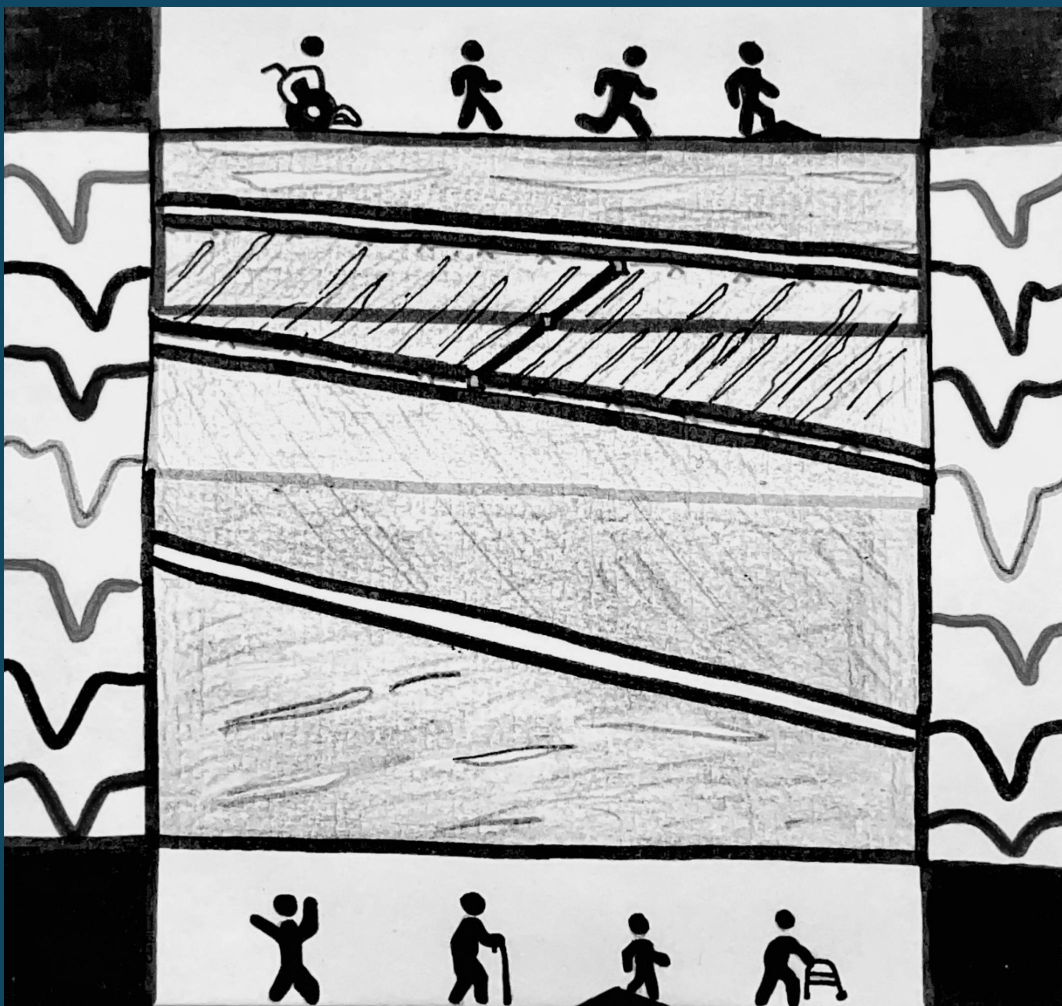
Validation of Automated Fascicle Tracking Algorithms for Gait Cycle Analysis Across Functional Gait Tasks in Children

Suvi Lotta van Hunen

15 April 2025



DELFT UNIVERSITY OF TECHNOLOGY

BIOMEDICAL ENGINEERING
MASTER THESIS

Validation of Automated Fascicle Tracking Algorithms for Gait Cycle Analysis Across Functional Gait Tasks in Children

Author:
Suvi Lotta van Hunen

In Partial Fulfilment of the Requirements of
Master of Science in Biomedical Engineering
track Medical Devices (Track II)
at the Delft University of Technology,
to be defended publicly on 15 April 2025

Thesis advisors:
Dr. ir. Winfred Mugge
Dr. Marjolein van der Krogt
Babette Mooijekind, MSc

Thesis Committee:
Dr. ir. Winfred Mugge (Chair), Department of BioMechanical Engineering, Delft University of Technology
Dr. Marjolein van der Krogt, Department of Rehabilitation at University Medical Center, Amsterdam (VUmc)
Babette Mooijekind, MSc., Department of Rehabilitation at University Medical Center, Amsterdam (VUmc)
Dr. ir. Arno Stienen, Department of BioMechanical Engineering, Delft University of Technology

Preface

Manual tracking is considered the gold standard for analysis of fascicle behavior. However, it is time-consuming, prone to human error, and inconsistent over long sequences. To address these limitations, it is essential to develop and validate automated tracking algorithms that are both accurate and computationally efficient.
This thesis evaluates the performance of two algorithms, the semi-automated Kanade-Lucas-Tomasi (KLT) and the fully automated UltraTimTrack (UTT), in tracking fascicle behavior relative to manual tracking. It also examines how functional tasks affect fascicle behavior across manual tracking and validated algorithm(s). While the UTT and KLT algorithms are expected to offer greater efficiency due to faster computational times, manual tracking is anticipated to perform better in more demanding tasks, such as functional tasks involving weights or an incline. Validating automated tracking algorithms could enhance clinical interventions and enable fast, real-time fascicle tracking in clinical settings.
This research was conducted as part of my Master of Science in Biomedical Engineering at Delft University of Technology. Understanding pediatric biomechanics is especially important, as conditions like cerebral palsy significantly affect a child's ability to move and play. Improving tracking algorithms could contribute to better interventions that help children move more freely and enjoy a more active childhood. Conducting research in a hospital setting and interacting with these children has made this work particularly meaningful. The study was conducted under the supervision of Dr. Marjolein van der Krogt and Babette Mooijekind from the Department of Rehabilitation at Amsterdam University Medical Center, and Dr. ir. Winfred Mugge from the Department of BioMechanical Engineering at Delft University of Technology.

Suvi Lotta van Hunen
Delft, April 2025

Table of Contents

Abstract—**Manual fascicle tracking from ultrasound images is the gold standard for assessing fascicle behavior. However, manual tracking is labor-intensive, prone to human error, and often inconsistent. The semi-automated Kanade-Lucas-Tomasi (KLT) and fully automated UltraTimTrack (UTT) algorithms offer potential alternatives, but their accuracy and computational efficiency require further evaluation. The current study aimed to determine which algorithm provides the most accurate fascicle tracking relative to manual tracking. Furthermore, how functional tasks including walking and running, while carrying weights or walking on an incline influence fascicle behavior across manual tracking and validated tracking algorithm(s). For the current research, treadmill walking and running data from typically developing (TD) children (n = 3) was used. The reliability of manual tracking was assessed using Root Mean Square Error (RMSE) and Intraclass Correlation Coefficients (ICC) in five trials, while computational efficiency was evaluated based on the time ratio. Gait cycles were time-normalized, and Statistical Parametric Mapping (SPM) two-way repeated measures ANOVA with post-hoc comparisons was conducted to analyze fascicle length differences across manual tracking, KLT, UTT and functional tasks. Repeated measures analysis with post-hoc comparisons in SPSS was conducted to analyze range values (max-min). Results demonstrated high reproducibility (RMSE = 4.63 mm and r = 0.78) of manual tracking, confirming its reliability as the gold standard. Compared to manual tracking, UTT exhibited superior computational efficiency and accuracy, while KLT showed greater limitations in tracking performance. Further analysis was conducted using manual tracking and UTT, where functional tasks were compared. The medial gastrocnemius muscle adapted its neuromechanical behavior across functional tasks, with walking relying on passive elastic energy storage, running emphasizing active fascicle shortening, and tasks involving weights increasing active force generation. Future research should expand participant cohorts, refine tracking algorithms, and evaluate the applicability of automated algorithms in clinical populations.**

1 Introduction

Gaining insight into the mechanics of skeletal muscle fascicles using ultrasound imaging is essential for enhancing our understanding of human movement and improving the effectiveness of clinical interventions by allowing real-time analysis of fascicle length, pennation angle, and muscle-tendon dynamics during movement [1]. The medial gastrocnemius (MG) muscle plays an essential role in walking and running and has been a key focus in studies on muscle-tendon dynamics under various conditions [2, 3]. The human gait cycle, which spans from 0 % to 100 %, is divided into two main phases: the stance phase and the swing phase, as shown in Fig. 1. The fascicle behavior of the MG in response to changes in walking speed and horizontal forces is displayed in Fig. 2a & Fig. 2b. Accurate analysis of fascicle behavior is crucial for understanding mobility outcomes in pediatric and clinical populations, particularly in children with Cerebral Palsy (CP), where a toe-walking gait pattern (CP equinus) is often present [4, 5, 6, 7]. The fascicle behavior of the MG for CP non equinus, CP equinus and for TD children is displayed in Fig. 3a, Fig. 3b & Fig. 3c. Manual fascicle tracking of the MG muscle in ultrasound images is considered a benchmark method due to its reliance on human interpretation, eliminating algorithmic biases and allowing for nuanced identification of fascicle boundaries [8, 9]. However, it is labor-intensive, prone to human error, and lacks consistency over extended sequences, making it less suitable for large-scale studies [1]. Recent advancements in ultrasound imaging and automated fascicle tracking algorithms have revolutionized this field, yet significant challenges remain in ensuring the efficiency and accuracy of these algorithms across diverse populations and movement paradigms [8, 9, 10, 11, 12].

Dynamic muscle fascicle length changes are typically estimated using B-mode ultrasound imaging, with studies leveraging both manual tracking and automated tracking algorithms to capture their real-time behavior during functional tasks [5, 15, 16]. Semi-automated approaches, using both manual fascicle length determination in the first frames and automated Kanade-Lucas (KL) optical flow algorithm in the other frames, have demonstrated utility. However, their reliance on the KL optical flow algorithm introduces variability, particularly during complex movements [8, 9]. The Kanade-Lucas-Tomasi (KLT) algorithm improves feature tracking reliability and robustness to large motions compared to the KL algorithm, thanks to its pyramidal implementation. It tracks localized tissue motion by identifying prominent features and estimating their displacement across image frames, enabling fine-grained motion analysis and automatic selection of reliable features [10, 11]. Concurrently, the active shape model (ASM) statistically models muscle shape variations using annotated landmarks and principal component analysis, ensuring anatomically accurate region-of-interest identification and dynamic adaptation to muscle deformations [10]. Together, these

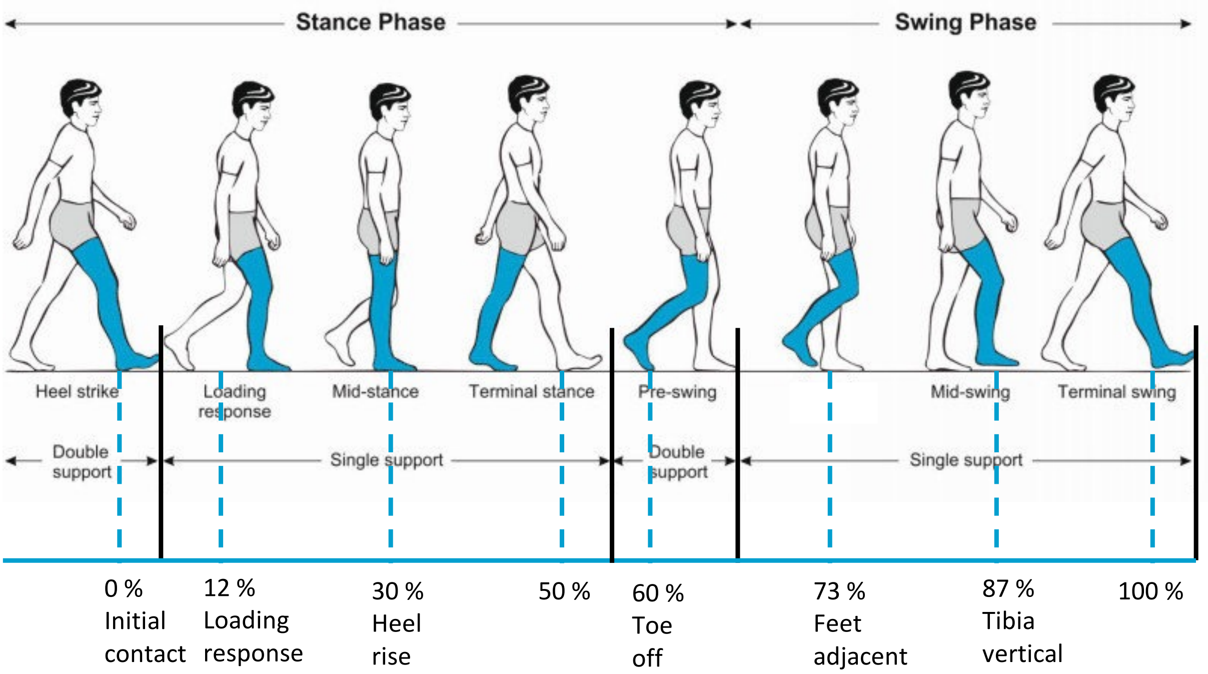


Fig. 1. One complete gait cycle (1–100 %) with corresponding phases [13, 14].

techniques provide a robust framework for capturing both detailed motion patterns and global shape dynamics, advancing automated analysis of muscle function [10]. However, the KLT algorithm requires well-defined manual initialization for accurate results and is both complex to implement and computationally demanding [10, 11]. The fully automated UltraTrack algorithm has made strides in reducing observer bias [2, 17, 18]. However, issues such as drift during large displacements or inaccuracies in regions with non-uniform muscle deformation remain unresolved [18, 19]. To mitigate drift, the fully automated TimTrack algorithm was designed to function independently of image history, effectively eliminating drift while remaining sensitive to noise. TimTrack is based on line detection techniques using filters like Frangi vessel enhancement filter and Hough Transform [19]. A recently proposed tracking algorithm, UltraTimTrack (UTT), a hybrid approach using a Kalman filter combining the drift-free line detection of TimTrack with the noise-insensitivity optical flow of UltraTrack [12, 20, 21, 22, 23]. A Kalman filter balances the strengths and weaknesses of predictions and measurements to achieve accurate and robust tracking over long sequences, estimating the state of a dynamic system over time despite noise and uncertainty [23]. These advancements in semi and fully automated tracking algorithms offer increasingly robust and accurate solutions while addressing challenges such as drift, noise, computational time and operator dependency.

While mathematical algorithms for fascicle tracking have shown promise in adult populations, their translation to pediatric applications, particularly in pathological contexts, remains limited [1, 11, 19, 23]. Therefore, identifying the most accurate fascicle tracking algorithm and understanding how functional tasks affect fascicle behavior in children hold profound medical and technological implications, including improved access to high-quality diagnostic tools [11]. Medically, the development and validation of advanced tracking algorithms will enable accurate and dynamic assessments of muscle-tendon behavior during functional tasks. As well as, facilitate early detection of musculoskeletal abnormalities and the implementation of personalized rehabilitation strategies, particularly for children with CP, who exhibit altered muscle-tendon mechanics and atypical fascicle behaviors [4, 5, 6]. Technologically, the integration of automated algorithms represents a major advancement in ultrasound-based motion analysis by reducing observer bias, improving computational efficiency, and enabling real-time applications in both research and clinical settings. Many existing

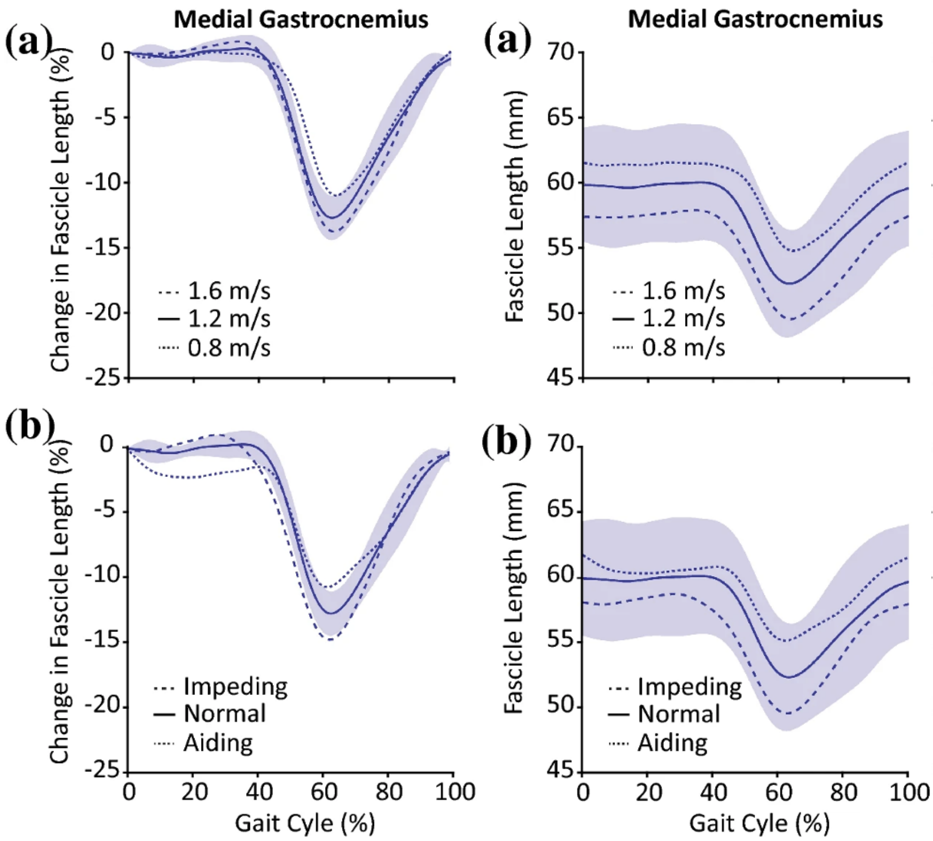


Fig. 2. The time series illustrate medial gastrocnemius fascicle behavior, change in fascicle length (left) and fascicle length (right), in response to changes in walking speed, 1.6 m/s, 1.2 m/s and 0.8 m/s (a) and horizontal forces, Impeding, Normal and Aiding (b) during a complete gait cycle (0–100 %). Fascicle length changes were normalized to their length at heel strike. The shaded region (purple) represents the standard error during normal walking (1.2 m/s) across all associated experimental trials and ten adults [2].

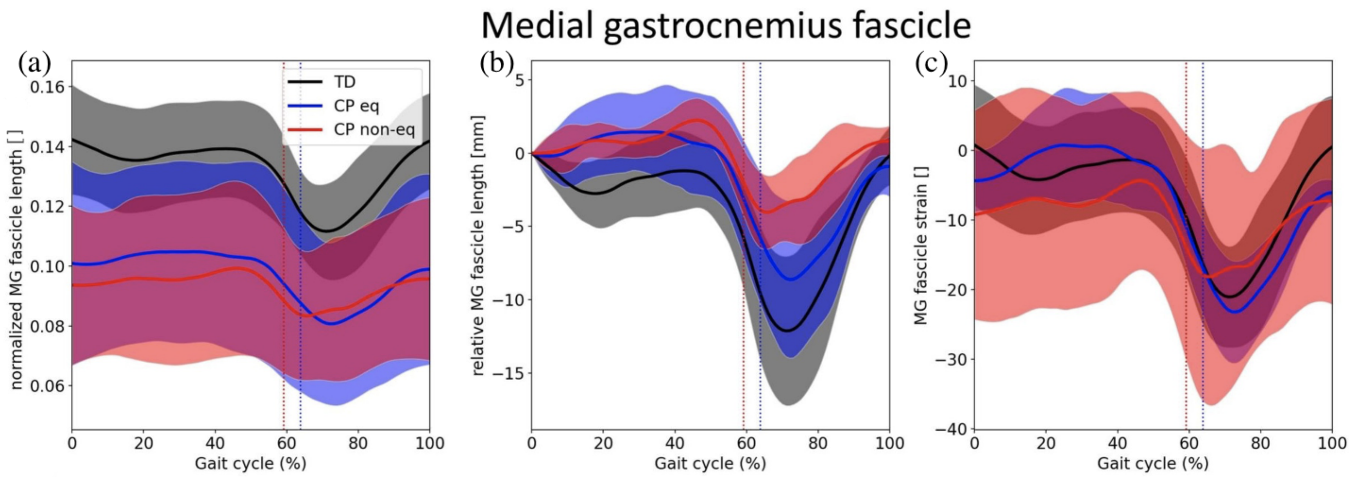


Fig. 3. The time series illustrate medial gastrocnemius fascicle behavior with the normalized (to corresponding tibia length) length in (a), relative fascicle length (absolute fascicle lengths at initial contact subtracted from the corresponding values at each point in the gait cycle) (b) and strain in (c) with calculated mean and standard deviation. Data is shown for TD (n = 12) (grey), CP equinus (n = 6) (blue) and CP non-equinus (n = 6) (red) groups during a complete gait cycle (0–100 %) [7].

automated algorithms rely on static or passive movement validation, and high-quality ultrasound images, often requiring manual initialization and feature-point tracking, making them sensitive to user input and image quality (Supplementary Table 5, 6, 7, 8, 9 and 10). Given these limitations, there is a critical need to develop and validate robust tracking algorithms that can accommodate the distinctive muscle-tendon dynamics of pediatric populations while addressing the shortcomings of manual tracking.

The current study aims to assess the accuracy and efficiency of the semi-automated Kanade-Lucas-Tomasi (KLT) and fully automated UltraTimTrack (UTT) fascicle tracking algorithms compared to manual tracking during treadmill-based walking and running, with and without weights or on an incline, in Typically Developing (TD) children. This will pave the way for more accurate and efficient tools tailored to pediatric populations and could potentially enhance diagnostic precision, improve intervention outcomes, and foster a comprehensive understanding of muscle-tendon dynamics. Key research points include: (1) assessing the intra-rater reliability of manual tracking in TD children to establish its reliability as a gold standard, (2) comparing the accuracy and efficiency of the semi-automated KLT and fully automated UTT fascicle tracking algorithms with manual tracking across functional tasks and (3) comparing the effects of various functional tasks on fascicle behavior using manual tracking and validated tracking algorithm(s).

It is hypothesized that the UTT fascicle tracking algorithm will demonstrate reduced computational time compared to manual tracking, as well as the semi-automated KLT algorithm. While the UTT and KLT algorithms are expected to be more efficient due to reduced computational time, manual tracking is anticipated to outperform in challenging cases. For instance, tasks that involve weights or an incline, as it allows for customized adjustments that the automated algorithm may not accommodate. In the second part of the research where functional tasks are compared to each other, it is hypothesized that during running and tasks involving weights or an incline, fascicle lengths in children will be shorter than in walking tasks. To test these hypotheses, several key outcome measures will be analyzed, including intra-reliability of manual tracking across five trials, fascicle length, relative fascicle length, range values and computational time. Fascicle length is directly measured from ultrasound images over multiple frames, the range is determined by identifying the maximum and minimum values within the gait cycle and the relative fascicle length is determined by subtracting the fascicle length values with the corresponding initial contact fascicle length values. Together these outcome measures provide insights into muscle adaptation, force generation, and the performance of automated tracking algorithms compared to manual tracking. Computational time, defined as the time required for manual tracking and the two algorithms, reflects the efficiency of automated algorithms relative to manual tracking. To assess differences between manual tracking, KLT and UTT algorithms and functional tasks, Statistical Parametric Mapping (SPM) repeated-measures will be conducted for time-dependent data and conventional repeated-measures will be conducted for time-independent data. This analysis will help evaluate the accuracy and efficiency of the tracking algorithms and assess how functional tasks affect fascicle behavior in children.

2 Materials & Methods

The current study utilized previously collected data from dynamic ultrasound measurements of Typically Developing (TD) children (details on participant demographics: Supplementary Table 1). Based on a previous study, a power analysis was conducted using G*Power [25], with an effect size of 0.60, a power of 0.95, two groups (TD/CP), and three conditions (with or without weights or an incline) [4]. This analysis indicated that a sample size of ten participants per group is sufficient. Each participant completed a series of functional tasks which are presented in Table 1. This study was conducted in accordance with the Declaration of Helsinki under approval number ONZ-2023-0356 and both the children and their parents provided written informed consent prior to participation.

Inclusion Criteria

- Typically developing children: Healthy children with no history of neurological disorders.

Table 1. Functional tasks performed by each child.

Functional tasks	Description
Walking	Performed at 70 % of their maximum walking speed
Running	Performed at 70 % of their maximum running speed
Pulley Walking	Weights (45 % of their body weight) attached to a pulley system connected to the child (safety harness) while walking
Pulley Running	Weights (25 % of their body weight) attached to a pulley system connected to the child (safety harness) while running
Slope Walking	Performed at 6.9° inclined treadmill and at 70 % of their maximum walking speed

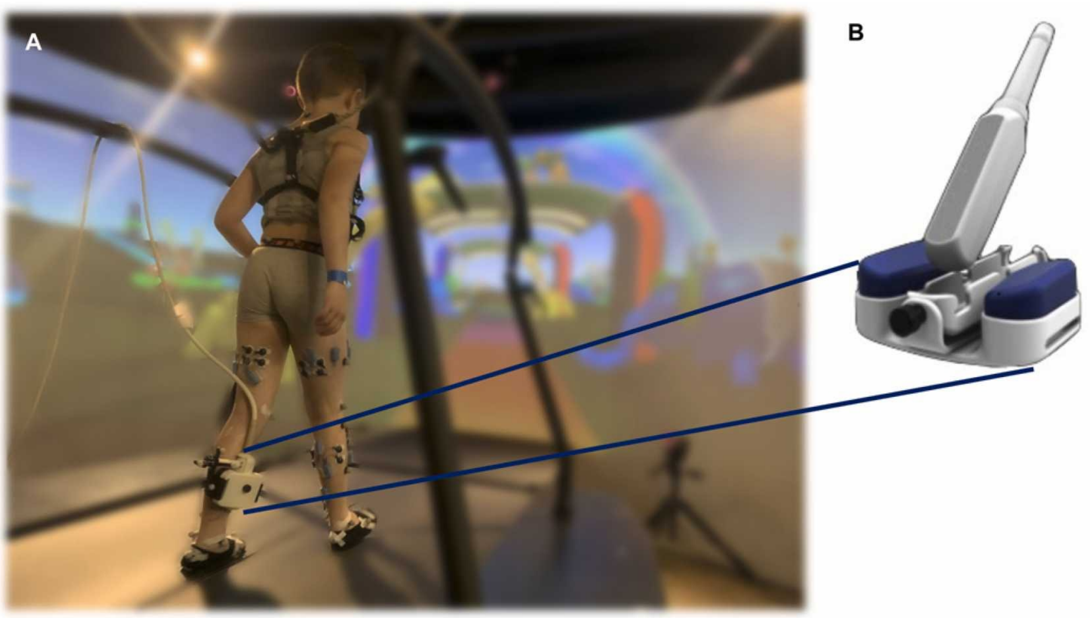


Fig. 4. Part (a) shows the setup for the measurements, where a child can walk and run on a treadmill. Part (b) displays the ultrasound probe (Probefix Dynamic T, USONO, The Netherlands) positioned at the gastrocnemius muscle-tendon junction [24].

Exclusion Criteria

- Severe joint deformities or physical conditions preventing walking or running.
- Severe cognitive impairments that interfere with task completion.

2.1 Data collection protocol

Dynamic ultrasound imaging was utilized to capture real-time images of the medial gastrocnemius muscle during dynamic movements (Fig. 4a & Fig. 4b). The ultrasound probe was positioned longitudinally along the mid-calf region, oriented parallel to the muscle fibers to optimize visualization of fascicle dynamics, allowing movements in five degrees-of-freedom. The probe (111 g) and probe-holder (219 g) have a total mass of 330 g. A coupling gel was applied to ensure acoustic transmission and minimize impedance. Imaging was performed using Dynamic 2D B-mode Ultrasonography, with a 59 mm linear US probe (Telemed SmartUS, Lithuania), a frequency of 8 MHz, a low line density and a frame rate of 60 fps optimized for dynamic musculoskeletal imaging. Participants performed a set of functional tasks on a split-belt treadmill (GRAIL, Motek ForceLink BV, The Netherlands) listed in Table 1. Each trial lasted approximately 20 s, ensuring at least ten valid steps, five selected steps were randomly selected. Atypical steps showing large variations from the expected gait cycle were excluded as well as the first and last steps as they might reflect acceleration or deceleration. Ultrasound images were recorded continuously during each trial and securely stored for subsequent analysis. Changes to probe placement and parameter calibration adjustments were made to accommodate individual anatomical variations, ensuring high-quality data acquisition across all participants.

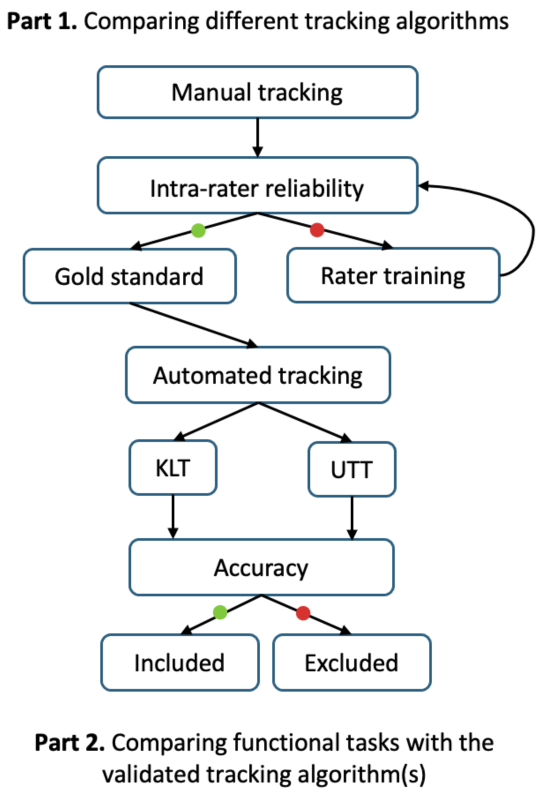


Fig. 5. Workflow for validating automated fascicle tracking algorithms across functional tasks in children. Manual tracking serves as the gold standard (green dot), with criteria of $r \geq 0.68$ and $RMSE \leq 10$ %. Automated tracking algorithms were included (green dot), if they showed no significant difference from manual tracking across multiple (at least two) tasks, while also demonstrating time efficiency. Red dots indicate instances where the specified criteria are not met.

2.2 *Data processing*

Ultrasound images were analyzed using manual tracking, the semi-automated Kanade-Lucas-Tomasi (KLT) algorithm and the fully automated Ultra-TimTrack (UTT) algorithm. For manual tracking, fascicle lengths were marked every two frames using the ImageJ 1.54g software [26]. Key considerations for manually tracking fascicle length include the placement of tracking lines along the inner edge of the aponeurosis. Additionally, utilizing bright white reflections from the perimysium along the predominantly dark fascicle edges, caused by low echogenicity, ensures consistent tracking across frames. Finally, observing the upper aponeurosis before the lower aponeurosis improves the identification of medial gastrocnemius movement direction, thereby enhancing fascicle tracking precision [27]. For the semi-automated algorithm, a point distribution model was created using data trained by manually labeling the first frames of the ultrasound image sequences collected from each participant [10]. Image segmentation was performed using the ASM model, which identified regions containing muscle fascicles, located the aponeuroses, and divided the region into longitudinal segments. The final step involved a dynamical model with Gaussian noise, which predicts shape changes across frames, with noise variance estimated from local movements tracked using the KLT feature tracker [10, 28]. The fully automated UTT refines estimates of fascicle length correcting them based on Kalman gain for improved accuracy [23]. Before data analysis, preprocessing was performed to ensure consistency in statistical evaluation. Fascicle length data were time-normalized and segmented into gait cycles using the Gait Off-line Analysis Tool (GOAT v4.2, Motek Medical, The Netherlands) and MATLAB (R2024b) algorithms.

2.3 *Data analyses*

Data analyses assessed the accuracy and efficiency of manual tracking, the semi-automated

Table 2. Characteristics of participants whose data were successfully processed without errors.

Characteristics	**TD (n = 3)**
Age (years)	11.00 ± 1.00
Weight (kg)	34.47 ± 4.97
Height (cm)	148.67 ± 1.86
Max treadmill walking speed (m/s)	2.35 ± 0.21
Max treadmill running speed (m/s)	3.90 ± 0.61

Kanade-Lucas-Tomasi (KLT) algorithm, and the fully automated UltraTimTrack (UTT) algorithm. The workflow for the current research is displayed in Fig. 5. Intra-reliability of manual tracking was determined by assessing the consistency of fascicle length measurements across the entire gait cycle obtained by the same operator across five trials using Intraclass Correlation Coefficient (ICC) and Root Mean Square Error (RMSE). Previous studies have classified correlations as weak ($r \leq 0.35$), moderate (0.36 to 0.67), and strong (0.68 to 1), a classification that will also be applied in the current study [1, 29]. To evaluate performance, previous studies that analyzed fascicle images of similar lengths (e.g., the human soleus, approximately 30 to 50 mm at rest) reported RMSE values below 10 % of the mean fascicle length. Based on this, 10 % RMSE, considered a low error, was applied in the current study [11, 30]. Each trial was analyzed twice, with an additional analysis performed on randomly selected frames, approximately on flexion points of the gait cycle, to measure fascicle length independently. Tracking efficiency was evaluated by recording the time required to complete fascicle tracking from pre-processing to final analysis. Data was checked for normality, before conducting SPM two-way repeated measures ANOVA to compare absolute and relative fascicle lengths across manual tracking and the two algorithms and functional tasks. Post-hoc pairwise comparisons were performed when significant differences were detected ($p < 0.05$). To account for temporal correlations across the time series data, Random Field Theory (RFT) correction is applied in SPM. For post-hoc comparisons, a False Discovery Rate (FDR) correction (time-dependent data) was applied to control for Type I errors while maintaining an appropriate balance between sensitivity and statistical power. ICC and RMSE were also calculated for the two automated tracking algorithms compared to manual tracking and across functional tasks. Repeated measures analysis was conducted in SPSS 29.0.2.0 (IBM Corp. in Armonk, NY) on the range values (defined as the difference between the maximum and minimum absolute fascicle lengths). All range values are displayed in Supplementary Table 2, 3 and 4. Post-hoc pairwise comparisons, with Bonferroni correction (time-independent data), were performed when significant differences were detected ($p < 0.05$).

3 Results

The results evaluate the accuracy and efficiency of manual tracking, KLT, and UTT, as well as differences in fascicle behavior across various functional tasks. Participant characteristics used in the current study are presented in Table 2.

3.1 *Reliability of manual tracking and efficiency of automated tracking algorithms*

Ultrasound data from seven TD children and slope running data (running with an incline) were excluded from the analysis due to the KLT algorithm failing to produce reliable fascicle tracking data, resulting in errors and poor tracking quality.

Manual fascicle tracking was assessed for intra-rater reproducibility to establish a reliable gold standard. The method demonstrated high reproducibility, with an average correlation coefficient of 0.78 and an average Root Mean Square Error (RMSE) of 4.63 mm (10 % of the mean fascicle length) across five trials. Computational efficiency was evaluated by calculating the time ratio between the computational time of the algorithms (KLT: 60 ± 10 min; UTT: 35 ± 5 s) and the time required for manual tracking (60 ± 15 min) per trial. Table 3 presents the computational efficiency of the algorithms, illustrating their substantial time savings compared to manual tracking.

Table 3. Presents a comparison of the processing speeds of KLT and UTT relative to manual tracking.

Time	**Manual Tracking/KLT**	**Manual Tracking/UTT**
Time ratio	1.00 ± 0.44	102.86 ± 29.60

Table 4. Average ranges (max-min) across manual tracking, KLT, and UTT algorithms for functional tasks where significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$.

Functional Tasks	**Manual Tracking Range** ($Mean \pm SD$) **(mm)**	**KLT Range** ($Mean \pm SD$) **(mm)**	**UTT Range** ($Mean \pm SD$) **(mm)**
Walking	13.71 (4.10)	20.51 (3.81)***	21.76 (3.31)***
Running	15.34 (4.19)	19.93 (4.38)*	21.28 (5.19)
Pulley Walking	15.27 (5.19)	17.17 (4.98)	20.91 (3.52)**
Pulley Running	10.53 (2.19)	13.87 (4.24)	19.67 (15.31)
Slope Walking	11.96 (2.49)	17.45 (4.80)*	19.78 (2.50)***

3.2 Accuracy across KLT and UTT compared to manual tracking

A significant overall difference ($p < 0.001$) in fascicle length and range values were found across manual tracking and the two algorithms. Further analysis revealed that the KLT algorithm significantly underestimated fascicle length and exhibited greater shortening values compared to manual tracking. Comparisons between the UTT algorithm and manual tracking showed no significant differences in fascicle length for the running and pulley running tasks showing the accuracy of UTT for dynamic movements.

Table 4 displays the average ranges, while Fig. 6 shows the average fascicle length, across manual tracking, KLT, and UTT algorithms for functional tasks. The fascicle length of each child across manual tracking and the two algorithms can be found in Supplementary Fig. 1, 2 and 3.

3.2.1 Walking trials

When comparing manual and KLT data, significant differences in fascicle length were found between the loading response and terminal stance phase (8–41 % of the gait cycle; $p < 0.001$; RMSE = 7 and r = 0.61). A significant difference was also found when comparing the range values ($p = 0.001$).

When comparing manual and UTT data, significant differences were found from heel rise to toe-off (31–68 % of the gait cycle; $p < 0.001$; RMSE = 4.9 and r = 0.82). A significant difference was also found when comparing the range values ($p < 0.001$).

3.2.2 Running trials

When comparing manual and KLT data, significant differences in fascicle length were found throughout the loading response and terminal stance phase (14–51 % of the gait cycle; $p < 0.001$; RMSE = 6.8 and r = 0.61). A significant difference was also found when comparing the range values ($p = 0.03$).

3.2.3 Pulley walking trials

When comparing manual and KLT data, significant differences in fascicle length were found from initial contact to loading response (3–15 % of the gait cycle; $p < 0.001$) and from heel rise to toe-off (30–64 % of the gait cycle; $p < 0.001$) (RMSE = 6.7 and r = 0.74).

When comparing manual and UTT data, significant differences were found from the terminal stance phase to feet adjacent (44–67 % of the gait cycle; $p < 0.001$; RMSE = 5.2 and r = 0.83). A significant difference was also found when comparing the range values ($p = 0.008$).

3.2.4 Pulley running trials

When comparing manual and KLT data, significant differences in fascicle length were found throughout the entire stance phase (11–66 % of the gait cycle; $p = 0.002$; RMSE = 6 and r = 0.62).

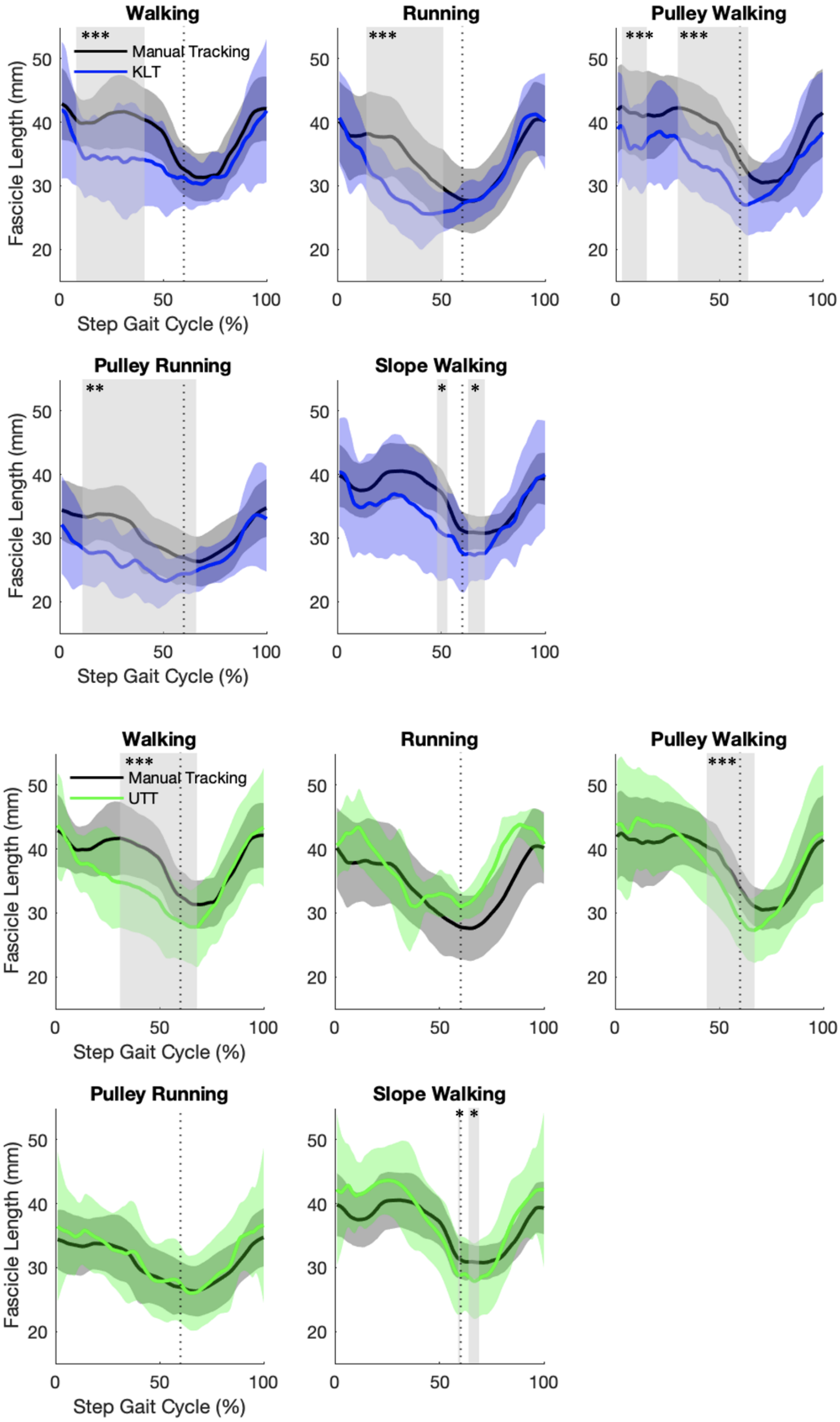


Fig. 6. Average fascicle length across the gait cycle for three children, with standard deviation calculated across five steps per child. Data are shown for five functional tasks for manual tracking compared to KLT (above) and UTT (below). Significantly different parts of the curves are highlighted in light grey. Significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$. The x-axis represents the normalized gait cycle (0–100 %), while the y-axis indicates the fascicle length. The dotted vertical line at 60 % of the gait cycle marks the separation between the stance and swing phase.

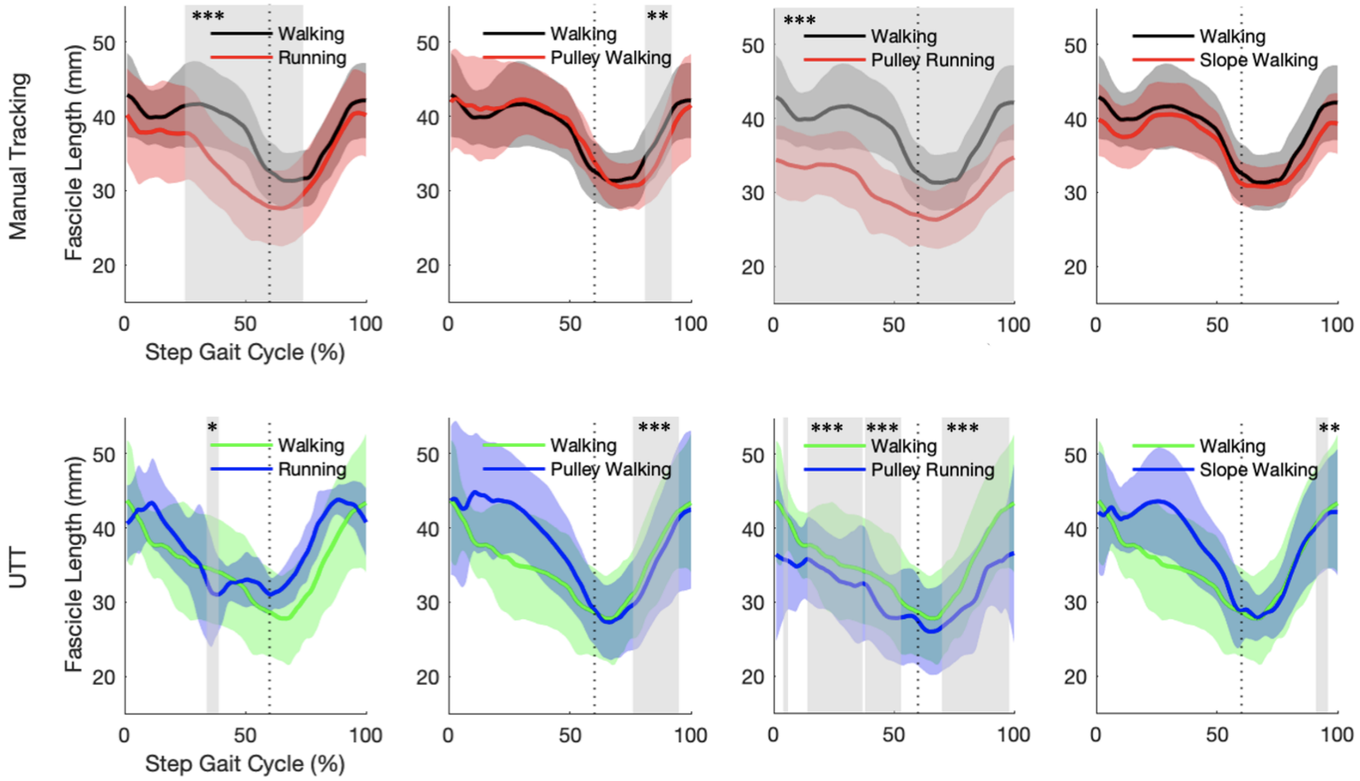


Fig. 7. Average fascicle length across the gait cycle for three children, with standard deviation calculated across five steps per child. Data are shown for four different comparisons between functional tasks for manual tracking and UTT. Significantly different parts of the curves are highlighted in light grey. Significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$. The x-axis represents the normalized gait cycle (0–100 %), while the y-axis indicates the fascicle length. The dotted vertical line at 60 % of the gait cycle marks the separation between the stance and swing phase.

3.2.5 *Slope walking trials*

When comparing manual and KLT data, significant differences in fascicle length were found during the terminal stance phase (48–53 % of the gait cycle; $p = 0.038$) and between toe-off and feet adjacent (63–71 % of the gait cycle; $p = 0.016$) (RMSE = 7.7 and r = 0.52). A significant difference was also found when comparing the range values ($p = 0.02$).

When comparing manual and UTT data, significant differences were found at toe-off (59–60 % of the gait cycle; $p = 0.047$) and when the feet were adjacent (64–69 % of the gait cycle; $p = 0.03$) (RMSE = 4.9 and r = 0.81). A significant difference was also found when comparing the range values ($p < 0.001$).

3.3 *Gait cycle analysis across various functional tasks*

The KLT fascicle tracking algorithm was excluded from the gait cycle analysis across the functional tasks due to its low tracking accuracy, as demonstrated in Section 3.2.

A significant overall difference ($p < 0.001$) in fascicle length between the functional tasks across both manual and UTT tracking was found. However, range values between the functional tasks did not show any significant difference. Further analysis of the manual tracking data revealed significant differences in fascicle length between functional tasks, except for the comparisons between walking and slope walking, running and pulley walking, running and slope walking, and finally pulley running and slope walking. UTT results revealed significant differences in fascicle length between functional tasks, except for the comparisons between pulley walking and slope walking and lastly pulley running and slope walking. Fig. 7 and Fig. 8 show the average fascicle length using manual tracking and UTT algorithm for the comparison between the functional tasks.

3.3.1 Walking compared to running

When looking at manual tracking data, significant differences in fascicle length were found from heel rise to feet adjacent (25–74 % of the gait cycle; $p < 0.001$; RMSE = 5.9 and r = 0.79), whereas UTT data, revealed significant differences at heel rise only (34–39 % of the gait cycle; $p = 0.026$; RMSE = 7.4 and r = 0.51). Fascicle length values are consistently higher during walking than during running.

3.3.2 Walking compared to pulley walking

When looking at manual tracking data, significant differences in fascicle length were found at mid-swing phase (81–92 % of the gait cycle; $p = 0.003$; RMSE = 3.9 and r = 0.8) whereas UTT data, revealed significant differences during the entire swing phase (76–95 % of the gait cycle; $p < 0.001$; RMSE = 5 and r = 0.84). Fascicle length values are consistently higher during walking than during pulley walking.

3.3.3 Walking compared to pulley running

When looking at manual tracking and UTT data, significant differences in fascicle length were found during the entire gait cycle (manual: 0–100 %; $p < 0.001$; RMSE = 8.2 and r = 0.79; UTT: 4–6 %, 14–37 %, 38–53 % and 70–98 %; $p = 0.04$, $p < 0.001$, $p < 0.001$, $p < 0.001$; RMSE = 7.8 and r = 0.75). Fascicle length values are consistently higher during walking than during pulley running.

3.3.4 Walking compared to slope walking

When looking at UTT data, significant differences in fascicle length were found at terminal swing phase (91–96 % of the gait cycle; $p = 0.007$; RMSE = 4.2 and r = 0.86). Fascicle length values at terminal swing phase are higher during walking than during slope walking.

3.3.5 Running compared to pulley walking

When looking at UTT data, significant differences in fascicle length were found from toe-off to mid-swing phase (67–88 % of the gait cycle; $p < 0.001$; RMSE = 8.7 and r = 0.42). Fascicle length values are higher during running than during pulley walking from toe-off to mid-swing phase.

3.3.6 Running compared to pulley running

When looking at manual tracking data, significant differences in fascicle length were found from initial contact to loading response (0–16 % of the gait cycle; $p < 0.001$), at mid-stance phase (24–27 % of the gait cycle; $p = 0.042$) and at terminal stance phase (75–100 % of the gait cycle; $p < 0.001$) (RMSE = 5.3 and r = 0.82) whereas UTT data, also revealed significant differences from initial contact to loading response (4–16 % of the gait cycle; $p < 0.001$) and at terminal stance phase (64–96 % of the gait cycle; $p < 0.001$) (RMSE = 9 and r = 0.47). Fascicle length values are consistently higher during running than during pulley running.

3.3.7 Running compared to slope walking

When looking at UTT data, significant differences in fascicle length were found from toe-off to mid stance phase (64–80 % of the gait cycle; $p < 0.001$; RMSE = 7.5 and r = 0.46). Fascicle length values are higher during running than during slope walking from toe-off to mid stance phase.

3.3.8 Pulley walking compared to pulley running

When looking at manual tracking data, significant differences in fascicle length were found during the entire gait cycle (0–100 %; $p < 0.001$; RMSE = 8.5 and r = 0.72), whereas UTT data, revealed significant differences from loading response to terminal stance phase (7–54 % of the gait cycle; $p < 0.001$) and at mid-swing phase (82–86 % of the gait cycle; $p = 0.036$) (RMSE = 9.1 and r = 0.66). Fascicle length values are consistently higher during pulley walking than during pulley running.

3.3.9 Pulley walking compared to slope walking

When looking at manual tracking data, significant differences in fascicle length were found from terminal stance phase to toe-off (55–61 % of the gait cycle; $p = 0.030$; RMSE = 4.5 and r = 0.76).

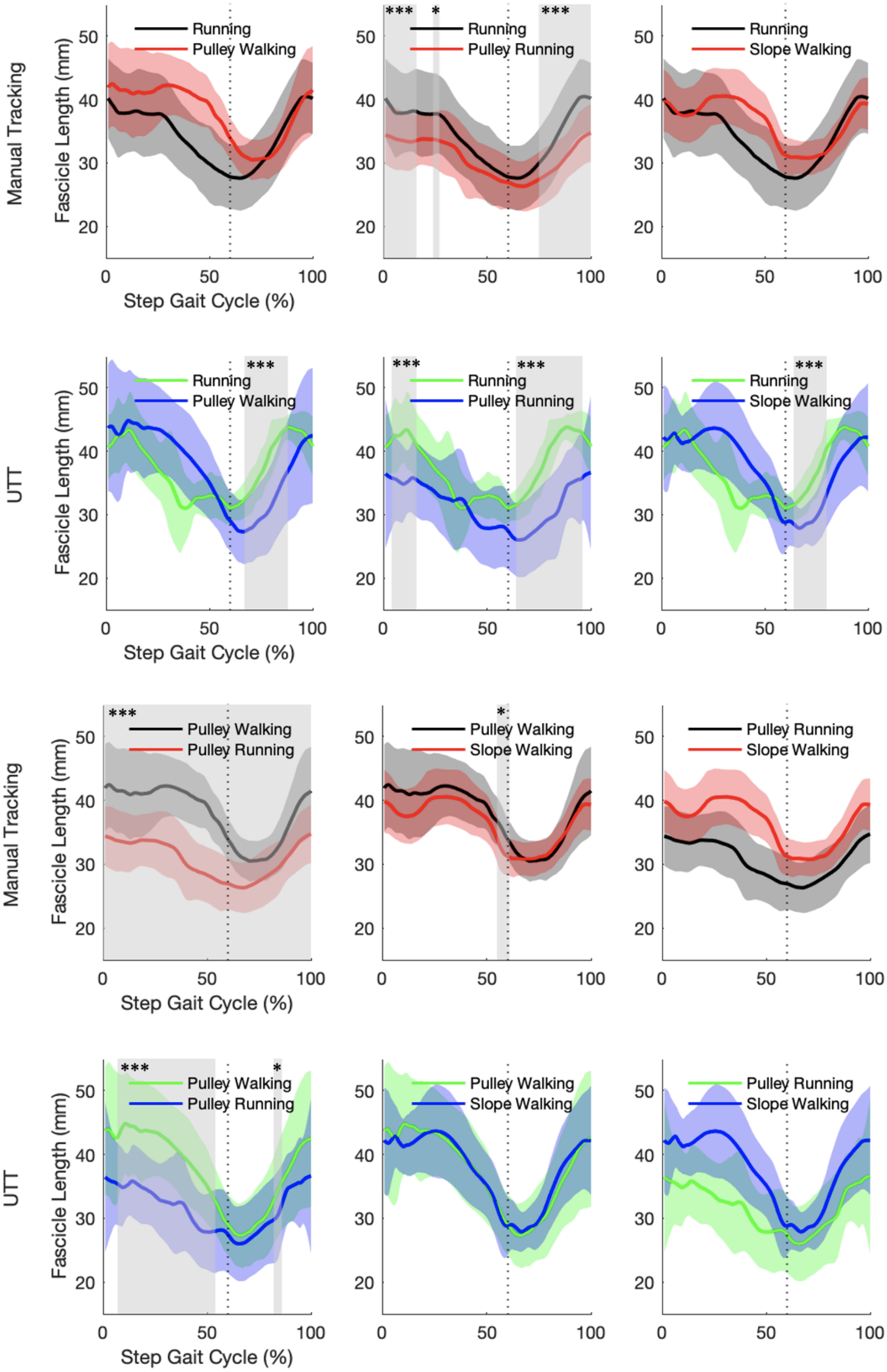


Fig. 8. Average fascicle length across the gait cycle for three children, with standard deviation calculated across five steps per child. Data are shown for six different comparisons between functional tasks for manual tracking and UTT. Significantly different parts of the curves are highlighted in light grey. Significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$. The x-axis represents the normalized gait cycle (0–100 %), while the y-axis indicates the fascicle length. The dotted vertical line at 60 % of the gait cycle marks the separation between the stance and swing phase.

4 Discussion

4.1 KLT and UTT compared to manual tracking

As hypothesized, the findings indicate that the UTT algorithm demonstrates significantly greater computational efficiency while maintaining a high level of agreement with manual tracking (no significant difference for the running trials). However, discrepancies in fascicle length measurements suggest inherent limitations within both UTT and KLT algorithms, that warrant further investigation.

Manual tracking remains the gold standard for fascicle length estimation due to its ability to incorporate human interpretation and account for challenging tracking conditions [8, 9]. The results confirm high intra-rater reliability (RMSE = 4.63 mm (10 %) and r = 0.78), reinforcing its validity as a benchmark method [8, 9]. However, its time-intensive nature limits its feasibility for large datasets. Additionally, it remains susceptible to human error and some variability across trials [1]. It is recommended that future research incorporates advanced validation techniques. For instance, incorporating a manual tracking analysis conducted by a second rater could help reduce potential biases and improve inter-rater reliability. Additionally, the use of statistical tests to quantify statistical difference, for example Coefficient of Multiple Determination (CMD) which is the square of the coefficient of multiple correlation (CMC), would provide robust validation of manual tracking [9].

The performance of KLT compared to manual tracking revealed a significant underestimation of fascicle length during the stance phase (0–60 % of the gait cycle; $p < 0.05$) across all functional tasks, including walking (RMSE = 7 mm, r = 0.61), running (RMSE = 6.8 mm, r = 0.61), pulley walking (RMSE = 6.7 mm, r = 0.74), pulley running (RMSE = 6 mm, r = 0.62), and slope walking (RMSE = 7.7 mm, r = 0.52). The range of fascicle length (max-min) revealed significant differences between manual tracking and KLT, with values of 13.71 ± 4.10 and 20.51 ± 3.81 mm for walking, 15.34 ± 4.19 and 19.93 ± 4.38 mm for running and 11.96 ± 2.49 and 17.45 ± 4.80 mm for slope walking. KLT showed greater fascicle excursion compared to manual tracking. This indicates its tendency to underestimate fascicle length during the stance phase due to smaller minimum values, despite exhibiting relatively high or comparable maximum values to manual tracking. This finding reinforces the systematic underestimation of fascicle length by KLT throughout the gait cycle, as not all significant different clusters are detected when evaluating fascicle range, excluding offset values (smaller minimum and maximum values). The underestimation of fascicle length by KLT during the stance phase is likely due to a combination of factors, including its sensitivity to rapid fascicle movements, increased pennation angles and tracking drift [28]. During the stance phase, the higher muscle activation and rapid fascicle shortening may cause feature loss, leading to tracking errors [28, 31]. In contrast, during the swing phase, where forces are lower and movement is smoother, tracking remains more stable, reducing the likelihood of underestimation of fascicle length [31].

The exclusion of seven TD children (out of the ten available data from TD children) due to unreliable KLT tracking further underscores its limitations in robustness and accuracy. It is important to note that correlation coefficients provide information about the overall strength and direction of the linear relationship between manual tracking and KLT algorithm, while RMSE quantifies the absolute differences between the two by measuring the magnitude of error. While these numerical metrics are essential, analyzing the significant differences along the gait cycle provides additional biomechanical insights by identifying specific clusters where the two diverge significantly. Although KLT demonstrates acceptable correlation and RMSE values (average for functional tasks RMSE = 6.84 mm and r = 0.62) when compared to manual tracking across all functional tasks, significant clusters in fascicle length are evident, highlighting limitations in the accuracy of KLT.

The KLT algorithm showed high tracking error, leading to the exclusion of datasets and a significant underestimation of fascicle length, highlighting the need for further optimization. The KLT algorithm has several limitations, including its reliance on a weak Gaussian Process prior, which assumes that fascicle shapes are locally smooth. Although this assumption facilitates stable tracking, it may reduce accuracy in cases of sharp or abrupt shape changes. Its high computational time may also limit its suitability for real-time applications without further optimization [28]. Additionally, the low line density used in the current study likely resulted in poor texture and low contrast in ultrasound images, further impairing the performance of KLT and contributing to increased tracking drift over time [28]. Future work could integrate more adaptive shape constraints to improve flexibility in capturing rapid or irregular fascicle deformations in low-quality ultrasound images. Approaches such as deep neural networks (e.g., U-net architectures) or Kalman filtering could enhance tracking robustness [23, 32].

The performance of UTT compared to manual tracking revealed significant differences in fascicle length from heel rise to toe-off (30–60 % of the gait cycle; $p < 0.05$) in walking (RMSE = 4.9 mm and r = 0.82), pulley walking (RMSE = 5.2 mm and r = 0.83), and, to a lesser extent, slope walking (RMSE mm = 4.9 and r = 0.81). The range values with significant differences between manual tracking and UTT were 13.71 ± 4.10 and 21.76 ± 3.31 for walking, 15.27 ± 5.19 and 20.91 ± 3.52 for pulley walking and 11.96 ± 2.49 and 19.78 ± 2.50 for slope walking. UTT captured greater fascicle excursion than manual tracking, emphasizing its tendency to underestimate fascicle length from heel rise to toe-off due to smaller minimum values and relatively high maximum values compared to manual tracking. During heel rise and toe-off, relatively high muscle contractions lead to a rapid increase in pennation angle and fascicle shortening, which may have contributed to feature loss in walking trials. During dynamic movements fascicles contract at higher speeds, leading to more extreme changes in length compared to walking trials which may have amplified tracking noise in UTT (average range values: 21.28 ± 5.19 for running and 19.67 ± 15.31 for pulley running) [31]. UTT may inconsistently track fascicle lengths, resulting in measurement fluctuations, particularly when low line density ultrasound images are used [23]. Furthermore, the constant noise assumptions in the UTT algorithm can lead to tracking instability. During dynamic tasks, where fascicle behavior is highly variable, these assumptions may reduce the reliability of feature tracking. As a result, the algorithm may overcompensate for noise, causing exaggerated variations in measured fascicle length [23].

UTT demonstrates strong correlation and low RMSE values ($r \geq 0.68$ and RMSE close to 10 %) when compared to manual tracking across all functional tasks but significant clusters in fascicle length are evident in walking trials, highlighting limitations in its accuracy. In past studies, RMSE and correlation coefficients were used to evaluate the accuracy of machine learning (ML) models and personalized musculoskeletal models in tracking muscle fascicle length changes and mechanical behavior [1, 33]. RMSE and correlation coefficients were used to validate an ML-based approach for real-time muscle fascicle tracking from ultrasound images [1]. The best-performing model, Support Vector Machine (SVM), achieved an average RMSE of 2.86 ± 2.55 mm for direct training and 3.28 ± 2.64 mm for cross-participant conditions, with correlation coefficients of 0.70 ± 0.34 and 0.65 ± 0.35, respectively, indicating strong agreement with ground-truth data from UltraTrack tracking software [1]. Similarly, RMSE and correlation coefficients were used to assess the ability of a personalized musculoskeletal model to capture plantarflexor mechanics in children with Cerebral Palsy (CP) [33]. RMSE measured discrepancies between simulated and experimental torque–angle and fascicle length–angle curves, with initial RMSE values of 0.12 Nm/kg for torque in CP and 0.31 Nm/kg in TD children, while for fascicle length, RMSE was 0.035 in CP and 0.021 in TD (normalized to tibia length). After personalization, RMSE decreased by over 88 % for torque and 98 % for fascicle length, demonstrating significant model accuracy improvements [33]. In both studies, RMSE and correlation coefficients were essential validation metrics, ensuring that computational models accurately reflected real-world muscle dynamics. While significant clusters were detected in UTT fascicle length data, the relevance of strong correlations and low RMSE values remains important.

UTT demonstrated superior computational efficiency (35 ± 5 s per trial) and greater accuracy in fascicle length estimation compared to KLT (60 ± 10 min). While discrepancies arose in specific gait phases (walking trials), UTT closely approximated manual tracking (no significant difference) in running and pulley running trials (RMSE = 6.7 mm, r = 0.62 and RMSE = 4.4 mm, r = 0.81), suggesting its effectiveness for faster movements. Since UTT assumes that aponeuroses are straight lines, it may not account for their curvature in certain tasks, such as at rest or during low muscle activation levels. This assumption may have affected the accuracy of walking trials, making them less precise compared to manual tracking [23].

Although the UTT algorithm demonstrated strong agreement with manual tracking in dynamic functional tasks, high tracking noise suggests that further refinement is necessary. Incorporating dynamic noise estimation based on Hough angle variance would allow the UTT algorithm to adapt in real-time to changing conditions, improving its robustness to high variability during fast movements [23]. Future research should explore additional optimization strategies such as develop more sophisticated finite element models that account for fascicle and aponeurosis curvature, to enhance the robustness and adaptability of the UTT algorithm for a wider range of movement types [23, 28]. This includes slower or more complex gait patterns, such as those observed in neurological disorders (e.g., Parkinson disease and CP), where altered gait mechanics and muscle coordination may challenge tracking accuracy [7, 34]. The computational efficiency of UTT compared to KLT and manual tracking suggests strong potential for real-time clinical applications, particularly in settings that require immediate feedback, such as rehabilitation and sports science [1, 35]. Future research should investigate its integration into wearable ultrasound systems, facilitating continuous muscle monitoring during rehabilitation protocols, neuromuscular assessments, and biofeedback-based therapies. This would enable real-time tracking of muscle dynamics, helping clinicians and researchers better understand muscle function in both clinical and high-performance contexts [1, 35].

The first part of the current study compared the accuracy and efficiency of KLT and UTT with manual tracking for fascicle length estimation across the gait cycle. UTT showed more computationally efficiency while maintaining strong agreement with manual tracking, though some noise in tracking was observed. KLT, however, showed systematic underestimation of fascicle length, likely due to tracking drift and sensitivity to rapid movements, limiting its accuracy. UTT performed well in dynamic tasks, whereas manual tracking remains the most accurate but time-intensive tracking.

4.2 *Manual and UTT tracking across various functional tasks*

The present findings provide valuable insights into the neuromechanical behavior of the Medial Gastrocnemius (MG) under various functional tasks, reinforcing and extending existing evidence on muscle-specific contributions to propulsion [2]. During walking, fascicles in the MG operate near their resting length at initial contact and remain relatively isometric throughout the stance phase, facilitating efficient absorption and release of elastic energy by the tendons [2, 7, 36]. In contrast, during running, fascicles are shorter at initial contact and actively shorten throughout the stance phase, increasing strain on the series elastic elements and enabling rapid force production [2]. The greater fascicle length during walking compared to running, particularly from heel rise to toe-off based on manual tracking data, confirms that walking relies more on passive elastic energy storage and release, whereas running depends more on active fascicle shortening for propulsion [2]. However, analysis using UTT data revealed significant differences only at heel rise. This discrepancy between manual tracking and UTT underscores potential differences in the sensitivity and resolution of manual and automated techniques in detecting subtle variations in fascicle behavior across gait tasks (Section 4.1). Range values did not differ significantly between functional tasks for both manual tracking and UTT, suggesting that the muscle-tendon unit adapts to maintain consistent fascicle behavior despite variations in speed and force. This highlights the conserved strategies that regulate fascicle length changes across functional tasks [2, 7, 36].

Differences in fascicle length between walking and running, and pulley-based tasks support the hypothesis that increased loading elicits a task-specific neuromechanical response [2]. The significantly shorter fascicle length during pulley-based tasks compared to walking and running, particularly during the stance and/or swing phase (detected using both manual tracking and UTT), suggests that pulley-based tasks increases mechanical demand for active force generation. This highlights the adaptive capacity of the MG to accommodate increased propulsive requirements [2, 7, 36]. However, the absence of significant differences in shortening and lengthening patterns (Supplementary Fig. 4, 5, 6, 7, 8 and 9), across most pulley-based tasks comparisons suggests that motor strategies governing fascicle length change (neural activation, muscle-tendon interaction, and elastic energy) remain conserved despite increased mechanical demand [7]. These findings have important implications for gait rehabilitation and assistive device design. Pulley-based tasks may serve as a targeted intervention to enhance MG function and propulsion, particularly in populations with gait impairments, where increasing loading capacity and active force generation are critical for functional mobility [7].

The second part of the current study provides insights into the neuromechanical behavior of the MG across functional tasks. Walking relies more on passive elastic energy storage, while running emphasizes active fascicle shortening for propulsion. Pulley-based tasks further alter contractile behavior, increasing the demand for active force generation. Despite differences in fascicle length across tasks, motor strategies governing fascicle behavior remain largely conserved. These findings highlight the adaptive capacity of the gastrocnemius and have implications for gait rehabilitation and assistive device design.

Finally, while the study included a diverse range of functional tasks, only a small number (see power analysis in Section 2) of TD children were analyzed and no clinical populations. The small sample size may have failed to capture the variability across different ages, genders and measured leg, which makes it difficult to control for individual differences such as body composition, muscle fiber distribution, or motor ability. All could influence fascicle behavior during functional tasks and introduce variability into the data, this is especially varying in a pediatric population [37]. KLT should have been excluded from the analysis earlier due to its high computational time and tracking errors, which limited the inclusion of more children and slope running data for the comparison between manual tracking, UTT, and functional tasks. Consequently, the study lacks an explicit validation framework for its key findings. Comparing data between TD and CP children would provide a valuable assessment of the accuracy of UTT in both dynamic tasks and clinical populations while offering insight into muscle dynamics differences between these groups. Furthermore, important to note is the dynamic nature of treadmill-based tasks which may have introduced motion artifacts, such as probe displacement or participant movement, potentially degrading the tracking quality data [24]. Even slight misalignment or perturbations during movement could affect the accuracy of fascicle length data, further contributing to potential measurement errors. In summary, key limitations included KLT tracking errors, UTT discrepancies in walking trials, high tracking noise in running trials and small sample size. Future research should focus on refining tracking algorithms, increasing sample diversity and size, and investigating real-time clinical applications.

5 Conclusion

The current study evaluates the accuracy and efficiency of the Kanade-Lucas-Tomasi (KLT) and UltraTimTrack (UTT) algorithms compared to manual tracking for assessing muscle fascicle behavior. While manual tracking remains the gold standard in accuracy, its time-intensive nature limits real-time applications.

Key Findings

- Manual Tracking: Highly accurate but time-consuming.
- KLT: High tracking error and time-inefficient.
- UTT: Time-efficient but shows discrepancies in walking trials and high tracking noise in running trials.

Specific Findings when comparing various functional tasks

- UTT maintains trend consistency with manual tracking across functional tasks, highlighting its potential for real-time medical and biomechanical applications.
- Walking: Associated with longer fascicle lengths and greater reliance on passive elastic energy storage.
- Running: Emphasizes active fascicle shortening.
- Pulley-Based Tasks: Further increase the demand for active fascicle shortening.

UTT emerges as the most accurate and efficient alternative to manual tracking, particularly for dynamic movements (running and pulley-based running tasks). However, further validation is necessary to enhance its accuracy in biomechanics research. Future studies should focus on expanding participant cohorts, refining tracking algorithms, and exploring automated tracking applications in clinical populations.

References

[1] L. G. Rosa, J. S. Zia, O. T. Inan, and G. S. Sawicki, "Machine learning to extract muscle fascicle length changes from dynamic ultrasound images in real-time," *PLoS ONE*, vol. 16, no. 5, p. e0246611, 5 2021. [Online]. Available: https://doi.org/10.1371/journal.pone.0246611

[2] W. H. Clark, R. E. Pimentel, and J. R. Franz, "Imaging and simulation of inter-muscular differences in triceps SURAE contributions to forward propulsion during walking," *Annals of Biomedical Engineering*, vol. 49, no. 2, pp. 703–715, 9 2020. [Online]. Available: https://doi.org/10.1007/s10439-020-02594-x

[3] S. May, S. Locke, and M. Kingsley, "Reliability of ultrasonographic measurement of muscle architecture of the gastrocnemius medialis and gastrocnemius lateralis," *PLoS ONE*, vol. 16, no. 9, p. e0258014, 9 2021. [Online]. Available: https://doi.org/10.1371/journal.pone.0258014

[4] M. Hösl, H. Böhm, A. Arampatzis, A. Keymer, and L. Döderlein, "Contractile behavior of the medial gastrocnemius in children with bilateral spastic cerebral palsy during forward, uphill and backward-downhill gait," *Clinical Biomechanics*, vol. 36, pp. 32–39, 5 2016. [Online]. Available: https://doi.org/10.1016/j.clinbiomech.2016.05.008

[5] L. Barber, C. Carty, L. Modenese, J. Walsh, R. Boyd, and G. Lichtwark, "Medial gastrocnemius and soleus muscleâtendon unit, fascicle, and tendon interaction during walking in children with cerebral palsy," *Developmental Medicine Child Neurology*, vol. 59, no. 8, pp. 843–851, 4 2017. [Online]. Available: https://doi.org/10.1111/dmcn.13427

[6] B. M. Kalkman, L. Bar-On, F. Cenni, C. N. Maganaris, A. Bass, G. Holmes, K. Desloovere, G. J. Barton, and T. D. O' Brien, "Muscle and tendon lengthening behaviour of the medial gastrocnemius during ankle joint rotation in children with cerebral palsy," *Experimental Physiology*, vol. 103, no. 10, pp. 1367–1376, 8 2018. [Online]. Available: https://doi.org/10.1113/ep087053

[7] F. Cenni, N. Alexander, M. Sukanen, A. Mustafaoglu, Z. Wang, R. Wang, and T. Finni, "ISB clinical biomechanics award winner 2023: Medial gastrocnemius muscle and Achilles tendon interplay during gait in cerebral palsy," *Clinical Biomechanics*, vol. 111, p. 106158, 12 2023. [Online]. Available: https://doi.org/10.1016/j.clinbiomech.2023.106158

[8] N. J. Cronin, C. P. Carty, R. S. Barrett, and G. Lichtwark, "Automatic tracking of medial gastrocnemius fascicle length during human locomotion," *Journal of Applied Physiology*, vol. 111, no. 5, pp. 1491–1496, 8 2011. [Online]. Available: https://doi.org/10.1152/japplphysiol.00530.2011

[9] J. G. Gillett, R. S. Barrett, and G. A. Lichtwark, "Reliability and accuracy of an automated tracking algorithm to measure controlled passive and active muscle fascicle length changes from ultrasound," *Computer Methods in Biomechanics Biomedical Engineering*, vol. 16, no. 6, pp. 678–687, 1 2012. [Online]. Available: https://doi.org/10.1080/10255842.2011.633516

[10] J. Darby, E. F. Hodson-Tole, N. Costen, and I. D. Loram, "Automated regional analysis of B-mode ultrasound images of skeletal muscle movement," *Journal of Applied Physiology*, vol. 112, no. 2, pp. 313–327, 10 2011. [Online]. Available: https://doi.org/10.1152/japplphysiol.00701.2011

[11] J. F. Drazan, T. J. Hullfish, and J. R. Baxter, "An automatic fascicle tracking algorithm quantifying gastrocnemius architecture during maximal effort contractions," *PeerJ*, vol. 7, p. e7120, 7 2019. [Online]. Available: https://doi.org/10.7717/peerj.7120

[12] L. Gionfrida, R. W. Nuckols, C. J. Walsh, and R. D. Howe, "Age-Related reliability of B-Mode analysis for tailored exosuit assistance," *Sensors*, vol. 23, no. 3, p. 1670, 2 2023. [Online]. Available: https://doi.org/10.3390/s23031670

[13] W. Pirker and R. Katzenschlager, "Gait disorders in adults and the elderly," *Wiener klinische Wochenschrift*, vol. 129, no. 3-4, pp. 81–95, 10 2016. [Online]. Available: https://doi.org/10.1007/s00508-016-1096-4

[14] R. Di Gregorio and L. Vocenas, "Identification of Gait-Cycle phases for prosthesis control," *Biomimetics*, vol. 6, no. 2, p. 22, 3 2021. [Online]. Available: https://doi.org/10.3390/biomimetics6020022

[15] J. Duclay, A. Martin, A. Duclay, G. Cometti, and M. Pousson, "Behavior of fascicles and the myotendinous junction of human medial gastrocnemius following eccentric strength training," *Muscle Nerve*, vol. 39, no. 6, pp. 819–827, 3 2009. [Online]. Available: https://doi.org/10.1002/mus.21297

[16] J. N. Maharaj, L. Barber, H. P. Walsh, and C. P. Carty, "Flip-flops do not alter the neuromuscular function of the gastrocnemius muscle and tendon during walking in children," *Gait Posture*, vol. 77, pp. 83–88, 1 2020. [Online]. Available: https://doi.org/10.1016/j.gaitpost.2019.12.032

[17] J. Aeles, G. A. Lichtwark, S. Lenchant, L. Vanlommel, T. Delabastita, and B. Vanwanseele, "Information from dynamic length changes improves reliability of static ultrasound fascicle length measurements," *PeerJ*, vol. 5, p. e4164, 12 2017. [Online]. Available: https://doi.org/10.7717/peerj.4164

[18] W. H. Clark and J. R. Franz, "Triceps surae muscleâsubtendon interaction differs between young and older adults," *Connective Tissue Research*, vol. 61, no. 1, pp. 104–113, 5 2019. [Online]. Available: https://doi.org/10.1080/03008207.2019.1612384

[19] T. J. Van Der Zee and A. D. Kuo, "TimTrack: A drift-free algorithm for estimating geometric muscle features from ultrasound images," *PLoS ONE*, vol. 17, no. 3, p. e0265752, 3 2022. [Online]. Available: https://doi.org/10.1371/journal.pone.0265752

[20] E. S. Matijevich, L. M. Branscombe, and K. E. Zelik, "Ultrasound estimates of Achilles tendon exhibit unexpected shortening during ankle plantarflexion,"

Journal of Biomechanics, vol. 72, pp. 200–206, 3 2018. [Online]. Available: https://doi.org/10.1016/j.jbiomech.2018.03.013

[21] W. Swinnen, W. Hoogkamer, T. Delabastita, J. Aeles, F. De Groote, and B. Vanwanseele, "Effect of habitual foot-strike pattern on the gastrocnemius medialis muscle-tendon interaction and muscle force production during running," *Journal of Applied Physiology*, vol. 126, no. 3, pp. 708–716, 1 2019. [Online]. Available: https://doi.org/10.1152/japplphysiol.00768.2018

[22] C. Zhang, L. Deng, X. Zhang, K. Wu, J. Zhan, W. Fu, and J. Jin, "Effects of 12-week gait retraining on plantar flexion torque, architecture, and behavior of the medial gastrocnemius in vivo," *Frontiers in Bioengineering and Biotechnology*, vol. 12, 3 2024. [Online]. Available: https://doi.org/10.3389/fbioe.2024.1352334

[23] T. J. Van Der Zee, P. Tecchio, D. Hahn, and B. J. Raiteri, "UltraTimTrack: a Kalman-filter-based algorithm to track muscle fascicles in ultrasound image sequences," *PeerJ Computer Science*, vol. 11, p. e2636, 1 2025. [Online]. Available: https://doi.org/10.7717/peerj-cs.2636

[24] B. Mooijekind, E. Flux, A. I. Buizer, M. M. Van Der Krogt, and L. Bar-On, "The influence of wearing an ultrasound device on gait in children with cerebral palsy and typically developing children," *Gait Posture*, vol. 101, pp. 138–144, 2 2023. [Online]. Available: https://doi.org/10.1016/j.gaitpost.2023.02.007

[25] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, "G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences," *Behavior Research Methods*, vol. 39, no. 2, pp. 175–191, 5 2007. [Online]. Available: https://doi.org/10.3758/bf03193146

[26] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nature Methods*, vol. 9, no. 7, pp. 671–675, 6 2012. [Online]. Available: https://doi.org/10.1038/nmeth.2089

[27] G. Lichtwark, *Ultrasound technology for examining the mechanics of the muscle, tendon, and ligament*, 1 2017. [Online]. Available: https://doi.org/10.1007/978-3-319-30808-1_156-1

[28] J. Darby, N. B. Li, N. Costen, I. Loram, and E. Hodson-Tole, "Estimating skeletal muscle fascicle curvature from B-Mode ultrasound image sequences," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 7, pp. 1935–1945, 2 2013. [Online]. Available: https://doi.org/10.1109/tbme.2013.2245328

[29] R. Taylor, "Interpretation of the correlation Coefficient: A Basic review," *Journal of diagnostic medical sonography*, vol. 6, no. 1, pp. 35–39, 1 1990. [Online]. Available: https://doi.org/10.1177/875647939000600106

[30] L. K. Kwah, R. Z. Pinto, J. Diong, and R. D. Herbert, "Reliability and validity of ultrasound measurements of muscle fascicle length and pennation in humans: a systematic review," *Journal of Applied Physiology*, vol. 114, no. 6, pp. 761–769, 1 2013. [Online]. Available: https://doi.org/10.1152/japplphysiol.01430.2011

[31] R. Hamard, J. Aeles, N. Y. Kelp, R. Feigean, F. Hug, and T. J. M. Dick, "Does different activation between the medial and the lateral gastrocnemius during walking translate into different fascicle behavior?" *Journal of Experimental Biology*, vol. 224, no. 12, 6 2021. [Online]. Available: https://doi.org/10.1242/jeb.242626

[32] P. Ritsche, M. V. Franchi, O. Faude, T. Finni, O. Seynnes, and N. J. Cronin, "Fully Automated Analysis of Muscle Architecture from B-Mode Ultrasound Images with DL-Track-US," *Ultrasound in Medicine Biology*, vol. 50, no. 2, pp. 258–267, 11 2023. [Online]. Available: https://doi.org/10.1016/j.ultrasmedbio.2023.10.011

[33] K. Veerkamp, M. M. Van Der Krogt, J. Harlaar, T. D. O' Brien, B. Kalkman, A. Seth, and L. Bar-On, "Personalisation of Plantarflexor Musculotendon Model Parameters in Children with Cerebral Palsy," *Annals of Biomedical Engineering*, vol. 51, no. 5, pp. 938–950, 11 2022. [Online]. Available: https://doi.org/10.1007/s10439-022-03107-8

[34] R. R. Smart, C. M. Richardson, D. J. Wile, B. H. Dalton, and J. M. Jakobi, "Importance of Maximal Strength and Muscle-Tendon Mechanics for Improving Force Steadiness in Persons with Parkinsonâs Disease," *Brain Sciences*, vol. 10, no. 8, p. 471, 7 2020. [Online]. Available: https://doi.org/10.3390/brainsci10080471

[35] Z.-H. Huang, C. Z-h, MA, L.-K. Wang, X.-Y. Wang, S.-N. Fu, and Y.-P. Zheng, "Real-Time visual biofeedback via wearable ultrasound imaging can enhance the muscle contraction training outcome of young adults," *The Journal of Strength and Conditioning Research*, vol. 36, no. 4, pp. 941–947, 2 2022. [Online]. Available: https://doi.org/10.1519/jsc.0000000000004230

[36] S. Bohm, R. Marzilger, F. Mersmann, A. Santuz, and A. Arampatzis, "Operating length and velocity of human vastus lateralis muscle during walking and running," *Scientific Reports*, vol. 8, no. 1, 3 2018. [Online]. Available: https://doi.org/10.1038/s41598-018-23376-5

[37] T. D. O' Brien, N. D. Reeves, V. Baltzopoulos, D. A. Jones, and C. N. Maganaris, "Muscleâtendon structure and dimensions in adults and children," *Journal of Anatomy*, vol. 216, no. 5, pp. 631–642, 3 2010. [Online]. Available: https://doi.org/10.1111/j.1469-7580.2010.01218.x

Acknowledgements

First, I would like to thank Dr. ir. Winfred Mugge for making this thesis possible, supporting me throughout the entire process and providing constructive feedback. The insightful lunch meetings at Neuro Muscular Control lab (NMClab) provided a valuable space to share and connect with other students.

Furthermore, I would like to thank Dr. Marjolein van der Krogt and Babette Mooijekind for giving me the opportunity to work on this interesting topic, providing support and creating a welcoming and pleasant atmosphere in the rehabilitation department at Amsterdam Medical Center VUmc.

A special thanks to Dr. Liesbeth van Vulpen for giving me the opportunity to participate as a trainee in the MegaPower training at Reade Amsterdam. Being able to engage, laugh, and train children with Cerebral Palsy (CP) has taught me so much!

To Rijk Dersjant for giving me the opportunity to assist with his PhD project by conducting measurements across the Netherlands on children with CP. This experience allowed me to deepen my understanding of CP and explore various medical centers throughout the country.

A big thank you to Emile Jongejans for his support, enjoyable coffee breaks, engaging conversations, and help in resolving MATLAB errors.

To my study friends, Elena and Eline, for their support throughout this entire master's journey. Sharing this rollercoaster of emotions with you made me feel so much stronger. Through our laughter, tears, and moments of relief, we lifted each other up. I will miss our time together at TU Delft!

Thank you to Hayo, Charlot, Kaitlynn and Boy for making me feel supported and at home in the Netherlands. Knowing that I can always count on you means a lot to me. My entire experience since moving here would not have been the same without you.

To my Canada exchange roommates, Marthe-Andrea and Annie, despite being spread across Europe, I always know I can count on you. No matter the distance, we are just one FaceTime away, which feels like a huge virtual hug. *Piep piep piep!*

To all of my childhood and lifelong friends in Switzerland, thank you for your support and love. No matter how far I go, you always welcome me back as if I never left.

Lastly, infinite thanks to my parents and family for supporting and advising me through all the ups and downs since the moment I decided to study in the Netherlands. I would not be where I am today without your trust, love, and confidence. You are truly an inspiration to me and I am deeply grateful.

Kiitos, merci & dankjewel!

Supplementary materials

Supplementary Table 1. Details on participant demographics of data collected in Gent.

Number	Sex	Date of Birth	Age (years)	Weight (kg)	Height (cm)	Tibial Length (cm)	Non-preferred leg	Weight walk (kg)	Weight run (kg)
TD02	V	12/19/16	7.15	21.20	127.10	30.00	L	4.25	2.25
TD03	M	12/2/13	10.20	40.10	150.00	38.00	L	8.25	6.25
TD04	V	9/20/13	10.41	37.90	143.50	32.00	L	8.25	6.25
TD05	V	2/7/13	13.00	29.50	140.50	32.00	L	8.25	4.75
TD06	M	1/29/12	12.05	32.90	149.50	36.00	R	8.25	6.25
TD07	V	3/18/13	10.92	32.20	149.50	34.50	L	8.25	6.25
TD08	V	3/18/13	10.92	30.40	146.50	33.00	L	8.25	6.25
TD09	M	1/6/14	10.27	40.30	153.50	40.30	L	10.25	8.25
TD10	M	9/16/14	9.58	30.70	138.00	30.00	L	8.25	6.25
TD11	M	12/11/12	11.35	49.50	149.40	34.00	L	10.25	8.25

Supplementary Table 2. Range values (mm) using manual tracking data.

Number	Steps	Walking	Running	Pulley Walking	Pulley Running	Slope Walking
TD03	1	14.54	12.20	9.88	10.74	11.46
TD03	2	14.35	14.70	12.39	14.34	9.37
TD03	3	12.16	13.27	12.25	14.54	9.34
TD03	4	13.86	16.98	9.36	6.47	11.03
TD03	5	11.97	21.32	8.49	9.32	9.34
TD06	1	14.21	12.44	20.53	11.73	16.66
TD06	2	9.78	16.16	11.97	11.62	10.35
TD06	3	5.90	11.71	8.80	9.43	10.56
TD06	4	8.51	12.97	17.88	9.11	15.90
TD06	5	9.56	6.60	21.56	13.15	10.38
TD08	1	18.88	21.54	17.17	9.65	11.27
TD08	2	18.05	20.32	19.36	9.29	16.26
TD08	3	18.97	13.19	20.20	10.02	12.60
TD08	4	15.70	19.39	24.08	8.86	11.75
TD08	5	19.20	17.25	15.08	9.73	13.10

Supplementary Table 3. Range values (mm) using KLT data.

Number	Steps	Walking	Running	Pulley Walking	Pulley Running	Slope Walking
TD03	1	14.04	22.17	12.64	11.00	15.10
TD03	2	17.65	20.33	15.23	10.84	11.76
TD03	3	17.66	21.30	9.33	14.92	11.24
TD03	4	16.29	17.32	12.16	12.33	15.40
TD03	5	17.65	23.53	10.71	10.12	16.32
TD06	1	17.03	20.47	13.62	11.04	10.33
TD06	2	22.38	14.94	20.79	10.31	14.23
TD06	3	18.92	13.98	14.43	9.20	17.84
TD06	4	21.98	11.08	18.03	12.29	16.34
TD06	5	21.62	15.88	18.02	10.52	19.13
TD08	1	24.55	23.72	25.21	17.27	25.42
TD08	2	28.18	23.63	21.09	17.88	26.17
TD08	3	23.12	23.69	20.68	18.03	19.82
TD08	4	23.16	20.31	24.31	18.99	21.64
TD08	5	23.44	26.55	21.24	23.33	20.95

Supplementary Table 4. Range values (mm) using UTT data.

Number	Steps	Walking	Running	Pulley Walking	Pulley Running	Slope Walking
TD03	1	24.04	28.59	15.40	13.88	18.02
TD03	2	18.73	22.94	19.97	56.81	16.77
TD03	3	21.84	29.56	15.97	55.93	18.75
TD03	4	25.78	24.46	16.53	15.09	22.54
TD03	5	25.16	24.33	17.13	22.42	17.43
TD06	1	19.38	14.41	21.28	16.80	18.54
TD06	2	18.81	16.35	24.52	17.70	16.47
TD06	3	13.18	13.74	23.68	12.71	21.28
TD06	4	21.53	17.23	23.28	10.27	19.21
TD06	5	20.40	20.36	18.28	13.06	16.63
TD08	1	23.60	19.11	22.42	12.00	20.38
TD08	2	24.27	19.10	23.86	7.57	22.60
TD08	3	21.75	19.57	26.31	11.78	23.82
TD08	4	22.94	19.70	24.14	11.79	21.84
TD08	5	24.98	29.80	20.90	17.31	22.48

Supplementary Table 5. First part of the comparison between fascicle tracking techniques including manual tracking, semi-automated and fully automated algorithms (each color representing one method approach).

Reference	Study design	Sample size	Population	Tracking technique	Data processing	Outcome measures
(Aeles et al., 2017)	Comparative study	28 participants	Healthy adults (15 males, 13 females)	Ultrasound Imaging, fascicle tracking	UltraTrack: software for semi-automated tracking (MATLAB-based custom scripts)	Fascicle length
(Barber et al., 2017)	Comparative study	14 CP, 10 TD	Approximately 10 years old, GMFCS I-II	US, 3D motion capture, EMG	Manual fascicle length determination in the first frames, Kanade-Lucas optical flow algorithm method in MATLAB	Length changes in MTU, fascicle, and tendon components
(Clark & Franz, 2018)	Experimental study	9 adults	5 males, 4 females; average age 25.1 ± 5.6 years	Dual-probe ultrasound imaging	Speckle tracking algorithms	Differences in muscle shortening and tendon tissue displacement
(Clark & Franz, 2019)	Comparative experimental study	19 adults	9 adults (average age: 25.1 years, 4 females) and 10 older adults (average age: 74.3 years, 4 females)	Dual-Probe Ultrasound Imaging, motion capture, biodex dynamometer	UltraTrack: software for semi-automated tracking (MATLAB-based custom scripts) and 2D speckle tracking	Peak muscle shortening (GAS and SOL), Subtendon displacement within AT
(Cronin et al., 2011)	Methodological development and validation study	8 healthy adults	Age: 30 ± 6 years	Ultrasound Imaging, motion capture, force sensors	Manual fascicle length determination in the first frames, Kanade-Lucas optical flow algorithm method in MATLAB	Fascicle length
(Cronin & Finni, 2013)	Comparative experimental study	10 healthy male volunteers	29 ± 4 years	Ultrasound, force sensors and overground motion analysis	Manual fascicle length determination in the first frames, Kanade-Lucas optical flow algorithm method in MATLAB	Fascicle length change and velocity
(Darby et al., 2011)	Methodological development and validation study	8 healthy adults	Not specified	Ultrasound imaging, feature tracking	Kanade-Lucas-Tomasi (KLT) and multiresolution active shape model (ASM)	Fascicle and aponeurosis movement
(Drazan et al., 2019)	Methodological development and validation study	5 adults	Healthy individuals	Ultrasound Imaging, dynamometry	Kanade-Lucas-Tomasi (KLT) point tracking method in MATLAB	ascicle length and pennation angle

Supplementary Table 6. First part of the comparison between fascicle tracking techniques including manual tracking, semi-automated and fully automated algorithms (each color representing one method approach).

(Duclay et al., 2009)	Controlled interventional study	10 training group (TG) and 8 control group (CG)	Age: 23 ± 3 years (TG) and 24 ± 3 years (CG)	Ultrasound imaging and isokinetic Dynamometer	Digitizing software to measure fascicle length	Fascicle length (Fl), angle (θ), and thickness (t)
(Finni et al., 2022)	Comparative study	22 participants	42 ± 9 years, experienced unilateral Achilles tendon rupture	Ultrasound Imaging	Tracker 5.1.3 Software	Displacement and strain: Differences between MTJ and proximal tracking points
(Frisk et al., 2019)	Experimental study	14 adults with CP, 17 neurologically intact adults	36 ± 10.6 years, GMFCS levels I–III	Torque Measurements, ultrasound Imaging	Manual fascicle length determination in the first frames, Kanade-Lucas optical flow algorithm method in MATLAB	Plantar flexor torque, MG fascicle length
(Gillett et al., 2012)	Methodological validation study	6 healthy male volunteers	Age: 30 ± 4 years	Ultrasound Imaging, dynamometry	Manual fascicle length determination in the first frames, Kanade-Lucas optical flow algorithm method in MATLAB	Fascicle length changes
(Gionfrida et al., 2023)	Feasibility study combining machine learning and biomechanics	13 adults	3 females, 6 males; average age: 29.1 ± 4.04 years, 1 female, 3 males; average age: 75.6 ± 6 years	Ultrasound Imaging, Ground Reaction Force Detection	UltraTrack: software for semi-automated tracking (MATLAB-based custom scripts), ResNet + 2x-LSTM model predicted fascicle lengths over continuous ultrasound video frames	Model performance metrics: R^2 (coefficient of determination), RMSE (root-mean-square error), and MAPE (mean absolute percentage error), Fascicle length predictions compared to manual labels, average detection time for fascicle lengths
(Hauraix et al., 2017)	Experimental study	Fifteen healthy, physically active men (athletes)	Mean age: 22.8 ± 3.2 years	Ultrasound Imaging, dynamometry, 3D motion capture	Manual fascicle length determination in the first frames, Kanade-Lucas optical flow algorithm method in MATLAB	Fascicle shortening, velocity and torque
(Heieis et al., 2023)	Crossover experimental study	12 healthy participants	18–30 years and 60–70 years	Ultrasound Imaging & marker-based motion capture	Manually processed using custom Python scripts to outline fascicle and aponeurosis architecture	Fascicle curvature, MTC strain, pennation angle, fascicle length, and strain
(Hollville et al., 2020)	Experimental study	12 active males	Aged 24.2 ± 2.0	Ultrasound Imaging, EMG, force sensors and kinematic analysis	Manual fascicle length determination in the first frames, Kanade-Lucas optical flow algorithm method in MATLAB	Fascicle shortening velocity, MTU shortening velocity, and muscle activation levels
(Hösl et al., 2016)	Experimental study	15 SCP and 17 TD children	(GMFCS Levels I–II, bilateral mean age: 11 ± 2.8 years) and (mean age: 12.2 ± 2.3 years)	Ultrasound imaging, 3D gait analysis and EMG	Manual fascicle length determination in the first frames, Kanade-Lucas optical flow algorithm method in MATLAB	Fascicle length, SEE, muscle activity and joint kinematics

Supplementary Table 7. First part of the comparison between fascicle tracking techniques including manual tracking, semi-automated and fully automated algorithms (each color representing one method approach).

(Kalkman et al., 2018)	Comparative study	15 CP, 16 TD	Age: 6–16 years, GMFCS I-II	Ultrasound Imaging, motion capture, force sensors	Manual fascicle length determination in the first frames, Kanade-Lucas optical flow algorithm method in MATLAB	Fascicle, muscle, and tendon lengthening during ankle rotation
(Maharaj et al., 2020)	Comparative study	8 children	Aged 10 ± 2 years	3D Gait Analysis, US, EMG, force platforms	UltraTrack: software for semi-automated tracking (MATLAB-based custom scripts)	Angles, lengths, activation patterns
(Matijevich et al., 2018)	Comparative study	8 adults	5 males, 3 females; mean age: 21 ± 2 years	Ultrasound Imaging	UltraTrack: software for semi-automated tracking (MATLAB-based custom scripts)	Length changes in the AT, MG muscle, and MTU
(May et al., 2021)	Reliability study	87 volunteers	44 males, 43 females, Age range: 13–63 years	Ultrasound Imaging	Manual digitization of images using LabVIEW software	Reliability of FL, θ, and MT for GM and GL
(Rosa et al., 2021)	Feasibility study	6 healthy adults	Aged 24 ± 4 years	B-mode ultrasound	Downsampled using Python scripts, UltraTrack software and ML models	Fascicle length changes estimated by ML models compared to UltraTrack
(Swinnen et al., 2019)	Experimental comparative study	19 trained runners, 10 habitual mid-/forefoot strikers and 9 habitual rearfoot strikers	Runners training >30 km per week	Dynamic Ultrasound Imaging, 3D motion capture, force plates, EMG	UltraTrack: software for semi-automated tracking (MATLAB-based custom scripts)	Muscle fascicle contraction velocity and shortening, Series elastic element and forces
(Van Der Zee & Kuo, 2022)	Methodological validation study	Dataset 1: Vastus lateralis from 9 subjects (6 males, 3 females). Dataset 2: Gastrocnemius lateralis from 1 subject. Dataset 3: Gastrocnemius medialis from 1 subject (previously collected data). Dataset 4: Gastrocnemius medialis	Healthy adults	Ultrasound Imaging	TimTrack algorithm combined vessel enhancement filtering, Hough transforms, and line detection techniques	Speed of processing compared to other state-of-the-art algorithms
(Van Der Zee et al., 2024)	Evaluation study	8 healthy adults	26 ± 3 years	Ultrasound imaging, EMG	UltraTrack (optical flow-based) and TimTrack (line detection-based) using a Kalman filter	Tracking accuracy, noise and drift
(Zhang et al., 2024)	Controlled intervention study	30 healthy male	15 participants trained to use a forefoot strike (FFS) pattern, 15 participants maintained their regular running habits	Ultrasound Imaging, Dynamometer, Video Analysis	UltraTrack: software for semi-automated tracking (MATLAB-based custom scripts), ImageJ software	Peak plantar flexion torque, Normalized fascicle length

Supplementary Table 8**.** Second part of the comparison between fascicle tracking techniques including manual tracking, semi-automated and fully automated algorithms.

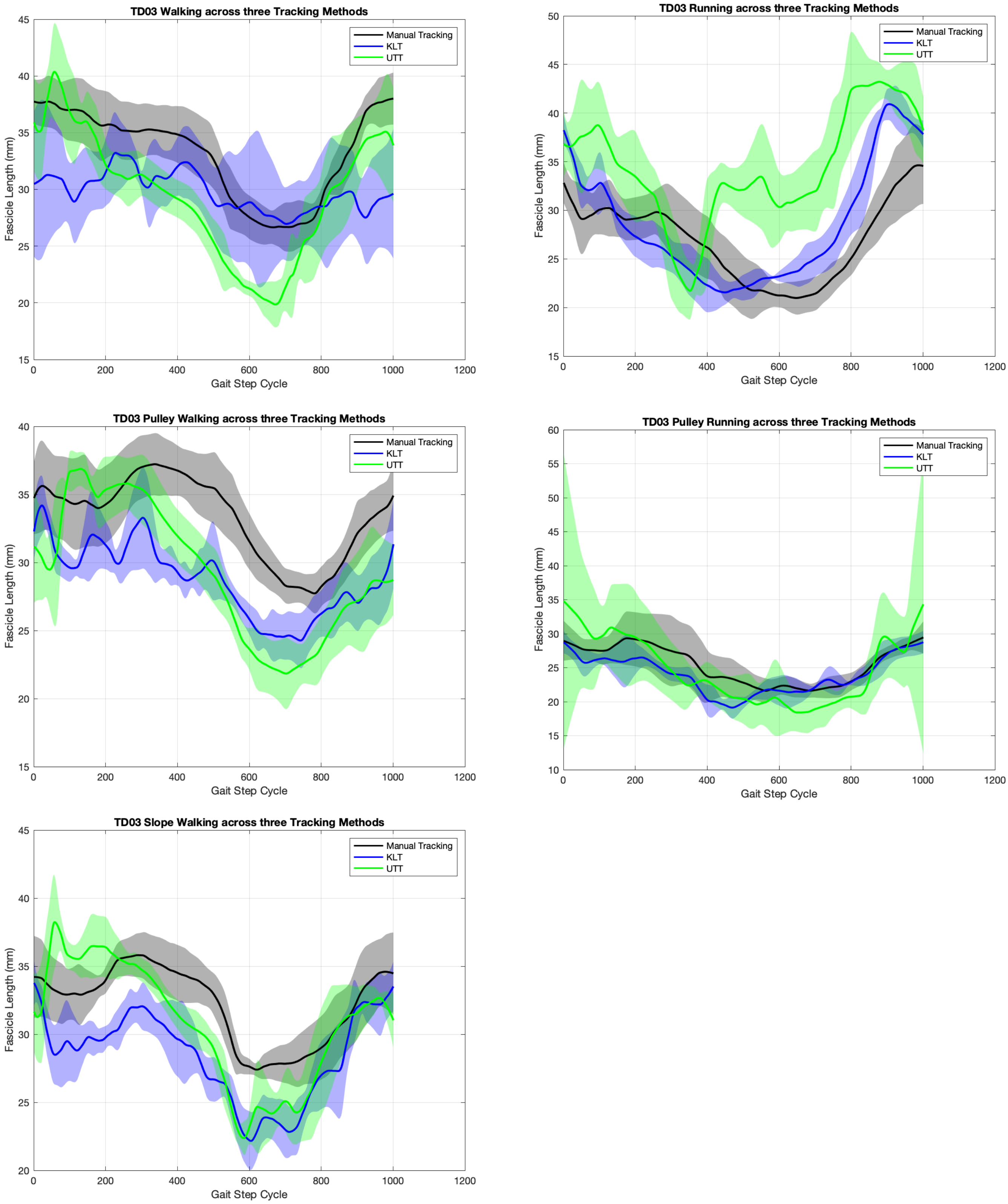
Reference	**Repeatbaility/consistency**	**Accuracy**	**Statistical Analysis**	**Movement type**	**Clinical relevance**	**Limitations**
(Aeles et al., 2017)	Intra-rater ICC: Ranged from good to excellent (minimum 0.78, maximum 0.98), Inter-rater ICC: Improved for DIT (data-informed tracking) (0.857 for initial-short, 0.818 for initial-long) compared to UFI (uninformed fascicle identification) (0.666)	DIT demonstrated higher inter-rater consistency compared to UFI	ICC, SEM, MSE and CMD, Bland-Altman plots	Passive ankle joint rotations	DIT improves reliability for measuring fascicle length changes	Small sample size, passive movement
(Barber et al., 2017)	Time-normalized gait cycles ensured comparison consistency	Inter-participant variability by normalizing kinematic and length data, Limitations in ultrasound probe alignment and modeling simplifications	Two-tailed unpaired t-tests	Walking self-selected speed over a flat, level surface	Increased reliance on passive tendon recoil for propulsion in CP gait	Small sample size
(Clark & Franz, 2018)	Strong correlations and consistency in the observed non-uniform displacement patterns	Non-uniform displacement patterns correlated well with muscle behavior	Two-way repeated measures ANOVA, Pearson correlation coefficients	Isometric contractions	Mechanistic insight into Achilles tendon behavior and its interaction with triceps surae muscles	Two-dimensional imaging, Speckle tracking and pennation angle estimation involve limitations in spatial accuracy, requires good-quality ultrasound images, struggle with fast moving images
(Clark & Franz, 2019)	Time-series data ensured consistent measurements	Limitations included possible misalignment and out-of-plane motion artifacts	Three-way mixed factorial ANOVA, Pearson's correlations	Ramped isometric voluntary contractions	Age-related subtendon changes linked to smaller ankle moments during walking	Small sample size, 2D imaging for 3D tendon structures
(Cronin et al., 2011)	Automated algorithm (CMC ranged from 0.88 ± 0.08, repeatability across conditions to 0.98 ± 0.02)	Agreement between automated and manual methods (CMC values ranged from 0.84 (jogging at 7 km/h) to 0.94 (walking at 4 km/h)) Fascicle length offsets (0.60–2.11 mm)	CMC, One-way repeated measures ANOVA	Walked at speeds of 4, 5, 6, and 7 km/h and jogged at 7 km/h on a treadmill	The automated tracking method provided reliable, time-efficient measurements comparable to manual analysis	Small sample size, manual initialization, straight-line fascicle length
(Cronin & Finni, 2013)	Step speeds in walking varied by 0.2 ± 0.1%, and in running by 1.8 ± 1.1%, ensuring consistent gait cycles	Minimal variability	Three-factor ANOVA, Bonferroni-corrected pairwise comparisons	walking (5.2 km/h) and running (10.3 km/h)	No significant differences in triceps surae fascicle behavior (length changes or velocity) between treadmill and overground or between barefoot and shod conditions	Small sample size
(Darby et al., 2011)	High repeatability (under 1 mm segmentation error) and consistent	Automated methods matched or exceeded the accuracy of manual and cross-correlation-based tracking methods (error 0.5–1.0 mm)	Pearson correlation coefficients and absolute error measures	Ankle joint rotations, Isometric contractions and deep knee bends	Enables detailed analysis of muscle and aponeurosis behavior (during dynamic tasks)	Limited sample size, tracking performance was reduced during larger, rapid movements due to feature loss
(Drazan et al., 2019)	ICCs > 0.74 for both automatic and manual approaches (repeatability), ICCs > 0.76 for manual measurements (reproducibility)	Agreement between automatic and manual measurements was strong (corrected CMC > 0.93), RMSE for fascicle length: 3.31 ± 1.53 mm, RMSE for pennation angle: 4.12 ± 1.98°	ICCs, CMC, Bland-Altman plots assessed bias and limits of agreement	Controlled maximal effort plantar flexion contractions	The algorithm offers a reliable and time-efficient method for studying muscle mechanics	Small sample size, depended on ultrasound image quality and initial user-defined tracking points, no movements

Supplementary Table 9. Second part of the comparison between fascicle tracking techniques including manual tracking, semi-automated and fully automated algorithms.

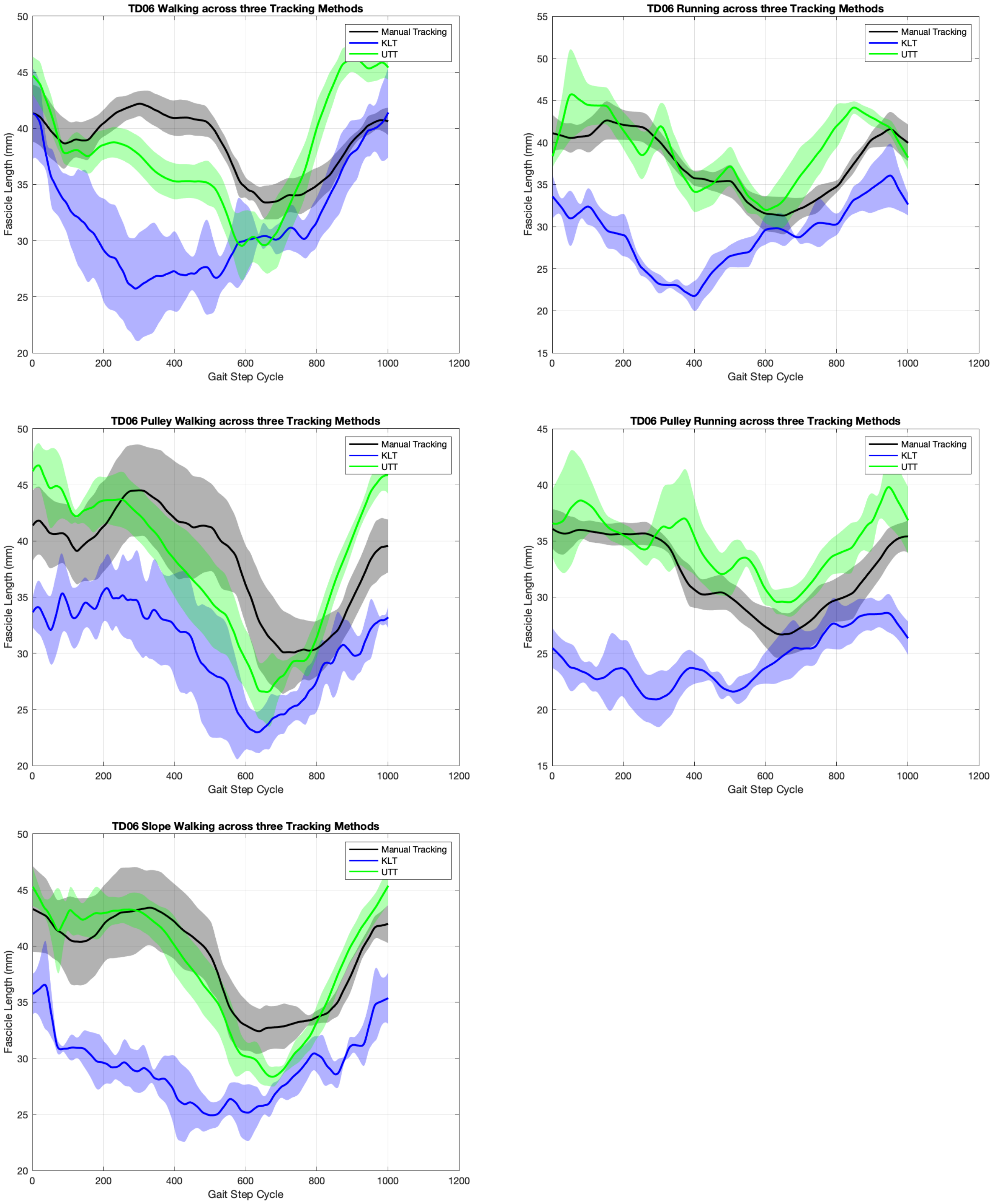
(Duclay et al., 2009)	Coefficients of variation (CVs) for fascicle and tendon measurements: 3.5–8.7%, (ICCs): 0.81–0.99	Manual measurement errors estimated at 2.4%	Two-factor repeated-measures ANOVA, regression analysis and post-hoc tests	Isometric plantar flexion contractions	TG showed significant increases in fascicle length	Small sample size
(Finni et al., 2022)	Displacement tracking (intraclass correlation of 0.967)	Tracking at the MTJ provided more accurate displacement and strain estimates. Tracking points on the aponeurosis exaggerated these measures due to the additional compliance of the aponeurosis	Two-way repeated measures ANOVA, pairwise comparisons	Isometric plantar flexion contractions	Tracking directly at the MTJ to avoid overestimating tendon displacement and strain	two-dimensional tracking, small sample size
(Frisk et al., 2019)	Consistent protocols for torque and fascicle length measurements	Direct measurement of fascicle length and torque generation	Mixed-model regression analysis, Post-hoc pairwise comparisons with Tukey adjustments identified specific group differences	Torque generation and fascicle behavior under static conditions	Impaired contractile properties in the plantar flexor muscles of adults with CP, contributing to reduced torque generation	Limited sample size, static testing conditions
(Gillett et al., 2012)	CMD = 0.86 (Intra-examiner reliability), CMD = 0.70 (Inter-examiner reliability)	High agreement between automated and manual methods	CMC, SE, CMC	Controlled joint movements	Automated tracking is reliable and accurate for analyzing MG fascicle behavior during controlled tasks	Small sample size, straight fascicle lines and not dynamic activities
(Gionfrida et al., 2023)	High repeatability for young adults, lower consistency for older adults	Variability of ~2 mm in fascicle length estimates across tasks	R^2, RMSE, MAPE, and average absolute frame difference, Statistical comparisons conducted across walking speeds and inclines	Level walking at 0.75, 1.25, and 1.5 m/s, inclined walking at 1.25 m/s with a 10% slope	Feasibility of integrating machine learning into exoskeleton systems for real-time feasibility of integrating machine learning into exoskeleton systems for real-time assistance	Small sample size, Intrinsic errors in the training dataset (UltraTrack labels) may have affected model performance,
(Hauraix et al., 2017)	R^2 (0.93 ± 0.08)	High temporal resolution	One-way ANOVA, Shapiro-Wilk tests, Hill's hyperbolic model	Maximal knee extensions	Understanding of muscle-tendon interactions	Ultrasound field-of-view limitations required fascicle length extrapolation
(Heieis et al., 2023)	Sex-specific differences but no significant age-related effects	The risk of measurement errors due to 3D effects	Linear mixed models	Isometric contractions	Relationship between fascicle curvature and biomechanical variables	Small sample size
(Hollville et al., 2020)	Good repeatability	Medium to large effect sizes	Repeated-measures ANOVA and Kruskal-Wallis tests	Maximal countermovement jumps	How surface compliance and damping affect GM and VL neuromechanics	Small sample size, technical issues
(Hösl et al., 2016)	Measurements followed standardized protocols, including repeated trials and manual inspection of ultrasound fascicle tracking	Precise and manual frame-wise inspection	Shapiro–Wilk tests, ANOVA, Mann–Whitney U tests, unpaired t-tests and Post-hoc tests included paired t-tests or Wilcoxon tests	Treadmill walking under three conditions: flat-forward, uphill, and backward-downhil	SCP-affected children have shorter gastrocnemius fascicles, reduced concentric fascicle shortening, and distinct muscle activity patterns during gait compared to TD controls	Treadmill gait may not fully replicate overground walking dynamics, fascicle behavior was normalized to resting length

Supplementary Table 10**.** Second part of the comparison between fascicle tracking techniques including manual tracking, semi-automated and fully automated algorithms.

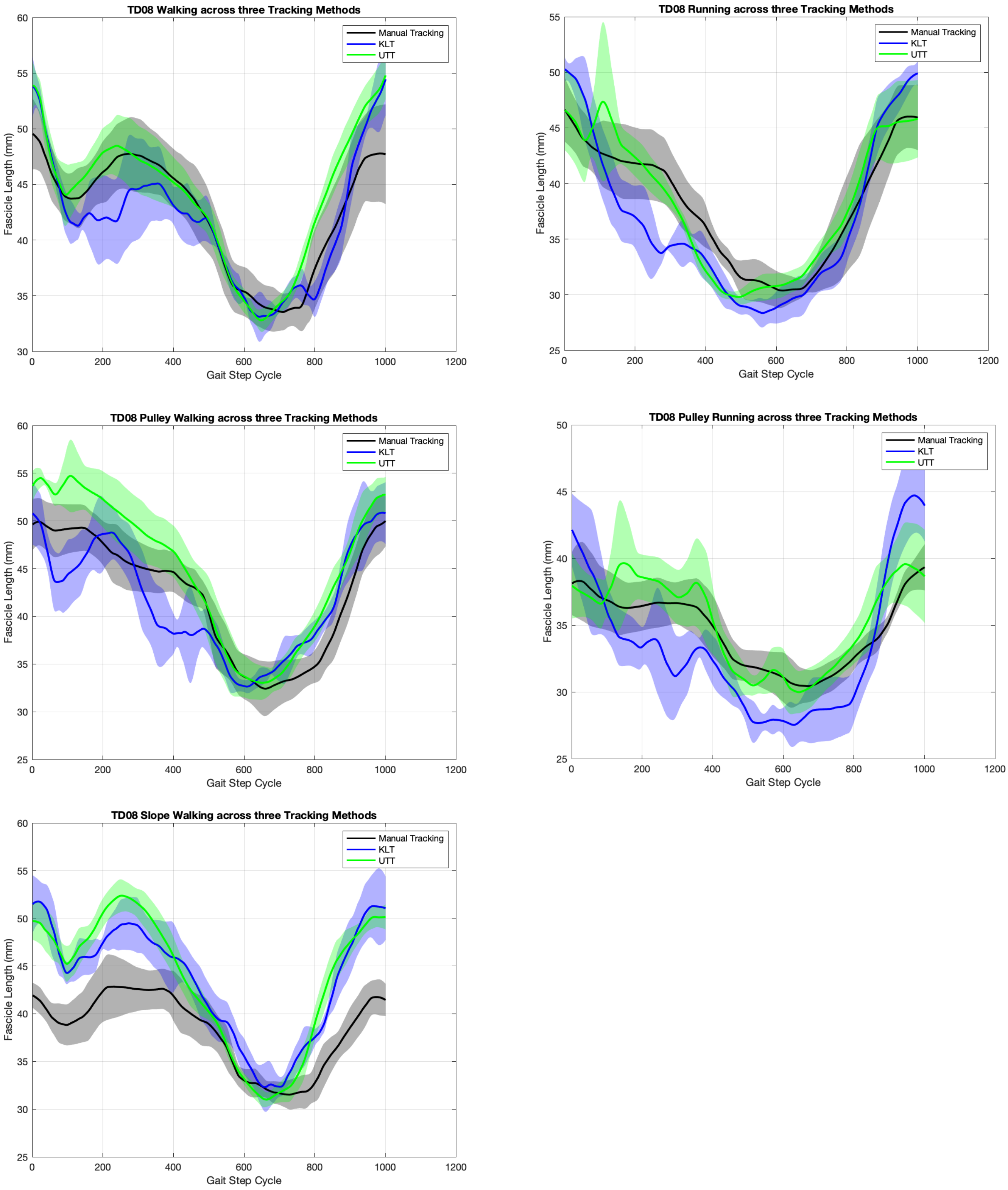
(Kalkman et al., 2018)	ICC 0.7 - 0.9	Limitations in ultrasound resolution, overestimation of tendon lengthening due to neglecting tendon curvature	Two-way mixed ICC and standard error	Passive ankle joint rotations in the sagittal plane	Muscle contributes less to MTU lengthening in CP (50%) compared to TD (63%), Tendon compensates with a higher contribution in CP children	Small sample size, Manual rotation of the joint
(Maharaj et al., 2020)	Reliable and consistent	Normalized and consistent	Paired t-tests, Normality tests	Level walking at a self-selected speed on a 10-meter walkway	Walking with flip-flops does not increase the stress or strain on MG or Achilles tendon	Small sample size, may not extend to unstable walking surfaces or children with foot abnormalities
(Matijevich et al., 2018)	AT length changes were strongly correlated between MTJ and MF methods for certain tasks, but differences emerged during heel raises	Inconsistencies between experimental estimates and model expectations	Correlation coefficients	Calf contractions, ankle dorsiflexion, heel raises	The need for improved methods to estimate tendon kinematics	No ground truth data for tendon length changes were collected, oversimplify tendon-muscle dynamics, AT dynamics were only assessed in two dimensions
(May et al., 2021)	Intra-rater and inter-rater reliability (Excellent for FL, θ, and MT (ICC ≥ 0.91)), Test-retest reliability moderate/good	17-30% from median values	ICCs, Bland-Altman plots, SEM and MDC	Static measurements	Useful for detecting population differences but less sensitive for tracking individual adaptations post-intervention	Variability in probe repositioning, dynamic tasks not tested
(Rosa et al., 2021)	Direct training: same subject and task (correlation 0.9)	Average r = 0.7 (RMSE 2.86) direct training Moderate r = 0.65 (RMSE 3.28) cross-participant	Pearson's correlation coefficient (r) and RMSE	Walking at 1.5 m/s, seated ankle flexion, calf raises, and a mix of foot movements	The feasibility of ML models	Small sample size, depends on UltraTrack software accuracy
(Swinnen et al., 2019)	Consistent across multiple strides	100 data points per stride to ensure accuracy in comparisons	Mixed ANOVA, Statistical parametric mapping	Running on a treadmill at two speeds: 10 km/h and 14 km/h	Highlights how foot-strike patterns may influence metabolic energy expenditure and running mechanics	Small sample size, stiffness was assumed constant, Muscle force estimates were model-dependent
(Van Der Zee & Kuo, 2022)	Fascicle length RMSD: 0.52 cm, (ICCs) ranged from 0.93 to 0.99	CMC for fascicle length = 0.98, Performed better than or equal to optic flow-based algorithms like UltraTrack	Linear regressions, RMSD and MAD and ICC values	Static and dynamic muscle tasks	Eliminates drift sensitivity issues	Curvature Limitations and image Quality Sensitivity
(Van Der Zee et al., 2024)	High consistency, with low variability across contraction cycles and image acquisition settings	Fascicle length RMSE: 2.7 mm, fascicle angle RMSE: 0.7°	Paired t-tests, linear regression models	Submaximal fixed-end plantar flexion contractions, maximal voluntary contractions and passive ankle rotations at varying velocities	Generalizability to other muscles is uncertain, controlled experimental conditions	Improves the precision and reliability of fascicle tracking in ultrasound imaging
(Zhang et al., 2024)	Consistent supervised intervention	Precise in vivo measurements	Two-way repeated measures ANOVA, Shapiro–Wilk	Walking and running	Gait retraining as a viable strategy for improving muscle function, mechanical efficiency, and reducing injury risk	Small sample size, no long-term follow-up



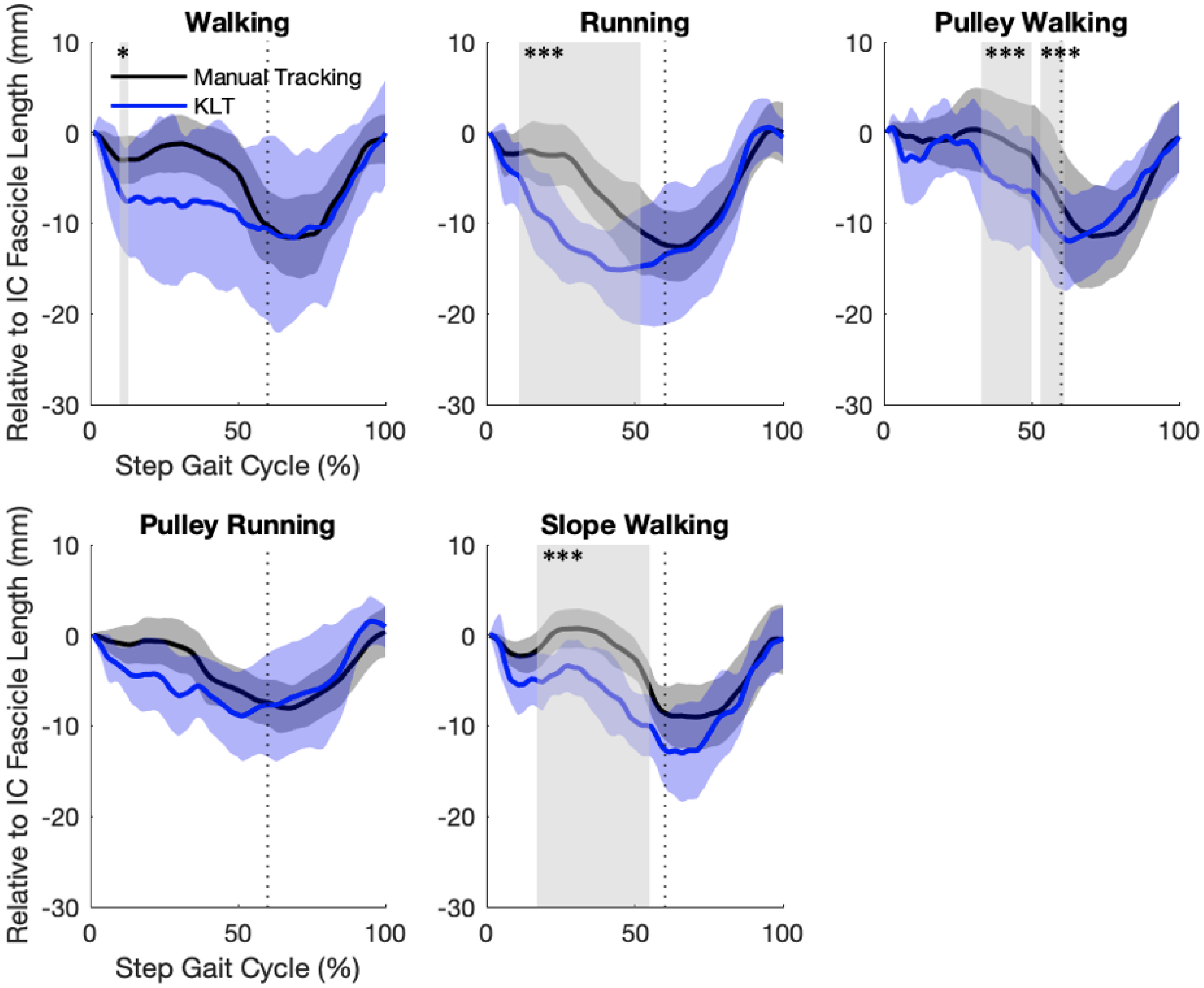
Supplementary Fig. 1. Graphs above show the fascicle length of child TD03 across all functional tasks, manual tracking, KLT and UTT.



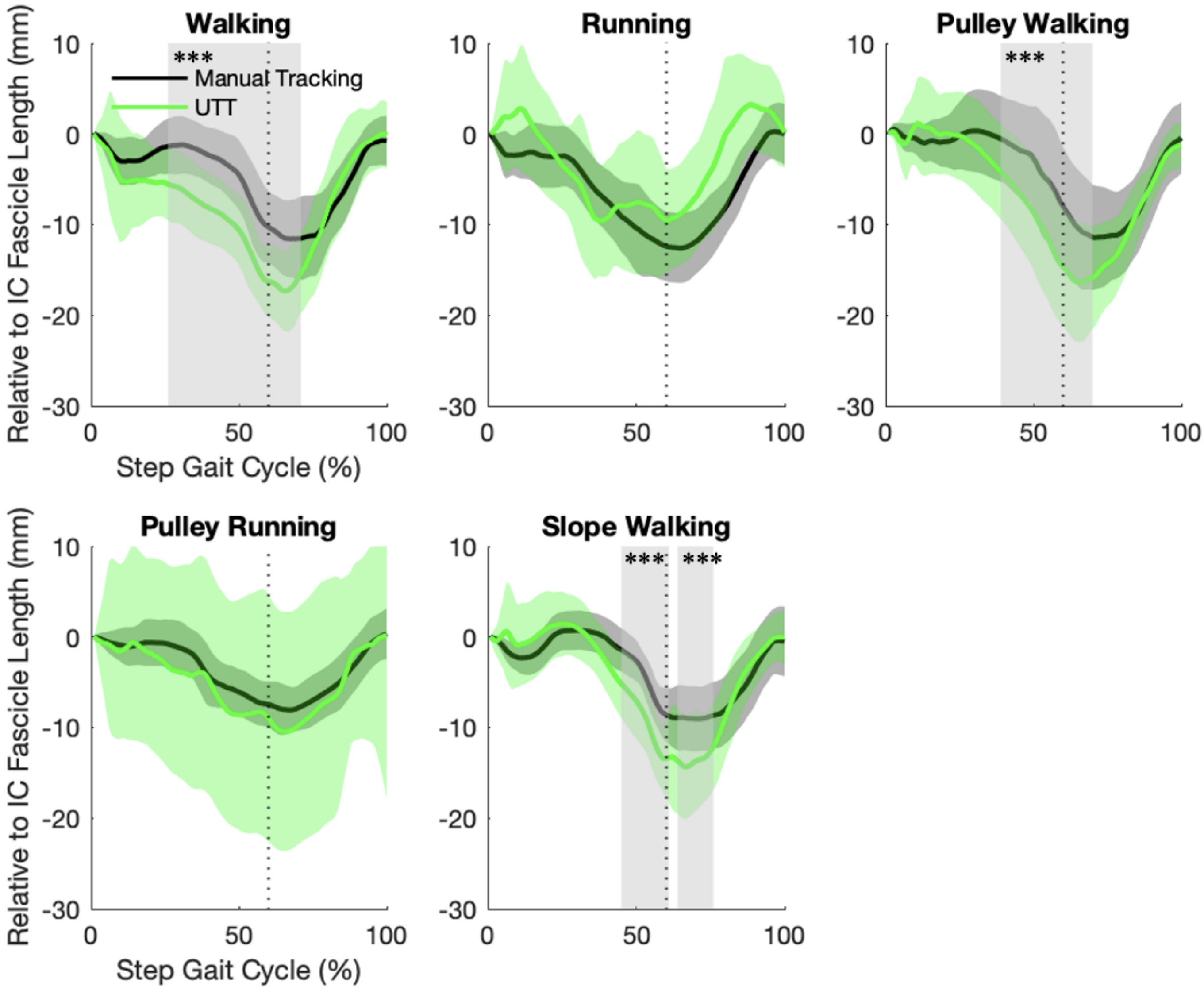
Supplementary Fig. 2. Graphs above show the fascicle length of child TD06 across all functional tasks, manual tracking, KLT and UTT.



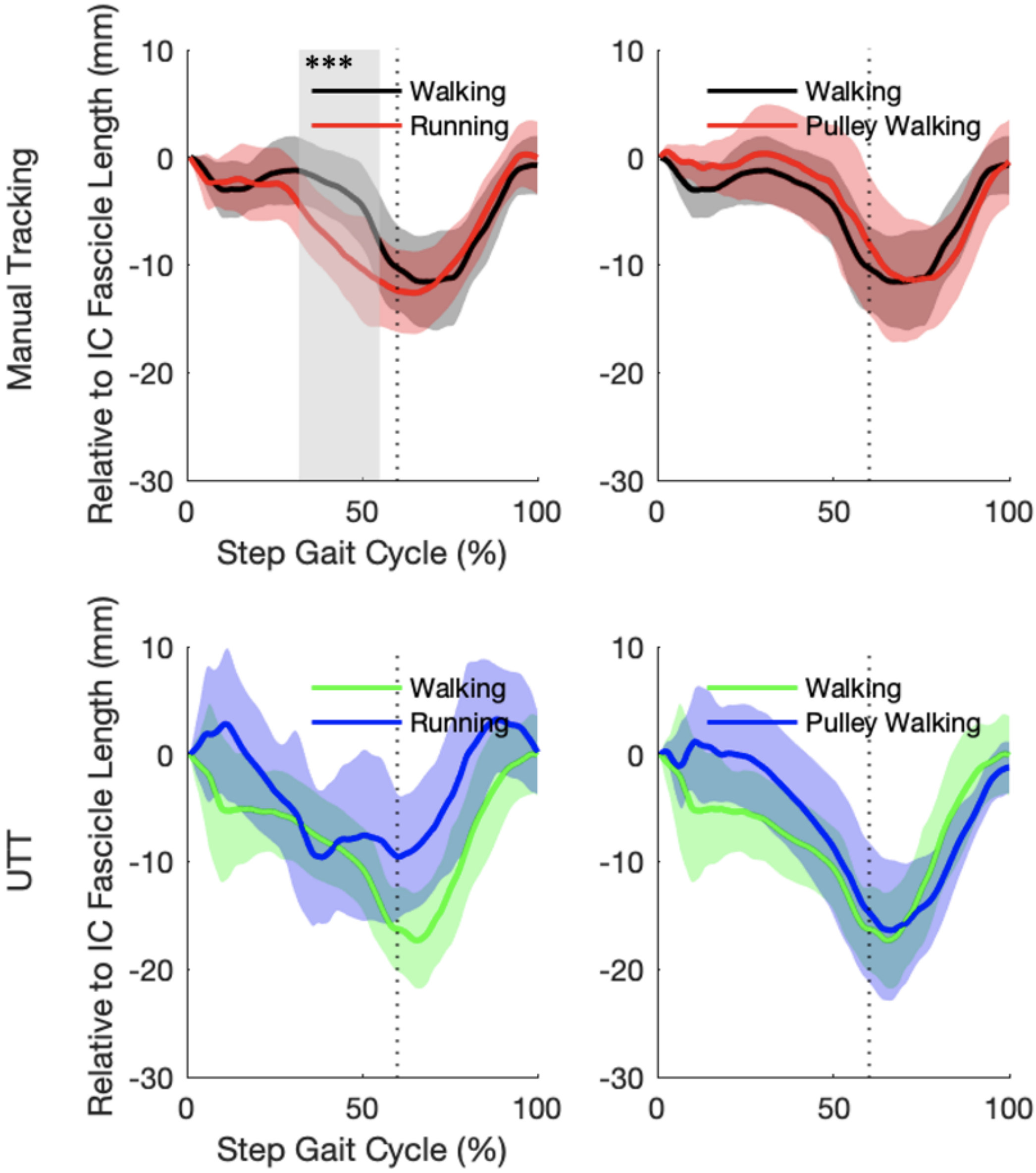
Supplementary Fig. 3. Graphs above show the fascicle length of child TD08 across all functional tasks, manual tracking, KLT and UTT.



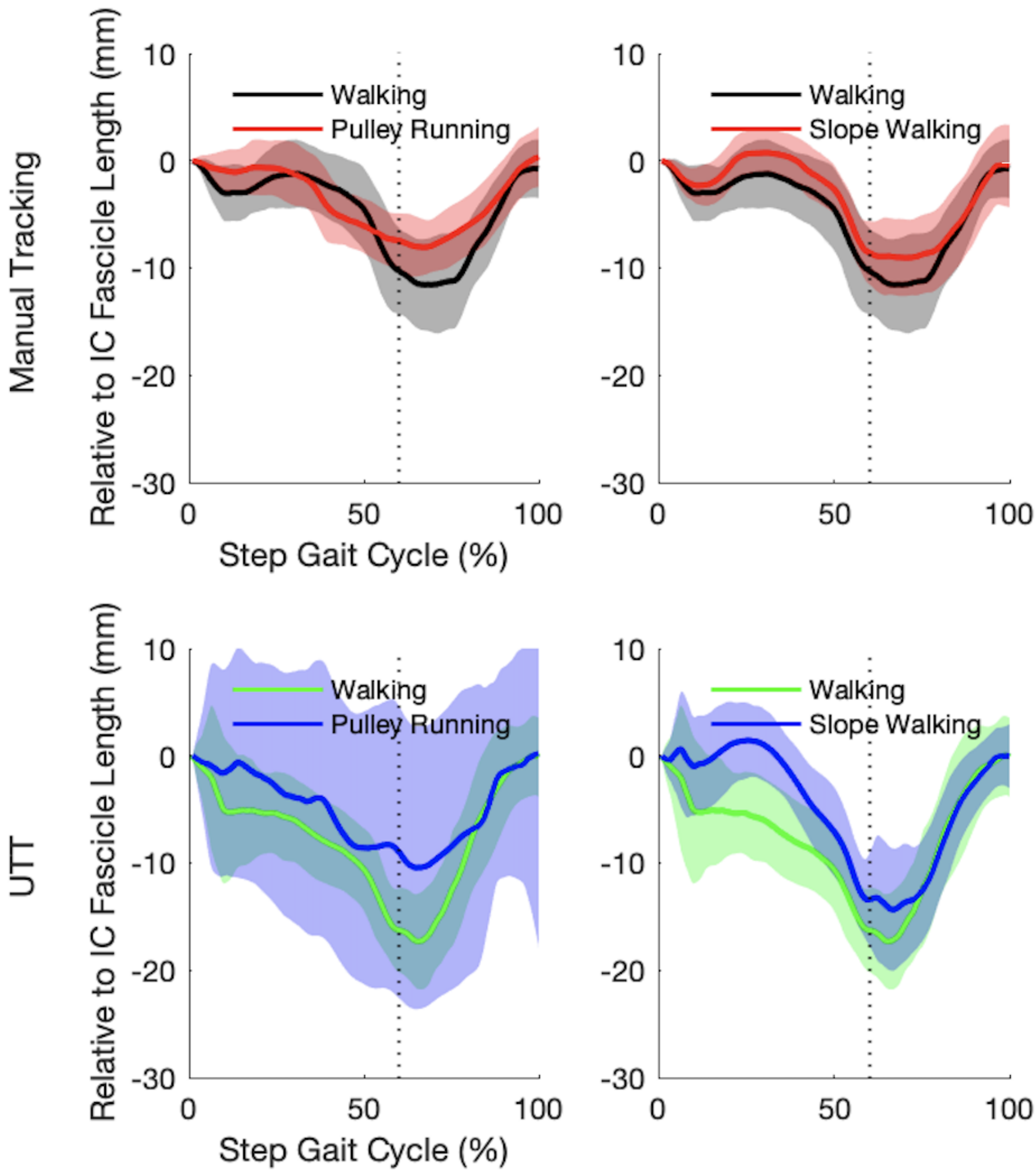
Supplementary Fig. 4. Relative to Initial Contact (IC) fascicle length across the gait cycle for three children, with standard deviation calculated across five steps per child. Data are shown for five functional tasks for manual tracking and KLT. Significantly different parts of the curves are highlighted in light grey. Significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$. The x-axis represents the normalized gait cycle (0–100 %), while the y-axis indicates the fascicle length. The dotted vertical line at 60 % of the gait cycle marks the separation between the stance and swing phase.



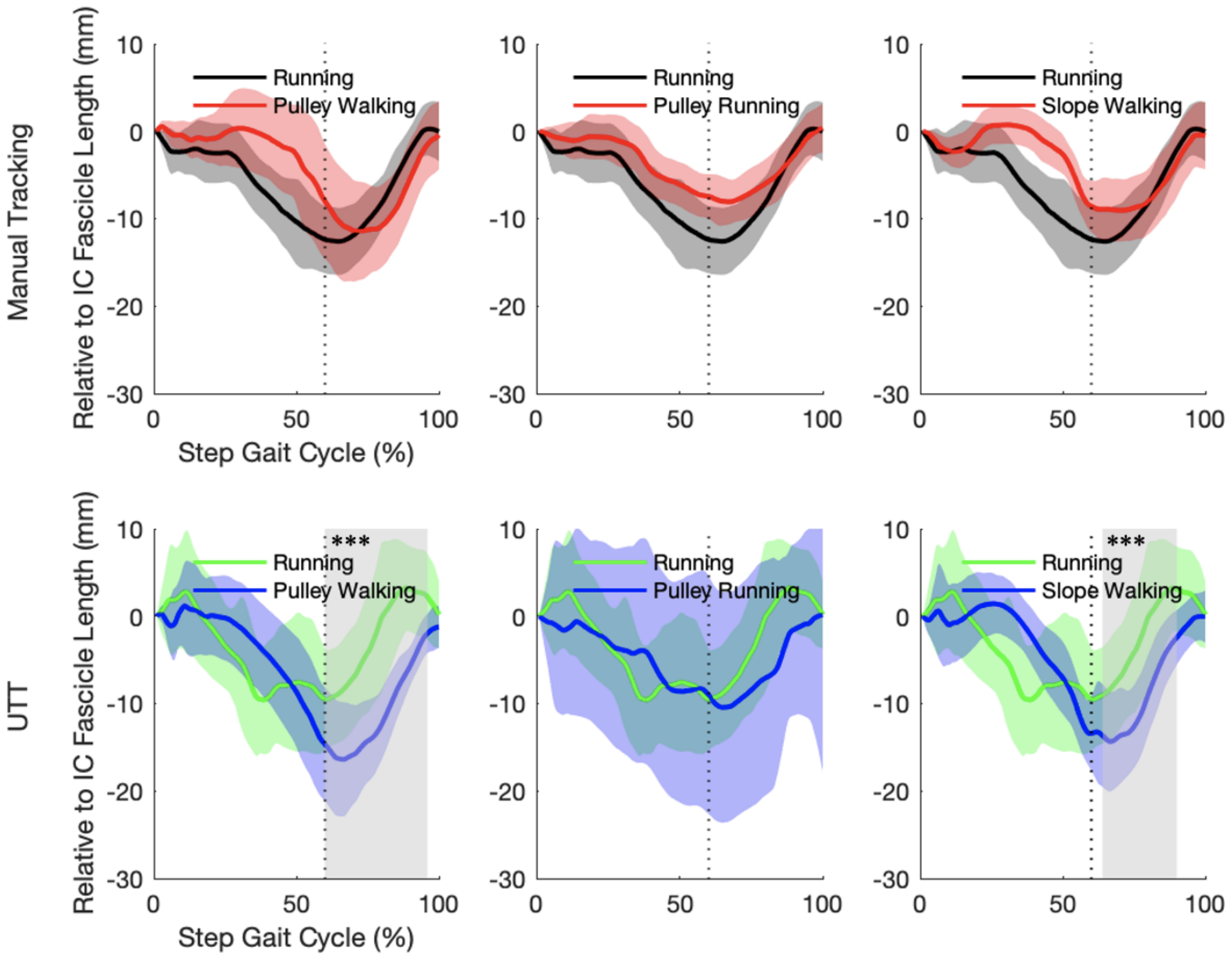
Supplementary Fig. 5. Relative to Initial Contact (IC) fascicle length across the gait cycle for three children, with standard deviation calculated across five steps per child. Data are shown for five functional tasks for manual tracking and UTT. Significantly different parts of the curves are highlighted in light grey. Significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$. The x-axis represents the normalized gait cycle (0–100 %), while the y-axis indicates the fascicle length. The dotted vertical line at 60 % of the gait cycle marks the separation between the stance and swing phase.



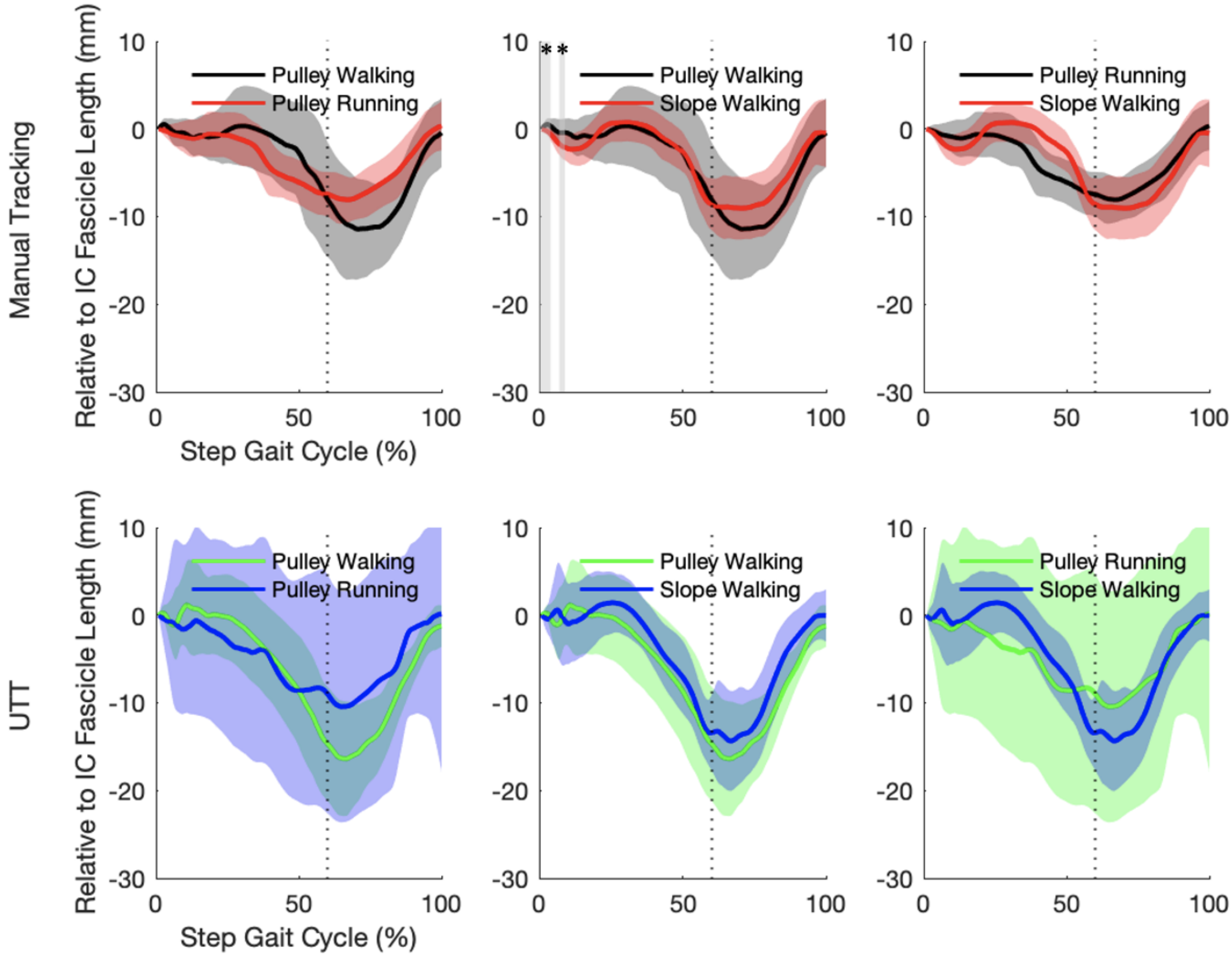
Supplementary Fig. 6. Relative to Initial Contact (IC) fascicle length across the gait cycle for three children, with standard deviation calculated across five steps per child. Data are shown for two different comparisons between functional tasks for manual tracking and UTT. Significantly different parts of the curves are highlighted in light grey. Significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$. The x-axis represents the normalized gait cycle (0–100 %), while the y-axis indicates the fascicle length. The dotted vertical line at 60 % of the gait cycle marks the separation between the stance and swing phase.



Supplementary Fig. 7. Relative to Initial Contact (IC) fascicle length across the gait cycle for three children, with standard deviation calculated across five steps per child. Data are shown for two different comparisons between functional tasks for manual tracking and UTT. Significantly different parts of the curves are highlighted in light grey. Significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$. The x-axis represents the normalized gait cycle (0–100 %), while the y-axis indicates the fascicle length. The dotted vertical line at 60 % of the gait cycle marks the separation between the stance and swing phase.



Supplementary Fig. 8. Relative to Initial Contact (IC) fascicle length across the gait cycle for three children, with standard deviation calculated across five steps per child. Data are shown for three different comparisons between functional tasks for manual tracking and UTT. Significantly different parts of the curves are highlighted in light grey. Significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$. The x-axis represents the normalized gait cycle (0–100 %), while the y-axis indicates the fascicle length. The dotted vertical line at 60 % of the gait cycle marks the separation between the stance and swing phase.



Supplementary Fig. 9. Relative to Initial Contact (IC) fascicle length across the gait cycle for three children, with standard deviation calculated across five steps per child. Data are shown for three different comparisons between functional tasks for manual tracking and UTT. Significantly different parts of the curves are highlighted in light grey. Significant differences are denoted as follows: $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$. The x-axis represents the normalized gait cycle (0–100 %), while the y-axis indicates the fascicle length. The dotted vertical line at 60 % of the gait cycle marks the separation between the stance and swing phase.