

# Fall Risk and Gait Analysis in Community-Dwelling Older Adults using World-Spaced 3D Human Mesh Recovery

Chitra Banarjee<sup>1</sup> Patrick Kwon<sup>2</sup> Ania Lipat<sup>3</sup> Rui Xie<sup>4</sup> Chen Chen<sup>2</sup> Ladda Thiamwong<sup>3</sup>

<sup>1</sup>College of Medicine, UCF <sup>2</sup>Institute of AI, UCF

<sup>3</sup>College of Nursing, UCF <sup>4</sup>College of Sciences, UCF

chitra.banarjee@ucf.edu yo564250@ucf.edu ania.lipat@ucf.edu  
rui.xie@ucf.edu chen.chen@ucf.edu ladda.thiamwong@ucf.edu

## Abstract

*Gait assessment is a key clinical indicator of fall risk and overall health in older adults. However, standard clinical practice is largely limited to stopwatch-measured gait speed. We present a pipeline that leverages a 3D Human Mesh Recovery (HMR) model to extract gait parameters from recordings of older adults completing the Timed Up and Go (TUG) test. From videos recorded across different community centers, we extract and analyze spatiotemporal gait parameters, including step time, sit-to-stand duration, and step length. We found that video-derived step time was significantly correlated with IMU-based insole measurements. Using linear mixed effects models, we confirmed that shorter, more variable step lengths and longer sit-to-stand durations were predicted by higher self-rated fall risk and fear of falling. These findings demonstrate that our pipeline can enable accessible and ecologically valid gait analysis in community settings. Project website : [https://yj7082126.github.io/fallrisk\\_gvhmr/](https://yj7082126.github.io/fallrisk_gvhmr/)*

## 1. Introduction

Movement is central to human well-being across the lifespan [7, 11, 17]. For older adults, movement is an indicator of overall health in the physical, mental, and cognitive domains [10, 12, 55]. As such, gait evaluation is a key clinical measure for identifying older adults at risk for several physiological and pathological aging conditions, commonly used in primary care visits and geriatric assessments to screen participants at risk for frailty or falls [30, 51].

However, clinical gait evaluation remains largely limited to stopwatch-measured gait speed, as comprehensive assessments are constrained by limited access to technology and specialized training. Although the biomechanical correlates of fall risk are well-established, existing method-

ologies, such as wearable inertial sensors, optoelectronic marker-based systems, and multi-camera markerless motion capture, require dedicated infrastructure and considerable technical training, restricting their deployment beyond controlled clinical or research settings.

To solve this, we present a novel computer vision-informed approach that evaluates 3D gait measures using a Human Mesh Recovery (HMR) method, from a single monocular video camera that records the Timed Up and Go (TUG) test for older adults. While prior studies relied on 2D keypoints [9, 16, 58] or multi-camera setups, we leverage Ground View HMR (GVHMR) [53] as the foundation of our pipeline to recover full human body motion anchored in *world coordinates*. Unlike camera-relative methods that conflate human and camera motion, GVHMR reconstructs the participant’s true trajectory through space, enabling the extraction of spatiotemporal gait parameters in *absolute metric units*. Also, since older adult populations are sensitive to environments where unfamiliar equipment can limit participation, our approach is well-suited for deployment in real-world community settings.

Using GVHMR, we extract spatiotemporal variables from videos of older adults completing the TUG taken across several community and senior centers, which are compared to in-shoe pressure insoles to evaluate validity. We also explore how measures of fall risk predict various aspects of gait, and evaluate the use of HMR methods to analyze movement in community settings.

Our main contributions are summarized as follows:

- A gait acquisition protocol that uses a single camera to record older adults completing the TUG test across multiple community centers, only requiring limited clinical infrastructure and supervision.
- A novel application of world-grounded 3D HMR to clinical gait analysis, leveraging GVHMR to reconstruct full-body trajectories and spatiotemporal gait parameters.

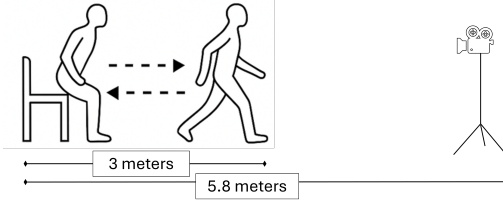


Figure 1. Schematic of Timed Up and Go (TUG) test set up

- An automated pipeline extracting spatiotemporal parameters of gait and TUG transitions via signal processing and peak detection.
- A correlation analysis comparing video-derived step time against insole measurements, and a linear mixed effects (LME) framework evaluating how fall risk factors predict spatiotemporal gait in community-dwelling older adults.

## 2. Related Work

Several studies have demonstrated the relevance of gait metrics for fall risk in older adults. For instance, [8, 27] showed that shorter step length and increased step width variability are associated with adverse outcomes, while pressure-based measures have been linked to both fall risk and cognitive impairment [1, 2, 40]. Beyond spatial measures, segmentation of TUG subtasks (e.g. standing up and turning) has demonstrated that increased turn duration and sit-to-stand times are associated with higher fall risk [34, 36, 45].

Recently, computer vision has been used for various healthcare applications, for both medical images [15, 18, 23, 26, 62] and human recordings [22, 32, 48, 50, 52]. With the advent of deep learning, human pose estimation methods have greatly evolved over the past decades [39, 65], advancing research in medical fields such as gait analysis and sports optimization [35, 41, 49]. Several works in gait analysis further explored how these methods can increase the accessibility of fall risk assessment in labs and assisted living centers [5, 21, 37, 42, 60], though deployment in uncontrolled community environments remains limited [14, 66]. These few studies in community settings rely on 2D skeleton-based models [9, 16, 58, 61], which cannot recover depth information or separate camera perspectives from human pose, making them ill-suited for extracting spatial measures such as step length. We aim to resolve this by using GVHMR [53], which, unlike other pose estimation or human mesh recovery methods [20, 28, 54, 64], efficiently predicts human poses in a gravity-aware world coordinate system, enabling extraction of spatiotemporal gait parameters from a single monocular camera.

## 3. Methodology

### 3.1. Data Collection

*Video Timed Up and Go (vTUG).* Participants were instructed to complete the Timed Up and Go (TUG) test three times. The trials were recorded using a conventional

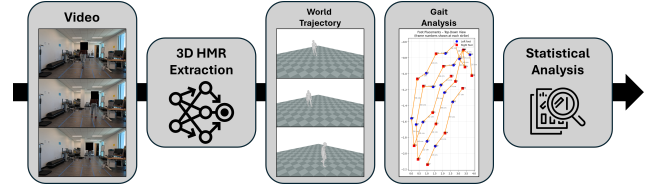


Figure 2. The overall pipeline of our study. Basing on the world trajectory extraction results from the TUG test using GVHMR [53] (Sec. 3.2), we analyze the gait factors (Sec. 3.3) and conduct statistical analysis (Sec. 3.4).

monocular video camera, GoPro Hero 12, set up approximately 280 cm from the 3-meter mark (Fig. 1). The trials were then edited by trained URAs using Adobe Premiere Pro to trim and crop each video (30 fps).

*Instrumented Timed Up and Go (iTUG).* In addition to the video, we used the wearable XSENSOR insole system [46] (XSENSOR Technology Corporation, Calgary, Canada, shown in **Appendix** Fig. 5). The sensors are placed on the lateral malleoli of the ankle and collected data at 60 fps for accelerometer and gyroscope in three dimensions ( $x, y, z$ ). These sensors served as a comparative reference for evaluating the video-derived gait measures. While not a perfect ground truth, due to differences in coordinate systems, sampling rates, and our own processing choices, the insole data provides an independent sensor-based benchmark to assess the temporal performance of the video pipeline.

*Fall risk measures.* Participants were given questionnaires to evaluate subjective fall risk factors and a postural sway assessment to quantify balance (Table 1).

### 3.2. Human Pose Estimation

The TUG (Fig. 1) consists of standing up from a chair, walking, turning, and coming back to sit in the chair. Each TUG video is processed through GVHMR [53], which recovers human pose separately for camera-space and world-space. Specifically, given a video  $\{I^t\}_{t=0}^T$ , GVHMR predicts the local body poses  $\{\theta^t \in \mathbb{R}^{21 \times 3}\}_{t=0}^T$ , body shape coefficients  $\beta \in \mathbb{R}^{10}$  of SMPL-X [47], the orientation  $\{\Gamma_c^t \in \mathbb{R}^3\}_{t=0}^T$  and translation  $\{\tau_c^t w \in \mathbb{R}^3\}_{t=0}^T$  corresponding to the camera-space trajectory, and the orientation  $\{\Gamma_w^t \in \mathbb{R}^3\}_{t=0}^T$  and translation  $\{\tau_w^t \in \mathbb{R}^3\}_{t=0}^T$  corresponding to the gravity-aligned world-space trajectory. Here, we mainly focus on the world-space trajectory information. The gait parameters are then derived from the 3D joint positions  $\{J^t \in \mathbb{R}^{24 \times 3}\}_{t=0}^T$ , regressed from the SMPL-X kinematic model in world space. We present the TUG test setup in Fig. 1, and an example prediction result in Fig. 2. Further details are provided in **Appendix**.

### 3.3. Gait Analysis

Using the world-space joint trajectories, we derive a set of spatiotemporal gait parameters that characterize participants' mobility during the TUG test. These include tempo-

Variable	Assessment	Significance/Interpretation
Self-Rated Fall Risk	Centers for Disease Control STEADI Score	12-item questionnaire quantifying self-rated fall risk, summed to produce the STEADI score, with higher scores indicating higher risk. Validated as indicator of fall risk [4, 43].
Fear of Falling (FoF)	Short Falls Efficacy Scale-International (FES-I) Score	7-item questionnaire evaluating FoF, with higher scores indicating higher FoF. Validated as indicator of fall risk [13, 25, 63].
Postural Sway	BTrackS Balance Test Score	Static balance assessment in which participants stand on a balance plate for 20 seconds three times to evaluate postural sway, with higher measures indicating worse static balance. Validated as indicator of fall risk [19, 31, 44].

Table 1. The measures in this table were completed by the participants to provide information about subjective and objective fall risk, specifically self-rated fall risk, fear of falling, and postural sway.

ral and spatial gait measures, as well as transition durations between subtasks (e.g., sit-to-stand and turning). The metrics are computed from the video-derived joint trajectories, while wearable insole sensors are used as an independent reference to compare temporal gait measurements (Fig. 2).

Prior to analysis, we reduce high-frequency noise by applying Gaussian smoothing to each time series of 3D body keypoints from video data and to the sensor signals. A discrete Gaussian kernel was constructed with  $\sigma = 3$  samples. The kernel support was limited to  $\pm 3\sigma$ , resulting in a 19-point symmetric filter. The kernel was then normalized to unit sum prior to convolution.

*Step Time (ST)*. For the video-derived data, ST was calculated using GVHMR joint positions. For the insole-derived data, the raw data was exported through XSENSOR Pro Foot & Gait System, and was calculated using peak detection with angular velocity (gyroscope  $z$ -axis) and vertical acceleration of ankles.

*Step Length (SL) and Step Width (SW)*. These were extracted by detecting the local minima of the GVHMR ankle joint positions to identify steps, then calculating the forward and lateral directions for SL and SW, respectively.

*Sitting, Standing, Turning Transitions*. Similarly, peak detection for sit-to-stand (STS-1) and stand-to-sit (STS-2) was computed by calculating a composite sit-to-stand (STS, Eq. 1) and hip line signal (Eq. 2), which utilize the hip and shoulder coordinates to segment subtasks (Table 2).

$$\text{STS} = 1.0 \cdot \dot{y}_{\text{hip}} + 0.7 \cdot \dot{z}_{\text{shoulder}} + 0.5 \cdot \dot{\theta}_{\text{trunk}} \quad (1)$$

$$\text{Hip Line Signal} = x_{R\text{hip}} - x_{L\text{hip}} \quad (2)$$

Further details are provided in **Appendix**.

### 3.4. Statistical Analysis

To evaluate the reliability and clinical relevance of the extracted gait metrics, we conducted two sets of statistical analyses. First, we assessed the agreement between video-derived gait measures and wearable sensor measurements to validate the temporal accuracy of the video pipeline. Second, we examined whether the extracted gait parameters were associated with established fall risk indicators.

*Correlation Analysis*. For this analysis, we aim to measure the correlation between ground-truth insole-derived sensor data and HMR-based video-derived data, to validate

Component	Indicator
$\dot{y}_{\text{hip}}$	Vertical hip trajectory indicates initiation of STS, weighted the highest.
$\dot{z}_{\text{shoulder}}$	Anterior translation of the upper trunk during forward flexion; reliably marks the preparatory lean phase that precedes vertical rise.
$\dot{\theta}_{\text{trunk}}$	Angular trunk flexion–extension is noisier and more sensitive to small movements, thus assigned a lower weight.
Hip Line Signal	Alternating signed extrema in hip-line velocity reflect rapid angular deceleration and reversal of pelvic rotation during turning.

Table 2. Components of signals for segmentation of TUG subtask (sit-to-stand, turns, stand-to-sit).

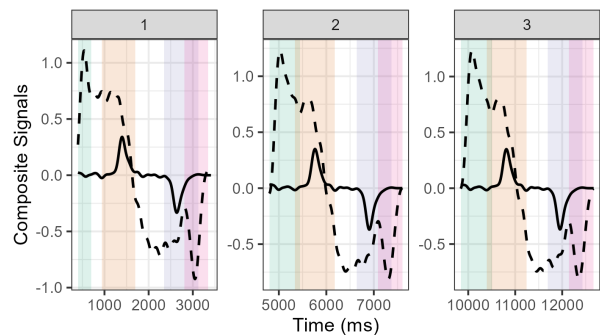


Figure 3. Automated Subtask Segmentation for three trials. Solid line refers to the hip line velocity and dashed line refers to sit-to-stand and stand-to-sit (STS) signal.

the usability of HMR. To improve the accuracy of this comparison and account for differences in each trial, only participants with three valid trials of the iTUG and vTUG were included for this analysis. The correlation analysis was conducted for each participant and trial from the 90 trials from 30 participants through a Spearman’s correlation test after evaluation of normality using the Shapiro–Wilks test.

*Linear Mixed Effects Models (LME)*. Because each participant completed three trials of the TUG, the observations are not independent. As such, LMEs were used to model the data. Fixed effects consisted of fall risk factors (Tab. 1) and age, with a random effect of participant to account for within-person variance (see Eq. (3)). Models predicted gait parameters (mean STS and durations, SL and SW mean, variability, and symmetry) for each trial.

## 4. Experiments

### 4.1. Dataset Characteristics

The final sample included 207 video recordings of the vTUG (video) from 52 older adults ( $M_{\text{age}} = 74.2 + 7.81$  years), after removing participants without 3 valid trials. Participant demographics, including average self-rated fall risk, fear of falling, and postural control are shown in Tab. 4 of Appendix. Of those 207 trials, 30 assessments with 90 trials contained simultaneously collected iTUG. Further details are provided in Appendix.

### 4.2. Insole and Video Comparison

The Spearman’s rank correlation analysis indicates moderate agreement between insole-derived and video-derived ST ( $n = 90, \rho = 0.673, p < 0.001$ ). However, video-derived ST systematically underestimated the insole measurements.

### 4.3. LME Analysis with Fall Risk Factors

Self-rated fall risk (STEADI Score) significantly predicted STS duration. Turn durations were not predicted by any of the fall risk factors. Spatial measures of gait involved SL and SW, but only SL and SL variability were predicted by a fixed effect of self-rated fall risk (STEADI score) and fear of falling (FoF, short FES-I, Table 3). SW variability was predicted by age ( $p < 0.01$ ) but not fall risk factors (Tab. 7). The other measures, SW and symmetry measures, were not predicted by age or fall risk factors (Fig. 3, Table 3). SL (Tab. 6) and SL Variability (Tab. 8) showed far stronger between-participant consistency ( $\text{ICC} = 0.81$ ) and a substantially better model fit (conditional  $R^2 = 0.85$ ) than STS durations (Tab. 9). Additional details about the statistical measures are described in Appendix.

## 5. Discussion

We demonstrate the novel application of GVHMR to extract clinically meaningful gait metrics from monocular videos of older adults completing the TUG, with derived measures showing moderately strong correlations with simultaneously collected wearable-sensor data. Using LME models, we found that lower STEADI scores predicted longer sit-to-stand durations, though not turning durations (Tab. 9). This result is in line with previous findings indicating a relationship between STS and fall risk. However, the null turning result contrasts with other literature that demonstrates a strong relationship with fall risk [29, 36, 45]. Turn strategy (distinct or overlapping) varies widely between individuals, which may result in imprecise turn segmentation [59]. Future work can utilize GVHMR to evaluate turning strategy as an individual factor. Shorter and more variable step length was associated with greater self-rated fall risk and fear of falling (Tab. 6), consistent with literature

	STEADI Score	
	Estimate (CI)	$p$
STS Duration	1.23 (0.45, 2.01)	<b>0.002</b>
Step Length (SL)	-1.36 (-2.03, -0.68)	<b>&lt; 0.001</b>
SL Variability	-19.62 (-30.44, -8.80)	<b>&lt; 0.001</b>
	Short FES-I Score	
	Estimate (CI)	$p$
STS Duration	0.71 (-0.01, 1.42)	0.052
Step Length (SL)	-1.04 (-1.65, -0.43)	<b>0.001</b>
SL Variability	-12.76 (-22.43, -3.09)	<b>0.010</b>

Table 3. Linear mixed effect model results of STEADI Score (self-rated fall risk) and short FES-I (fear of falling). ST refers to the video-derived step time. Estimate refers to the  $\beta$  coefficient, CI = confidence interval,  $p = p$ -value.

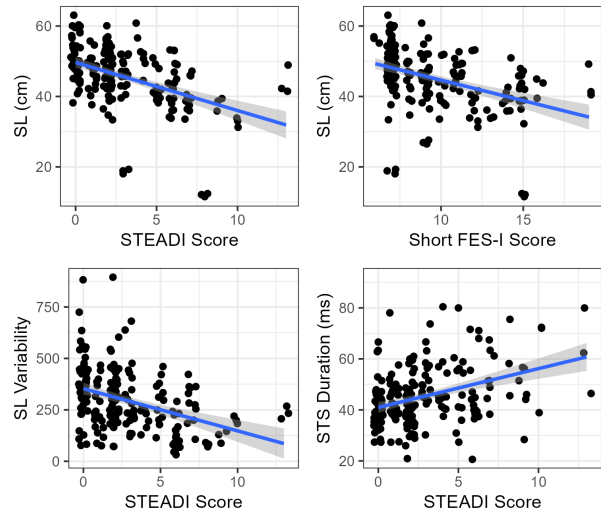


Figure 4. Scatterplots of mean step length (SL), SL variability, and sit-to-stand and stand-to-sit (STS) duration with fall risk factors (STEADI: self-rated fall risk, Short FES-I: fear of falling).

linking cautious gait patterns to both perceived and physiological instability [6, 27, 38]. Greater between-participant consistency and model fit indicate that gait characteristics such as SL may be more stable and more strongly tied to fall-risk status than STS performance, which exhibited high within-participant variability (Tabs. 6 and 9). These results suggest that older adults with poorer balance or reduced confidence tend to take smaller, more inconsistent steps and require more time to complete sit-to-stand transitions. These patterns may reflect compensatory strategies related to muscle weakness or fear of falling.

## 6. Conclusion

We offer a feasible and clinically relevant method for extracting gait and transitional movement metrics from a single TUG recording. GVHMR’s ability to segment short, meaningful subtasks and produce measures that align with wearable-derived metrics highlights its potential as an accessible assessment tool. Associations between fall risk factors and specific gait features—particularly step length

variability and sit-to-stand duration-underscore the value of video-based analysis for identifying subtle mobility changes. Future work should evaluate whether the derived metrics can enhance fall risk screening, support remote assessments, or predict prospective functional decline.

## References

- [1] S. Akter, T. M. Guess, S. Sarker, et al. Explainable machine learning for early detection of mild cognitive impairment, fall risk, and frailty using sensor-based motor function data. *medRxiv*, 2025. 2025.12.23.25342943. [2](#)
- [2] C. Banarjee, P. M. Maldonado, M. S. B. Hossain, H. Choi, R. Xie, and L. Thiamwong. Associations of gait characteristics with fall risk and frailty in older women. *Innovation in Aging*, 9(Supplement\_2):igaf122.3357, 2025. [2](#)
- [3] D. Bates. lme4: Linear mixed-effects models using Eigen and S4. *R Package Version*, 1(1), 2016. [2](#)
- [4] G. Bergen and I. Shakya. CDC STEADI: Evaluation guide for older adult clinical fall prevention programs. Technical report, Centers for Disease Control and Prevention, 2019. [3](#)
- [5] J. D. Blasco-García, G. García-López, M. Jiménez-Muñoz, et al. A computer vision-based system to help health professionals to apply tests for fall risk assessment. *Sensors*, 24(6): 2015, 2024. [2](#)
- [6] A. Brachman, V. Hadyk, A. Kamieniarz-Olczak, and A. Nawrat-Szołtysik. Examining the role of fear of falling on gait parameters and short-term gait adaptation in older adults. *Journal of Clinical Medicine*, 14(23):8311, 2025. [4](#)
- [7] A. S. Buchman, P. A. Boyle, R. S. Wilson, D. A. Fleischman, S. Leurgans, and D. A. Bennett. Association between late-life social activity and motor decline in older adults. *Archives of Internal Medicine*, 169(12):1139–1146, 2009. [1](#)
- [8] I. Bytyci and M. Y. Henein. Stride length predicts adverse clinical events in older adults: a systematic review and meta-analysis. *Journal of Clinical Medicine*, 10(12):2670, 2021. [2](#)
- [9] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh. OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(1):172–186, 2019. [1](#), [2](#)
- [10] J. Chang, J. Nathalie, M. Nguyenhuy, R. Xu, S. A. Virk, and A. Saxena. Slow gait speed is associated with worse post-operative outcomes in cardiac surgery: A systematic review and meta-analysis. *Journal of Cardiac Surgery*, 37(1):197–204, 2022. [1](#)
- [11] T. Cuignet, C. Perchoux, G. Caruso, et al. Mobility among older adults: Deconstructing the effects of motility and movement on wellbeing. *Urban Studies*, 57(2):383–401, 2020. [1](#)
- [12] M. Á. De la Cámara, S. Higuera-Fresnillo, K. P. Sadarangani, I. Esteban-Cornejo, D. Martinez-Gomez, and Ó. L. Veiga. Clinical and ambulatory gait speed in older adults: associations with several physical, mental, and cognitive health outcomes. *Physical Therapy*, 100(4):718–727, 2020. [1](#)
- [13] K. Delbaere, J. C. Close, A. S. Mikolaizak, P. S. Sachdev, H. Brodaty, and S. R. Lord. The falls efficacy scale international (FES-I): A comprehensive longitudinal validation study. *Age and Ageing*, 39(2):210–216, 2010. [3](#)
- [14] G. Delfi, M. Kamachi, and T. Dutta. Development of an automated minimum foot clearance measurement system: Proof of principle. *Sensors*, 21(3):976, 2021. [2](#)
- [15] E. Elyan, P. Vuttipittayamongkol, P. Johnston, et al. Computer vision and machine learning for medical image analysis: recent advances, challenges, and way forward. *Artificial Intelligence Surgery*, 2(1):24–45, 2022. [2](#)
- [16] H.-S. Fang, J. Li, H. Tang, et al. AlphaPose: Whole-body regional multi-person pose estimation and tracking in real-time. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(6):7157–7173, 2022. [1](#), [2](#)
- [17] A. Fogel. Movement and communication in human infancy: The social dynamics of development. *Human Movement Science*, 11(4):387–423, 1992. [1](#)
- [18] M. U. Friedrich, S. Relton, D. Wong, and J. Alty. Computer vision in clinical neurology: a review. *JAMA Neurology*, 2025. [2](#)
- [19] D. J. Goble. The BTrackS Balance Test is a valid predictor of older adult falling, 2018. [3](#)
- [20] Shubham Goel, Georgios Pavlakos, Jathushan Rajasegaran, Angjoo Kanazawa, and Jitendra Malik. Humans in 4d: Reconstructing and tracking humans with transformers. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 14737–14748, 2023. [2](#), [1](#)
- [21] L. Hou, X. Liu, Y. Zhang, et al. Cohort profile: West china health and aging trend (WCHAT). *The Journal of Nutrition, Health and Aging*, 25(3):302–310, 2021. [2](#)
- [22] K. Jakubowski, T. Eerola, P. Alborno, G. Volpe, A. Camurri, and M. Clayton. Extracting coarse body movements from video in music performance: A comparison of automated computer vision techniques with motion capture data. *Frontiers in Digital Humanities*, 4:9, 2017. [2](#)
- [23] M. Javaid, A. Haleem, R. P. Singh, and M. Ahmed. Computer vision to enhance healthcare domain: An overview of features, implementation, and opportunities. *Intelligent Pharmacy*, 2(6):792–803, 2024. [2](#)
- [24] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. Ultralytics YOLOv8, 2023. [1](#)
- [25] G. I. Kempen, L. Yardley, J. C. Van Haastregt, et al. The short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. *Age and Ageing*, 37(1):45–50, 2008. [3](#)
- [26] A. Khang, V. Abdullayev, E. Litvinova, S. Chumachenko, A. V. Alyar, and P. Anh. Application of computer vision (CV) in the healthcare ecosystem. In *Computer Vision and AI-Integrated IoT Technologies in the Medical Ecosystem*, pages 1–16. CRC Press, 2024. [2](#)
- [27] U. Kim, J. Lim, Y. Park, and Y. Bae. Predicting fall risk through step width variability at increased gait speed in community dwelling older adults. *Scientific Reports*, 15(1): 16915, 2025. [2](#), [4](#)
- [28] Muhammed Kocabas, Nikos Athanasiou, and Michael J. Black. Vibe: Video inference for human body pose and shape estimation. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5252–5262, 2019. [2](#), [1](#)

- [29] B. Langhammer, B. Lindmark, and J. Stanghelle. The relation between gait velocity and static and dynamic balance in the early rehabilitation of patients with acute stroke. *Advances in Physiotherapy*, 8(2):60–65, 2006. 4
- [30] L. Lee, T. Patel, A. Costa, et al. Screening for frailty in primary care: accuracy of gait speed and hand-grip strength. *Canadian Family Physician*, 63(1):e51–e57, 2017. 1
- [31] S. S. Levy, K. J. Thralls, and S. A. Kviatkovsky. Validity and reliability of a portable balance tracking system, BTrackS, in older adults. *Journal of Geriatric Physical Therapy*, 41(2):102–107, 2018. 3
- [32] R. Li, R. J. St George, X. Wang, et al. Moving towards intelligent telemedicine: Computer vision measurement of human movement. *Computers in Biology and Medicine*, 147:105776, 2022. 2
- [33] Zhihao Li, Jianzhuang Liu, Zhensong Zhang, Songcen Xu, and Youliang Yan. Cliff: Carrying location information in full frames into human pose and shape estimation. In *European Conference on Computer Vision*, 2022. 1
- [34] N. Löfgren, L. Berglund, V. Giedraitis, E. Rosendahl, and A. C. Åberg. Can turn duration and step parameters during the timed up and go test with and without a dual-task discriminate between individuals with different cognitive abilities? An explorative study. *Assessment*, page 10731911251410337, 2026. 2
- [35] C.-T. Lu, Y.-C. Liu, and Y.-C. Pan. An intelligent playback control system adapted by body movements and facial expressions recognized by OpenPose and CNN. *Multimedia Tools and Applications*, 83(10):31139–31160, 2024. 2
- [36] U. M. U. Luis, M. S. Rodrigo, S. S. Cristhian, and S. M. C. Mauricio. Beyond timing: A critical review of the iTUG test and its implementation challenges for fall risk assessment in community-dwelling older adults. *Health Policy and Technology*, page 101166, 2026. 2, 4
- [37] Z. Luo, J.-T. Hsieh, N. Balachandar, et al. Computer vision-based descriptive analytics of seniors’ daily activities for long-term health monitoring. *Machine Learning for Healthcare (MLHC)*, 2(1), 2018. 2
- [38] M. Marano, G. Sergi, A. Magliozzi, et al. Fear of falling impairs spatiotemporal gait parameters, mobility, and quality of life in Parkinson’s disease: a cross-sectional study. *Neurological Sciences*, 46(6):2655–2663, 2025. 4
- [39] D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press, 2010. 2
- [40] Y. Mawarikado, Y. Uchihashi, Y. Inagaki, et al. Relationship between history of falls and foot pressure centre parameters during gait and stance in patients with lower-limb osteoarthritis. *Scientific Reports*, 15(1):26723, 2025. 2
- [41] S. Mroz, N. Baddour, C. McGuirk, et al. Comparing the quality of human pose estimation with BlazePose or OpenPose. In *2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART)*, 2021. 2
- [42] K.-D. Ng, S. Mehdizadeh, A. Iaboni, A. Mansfield, A. Flint, and B. Taati. Measuring gait variables using computer vision to assess mobility and fall risk in older adults with dementia. *IEEE Journal of Translational Engineering in Health and Medicine*, 8:1–9, 2020. 2
- [43] R. W. Nithman and J. L. Vincenzo. How steady is the STEADI? inferential analysis of the CDC fall risk toolkit. *Archives of Gerontology and Geriatrics*, 83:185–194, 2019. 3
- [44] S. M. O’Connor, H. S. Baweja, and D. J. Goble. Validating the BTrackS balance plate as a low cost alternative for the measurement of sway-induced center of pressure. *Journal of Biomechanics*, 49(16):4142–4145, 2016. 3
- [45] P. Ortega-Bastidas, B. Gomez, P. Aqueveque, S. Luarte-Martinez, and R. Cano-de-la Cuerda. Instrumented timed up and go test (iTUG)—more than assessing time to predict falls: a systematic review. *Sensors*, 23(7):3426, 2023. 2, 4
- [46] D. Parker, J. Andrews, and C. Price. Validity and reliability of the XSENSOR in-shoe pressure measurement system. *PLoS One*, 18(1):e0277971, 2023. 2
- [47] Georgios Pavlakos, Vasileios Choutas, Nima Ghorbani, Timo Bolkart, Ahmed A. A. Osman, Dimitrios Tzionas, and Michael J. Black. Expressive body capture: 3d hands, face, and body from a single image. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10967–10977, 2019. 2
- [48] R. Ranjan, D. Ahmedt-Aristizabal, M. A. Armin, and J. Kim. Computer vision for clinical gait analysis: A gait abnormality video dataset. *IEEE Access*, 13:45321–45339, 2025. 2
- [49] F. Ren, C. Ren, and T. Lyu. IoT-based 3D pose estimation and motion optimization for athletes: Application of C3D and OpenPose. *Alexandria Engineering Journal*, 115:210–221, 2025. 2
- [50] E. Salcedo. Computer vision-based gait recognition on the edge: A survey on feature representations, models, and architectures. *Journal of Imaging*, 10(12):326, 2024. 2
- [51] L. Seematter-Bagnoud and C. Büla. Brief assessments and screening for geriatric conditions in older primary care patients: a pragmatic approach. *Public Health Reviews*, 39(1): 8, 2018. 1
- [52] C. Shen, S. Yu, J. Wang, G. Q. Huang, and L. Wang. A comprehensive survey on deep gait recognition: Algorithms, datasets, and challenges. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2024. 2
- [53] Zehong Shen, Huaijin Pi, Yan Xia, Zhi Cen, Sida Peng, Zechen Hu, Hujun Bao, Ruizhen Hu, and Xiaowei Zhou. World-grounded human motion recovery via gravity-view coordinates. *SIGGRAPH Asia 2024 Conference Papers*, 2024. 1, 2
- [54] Soyong Shin, Juyong Kim, Eni Halilaj, and Michael J. Black. Wham: Reconstructing world-grounded humans with accurate 3d motion. *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2070–2080, 2023. 2, 1
- [55] S. Studenski. Gait speed reveals clues to lifelong health. *JAMA Network Open*, 2(10):e1913112, 2019. 1
- [56] Jianlin Su, Yu Lu, Shengfeng Pan, Bo Wen, and Yunfeng Liu. Roformer: Enhanced transformer with rotary position embedding. *ArXiv*, abs/2104.09864, 2021. 1
- [57] Zachary Teed, Lahav Lipson, and Jia Deng. Deep patch visual odometry. *ArXiv*, abs/2208.04726, 2022. 1

- [58] A. Toshev and C. Szegedy. DeepPose: Human pose estimation via deep neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014. [1](#), [2](#)
- [59] Aner Weiss, Anat Mirelman, Nir Giladi, Lisa L Barnes, David A Bennett, Aron S Buchman, and Jeffrey M Hausdorff. Transition between the timed up and go turn to sit subtasks: is timing everything? *Journal of the American Medical Directors Association*, 17(9):864–e9, 2016. [4](#)
- [60] C. W. Won, S. Lee, J. Kim, et al. Korean frailty and aging cohort study (KFACS): cohort profile. *BMJ Open*, 10(4): e035573, 2020. [2](#)
- [61] Yufei Xu, Jing Zhang, Qiming Zhang, and Dacheng Tao. ViTPose: Simple vision transformer baselines for human pose estimation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022. [2](#), [1](#)
- [62] S. Yan, Z. Yu, C. Primiero, et al. A multimodal vision foundation model for clinical dermatology. *Nature Medicine*, pages 1–12, 2025. [2](#)
- [63] L. Yardley, N. Beyer, K. Hauer, G. Kempen, C. Piot-Ziegler, and C. Todd. Development and initial validation of the falls efficacy scale-international (FES-I). *Age and Ageing*, 34(6): 614–619, 2005. [3](#)
- [64] Vickie Ye, Georgios Pavlakos, Jitendra Malik, and Angjoo Kanazawa. Decoupling human and camera motion from videos in the wild. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. [2](#), [1](#)
- [65] C. Zheng, W. Wu, C. Chen, et al. Deep learning-based human pose estimation: A survey. *ACM Computing Surveys*, 56(1):1–37, 2023. [2](#)
- [66] D. Zhu, J. Zhao, T. Wu, B. Zhu, M. Wang, and T. Han. Effects of a computer vision-based exercise application for people with knee osteoarthritis: Randomized controlled trial. *JMIR mHealth and uHealth*, 13(1):e63022, 2025. [2](#)

# Fall Risk and Gait Analysis in Community-Dwelling Older Adults using World-Spaced 3D Human Mesh Recovery

## Appendix

### 7. Population Characteristics

The sample consisted of community-dwelling older adults. Participants were recruited within Orlando, Florida using various strategies, including flyers, word-of-mouth, and collaboration with community partners. The inclusion criteria were that participants must be aged  $\geq 60$  years, be able to walk (with or without assistive devices but not requiring assistance from another person), live in their own homes or apartments, and be fluent in English or Spanish. The exclusion criteria were (1) having a medical condition that may preclude engagement in PA (including shortness of breath, dizziness, tightness or chest pain, and unusual fatigue at rest or with light exertion) and (2) currently receiving treatment from a rehabilitation facility. Participants provided written consent upon enrollment. All study procedures were approved by the University of Central Florida Institutional Review Board and all participants provided written informed consent prior to participation.

Variable	Mean $\pm$ Standard Deviation
Age	74.2 $\pm$ 7.81 years
STEADI Score	2.95 $\pm$ 2.83
Short FES-I Score	9.40 $\pm$ 3.02
BTrackS Balance Score	29.3 $\pm$ 11.2 cm

Table 4. Population characteristics.

### 8. Specifics on GVHMR

*Comparisons to previous methods* Skeleton-based pose estimation methods such as OpenPose [9] and AlphaPose [16] extract two dimensional pixel-level joint positions from RGB images, which are fundamentally ill-suited for gait analysis; depth is unrecoverable from a single 2D projection, and measures such as step length and step width are conflated with camera perspective, rendering them neither metrically accurate nor view-invariant. While 3D HMR methods can reconstruct full-body geometry in three dimensions, most 3D HMR methods, such as HMR2.0 [20] and VIBE [28] still estimate body pose only in the camera frame. Optimization-based methods, such as SLAHMR [64], recover global trajectories but are computationally prohibitive and prone to convergence failures on long inputs, while regression-based alternatives, such as WHAM [54], require good initialization including a standardized environment with calibration to maintain consistent, accurate predictions. GVHMR [53] addresses these limitations by

Predictor	Estimate	CI	p
(Intercept)	296.47	[160.02, 432.92]	< <b>0.001</b>
ste8	-0.14	[-5.34, 5.05]	0.956
age	-2.79	[-4.73, 0.86]	<b>0.005</b>

Table 5. Step Width Variability



Figure 5. XSENSOR Insole and IMU System

predicting pose in a per-frame *Gravity-View coordinate system*, directly outputting gravity-aligned world-space trajectories without additional refinement networks, making it the most suitable method for processing extended, in-the-wild clinical video such as the TUG.

*Calculation specifics* Each TUG video is processed through GVHMR [53], which preprocesses each frame using human bounding boxes [24, 33], keypoints [61], image features [20], and estimated camera rotation [57]. These per-frame features are fused into tokens and passed through the model, recovering human pose within a **Gravity-View (GV) coordinate system** defined by the world gravity and camera view directions, which is naturally gravity-aligned and uniquely defined for each individual frame, reducing the ambiguity of the image-to-pose mapping. The model is based on a transformer architecture with Rotary Positional Embeddings (RoPE) [56], which enables processing of sequences of arbitrary length without the need for a sliding window, making it suitable for application to TUG videos of varying lengths.

### 9. Gait Event Detection

#### 9.1. Transition Duration Computation

##### 9.1.1. Sit-to-Stand and Stand-to-Sit (STS)

To extract STS and turning durations from monocular video, we first computed a set of biomechanically meaningful signals from GVHMR joint trajectories. Hip and shoulder midpoints were used to derive vertical and anterior-posterior motion, and a trunk angle signal was computed from the vector connecting the hip midpoint to the

shoulder midpoint. These raw trajectories were differentiated to obtain velocities and then smoothed using a Gaussian kernel ( $\sigma = 3$ ) and a 4th-order low-pass Butterworth filter ( $f_s = 30\text{Hz}$ ,  $f_c = 2\text{ Hz}$ ) to reduce frame-to-frame jitter inherent in video-based pose estimation. These velocities were utilized to create a weighted composite signal (STS) defined in Eq. (1). STS-1 (sit-to-stand) was identified as the largest positive peak, and STS-2 (stand-to-sit) as the largest negative peak. Weights were assigned based on the presumed biomechanical contribution of each feature to sit-to-stand performance, with vertical hip displacement (1.0) reflecting primary task completion, anterior shoulder translation (0.7) indexing forward momentum generation, and trunk flexion angle (0.5) capturing compensatory strategy rather than core task success. The `findpeaks` function in R provided start frame, peak, and end frame for each point. Parameters for peak detection were tuned to each participant and trial, such that minimum peak height was defined as  $\overline{STS} \pm 0.8 * \sigma_{STS}$  and minimum peak distance as  $0.7 * (Frame_{max} - Frame_{min})$ .

### 9.1.2. Turns

Turning was detected using the filtered hip-line signal, as defined in Eq. (2). This hip line signal was differentiated to extract rotational motion of the pelvis. After smoothing with the Gaussian kernel and 4th-order low-pass Butterworth filter, as in the STS signal,  $Turn_1$  is defined as a peak and  $Turn_2$  is defined as a trough. While the minimum peak height was not changed, the number of peaks was defined as 1 for each trial to limit erroneous peak detection (see Fig. 3).

## 10. LME Models

Linear mixed effects (LME) models were utilized in this study to account for the repeated measures model of the data.  $Y_{ij}$  is the outcome for participant  $i$  on trial  $j$ ,  $\beta_0$  is the fixed intercept,  $\mathbf{x}_{ij}\beta$  represents the fixed effects (fall risk factors and age) and  $b_{0i}$  is the random intercept of participant  $i$ , to capture individual-level differences. The error residuals are represented by  $\varepsilon_{ij}$ . LMEs were constructed in R using the `lmer` function from the `lme4` library[3]. The LME output has several measures in addition to the estimate ( $\beta$ ) and the  $p$ -value. The full output is presented in the tables below. The estimate is shown alongside the 95% confidence interval, and the  $p$ -value. Below these are the random effects evaluating between- and within-participant variance, model structure, input, and fit.

$$Y_{ij} = \beta_0 + \mathbf{x}_{ij}\beta + \mathbf{Age}_{ij}\beta + b_{0i} + \varepsilon_{ij} \quad (3)$$

$\sigma^2$  refers to the within-participant variability; for instance, the amount of variability of step length trial-to-trial when accounting for the fixed effects (Age and a Fall Risk Factor) in the model.  $\tau_{00}$  refers to the between-participant

variability; for instance, how variable step lengths are between different participant after accounting for the fixed effects. ICC, or the intra-class correlation coefficient, quantifies the proportion of total variance attributable to differences between participants.  $N$  refers to the total number of observations in the clustered model, this may vary depending on missing data for each fall risk factor. In models involving the STEADI, missing data let to a decreased number of participants. Observations represents total number of trials utilized in the model. The  $R^2$  values provide measure of model fit. The Marginal  $R^2$  represents the variance explained only by the fixed effects—in this case, STEADI and age. Marginal  $R^2$  explains how much of the outcome (e.g. step length) can be predicted without considering participant-level differences. Conditional  $R^2$  is the variance explained by the full model, including both fixed effects and random effects. Higher values indicates greater model fit, and greater gap between the two  $R^2$  values reveals if outcome variability is due to stable, between-person differences, not trial-to-trial fluctuations.

## 11. Limitations

This pilot study is limited by a relatively small sample of 207 videos from 52 older adults from an ongoing study, which may constrain generalizability. Nevertheless, the dense within-participant data, spanning multiple trials and complementary measures including insole-derived step times, self-rated fall risk, fear of falling, and postural sway, strengthen the interpretability of the video-based metrics. Additionally, 18 participants completed assessments at two timepoints at least two months apart, which were treated as independent observations given the temporal separation.

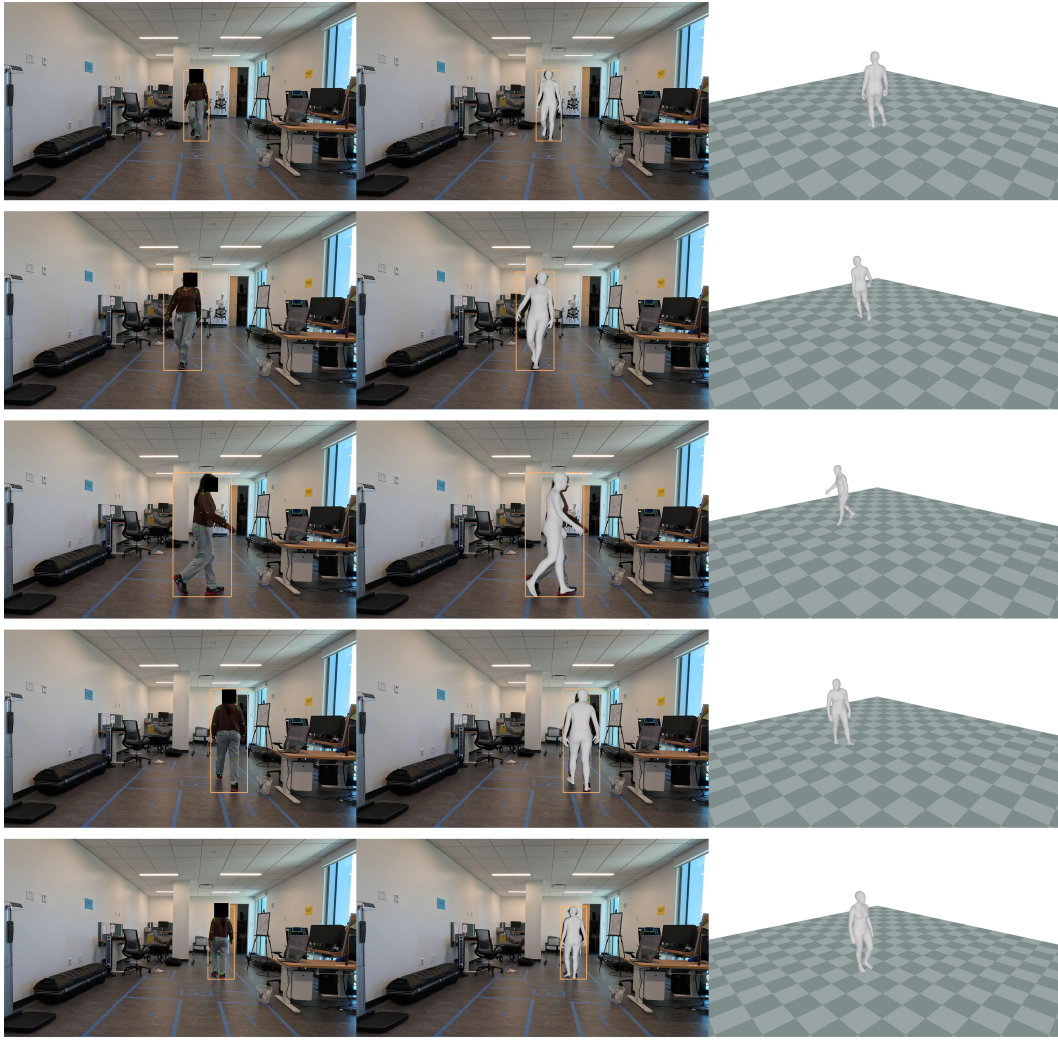


Figure 6. Snapshots of research assistant completing the TUG (left), with the camera-centric view overlay of the HMR (middle), and the global-centric view (right).

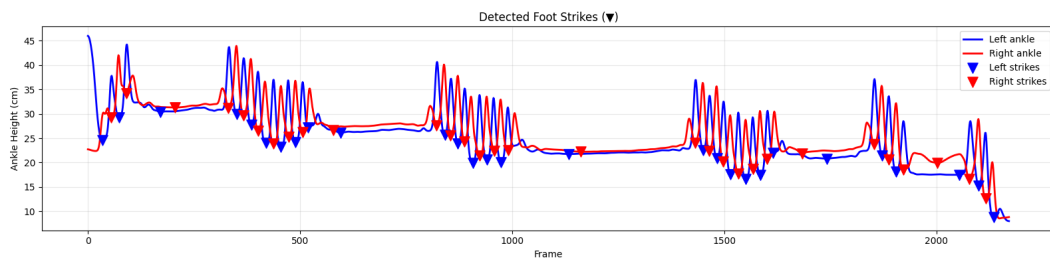


Figure 7. Detection of heel strike using peak detection, according to the detected joints using GVHMR.

Predictors	Model 1			Model 2			Model 3		
	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	49.74	[32.01, 67.46]	< 0.001	65.08	[48.45, 81.72]	< 0.001	65.69	[47.95, 83.44]	< 0.001
STEADI	-1.36	[-2.03, -0.68]	< 0.001	—	—	—	—	—	—
Age	-0.00	[-0.25, 0.25]	0.988	-0.14	[-0.37, 0.10]	0.264	-0.23	[-0.48, -0.01]	0.065
Short FES-I	—	—	—	-1.04	[-1.65, -0.43]	0.001	—	—	—
BTrackS	—	—	—	—	—	—	-0.11	[-0.27, -0.05]	0.176
Random Effects									
$\sigma^2$		10.7			10.66			10.66	
$\tau_{00}$		45.1			49.72			57.08	
ICC		0.81			0.82			0.84	
N		66			70			70	
Observations		195			207			207	
Marginal R <sup>2</sup>		0.210			0.180			0.084	
Conditional R <sup>2</sup>		0.848			0.855			0.856	

Table 6. Step Length (SL) Analysis

Predictors	Model 1			Model 2			Model 3		
	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	296.47	[160.02, 432.92]	< 0.001	301.74	[182.35, 421.13]	< 0.001	302.56	[182.97, 422.14]	< 0.001
STEADI	-0.14	[-5.34, 5.05]	0.956	—	—	—	—	—	—
Age	-2.79	[-4.73, -0.86]	0.005	-2.72	[-4.43, -1.02]	0.002	-2.95	[-4.61, -1.29]	0.001
Short FES-I	—	—	—	-1.24	[-5.61, 3.13]	0.577	—	—	—
BTrackS	—	—	—	—	—	—	0.14	[-0.93, 1.22]	0.794
Random Effects									
$\sigma^2$		960.87			949.11			949.11	
$\tau_{00}$		2563.93			2426.50			2436.54	
ICC		0.73			0.72			0.72	
N		66			70			70	
Observations		195			207			207	
Marginal R <sup>2</sup>		0.117			0.133			0.130	
Conditional R <sup>2</sup>		0.759			0.756			0.756	

Table 7. SW Variability Analysis

Predictors	Model 1			Model 2			Model 3		
	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	393.99	[110.09, 677.89]	0.007	594.80	[330.99, 858.61]	< 0.001	602.10	[330.27, 873.93]	< 0.001
STEADI	-19.62	[-30.44, -8.80]	< 0.001	—	—	—	—	—	—
Age	-0.59	[-4.61, 3.43]	0.772	-2.50	[-6.27, 1.27]	0.193	-3.44	[-7.21, 0.34]	0.074
Short FES-I	—	—	—	-12.76	[-22.43, -3.09]	0.010	—	—	—
BTrackS Balance Score	—	—	—	—	—	—	-1.97	[-4.42, 0.47]	0.113
Random Effects									
$\sigma^2$		11382.63			11061.41			11055.43	
$\tau_{00}$		8672.18			9695.94			10532.24	
ICC		0.43			0.47			0.49	
N		66			70			70	
Observations		195			207			207	
Marginal R <sup>2</sup>		0.143			0.103			0.069	
Conditional R <sup>2</sup>		0.513			0.522			0.523	

Table 8. SL Variability Analysis

Predictors	Model 1			Model 2			Model 3		
	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	26.82	[5.86, 47.79]	<b>0.012</b>	13.39	[-7.16, 33.95]	0.200	13.69	[-7.46, 34.83]	0.203
STEADI	1.23	[0.45, 2.01]	<b>0.002</b>	—	—	—	—	—	—
Age	0.21	[-0.08, 0.51]	0.157	0.36	[0.07, 0.65]	<b>0.016</b>	0.44	[0.15, 0.74]	<b>0.003</b>
Short FES-I	—	—	—	0.71	[-0.01, 1.42]	0.052	—	—	—
BTrackS Balance Score	—	—	—	—	—	—	0.00	[-0.19, 0.19]	0.989
Random Effects									
$\sigma^2$		172.17			175.44			175.51	
$\tau_{00}$		15.18			21.18			25.47	
ICC		0.08			0.11			0.13	
N		71			73			73	
Observations		211			217			217	
Marginal R <sup>2</sup>		0.092			0.071			0.051	
Conditional R <sup>2</sup>		0.165			0.171			0.172	

Table 9. Sit-to-Stand and Stand-to-Sit (STS) Duration Analysis