A peer-reviewed version of this preprint was published in PeerJ on 2 July 2019.

View the peer-reviewed version (peerj.com/articles/7120), which is the preferred citable publication unless you specifically need to cite this preprint.

Drazan JF, Hullfish TJ, Baxter JR. 2019. An automatic fascicle tracking algorithm quantifying gastrocnemius architecture during maximal effort contractions. PeerJ 7:e7120 https://doi.org/10.7717/peerj.7120
An automatic fascicle tracking algorithm quantifying gastrocnemius architecture during maximal effort contractions

John F Drazan Corresp., 1, Todd J Hullfish 1, Josh R Baxter Corresp. 1

1 Department of Orthopaedic Surgery, University of Pennsylvania, Philadelphia, Pennsylvania, United States
Corresponding Authors: John F Drazan, Josh R Baxter
Email address: john.drazan@uphs.upenn.edu, josh.baxter@uphs.upenn.edu

Background. Ultrasound has become the gold-standard for making dynamic measurements of muscle structure during functional movements in biomechanical studies. Manual measurements of fascicle length and pennation angle are time intensive which limits the clinical utility of this approach while also limiting sample sizes. The purpose of this study was to develop a novel tracking paradigm to quantify individual fascicle length and pennation measurements during maximal voluntary contractions and demonstrate is repeatability between days and reproducibility between different examiners.

Methods. Five healthy young adults performed maximal isokinetic contractions at 0, 30, 120, 210, and 500 degrees about their ankle on an isokinetic dynamometer while their gastrocnemius muscle was observed using ultrasound. Individual muscle fascicles were identified in the first frame, and tracked using the automatic fascicle tracking algorithm and a manual approach by three observers on three separate days. Repeatability within examiners across days and reproducibility across examiners and days was evaluated using intraclass correlation coefficients. Agreement between manual and automatic tracking was evaluated using the coefficient of multiple correlations. Supervised automatic tracking was performed on all videos by one examiner to evaluate the fidelity of automatic tracking in practice.

Results. We found both manual and automatic measurements of fascicle length and pennation angle to be strongly repeatable within examiners and strongly reproducible across examiners and days (ICCs>0.76). There was greater agreement between manual and automatic measurements of fascicle length than pennation angle, however the mean CMC value for both was still found to be strong in both cases (CMC>0.8). Supervision of automatic tracking greatly showed very strong agreement between manual and automatic measurements of fascicle length and pennation angle (CMC>0.94).

Conclusions. We have developed a novel automatic fascicle tracking algorithm that quantifies fascicle length and pennation angle of individual muscle fascicles during dynamic contractions across a range of velocities. We demonstrated that this fascicle tracking algorithm is repeatable and reproducible across different examiners and different days and showed strong agreement with manual measurements, especially when tracking is supervised by the user so that tracking can be reinitialized if poor tracking fidelity is observed.
An automatic fascicle tracking algorithm quantifying gastrocnemius architecture during maximal effort contractions

John F Drazan¹, Todd J Hullfish¹, Josh R Baxter¹

¹Department of Orthopedic Surgery, University of Pennsylvania, Philadelphia, Pennsylvania, United States

Corresponding Author:
Josh R Baxter¹
Human Motion Lab, 3737 Market Street, Philadelphia, Pennsylvania, 19104, United States
Email address: josh.baxter@uphs.upenn.edu

Abstract

Background. Ultrasound has become the gold-standard for making dynamic measurements of muscle structure during functional movements in biomechanical studies. Manual measurements of fascicle length and pennation angle are time intensive which limits the clinical utility of this approach while also limiting sample sizes. The purpose of this study was to develop a novel tracking paradigm to quantify individual fascicle length and pennation measurements during maximal voluntary contractions and demonstrate is repeatability between days and reproducibility between different examiners.

Methods. Five healthy young adults performed maximal isokinetic contractions at 0, 30, 120, 210, and 500 degrees about their ankle on an isokinetic dynamometer while their gastrocnemius muscle was observed using ultrasound. Individual muscle fascicles were identified in the first frame, and tracked using the automatic fascicle tracking algorithm and a manual approach by three observers on three separate days. Repeatability within examiners across days and reproducibility across examiners and days was evaluated using intraclass correlation coefficients. Agreement between manual and automatic tracking was evaluated using the coefficient of multiple correlations. Supervised automatic tracking was performed on all videos by one examiner to evaluate the fidelity of automatic tracking in practice.

Results. We found both manual and automatic measurements of fascicle length and pennation angle to be strongly repeatable within examiners and strongly reproducible across
examiners and days (ICCs>0.76). There was greater agreement between manual and automatic measurements of fascicle length than pennation angle, however the mean CMC value for both was still found to be strong in both cases (CMC>0.8). Supervision of automatic tracking greatly showed very strong agreement between manual and automatic measurements of fascicle length and pennation angle (CMC>0.94).

**Conclusions.** We have developed a novel automatic fascicle tracking algorithm that quantifies fascicle length and pennation angle of individual muscle fascicles during dynamic contractions across a range of velocities. We demonstrated that this fascicle tracking algorithm is repeatable and reproducible across different examiners and different days and showed strong agreement with manual measurements, especially when tracking is supervised by the user so that tracking can be reinitialized if poor tracking fidelity is observed.

**Introduction**

Muscle shortening dynamics govern locomotor function in athletes, the elderly, and many patient populations (Kumagai et al., 2000; Abe, Kumagai & Brechue, 2000; Suzuki, Bean & Fielding, 2001; Mulier et al., 2003; Randhawa & Wakeling, 2013). The plantarflexors, despite their relatively small size compared to the muscles of the hip and knee, play a critical role in ambulation. These muscles behave in a variety of different ways to minimize the energy expenditure needed to complete functional activities. During the stance phase of walking, the plantarflexors act isometrically to facilitate elastic energy storage and return in the Achilles tendon (Fukunaga et al., 2001). However, running requires increased rates of shortening to do the positive work necessary to accelerate the body (Lichtwark, Bougoulias & Wilson, 2007). While computational models provide critical insight into muscle-tendon dynamics in response to small changes to the system (Nagano et al., 2007; Baxter & Hast, 2019), coupled muscle shortening and rotation is described to have a complex ‘gearing’ that is dependent on both load and speed (Azizi, Brainerd & Roberts, 2008). Therefore, experimental measurements of muscle shortening dynamics are critical for both understanding the movement biomechanics of different populations while also serving to improve and validate musculoskeletal models.

Structure-function relationships can be described by quantifying both muscle shortening dynamics and joint kinetics (Maganaris, 2003; Arampatzis et al., 2005). Ultrasound imaging is the gold-standard for quantifying skeletal muscle structure and shortening dynamics in human subjects
during functional tasks (Franchi et al., 2018). Isokinetic dynamometry provides a unique framework for measuring joint torques generated during isolated movements while controlling for load or velocity. The combination of these two measurement techniques enable researchers to study the link between muscle structure and function that underpins musculoskeletal modeling (Reeves & Narici, 2003; Urayama et al., 2009; Blazevich et al., 2009; O’Brien et al., 2010; Randhawa, Jackman & Wakeling, 2013).

Muscle shortening dynamics in pennate muscles are quantified using ultrasound by measuring changes in length and pennation of individual fascicles in a muscle bell. However, analyzing ultrasound images acquired during functional tasks has proven to be time intensive and technically challenging. Fascicle length and pennation have traditionally been manually digitized using custom-written computer software (Maganaris, 2001; Magnusson et al., 2001; Loram, Maganaris & Lakie, 2006; Barber, Barrett & Lichtwark, 2011), but this approach is a time intensive process when analyzing high frame rate ultrasound data. Automatic tracking routines have been developed to make these measurements using image processing algorithms (Cronin et al., 2011; Gillett, Barrett & Lichtwark, 2013; Zhou, Chan & Zheng, 2015). Cronin and colleagues (2011) leveraged an optical flow algorithm to quantify fascicle length and made the analysis software freely available (Farris & Lichtwark, 2016). This contribution has been a major advancement in the field of functional muscle imaging but improvements can still be made to this paradigm to increase its usability and efficiency.

Therefore, the purpose of this study was to develop and validate a novel fascicle tracking program. We manually and automatically tracked the aponeuroses and an individual muscle fascicle during maximal voluntary contractions performed in an isokinetic dynamometer throughout a range of angular velocities. We used this data to evaluate the repeatability, reproducibility, and agreement between automatic and manual measurements of fascicle length and pennation angle across three examiners across three days.

**Materials & Methods**

**Study Overview**

This study had two discrete activities: first, we developed a fascicle tracking algorithm; and second, we established the validity of this tracking algorithm for quantifying muscle structure during maximal effort contractions. We evaluated the performance of automatic tracking...
compared to manual tracking during a variety of plantarflexion contractions in healthy young adults. We acquired dynamometer and ultrasound data while subjects performed several isometric and isokinetic maximal-effort plantarflexion contractions. Three examiners then analyzed the ultrasound data using both the automatic tracking program and a manual tracking program multiple times across different days. We compared the measurements made between days and investigators to test the intra-examiner repeatability and inter-examiner reproducibility of quantifying fascicle length and pennation angle for each approach. We tested the agreement between the automatic and manual tracking approaches by using coefficients of multiple correlation (CMC). In this study, we define repeatability as the agreement between repeated measurements of the same data using identical methods and reproducibility as the agreement of repeated measures of the same data made by different observers and/or different methods (Bartlett & Frost, 2008).

**Fascicle Tracking Algorithm**

We developed a custom software tool to identify and track a single fascicle during a contraction (Supplemental Material) using MATLAB and the Computer Vision Toolbox (Nanick MA, Mathworks). The workflow of this tool can be broken into two phases: manual identification and automatic tracking. The first phase of automatic tracking required a user to identify the deep and superficial aponeuroses of the muscle belly as well as a single muscle fascicle in the first frame of a trial. The user did this by drawing a line over each of these structures. After the user confirmed that these drawn lines correctly identify the structures of interest the automatic tracking phase begins. A region of interest was defined around each of the lines and 100 trackable points were identified within each region (detectMinEigenFeatures). For the fascicle, the region of interest was limited to the middle 80% of the line to avoid identifying points that could be part of the aponeuroses. The position of these points was then tracked through each frame using a Kanade-Lucas-Tomasi algorithm. When a frame advances, the algorithm attempted to identify the new positions of each point in the new frame. The tracking quality of each point was ranked between 0 and 1 based on how similar the point was compared to the previous frame as well as the maximum distance between the point’s position in the two frames. If the quality dropped below 0.9, the point is removed from the set of trackable points. In the event that more than 10% of the points were no longer trackable, 100 new points were identified in the last high quality frame and tracking continues. A line was fit to these points in each frame and was used to describe the position and orientation of each of the identified structures (**Figure 1**). The intersections of the fascicle line and
the aponeurotic lines were defined to be the tracked fascicles insertions. Fascicle length was defined as the point to point distance between the insertion of the fascicle line into the superficial and deep aponeuroses. Pennation angle was calculated as the angle between the fascicle and the deep aponeurosis. The software visualized these lines overlaid on the ultrasound video and allowed users to accept the automatic tracking results, reprocess the trial, or validate the automatic tracking results by manually identifying the same fascicle throughout a user-defined set of frames.

In this study, we performed both automatic and manual tracking. Automatic tracking could be performed in two ways, either supervised in which the examiner can reinitialize the program if the observed fidelity of the track was poor, or unsupervised where the first tracking attempt was accepted, regardless of the observed fidelity of the track. Manual measurements of fascicle length and pennation angle were collected for six, evenly spaced frames across each trial to serve as a comparison for the automatic tracking.

Data Acquisition

Five healthy adults participated and provided written-informed consent in this study approved by the University of Pennsylvania IRB (828374). Subjects were positioned prone on a treatment table rigidly attached to a multi-mode dynamometer (System 4, Biodex, Shirley, NY, USA). The subject’s right foot was secured to a foot plate with the medial malleolus of the ankle aligned with the dynamometer’s spindle. An ultrasound transducer was affixed to the lower leg over the mid-belly of the medial gastrocnemius using a custom-made cast (Supplemental Material). Ultrasounds frames were collected with a 6 cm transducer (LV7.5/60/128Z-2, SmartUs, TELEMED) at a rate of 60 Hz. Dynamometer and ultrasound data were acquired simultaneously while subjects performed isometric and isokinetic maximal contractions at 30, 120, 210, and 500 degrees per second. We provided subjects verbal encouragement and visual feedback to help maximize effort during each contraction. Subjects were asked to perform multiple maximal effort contractions during each velocity condition until they produced similar peak torques for three consecutive contractions, which typically took 3-5 trials. We analyzed the final three trials in this study.

Evaluation of Repeatability, Reproducibility, and Agreement Between Manual and Automatic tracking
Three examiners tracked the ultrasound data for the five subjects on three separate days to
test intra-examiner repeatability and inter-examiner reproducibility for both manual and automatic
approaches. In this portion of the study, unsupervised automatic tracking was used to test the
program as a worst case scenario. Each examiner was trained on how to identify fascicles and how
use the program on a sample ultrasound video. Each examiner independently performed both
manual and automatic tracking during each session to provide a comparison between each method.
Manual tracking of took place following automatic tracking a given video and the examiner was
provided with a visual marker indicating the location of the deep insertion of the automatically
tracked fascicle. Examiners were blinded to the identifying data for each trial within a subject to
prevent bias. Automatic measurements of fascicle length and pennation angle were then extracted
at the indices corresponding to manual measurements for comparison between the two methods.
Manual tracking was an intensive process and mistakes in manual tracking occurred very rarely
due to errant mouse clicks. Eight instances (out of 4,050 individual fascicle measurements) of
extreme user error (difference between manual and automatic measurements of fascicle length
greater than 40 mm) were identified and removed from analysis.

Similar studies have observed a persistent offset between repeated automatic
measurements of fascicle length which reflect differences in user identified fascicles in the first
frame (Cronin et al., 2011; Gillett, Barrett & Lichtwark, 2013). We corrected for this offset by
subtracting the offset from each measurement so that measurements shared the same initial value.
As such, reproducibility of automatic measurements of fascicle length and pennation angle were
reported for both uncorrected and corrected CMC values.

We performed several correlation analyses to quantify the repeatability within examiners
between days and reproducibility across examiners. The intra-examiner repeatability between days
and inter-examiner reproducibility across days was calculated for both manual and automatic
measurements using intra-class correlation coefficients (ICC) in a two way mixed effects model
(McGraw & Wong, 1996). We tested the absolute agreement between individual measurements
(A-1 formulation) of fascicle length and pennation angle. To test intra-examiner repeatability,
ICCs were calculated within examiners across days. To test inter-examiner reproducibility, ICCs
were calculated across examiners and across days. We also tested the reproducibility of automatic
tracking by calculating the CMC values for each trial across examiners and days for both
uncorrected and corrected values. To test the agreement between manual and automatic tracking,
we calculated the CMC value for each individual trial comparing manual to automatic tracking for
measurements of fascicle length and pennation angle (Kadaba et al., 1989; Queen, Gross & Liu,
2006). Both ICC and CMC tests produce r values ranging between 0 and 1 where higher values
representing greater agreement between measurement methods. Specifically, r values between 0
to 0.36, 0.36 to 0.67, 0.67 to 0.9, and 0.9 to 1.0 represent poor, moderate, strong, and very strong
correlations respectively (Taylor, 1990).

The evaluate whether the supervised automatic tracking algorithm was a suitable
replacement for manual fascicle tracking, examiner 1 used the program to track all five subjects
with the ability to re-initialize the tracking if poor tracking was observed. Agreement between
automatic and manual tracking during the supervised approach was evaluated using CMCs. The
Bland-Altman method of differences was used to evaluate the agreement between the two
approaches (Bland & Altman, 1986). Bland-Altman analysis does not have a defined threshold
for statistical acceptance and instead relies on establishing an acceptable threshold a priori based
on application specific requirements (Giavarina, 2015). A recent review reported coefficient of
variation ranging from 0 to 9.8% for measurements of either fascicle length and pennation angle
in the human gastrocnemius across a range of measurement approaches (Kwah et al., 2013). As
such, we established an a priori coefficient of variation value of 10% for both fascicle length and
pennation angle.

Results

Manual and automatic measurements of fascicle length and pennation angle were found to
be repeatable with strong intra-examiner agreement (ICC>0.84, Table 1). These measurements
were also found to be reproducible with strong inter-examiner agreement across all examiners and
days (ICC>0.76, Table 1). Fascicle length measurements were more reproducible than
measurements of pennation angle. The automatic measurements were less reproducible than the
manual measurements, however, automatic tracking reproducibility remained strong across
different examiners and different days (ICC > 0.76).

We found that the reproducibility of automatic tracking was strong for uncorrected
measurements and very strong in corrected measurements of fascicle length (Figure 2 A/B) and
pennation angle measurements (Figure 3 A/B). Corrected values for fascicle measurements had a
higher mean CMC value with a smaller standard deviation (CMC = 0.98 ± 0.02) relative to the
uncorrected values (CMC = 0.88 ± 0.09). Automatic tracking reproducibility of pennation angle were also improved following initial bias correction (CMC = 0.92 ± 0.04) compared to uncorrected values (CMC = 0.84± 0.1).

Supervised tracking improved and manual measurements of both fascicle length and pennation angle across all individual trials were in strong agreement (Table 2). Supervised automatic tracking demonstrated very strong agreement between manual and automatic measurements of both fascicle length and pennation angle (Table 2). Supervision sharply reduced the incidence of tracking trials with poor and moderate agreement with over 90% of fascicle length measurements and over 80% of pennation angle measurements having very strong agreement between manual and automatic approaches (Figure 4).

The coefficient of variation for supervised measurement of fascicle length and pennation angle was 11% and 8.6% respectively (Figure 5). The coefficient of variation for pennation angle fell slightly outside of our a priori threshold of 10% while fascicle length met this criterion. Despite the strong agreement between the automatic and manual tracking approaches, the automatic tracking under-approximated fascicle lengths by 5.6% and over approximated pennation angles by 10.1% across the entire range of motion. Supervision of automatic tracking decreased these errors to less than -1.1% for fascicle length and 5% for pennation angle.

**Discussion**

The purpose of this study was to establish the reproducibility, repeatability, and agreement of an automatic fascicle tracking algorithm developed by our group. This algorithm was tested with five subjects during maximal voluntary contractions performed in an isokinetic dynamometer throughout a range of speeds. Our results indicate that this automatic tracking approach is repeatable, reproducible, and had strong agreement with manual measurements for three different examiners across three different days. We showed that supervision of the automatic tracking and reinitializing the program as needed provided accurate measurements of both fascicle length and pennation angle in one tracking session. The reliability of the automatic tracking coupled with its speed relative to manual tracking makes this approach an attractive method for studying muscle geometry during maximal effort contractions.

The agreement between automatic and manual fascicle length measurements compared well with the literature. Our repeatability measurements fall within previously reported values.
(0.62 to 0.99) for repeatability of measuring muscle architecture parameters both manually and automatically (Kwah et al., 2013). Previous groups have reported mean uncorrected CMC values for fascicle length measurements of 0.88 and 0.9 which compares well with our value of 0.89 (Cronin et al., 2011; Gillett, Barrett & Lichtwark, 2013). We observed that measurements of pennation angle had less agreement than measurements of fascicle length which has also been observed by previous groups (Aggeloussis et al., 2010; Kwah et al., 2013). The repeatability of automatic fascicle length measurements for uncorrected and corrected values was 0.88 ± 0.09 and 0.98 ± 0.02 respectively. This compare well with previously reported values for a similar Kanade-Lucas tracking approach 0.88 ± 0.08 and 0.98 ± 0.02 for uncorrected and corrected values respectively (Cronin et al., 2011). While there are fewer examples of automatic pennation angle measurement in the medial gastrocnemius, one group (Zhou, Chan & Zheng, 2015) reported an average correlation of r=0.935 which compares well with our value of r=0.94 for supervised pennation measurement. While the coefficient of variation for pennation measurements fell slightly outside of our a priori value of 10%, an 11% coefficient of variation still provides researchers with a useful tool for the automatic tracking of fascicles.

We found that the automatic tracking was reproducible across days and examiners as demonstrated by strong agreement (ICC values>0.76). While this value was lower than the reproducibility of our manual measurements (>0.9), it represents the worst case scenario as the program was unsupervised. While ICCs were not calculated for the supervised algorithm, the high correlation between manual and automatic measurements (CMC>0.94) provides evidence supervising automatic tracking will enhance reproducibility. We should note that two out of the three examiners who analyzed the ultrasound images were relative novices. While these examiners were trained by an experienced examiner, novices have been shown to be less reliable than expert and automatic tracking approaches (Miyoshi et al., 2009). However, our findings show that even novice examiners have very strong reproducibility when manually measuring fascicle length (ICC>0.9) and pennation angle (ICC=0.9).

The automatic measurements were less reproducible than manual measurements in contrast to a previous study (Miyoshi et al., 2009). Automatic measurements were observed to be effected by the initial fascicle that was identified in the first frame. At times, the selected fascicle would move over a vein or another stationary structure, or the fascicle would appear to move out of frame. This would cause the tracked fascicle to “lag” and at times track the fascicle on either side of the...
initial fascicle. During unsupervised tracking, the first tracking attempt was accepted regardless of fidelity, which we believe contributed to the lower agreement between manual and automatic tracking (CMC>0.83). Supervision removes this issue because it allows the user to simply identify a different fascicle that does not have this issue as demonstrated by the very strong agreement between manual and automatic measurements (CMC>0.94).

Tracking fidelity was dependent on ultrasound acquisition rate and quality. During pilot testing, we acquired images at nearly half the frame-rate of our reported data and found that the increased motion artifact during faster contractions increased tracking errors. Based on these findings, we suggest that ultrasound images should be acquired as quickly as possible based on specific hardware specifications. Recent advances in low-cost ultrasonography have increased the possible capture rates to well above 100 Hz, which we posit will further improve tracking fidelity. Additionally, we found that inter-subject muscle variability affected tracking fidelity. Most notably, subjects with clearly identifiable fascicles that remained visible throughout the entire contraction tracked better than trials that became less clear in deeper contraction. Tracking fidelity was partly dependent on defining the fascicle in the first frame. A potential improvement to this approach would be the development of a process to automatically identify the fascicle in the first frame to reduce variability.

This study was affected by several limitations. Each trial contained between 50 and 150 frames which was effectively down sampled to six measurements. Our rational for the limited number of manual measurements was to prioritize the number of subjects (n=5), rotational velocities (n=5), number of trials (n=3), and number of repeated measurements (n=3) that each examiner (n=3) analyzed. While other validation studies have had a larger number of manual measurements for each ultrasound video, these studies had a limited number of subjects (Miyoshi et al., 2009), a limited number of conditions (Zhou, Chan & Zheng, 2015), or digitized only a subset of the ultrasound videos (Cronin et al., 2011; Gillett, Barrett & Lichtwark, 2013). To our knowledge, this presented work is one of the most rigorous examinations on the repeatability and reproducibility of fascicle length and pennation angle measurements for both manual and automatic methods in the medial gastrocnemius (Kwah et al., 2013).

Conclusions

Accurately quantifying muscle architecture from ultrasound imaging during maximal effort tasks provides researchers and clinicians with an important tool for understanding the
structure-function relationship that underpin muscular injury, disease, and recovery. This study
introduced a novel tracking algorithm that quantifies architectural parameters of individual muscle
fascicles. We demonstrated that our proposed automatic fascicle tracking algorithm is repeatable
and reproducible across different examiners and different days and showed strong agreement with
manual measurements, especially when used in a supervised manner.
References

Abe T, Kumagai K, Brechue WF. 2000. Fascicle length of leg muscles is greater in sprinters than distance runners. Medicine and science in sports and exercise 32:1125–1129.

Aggeloussis N, Giannakou E, Albracht K, Arampatzis A. 2010. Reproducibility of fascicle length and pennation angle of gastrocnemius medialis in human gait in vivo. Gait and Posture 31:73–77. DOI: 10.1016/j.gaitpost.2009.08.249.

Arampatzis A, Stafilidis S, DeMonte G, Karamanidis K, Morey-Klapsing G, Brüggemann GP. 2005. Strain and elongation of the human gastrocnemius tendon and aponeurosis during maximal plantarflexion effort. Journal of Biomechanics 38:833–841. DOI: 10.1016/j.jbiomech.2004.04.031.

Azizi E, Brainerd EL, Roberts TJ. 2008. Variable gearing in pennate muscles. Proceedings of the National Academy of Sciences of the United States of America 105:1745–1750. DOI: 10.1073/pnas.0709212105.

Barber L, Barrett R, Lichtwark G. 2011. Passive muscle mechanical properties of the medial gastrocnemius in young adults with spastic cerebral palsy. Journal of Biomechanics 44:2496–2500. DOI: 10.1016/j.jbiomech.2011.06.008.

Bartlett JW, Frost C. 2008. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound in Obstetrics and Gynecology 31:466–475. DOI: 10.1002/uog.5256.

Baxter JR, Hast MW. 2019. Plantarflexor metabolites are sensitive to resting ankle angle and optimal fiber length in computational simulations of gait. Gait & Posture 67:194–200. DOI: 10.1016/j.gaitpost.2018.10.014.

Bland JM, Altman DG. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307–310.

Blazevich AJ, Coleman DR, Horne S, Cannavan D. 2009. Anatomical predictors of maximum isometric and concentric knee extensor moment. European journal of applied physiology 105:869–878. DOI: 10.1007/s00421-008-0972-7.

Cronin NJ, Carty CP, Barrett RS, Lichtwark G. 2011. Automatic tracking of medial gastrocnemius fascicle length during human locomotion. Journal of Applied Physiology 111:1491–1496. DOI: 10.1152/japplphysiol.00530.2011.

Farris DJ, Lichtwark GA. 2016. UltraTrack: Software for semi-automated tracking of muscle fascicles in sequences of B-mode ultrasound images. Computer Methods and Programs in Biomedicine 128:111–118. DOI: 10.1016/j.cmpb.2016.02.016.

Franchi MV, Raiteri BJ, Longo S, Sinha S, Narici MV, Csapo R. 2018. Muscle Architecture Assessment: Strengths, Shortcomings and New Frontiers of in Vivo Imaging Techniques. Ultrasound in Medicine & Biology. DOI: 10.1016/j.ultrasmedbio.2018.07.010.

Fukunaga T, Kubo K, Kawakami Y, Fukashiro S, Kamehisa H, Maganaris CN. 2001. In vivo behaviour of human muscle tendon during walking. Proceedings. Biological sciences / The Royal Society 268:229–233. DOI: 10.1098/rspb.2000.1361.
Giavarina D. 2015. Understanding Bland Altman analysis. *Biochemia Medica* 25:141–151. DOI: 10.11613/BM.2015.015.

Gillett JG, Barrett RS, Lichtwark GA. 2013. Reliability and accuracy of an automated tracking algorithm to measure controlled passive and active muscle fascicle length changes from ultrasound. *Computer Methods in Biomechanics and Biomedical Engineering* 16:678–687. DOI: 10.1080/10255842.2011.633516.

Kadaba MP, Ramakrishnan HK, Wootten ME, Gainey J, Gorton G, Cochran GVB. 1989. Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. *Journal of Orthopaedic Research* 7:849–860. DOI: 10.1002/jor.1100070611.

Kumagai K, Abe T, Brechue WFWF, Ryushi T, Takano S, Mizuno M. 2000. Sprint performance is related to muscle fascicle length in male 100-m sprinters. *Journal of Applied Physiology* 88:811–816. DOI: 10.1152/japplphysiol.01430.2011.

Lichtwark Ga., Bougoulias K, Wilson a. M. 2007. Muscle fascicle and series elastic element length changes along the length of the human gastrocnemius during walking and running. *Journal of Biomechanics* 40:157–164. DOI: 10.1016/j.jbiomech.2005.10.035.

Loram ID, Maganaris CN, Lakie M. 2006. Use of ultrasound to make noninvasive in vivo measurement of continuous changes in human muscle contractile length. *Journal of Applied Physiology* 100:1311–1323. DOI: 10.1152/japplphysiol.01229.2005.

Maganaris CN. 2001. Force-length characteristics of in vivo human skeletal muscle. *Acta Physiol Scand* 172:279–285.

Maganaris CN. 2003. Force-length characteristics of the in vivo human gastrocnemius muscle. *Clinical Anatomy* 16:215–223. DOI: 10.1002/ca.10064.

Magnusson SP, Aagaard P, Dyhre-Poulsen P, Kjaer M. 2001. Load-displacement properties of the human triceps surae aponeurosis in vivo. *The Journal of physiology* 531:277–288.

McGraw KO, Wong SP. 1996. Forming inferences about some intraclass correlation coefficients. *Psychological methods* 1:30.

Miyoshi T, Kihara T, Koyama H, Yamamoto S-I, Komeda T. 2009. Automatic detection method of muscle fiber movement as revealed by ultrasound images. *Medical Engineering & Physics* 31:558–564. DOI: 10.1016/j.medengphy.2008.11.004.

Mulier T, Pienaar H, Dereymaeker G, Reynders P, Broos P. 2003. The management of chronic achilles tendon ruptures: gastrocnemius turn down flap with or without flexor hallucis longus transfer. *Foot and Ankle Surgery* 9:151–156. DOI: 10.1016/S1268-7731(03)00048-1.

Nagano A, Yoshioka S, Komura T, Fukashiro S. 2007. Influence of the Fascicle Length and Physiological Cross-sectional Area of M. soleus on Ankle Joint Kinetics. *International Journal of Sport and Health Science* 5:98–104. DOI: 10.5432/ijshs.5.98.
O’Brien TD, Reeves ND, Baltzopoulos V, Jones DA, Maganaris CN. 2010. In vivo measurements of muscle specific tension in adults and children. Experimental Physiology 95:202–210. DOI: 10.1113/expphysiol.2009.048967.

Queen RM, Gross MT, Liu H-Y. 2006. Repeatability of lower extremity kinetics and kinematics for standardized and self-selected running speeds. Gait & Posture 23:282–287. DOI: 10.1016/j.gaitpost.2005.03.007.

Randhawa A, Jackman ME, Wakeling JM. 2013. Muscle gearing during isotonic and isokinetic movements in the ankle plantarflexors. European Journal of Applied Physiology 113:437–447. DOI: 10.1007/s00421-012-2448-z.

Randhawa A, Wakeling JM. 2013. Associations between muscle structure and contractile performance in seniors. Clinical Biomechanics 28:705–711. DOI: 10.1016/j.clinbiomech.2013.04.010.

Reeves ND, Narici MV. 2003. Behavior of human muscle fascicles during shortening and lengthening contractions in vivo. Journal of Applied Physiology 95:1090–1096. DOI: 10.1152/japplphysiol.01046.2002.

Suzuki T, Bean JF, Fielding RA. 2001. Muscle Power of the Ankle Flexors Predicts Functional Performance in Community-Dwelling Older Women. Journal of the American Geriatrics Society 49:1161–1167. DOI: 10.1046/j.1532-5415.2001.49232.x.

Taylor R. 1990. Interpretation of the Correlation Coefficient: A Basic Review. Journal of Diagnostic Medical Sonography 6:35–39. DOI: 10.1177/875647939000600106.

Urayama KY, Buffler PA, Gallagher ER, Ayoob JM, Ma X. 2009. Dynamic measurement of pennation angle of gastrocnemius muscles during contractions based on ultrasound imaging. Zhou GQ, Chan P, Zheng YP. 2015. Automatic measurement of pennation angle and fascicle length of gastrocnemius muscles using real-time ultrasound imaging. Ultrasonics 57:72–83. DOI: 10.1016/j.ultras.2014.10.020.
Figure 1 (on next page)

A visual of a frame from the automated tracking algorithm.

The superior and deep aponeuroses are marked with point clouds and fit with a best fit lines in white. The identified fascicle with a point cloud in red and a best fit line it white.
Fascicle Length: 74.2 mm
Pennation Angle: 23 degrees

Blinded Review
Deep Aponeurosis: 100% (100/100)
Superficial Aponeurosis: 100% (100/100)
Fascicle: 100% (100/100)
Figure 2 (on next page)

Representative data for fascicle length measurements from a single subject

a) overlay of fascicle length measurements in one trial for all examiners across all days prior to correction for initial offset. B) overlay of fascicle length measurements in one trial for all examiners across all days after to correction for initial offset. C) mean fascicle length measurements across all examiners and days for uncorrected, corrected, and manual measurements with standard deviation.
Representative data of pennation angle measurements from a single subject

a) overlay of pennation angle measurements in one trial for all examiners across all days prior to correction for initial offset. B) overlay of pennation angle measurements in one trial for all examiners across all days after to correction for initial offset. C) mean pennation angle measurements across all examiners and days for uncorrected, corrected, and manual measurements with standard deviation.
Figure 4 (on next page)

Count density histograms for all CMC values comparing manual to automatic tracking.

Distributions of CMC values comparing manual and automatic measurements. Values were grouped by poor, moderate, strong, and very strong from left to right. A) distribution of CMC values for uncorrected fascicle lengths measurements. B) distribution of CMC values for corrected fascicle lengths measurements. C) distribution of CMC values for uncorrected pennation angle measurements. D) distribution of CMC values for corrected pennation angle measurements.
Distribution of CMC Values for Fascicle Length in Unsupervised Tracking

Distribution of CMC Values for Fascicle Length in Supervised Tracking

Distribution of CMC Values for Pennation Angle in Unsupervised Tracking

Distribution of CMC Values for Pennation Angle in Supervised Tracking
**Figure 5** (on next page)

Bland-Altman plots comparing supervised automatic tracking to manual tracking.

(A) Comparison between manual and automatic tracking for fascicle length (B) Comparison between manual and automatic tracking for pennation angle
**Fascicle Length**

- Equation: $y = 0.89x + 5.27$
- $r^2 = 0.95$
- CV: 8.6%

**Pennation Angle**

- Equation: $y = 0.79x + 6.48$
- $r^2 = 0.88$
- CV: 11%

### Comparison between Automatic and Manual Measurements

**Fascicle Length**
- Mean: 7.9 (+1.96SD)
- Mean: -7.0 (-1.96SD)

**Pennation Angle**
- Mean: 6.5 (+1.96SD)
- Mean: -10 (-1.96SD)
ICC comparisons for intra-examiner repeatability and inter-examiner reproducibility.

ICC measurements indicated strong intra-examiner repeatability and strong inter-examiner reproducibility across the three examiners and three days.
|               | Intra-Examiner ICC | Inter-Examiner ICC |
|---------------|-------------------|--------------------|
|               | Automatic | Manual | Automatic | Manual |
| Fascicle Length
| Examiner 1 | 0.915    | 0.936    | 0.863      | 0.922   |
| Examiner 2 | 0.909    | 0.949    | 0.837      | 0.933   |
| Examiner 3 | 0.837    | 0.933    |            |         |
| Pennation Angle
| Examiner 1 | 0.863    | 0.939    | 0.764      | 0.899   |
| Examiner 2 | 0.836    | 0.915    |            |         |
| Examiner 3 | 0.744    | 0.925    |            |         |
Table 2 (on next page)

CMC calculations to test agreement between manual and automatic tracking approaches.

CMC values for agreement between manual and automatic measurements for individual tracking trials of fascicle length and pennation angle for unsupervised and supervised tracking approaches.
|                      | Mean CMC (±S.D.) |
|----------------------|-----------------|
|                      | Unsupervised    | Supervised     |
| Fascicle Length      | 0.90(±0.13)     | 0.97(±0.5)     |
| Pennation Angle      | 0.83 (±0.18)    | 0.94(±0.07)    |