

OPEN

Open-Source Remote Gait Analysis: A Post-Surgery Patient Monitoring Application

Reed D. Gurchiek^{1*}, Rebecca H. Choquette², Bruce D. Beynon², James R. Slauterbeck², Timothy W. Tourville³, Michael J. Toth^{2,4} & Ryan S. McGinnis¹

Critical to digital medicine is the promise of improved patient monitoring to allow assessment and personalized intervention to occur in real-time. Wearable sensor-enabled observation of physiological data in free-living conditions is integral to this vision. However, few open-source algorithms have been developed for analyzing and interpreting these data which slows development and the realization of digital medicine. There is clear need for open-source tools that analyze free-living wearable sensor data and particularly for gait analysis, which provides important biomarkers in multiple clinical populations. We present an open-source analytical platform for automated free-living gait analysis and use it to investigate a novel, multi-domain (accelerometer and electromyography) asymmetry measure for quantifying rehabilitation progress in patients recovering from surgical reconstruction of the anterior cruciate ligament (ACL). Asymmetry indices extracted from 41,893 strides were more strongly correlated ($r = -0.87$, $p < 0.01$) with recovery time than standard step counts ($r = 0.25$, $p = 0.52$) and significantly differed between patients 2- and 17-weeks post-op ($p < 0.01$, effect size: 2.20–2.96), and controls ($p < 0.01$, effect size: 1.74–4.20). Results point toward future use of this open-source platform for capturing rehabilitation progress and, more broadly, for free-living gait analysis.

The digital medicine revolution is driven by advances in wearable sensor technology and the algorithms for analyzing and interpreting their data. These mobile health technologies enable improved remote patient monitoring, personalized intervention, and could be used to provide improved continuity across care transitions. While transitional care has been recognized as a national priority¹, and has been shown to improve outcomes, increase the efficient use of health care resources, and decrease health care costs^{2,3}, current interventions are resource and personnel intensive⁴. Digital medicine innovations that harness existing cyber infrastructures, wearable sensors, and mobile devices may improve the efficiency and effectiveness of transitional care interventions. This approach could be transformative for a broad range of clinical domains, including for neurological⁵, musculoskeletal⁶, and mental health^{7,8} conditions.

Concerning current techniques for remote patient monitoring, physical activity is the most targeted health behavior, with a large commercial market. However, these measures are too general for most clinical applications as they provide minimal biomechanical or physiological insight at a joint- or limb-specific level. For both neurological and musculoskeletal disorders, physical activity (e.g. step counts) is often the primary outcome measure from free-living wearable sensor data^{9,10}. In these clinical populations, traditional gait analysis provides far more valuable information concerning neuromusculoskeletal health and, in the aftermath of various clinical interventions, recovery of physical function¹¹. However, these traditional assessments are constrained to specialized motion analysis laboratories which may not accurately reflect an individual's free-living gait¹². This motivates the pursuit of remote gait analysis techniques able to capture more clinically relevant biomechanics including quantification of motor control indices (e.g. muscle activation patterns¹³) and musculoskeletal dynamics¹⁴.

To answer this unmet need, several groups have started to explore methods for tracking free-living gait biomechanics^{12,15–21}. The general framework for remote gait analysis shared by these efforts has three steps: (1) identification of walking bouts, (2) stride detection, and (3) analysis. This approach has been used to detect and characterize bradykinesia in Parkinson's disease patients¹⁵ and demonstrate that in-lab observations of gait

¹M-Sense Research Group, University of Vermont, Burlington, VT, 05405, USA. ²Department of Orthopaedics and Rehabilitation, University of Vermont, Burlington, VT, 05405, USA. ³Department of Rehabilitation and Movement Science, University of Vermont, Burlington, VT, 05405, USA. ⁴Department of Medicine, University of Vermont, Burlington, VT, 05405, USA. *email: reed.gurchiek@uvm.edu

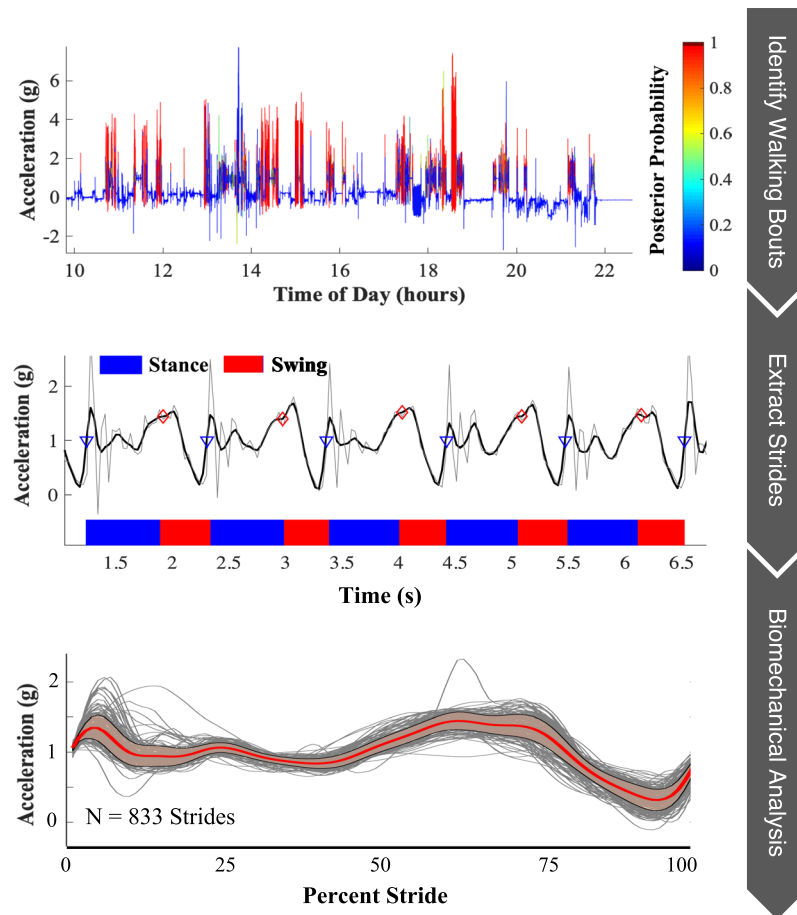


Figure 1. Graphical summary of the proposed remote gait analysis. The proposed approach is comprised of three basic steps: (1) walking bout identification, (2) stride extraction and gait phase segmentation, and (3) biomechanical analysis of individual strides.

speed¹² and gait asymmetry^{20,21} differ from daily-life. The components of this framework reflect recent advances in wearable sensor-based activity identification²², event detection²³, and biomechanical analysis^{24–29}. The development and application of these solutions are multi-disciplinary efforts requiring research teams with expertise in medicine, data science, and engineering. Open-source solutions may allow analysis of wearable sensor data collected in free-living conditions by teams without expertise in signal processing or machine learning. This reduces barriers to applying digital health solutions for remote patient monitoring.

To this end, we present an open-source analytical platform which captures this general framework (Fig. 1) and could be applied for remote biomechanical analysis of any task (see <https://github.com/M-SenseResearchGroup/RemoteBMX>). It has been designed with a modular structure to enable flexibility and to encourage improvements and customization from members of the scientific community. For example, customization could allow for the analysis of other tasks (e.g. stair ascent, running), the extraction of alternative clinical information, or the utilization of different wearable sensors. These modifications are dependent on the patient population being monitored.

To demonstrate how the platform may be tailored for a specific patient population, it was deployed as a rehabilitation monitoring application in patients recovering from reconstructive surgery of the anterior cruciate ligament (ACL-R). ACL-R is the most effective treatment for ACL rupture³⁰. However, approximately 50% of patients who undergo ACL-R will go on to develop post-traumatic osteoarthritis (PTOA)³¹. Previous research suggests that altered gait biomechanics following ACL-R are responsible for this phenomenon^{31,32}. Pathological gait is often characterized by inter-limb asymmetries which manifest early in the post-surgical period^{33–35}, and develop into compensatory gait patterns over time^{11,36}. It is imperative that these maladaptations are identified early so that corrective rehabilitative interventions can be pursued³⁷. Thus, this application presents an ideal candidate for which to demonstrate the clinical utility of the proposed platform.

Platform modifications were context-specific and driven by the current understanding of the clinical problem. In the context of ACL-R, inter-limb gait asymmetry has been identified as a biomarker for recovery^{33,34,38}. These asymmetries represent altered gait kinematics that are linked to PTOA-related knee joint contact forces in this population^{11,39,40}. Quadriceps muscle dysfunction is common following ACL-R^{13,38,41,42} and may be responsible for the development of pathological gait. Thus, direct observation of asymmetric gait kinematics and quadriceps muscle activity in the early post-operative period could identify pathological gait and signal the need for remedial

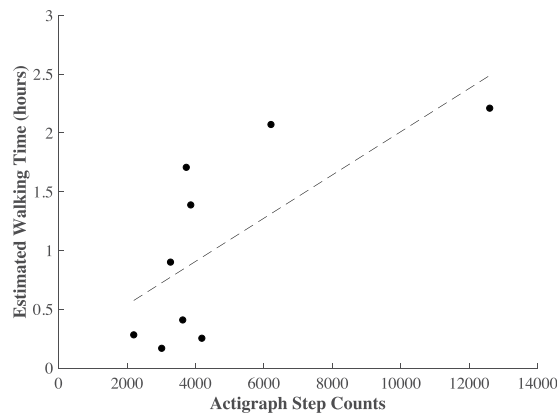


Figure 2. Scatter plot of the total time spent walking from the proposed method vs step counts estimated by Actigraph activity monitors.

actions to prevent developing PTOA. However, algorithms for extracting this information from wearable sensor data captured in free-living conditions do not yet exist.

To this end, patients recovering from ACL-R present an ideal population for which to specify and test the use of the proposed platform. Our analysis was tailored for monitoring gait asymmetries relating to spatiotemporal variables, and kinematic and muscle activation time-series. Our novel analysis quantifies asymmetry in the following gait measures: duty factor (DF); mean surface electromyography (sEMG) amplitude of the rectus femoris during the stance and swing phase of gait; muscle activation time-series; as well as cranial-caudal (CC), medio-lateral (ML), and antero-posterior (AP) thigh acceleration time-series. A composite asymmetry score was also defined as the average of these seven asymmetries. We show the ability of these asymmetry indices to discriminate between patients at different time points in the recovery process and compared to healthy controls. Step count estimates from a commercially available activity monitor worn by some of the patients also enables a comparison of the proposed analysis to current techniques for monitoring recovery.

Results

Subjects were categorized as either T1 (less than 6 weeks post-surgery), T2 (greater than 6 weeks post-surgery) or C (healthy controls). Data were collected for 20.24 ± 6.28 hours on average for each subject (except one T1 patient for whom no walking bouts met the criteria for analysis) and a total of 41,893 strides were analyzed (T1: 1,743 strides, T2: 9,616 strides, C: 20,939 strides). Estimated total time spent walking (hours) was strongly correlated ($r = 0.71$, $p = 0.03$) with Actigraph step counts (Fig. 2) but showed a stronger correlation with recovery time ($r = 0.63$, $p = 0.08$) than Actigraph step counts ($r = 0.25$, $p = 0.52$), although neither was statistically significant. Importantly, the composite asymmetry score ($r = -0.87$, $p < 0.01$) and stride time ($r = -0.91$, $p < 0.01$) were both strongly associated with recovery time (see Fig. 3).

Stride times and most gait asymmetry measures decreased across groups (i.e. T1 > T2 > C). Pairwise significant differences were found between C and T1 as well as T1 and T2 for stride times and all asymmetries (all with large effect sizes) except mean normalized sEMG which was trending towards significance (stance: $p = 0.05$, swing $p = 0.12$) (Table 1). All time-series (sEMG, CC, AP, ML), duty factor, and composite asymmetries were significantly different ($p = 0.01$) between C and T1, but only the CC acceleration time-series, duty factor, and composite asymmetries were significantly different between T1 and T2 (Table 1). There were no differences between C and T2 for any asymmetries.

The proposed framework enables the investigation of gait biomechanics continuously throughout the day. To demonstrate the utility of this application, composite asymmetry scores averaged over each 15-minute bin are illustrated in Fig. 4 for one patient with longitudinal observations at about 2 weeks post-surgery (red dashed line) and again 17 weeks later (blue dashed line). The solid lines illustrate the average trend of the other groups for comparison (T1: red, T2: blue, C: black).

Discussion

Herein, we present a new computational platform that enables multi-modal gait analysis in free-living conditions based on data from commercially available wearable devices. We demonstrate the utility of this platform in a sample of patients recovering from ACL-R. Results suggest that platform-derived gait measures agree with gold standard actigraphy but better discriminate between patients' gait at different time points in the recovery process. Similarly, platform-derived measures are able to detect gait differences between patients at different stages of recovery, and with large effect sizes. We further discuss the implications of these results and how this new platform could be deployed for remote gait analysis in a variety of clinical populations.

Currently, actigraphy technologies are broadly used for characterizing free-living physical activity and gait^{9,10}. The statistically significant association we observe (Fig. 2; $r = 0.71$, $p = 0.03$) between actigraphy-derived daily step counts and platform-derived walking time suggest that the proposed analysis platform is valid for capturing free-living physical activity. Similar findings of agreement between different indices of physical activity have been reported elsewhere⁹. Although the correlation between recovery time and the platform-derived measure of

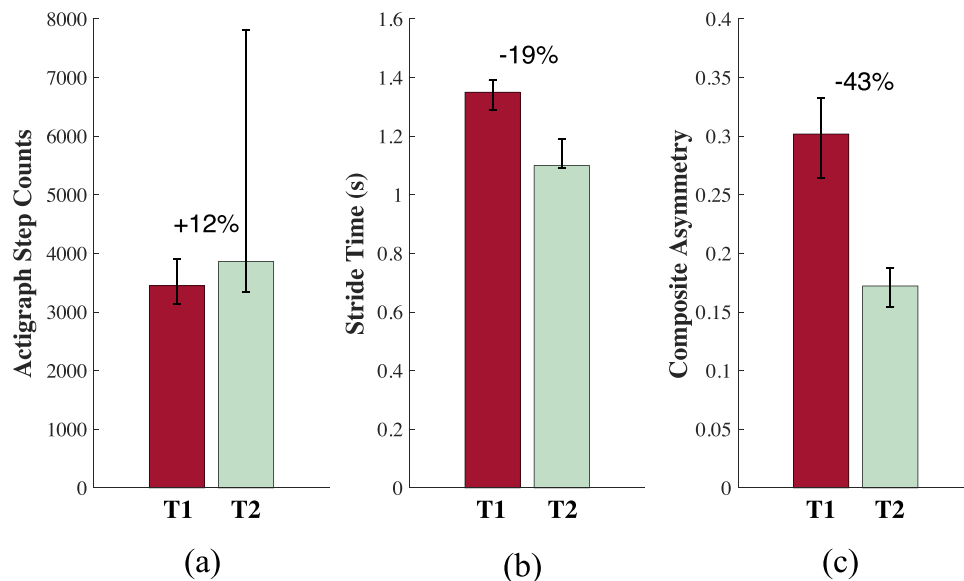


Figure 3. Percent difference in the median Actigraph step counts (a), strides times (b), and composite asymmetry scores (c) between the T1 (red) and T2 (green) groups. Error bars denote the 25th and 75th quantiles.

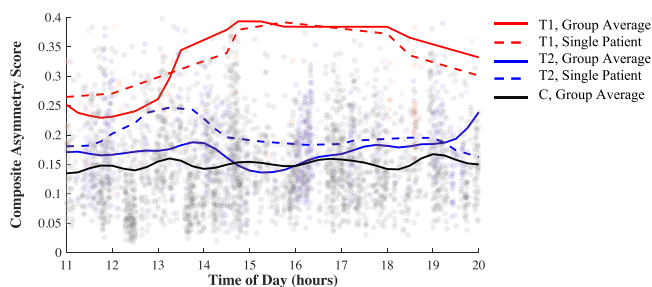


Figure 4. Composite asymmetry score throughout the day (averaged over every 15-minute bin) for a patient with longitudinal observations: 2.1 weeks post-surgery (red dashed line) and 19.1 weeks post-surgery (blue dashed line). The solid lines illustrate the average trends for the T1 (red), T2 (blue), and C (black) groups. The longitudinal patient’s data was not included in the group means.

	Mean (SD)			Pairwise Comparisons		
	T1	T2	C	C-T1	C-T2	T1-T2
ST	1.34 (0.08)	1.14 (0.07)	1.10 (0.05)	4.08**	0.34	2.70**
DF	0.13 (0.05)	0.04 (0.02)	0.03 (0.01)	4.20**	0.61	2.48**
EMG St	0.22 (0.12)	0.13 (0.03)	0.15 (0.04)	ANOVA $p = 0.05$		
EMG Sw	0.41 (0.14)	0.27 (0.04)	0.29 (0.12)	ANOVA $p = 0.12$		
EMG(t)	0.41 (0.10)	0.29 (0.11)	0.25 (0.07)	2.10**	0.52	0.07
AP(t)	0.32 (0.13)	0.10 (0.05)	0.07 (0.04)	3.57**	0.49	0.08
ML(t)	0.38 (0.13)	0.28 (0.04)	0.24 (0.06)	1.74**	0.51	0.08
CC(t)	0.20 (0.08)	0.06 (0.04)	0.03 (0.02)	4.19**	0.26	2.20**
Comp.	0.29 (0.05)	0.17 (0.03)	0.15 (0.03)	3.78**	0.65	2.96**

Table 1. Comparison of daily average stride times and asymmetries. ST: Stride Time; units seconds. Duty Factor (DF), EMG Stance (EMG St), and EMG Swing (EMG Sw) asymmetry scores are the percent difference between the healthy and injured leg (i.e. 0.5 indicates that the between leg difference is 50% that of the healthy leg). EMG(t), AP(t), ML(t), and CC(t) are pattern asymmetries for the sEMG time-series and the antero-posterior, medio-lateral, and cranial-caudal thigh acceleration time-series respectively. Composite asymmetry score (Comp.) is the average value of the other seven asymmetry scores. Bold numbers in the pairwise comparisons are effect sizes ($*p \leq 0.05$, $**p \leq 0.01$) and non-bold numbers are the p values for non-significant pairwise differences.

walking time ($r = 0.63, p = 0.08$) appears stronger than that for actigraphy-derived step counts ($r = 0.25, p = 0.52$), neither was statistically significant, which also agrees with previous work⁹. This finding is intuitive as there are a variety of extraneous variables that affect daily physical activity and have nothing to do with rehabilitation progress (e.g. weather, day of the week). These results suggest a fundamental limitation of physical activity indices for monitoring rehabilitation progression. The vast majority of research investigating pathological biomechanics characteristic of gait following ACL-R suggests the existence of more sensitive metrics including quadriceps activation⁴³, ground reaction force³³, and joint work³⁴. However, these metrics currently require laboratory-based methodologies that are not widely transferable to clinical use.

The proposed platform answers this unmet need by capturing more subtle biomechanical changes in gait associated with recovery under free-living conditions. Our results demonstrate that platform-derived stride times are more strongly associated with recovery time ($r = -0.91, p < 0.01$) than both actigraphy-derived step counts ($r = 0.25, p = 0.52$) and platform-derived estimates of total walking time ($r = 0.63, p = 0.08$). These results are supported by previous work where decreased walking speeds have been associated with decreased joint health following ACL-R⁴⁴ (stride times are associated with walking speed⁴⁵). These results provide additional evidence in support of the utility of platform-derived measures of free-living gait biomechanics.

Stride time, walking speed, and other spatiotemporal parameters are informative biomechanical measures related to recovery in this population, and more broadly are important indicators of mobility impairment^{12,27,28,44}. Nevertheless, a need has been recognized for more sensitive biomarkers. For example, it has been shown that spatiotemporal *symmetry* may manifest even in the presence of a true gait abnormality⁴⁶. Further, research suggests the pursuit of metrics that characterize the *full waveform pattern* of mechanical variables throughout the gait cycle⁴⁷. In light of these results, the proposed platform has been designed to capture additional measures that may be more indicative of, and sensitive to, rehabilitation progress. Figure 3 and Table 1 report results from our efforts to define a novel gait asymmetry analysis that incorporates temporal (duty factor), kinematic (ML, AP, CC acceleration time-series), and neuromuscular (sEMG Swing, sEMG Stance, and sEMG time-series) measures. The effect sizes of the differences between groups were largest for the asymmetries relating to kinematic measures and duty factor. The mean normalized sEMG during stance and swing asymmetries were the only variables for which no significant difference was found, which may reflect increased within subject and group variance. The composite measure that captures asymmetries within each of these domains demonstrates significant differences between T1 and T2/C with large effect sizes ($d = 2.96$ and $d = 3.78$ respectively) suggesting that this measure may be useful for tracking biomechanical changes associated with rehabilitation progress.

To this end, we further examine the association of these measures of free-living gait biomechanics with time since surgery. Both stride time and our new composite asymmetry score present strong relationships with recovery time ($r = -0.91, p < 0.01$ and $r = -0.87, p < 0.01$ respectively), and are noticeably larger than that observed for any gross index of physical activity. Further, when comparing T1 and T2, the largest effect sizes were observed for the composite asymmetry ($d = 2.96$) which was also responsible for the largest percent difference between T1 and T2 (-43%) when compared to both stride time (-19%) and step counts ($+12\%$) (Fig. 3). This suggests that these more detailed biomechanical measures may provide increased sensitivity to recovery time in this population and thus may be suitable candidates to pursue in developing novel *digital biomarkers* for tracking rehabilitation progress and gait asymmetries that have pathological consequences^{33,34}. The proposed asymmetry analysis also provides the clinician with insight into patient-specific adaptations and their particular mechanistic origin since it captures indices of both muscle activation and limb kinematics.

Having established the improved association between free-living measures of gait asymmetry and recovery time, we further examine how this new measure changes over the entire wear time for patients monitored in this study. Figure 4 reports the composite asymmetry score captured during every gait bout identified between 11 am and 8 pm for a single patient with longitudinal observations and with the group average trend for subjects in the T1, T2, and C groups for comparison. The difference between groups based on asymmetry magnitudes alone reflects the same general pattern observed in Table 1, namely a convergent trend toward decreased gait asymmetries with increased recovery time. However, Fig. 4 provides additional insight whereby gait asymmetries appear more variable throughout the day for the early post-surgery time point when compared to the later time point and that of the control group. This supports the need for remote gait monitoring, as this variability would not be captured in a single gait assessment. Further investigation is necessary to understand the origins of this observation which may, for example, indicate an increased susceptibility to fatigue early in rehabilitation. The similar trends observed between the dashed lines (single patient, different time points) and the solid lines (average of all other patients within respective group) of like colors suggests the results of our cross-sectional design may mirror what would be seen in a longitudinal study. Further investigation using a longitudinal design is necessary to confirm this conjecture.

Far fewer strides were identified and analyzed for patients in the T1 group (≈ 350 strides/day per patient) than for those in the T2 ($\approx 1,600$ strides/day per patient) and C ($\approx 1,310$ strides/day per patient) groups. One explanation is that the T1 patients simply walk less, which is also supported by the Actigraph step counts. Intuitively, this may reflect a natural tendency for an individual to avoid an activity like walking which loads the recently reconstructed knee. Despite the difference in the number of strides analyzed for the T1 group compared to T2/C, we see this as an acceptable limitation as the availability of gait biomechanics characterizing even 350 strides per day is already a substantial improvement over the current standard which often yields fewer observed strides and with limited ecological validity.

In the approach to remote patient monitoring proposed herein, we chose sensor locations that would minimize user burden while also providing the kinematic and muscle activity data key to the presented analysis. A minimum of one sensor per leg was required to extract inter-leg kinematic and muscle activation asymmetries. Future work could consider additional sensor locations in an effort to provide improved clinical utility. For example, if the muscle activity from a quadriceps antagonist (e.g. a knee flexor) were available it may provide insight

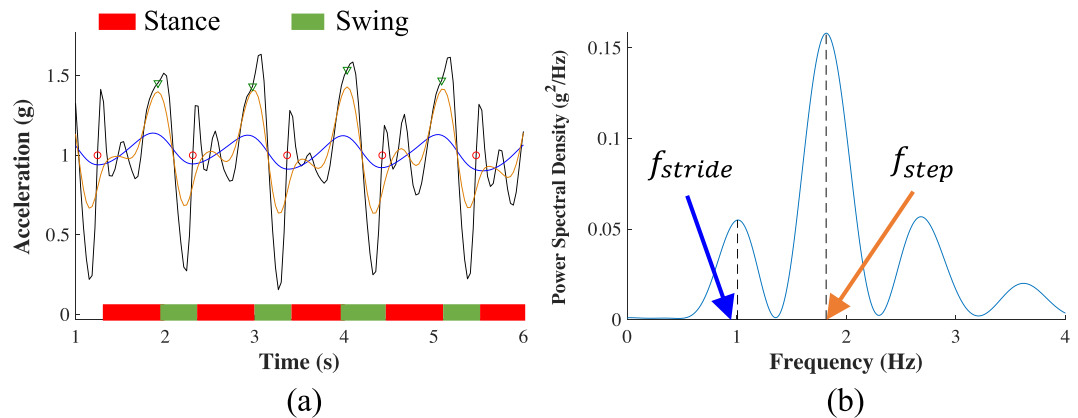


Figure 5. Stride detection and segmentation example. (a) Foot contact (red circles) and foot off (green triangles) events are identified using the CC-axis accelerometer time-series lowpass filtered with a 5 Hz cutoff (black trace) and with cutoff frequencies equal to the approximate step frequency (orange trace) and stride frequency (blue trace). Step and stride frequencies are approximated using the power spectral density of the raw accelerometer signal (b).

into co-contraction indices and would perhaps make the knee extensor muscle activation time-series more interpretable. Inclusion of a sensor on the shank could allow for the extraction of knee joint kinematics^{48,49} that could add additional information indicative of recovery. However, current hardware constraints on the capacity of on-board memory and battery prevented use of gyroscope data and required a relatively low accelerometer sampling frequency (31.25 Hz) in order to enable sEMG data collection over the recording durations considered herein. While higher sampling frequencies and additional sensor modalities would be useful, accelerometer data recorded at 31.25 Hz is likely sufficient for capturing the segment kinematics of interest which is supported given the accelerometer signal power spectrum (see Fig. 5) and the fact that traditional laboratory-based gait analysis often employs cutoff frequencies ≤ 8 Hz for lowpass filtering kinematic data^{45,50,51}.

The results presented herein demonstrate the clinical utility of the proposed analysis platform for remote monitoring of patients recovering from ACL reconstruction surgery. Our analysis describes an unprecedented evaluation of muscle activation concurrently with kinematical time-series during free-living gait. These constitute a novel asymmetry analysis which presents a substantial improvement over current techniques utilized for remote monitoring. Further investigation is necessary to fully realize the practical application of the proposed approach for continuous patient monitoring with a larger sample, longitudinal design, and more frequent observation (i.e. more than just two time points) during the recovery process. The underlying MATLAB code has been open-sourced and structured in a modular fashion so that users can easily add/edit functionality for their specific use case. For example, different activity classification models or stride segmentation algorithms that incorporate data from other wearable sensors located on different body segments could be inserted that are more precisely tuned for a given patient population. There are certainly algorithmic improvements and additions that could be made for the proposed remote gait analysis such as automatic anatomical calibration or estimation of other biomechanical variables. The modularity of the platform is intended to promote these improvements so that the platform may serve as a basis upon which to build a comprehensive approach for remote gait analysis. This contribution comes at a critical time in the use of wearable sensors for providing free-living patient monitoring capabilities. An open-source platform will promote cross-disciplinary efforts to further advance remote patient monitoring paradigms and digital medicine. Future research should investigate similar applications in other clinical contexts and begin the difficult task of understanding how to translate the detailed, big-data evaluations enabled by continuous monitoring paradigms into optimizing patient-specific interventions.

Methods

Study design. Our system was deployed to evaluate gait in three groups of subjects: (1) subjects less than six weeks after ACL-R (T1: 3 male, 3 female, recovery time (mean \pm SD) = 2.1 ± 1.6 weeks, age = 26 ± 11 yo, height = 1.74 ± 0.11 m, mass = 70.52 ± 16.21 kg), (2) subjects at least six weeks after ACL-R (T2: 2 male, 4 female, recovery time = 17.2 ± 2.0 weeks, age = 26 ± 6 yo, height = 1.70 ± 0.13 m, mass = 77.82 ± 15.44 kg), and (3) healthy controls (C: 8 male, 8 female, age = 23 ± 5 yo, height = 1.74 ± 0.11 m, mass = 70.51 ± 13.17 kg). Each subject wore a single sensor (BioStamp, MC10 Inc., Lexington, MA) on the muscle belly of the rectus femoris of each thigh, which recorded tri-axial accelerometer (sampling frequency: 31.25 Hz, range: ± 16 g) and surface electromyography (sEMG) (sampling frequency: 250 Hz) data during daily life for one day. Daily step counts from a waist worn activity monitor (Actigraph, Pensacola, FL) were also available for nine ACL-R patients (T1: N = 4, T2: N = 5). These enable a comparison between the proposed analysis and current standards (e.g step counts⁹) in their respective sensitivity to patient recovery. This study was approved by the University of Vermont Institutional Review Board. All study activities were in accordance with the relevant guidelines and informed consent was obtained from all participants.

Remote gait analysis platform. The general approach to biomechanical analysis of any task in daily life using the proposed platform takes on the following form: (1) activity identification, (2) event detection, and (3) analysis. The specifications for the current ACL-R patient monitoring application are described graphically in Fig. 1. This framework easily allows future users to make application-specific modifications to analyze different tasks (e.g. crutching, stair ascent) or analyze differently the same tasks presented in the current study. For monitoring ACL-R patients, the proposed analysis requires, at a minimum, raw accelerometer data to operate. Herein, we also consider sEMG data to provide a more complete picture of free-living gait biomechanics in this population.

Identification of walking bouts. A support vector machine (SVM) binary classifier (Gaussian kernel) was used to identify walking bouts. The model was trained using annotated tri-axial accelerometer data from the healthy control group recorded during various activities including multi-speed walking and running over ground and on a treadmill, stair ascent and descent, crutching, sitting, standing, and lying down. Annotated walking, crutching, lying down, sitting, and standing data were also available from three ACL-R patients. Data from each activity were partitioned into four-second, non-overlapping windows. This yielded a total of 3,102 observations (1,318 walk; 1,784 not walk) for training the walking classifier.

The constant thigh relative sensor attitude was determined for each leg by considering accelerometer data during a standing calibration trial (as in²⁴) and used to project raw accelerometer data onto the thigh-fixed cranial-caudal axis (referred to as CC_i , $i \in \{R, L\}$, R : right, L : left, directed proximally). The first principal component of acceleration in the transverse plane during each four-second window was assumed to coincide with the antero-posterior axis of the thigh (referred to as AP_i , points anteriorly, see⁴⁸), from which the medio-lateral axis (referred to as ML_i , points laterally) can also be determined. Accelerometer data were projected onto these three anatomical axes and used to extract a total of 152 time and frequency domain features. Performance of the classifier was evaluated using leave-one-subject-out (LOSO) cross-validation where, for each iteration, data from one subject was removed for testing and a classifier was trained on the remaining data using only those features from the training set with a Davies-Bouldin index⁵² (DBI) less than two. From the LOSO validation, the classifier achieved 98.32% accuracy, 97.04% sensitivity, 99.27% specificity, and the area under the ROC curve (AUC) was 1.00. The classifier made no misclassification errors (100% accuracy) on the 132 observations from the three ACL-R patients during the LOSO validation, which supports its use to identify gait in this population.

Following this performance characterization, annotated data from all sixteen healthy controls and the three ACL-R subjects were used to train the classifier deployed for the proposed analysis. Davies-Bouldin feature selection ($DBI < 2$) indicated six features to use as input to the classifier: (i) average 25th quantile of CC_R and CC_L , (ii) average correlation between CC_R with AP_L and CC_L with AP_R , (iii) correlation between CC_R with CC_L , (iv) average skewness of AP_R and AP_L , (v) average percentage of signal power contained below 0.25 Hz for AP_R and AP_L , and (vi) average median value of AP_R and AP_L .

These six features were extracted from each four-second window of raw accelerometer data collected during daily life from all subjects and were used as input to the SVM classifier to identify each window as *walking* or *not walking*. This enabled an estimate of the total amount of time spent walking which we used as an indicator of gross physical activity. Windows labeled as walking were aggregated into walking bouts if (1) at least two consecutive windows (8 seconds) were classified as walking, and (2) the posterior probability of the window's walking classification was at least 0.8. This threshold was determined as the point that minimized the distance between the classifier's ROC curve and the point (0,1) on the true positive rate-false positive rate plane. A very small subset (451 observations or approximately 0.19%) of bouts containing clipped sEMG (e.g. due to sensor delamination from skin) and/or accelerometer signals were removed.

All aspects of the walking classifier presented herein (feature extraction, training, validation, etc.) were performed using the MATLAB R2018a Statistics and Machine Learning Toolbox (Version 11.3) and our open-source platform specifically designed for wearable sensor-based activity identification. This platform streamlines the development of activity classifiers by enabling the building of population-specific feature sets, the extraction of novel features, feature manipulation (e.g. PCA, DBI), training and testing of various classification models, and automated leave-one-subject-out cross-validation with detailed error analysis. For more details see <https://github.com/M-SenseResearchGroup/ActivityIdentification>.

Stride extraction and gait phase segmentation. Gait events were identified for each walking bout by considering the CC accelerometer signal passed through a bank of lowpass filters with cutoff frequencies at 5 Hz and the stride and step frequencies (f_{stride} , f_{step}). Stride and step frequencies were estimated for each walking bout from Welch's power spectral density (PSD) of the signal (see Fig. 5). As illustrated in Fig. 5, local minima and maxima in the f_{stride} - and f_{step} -filtered signals indicate the initiation of the swing phase of gait (i.e. foot off). The first transition in the 5 Hz-filtered signal from below to above 1 g following foot off indicates the initiation of the stance phase of gait (i.e. foot contact). Constraints were placed on the stride time (0.91 s–1.57 s) and duty factor (DF, percentage of stride cycle spent in stance) (0.44–0.73) of identified strides to avoid accidental analysis of non-walking data⁴⁵. To be considered for further analysis, an eight-second walking bout had to include at least two strides.

Biomechanical analysis. Accelerometer and sEMG signals from each extracted stride were used to compute discrete biomechanical variables to evaluate gait. We quantify asymmetries in gait kinematics and muscle activity between legs during each stride using accelerometer data in each anatomical direction lowpass filtered at a frequency of 6 Hz (as per⁵¹) and the envelope of the sEMG data (computed as per⁵¹). These four time-series (sEMG, CC, AP, ML) were normalized by stride time so that each sample corresponds to a percentage of the gait cycle.

Eight indices of asymmetry (referred to as asymmetries) were computed for each eight-second walking bout. Three of these asymmetries ($a_k, k \in \{1, 2, 3\}$) were the relative difference between discrete biomechanical variables of each leg as per

$$a_k = \left| \frac{I_k - H_k}{H_k} \right| \quad (1)$$

where I_k and H_k are the values of discrete variable k for the injured and healthy leg respectively. The three discrete variables for this analysis were mean sEMG during stance, mean sEMG during swing, and DF. The other four asymmetries (a_m) are defined as per

$$a_m = \frac{1}{2}(1 - r_m). \quad (2)$$

where $m \in \{\text{sEMG, CC, AP, ML}\}$, and r_m is the correlation coefficient of the ensemble means of like time-series between the injured and healthy legs. The final index of asymmetry was a composite asymmetry score equal to the mean of the aforementioned asymmetries.

Statistical analysis. Outlier asymmetries were identified for each patient and removed after the biomechanical analysis as a final check to remove potentially errant data. A one-way analysis of variance (ANOVA) was used to compare the daily average of gait asymmetries and stride times between the three groups. Normality was checked using the Kolmogorov-Smirnov test. If the assumption of normality was violated, group distributions were compared using the Kruskal-Wallis test. If a significant difference was detected, post-hoc pairwise comparisons were made using Tukey's honest significant difference criterion. Effect sizes (Cohen's d) were computed where the ANOVA revealed significant differences and were interpreted qualitatively as weak ($d < 0.25$), small ($0.25 \leq d < 0.5$), medium ($0.5 \leq d < 1.0$), and large ($d \geq 1.0$)⁵³. The agreement between our estimate of gross physical activity (total time spent walking) and the Actigraph step counts estimate was evaluated using Pearson's correlation.

To compare the sensitivity of physical activity with that of our composite asymmetry score and estimated stride times to time spent in recovery, we also determine the correlation between recovery time with composite asymmetry, stride time, Actigraph step counts, and estimated time spent walking for each ACL-R patient using Spearman's rank correlation. The level of significance was set to 0.05 for all statistical tests.

Received: 16 July 2019; Accepted: 14 November 2019;

Published online: 29 November 2019

References

1. Fisher, E. S. *et al.* Evaluating Efficiency Across Patient-Focused Episodes of Care (2009).
2. Naylor, M. D., Aiken, L. H., Kurtzman, E. T., Olds, D. M. & Hirschman, K. B. The care span: The importance of transitional care in achieving health reform. *Health Aff. Proj. Hope* **30**, 746–754 (2011).
3. Konetzka, R. T., Spector, W. & Limcangco, M. R. Reducing hospitalizations from long-term care settings. *Med. Care Res. Rev. MCRR* **65**, 40–66 (2008).
4. Stauffer, B. D. *et al.* Effectiveness and Cost of a Transitional Care Program for Heart Failure: A Prospective Study With Concurrent Controls. *Arch. Intern. Med.* **171**, 1238–1243 (2011).
5. Kourtis, L. C., Regele, O. B., Wright, J. M. & Jones, G. B. Digital biomarkers for Alzheimer's disease: the mobile/wearable devices opportunity. *Npj Digit. Med.* **2** (2019).
6. Tomkins-Lane, C., Norden, J., Sinha, A., Hu, R. & Smuck, M. Digital biomarkers of spine and musculoskeletal disease from accelerometers: Defining phenotypes of free-living physical activity in knee osteoarthritis and lumbar spinal stenosis. *Spine J.* **19**, 15–23 (2019).
7. McGinnis, E. W. *et al.* Movements indicate threat response phases in children at-risk for anxiety. *IEEE J. Biomed. Health Inform.* **21**, 1460–1465 (2017).
8. McGinnis, R. S. *et al.* Rapid detection of internalizing diagnosis in young children enabled by wearable sensors and machine learning. *PLOS ONE* **14**, e0210267 (2019).
9. Bell, D. R. *et al.* Objectively Measured Physical Activity in Patients After Anterior Cruciate Ligament Reconstruction. *Am. J. Sports Med.* **45**, 1893–1900 (2017).
10. Sparaco, M., Lavorgna, L., Conforti, R., Tedeschi, G. & Bonavita, S. The Role of Wearable Devices in Multiple Sclerosis. *Mult. Scler. Int.* **2018**, 1–7 (2018).
11. Kernozek, T., Torry, M., Shelburne, K., Durall, C. J. & Willson, J. From the Gait Laboratory to the Rehabilitation Clinic: Translation of Motion Analysis and Modeling Data to Interventions That Impact Anterior Cruciate Ligament Loads in Gait and Drop Landing. *Crit. Rev. Biomed. Eng.* **41**, 243–258 (2013).
12. Takayanagi, N. *et al.* Relationship between Daily and In-laboratory Gait Speed among Healthy Community-dwelling Older Adults. *Sci. Rep.* **9**, 3496 (2019).
13. Blackburn, J. T., Pietrosimone, B., Harkey, M. S., Luc, B. A. & Pamukoff, D. N. Quadriceps Function and Gait Kinetics after Anterior Cruciate Ligament Reconstruction. *Med. Sci. Sports Exerc.* **48**, 1664–1670 (2016).
14. Hart, H. F. *et al.* Knee kinematics and joint moments during gait following anterior cruciate ligament reconstruction: a systematic review and meta-analysis. *Br. J. Sports Med.* **50**, 597–612 (2016).
15. Samà, A. *et al.* Estimating bradykinesia severity in Parkinson's disease by analysing gait through a waist-worn sensor. *Comput. Biol. Med.* **84**, 114–123 (2017).
16. Rodríguez-Molinero, A. *et al.* Validation of a Portable Device for Mapping Motor and Gait Disturbances in Parkinson's Disease. *JMIR MHealth UHealth* **3**, e9 (2015).
17. Rodríguez-Molinero, A. *et al.* A Kinematic Sensor and Algorithm to Detect Motor Fluctuations in Parkinson Disease: Validation Study Under Real Conditions of Use. *JMIR Rehabil. Assist. Technol.* **5**, e8 (2018).
18. Gurchiek, R. D. *et al.* Remote Gait Analysis Using Wearable Sensors Detects Asymmetric Gait Patterns in Patients Recovering from ACL Reconstruction. In *2019 IEEE International Conference on Body Sensor Networks (BSN)* (2019).

19. McGinnis, R. S. *et al.* Wearable Sensors Capture Differences in Muscle Activity and Gait Patterns During Daily Activity in Patients Recovering from ACL Reconstruction. In *2018 IEEE International Conference on Body Sensor Networks (BSN)* (2018).
20. Del Din, S., Godfrey, A., Galna, B., Lord, S. & Rochester, L. Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length. *J. Neuroengineering Rehabil.* **13**, 46 (2016).
21. Prajapati, S. K., Gage, W. H., Brooks, D., Black, S. E. & McIlroy, W. E. A Novel Approach to Ambulatory Monitoring: Investigation Into the Quantity and Control of Everyday Walking in Patients With Subacute Stroke. *Neurorehabil. Neural Repair* **25**, 6–14 (2011).
22. Lee, K. Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results. **8** (2018).
23. Haji Ghassemi, N. *et al.* Segmentation of Gait Sequences in Sensor-Based Movement Analysis: A Comparison of Methods in Parkinson's Disease. *Sensors* **18** (2018).
24. Gurchiek, R. D., McGinnis, R. S., Needle, A. R., McBride, J. M. & van Werkhoven, H. The use of a single inertial sensor to estimate 3-dimensional ground reaction force during accelerative running tasks. *J. Biomech.* **61**, 263–268 (2017).
25. McGinnis, R. S. *et al.* Accuracy of Femur Angles Estimated by IMUs During Clinical Procedures Used to Diagnose Femoroacetabular Impingement. *IEEE Trans. Biomed. Eng.* **62**, 1503–1513 (2015).
26. McGinnis, R. S., Hough, J. & Perkins, N. C. Accuracy of Wearable Sensors for Estimating Joint Reactions. *J. Comput. Nonlinear Dyn.* **12**, 041010–041010–10 (2017).
27. McGinnis, R. S. *et al.* A machine learning approach for gait speed estimation using skin-mounted wearable sensors: From healthy controls to individuals with multiple sclerosis. *PLOS ONE* **12**, e0178366 (2017).
28. Moon, Y. *et al.* Monitoring gait in multiple sclerosis with novel wearable motion sensors. *PLOS ONE* **12**, e0171346 (2017).
29. Seel, T., Raisch, J. & Schauer, T. IMU-Based Joint Angle Measurement for Gait Analysis. *Sensors* **14**, 6891–6909 (2014).
30. Mather, R. C. *et al.* Societal and Economic Impact of Anterior Cruciate Ligament Tears. *J. Bone Jt. Surg.-Am.* **95**, 1751–1759 (2013).
31. Carbone, A. & Rodeo, S. Review of current understanding of post-traumatic osteoarthritis resulting from sports injuries. *J. Orthop. Res.* **35**, 397–405 (2017).
32. Andriacchi, T. P. & Mündermann, A. The role of ambulatory mechanics in the initiation and progression of knee osteoarthritis. *Curr. Opin. Rheumatol.* **18**, 514–518 (2006).
33. Pietrosimone, B. *et al.* Walking gait asymmetries 6 months following anterior cruciate ligament reconstruction predict 12-month patient-reported outcomes. *J. Orthop. Res. Off. Publ. Orthop. Res. Soc.* **36**, 2932–2940 (2018).
34. Sigward, S. M., Lin, P. & Pratt, K. Knee loading asymmetries during gait and running in early rehabilitation following anterior cruciate ligament reconstruction: A longitudinal study. *Clin. Biomech.* **32**, 249–254 (2016).
35. Wellsandt, E. *et al.* Decreased Knee Joint Loading Associated With Early Knee Osteoarthritis After Anterior Cruciate Ligament Injury. *Am. J. Sports Med.* **44**, 143–151 (2016).
36. Devita, P. *et al.* Gait adaptations before and after anterior cruciate ligament reconstruction surgery. *Med. Sci. Sports Exerc.* **29**, 853–859 (1997).
37. Decker, M. J., Torry, M. R., Noonan, T. J., Sterett, W. I. & Steadman, J. R. Gait retraining after anterior cruciate ligament reconstruction. *Arch. Phys. Med. Rehabil.* **85**, 848–856 (2004).
38. Lisee, C. Quadriceps Strength and Volitional Activation After Anterior Cruciate Ligament Reconstruction: A Systematic Review and Meta-analysis. *SPORTS Health* **17** (2019).
39. Farrokhi, S., Voycheck, C. A., Gustafson, J. A., Fitzgerald, G. K. & Tashman, S. Knee joint contact mechanics during downhill gait and its relationship with varus/valgus motion and muscle strength in patients with knee osteoarthritis. *The Knee* **23**, 49–56 (2016).
40. Slater, L. V., Hart, J. M., Kelly, A. R. & Kuenze, C. M. Progressive Changes in Walking Kinematics and Kinetics After Anterior Cruciate Ligament Injury and Reconstruction: A Review and Meta-Analysis. *J. Athl. Train.* **52**, 847–860 (2017).
41. Lepley, A. S. *et al.* Quadriceps neural alterations in anterior cruciate ligament reconstructed patients: A 6-month longitudinal investigation: ACLr neural alterations. *Scand. J. Med. Sci. Sports* **25**, 828–839 (2015).
42. Palmieri-Smith, R. M. & Lepley, L. K. Quadriceps Strength Asymmetry After Anterior Cruciate Ligament Reconstruction Alters Knee Joint Biomechanics and Functional Performance at Time of Return to Activity. *Am. J. Sports Med.* **43**, 1662–1669 (2015).
43. Hobson, A. The Etiology of Persistent Quadriceps Weakness Following Anterior Cruciate Ligament Reconstruction. *JBJS J. Orthop. Physician Assist.* **6**, e24 (2018).
44. Pfeiffer, S. *et al.* Associations Between Slower Walking Speed and T1ρ Magnetic Resonance Imaging of Femoral Cartilage Following Anterior Cruciate Ligament Reconstruction. *Arthritis Care Res.* **70**, 1132–1140 (2018).
45. Sun, D., Fekete, G., Mei, Q. & Gu, Y. The effect of walking speed on the foot inter-segment kinematics, ground reaction forces and lower limb joint moments. *PeerJ* **6**, e5517 (2018).
46. Sigward, S. M., Chan, M.-S. M. & Lin, P. E. Characterizing knee loading asymmetry in individuals following anterior cruciate ligament reconstruction using inertial sensors. *Gait Posture* **49**, 114–119 (2016).
47. van der Straaten, R., De Baets, L., Jonkers, I. & Timmermans, A. Mobile assessment of the lower limb kinematics in healthy persons and in persons with degenerative knee disorders: A systematic review. *Gait Posture* **59**, 229–241 (2018).
48. McGinnis, R. S. *et al.* Skin mounted accelerometer system for measuring knee range of motion. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the 5298–5302* (IEEE, 2016).
49. Chapman, R. M., Moschetti, W. E. & Van Citters, D. W. Stance and swing phase knee flexion recover at different rates following total knee arthroplasty: An inertial measurement unit study. *J. Biomech.* **84**, 129–137 (2019).
50. Ren, L., Jones, R. K. & Howard, D. Whole body inverse dynamics over a complete gait cycle based only on measured kinematics. *J. Biomech.* **41**, 2750–2759 (2008).
51. Lloyd, D. G. & Besier, T. F. An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments *in vivo*. *J. Biomech.* **36**, 765–776 (2003).
52. Davies, D. L. & Bouldin, D. W. A Cluster Separation Measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1**, 224–227 (1979).
53. Cohen, J. The statistical power of abnormal-social psychological research: A review. *J. Abnorm. Soc. Psychol.* **65**, 145–153 (1962).

Acknowledgements

This work was supported in part by the National Institutes of Health under Grant R21AR069199.

Author contributions

Study conception and design: B.D.B., J.R.S., T.W.T., M.J.T. and R.S.M. Data collection: R.H.C. and R.D.G. Data analysis and interpretation: R.D.G. and R.S.M. All authors critically reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.D.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019