Human Movement Science 66 (2019) 335–346

Contents lists available at ScienceDirect

Human Movement Science

journal homepage: www.elsevier.com/locate/humov

Classification of gait muscle activation patterns according to knee injury history using a support vector machine approach

Maurice Mohr[a,b,*], Vinzenz von Tscharner[a], Carolyn A. Emery[a,b,c], Benno M. Nigg[a]

[a] *Human Performance Laboratory, Faculty of Kinesiology, University of Calgary, Alberta, Canada*
[b] *Sport Injury Prevention Research Centre, Faculty of Kinesiology, University of Calgary, Alberta, Canada*
[c] *The Alberta Children's Hospital Research Institute and McCaig Institute for Bone and Joint Health, Cumming School of Medicine, University of Calgary, Alberta, Canada*

ARTICLE INFO

Keywords:
Anterior cruciate ligament (ACL)
Biomechanics
Electromyography
Knee osteoarthritis
Principal component analysis
Wavelet transform

ABSTRACT

Abnormal muscle activation patterns during gait following knee injury that persist past the acute injury and rehabilitation phase (> three years) are not well characterized but may be related to post-traumatic knee osteoarthritis. The aim was to characterize the abnormal muscle activity from electromyograms of five leg muscles that were recorded during treadmill walking for young adults with and without a previous knee injury 3–12 years prior. The wavelet transformed and amplitude normalized electromyograms yielded intensity patterns that reflect the muscle activity of these muscles resolved in time and frequency. Patterns belonging to the affected or unaffected leg in previously injured participants and patterns belonging to a previously injured vs. uninjured participant were grouped and then classified using a principal component analysis followed by a support vector machine. A leave-one-out cross-validation was used to test the model significance and generalization. The results showed that trained classifiers could successfully recognize whether muscle activation patterns belonged to the affected or unaffected leg of previously injured individuals. Classification rates of 83% were obtained for all subjects, 100% for females only, indicating sex-specific knee injury effects. In contrast, it was not possible to discriminate between patterns belonging to the previously injured legs or dominant legs of control subjects. For females, the injured leg showed a stronger muscle activity for hamstring muscles and a lower activity for the vastus lateralis. In conclusion, systematic knee injury effects on the neuromuscular control of the knee during gait were present 3–12 years later.

1. Introduction

Following a significant knee injury, the neuromuscular control system shows multifaceted adaptations in order to maintain movement performance despite compromised knee joint stability, reduced muscle strength, and joint pain related to the injury (Hart, Pietrosimone, Hertel, & Ingersoll, 2010; Pietrosimone, McLeod, & Lepley, 2012). During gait in the first year after a rupture of the anterior cruciate ligament (ACL), characteristic adaptations in lower extremity muscle activation have been observed frequently including reduced quadriceps muscle activation ('quadriceps avoidance') and/or increased hamstring muscle activation, and/or increased quadriceps-hamstring co-contraction during early stance. These alterations may be aimed at avoiding excessive anterior translation of the tibia in individuals who are missing the passive stability and sensory feedback provided by the ACL (Berchuck, Andriacchi, Bach, & Reider, 1990; Kvist & Gillquist, 2001; Li et al., 1999). At the same time, these neuromuscular adaptations

* Corresponding author at: University of Calgary, KNB 219, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada.
E-mail address: mmaurice@ucalgary.ca (M. Mohr).

https://doi.org/10.1016/j.humov.2019.05.006
Received 25 September 2018; Received in revised form 14 May 2019; Accepted 15 May 2019
Available online 28 May 2019
0167-9457/

permanently alter the forces that are acting across the articulating joint surfaces of the knee (Victor, Labey, Wong, Innocenti, & Bellemans, 2010). Thus, abnormalities in lower extremity muscle activation have been suggested as possible risk factors for the slow development of post-traumatic osteoarthritis (PTOA) of the knee over the course of 10–15 years following previous joint trauma (Palmieri-Smith & Thomas, 2009; Pietrosimone et al., 2012; Roos, Herzog, Block, & Bennell, 2011). However, long-term adaptations in muscle activation patterns past the first few years after knee injury are not well characterized and vary widely between studies (Hall, Stevermer, & Gillette, 2015; Limbird, Shiavi, Frazer, & Borra, 1988; Lindström, Felländer-Tsai, Wredmark, & Henriksson, 2010; Roberts, Rash, Honaker, Wachowiak, & Shaw, 1999), which hinders an analysis of their relationship with knee osteoarthritis. For example, Roberts and colleagues observed abnormal timing of muscle activity for the quadriceps and hamstrings muscles but not for lower leg muscles during the stance phase of gait in ACL-deficient individuals at 47 months post-injury (Roberts et al., 1999). In contrast, Lindström and colleagues showed differences in on-offset timing of the tibialis anterior and gastrocnemius but no differences for the thigh muscles between ACL-deficient individuals and a control group at 35 months post-injury (Lindström et al., 2010). Furthermore, conflicting evidence exists for the persistence of a 'quadriceps avoidance' gait more than one year post-knee injury (Limbird et al., 1988; Roberts et al., 1999). One underlying reason for the disagreement may be a low sensitivity to detect differences in the timing of quadriceps and hamstrings muscle activity as reported by Lindström and colleagues. Furthermore, pre-selected muscle activity variables (e.g. on-offset timing or average amplitudes) at pre-selected time intervals (e.g. 5% gait cycle increments or first/second half of stance) in previous studies may not capture the existing abnormalities in neuromuscular control in a previously injured leg (Chau, 2001).

Therefore, the aim of this study was to develop a sensitive analytic approach that enables a holistic investigation of abnormal lower extremity muscle activation patterns during gait persistent past the acute knee injury and rehabilitation period (> 3 years post-injury). We suggest that the methodological approach to classify 'normal' and 'abnormal' muscular control during gait should attempt to represent the multi-muscle activation pattern of the lower extremities as a whole, taking into account the time- and frequency-dependent changes during the gait cycle as well as the interactions between muscles (Nigg & Herzog, 2007). Such patterns can be used as an input to machine learning classification algorithms such as a support vector machine, which are highly sensitive to detecting systematic differences between multi-dimensional data sets (Begg, Palaniswami, & Owen, 2005; Lai, Levinger, Begg, Gilleard, & Palaniswami, 2009). Furthermore, once such classifiers have been trained with sufficient data from individuals with and without previous knee injuries, they could be used to decide whether or not a previously injured individual still shows abnormal neuromuscular control features during gait (Begg & Kamruzzaman, 2005). Ultimately, pattern recognition and machine learning could enable an objective data-driven approach to explore the relationship between abnormal muscular control and post-traumatic knee osteoarthritis and guide rehabilitation.

von Tscharner and Valderrabano (2010) combined a principal component analysis of lower leg EMG time-frequency patterns with a support vector machine (SVM) analysis and could successfully classify whether patterns belonged to legs with or without ankle osteoarthritis (von Tscharner & Valderrabano, 2010). The classification rate of this approach, however, only reached about 70%. This may have been because males and females were combined in their analysis. It is known that males and females show sex-specific muscle activation features during gait (Di Nardo, Mengarelli, Maranesi, Burattini, & Fioretti, 2015; von Tscharner & Goepfert, 2003) and may respond differently to musculoskeletal pathology (Ko, Simonsick, Husson, & Ferrucci, 2011; Yamazaki, Muneta, Ju, & Sekiya, 2010), which may have hindered a higher classification rate according to ankle osteoarthritis. Therefore, classifiers that are specific to either male or female subjects may be more accurate in predicting lower extremity pathologies based on muscle activation features.

The objective of this study was to test the application of a support vector machine analysis to classify multi-muscle EMG patterns of the lower extremities during gait between limbs with a previous knee injury more than three years ago and uninjured limbs. The following hypotheses were tested:

Hypothesis 1: Individuals who sustained a previous intra-articular knee injury 3–12 years ago show systematic differences in average multi-muscle patterns during gait between the affected and not affected leg.

Hypothesis 2: The affected legs of individuals who sustained a previous intra-articular knee injury 3–12 years ago show systematic differences in average multi-muscle patterns during gait compared to multi-muscle patterns legs of individuals with no history of knee injury.

Hypothesis 3: The separation and classification of average multi-muscle patterns according to knee injury history is more accurate when conducting sex-specific analyses rather than a combined analysis of males and females.

2. Methods

2.1. Study design and participants

Participants of this study were a subset of an ongoing longitudinal 3-year historical cohort study [The Alberta Youth Prevention of Early OA Study (AB Pre-OA Study)] (Toomey et al., 2017; Whittaker, Toomey, Nettel-Aguirre, et al., 2018; Whittaker, Toomey, Woodhouse, et al., 2018; Whittaker, Woodhouse, Nettel-Aguirre, & Emery, 2015). The Pre-OA study includes 200 youth/young adults (aged 15–26); 100 participants who have sustained a sport-related intra-articular knee injury three to ten years ago and 100 participants with no history of knee injury, matched for age, sex, and sport. Recruitment strategies and inclusion/exclusion criteria have been described in detail in previous publications (Toomey et al., 2017; Whittaker, Toomey, Nettel-Aguirre, et al., 2018; Whittaker, Toomey, Woodhouse, et al., 2018; Whittaker et al., 2015). This study was carried out in accordance with the guidelines of the University of Calgary's Conjoint Health Research Ethics Board (#E-25075) and with the Declaration of Helsinki.

Table 1
Participant characteristics.

Variable	Not injured		Injured	
n (all [female])	33	[23]	28	[16]
Age at Testing (median [range])	25	[18 29]	24.5	[19 30]
Height in cm (median [range])	169	[158 192]	167	[157 196]
Weight in kg (median [range])	71.8	[54.2 108.5]	68.2	[49.3 109.5]
Months since injury (median [range])	n/a		100	[38 150]
Primary knee injury surgery	n/a		16	
ACL reconstruction	n/a		15	
Knee injury with meniscus involvement	n/a		15	
Contralateral knee injury	n/a		5	
Stride duration [s] (median, range)	1.06	[0.96 1.18]	1.04	[0.95 1.19]

For this sub-study, a convenience sample of 71 individuals (34 injured, 37 uninjured) of the Pre-OA study volunteered and gave written informed consent to participate in an analysis of their leg muscle activity during gait. Participant characteristics are summarized in Table 1. After visual analysis of the raw EMG data, ten participants had to be excluded due to defects of the recording equipment (n = 6) or insufficient signal-to-noise ratio for one or two muscles (n = 4). Due to these reasons 61 individuals (28 injured, 33 uninjured) were included in the EMG pattern analysis. For the within-subject comparisons in the injury group, five additional individuals had to be excluded because they had also previously injured their contralateral knee. The final sample size for this comparison was therefore n = 23 (INJ_X vs. INJ_C in Table 2). Fifteen individuals had sustained a full tear of the anterior cruciate ligament and all of them underwent surgical treatment. Other injuries included isolated meniscal injuries, medial or lateral collateral ligament injuries, and one fracture.

2.2. Gait measurements

Biomechanical and surface EMG measurements were conducted at the University of Calgary's Human Performance Laboratory. Locations for bipolar surface EMG sensors (Ag/AgCl, 10 mm diameter, 20 mm inter-electrode distance, Norotrode Myotronics-Noromed Inc., US) were identified on the vastus lateralis (VL), biceps femoris (BF), medial hamstrings (MH), gastrocnemius lateralis (GL), and gastrocnemius medialis (GM) of both legs according to standardized landmarks described in the SENIAM guidelines (Hermens, Freriks, & Merletti, 1999). Before sensor placement, the skin around the identified locations was shaved, lightly abraded, and cleaned with alcohol wipes to reduce skin impedance. After sensor placement, the EMG signal quality was visually inspected and validated during a series of test movements, i.e. squats for vastus lateralis, hamstring curls for hamstrings, and calf rises for the gastrocnemii. The acceleration of a 1D-accelerometers that was taped to the lateral aspect of the right and left heel and was used to detect the time of heel strike.

Bipolar, differential EMG signals and the accelerometer signal were recorded simultaneously at 2400 samples per second. The EMG signals were amplified by a factor of 1000 and bandpass-filtered between 10 and 500 Hz (Biovision, Wehrheim, Germany). The measuring system was grounded by connecting the ground electrode of the system to the right tibial tuberosity. The analog signals were digitized by a 12-bit A/D converter (National Instruments, Austin, TX).

After the sensor set-up, participants walked on a treadmill (Quinton Q55, Mortara Instrument Inc., Milwaukee, WI, USA) for a duration of two minutes at a speed of 4.5 km/h. The first minute of walking was considered a warm-up and familiarization period whereas the second minute was used for EMG data acquisition.

Table 2
SVM algorithm and feature selection output for all comparisons.

Comparison (G1, G2)	Participant Group	Sample size (G1, G2)	Critical CR** [%]	Selected feature set	Explained variance [%]	SR [%]	CR** [%]	FS
INJ_X,	All Subjects	23, 23	74	4,8,10	15	78	83	65
INJ_C*	Females	13, 13	85	2,3,9	34	100	100	100
INJ_X,	All Subjects	28, 33	66	no solution	–	–	–	–
CON_D†	Females	16, 23	69	no solution	–	–	–	–
CON_N,	All Subjects	33, 33	70	2,9,10	22	79	70	55
CON_D*	Females	23, 23	74	0,2,3	56	78	78	61

SR – separation rate, CR – classification rate, FS – feature selection factor.

* within-subject design;

† between-subject design.

** classification rates were significant if they exceeded the 'Critical classification rate' as determined by a binomial test with a 99% confidence level.

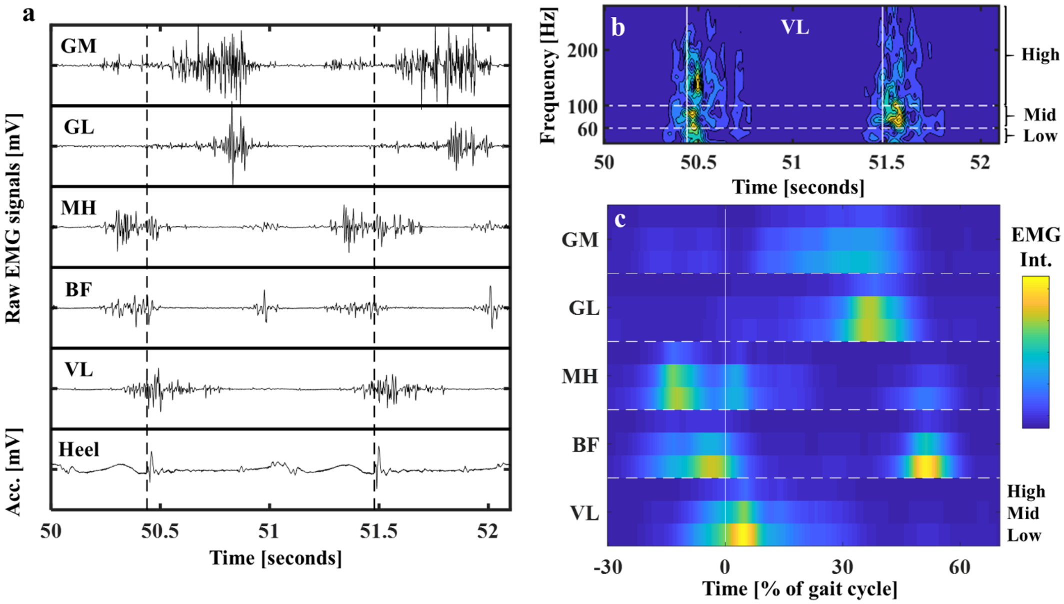

Fig. 1. Raw EMG signals and heel acceleration for two gait cycles. Dashed, vertical lines mark the heel strikes. All y-axes have the axis limits of −0.5 to 0.5 mV (a). Wavelet-transformed EMG intensity of vastus lateralis (VL) corresponding to the shown raw signal. Solid, vertical lines mark the heel strikes. Dashed, horizontal lines mark the separation of the EMG intensity into a low, mid, and high-frequency band (b). Final, averaged, and interpolated multi-muscle pattern. Solid, vertical line represents the heel strike. Dashed, horizontal lines separate the five muscles (c). Warm and cold colors represent higher and lower EMG intensity. All graphs are displayed for one leg of a representative participant.

2.3. EMG signal processing

All data processing steps were performed using a custom-written software in MATLAB 2017a (Mathworks, Natick, MA, USA). Fifty heel strikes (about 1 min of walking), thus 49 gait cycles were considered per subject for the analysis. The heel strikes were determined as the onset of high frequency oscillations in the heel acceleration signal (vertical dashed lines, Fig. 1a) (Meyer, Mohr, Falbriard, Nigg, & Nigg, 2018).

Step 1: The EMG wavelet transform

Raw EMG signals were transformed into time-frequency space using a wavelet analysis (Barandun, von Tscharner, Meuli-Simmen, Bowen, & Valderrabano, 2009; von Tscharner, 2000). Specifically, the raw EMG signals were convolved with 30 non-linearly scaled wavelets (scale = 1.6 in von Tscharner, 2000) spanning center frequencies between 1 and 500 Hz to extract the EMG power in 30 frequency bands as a function of time. The square root of the power yields the EMG intensity, which was used as a measure of muscle excitation. A typical wavelet intensity pattern of two steps is shown inf Fig. 1b and was used for further processing steps.

Step 2: Creation of low dimensional EMG intensity patterns

A time window of EMG intensities was selected starting 30% of gait cycle duration before the current heel strike and ending 70% of gait cycle duration after the heel strike. Within this window the EMG intensities were resampled to 250 time points which yielded a time resolution, which is shorter than the time resolution of the wavelets (von Tscharner, 2000). This procedure ensured that the main activation bursts during gait were not disrupted by the edges of the constructed patterns.

Based on a study of running, the EMG frequency content can be subdivided into three frequency bands (von Tscharner, Ullrich, Mohr, Comaduran Marquez, & Nigg, 2018; von Tscharner, Ullrich, Mohr, Marquez, & Nigg, 2018). The wavelet center frequencies were 32, 40.9, 50.9 Hz for the low frequency band, 62, 74.1, 87.4 Hz for the mid frequency band, and 101.6, 117, 133.4, 150.8, 169.3, 188.8, 209.4, 231, 253.6, 277.3 Hz for the high frequency band. Frequencies below 25 Hz were not considered as they may be contaminated by movement artifacts (Conforto, D'Alessio, & Pignatelli, 1999). EMG intensities at higher frequencies (> 300 Hz) showed negligible contributions (Fig. 1b) and were thus excluded from the analysis. The overall EMG intensity was obtained by summing the intensities across all 3 frequency bands and 250 time points and averaging across gait cycles. The patterns were normalized by dividing the EMG intensities by the overall EMG intensity. The normalization was necessary to make the EMG intensities comparable in magnitude between muscles and legs. Normalized EMG intensities were then averaged across the 49 gait cycles, yielding one EMG intensity pattern for each of the five muscles, representing 3 frequency bands and 250 time points. These five patterns were stacked on top of one another and can be represented as a multi-muscle pattern (MMP) shown in Fig. 1c. Thus, there are 122 MMPs, one for each leg of the 61 participants that were used for the pattern recognition analysis.

Each MMP was reshaped into a one-dimensional vector and appended row-wise into a matrix M. Each vector was normalized to a length of 1 so that the patterns of each leg and subject contributed equally to the analysis. M consisted of 122 rows and 3750 columns with odd rows representing right leg EMG patterns and even rows representing left leg EMG patterns. Next, M was used as an input to a principal component analysis to further reduce the dimensionality.

Step 3: Principal component analysis (PCA)

A PCA was applied to the data matrix *M* (von Tscharner et al., 2018). In summary, the mean of *M* across observations (rows) was subtracted and the eigenvectors (PC-vectors) and eigenvalues of the covariance matrix of *M* were computed. The number of PC-vectors that explained 70% of the variance in *M* and the residual mean vector, referenced as PC-vector#0, formed the base of the lower dimensional vector space. The residual mean was added because the discrimination may include the parts of the signal that are not explained by the PCs but are captured in the remaining part, which is represented by the residual mean (see von Tscharner et al., 2018 for a description of the residual mean vector). Because the PC-vectors represent points of the MMPs that change in a correlated way, the PC-vectors were called features (F0 = PC-vector#0, F1 = PC-vector#1, etc.). The number of features is much smaller than the points in the MMPs while containing most of their variance, which represents a successful reduction of dimensionality via the PCA. The PC-weights indicate the fraction with which each feature contributes to the MMP of a given observation (leg of a subject) and were obtained by projecting the corresponding row vector in *M* onto the PC-vectors. The weights are arranged in a new matrix *W* with 122 rows that represent the observations in *M*. The columns of *W* contain the weights, one for each of the 11 features. Thus, the first column corresponds to weights on F0, the second column to weights on F1, etc. The whitened weights of a feature (w_f) for the 122 observations were computed by dividing the weights by their standard deviation.

The matrix *W* with 122 whitened weights (rows) and 11 features (columns) was used for SVM comparisons between different conditions. Each previously injured participant has one injured leg (INJ_X) and one contralateral leg (INJ_C). Each uninjured participant has one dominant leg (CON_D) and one non-dominant leg (CON_N). The comparisons investigated in this study are shown in Table 2.

Step 4: SVM analyses

In Hypothesis 1, the assumption is that MMPs are specific to each person whereas there is a systematic side-to-side difference in MMPs that is not subject-specific as a result of the previous knee injury. To isolate side-to-side differences and reduce the influence of between-subject variation, the mean of the whitened weights across the right and left leg was subtracted for each subject so that each pair of weights was centered around zero. The centered weights were used as the input to the SVM when investigating within-subject comparisons (INJ_X vs. INJ_C and CON_N vs. CON_D, Table 2). If side-to-side differences in MMPs are randomly distributed (Null Hypothesis 1), there will be no successful SVM classification of the two conditions.

In Hypothesis 2, the assumption is that a previous knee injury results in systematic differences in MMPs between legs of injured und uninjured individuals despite between subject-variation. In this case, the whitened weights were centered by subtracting their overall mean across all subjects included in the comparison and then used as the input to the SVM (INJ_X vs. CON_D, Table 2).

Linear-kernel support vector machines (SVMs) were used to test the hypothesis that the centered weights are systematically different between legs. The SVM is a supervised learning algorithm, which determines a separating hyperplane that optimizes the margin between the two groups of the input data based on known group labels (e.g. injured vs. not injured) (Vapnik, 2000). The SVM was trained by using the *fitcsvm* function of the MATLAB 2017a Statistics and Machine Learning Toolbox with the default setting but without scaling of the input data as the whitening process was done manually. If observations of group 1 fall on one side of the optimized hyperplane and observations of group 2 on the other side, the SVM separation rate (SR) is 100%.

To assess the statistical significance and generalization of the trained SVM models, a leave-one-out (LOO) cross-validation method was used (von Tscharner & Valderrabano, 2010). During the LOO procedure, each individual of the analyzed comparison was once excluded from the data set to train the SVM. Each LOO iteration included the re-computation of the PC-vectors and corresponding weights (*Step 3* above) to ensure that the SVM training data does not include any information about the left-out observation. The MMPs of the left-out subject were then projected onto the re-computed LOO PC-vectors to determine the weights and use them as an input to the MATLB *predict* function to predict the group assignment. The number of correctly predicted test observations divided by the number of participants yields the generalized classification rate (CR).

Each assignment of a left-out observation can either be correct or wrong. Thus, a binomial test (MATLAB function *binoinv* with a probability of success $P = 0.5$) was used to assess whether the CR is statistically higher than chance at a confidence level of 99%.

Step 5: Feature selection

When using machine learning algorithms, a high risk of overfitting is typically present when the training data set contains a high number of features relative to the number of observations. To lower the risk of overfitting, we limited the number of our features using a feature selection method. SVM analyses have demonstrated that besides 3–5 important gait features that enable a successful classification, further gait features are often redundant and do not provide additional discriminatory information (Begg et al., 2005; Begg & Kamruzzaman, 2005). It was therefore decided that three features are sufficient for the classification. This bears the further advantages that 1) the SVM model can still be visually inspected in a 3D-plot, 2) each combination of three input features still explained a minimum of 10% of the original variance (see Results section). 3) it reduces the complexity of the analysis.

The selection of the three features was done using a trial and error process. There are 165 possible ways to select three features from 11 available features. Therefore, the product of separation rate and classification rate, which will be called the 'feature selection factor' (FS) was determined for all 165 feature combinations. Non-significant classification rates were set to zero, thus yielding FS = 0. The maximum FS was used to select the final input feature set. This feature selection approach is computationally costlier compared to the 'Hill-Climbing' approach described by Begg and colleagues (2005) who build up their feature set by iteratively adding features with discriminatory information in a feed-forward fashion. However, we noticed that the inclusion of one feature in the feature set may influence the discriminatory ability of another feature. Therefore, albeit computationally more demanding, this approach avoided erroneously excluding a feature with good discriminatory ability.

A flow chart of the algorithm is provided in Fig. A.1 (Appendix A).

Step 6: Discriminatory multi-muscle patterns

Each SVM comparison resulted in three features that best solved the classification problem. To interpret the discriminatory multi-muscle patterns (DMP) encoded in these features, the following analytical steps were taken. For a given comparison and the selected best set of features f = 1,2,3, the linear combinations of features F_f and corresponding input weights w_f were determined for each observation k considered in the comparison according to Eq. (1):

$$Z(k) = \sum_{f=1}^{3} w(k)'_f \cdot F(k)_f \tag{1}$$

Z(k) are reconstructed patterns with the structure of the original MMPs but contain only the muscle activation features that enabled the classification between two groups. Then, the DMP was determined as the average difference in Z between the individuals belonging to the first group, e.g. injured leg, and the individuals belonging to the second group, e.g. contralateral leg. DMPs can be displayed similarly to the MMP in Fig. 1c and indicate the location and direction of differences in muscle activation that were identified between the two groups.

2.4. Data sharing

Data underlying the analysis described above has been uploaded to the Mendeley public data repository (Mohr, von Tscharner, Emery, & Nigg, 2018). These data include 1) raw EMG and heel strike data, 2) the matrix M containing the wavelet-transformed, interpolated, and amplitude-normalized multi-muscle patterns, the principal component vectors PC#0–10 (=F0-F10), the matrix W containing the weights *before* whitening, and anonymized labels for the grouping of patterns according to injury history, leg dominance, contralateral knee injury, and sex.

3. Results

3.1. Principal components

Ten PC-vectors explained 70% of the variance contained in the original 122 input MMPs. The residual mean and the PC-vectors were called features 0–10 and were ordered according to their relative, explained variance. The residual mean (F0), PC-vectors (F1-

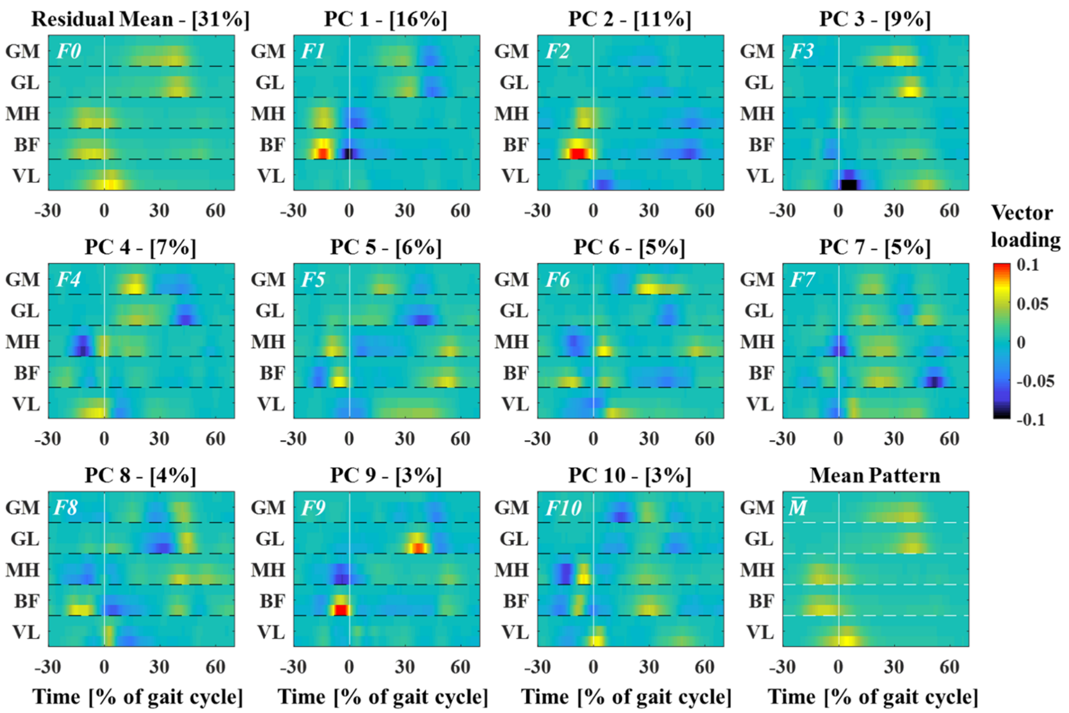

Fig. 2. Output of the PCA. The residual mean and the first ten principal component vectors are displayed in order of their relative explained variance, shown in brackets. Their contribution to the original MMP input patterns is indicated by their color-coded PC-vector loading. These 11 patterns were called features F0-F10 and used for machine learning analyses. The bottom right image shows the average MMP $\bar{M}$ obtained from the mean of all input MMPs (2 legs for 61 subjects = 122 MMPs).

F10), and the mean across all 122 input MMPs ($\bar{M}$) are displayed as heat maps in Fig. 2.

The mean MMP shows the characteristic activity bursts of the lower extremity muscles during gait, e.g. the main vastus activation burst shortly after heel strike, the main hamstrings activation burst right before and during heel strike, and the main gastrocnemius activation burst during the second half of the stance phase. The averaging procedures to calculate $\bar{M}$ removed most of the high-frequency components and EMG signal fine structure that were still visible in the wavelet patterns of individual steps in Fig. 1b. The residual mean vector largely resembles the mean pattern and is of importance for the full reconstruction of original MMPs from the reduced set of features. An example for the interpretation of features in Fig. 2 is as follows. If the weight on feature#1 is positive then the person exhibits higher hamstrings activation before heel strike (yellow area in Fig. 2, PC1) and lower hamstrings activation during heel strike (blue area in Fig. 2, PC1). This deviation in hamstrings activity before and during heel strike combined with a trade-off in the timing of gastrocnemius activity explained 16% of the total variance. Features 2 and 3 start to describe EMG intensities of the vastus lateralis as well as features that are related to muscle activity events outside of the main activation bursts, e.g. hamstrings and vastus lateralis activity during the late stance phase. Higher numbered features explain a smaller amount of variance, their interpretation becomes more complex, and they are less relevant. However, any set of three muscle activation features selected from the 11 features explained a minimum of 10% of the total variance. The whitened weights corresponding to the 11 features were used as the input for the SVM classification analysis.

3.2. SVM and feature selection

Table 2 summarizes the separation rate (SR), classification rate (CR), and feature selection factor (FS) of the selected best feature sets and their explained variance for all comparisons of this study. Classification rates above the critical classification threshold were obtained when comparing the injured and uninjured legs of the injury group (INJ_X vs. INJ_C) as well as the dominant and non-dominant legs of controls (CON_N vs. CON_D) but not when comparing individuals with and without a previous knee injury (INJ_X vs. CON_D). Therefore, feature sets that classify between a previously injured leg and the dominant leg of a control subject could not be obtained. The feature sets that led to the highest product of separation and classification rate (feature selection factor, FS) were different for all four successful classifications. The variance explained by selected features was generally higher when considering only female individuals (34% for INJ_X vs. INJ_C, 56% for CON_N vs. CON_D) compared to all individuals (15% for INJ_X vs. INJ_C, 22% for CON_N vs. CON_D). The classification rates and feature selection factors were higher for the comparison of injured vs. uninjured legs compared to non-dominant vs. dominant legs (Table 2). For 100% of female individuals with a knee injury history, the injured and uninjured legs could be separated and classified successfully by the SVM algorithm.

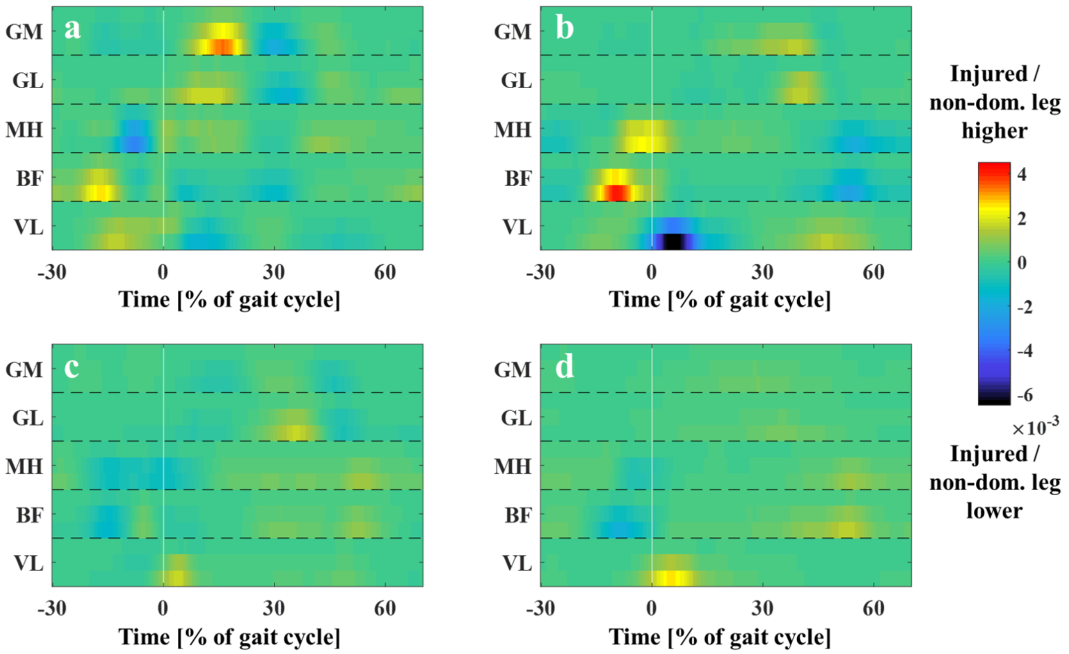

Fig. 3. Discriminatory multi-muscle patterns reconstructed from the combination of three features that enable the best SVM separation and classification of multi-muscle EMG patterns between a) INJ_X vs. INJ_C (n = 23), b) INJ_X vs. INJ_C, female only (n = 13), c) CON_N vs. CON_D (n = 33), d) CON_N vs. CON_D, female only (n = 23). The color scale is constant in all images and was selected based on the highest and lowest values across all four comparisons.

3.3. Discriminatory multi-muscle patterns

Fig. 3 shows the discriminatory multi-muscle patterns (DMPs) that enabled to separate and classify the original MMPs between injured and uninjured legs of individuals with a previous knee injury (Fig. 3a,b) and non-dominant and dominant legs of healthy controls (Fig. 3c,d). The discrimination was most clear when comparing the previously injured and contralateral legs in female individuals. Compared to the contralateral leg, muscle activation of the previously injured leg was mostly characterized by a lower vastus lateralis activation during the first 10% of stance combined with a stronger biceps femoris activation before heel strike and stronger medial hamstrings activation during heel strike (Fig. 3b). When considering both male and female injured individuals, differences in muscle activation were less clear but included a lower medial hamstring and higher biceps femoris activity before heel strike and a difference in timing of medial gastrocnemius activity for the injured compared to contralateral leg (Fig. 3a). DMPs were not pronounced for the comparison of the non-dominant and dominant legs of healthy controls. For the whole group as well as for female subjects separately, the non-dominant leg showed slightly higher vastus lateralis activation post-heel strike and lower hamstring activity pre-heel strike compared to the dominant leg (Fig. 3c,d).

4. Discussion

In this study, a pattern analysis of lower extremity muscle activation patterns during gait was applied and combined with a support vector machine analysis to distinguish patterns from individuals with and without a previous knee injury 3–12 years ago. The aim of this analysis was to test whether abnormal muscle activation features still exist in individuals more than three years following intra-articular knee injury and whether these features can be used to predict if a newly measured pattern belongs to an injured or uninjured leg. The findings support our first hypothesis that muscle activation features can be used to predict whether an activation pattern recorded from a previously injured subject belongs to the affected or unaffected leg. In contrast, muscle activation features could not be used to successfully classify whether an activation pattern belonged to a previously injured individual versus a control subject with no knee injury history and thus falsifying our second hypothesis. Finally, and in support of our third hypothesis, the classification according to knee injury history was more accurate when conducting a sex-specific analysis of females compared to an overall analysis of males and females.

4.1. Injured leg vs. contralateral leg

When comparing muscle activation features during gait between the previously injured leg and contralateral leg of individuals with a knee injury 3–12 years ago, the pattern recognition approach presented in this study classified over 80% of the injured and contralateral legs correctly. This indicates that the previously injured leg still shows systematic differences in muscle activity surrounding the knee joint more than three years post-injury corroborating earlier results (Hall et al., 2015; Limbird et al., 1988; Lindström et al., 2010; Roberts et al., 1999). The discriminatory pattern consisted of many subtle differences in activation distributed over all measured muscles at different time intervals, since they only explained a small portion of the variance contained in the original MMPs. In comparison, the separate analysis of females resulted in a higher classification rate of 100%, more explained variance by the selected features, and a different, more obvious discriminatory multi-muscle pattern. In support of previous reports (Yamazaki et al., 2010), this indicates that the neuromuscular adaptations observed for female individuals were sex-specific and were not shared by all male individuals. In turn, the training of a linear classifier for the assignment of lower extremity muscle activation patterns to the injured vs. uninjured leg was less successful when both sexes were combined in the analysis. Consequently, the discriminatory muscle activation features presented in Fig. 3a are likely a mix of features that are characteristic for male and female gait adaptations and should thus not be generalized. Instead, these findings suggest that future analyses should always be designed to examine sex-specific gait patterns following a knee injury.

The specific neuromuscular differences that were observed for previously injured female individuals were characterized by reduced quadriceps (VL) activity during early stance and an increase in hamstring activity before and during heel strike, thus not at the same time in the gait cycle. In contrast, a simultaneous increase in quadriceps and hamstring activation, i.e. co-contraction, was not observed. Thus, we can confirm that in female individuals with a previous knee injury, ‘quadriceps avoidance’ and ‘increased hamstrings activity’ but not ‘increased quadriceps-hamstrings co-contraction’ persist more than 3–12 years following knee injury (Beard, Soundarapandian, O'connor, & Dodd, 1996; Bulgheroni, Bulgheroni, Andrini, Guffanti, & Castelli, 1997; Bulgheroni, Bulgheroni, Andrini, Guffanti, & Giughello, 1997; Ferber, Osternig, Woollacott, Wasielewski, & Lee, 2002; Hall et al., 2015; Hurd & Snyder-Mackler, 2007; Knoll, Kiss, & Kocsis, 2004; Limbird et al., 1988). The above result was obtained for a heterogeneous sample that included individuals with varying types of knee injuries and time since injury. This may suggest that ‘quadriceps avoidance’ and ‘increased hamstrings activity’ represent neuromuscular adaptations that are not specific to ACL reconstructions but can also result from other ligamentous injuries and/or traumas to the knee joint. The sample size of this study, however, did not allow sub-analyses of homogeneous injury groups and future studies are needed to test the above suggestion.

Since this study used gait cycle-averaged, amplitude-normalized MMPs as the input to the pattern recognition and machine learning analysis, the identified differences in muscle activity have to be interpreted from a probabilistic point of view (von Tscharner

& Valderrabano, 2010). For example, increased biceps femoris EMG intensity pre-heel strike on the injured side could represent 1) a more consistent biceps femoris EMG intensity peak at this time point across gait cycles, and/or 2) higher biceps femoris EMG intensity relative to the average intensity during walking. The first scenario may result from a less variable neuromuscular control program for the biceps femoris whereas the second scenario may result from an increased number of recruited motor units or a more synchronized activation of multiple motor units of the biceps femoris (Asmussen, Von Tscharner, & Nigg, 2018; De Luca, 1997). This analysis, however, could not discern the neurophysiological origin of differences in EMG intensity, which provides targets for future studies. Regardless, changes in quadriceps and hamstring muscle activity most likely alter tibio-femoral kinematics and articular contact forces, which may facilitate the degradation of knee joint tissues and increase the risk for future PTOA (Andriacchi, Briant, Bevill, & Koo, 2006; Wellsandt et al., 2016). Previous analyses of previously injured individuals in this cohort have reported reduced self-reported knee joint health and function as well as a higher incidence of MRI-defined knee OA, which is consistent with a trajectory towards clinical knee OA (Whittaker et al., 2015; Whittaker, Toomey, Nettel-Aguirre, et al., 2018; Whittaker, Toomey, Woodhouse, et al., 2018). The methodological approach presented here may be a useful tool to assist the monitoring of progress in neuromuscular rehabilitation following a knee injury or to assess the risk of developing PTOA. Specifically, a classification model trained by a sufficiently large data set ($n > 100$) could be used to decide if the previously injured leg still shows abnormal muscle activation patterns during gait in comparison to the uninjured leg. The further assessment of PTOA risk via this approach, however, necessitates that the relationship between sEMG-based leg muscle activation patterns, knee mechanics, and joint tissue adaptation is better understood.

When comparing the MMPs between the non-dominant and dominant legs of individuals with no knee injury history, the feature based-SVM approach also resulted in successful classifications regarding leg dominance. This was not necessarily surprising since limb dominance has been shown to systematically influence muscle activity of the lower extremities during gait (Õunpuu & Winter, 1989). Even though scientific debate remains related to specific biomechanical functions of the dominant and non-dominant legs during gait (Sadeghi, Allard, Prince, & Labelle, 2000), the SVM discrimination between the two in the current study helps to put the knee injury analysis into perspective. For comparisons involving all subjects or female subjects separately, the classification and separation rates were lower when classifying according to leg dominance compared to knee injury history. Similarly, the discriminatory patterns in Fig. 3 related to leg dominance demonstrated less pronounced average differences in muscle activation features compared to patterns related to knee injury history. Therefore, systematic side-to-side differences in muscle activation between the injured and uninjured leg of individuals with a knee injury 3–12 years ago, specifically females, are not observed to the same degree in individuals with no knee injury history, thus emphasizing the relevance of these neuromuscular differences in injured subjects.

4.2. Injured leg vs. healthy control

In contrast to the successful within-subject classifications of injured and uninjured legs, no significant classification rates were obtained when comparing the muscle activation patterns between the previously injured leg with the dominant leg of healthy controls. This was true for analyses including all subjects and only female subjects and is in contrast to other previous machine learning analyses, which resulted in successful classifications of kinematic or electromyographic gait patterns between individuals with and without ankle osteoarthritis (von Tscharner & Valderrabano, 2010), idiopathic hip osteoarthritis or rheumatoid arthritis (Nair, French, Laroche, & Thomas, 2010), and patellofemoral pain (Lai et al., 2009). This suggests that the natural between-subject variation in the investigated muscle activation features among the previously injured and uninjured individuals was larger than any potential systematic variation due to knee injury history, thus prohibiting the training of a successful classifier. The discrepancy to former studies likely results from the fact that these authors investigated musculoskeletal conditions such as end-stage osteoarthritis, which lead to much more dramatic effects on neuromuscular control during movement that are larger than natural between-subject variation and can thus be detected by machine learning analyses. A second methodological reason for the successful classifiers in previous studies may be that non-linear machine learning techniques such as neural networks (Nair et al., 2010) or SVMs with non-linear kernel functions (von Tscharner & Valderrabano, 2010) were applied. Such techniques can improve the classification rates between two classes but the exact features responsible for the classification cannot easily be reconstructed (Stirling, von Tscharner, Kugler, & Nigg, 2011; von Tscharner & Valderrabano, 2010). While such methods could likely improve the classification rates in this study, it was our intention to keep the machine learning procedure as simple as possible so that identified differences can still be displayed and potentially be used by clinicians to guide the rehabilitation process. The findings of this study suggest that either systematic differences due to a previous knee injury 3–12 years ago were too small to be resolved by the applied technique or long-lasting systematic effects of a previous knee injury on the neuromuscular system are resolved in this young population more than three years post-injury.

4.3. Methodological considerations

In this study, the decision was made to only include the first 10 PCs to explain 70% of the original variance. Furthermore, the fine-structure of the EMG signal particularly in the high-frequency band was largely reduced due to the averaging across gait cycles.

Finally, we set the significance level for the assessment of classification rates to 1% to lower the risk of type 2 error when testing 165 feature sets. These three steps represent a conservative approach and were aimed at 1) avoiding misleading conclusions due to overfitting of the data and 2) ensuring that identified feature sets still explained a minimum of 10% of the original variance. Despite the conservative approach, high classification rates of 80–100% were achieved, which emphasizes the sensitivity of the applied techniques. With the present sample size, this study already belongs to the largest EMG investigations during gait following knee injuries (Shanbehzadeh, Bandpei, & Ehsani, 2017). Nevertheless, larger sample sizes would enable a direct comparison of sex-specific discriminatory multi-muscle patterns and inform the most important aspects of neuromuscular rehabilitation in males and females following a knee injury.

Other studies that focus on a more homogeneous injury history than this study, e.g. isolated ACL rupture with hamstring tendon graft reconstruction may observe more systematic injury dependent differences. Furthermore, the findings of this study may not be generalizable to older, more sedentary populations. A previous analysis of this PrE-OA cohort has shown that despite lower physical activity levels compared to controls, the previously injured sample is highly physically active with over 87% participation in sports at the time of testing (Toomey et al., 2017). Participation in physical activity may have helped this population to resolve abnormal neuromuscular control patterns more quickly. Nevertheless, it may be the more physically active populations for which abnormal neuromuscular control following knee injuries represents a more relevant problem in the context of risk for re-injury or post-traumatic osteoarthritis.

5. Conclusions

Lower extremity muscle activation patterns during gait from individuals with and without a previous knee injury 3–12 years ago were used to train linear SVM classifiers. Systematic side-to-side differences in neuromuscular control still existed more than three years post-injury and thus enabled a successful recognition of affected and unaffected legs. Differences in patterns between injured and uninjured legs in female individuals were characterized by increased hamstring activity before heel strike and reduced vastus lateralis activity post-heel strike. Male individuals did not show the same neuromuscular adaptations to a knee injury and thus future studies should always conduct sex-specific analysis of knee injury effects on gait patterns. The presented machine learning approach may be used to guide post-injury rehabilitation and inform the presence or absence of abnormal neuromuscular patterns in the injured compared to the contralateral leg. In contrast, the trained classifiers were not able to successfully discriminate between muscle activation patterns that belonged to a previously injured leg vs. the dominant leg of healthy controls. As compared to controls, neuromuscular gait adaptations to a previous knee injury are either resolved or not systematic at more than three years post-knee injury.

Acknowledgements

The authors would like to acknowledge the assistance of research coordinators Gabriella Nasuti, Jamie Rishaug and Lisa Loos, and numerous PrE-OA study team members (including Kristin Lorenzen and Sagar Grewal (data collection support), Clodagh Toomey and Jackie Whittaker (personal communication) and research assistants/students, as well as the participants).

Funding statement

The Alberta PrE-OA cohort is funded by the Canadian Institutes of Health Research (MOP 133597), the Alberta Osteoarthritis Team supported by Alberta Innovates and the Alberta Children's Hospital Research Institute Chair in Pediatric Rehabilitation (Alberta Children's Hospital Foundation). MM holds an Alberta Innovates Graduate Studentship and previously a Izaak Walton Killam Pre-Doctoral Scholarship for this doctoral research related to long-term consequences of knee joint injuries on lower extremity muscle activation.

Declaration of Competing Interest

All other authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. The sponsors had no involvement with respect to design, collection or data, analyses, interpretation writing or submission. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

Appendix

See Fig. A.1.

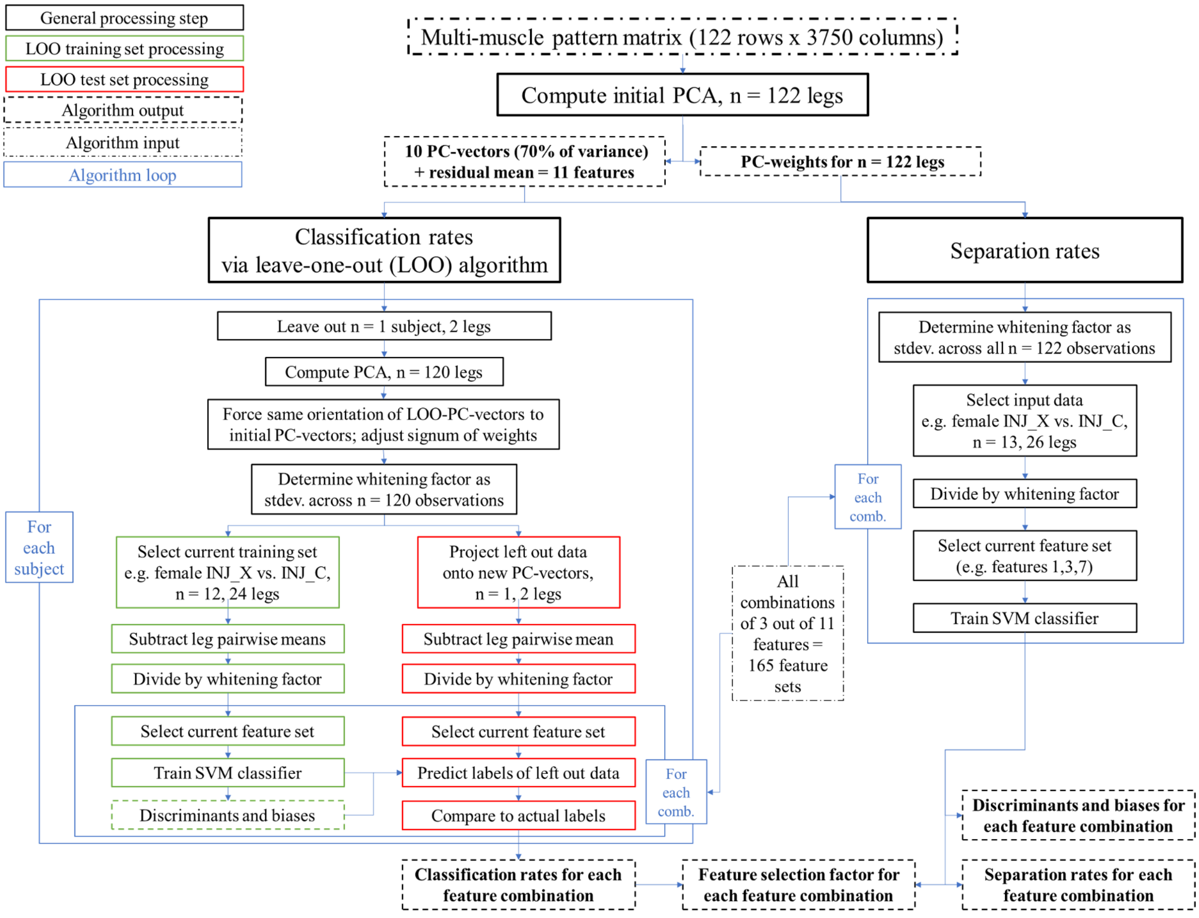

Fig. A.1. Flow chart of SVM algorithm.

References

Andriacchi, T., Briant, P., Bevill, S., & Koo, S. (2006). Rotational changes at the knee after ACL injury cause cartilage thinning. *Clinical Orthopaedics & Related Research*, 39–44.

Asmussen, M. J., Von Tscharner, V., & Nigg, B. (2018). Motor unit action potential clustering – Theoretical consideration for muscle activation during a motor task. *Frontiers in Human Neuroscience, 12*. https://doi.org/10.3389/fnhum.2018.00015.

Barandun, M., von Tscharner, V., Meuli-Simmen, C., Bowen, V., & Valderrabano, V. (2009). Frequency and conduction velocity analysis of the abductor pollices brevis muscle during early fatigue. *Journal of Electromyography and Kinesiology, 19*(1), 65–74.

Beard, D. J., Soundarapandian, R. S., O'connor, J. J., & Dodd, C. A. F. (1996). Gait and electromyographic analysis of anterior cruciate ligament deficient subjects. *Gait & Posture, 4*(2), 83–88.

Begg, R. K., Palaniswami, M., & Owen, B. (2005). Support vector machines for automated gait classification. *IEEE Transactions on Biomedical Engineering, 52*(5), 828–838. https://doi.org/10.1109/TBME.2005.845241.

Begg, R. K., & Kamruzzaman, J. (2005). A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data. *Journal of Biomechanics, 38*(3), 401–408. https://doi.org/10.1016/j.jbiomech.2004.05.002.

Berchuck, M., Andriacchi, T. P., Bach, B. R., & Reider, B. (1990). Gait adaptations by patients who have a deficient anterior cruciate ligament. *Journal of Bone and Joint Surgery-American, 72*(6), 871–877.

Bulgheroni, P., Bulgheroni, M. V., Andrini, L., Guffanti, P., & Castelli, C. (1997). Walking in anterior cruciate ligament injuries. *The Knee, 4*(3), 159–165. https://doi.org/10.1016/S0968-0160(97)00259-7.

Bulgheroni, P., Bulgheroni, M. V., Andrini, L., Guffanti, P., & Giughello, A. (1997). Gait patterns after anterior cruciate ligament reconstruction. *Knee Surgery, Sports Traumatology, Arthroscopy, 5*(1), 14–21. https://doi.org/10.1007/s001670050018.

Chau, T. (2001). A review of analytical techniques for gait data. Part 1: Fuzzy, statistical and fractal methods. *Gait & Posture, 13*(1), 49–66. https://doi.org/10.1016/S0966-6362(00)00094-1.

Conforto, S., D'Alessio, T., & Pignatelli, S. (1999). Optimal rejection of movement artefacts from myoelectric signals by means of a wavelet filtering procedure. *Journal of Electromyography and Kinesiology, 9*(1), 47–57. https://doi.org/10.1016/S1050-6411(98)00023-6.

De Luca, C. J. (1997). The use of surface electromyography in biomechanics. *Journal of Applied Biomechanics, 13*(2), 135.

Di Nardo, F., Mengarelli, A., Maranesi, E., Burattini, L., & Fioretti, S. (2015). Gender differences in the myoelectric activity of lower limb muscles in young healthy subjects during walking. *Biomedical Signal Processing and Control, 19*, 14–22. https://doi.org/10.1016/j.bspc.2015.03.006.

Ferber, R., Osternig, L., Woollacott, M., Wasielewski, N., & Lee, J. (2002). Gait mechanics in chronic ACL deficiency and subsequent repair. *Clinical Biomechanics, 17*(4), 274–285.

Hall, M., Stevermer, C. A., & Gillette, J. C. (2015). Muscle activity amplitudes and co-contraction during stair ambulation following anterior cruciate ligament reconstruction. *Journal of Electromyography and Kinesiology, 25*(2), 298–304. https://doi.org/10.1016/j.jelekin.2015.01.007.

Hart, J. M., Pietrosimone, B., Hertel, J., & Ingersoll, C. D. (2010). Quadriceps activation following knee injuries: A systematic review. *Journal of Athletic Training, 45*(1), 87–97. https://doi.org/10.4085/1062-6050-45.1.87.

Hermens, J., Freriks, B., & Merletti, R. (1999). *SENIAM: European recommendations for surface electromyography* (2nd ed.). Netherlands: Roessingh Research and Development.

Hurd, W., & Snyder-Mackler, L. (2007). Knee instability after acute ACL rupture affects movement patterns during the mid-stance phase of gait. *Journal of Orthopaedic Research, 25*(10), 1369–1377.

Knoll, Z., Kiss, R., & Kocsis, L. (2004). Gait adaptation in ACL deficient patients before and after anterior cruciate ligament reconstruction surgery. *Journal of Electromyography, 14*(3), 287–294.

Ko, S., Simonsick, E. M., Husson, L. M., & Ferrucci, L. (2011). Sex-specific gait patterns of older adults with knee osteoarthritis: Results from the baltimore longitudinal study of aging. *Current Gerontology and Geriatrics Research*. https://doi.org/10.1155/2011/175763.

Kvist, J., & Gillquist, J. (2001). Anterior positioning of tibia during motion after anterior cruciate ligament injury. *Medicine & Science in Sports & Exercise, 33*(7), 1063–1072.

Lai, D. T. H., Levinger, P., Begg, R. K., Gilleard, W. L., & Palaniswami, M. (2009). Automatic recognition of gait patterns exhibiting patellofemoral pain syndrome using a support vector machine approach. *IEEE Transactions on Information Technology in Biomedicine, 13*(5), 810–817. https://doi.org/10.1109/TITB.2009.2022927.

Li, G., Rudy, T. W., Sakane, M., Kanamori, A., Ma, C. B., & Woo, S. L.-Y. (1999). The importance of quadriceps and hamstring muscle loading on knee kinematics and in-situ forces in the ACL. *Journal of Biomechanics, 32*(4), 395–400. https://doi.org/10.1016/S0021-9290(98)00181-X.

Limbird, T. J., Shiavi, R., Frazer, M., & Borra, H. (1988). EMG profiles of knee joint musculature during walking: Changes induced by anterior cruciate ligament deficiency. *Journal of Orthopaedic Research, 6*(5), 630–638. https://doi.org/10.1002/jor.1100060503.

Lindström, M., Felländer-Tsai, L., Wredmark, T., & Henriksson, M. (2010). Adaptations of gait and muscle activation in chronic ACL deficiency. *Knee Surgery, Sports Traumatology, Arthroscopy, 18*(1), 106–114. https://doi.org/10.1007/s00167-009-0886-x.

Meyer, C., Mohr, M., Falbriard, M., Nigg, S. R., & Nigg, B. M. (2018). Influence of footwear comfort on the variability of running kinematics. *Footwear Science, 10*(1), 29–38. https://doi.org/10.1080/19424280.2017.1388296.

Mohr, Maurice, von Tscharner, V., Emery, C. A., & Nigg, B. M. (2018). Data for: Surface EMG muscle activation patterns of the lower extremities during gait in individuals with and without a knee injury history. *Mendeley Data, v1*. https://doi.org/10.17632/f2fv7gb577.1.

Nair, S. S., French, R. M., Laroche, D., & Thomas, E. (2010). The application of machine learning algorithms to the analysis of electromyographic patterns from arthritic patients. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 18*(2), 174–184. https://doi.org/10.1109/TNSRE.2009.2032638.

Nigg, B. M., & Herzog, W. (2007). *Biomechanics of the musculo-skeletal system* (3rd ed). New Jersey: John Wiley & Sons.

Õunpuu, S., & Winter, D. A. (1989). Bilateral electromyographical analysis of the lower limbs during walking in normal adults. *Electroencephalography and Clinical Neurophysiology, 72*(5), 429–438. https://doi.org/10.1016/0013-4694(89)90048-5.

Palmieri-Smith, R. M., & Thomas, A. C. (2009). A neuromuscular mechanism of posttraumatic osteoarthritis associated with ACL injury. *Exercise and Sport Sciences Reviews, 37*(3), 147–153.

Pietrosimone, B. G., McLeod, M. M., & Lepley, A. S. (2012). A theoretical framework for understanding neuromuscular response to lower extremity joint injury. *Sports Health: A Multidisciplinary Approach, 4*(1), 31–35. https://doi.org/10.1177/1941738111428251.

Roberts, C. S., Rash, G. S., Honaker, J. T., Wachowiak, M. P., & Shaw, J. C. (1999). A deficient anterior cruciate ligament does not lead to quadriceps avoidance gait. *Gait & Posture, 10*(3), 189–199. https://doi.org/10.1016/S0966-6362(99)00038-7.

Roos, E. M., Herzog, W., Block, J. A., & Bennell, K. L. (2011). Muscle weakness, afferent sensory dysfunction and exercise in knee osteoarthritis. *Nature Reviews Rheumatology, 7*(1), 57–63. https://doi.org/10.1038/nrrheum.2010.195.

Sadeghi, H., Allard, P., Prince, F., & Labelle, H. (2000). Symmetry and limb dominance in able-bodied gait: A review. *Gait & Posture, 12*(1), 34–45. https://doi.org/10.1016/S0966-6362(00)00070-9.

Shanbehzadeh, S., Bandpei, M. A. M., & Ehsani, F. (2017). Knee muscle activity during gait in patients with anterior cruciate ligament injury: A systematic review of electromyographic studies. *Knee Surgery, Sports Traumatology, Arthroscopy, 25*(5), 1432–1442. https://doi.org/10.1007/s00167-015-3925-9.

Stirling, L. M., von Tscharner, V., Kugler, P. F., & Nigg, B. M. (2011). Classification of muscle activity based on effort level during constant pace running. *Journal of Electromyography and Kinesiology, 21*(4), 566–571. https://doi.org/10.1016/j.jelekin.2011.02.005.

Toomey, C. M., Whittaker, J. L., Nettel-Aguirre, A., Reimer, R. A., Woodhouse, L. J., Ghali, B., ... Emery, C. A. (2017). Higher fat mass is associated with a history of knee injury in youth sport. *The Journal of Orthopaedic and Sports Physical Therapy, 47*(2), 80–87. https://doi.org/10.2519/jospt.2017.7101.

Vapnik, V. N. (2000). *The nature of statistical learning theory* (2nd ed). New York: Springer.

Victor, J., Labey, L., Wong, P., Innocenti, B., & Bellemans, J. (2010). The influence of muscle load on tibiofemoral knee kinematics. *Journal of Orthopaedic Research, 28*(4), 419–428. https://doi.org/10.1002/jor.21019.

von Tscharner, Vinzenz, & Goepfert, B. (2003). Gender dependent EMGs of runners resolved by time/frequency and principal pattern analysis. *Journal of Electromyography and Kinesiology: Official Journal of the International Society of Electrophysiological Kinesiology, 13*(3), 253–272.

von Tscharner, Vinzenz (2000). Intensity analysis in time-frequency space of surface myoelectric signals by wavelets of specified resolution. *Journal of Electromyography and Kinesiology: Official Journal of the International Society of Electrophysiological Kinesiology, 10*(6), 433–445.

von Tscharner, Vinzenz, Ullrich, M., Mohr, M., Comaduran Marquez, D., & Nigg, B. M. (2018). Beta, gamma band, and high-frequency coherence of EMGs of vasti muscles caused by clustering of motor units. *Experimental Brain Research*. https://doi.org/10.1007/s00221-018-5356-6.

von Tscharner, Vinzenz, Ullrich, M., Mohr, M., Marquez, D. C., & Nigg, B. M. (2018). A wavelet based time frequency analysis of electromyograms to group steps of runners into clusters that contain similar muscle activation patterns. *PLOS ONE, 13*(4), e0195125. https://doi.org/10.1371/journal.pone.0195125.

von Tscharner, Vinzenz, & Valderrabano, V. (2010). Classification of multi muscle activation patterns of osteoarthritis patients during level walking. *Journal of Electromyography and Kinesiology, 20*(4), 676–683. https://doi.org/10.1016/j.jelekin.2009.11.005.

Wellsandt, E., Gardinier, E. S., Manal, K., Axe, M. J., Buchanan, T. S., & Snyder-Mackler, L. (2016). Decreased knee joint loading associated with early knee osteoarthritis after anterior cruciate ligament injury. *The American Journal of Sports Medicine, 44*(1), 143–151. https://doi.org/10.1177/0363546515608475.

Whittaker, Jackie L., Toomey, C. M., Nettel-Aguirre, A., Jaremko, J. L., Doyle-Baker, P. K., Woodhouse, L. J., & Emery, C. A. (2018). Health-related outcomes following a youth sport–related knee injury. *Medicine & Science in Sports & Exercise*. https://doi.org/10.1249/MSS.0000000000001787 Publish Ahead of Print.

Whittaker, Jackie L., Toomey, C. M., Woodhouse, L. J., Jaremko, J. L., Nettel-Aguirre, A., & Emery, C. A. (2018). Association between MRI-defined osteoarthritis, pain, function and strength 3–10 years following knee joint injury in youth sport. *British Journal of Sports Medicine, 52*(14), 934–939. https://doi.org/10.1136/bjsports-2017-097576.

Whittaker, Jackie L., Woodhouse, L. J., Nettel-Aguirre, A., & Emery, C. A. (2015). Outcomes associated with early post-traumatic osteoarthritis and other negative health consequences 3–10 years following knee joint injury in youth sport. *Osteoarthritis and Cartilage, 23*(7), 1122–1129. https://doi.org/10.1016/j.joca.2015.02.021.

Yamazaki, J., Muneta, T., Ju, Y. J., & Sekiya, I. (2010). Differences in kinematics of single leg squatting between anterior cruciate ligament-injured patients and healthy controls. *Knee Surgery, Sports Traumatology, Arthroscopy, 18*(1), 56. https://doi.org/10.1007/s00167-009-0892-z.