreement and repeatability characteristics were observed for the patellar compared to the quadriceps tendon, and colour score (CS) method over the elasticity ratio (ER). Intra-operator agreement was better for the experienced operator. Intra-operator repeatability was achieved in 55% of ER (ICC, 0.40-0.91, $p < 0.05$) and 77-85% (k, -0.25-1) of CS measures. Inter-operator repeatability was achieved in 35% (t/z, -2.93-7.94, p , 0.000-0.048) of all ER measures. No significant difference in proximal (z, -0.13- -0.78) and distal patellar (z, -1.52- 2.26, $p > 0.5$) ER measures were observed. 74-75% mean agreement (k, 0-0.5) for CS measures comparable across both ultrasound systems was observed. Inter-machine ER associations were poor (r, -0.39-0.13, $p > 0.05$), whereas greater than 70% agreement (k, -0.87-0.53) for CS was achieved.

Conclusions: Reproducibility of knee tendon strain elastography measurements is influenced by operator experience, ultrasound system, and tendon site.

Keywords: *elastography; ultrasound; comparison; knee tendon; healthy subjects.*

Introduction

Ultrasound elastography is an imaging technique offered by several commercial ultrasound equipment manufacturers, and is used to assess tissue elasticity. Two common types of ultrasound elastography are strain elastography (SE) and shear wave elastography (SWE). SE was the first to be introduced and is performed by manual compression of the ultrasound transducer to produce tissue stress. Tissue responses to applied external stress enables qualitative and semi- quantitative assessment of tissue strain; where large strain equates to softer tissue, and low strain corresponds to harder tissue¹. SWE makes use of shear wave propagation speed to quantify tissue stiffness². Elastography techniques can be used to investigate the mechanical properties of tissue to determine tissue status³. However, ultrasound elastography presents inherent technical challenges. SE is the one of the most common techniques⁴ and primarily suffers from operator dependency, while SWE is subject to tissue non-uniformity and boundary artifacts⁵. Despite technical limitations, both techniques continue to be employed in research and clinical practice³.

SE is widely available and is considered clinically appropriate in breast, prostate, lymph node and thyroid imaging⁶⁻¹⁰. Over recent years there has been increased interest in SE of tendons; where validity¹¹ and reliability of the technique has been demonstrated in the Achilles¹² and patellar tendons^{13,14}. Robust measurements of mechanical properties of tendons may provide valuable functional information for identification of abnormal tendon stiffness that could inform therapeutic intervention^{11,15}.

Application measurement packages for SE vary by manufacturer, but include qualitative colour score mapping (CS), where colour represents the degree of tissue stiffness and is visually scored. Semi-quantitative analyses such as histograms demonstrating tissue stiffness

distribution based on a colour hue scale and an elasticity index (EI) calculating the average stiffness value inside a region of interest (ROI) may also be available. Additionally, the elasticity strain ratio (ER) can be utilised to calculate the stiffness ratio between two ROI's (normal tissue and tissue under investigation)^{16,17}. SE imaging protocols are currently not standardised for the assessment and interpretation of tendon examinations. As such, previous studies have adopted a variety of heterogeneous methodologies^{13,18–20}, making comparison of results challenging.

As SE use increases within the musculoskeletal field, it is important to establish sources of error, to ensure that clinically significant differences can be detected. There may be three main potential sources of measurement variation. The first includes biologic alteration²¹ where tendon composition may fluctuate normally under the influence of changing physical activity levels, muscle strength, pathology, and hormones²². The second is measurement environment influenced by imaging protocol and the clinician performing and interpreting the examination (intra- and inter-operator discrepancy)²¹. The third is the measurement instrument²¹. While factors influencing tendon properties^{23,24} and operator reproducibility^{12–14,18,25} have been somewhat investigated, there is a scarcity of inter-machine reproducibility research. Imaging services often host a variety of equipment types leading to recurrent service users being examined on more than one ultrasound system. Similarly, different operators may perform follow up examinations, which highlights the importance of reproducibility of the technique and equipment for accurate imaging outcomes. Despite previous findings of high intra and inter-operator reproducibility in Achilles and patellar tendon SE studies^{12,13}. Inter-machine reproducibility has not been evaluated. SE patellar tendon reproducibility has been reported as good-to-excellent (ICC; 0.690-0.919) based on two experienced operators performing and interpreting scan findings¹³. We have previously found that reproducibility was experience-dependent when operators performed

measurements on scans they acquired¹⁴. Other studies are largely limited to observers' review of previously recorded static images or cine clips^{18–20,26–28} and fails to capture the influence of scan performance on reproducibility. There are currently no studies involving human tendons evaluating reproducibility and agreement of SE measures from different ultrasound systems across different operators.

Accordingly, this study aims to assess 1) the inter-machine agreement and association of SE measures of the quadriceps and patellar tendon performed using two commercially available ultrasound systems 2) and the difference in repeatability and reproducibility within (intra-operator) 3) and across (inter-operator) two operators with different levels of experience.

Materials and Methods

Participants

The study was approved by the institutional ethics committee (HLS/PSWAHP/16/203) and conducted within the imaging suite of the research institution. Written informed consent was obtained from all participants included in the study. Participants were included if they met all of the following criteria; they were able to consent to study participation, aged between 18 and ≤ 40 , had no history of knee osteoarthritis or history of knee pain (in last 3 months), no history of knee injury or surgery, and a body fat content that was considered within normal limits (male $<25\%$, female $<39\%$)²⁹. Participants were excluded if they met any of the following; presence of a condition affecting the quadriceps or patellar tendon including; tendinopathy, rheumatological/musculoskeletal condition, knee pain/ instability, or history of knee surgery. The presence of any abnormalities including altered shape, fibrillary pattern or echo texture, detected using B-Mode ultrasound during screening; self-reported presence or

history of autoimmune or connective tissue disorders; or currently in receipt of oestrogen or steroid medication due to previous association with tendon abnormalities²³.

Examination protocol

Equipment

Participants were examined using clinical practice standard ultrasound equipment with linear array transducers; GE Logic S8, (software version R2, revision 1.1, GE Healthcare, USA) (L6-15MHz) and esaote Mylab 70 XVG, (version EVO 13.60M, esaote, Italy) (LA523, L4-13MHz). The scanning sequence is summarised in Figure 1.

Location for Figure 1

Operators

Participants were scanned consecutively and independently by two operators with different levels of ultrasound and SE experience (operator 1, formally trained with 12 years' clinical US experience and > 50 SE examinations and operator 2, no US experience and 1-day protocol-specific SE application training <5 SE examinations). A pictorial protocol for standardised anatomical and technical assessment was developed for operator reference.

Scan protocol

Participants' self-reported dominant lower limb (defined as leg used to kick a ball) was scanned. Participants were scanned in a seated or lying position with the knee supported in flexion using a standardised 30 degree pad³⁰. Using B-mode, the quadriceps tendon was located in longitudinal orientation based on previous recommendations for higher reproducibility¹⁸, adopting a perpendicular approach. A SE map was applied at defined anatomical areas (PQT; myotendintendous junction used as superior landmark for region

standardisation, DQT; base of patella used as the distal most structure, PPT; patella apex as superior landmark, MPT; middle third of patellar tendon between myotendinous junction and distal insertion and DPT; distal patellar tendon insertion as distal image landmark, Figure 2). Elastograms were produced using previously published optimal settings³¹. The elastography box was standardised to full sector width and 2.5cm depth to allow full coverage of knee tendons whilst maintaining 25%-50% of colour map box region of interest (ROI)^{18,27,32}. Using the onscreen quality indicator to demonstrate sufficient stress, light transducer compression was applied and a minimum 5 second duration cine loop recorded as recommended previously^{11,12,17,18}. This technique was performed and repeated to obtain two scan series for all anatomical areas (Figure 1) by each operator, and ultrasound system, successively and separated by at least 15 minutes, before the second operator repeated the process.

Image analysis

Each operator selected a representative static image from their previously captured cine loop during the middle of the compression-relaxation cycle, which demonstrated sufficient stress represented by green quality indication bar. Analysis was performed directly from the US equipment ensuring standardisation of viewing conditions. Respective scans and analyses were undertaken independently by each operator and recorded separately. To eliminate recall bias, scoring analysis of scan measure 1 and scan measure 2 was performed over multiple dates, commencing at least one month following scan series, and repeated scans were not scored in succession.

Strain elastography measures

CS and ER were included for the evaluation of agreement and repeatability of SE measurements of knee tendons due to availability on both ultrasound systems.

168 Colour map score

169 Visual grading of respective elastography colour maps were performed and recorded by each
170 operator. Using a similar three-point scale employed in previous studies^{11,20,25}, Grade 1, No
171 strain/hard = blue (no green colour evident); Grade 2, Average strain/intermediate =
172 green/yellow or green (no red colour evident); and Grade 3, Greatest strain/soft = red
173 (positive for red colour).

174 Elasticity ratio

175 To perform ER measurements, the Q-analysis ratio option within the elastography package
176 (GE) and elastography measurement tab, option ElaXto Ratio (ELX-T-RAT, esaote) were
177 selected. Using the corresponding reference site (quadriceps tendon = pre-femoral fat pad
178 and patellar tendon = Hoffa's fat pad), a small fixed size reference ROI of 5mm was
179 positioned within homogenous fat pad tissue (Figure 2). Reference tissue ROI is ideally
180 positioned adjacent to, and at the same depth as the tissue area of interest, however this is not
181 anatomically possible for knee tendon studies. Commonly used reference sites from previous
182 studies include the Kager's fat pad, positioned deep to the Achilles tendon, or within patellar
183 and quadriceps studies, the Hoffa's and pre-femoral fat pads^{19,33,34}. These sites are both
184 accessible and facilitate easy placement of ROI's within one sector width, thus minimizing
185 placement error, and provide opportunity for both the tissue of interest and reference tissue to
186 undergo similar stress³⁵. Due to the heterogeneous composition of fat pad, a small
187 homogeneous soft (red) tissue area was selected to provide a standardised and easily
188 identifiable reference ROI. An anatomical site ROI was freehand traced with US machine
189 tracker-ball, and an ER value was automatically calculated, and recorded.

Statistical analysis

Intra-, inter-operator and inter-machine agreement of CS (grade 1-3) were estimated using the percentage of exact agreement³⁶ and Cohen's quadratic weighted kappa³⁷. Two way random, single measures, intra-class correlation coefficients (ICC_{2,1}) were performed to assess intra operator repeatability of ER for each US system. Inter-operator ER measurements were compared using paired sample T- or Wilcoxon signed rank tests according to data distributions, and were reported for each US system. Evaluation of ER inter-machine associations were performed using Pearson's or Spearman correlations, as the ER values for each system were expressed using different scales. The minimum accepted statistical threshold levels for Kappa and ICC tests were defined as ≥ 0.40 , (poor <0.40 , fair 0.40-0.59, good 0.60-0.75, excellent 0.75-1.00) with the level of minimum clinical acceptability defined as fair or above³⁸. The minimum accepted correlation statistical threshold levels were defined as ≤ 0.29 , small; 0.30-0.49, medium and ≥ 0.50 large, with the level of minimum clinical acceptability defined as medium or above³⁹. Two-tailed statistical significance level was defined as $p \leq 0.05$. Statistical analyses were performed using SPSS, version 24⁴⁰.

Results

Participant characteristics

Twenty healthy adult volunteers (5 males, 15 females; mean (range) age 29.3 (21-39) years; Body Mass Index 23.2 (17.9-29.0) kg/m²; body fat percentage 23.8 (13-39) %) were included in this study. Dominant lower limbs were self-reported (Left n=5, Right n=15). 40 tendons were examined twice with both ultrasound systems performed by two operators; a total of 400 SE examinations and measurements were performed by each operator (5 anatomical sites, scanned twice on each system for 20 participants).

Intra-operator agreement and repeatability

Total percentage agreement for CS ranged from a mean of 97% to 73% agreement for the most to the least experienced operator across two measurements and ultrasound systems (Table 1). Percentage agreement and kappa scores were better for operator 1 using the GE equipment (90-100%; k , 0.78-1.0) compared to the esaote (70-95%; k , 0-0.79). Operator 2 achieved similar CS agreement from both equipment types (Table 1). Operators mean percentage agreement scores across all anatomical tendon sites were higher for the GE equipment (GE; 85% and esaote; 77%). Both systems produced higher mean percentage agreement scores for the patellar tendon compared to the quadriceps tendon (Table 1). Operator 1 achieved the same (distal patellar tendon (DPT)) or greater percentage agreement scores than operator 2, across CS measures (Table 1) for the GE system. Operator 2 produced higher percentage agreement scores in two patellar tendon sites compared to operator 1 using the esaote equipment (PPT and DPT). Percentage agreement scores were higher at the DPT for the GE equipment and the mid patellar tendon (MPT) for the esaote equipment (mean % agreement, 95% and 88% respectively).

ER ICC's were better for operator 1 compared to operator 2 (Table 2). Each operator achieved ICCs >0.40 (fair to good) at three anatomical sites using the esaote equipment (Table 2). Operator 1 achieved intra-operator repeatability (ICCs >0.40) at four tendon sites (fair to excellent), and operator 2, one tendon site, using the GE system (fair). Only the PPT site was demonstrated to be a repeatable ER measure (ICCs >0.40) by each operator and both systems. Operator 1 demonstrated fair to excellent repeatability across all anatomical sites but failed to demonstrate repeatability across all sites using a single ultrasound system.

Location for Table 1 and Table 2

Inter-operator agreement and repeatability

Mean percentage agreement over two CS measures between operators were 74% and 75% for the GE and esaote equipment respectively. The patellar tendon yielded better CS agreement (range 70-95%, Table 1) compared to the quadriceps tendon (range 53-68%). Mean percentage agreement was greatest at DPT for the GE equipment and MPT for the esaote equipment (both 95%). No significant differences were observed between operator 1 and operator 2 for ER measurements for tendon sites PPT and DPT using the GE equipment (Table 2). All other GE ER tendon inter-operator measures demonstrated significant differences ($p<0.05$). For the esaote system, significant difference was recorded between operators for ER measurement of the MPT ($p<0.05$), no other tendon site measures were significantly different.

Inter-machine agreement and association

Operator 1 produced 100% CS agreement between machines at two tendon sites (PPT and MPT). Across all tendon sites, 71% mean inter-machine agreement was achieved (Table 3). Two tendon sites (PPT and MPT) achieved CS inter-machine agreement $\geq 80\%$. The mean inter-machine CS agreement of the PPT was 58%. Higher inter-machine percentage agreement scores of patellar tendon CS were obtained compared to the quadriceps tendon (patellar tendon; 77 and quadriceps tendon; 61% mean agreement). Inter-machine ER correlations were poor (<0.40) for both operators and all tendon sites (Table 4).

Location for Table 3 and Table 4

Discussion

This study is the first to report agreement and repeatability of knee tendon SE measures using different ultrasound systems and different operators, reflecting clinical practice. Findings demonstrate high mean inter-machine agreement (71%) for CS and poor inter-machine

association of ER measures. Repeatability of SE measures were better for the experienced operator, and inter-operator agreement was better using the esaote compared to the GE system.

SE performed using the GE equipment demonstrated better CS intra-operator agreement compared with the esaote (mean % agreement, 85% and 77%). This variance may be the result of the higher broadband frequency of the GE system transducer (SE centre frequency 15MHz) compared to the esaote (12MHz). Higher transducer frequency offers higher strain resolution⁴¹, therefore may have contributed to better agreement in visually determined CS for the experienced operator and supports use of the higher frequency transducer in clinical practice.

SE is an individually performed freehand technique, consequently some variability is inevitable. A recent phantom based study found greater ER accuracy over CS method³². Accuracy of lesions of known stiffness improved by 5.7-16.9% for experienced and inexperienced operators³². Clinical tests must be both accurate and reliable⁴². However, our clinical study found ER to be less repeatable than the CS method. Another *in vitro* study evaluated two SE systems (GE and Hitachi) within an elastography equipment comparison study. Excellent intra-(ICC; 0.977-0.998) and inter-operator (ICC; 0.845-0.995) reliability was reported of 10 repeated ER measures performed by operators with SE experience ranging from 1 month to 7 years⁴³. Our poorer ICC ranges may be result of both methodological differences and the influence of patient factors compared to *in vitro* investigation. Employing a fixed size ROI provides methodological standardisation and reduced operator input, where excellent repeatability was achieved⁴³, despite including a relatively inexperienced operator. However, this method restricts the autonomy of the operator and may not provide adequate coverage to fully investigate the tissue of interest. Replicating clinical practice and enabling the clinical advantages of dynamic ultrasound techniques, our study permits operators to

define the measurement ROI based on examination findings⁹. Although our findings suggest *in vivo* ER measurements are highly operator dependent, no significant difference in inter-operator ER measures were recorded at the proximal patellar tendon (PPT; p ; 0.455-0.896) and distal patellar tendon (DPT; p ; 0.09-0.36, Table 2) using either system. Furthermore, the PPT is identified as an anatomical site demonstrating fair to excellent intra-operator agreement of SE measures using both systems, advancing current knowledge and suggesting feasibility as a ROI for repeat clinical measures.

Consistent with previous B-mode ultrasound⁴⁴ and SE literature¹⁴, influence of operator experience is demonstrated within our study. Our findings further verify the operator dependency of ultrasound techniques, and highlight the importance of alignment to current professional guidance which recognises the requirement for appropriate education and training standards⁴⁵.

For both operators, discrepancy was observed between systems in the ER repeatability of examination of the MPT. Our findings suggest that despite operator experience, the esaote equipment provides opportunity to obtain more repeatable measures of the MPT compared to the GE. Although no formal qualitative evaluation of user friendliness was included within this study, both operators reported that ER measurements were more straightforward to perform using the esaote equipment, which may account for this finding. In light of a lack of literature in this area, based on our results, we hypothesise that ease of machine use may influence repeatability of measures in operators with different experience levels, particularly in less experienced operators.

Regardless of system or operator, measurements of the proximal and distal patellar tendon were more repeatable compared to the quadriceps tendon. Anatomical identification of the more superficially positioned and well defined patellar tendon may contribute to this finding.

The gradual muscle to tendon attachment of the proximal quadriceps tendon on B-mode ultrasound, in comparison to the abrupt and clearly defined patellar tendon attachment to bone⁴⁶ may account for the agreement and repeatability variance exhibited.

We found that CS measures were generally agreeable, particularly for the patellar tendon (Table 1). This indicates that a standardised three-point visual colour mapping score can be employed to achieve high agreement of repeated measures across operators and the ultrasound systems evaluated within this study. Similarly, Yoon et al reported greater CS agreement ($k=0.37$) across experienced operators using a standardised score, compared to real-time ER measurements of breast lesions ($k=0.28$)⁴⁷. Future clinical studies are required to contextualise the clinical relevance of CS findings in pathological groups.

Inter-machine variations of semi-quantitative measures (ER) are expected due to the unit-less nature of SE values adopted by different manufacturers¹. A previous study involving patellar tendon length and cross sectional area measurement in healthy subjects, reported high inter-machine agreement (ICC; 0.90-0.96). However, power Doppler (a more advanced ultrasound application) inter-machine studies of rheumatological small joints, have reported variable reproducibility between three ultrasound machine pairings (k ; 0.434, 0.455, 0.709)⁴⁸, and of four different machines, across multiple experienced operators (k ; 0.20-0.80)⁴⁹. Advancing previous inter-machine literature, our results inform clinical SE practice in knee tendon. To ensure differences in results are attributed to pathology, rather than random and/or systematic error, we recommend repeated ER measures are performed by an experienced operator using the same equipment. Furthermore, we recommended that reporting of clinical examinations should clearly state which ultrasound system has been employed.

Ultrasound equipment manufacturers make clear that semi-quantitative SE measures provide a relative strain reference value only, and that this data should not be translated for diagnostic

purposes¹⁶. Furthermore, the SE technique is operator dependant¹⁴, with no gold standard method of obtaining ER measures, and ongoing investigation of optimal reference regions⁵⁰. Yet, SE is actively used in clinical practice, where clinicians should be fully aware of its limitations. Our study participants were scanned during one appointment successively on both ultrasound systems, with no perceived change in tendon status or anatomical reason to account for deviation in SE measures. To evaluate machine influence on SE measures, bias was minimised by performing measurements under the same conditions using the same subjects, and measurement methods which are currently used in the practice setting. Subsequently, we are the first to identify variation within and between operators and machines, for CS and ER measures of the healthy quadriceps and patellar tendons performed under standardised conditions. However, ER measurements performed by an experienced and inexperienced operator are repeatable (within operator) and not significantly different (between operators) for PPT and DPT sites using either system, further supporting machine influence on repeatability of ER measurements of knee tendon in healthy participants. Future clinical studies are required to assess the accuracy of SE findings at these regions to determine future clinical relevance.

There are several limitations to this study, including the relatively small sample size. Lack of prevalence across all three CS criteria resulting in non-calculated kappa values. SE is an emerging technique and as such no standardised method of examination is agreed. To minimise known bias's such as operator dependency and methodological variations, the same subjects under the same conditions using a standardised methodology with currently clinically applied reference sites, has been adopted. Only two ultrasound systems were directly compared despite other manufacturers offering SE. SE measurements CS and ER were evaluated however alternative semi-quantification methods are available. These were

not included in this study due to data produced on different scales, resulting in the inability to make semi-quantitative comparisons. Two operators of different levels of experience were included in this study (one experienced and one inexperienced). Future SE reliability studies may choose to include only experienced operators to reflect appropriate clinical practice users. Knee joint angle was determined using a standardised pad, placed under the knee in line with clinical practice; however, knee angle was not directly measured.

Conclusion

Comparing CS and ER measurements of knee tendons of two commercially available SE ultrasound systems, we have revealed intra-, inter-operator and inter-machine variation. Greater repeatability was exhibited by the experienced operator, and in patellar tendons sites compared to the quadriceps tendon. Inter-machine variability is evident, with poor association for ER measures. We recommend that knee tendon SE is performed by experienced operators where high intra-operator and inter-machine agreement of CS measures was observed. Our findings suggest that reproducible CS measures of the quadriceps and patellar tendon can be achieved using the ultrasound systems included in this study. Intra- and inter-operator repeatability was demonstrated at the PPT and DPT with no significant difference in ER measures, suggesting these tendon site suitable for repeated measures. Inter-machine association of ER measures at all tendon sites were poor. To reduce variability of SE ER measurements of knee tendon, it is recommended that experienced operators perform follow up examinations using a standardised protocol, employing the same ultrasound system. Ultrasound manufacturer standardisation is required to eliminate inter-machine variability of ER measurements, and clinical studies required to determine accuracy and significance of SE findings for wider clinical practice use.

382

383 **References**

- 384 1. Ozturk A, Grajo J, Dhyani M, et al. Principles of ultrasound elastography. *Abdom*
385 *Radiol* 2018; 25: 773–785.
- 386 2. Nowicki A, Dobruch-sobczak K. Introduction to ultrasound elastography. *J Ultrason*
387 2016; 16: 113–124.
- 388 3. Prado-Costa R, Rebelo J, Monteiro-Barroso J, et al. Ultrasound elastography:
389 compression elastography and shear-wave elastography in the assessment of tendon
390 injury. *Insights Imaging* 2018; 9: 791–814.
- 391 4. Drakonaki EE, Allen GM, Wilson DJ. Ultrasound elastography for musculoskeletal
392 applications. *Br J Radiol* 2012; 85: 1435–1445.
- 393 5. Bamber J, Cosgrove D, Dietrich CF, et al. EFSUMB guidelines and recommendations
394 on the clinical use of ultrasound elastography. Part 1: Basic principles and technology.
395 *Ultraschall der Medizin - Eur J Ultrasound* 2013; 34: 169–84.
- 396 6. Bojanic K, Katavic N, Smolic M, et al. Implementation of Elastography Score and
397 Strain Ratio in Combination with B-Mode Ultrasound Avoids Unnecessary Biopsies of
398 Breast Lesions. *Ultrasound Med Biol* 2017; 43: 804–816.
- 399 7. Barr RG. The Role of Sonoelastography in Breast Lesions. *Semin Ultrasound, CT MRI*
400 2018; 39: 98–105.
- 401 8. Barr RG, Cosgrove D, Brock M, et al. WFUMB Guidelines and Recommendations on
402 the Clinical Use of Ultrasound Elastography: Part 5. Prostate. *Ultrasound Med Biol*
403 2017; 43: 27–48.

- 404 9. Cosgrove D, Barr R, Bojunga J, et al. WFUMB Guidelines and Recommendations on
405 the Clinical Use of Ultrasound Elastography: Part 4. Thyroid. *Ultrasound Med Biol*
406 2017; 43: 4–26.
- 407 10. Ying L, Hou Y, Zheng HM, et al. Real-time elastography for the differentiation of
408 benign and malignant superficial lymph nodes: A meta-analysis. *Eur J Radiol* 2012;
409 81: 2576–2584.
- 410 11. Klauser AS, Miyamoto H, Tamegger M, et al. Achilles Tendon Assessed with
411 Sonoelastography: Histologic Agreement. *Radiology* 2013; 267: 837–842.
- 412 12. Schneebeil A, Grande F Del, Vincenzo G, et al. Real-time sonoelastography using an
413 external reference material: test-retest reliability of healthy Achilles tendons. *Skeletal*
414 *Radiol* 2016; 45: 1045–1052.
- 415 13. Porta F, Damjanov N, Galluccio F, et al. Ultrasound elastography is a reproducible and
416 feasible tool for the evaluation of the patellar tendon in healthy subjects. *Int J Rheum*
417 *Dis* 2014; 17: 762–766.
- 418 14. Dickson DM, Fawole HO, Newcombe L, et al. Reliability of ultrasound strain
419 elastography in the assessment of the quadriceps and patellar tendon in healthy adults.
420 *Ultrasound*. Epub ahead of print 2019. DOI: 10.1177/1742271X19859380.
- 421 15. Arya S, Kulig K. Tendinopathy alters mechanical and material properties of the
422 Achilles tendon. *J Appl Physiol* 2009; 108: 670–675.
- 423 16. esaote. *MyLab ElaXto option*. 2014.
- 424 17. GE. *Elastography: concept to clinical application*. 2012.
- 425 18. Drakonaki EE, Allen GM, Wilson DJ. Real-time ultrasound elastography of the normal
426 Achilles tendon: reproducibility and pattern description. *Clin Radiol* 2009; 64: 1196–

- 427 1202.
- 428 19. Ozcan ANS, Tan S, Tangal NG, et al. Real-time sonoelastography of the patellar and
429 quadriceps tendons: Pattern description in profesional athletes and healthy volunteers.
430 *Med Ultrason* 2016; 18: 299–304.
- 431 20. De Zordo T, Fink C, Feuchtner G, et al. Real-time sonoelastography findings in
432 healthy Achilles tendons. *Am J Roentgenol* 2009; 193: 134–138.
- 433 21. de Vet HCW, Terwee CB, Knol DL, et al. When to use agreement versus reliability
434 measures. *J Clin Epidemiol* 2006; 59: 1033–1039.
- 435 22. Bordini B, Varacallo M. *Anatomy , Tendons*. StatPearls Publishing LLC,
436 <https://www.ncbi.nlm.nih.gov/books/NBK513237/> (2018).
- 437 23. Capalbo E, Peli M, Stradiotti P. Sonoelastography of the distal third of the Achilles
438 tendon in asymptomatic volunteers: correlation with anthropometric data, ultrasound
439 findings and reproducibility of the method. *Radiol Med* 2016; 121: 667–674.
- 440 24. Taş S, Yilmaz S, Onur MR, et al. Patellar tendon mechanical properties change with
441 gender, body mass index and quadriceps femoris muscle strength. *Acta Orthop*
442 *Traumatol Turc* 2017; 51: 54–59.
- 443 25. De Zordo T, Chhem R, Smekal V, et al. Real-time sonoelastography: Findings in
444 patients with symptomatic achilles tendons and comparison to healthy volunteers.
445 *Ultraschall der Medizin* 2010; 31: 394–400.
- 446 26. Ooi CC, Schneider ME, Malliaras P, et al. Diagnostic Performance of Axial-Strain
447 Sonoelastography in Confirming Clinically Diagnosed Achilles Tendinopathy:
448 Comparison with B-Mode Ultrasound and Color Doppler Imaging. *Ultrasound Med*
449 *Biol* 2015; 1–11.

- 450 27. Tan S, Kudas S, Özcan AS, et al. Real-time sonoelastography of the Achilles tendon:
451 Pattern description in healthy subjects and patients with surgically repaired complete
452 ruptures. *Skeletal Radiol*. Epub ahead of print 2012. DOI: 10.1007/s00256-011-1339-
453 4.
- 454 28. Berko NS, Mehta AK, Levin TL, et al. Effect of knee position on the ultrasound
455 elastography appearance of the patellar tendon. *Clin Radiol* 2015; 70: 1083–1086.
- 456 29. Gallagher D, Heymsfield SB, Moonseong H, et al. Healthy Percentage Body Fat
457 Ranges an Approach for Developing Guideline Based on BMI. *Am J Clin Nutr* 2000;
458 72: 694–701.
- 459 30. Beggs I, Bueno A, Cohen M, et al. *Musculoskeletal Ultrasound Technical Guidelines*
460 *V. Knee*. 2010.
- 461 31. Havre RF, Elde E, Gilja OH, et al. Freehand Real-Time Elastography: Impact of
462 Scanning Parameters on Image Quality and In Vitro Intra- and Interobserver
463 Validations. *Ultrasound Med Biol* 2008; 34: 1638–1650.
- 464 32. Carlsen JF, Ewertsen C, Săftoiu A, et al. Accuracy of visual scoring and semi-
465 quantification of ultrasound strain elastography - A phantom study. *PLoS One*. Epub
466 ahead of print 2014. DOI: 10.1371/journal.pone.0088699.
- 467 33. Drakonaki E. Ultrasound elastography for imaging tendons and muscles. *J Ultrason*
468 2012; 12: 214–225.
- 469 34. Turan A, Teber MA, Yakut ZI, et al. Sonoelastographic assessment of the age-related
470 changes of the Achilles tendon. *Med Ultrason* 2015; 17: 58.
- 471 35. Havre RF, Waage JR, Gilja OH, et al. Real-Time Elastography: Strain Ratio
472 Measurements Are Influenced by the Position of the Reference Area. *Ultraschall der*

- 473 *Medizin*. Epub ahead of print 10 June 2011. DOI: 10.1055/s-001-1273247.
- 474 36. Cicchetti D V, Feinstein A. High Agreement but Low Kappa: II. Resolving the
475 Paradoxes. *J Clin Epidemiol* 1990; 43: 551–558.
- 476 37. VassarStats. Kappa as a measure of concordance in a categorical sorting. 1998-2019,
477 <http://vassarstats.net/> (2019, accessed 8 March 2019).
- 478 38. Cicchetti D V. Guidelines , Criteria , and Rules of Thumb for Evaluating Normed and
479 Standardized Assessment Instruments in Psychology. 1993; 284–290.
- 480 39. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. 2nd ed. New York:
481 Routledge, 1988.
- 482 40. IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 24.0. Armonk
483 NIC. SPSS.
- 484 41. Dietrich CF, Barr RG, Farrokh A, et al. Strain Elastography- How To Do It?
485 *Ultrasound Int Open* 2017; 3: 137–149.
- 486 42. Kha KS, Chien PFW. Evaluation of a clinical test. I: Assessment of reliability. *BJOG*
487 *An Int J Obstet Gynaecol* 2001; 108: 562–567.
- 488 43. Mulabecirovic A, Vesterhus M, Gilja OH, et al. In Vitro Comparison of Five Different
489 Elastography Systems for Clinical Applications, Using Strain and Shear Wave
490 Technology. *Ultrasound Med Biol* 2016; 42: 2572–2588.
- 491 44. Ohrndorf S, Naumann L, Grundey J, et al. Is musculoskeletal ultrasonography an
492 operator-dependent method or a fast and reliably teachable diagnostic tool? Interreader
493 agreements of three ultrasonographers with different training levels. *Int J Rheumatol*;
494 2010. Epub ahead of print 2010. DOI: 10.1155/2010/164518.

- 495 45. Royal College of Radiologists. *Standards for the provsion of an ultrasound service*.
496 2014.
- 497 46. Martinoli C, Bianchi S, Derchi LE. Tendon and nerve sonography. *Radiol Clin North*
498 *Am* 1999; 37: 691–711.
- 499 47. Yoon JH, Kim MH, Kim EK, et al. Interobserver variability of ultrasound
500 elastography: How it affects the diagnosis of breast lesions. *Am J Roentgenol* 2011;
501 196: 730–736.
- 502 48. Hauzeur J-P, Kaiser M-J, Bentin J, et al. Non omeract -expert rheumatologists
503 sonographers and inter-ultrasound machine reliability of the omeract ultrasound
504 scoring in rheumatoid arthritis in a clinical-based setting: a belgian study. 2017; 97:
505 736.1-736.
- 506 49. D’Agostino MA, Conaghan PG, Naredo E, et al. The OMERACT Ultrasound Task
507 Force - Advances and priorities. *J Rheumatol* 2009; 36: 1829–1832.
- 508 50. Ok N, Agladioglu K, Gungor HR, et al. Strain Ratio Measurements of Patellar and
509 Achilles Tendons With Different Reference Regions in Healthy Volunteers.
510 *Ultrasound Med Biol* 2019; 45: 2027–2033.

511

Table 1. Colour scoring intra and inter-operator agreement

Equipment	Operator	Measure	PQT		DQT		PPT		MPT		DPT		Total %
			%	k	%	k	%	k	%	k	%	k	
Intra-operator													
GE	1	1 + 2	100	c	100	1.0	100	1.0	90	0.78	95	0.79	97
GE	2	1 + 2	65	0.21	70	0.40	60	-0.25	75	0.38	95	0	73
Mean %			81		85		80		83		95		85
Inter-operator													
GE	1 + 2	1	65	nc	55	0.19	80	nc	70	0.12	95	0	73
GE	1 + 2	2	70	0	50	0.50	80	0	75	0	95	0	74
Mean %			68		53		80		73		95		74
Intra-operator													
esaote	1	1 + 2	75	0.5	90	0.79	70	0.34	95	0	70	0.34	80
esaote	2	1 + 2	50	0	60	0.20	90	0.45	80	0.22	90	0.46	74
Mean %			63		75		80		88		80		77
Inter-operator													
esaote	1 + 2	1	75	0.39	55	0.12	95	nc	90	0	70	0.34	77
esaote	1 + 2	2	50	0	50	0.07	90	nc	100	c	70	0.34	72
Mean %			63		53		93		95		70		75

* *PQT; proximal quadriceps tendon, DQT; distal quadriceps tendon, PPT; proximal patellar tendon, MPT; mid patellar tendon, DPT; distal patellar tendon, %; observed agreement, k; kappa, c; constant, nc; not calculated.*

Table 2. Elasticity ratio intra and inter-operator repeatability

Equipment	Operator	Measure	PQT		DQT		PPT		MPT		DPT				
			ICC	Sig	ICC	Sig	ICC	Sig	ICC	Sig	ICC	Sig			
Intra-operator															
GE	1	1 + 2	0.49	0.012	0.43	0.025	0.85	0.000	0.13	0.295	0.91	0.000			
GE	2	1 + 2	0.17	0.238	0.01	0.485	0.40	0.028	0.21	0.179	0.22	0.169			
esaote	1	1 + 2	0.27	0.112	0.51	0.009	0.48	0.015	0.67	0.001	0.15	0.265			
esaote	2	1 + 2	0.35	0.06	0.20	0.197	0.46	0.016	0.41	0.034	0.57	0.004			
Inter-operator			t	df	Sig	t	df	Sig	Z	Sig	z	Sig	t	df	Sig
GE	1 + 2	1	4.96	19	0.000	7.94	19	0.000	-0.13	0.896	-2.93	0.03	2.26	19	0.36
GE	1 + 2	2	1.81	19	0.086	5.85	19	0.000	-0.30	0.765	-1.98	0.048	1.78	19	0.09
esaote	1 + 2	1	-0.57	19	0.575	-1.29	19	0.213	-0.78	0.455	-2.11	0.035	-1.19	19	0.25
esaote	1 + 2	2	-0.64	19	0.533	-1.77	19	0.092	-0.50	0.615	-2.01	0.044	-1.52	019	0.145
* <i>PQT</i> ; proximal quadriceps tendon; <i>DQT</i> , distal quadriceps tendon, <i>PPT</i> ; proximal patellar tendon, <i>MPT</i> ; mid patellar tendon, <i>DPT</i> ; distal patellar tendon, <i>ICC</i> ; intra class coefficient, <i>sig</i> ; significance, <i>t</i> ; <i>T</i> -Test statistic, <i>df</i> ; degrees of freedom, <i>z</i> ; Wilcoxon signed-rank test statistic; Bold text indicates <i>p</i> <0.05.															

Table 3. Colour scoring inter-machine agreement

Equipment	Operator	Measure	PQT		DQT		PPT		MPT		DPT		Total %
			%	k	%	k	%	k	%	k	%	k	
GE + esaote	1	1	65	nc	75	0.53	100	nc	85	-0.71	65	nc	78
GE + esaote	1	2	50	nc	75	0.55	100	nc	100	nc	65	-0.09	78
GE + esaote	2	1	45	-0.34	45	-0.06	75	-0.87	65	0.5	60	-0.09	58
GE + esaote	2	2	70	0.12	60	0.2	70	-0.15	75	nc	65	nc	68
Mean %			58		64		86		81		64		71
* PQT; proximal quadriceps tendon, DQT; distal quadriceps tendon, PPT; proximal patellar tendon, MPT; mid patellar tendon, DPT; distal patellar tendon, %; percentage of observed agreement, k; kappa, c; constant, nc; not calculated.													

Table 4. Elasticity ratio inter-machine association

Equipment	Operator	Measure	PQT		DQT		PPT		MPT		DPT	
			<i>r</i>	Sig	<i>r</i>	Sig	<i>r</i>	Sig	<i>r</i>	Sig	<i>r</i>	Sig
GE + esaote	1	1	0.13	0.575	-0.06	0.794	-0.26	0.267	-0.09	0.711	-0.34	0.138
GE + esaote	1	2	-0.01	0.963	-0.23	0.324	-0.29	0.224	-0.21	0.369	-0.12	0.619
GE + esaote	2	1	0.13	0.574	-0.05	0.846	-0.36	0.122	-0.03	0.901	-0.163	0.493
GE + esaote	2	2	-0.01	0.983	0.11	0.964	-0.39	0.092	0.07	0.756	-0.18	0.441
* PQT; proximal quadriceps tendon; DQT, distal quadriceps tendon, PPT; proximal patellar tendon, MPT; mid patellar tendon, DPT; distal patellar tendon, <i>r</i> ; Pearson's or Spearman correlation coefficient, sig; significance.

intermediate colour score grade. C and D, Elasticity ratio measurement for proximal quadriceps tendon (GE system, C) and distal patellar tendon (esaote system, D).