Sex Differences in Lower-Limb Surface Electromyographic Characteristics During Prolonged Overground Walking

Murat Cenk CELEN[1]

[1]Department of Biophysics, Faculty of Medicine, Bursa Uludag University, Bursa, Republic of Türkiye

Received July 12, 2026
Accepted October 20, 2026

Summary

This study examined sex-related differences in lower-limb surface electromyographic (sEMG) characteristics during prolonged overground walking and their sensitivity to signal processing. Publicly available recordings from 31 healthy adults (19 females, 12 males), comprising 13,465 gait cycles, were analyzed. Root mean square amplitude (RMS), peak amplitude, integrated EMG (iEMG), and stride-to-stride RMS coefficient of variation (CV) were evaluated. Raw and band-pass-filtered features were compared with filtered features normalized to the peak of each participant-side-muscle ensemble-mean walking envelope. Two channels with signal-quality concerns were excluded from the QC-screened dataset and retained in an inclusive sensitivity analysis. Mixed-effects models adjusted for height and body weight included 28 participants, contributing 278 and 280 participant-side-muscle summaries, respectively. In the QC-screened dataset, males had higher tibialis anterior (TA) and gastrocnemius lateralis (GL) peak amplitudes and higher GL RMS across all processing variants after Holm correction. The TA RMS contrast was significant only in the raw analysis; a hamstring-region peak contrast was significant only after task-reference normalization. No muscle-specific iEMG comparison survived Holm correction. Including the two excluded channels produced limited changes in TA and GL amplitude estimates, whereas CV findings were sensitive to processing and channel exclusions. Sex-related sEMG associations depended on muscle, outcome, and processing approach. Selected amplitude contrasts persisted on a task-relative scale, but these findings do not establish differences in neural drive, maximal activation capacity, or walking function.

Key words

Surface electromyography • Gait analysis • Sex differences • Human locomotion • Lower limb

Corresponding author

M. C. Celen, Department of Biophysics, Faculty of Medicine, Bursa Uludag University, Bursa, Republic of Türkiye. E-mail: muratcenkcelen@uludag.edu.tr

Introduction

Effective human movement depends on the precise spatio-temporal coordination of multiple lower-limb muscles [1,2]. The continuous interaction between the nervous and musculoskeletal systems ensures stable, adaptive, and mechanically efficient gait under varying environmental and functional demands [1,2]. Surface electromyography (sEMG) provides a non-invasive method for characterizing these neuromuscular activation patterns and has therefore become an established technique for investigating locomotor control in both healthy individuals and clinical populations [3,4].

Previous studies have reported sex-related differences in lower-limb muscle activation during walking. These findings were not consistent for muscles, gait phases, experimental protocols, or signal processing approaches [5,6]. Interpreting between-participant differences in sEMG amplitude requires consideration of the normalization reference and of anatomical and recording-related factors that influence the measured signal [4,7-9]. In particular, non-normalized amplitudes cannot isolate differences in muscle activation from the effects of subcutaneous tissue thickness, electrode

 An open access article under the CC BY 4.0 license https://creativecommons.org/licenses/by/4.0/ © 2027 by the authors. Published by the Institute of Physiology, Czech Academy of Sciences, Prague, Czech Republic. E-mail: physres@fgu.cas.cz, www.biomed.cas.cz/physiolres

placement, and other determinants of signal magnitude [4,7-9]. Consequently, the extent to which sex influences lower-limb surface EMG characteristics during prolonged walking remains incompletely understood.

Previous studies have investigated sex-related differences in lower-limb sEMG during walking by examining muscle activation patterns, their occurrence across gait cycles, and ankle-muscle coactivation [5,6]. These studies included numerous strides per participant. The present study extends this line of investigation by jointly examining whole-cycle amplitude characteristics and stride-to-stride RMS variability across five bilateral muscle regions using participant-level mixed-effects models. Its added value lies in evaluating the muscle dependence and analytical robustness of these associations within a common dataset. The use of a publicly available dataset also enables independent reanalysis of the underlying recordings [10-12].

The present study aimed to examine whether recorded lower-limb sEMG amplitude characteristics and stride-to-stride RMS variability were associated with sex during prolonged overground walking, and whether these associations differed across muscles. Additional analyses undertaken during revision evaluated their sensitivity to digital filtering, task-reference amplitude normalization, and the exclusion of channels with identified signal-quality concerns. Between-sex comparisons accounted for repeated observations within participants and were adjusted for height and body weight. The analyses characterized recorded sEMG signals and were not designed to establish differences in neural drive, muscle force, or activation relative to maximal capacity.

Methods

Study design and data source

This study was designed as a retrospective observational analysis using a publicly available human movement database. The dataset was obtained from the PhysioNet repository (Surface Electromyographic Signals Collected During Long-Lasting Ground Walking of Young Able-Bodied Subjects, Version 1.0.1), with citation of the repository resource, the original related publication, and the standard PhysioNet reference as recommended by PhysioNet [10,11,13]. The database and its acquisition protocol were further described in the accompanying dataset publication [12]. The database includes synchronized lower-limb surface electromyographic (sEMG), bilateral footswitch and knee electrogoniometer recordings collected from healthy young adults during prolonged overground walking performed at a self-selected walking speed [11,12].

Since only anonymized, publicly available data was analyzed, no additional ethical approval or participant recruitment was required for this study. The overall analysis workflow is summarized in Figure 1.

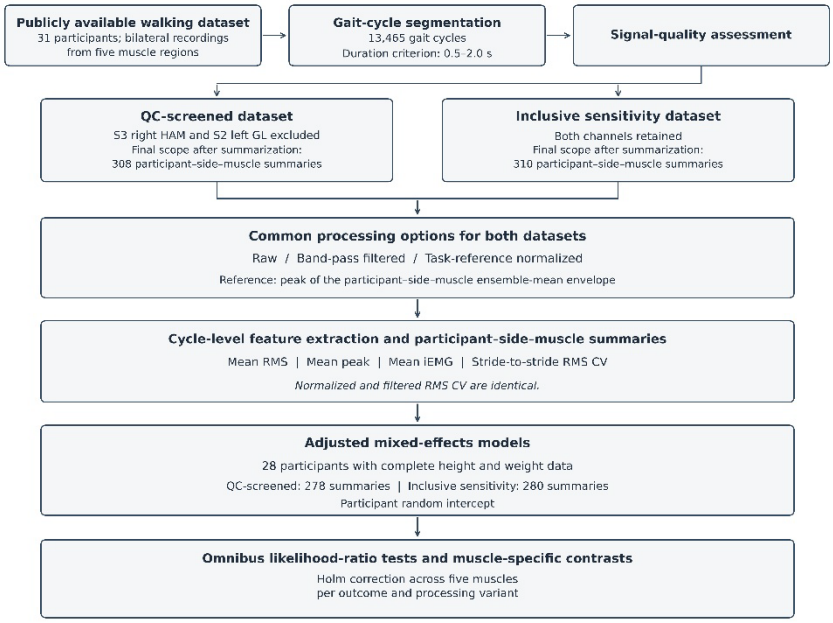


Fig. 1. Workflow of the analytical pipeline used in the present study. Following gait-cycle segmentation and signal-quality assessment, analyses were conducted using a QC-screened dataset excluding the S3 right HAM and S2 left GL channels and an inclusive sensitivity dataset retaining both channels. Raw, band-pass-filtered, and task-reference-normalized amplitude features were summarized at the participant-side-muscle level. Adjusted mixed-effects models included 28 participants with complete height and body-weight data. Omnibus effects were assessed using likelihood-ratio tests; Holm correction was applied to the five muscle-specific sex contrasts within each outcome, processing variant, and dataset. QC, quality control; HAM, hamstring-region recording; GL, gastrocnemius lateralis; RMS, root mean square; iEMG, integrated electromyography; CV, coefficient of variation.

Participants

The original data set consisted of 31 healthy adults (19 women and 12 men). Demographic information, including sex, age, height, and body weight, was obtained from the metadata provided with the database [11,12]. Height and/or body weight information was not available for three participants (S19, S25, and S31). Therefore, all adjusted models, including those used in the channel-inclusion sensitivity analyses, were fitted using complete-case observations from the 28 participants with available height and body-weight data.

Gait cycle segmentation

All physiological recordings were imported directly from the original waveform database (WFDB) records using the WFDB software tools for Python [14]. Individual gait cycles were determined from synchronized bilateral footswitch recordings using a dedicated analysis pipeline. Candidate gait events were identified from footswitch signals and subsequently analyzed using predefined physiological validity criteria. Gait cycles lasting only between 0.5 and 2.0 seconds were retained for further analysis. WFDB digital-to-physical scaling was used throughout feature extraction [14]. EMG amplitudes were interpreted as microvolts according to the repository documentation [11]. For S01, the version 1.0.1 header labels EMG channels as millivolts, conflicting with this documentation. The microvolt interpretation was retained because the data-file contents and muscle-specific gain and baseline values were unchanged from version 1.0.0, which specified microvolts. No additional factor-of-1000 conversion was applied; this metadata discrepancy remains unresolved by the data provider.

The application of these criteria resulted in the identification of 13,465 valid gait cycles among all participants. These gait cycles formed the basis for all subsequent EMG analyses.

EMG feature extraction

According to the original acquisition protocol, sEMG signals were sampled at 2 kHz, with a 10 Hz high-pass filter in the recording probes and a 450 Hz low-pass filter in the acquisition system [12]. The original investigators also reported checking signal quality and sensor positioning during acquisition [12]. These acquisition-stage procedures preceded the present secondary analysis and were distinct from the processing and exclusion criteria applied here. For each valid gait cycle, the corresponding raw EMG segments for the ipsilateral lower-limb muscles, including tibialis anterior (TA), gastrocnemius lateralis (GL), rectus femoris (RF), hamstrings, and vastus lateralis (VL), were extracted separately [11,12]. The designation HAM was retained from the original dataset. The dataset publication describes this channel as a hamstring-group recording rather than an isolated recording from biceps femoris, semitendinosus, or semimembranosus [12]. HAM recordings were obtained using fixed-geometry differential surface probes with 4 mm electrode diameter and 8 mm inter-electrode distance. The original investigators reported electrode positioning and orientation according to Surface EMG for Non-Invasive Assessment of Muscles (SENIAM) recommendations and illustrated the placement in Figure 3 of their publication [12]. Accordingly, the present HAM findings should be interpreted at the recording-region level and cannot be attributed to an individual hamstring muscle.

Three processing variants were compared using the same retained gait cycles and participant-side-muscle combinations. Variant A used the WFDB physical signal samples without additional digital filtering, mean subtraction, detrending, or amplitude normalization. Variant B used the signal after continuous-record filtering with a Butterworth band-pass filter (20-450 Hz; design order N=4), implemented in second-order sections with forward-backward filtering. For both variants, cycle-specific RMS was calculated as the square root of the mean squared signal, peak amplitude as the maximum absolute sample value, and iEMG as the sum of absolute sample values divided by the sampling frequency. These features were calculated at the original sampling frequency, before any temporal interpolation.

Variant C comprised the features from Variant B divided by a common task-specific amplitude reference for each participant-side-muscle combination. To obtain this reference, the band-pass-filtered continuous signal was full-wave rectified and low-pass filtered at 6 Hz using a Butterworth filter (N=4) with forward-backward filtering. Negative envelope values were set to zero. Each retained gait-cycle envelope was then linearly interpolated to 101 points spanning the gait cycle. The reference was defined as the maximum of the ensemble-mean envelope, calculated with equal weighting of all retained cycles. Thus, the reference was the peak of the average envelope, rather than the average of individual cycle peaks. The same positive reference was used to divide all cycle-specific RMS, peak, and iEMG values

within that participant-side-muscle combination. Normalized RMS and peak were expressed as dimensionless ratios, whereas normalized iEMG retained units of seconds. These values were not multiplied by 100 and did not represent percentages of MVC.

For each processing variant, cycle-specific amplitude features were averaged within each participant–side–muscle combination. Stride-to-stride RMS variability was calculated as the sample standard deviation of cycle-specific RMS values divided by their mean and expressed as a dimensionless coefficient of variation (CV). Because division by a common positive reference leaves this CV unchanged, Variants B and C yielded equivalent CV values; separate inferential tests were therefore performed only for the raw-signal and band-pass-filtered CV outcomes.

Because MVC recordings were unavailable, task-reference normalization was included to examine whether the observed between-sex associations persisted when amplitudes were expressed relative to each participant-side-muscle combination's walking reference [7]. Comparing raw, band-pass-filtered, and task-reference-normalized features allowed the effects of filtering and amplitude scaling to be examined separately. The normalized outcomes describe signal magnitude relative to the peak of the ensemble-mean walking envelope; they do not represent activation relative to maximal capacity. Task-reference normalization changes the interpretation of between-participant comparisons and does not establish that anatomical and recording-related influences have been eliminated [7,12]. Accordingly, all three processing variants were interpreted as recorded sEMG characteristics rather than direct measures of sex-related differences in neural drive or muscle activation.

The original raw-signal analysis did not apply artifact-based exclusions. During revision, additional signal-quality screening and targeted visual inspection identified repeated high-level plateaus and abrupt transitions in the right HAM channel of participant S3 and the left GL channel of participant S2. Before fitting the comparative models, these two participant-side-muscle channels were excluded from the QC-screened analysis set. Whole-channel exclusion was used to avoid selecting uncertain artifact boundaries and to prevent the affected channels from contributing to task-reference estimation. This decision did not imply that every segment of either channel was affected. All other channels from these participants were retained, and the analyses were repeated with both channels included as a sensitivity analysis. The QC-screened set comprised 66,882 stride-muscle observations and 308 participant-side-muscle summaries from 31 participants; the inclusive set comprised 67,325 observations and 310 summaries. The gait-cycle duration criterion of 0.5-2.0 s assessed temporal plausibility and was distinct from this signal-quality assessment.

Statistical analysis

Since several hundred gait cycles were available for each participant, individual strides were not treated as statistically independent observations. Instead, stride-level measurements were summarized at the participant-side-muscle level before inferential analyses were performed, thereby avoiding pseudoreplication.

The primary measure was the stridewise coefficient of variation of the RMS amplitude. Secondary outcome measures included mean RMS amplitude, integrated EMG, and peak EMG amplitude.

Associations between sex and EMG measurements were evaluated using linear mixed-effects models fitted by maximum likelihood [16]. Fixed effects comprised sex, muscle, body side, the sex-by-muscle interaction, height, and body weight. Participant identity was included as a random intercept to account for repeated participant-side-muscle observations. Female sex, tibialis anterior, and the left side were the reference categories. The same model structure was applied across processing variants and to both the QC-screened and inclusive sensitivity datasets. Complete-case analysis excluded participants S19, S25, and S31 because height and body weight were unavailable. The adjusted models therefore included 28 participants, contributing 278 participant-side-muscle summaries in the QC-screened dataset and 280 in the inclusive dataset.

The sex-by-muscle interaction was evaluated using a likelihood-ratio test comparing the full model with a reduced model retaining all main effects but omitting the four interaction terms (4 degrees of freedom). The overall sex-related component was evaluated by jointly removing the sex main effect and all four sex-by-muscle interaction terms (5 degrees of freedom). Each comparison used identical observations, the same participant-level random-intercept structure, and maximum-likelihood estimation. The individual sex coefficient in the full model represented the male-female difference in the reference muscle and was not interpreted as an overall sex effect. Sensitivity to signal processing was assessed by comparing raw, band-pass-filtered, and

task-reference-normalized amplitude outcomes using the same observations within each analysis dataset. Sensitivity to the channel exclusions was assessed by repeating these analyses with the S3 right HAM and S2 left GL channels included. Comparisons considered the direction and magnitude of the estimated differences and their confidence intervals, as well as the multiplicity-adjusted p-values. Because task-reference normalization changes the measurement scale, the numerical magnitudes of normalized and non-normalized contrasts were not compared directly.

Within each muscle, male-female contrasts were calculated from model-based estimated marginal means, with height and body weight set to their analysis-sample means and equal weighting of the left and right sides [17]. Contrasts were reported with unadjusted 95 % Wald confidence intervals. Within each processing variant and analysis dataset, p-values were adjusted across the five muscle-specific sex comparisons for each outcome using the Holm procedure [18]. Statistical significance for these comparisons was determined from the Holm-adjusted p-values; the confidence intervals were not adjusted for multiple comparisons.

Statistical analyses were performed using Python (Python Software Foundation, Wilmington, DE, USA) with the NumPy, Pandas, SciPy, Statsmodels, and WFDB libraries [14,19-22] . Statistical significance was accepted as $P<0.05$.

Results

Participant Characteristics

The study included 31 healthy adults (19 females and 12 males), with 13,465 gait cycles meeting the duration criterion. Before channel exclusions, these cycles provided 67,325 stride-muscle observations and 310 participant-side-muscle summaries. Excluding the S3 right HAM and S2 left GL channels removed 223 and 220 stride-muscle observations, respectively, leaving 66,882 observations and 308 summaries in the QC-screened dataset. All 31 participants remained represented, and the other channels from S2 and S3 were retained. Participants S19, S25, and S31 lacked height and body-weight data. Consequently, the adjusted models included 28 participants, contributing 278 summaries in the QC-screened dataset and 280 summaries in the inclusive sensitivity dataset. These summaries represented repeated observations within participants, rather than independent participants. Demographic characteristics are presented in Table 1.

Mixed-Effects Analyses

In the QC-screened dataset, likelihood-ratio tests supported an overall sex-related component and a sex-by-muscle interaction for RMS and peak amplitude across all three processing variants (Table 2). For task-reference-normalized RMS, the corresponding p-values were 0.0024 and 0.0067, respectively; for normalized peak amplitude, both were <0.001. For iEMG, both tests were significant in the raw and band-pass-filtered analyses, but neither was significant after task-reference normalization (sex-related component: p=0.143; interaction: p=0.102). For stride-to-stride RMS CV, neither the overall sex-related component nor the interaction was significant in the raw-signal analysis of the QC-screened dataset (p=0.111 and p=0.156, respectively). Both tests reached significance after band-pass filtering (p=0.029 and p=0.039). However, these findings were not reproduced when the two excluded channels were retained (p=0.292 and p=0.398), indicating sensitivity to the channel exclusions.

Post hoc comparisons

In the QC-screened dataset, muscle-specific male–female contrasts depended on the outcome and processing variant (Table 3). For RMS, males had higher values in GL across the raw, band-pass-filtered, and task-reference-normalized analyses (Holm-adjusted p=0.034, 0.039, and 0.011, respectively). The TA contrast remained positive across all three variants, but was significant only in the raw-signal analysis (Holm-adjusted p=0.034, 0.078, and 0.061, respectively). No other muscle-specific RMS comparison reached significance after Holm correction (Fig. 2A-C).

For peak amplitude, males had higher values in TA and GL across all three processing variants (all Holm-adjusted p<0.05). Task-reference-normalized peak amplitude also showed a positive male-female contrast in HAM (difference=1.769, unadjusted 95 % CI 0.876-2.662; Holm-adjusted p<0.001), whereas the HAM contrasts were not significant in the raw or band-pass-filtered analyses. RF and VL peak comparisons were not significant after Holm correction in any processing variant (Fig. 2D-F). No muscle-specific iEMG comparison reached significance after Holm correction in any processing variant (Fig. 2G-I). For normalized iEMG, the male-female differences were -0.023 s in TA (unadjusted 95 % CI -0.108 to 0.061) and 0.007 s in GL

(-0.078 to 0.092), with Holm-adjusted p=1.000 for both comparisons.

Sensitivity analyses

Repeating the analyses with the S3 right HAM and S2 left GL channels included produced limited changes in the estimated TA and GL amplitude contrasts. For example, inclusion changed the raw-signal RMS male-female contrasts by 0.241 μV in TA and 0.555 μV in GL, and the task-reference-normalized peak contrasts by 0.005 and 0.017, respectively. The directions of these contrasts were preserved. In contrast, the omnibus findings for stride-to-stride RMS CV depended on the channel exclusions, as described above. Across processing variants, GL RMS and TA/GL peak comparisons remained significant after Holm correction, whereas TA RMS did not reach significance after filtering or task-reference normalization. These findings indicate that sensitivity to processing and channel inclusion differed across outcomes. Omnibus test results for the inclusive sensitivity dataset are presented in Table 4.

Table 1. Participant characteristics.

Characteristic	**Total**	**Female**	**Male**	**Missings**
Age, (years)	24.2±1.9 (n=30)	23.9±1.4 (n=18)	24.7±2.4 (n=12)	S-31
Height, (cm)	170.7±10.3 (n=28)	163.4±4.8 (n=16)	180.4±7.0 (n=12)	S-19, 25, 31
Weight, (kg)	60.4±11.7 (n=28)	51.9±4.0 (n=16)	71.8±8.2 (n=12)	S-19,25, 31

Values are mean ± SD where applicable.

Table 2. Omnibus likelihood-ratio tests across processing variants in the QC-screened dataset.

Outcome	**Processing variant**	**Sex-related component p-value**	**Sex × muscle interaction p-value**
Mean RMS			
	A: Raw	**<0.001**	**0.0012**
	B: Band-pass filtered	**0.0030**	**0.0047**
	C: Task-reference normalized	**0.0024**	**0.0067**
Peak amplitude			
	A: Raw	**0.0052**	**0.0128**
	B: Band-pass filtered	**0.0062**	**0.0151**
	C: Task-reference normalized	**<0.001**	**<0.001**
iEMG			
	A: Raw	**<0.001**	**<0.001**
	B: Band-pass filtered	**0.0032**	**0.0018**
	C: Task-reference normalized	0.1431	0.1016
Stride-to-stride RMS CV			
	A: Raw	0.1109	0.1563
	B: Band-pass filtered	**0.0289**	**0.0392**

All models included 278 participant-side-muscle summaries from 28 participants and were fitted by maximum likelihood, with sex, muscle, side, sex × muscle, height, and body weight as fixed effects and participant as a random intercept. The sex-related component test jointly assessed the sex main effect and four sex × muscle interaction terms (5 degrees of freedom). The interaction test assessed the four interaction terms while retaining all main effects (4 degrees of freedom). Variant C comprised band-pass-filtered amplitude features divided by the peak of the participant–side–muscle ensemble-mean walking envelope. CV was identical for Variants B and C; therefore, Variant C CV was not tested separately. Bold values indicate p<0.05. Omnibus p-values are unadjusted; Holm correction was applied separately to the five muscle-specific contrasts for each outcome within each processing variant and dataset.

Table 3. Muscle-specific male-female contrasts.

Muscle	**Raw**	**Filtered**	**Normalized**
A. Mean RMS			
TA	12.457 (3.173, 21.742) *p=0.0342**	11.130 (1.791, 20.469) *p=0.0780*	0.092 (0.015, 0.170) *p=0.0607*
GL	12.872 (3.548, 22.196) *p=0.0341**	12.752 (3.373, 22.132) *p=0.0385**	0.121 (0.044, 0.199) *p=0.0112**
RF	1.478 (-7.806, 10.763) *p=1.0000*	0.951 (-8.388, 10.289) *p=1.0000*	0.011 (-0.067, 0.088) *p=0.7876*
HAM	1.535 (-7.759, 10.828) *p=1.0000*	2.454 (-6.894, 11.802) *p=1.0000*	0.096 (0.019, 0.173) *p=0.0607*
VL	9.397 (0.113, 18.682) *p=0.1418*	9.112 (-0.227, 18.450) *p=0.1675*	0.040 (-0.038, 0.117) *p=0.6305*
B. Peak EMG			
TA	122.542 (37.099, 207.984) *p=0.0247**	116.936 (33.295, 200.578) *p=0.0307**	1.410 (0.517, 2.302) *p=0.0059**
GL	112.224 (26.435, 198.012) *p=0.0414 **	110.513 (26.534, 194.492) *p=0.0396 **	1.524 (0.630, 2.419) *p=0.0033**
RF	25.969 (-59.474, 111.411) *p=0.6282*	24.218 (-59.423, 107.859) *p=0.5753*	0.615 (-0.277, 1.508) *p=0.1765*
HAM	43.924 (-41.598, 129.446) *p=0.6282*	45.417 (-38.302, 129.135) *p=0.5753*	1.769 (0.876, 2.662) *p<0.001**
VL	89.899 (4.456, 175.341) *p=0.1176*	87.998 (4.356, 171.639) *p=0.1176*	0.852 (-0.040, 1.745) *p=0.1223*
C. Integrated EMG			
TA	4.853 (0.067, 9.640) *p=0.1876*	3.498 (-1.226, 8.222) *p=0.5868*	-0.023 (-0.108, 0.061) *p=1.0000*
GL	5.234 (0.430, 10.039) *p=0.1637*	5.138 (0.394, 9.882) *p=0.1689*	0.007 (-0.078, 0.092) *p=1.0000*
RF	-0.963 (-5.749, 3.824) *p=1.0000*	-1.505 (-6.229, 3.220) *p=1.0000*	-0.058 (-0.142, 0.027) *p=0.8110*
HAM	-1.521 (-6.312, 3.270) *p=1.0000*	-0.679 (-5.408, 4.049) *p=1.0000*	-0.003 (-0.087, 0.082) *p=1.0000*
VL	2.236 (-2.550, 7.023) *p=1.0000*	2.055 (-2.669, 6.779) *p=1.0000*	-0.060 (-0.145, 0.024) *p=0.8110*

QC-screened dataset, 28 participants, 278 summaries. Cells: male-female difference; unadjusted 95 % CI; Holm-adjusted p. * p<0.05 after Holm correction across five muscles within each outcome and processing variant. Raw/Filtered units: RMS and peak, μV; iEMG, μV·s (S01 unit-metadata limitation applies). Normalized units: RMS and peak, dimensionless; iEMG, s. HAM: hamstring-region recording.

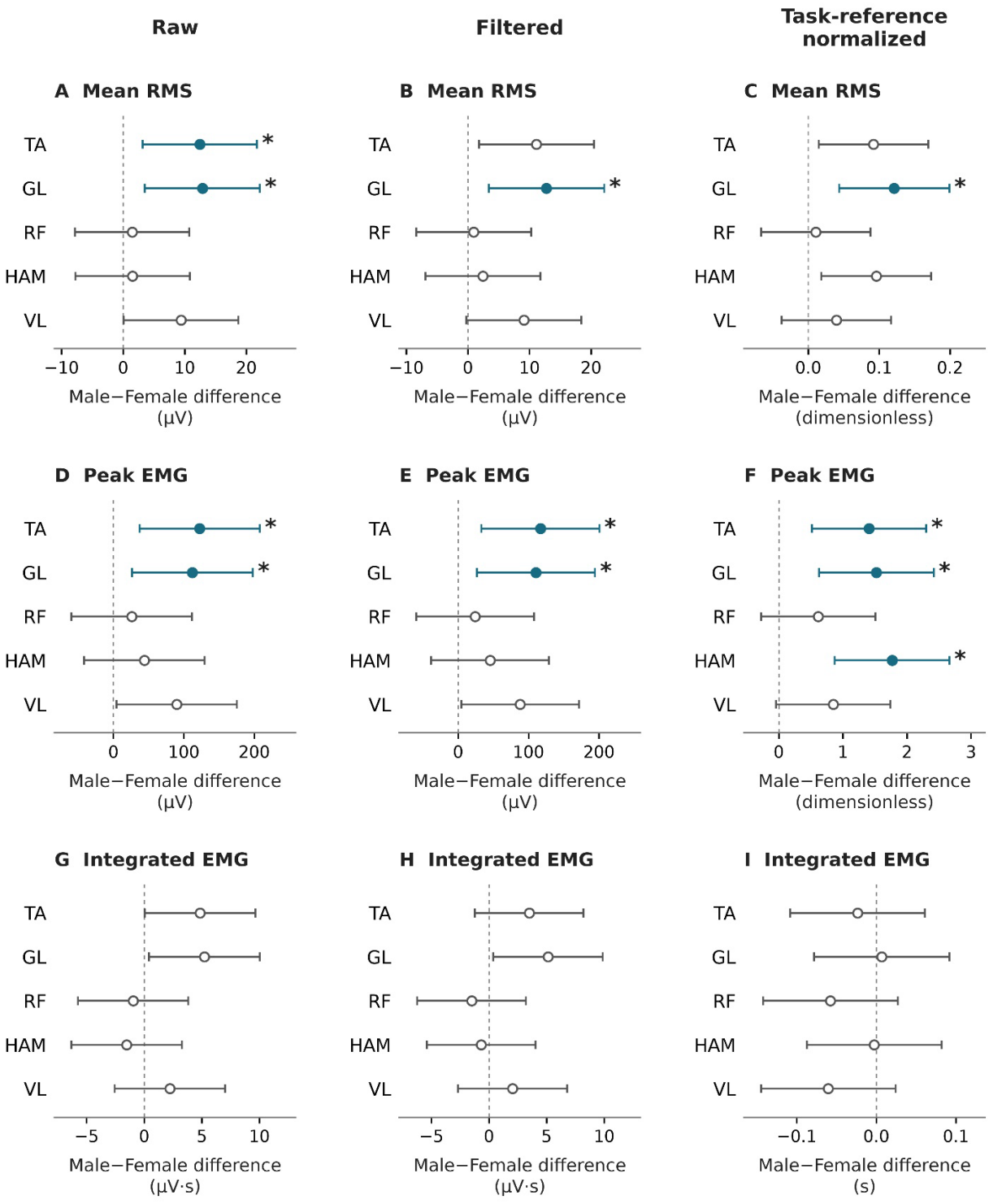


Fig. 2. Muscle-specific male–female contrasts across signal-processing variants in the QC-screened dataset. Panels show mean RMS (**A-C**), peak EMG (**D-F**), and integrated EMG (**G-I**), using raw signals, band-pass-filtered signals, and task-reference-normalized features, respectively. Points represent model-based contrasts; horizontal bars show unadjusted 95 % Wald confidence intervals. Positive values indicate higher values in males. Filled points and asterisks indicate Holm-adjusted p<0.05 across the five muscle comparisons within each outcome and processing variant. Dashed vertical lines indicate zero difference. Models included 278 participant-side-muscle summaries from 28 participants and were adjusted for height, body weight, and side, with participant-specific random intercepts. Normalization used the peak of each participant-side-muscle ensemble-mean walking envelope. Raw and filtered physical units are subject to the S01 metadata limitation described in Methods. TA, tibialis anterior; GL, gastrocnemius lateralis; RF, rectus femoris; HAM, hamstring-region recording; VL, vastus lateralis.

Table 4. Omnibus likelihood-ratio tests with the S3 right HAM and S2 left GL channels included.

	Processing variant	Sex-related component p-value	Sex × muscle interaction p-value
Mean RMS	Raw	**<0.001**	**<0.001**
Mean RMS	Filtered	**0.0021**	**0.0032**
Mean RMS	Normalized	**0.0021**	**0.0055**
Peak EMG	Raw	**0.0038**	**0.0096**
Peak EMG	Filtered	**0.0046**	**0.0114**
Peak EMG	Normalized	**<0.001**	**<0.001**
Integrated EMG	Raw	**<0.001**	**<0.001**
Integrated EMG	Filtered	**0.0016**	**<0.001**
Integrated EMG	Normalized	0.1434	0.0988
Stride-to-stride RMS CV	Raw	0.4532	0.5909
Stride-to-stride RMS CV	Filtered	0.2924	0.3978

All models included 280 participant-side-muscle summaries from 28 participants. Both channels excluded from the QC-screened dataset were retained. Model structure, covariate adjustment, processing variants, and likelihood-ratio comparisons were identical to those described for Table 2. The sex-related component and sex × muscle interaction tests had 5 and 4 degrees of freedom, respectively. Normalized and filtered RMS CV values were identical; therefore, normalized CV was not tested separately. Bold values indicate unadjusted omnibus $p<0.05$.

Discussion

The present study investigated sex-related differences in lower-limb surface EMG characteristics during prolonged overground walking using a publicly available dataset [11,12]. The findings depended on both the muscle examined and the signal-processing approach. In the QC-screened dataset, higher peak amplitudes in males for TA and GL, and higher RMS amplitudes for GL, were supported across raw, band-pass-filtered, and task-reference-normalized analyses. By contrast, the TA RMS comparison did not remain significant after filtering or normalization, and a HAM peak difference emerged only after task-reference normalization. No muscle-specific iEMG comparison was significant after Holm correction in any processing variant. Stride-to-stride RMS CV findings were sensitive to processing and channel exclusions. These results support muscle- and metric-specific associations with recorded sEMG characteristics, but do not establish sex-related differences in neural drive or muscle activation independent of anatomical and recording-related influences [7-9,15].

The persistence of the TA and GL peak contrasts and the GL RMS contrast across processing variants identifies these ankle-muscle findings as comparatively consistent within the present analysis. The distinct roles of TA and GL in ankle control, foot clearance, support, and propulsion provide a biomechanical context for considering this pattern [1,2,23]. However, the normalized HAM peak finding indicates that the observed associations were not confined to distal muscles. Moreover, task-reference normalization expresses each feature relative to the peak of its participant-specific ensemble-mean envelope, changing the quantity being compared. The HAM result therefore concerns a relative signal-amplitude characteristic and should not be interpreted as evidence of greater absolute hamstring activation in males. The present analyses did not test the biomechanical or neural mechanisms underlying these associations.

Previous studies provide context for the ankle-muscle findings. Di Nardo *et al.* reported sex-related differences in the occurrence of activation patterns in TA, GL, and VL [5], whereas Mengarelli *et al.* found more frequent TA-GL coactivation in females [6]. These observations concern recruitment timing and occurrence rather than whole-cycle signal magnitude and therefore do not directly confirm the higher RMS and peak amplitudes observed in males here. They suggest that sex-related associations may depend on both the muscle examined and the EMG characteristic selected. In the present study, RF and VL comparisons did not reach significance after Holm correction across the three

amplitude-processing variants, whereas the HAM peak comparison was significant only after task-reference normalization. This pattern does not establish a general distinction between distal and proximal muscles, and non-significant comparisons should not be interpreted as evidence of equivalence between sexes.

Interpretation of both non-normalized and task-reference-normalized surface EMG requires caution. Recorded amplitudes can be influenced by electrode-muscle distance, subcutaneous tissue thickness, muscle geometry, and cross-talk [4,8,9]. Task-reference normalization expresses the measured features relative to an internal walking reference and therefore addresses a different comparison from that based on physical signal amplitudes [7] . In the present analysis, the persistence of selected contrasts after normalization indicates that these associations were also detectable on the chosen relative scale; it does not establish that anatomical or recording-related influences were eliminated. Furthermore, the reference was derived from the same walking recordings and did not provide an independent measure of maximal activation capacity. Accordingly, neither persistence nor loss of statistical significance after normalization identifies the physiological mechanism underlying a sex-related association, and the findings should not be interpreted as direct evidence of greater neural drive in males.

The functional significance of the observed amplitude differences remains undetermined. Although synchronized footswitch and knee electrogoniometer recordings were available in the source dataset [11,12], the present analysis did not examine sex-related differences in spatiotemporal gait measures or joint kinematics, or their associations with EMG amplitude. The amplitude findings therefore cannot establish differences in walking performance, stability, or efficiency. Likewise, the absence of a robust sex association in stride-to-stride RMS CV concerns variability in this EMG feature and does not demonstrate equivalence in locomotor behavior.

The amplitude outcomes quantify related but distinct signal characteristics. RMS gives greater weight to larger sample amplitudes through squaring, peak amplitude captures the largest absolute sample within a cycle, and iEMG accumulates absolute amplitude over time and therefore also depends on cycle duration. In the QC-screened analyses, TA and GL peak contrasts and the GL RMS contrast were supported across processing variants, whereas the TA RMS contrast was sensitive to processing. No muscle-specific iEMG contrast remained significant after Holm correction. Task-reference normalization additionally supported a HAM peak contrast, illustrating that the pattern depended on both the selected feature and its reference scale. Because the normalization reference was obtained from a smoothed ensemble-mean envelope, normalized peak values describe individual-cycle signal maxima relative to that reference, rather than the peak of the normalized envelope itself. Normalized iEMG retains units of time and should not be interpreted as a dimensionless activation measure. These differences in statistical findings do not establish distinct physiological mechanisms; their explanation would require further examination of cycle durations, within-cycle signal profiles, and the uncertainty of the estimated contrasts.

The stride-to-stride RMS CV findings were sensitive to both signal processing and channel inclusion. In the QC-screened dataset, the omnibus sex-related component and sex-by-muscle interaction were significant after band-pass filtering, but not in the raw-signal analysis. Neither finding was reproduced when the two excluded channels were included. This dependence on analytical conditions does not support a consistent sex-related difference in stride-to-stride RMS variability. Task-reference normalization left the filtered RMS CV unchanged because all cycles within each participant-side-muscle combination were divided by the same positive reference. This mathematical invariance should not be interpreted as independent confirmation of robustness. Finally, variability in cycle-specific RMS amplitude is distinct from variability in gait timing or joint motion; the present CV findings therefore do not establish either differences or equivalence in locomotor stability.

The practical relevance of these findings lies primarily in the interpretation and design of gait sEMG studies [24]. The muscle-dependent amplitude associations suggest that sex should be considered alongside the recording site and the selected EMG outcome when planning comparisons. However, the present estimates are not validated clinical reference limits and do not justify sex-specific rehabilitation targets. For gait assessment, higher non-normalized amplitude should not automatically be interpreted as greater muscle activation or better function. Future studies should use normalization procedures appropriate to the research question and clearly document acquisition and processing methods [7,15]. Larger cohorts incorporating local tissue measurements, walking speed, joint kinematics, and functional outcomes could

determine whether the observed signal differences have relevance to impairment or rehabilitation response.

Several strengths of the present study should be noted. First, the prolonged recordings provided repeated observations for characterizing each participant's sEMG features. However, the independent sample comprised 31 participants, with 28 included in the anthropometrically adjusted models; the 13,465 gait cycles did not constitute independent observations or increase the number of participants available for between-sex comparisons. Second, the use of a publicly available dataset improves transparency and reproducibility [10-12]. Third, the statistical approach accounted for repeated observations by summarizing stride-level features at the participant-side-muscle level and using mixed-effects models [16]. Finally, comparisons of raw, band-pass-filtered, and task-reference-normalized features allowed sensitivity to signal processing to be examined using matched observations within each dataset. Repeating the analyses with the two excluded channels included also allowed the influence of the channel exclusions on the estimated associations to be assessed.

Several limitations should be considered. The modest participant sample (31 overall and 28 in the adjusted models) limits the precision and generalizability of between-sex comparisons. Repeated gait cycles do not increase the number of independent participants, and non-significant comparisons should not be interpreted as evidence of equivalence. Height and body weight were unavailable for three participants, who were excluded from the adjusted models. The cohort comprised healthy young adults, limiting generalization to older adults and clinical populations. MVC recordings were unavailable; although task-reference normalization was performed, the resulting ratios do not quantify activation relative to maximal capacity [7,15]. Neither anthropometric adjustment nor task-reference normalization establishes independence from local tissue thickness, electrode-muscle geometry, or other recording-related influences [7-9]. As a secondary analysis, this study relied on the acquisition procedures and documentation of the original dataset [11,12]. Electrode placement, inter-electrode spacing, and skin preparation could not be independently verified retrospectively. The HAM channel could not be attributed to an individual hamstring muscle. Additional screening and targeted visual inspection identified two channels with signal-quality concerns, and their influence was evaluated through channel-exclusion sensitivity analyses. However, this assessment did not establish that the retained recordings were free of artifacts. Whole-channel exclusion also removed potentially usable segments, and the CV findings were sensitive to these exclusions. Comparisons across the selected processing variants do not establish robustness to all possible filtering or normalization choices. Finally, the unresolved S01 header-unit discrepancy limits certainty about the physical-unit interpretation of its non-normalized amplitudes.

In conclusion, sex-related differences in recorded lower-limb sEMG during prolonged overground walking depended on the muscle, outcome, and processing approach. In the QC-screened dataset, higher peak amplitudes in males for TA and GL and higher RMS amplitudes for GL were supported across raw, band-pass-filtered, and task-reference-normalized analyses. The TA RMS comparison was sensitive to processing, whereas a HAM peak difference emerged only after task-reference normalization. No muscle-specific iEMG comparison was significant after Holm correction, and stride-to-stride RMS CV findings were sensitive to processing and channel exclusions. These findings characterize differences in recorded signal magnitude on physical and task-relative scales; they do not establish differences in neural drive, maximal activation capacity, or walking function.

Conflict of Interest

There is no conflict of interest.

Acknowledgements

The author gratefully acknowledges the creators of the publicly available PhysioNet dataset used in this study. The dataset was originally collected and made publicly available by Di Nardo *et al.*, whose contribution made this secondary analysis possible [11,12].

References

1. Perry J, Burnfield JM, Cabico LM. Gait Analysis: Normal and Pathological Function. J Sports Sci Med 2010;9:353.

2. Neptune RR, Kautz SA, Zajac FE. Contributions of the Individual Ankle Plantar Flexors to Support, Forward Progression and Swing Initiation during Walking. J Biomech 2001;34:1387-1398. https://doi.org/10.1016/S0021-9290(01)00105-1
3. Agostini V, Ghislieri, M, Rosati S, Balestra G, Knaflitz M. Surface Electromyography Applied to Gait Analysis: How to Improve Its Impact in Clinics? Front Neurol 2020:11:994. https://doi.org/10.3389/fneur.2020.00994
4. Hermens HJ, Freriks B, Disselhorst-Klug C, Rau G. Development of Recommendations for SEMG Sensors and Sensor Placement Procedures. J Electromyogr Kinesiol 2000;10:361-74. https://doi.org/10.1016/S1050-6411(00)00027-4
5. Di Nardo F, Mengarelli A, Maranesi E, Burattini L, Fioretti S. Gender Differences in the Myoelectric Activity of Lower Limb Muscles in Young Healthy Subjects during Walking. Biomed Signal Process Control 2015;19:14-22. https://doi.org/10.1016/j.bspc.2015.03.006
6. Mengarelli A, Maranesi E, Burattini L, Fioretti S, Di Nardo F. Co-Contraction Activity of Ankle Muscles during Walking: A Gender Comparison. Biomed Signal Process Control 2017;33:1-9. https://doi.org/10.1016/j.bspc.2016.11.010
7. Besomi M, Hodges PW, Clancy EA, Van Dieën J, Hug F, Lowery M, et al. Consensus for Experimental Design in Electromyography (CEDE) Project: Amplitude Normalization Matrix. J Electromyogr Kinesiol 2020:53:102438. https://doi.org/10.1016/j.jelekin.2020.102438
8. Nordander C, Willner J, Hansson GÅ, Larsson B, Unge J, Granquist L, Skerfving S. Influence of the Subcutaneous Fat Layer, as Measured by Ultrasound, Skinfold Calipers and BMI, on the EMG Amplitude. Eur J Appl Physiol 2003;89:514-519. https://doi.org/10.1007/s00421-003-0819-1
9. Kuiken TA, Lowery MM, Stoykov NS. The Effect of Subcutaneous Fat on Myoelectric Signal Amplitude and Cross-Talk. Prosthet. Orthot Int 2003;27:48-54. https://doi.org/10.3109/03093640309167976
10. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000;101:E215-E220. https://doi.org/10.1161/01.CIR.101.23.e215
11. Di Nardo F, Morbidoni, C, Fioretti S. Surface Electromyographic Signals Collected during Long-Lasting Ground Walking of Young Able-Bodied Subjects. PhysioNet 2024. https://doi.org/10.13026/grdj-qx06
12. Di Nardo F, Morbidoni C, Iadarola G, Spinsante S, Fioretti S. Long Duration Multi-Channel Surface Electromyographic Signals during Walking at Natural Pace: Data Acquisition and Analysis. PLoS One 2025;20:e0318560. https://doi.org/10.1371/journal.pone.0318560
13. Di Nardo F, Morbidoni C, Cucchiarelli A, Fioretti S. Influence of EMG-Signal Processing and Experimental Set-up on Prediction of Gait Events by Neural Network. Biomed Signal Process Control 2021;63:102232. https://doi.org/10.1016/j.bspc.2020.102232
14. Xie C, McCullum L, Johnson A, Pollard T, Gow B, Moody B. Waveform Database Software Package (WFDB) for Python. PhysioNet 2023. https://doi.org/10.13026/9njx-6322
15. Besomi M, Devecchi V, Falla D, McGill K, Kiernan MC, Merletti R, et al. Consensus for Experimental Design in Electromyography (CEDE) Project: Checklist for Reporting and Critically Appraising Studies Using EMG (CEDE-Check). J Electromyogr Kinesiol 2024;76:102874. https://doi.org/10.1016/j.jelekin.2024.102874
16. Laird NM, Ware JH. Random-Effects Models for Longitudinal Data. Biometrics 1982;38:963-974. https://doi.org/10.2307/2529876
17. Lenth RV. Least-Squares Means: The R Package Lsmeans. J Stat Softw 2016;69:1-33. https://doi.org/10.18637/jss.v069.i01
18. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scand J Statist 1979;6:65-70. https://www.jstor.org/stable/4615733
19. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array Programming with NumPy. Nature 2020;585:357-362. https://doi.org/10.1038/s41586-020-2649-2
20. McKinney W. Data Structures for Statistical Computing in Python. SciPy 2010;56-61. https://doi.org/10.25080/Majora-92bf1922-00a

21. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nat Methods 2020;17:261-272. https://doi.org/10.1038/s41592-019-0686-2
22. Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. SciPy 2010;92-96. https://doi.org/10.25080/Majora-92bf1922-011
23. Di Nardo F, Ghetti G, Fioretti S. Assessment of the Activation Modalities of Gastrocnemius Lateralis and Tibialis Anterior during Gait: A Statistical Analysis. J Electromyogr Kinesiol 2013;23:1428-33. https://doi.org/10.1016/j.jelekin.2013.05.011
24. Žarković D, Šorfová M, Tufano JJ, Kutílek P, Vítečková S, Ravnik D, et al. Gait Changes Following Robot-Assisted Gait Training in Children with Cerebral Palsy. Physiol Res 2021;70(Suppl 3):S397-S408. https://doi.org/10.33549/physiolres.934840